Cisco HyperFlex 4.0 for VMware Horizon VDI with VMware ESXi for up to 5000 Users

Design and Deployment Guide for a Virtual Desktop Infrastructure Deployment of 5000 Users with VMWare Horizon 7.12, Cisco HyperFlex Data Platform 4.0.2a, and VMware ESXi 6.7

Published: October 2020

About the Cisco Validated Design Program

The Cisco Validated Design (CVD) program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information, go to:

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

Contents

Executive Summary

To keep pace with the market, you need systems that support rapid, agile development processes. Cisco HyperFlex™ Systems let you unlock the full potential of hyper-convergence and adapt IT to the needs of your workloads. The systems use an end-to-end software-defined infrastructure approach, combining software-defined computing in the form of Cisco HyperFlex HX-Series Nodes, software-defined storage with the powerful Cisco HyperFlex HX Data Platform, and software-defined networking with the Cisco UCS fabric that integrates smoothly with Cisco® Application Centric Infrastructure (Cisco ACI™).

Together with a single point of connectivity and management, these technologies deliver a pre-integrated and adaptable cluster with a unified pool of resources that you can quickly deploy, adapt, scale, and manage to efficiently power your applications and your business

This document provides an architectural reference and design guide for up to 5000 RDS (Remote Desktop Server Sessions) and 4000 (Windows 10) VDI sessions workload on a 24-node (8x Cisco HyperFlex HXAF220C-M5SX server and 8x Cisco UCS C220M5 + 8x B200 M5 Compute only nodes) Cisco HyperFlex system. This is our first CVD featuring the ability for customers to subscribe 2 compute only nodes for every 1 hyperconverged node. This is an exclusive Cisco HyperFlex capability. No other hyperconverged platform vendor offers this significant money saving capability.

We provide deployment guidance and performance data for VMware Horizon virtual desktops running Microsoft Windows 10 with Office 2016 and Windows Server 2019 for Remote Desktop Server Hosted (RDSH). The solution is a pre-integrated, best-practice data center architecture built on Cisco Unified Computing System (Cisco UCS), the Cisco Nexus® 9000 family of switches, and Cisco HyperFlex Data Platform software version 4.0.2a.

The solution payload is 100 percent virtualized on Cisco HyperFlex HXAF220C-M5SX hyperconverged nodes, Cisco UCS C220 M5 Compute-Only and Cisco UCS B200 M5 Compute-Only Nodes booting through on-board M.2 SATA SSD drive running VMware ESXi hypervisor and the Cisco HyperFlex Data Platform storage controller virtual machine. The virtual desktops are configured with Horizon 7.12, which incorporates both traditional persistent and non-persistent virtual Windows 10 desktops, hosted applications, and remote desktop service (RDS) Microsoft Server 2019 based desktops. The solution provides unparalleled scale and management simplicity. VMware Horizon Windows 10 full clone desktops or RDSH server-based sessions can be provisioned on an eight node Cisco HyperFlex cluster. Where applicable, this document provides best practice recommendations and sizing guidelines for customer deployment of this solution.

Solution Overview

Introduction

The current industry trend in data center design is towards small, granularly expandable hyperconverged infrastructures. By using virtualization along with pre-validated IT platforms, customers of all sizes have embarked on the journey to "just-in-time capacity" using this new technology. The Cisco HyperFlex hyperconverged solution can be quickly deployed, thereby increasing agility, and reducing costs. Cisco HyperFlex uses best of breed storage, server, and network components to serve as the foundation for desktop virtualization workloads, enabling efficient architectural designs that can be quickly and confidently deployed and scaled-out.

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, VMware and Microsoft specific technologies, infrastructure concepts, networking connectivity, and security policies of the customer installation.

Purpose of this Document

This document provides a step-by-step design, configuration, and implementation guide for the Cisco Validated Design for a Cisco HyperFlex All-Flash system running VMware Remote Desktop Server Hosted (RDSH) server sessions two different Windows 10 workloads with Cisco UCS 6400 series Fabric Interconnects and Cisco Nexus 9000 series switches.

What's New in this Release?

This is the first Cisco Validated Design with Cisco HyperFlex All-Flash system running Virtual Desktop Infrastructure on Intel Xeon Scalable Family processor-based, fifth generation Cisco UCS HyperFlex system with these features:

- Cisco HyperFlex support for 2:1 compute only: hyperconverged nodes at large scale

- Support for the Cisco UCS 4.0(1b) release

- VMware vSphere 6.7 U3 Hypervisor

- VMware Horizon 7.12 Remote Desktop Sever Hosted Sessions

- VMware Horizon 7.12 Horizon Instant Clones and Persistent Full Desktops

New Features for Version 4.0

The Cisco HX Data Platform, release 4.0 provides the following features. These features, including the Invisible Cloud Witness for HyperFlex Edge clusters, are supported both on the Intersight Virtual Appliance and on Intersight.com.

New Features in Release 4.0(2c)

The following new features are in release 4.0(2c):

- HyperFlex Edge Short Depth Servers–New, short depth server offerings are now available for HyperFlex Edge. Both All-flash (HXAF240c-M5SD) and Hybrid (HX240c-M5SD) configuration options are available. See the [HyperFlex HX*240 SD Short Depth Edge Nodes](#) spec sheet for full details.

- New Drive support–New 7.12TB SED SSD capacity point introduced (HX-SD76TBEM2NK9). Also new drives SKUs for existing drive capacity points 3.8TB (HX-SD38TBEM2NK9) and 960G (HX-SD960GBM2NK9).

New Features in Release 4.0(2b)

The following new features are in release 4.0(2b):

- 7.12TB SSD data drive–Support on HX Edge configurations.

- Support for Cisco UCS Fabric Interconnects with limited cluster—Cisco UCS Fabric Interconnect (FI-64108) with limited scale of 32 nodes now supported.
- All NVMe and All Flash limits increase:
  ◦ Maximum cluster size for All NVMe (with 1TB, 4TB, or 8TB data drives) increased to 32 nodes.
  ◦ Maximum cluster size for HX220 All Flash with 7.12TB data drive increased to 32 nodes.
  ◦ HX240 All Flash increased to 16 nodes for full drive population (23 data drives/node) or 32 nodes at up to 12 drives/node.
- Cluster Scale Limits increase—Support for maximum scale limit increase on a cluster. See [Cisco HX Data Platform Compatibility and Scalability Details](#)
- HW Offload option—Support for Hardware Offload option with Stretched cluster configurations.
- Cisco Overlay Transport Virtualization for Stretched Cluster—Support for OTV as an overlay for Stretched Cluster.

New Features in Release 4.0(2a)

The following new features are in release 4.0(2a):

- Boost Mode—This release introduces Boost Mode for the following configurations: All NVMe, All Flash C240, All Flash C220, and Hypervisor: ESX. Boost Mode allows the Cisco HyperFlex cluster to deliver higher IOPs by increasing the storage controller VM CPU resources by 4 vCPU. For configuration information, see the [Cisco HyperFlex Data Platform Administration Guide](#).
- Cisco HyperFlex HTML plug-in for VMWare vCenter—Enables virtualization administrator to manage and monitor the Cisco HyperFlex physical infrastructure by cross launching HyperFlex Connect from the vSphere Client UI and perform management actions in the HyperFlex Connect UI.
- 25GE networking for HX Edge—Support for 25GE networking for HX Edge.
- All NVMe with Stretched Cluster—Support for All NVMe with Stretched Cluster (ESX only).
- Cluster Upgrade Eligibility Test—This release adds the capability to perform a pre-upgrade test which checks for cluster readiness before upgrading. Example checks include validating cluster state, rebalance status, controller VM Free Space, ESXi version, and much more. The Eligibility test is intended to help avoid unexpected problems that may arise during the upgrade process. It is highly recommended to run the test before performing the Hyperflex upgrade.
- Registering Smart Software Licensing—This release adds support for software that allows easy tracking of the status of license and software usage trends and simplifies the three core licensing functions: Purchasing, Management, and Reporting.
- Dynamic self-signed certificate generation enhancements—This release adds support for Self-signed SSL certificates on the Controller VMs, which were static in prior releases. The static certificates are replaced with dynamically generated self-signed certificates upon upgrading to HXDP 4.0(2a) so that the certificates are unique per cluster. The new clusters installed with HXDP 4.0(2a) have dynamically generated self-signed certificates.
- Test Upgrade Eligibility—This release adds support for testing your cluster readiness and infrastructure compatibility for an upgrade. For more information, see the Test Upgrade Eligibility sections in the [Cisco](#)

[HyperFlex Systems Upgrade Guide for VMware ESXi, Release 4.0](#) or the [Cisco HyperFlex Upgrade Guide for Microsoft Hyper-V, Release 4.0](#).

- Disaster Recovery:
  - Recovery Settings Configuration—This release supports configuration of recovery settings to define global recovery parameters and mapping for resources across recovery sites. These parameters are used during recovery, test recovery and migrate operations.
  - HyperFlex DR PowerShell Runbook—PowerShell runbook functionality is extended to support the recovery configuration settings in the runbooks for all recovery scenarios. New-HXRunbook cmdlet can now be used to generate runbook for a single or multiple protection groups. In addition, two new cmdlets, Wait-HXTask and Get-HXTaskStatus, are introduced.

- Protected Virtual Machine Scalability—This release adds support for 1500 VMs across both clusters and 750 VMs per cluster in a bi-direction or any split between the two clusters without exceeding the limit of 1500 VMs. For more information, see [Cisco HyperFlex Data Platform Administration Guide](#).

- System Management REST API enhancements—Pause data replication actions briefly using REST API to explicitly inform users on the current status of replication actions.

- Cisco HyperFlex with Data Platform for Hyper-V

- Cluster-Aware Updating (CAU)—This is an automated feature that allows you to perform updates on windows servers in a failover cluster with little or no loss in availability during the upgrade process.

New Features in Release 4.0(1b)

The following new feature is in release 4.0(1b):

- Support for Second Generation Intel® Xeon® Scalable Processor Refresh—This release includes support for the 2nd Generation Intel® Xeon® Scalable processor refresh (formerly Cascade Lake).

New Features in Release 4.0(1a)

The following new features are in release 4.0(1a):

- Ultra-Light HyperFlex Edge Clusters—This release introduces support for two-node HyperFlex Edge clusters, enabling HyperFlex to run in environments requiring a small footprint. Cisco Intersight provides comprehensive lifecycle management and includes remote cloud-based installation, centralized upgrades, and invisible witnessing. Both 1GE and 10GE networking topology options are available.

- Scaled-Up HyperFlex Edge Clusters—This release adds support for four node HyperFlex Edge clusters, enabling a full range of size options for remote and branch offices. Size the branch office environments to suit current needs with a two, three, or four node HyperFlex Edge cluster. Cisco Intersight provides full-lifecycle management and 1GE, and 10GE networking options are available.

- Cisco Intersight Invisible Cloud Witness—For two node clusters, this feature eliminates the need for witness VMs, the infrastructure to run those VMs, and the management overhead to deploy, scale, and patch witnessing software. The Invisible Cloud Witness is responsible for maintaining cluster HA in the event of failure scenarios. This feature is included at no additional cost and is automatically deployed and managed by Cisco Intersight.

- Cloud-delivered HyperFlex Edge Upgrades—Powered by Cisco Intersight, this feature adds support for multi-site orchestrated remote upgrades of the HyperFlex Data Platform. This feature will be enabled with the next 4.0 patch release and will allow HyperFlex Edge clusters deployed via Intersight to perform orchestrated upgrades across one or many sites in parallel.

- All-NVMe HyperFlex—Starting with this release, a new, high-end performance node powered by all NVMe drives is available for HyperFlex clusters. Co-engineered with Intel to support Intel VMD for hot-plug and surprise removal, this offering represents an industry first: an enterprise-ready and fully validated all-NVMe HCI appliance. The all-NVMe offering is available in the 220 form factor (1RU) and is powered by Intel Optane cache drives for maximum performance and highest endurance.

- VMware Site Recovery Manager (SRM) Integration—This release brings support for a Cisco developed Storage Replication Adapter (SRA) for SRM. The SRA provides the ability to leverage HyperFlex native async replication with the powerful orchestration and runbook capabilities of SRM. The SRA includes the ability to perform test recoveries, planned migrations, and full disaster recovery.

⚠ HX SRA is certified by VMware and is available for download from VMWare SRM site.

- HyperFlex DR PowerShell Runbooks—New PowerShell cmdlets are included for automated runbook generation when using HyperFlex native disaster recovery. The New-HXRunbook cmdlet supports the following workflows: Test Recovery, Planned Migration, and Disaster Recovery. These runbooks can be used to orchestrate DR workflows without the requirement for any third-party software.

- Windows Server 2019 with Hyper-V—Support has been added in this release for the Windows Server 2019 operating system for Hyper-V based HyperFlex deployments.

- Kubernetes CSI Plugin—This release adds support for the HyperFlex CSI (HX-CSI) plugin based on the Kubernetes Container Storage Interface (CSI) specification. Customers can now use the HX-CSI plugin to provision and manage persistent volumes in Kubernetes version 1.13 and later.

⚠ Kubernetes 1.13 support for Cisco Container Platform and Openshift Container Platform is forthcoming in their respective future releases.

- Cisco UCS C480 ML Compute-only node—This release introduces support for Cisco UCS C480 ML as a new compute-only node for Deep Learning/Machine Learning Workloads. Data scientists can now use the power of up to eight NVidia SXM2 V100 GPUs to accelerate deep learning workloads. VMs running deep-learning workloads will need to use PCIe pass-through for access to GPUs.

- Higher Capacity Drives—A new 2.4TB 10k rpm SAS HDD option for SFF Hybrid HyperFlex clusters, and a 12TB 7.2K rpm SAS HDD option for LFF Hybrid are now available. Both HyperFlex and HyperFlex Edge support the 2.4TB capacity point for maximum density in this form factor. Note that HyperFlex Edge does not support LFF-drives. HyperFlex Hyper-V version does not yet support the new 12TB drive option. See the HyperFlex spec sheets for a full list of configurable options.

- New Cache and increased scale for Hyper-V —NVMe and Optane SSDs are now supported as cache drives for Hyper-V deployments. Furthermore, scale limits have been increased to 16+16 (Converged+Compute-only) for both SFF (AF and Hybrid) and LFF (Hybrid) clusters.

- Centralized Audit Log Export—This release adds support for audit logging via a remote syslog server. This capability enables customers to retain audit logs from all HyperFlex nodes in a centralized remote syslog server to meet retention and compliance requirements.
- DISA STIG Compliance—This release adds new HX REST APIs for setting, removing, and checking status of DISA STIGs for Controller VMs, ESXi hosts and vCenter. These APIs enable customers to meet DISA security requirements by centrally and securely applying STIGs, detecting, and correcting for drifts in any STIG settings.

New Supported Drives

New drives are qualified for the 4.0(2a) release. The new drives include new capacity points and new cache drive options. Several of the new drives are alternate drives to already qualified existing drives in function which are qualified in 4.0(2a). These drives are functionally compatible with the existing drives and are available as alternates in case of lack of availability of existing drives. For expansion of existing clusters or general information about interoperability of different drives, see Cisco HyperFlex Drive Compatibility.

> ⚠ The NVMe Caching SSD's slot information is unavailable from HX-Connect for all AF server PIDs except for the All-NVMe server PIDs. Please refer to UCSM management console for NVMe SSD slot information.

Table 1   Supported Drives

Drive Function	Drive PID	Applicable Platforms	Version
7.12TB SED SSD Capacity drive	HX-SD76TBEM2NK9	All existing HX M5 servers except All NVMe	4.0(2c)
SKUs for existing 3.8TB and 960GB Capacity drive capacity	HX-SD38TBEM2NK9, HX-SD960GBM2NK9	All existing HX M5 servers except All NVMe	4.0(2c)
Alternate system (or housekeeping) drive	HX-SD480G6I1X-EV	All existing HX M5 servers except All NVMe	4.0(2b)
Alternate system (or housekeeping) drive	HX-SD480GM1X-EV	All existing HX M5 servers except All NVMe	4.0(2b)
New 960G FIPS compliant SED SSD data drives	HX-SD960G2HTNK9	HXAF220C-M5SX, HXAF240C-M5SX	4.0(2b)
Alternate boot drive	HX-M2-960GB	All existing HX M5 servers	4.0(2a)
All NVMe 4TB Capacity drive	HX-NVME2H-I4000	All NVMe: HXAF220C-M5SN	4.0(2a)
New high density All NVMe 8TB Capacity drive	HX-NVMEHW-I8000	All NVMe: HXAF220C-M5SN	4.0(2a)

Drive Function	Drive PID	Applicable Platforms	Version
New high density All Flash 7.12TB Capacity drive			
Maximum number of drives per node is 12 drives on HX AF C240.	HX-SD76T61X-EV	All Flash Configuration – namely: HXAF220C-M5SX, HXAF240C-M5SX, HXAF-E-220M5SX	
ESX support only.	4.0(2a)		
New 3.8TB FIPS compliant SED SSD data drives	HX-SD38T2HTNK9	HXAF220C-M5SX, HXAF240C-M5SX	4.0(2a)
Alternate drive for 8TB LFF capacity	HX-HD8T7K4KAN	HX240C-M5L	4.0(2a)

Documentation Roadmap

For the comprehensive documentation suite, refer to the [Cisco HyperFlex Systems Documentation Roadmap](#).

⚠   A login is required for the Documentation Roadmap.

Cisco HX Data Platform requires specific software and hardware versions, and networking settings for successful installation. See the [Cisco HyperFlex Systems Getting Started Guide](#) for a complete list of requirements.

For a complete list of hardware and software inter-dependencies, refer to the Cisco UCS Manager release version of [Hardware and Software Interoperability for Cisco HyperFlex HX-Series.](#)

Hyperflex Cisco Validated Design Advantage for VDI

The data center market segment is shifting toward heavily virtualized private, hybrid and public cloud computing models running on industry-standard systems. These environments require uniform design points that can be repeated for ease if management and scalability.

These factors have led to the need predesigned computing, networking and storage building blocks optimized to lower the initial design cost, simply management, and enable horizontal scalability and high levels of utilization. The use cases include:

- Enterprise Data Center (small failure domains)
- Service Provider Data Center (small failure domains)
- Commercial Data Center
- Remote Office/Branch Office
- SMB Standalone Deployments
- Solution Summary

This Cisco Validated Design prescribes a defined set of hardware and software that serves as an integrated foundation for both Horizon Microsoft Windows 10 virtual desktops and Horizon RDSH server desktop sessions based on Microsoft Server 2019. The mixed workload solution includes Cisco HyperFlex hardware and Data Platform software, Cisco Nexus® switches, the Cisco Unified Computing System (Cisco UCS®), VMware Horizon and VMware vSphere software in a single package. The design is efficient such that the networking, computing, and storage components occupy 18-rack units footprint in an industry standard 42U rack. Port density on the Cisco Nexus switches and Cisco UCS Fabric Interconnects enables the networking components to accommo-date multiple HyperFlex clusters in a single Cisco UCS domain.

A key benefit of the Cisco Validated Design architecture is the ability to customize the environment to suit a cus-tomer's requirements. A Cisco Validated Design scales easily as requirements and demand change. The unit can be scaled both up (adding resources to a Cisco Validated Design unit) and out (adding more Cisco Validated Design units).

The reference architecture detailed in this document highlights the resiliency, cost benefit, and ease of deploy-ment of a hyper-converged desktop virtualization solution. A solution capable of consuming multiple protocols across a single interface allows for customer choice and investment protection because it truly is a wire-once architecture.

The combination of technologies from Cisco Systems, Inc. and VMware Inc. produced a highly efficient, robust, and affordable desktop virtualization solution for a virtual desktop, hosted shared desktop or mixed deployment supporting different use cases. Key components of the solution include the following:

- More power, same size. Cisco HX-series nodes, dual 18-core 2.3 GHz Intel Xeon (Gold 6230) Scalable Family processors with 768GB of 2666Mhz memory with VMware Horizon support more virtual desktop workloads than the previously released generation processors on the same hardware. The Intel Xeon Gold 6230 18-core scalable family processors used in this study provided a balance between increased per-server capacity and cost

- Fault-tolerance with high availability built into the design. The various designs are based on multiple Cisco HX-Series nodes, Cisco UCS rack servers and Cisco UCS blade servers for virtual desktop and infrastruc-ture workloads. The design provides N+1 server fault tolerance for every payload type tested

- Stress-tested to the limits during aggressive boot scenario. The 5000 user Remote Desktop Hosted Ses-sions (RDSH) and 4000 Virtual Desktops environment booted and registered with the Horizon 7 in under 15 minutes, providing our customers with an extremely fast, reliable cold-start desktop virtualization sys-tem.

- Stress-tested to the limits during simulated login storms. The 5000 user Remote Desktop Hosted Sessions (RDSH) and 4000 Virtual Desktops environment ready state in 48-minutes without overwhelming the pro-cessors, exhausting memory, or exhausting the storage subsystems, providing customers with a desktop virtualization system that can easily handle the most demanding login and startup storms.

- Ultra-condensed computing for the datacenter. The rack space required to support the initial 5000 user system is 26 rack units, including Cisco Nexus Switching and Cisco Fabric interconnects. Incremental seat Cisco HyperFlex clusters can be added one at a time to a total of 64 nodes.

- 100 percent virtualized This CVD presents a validated design that is 100 percent virtualized on VMware ESXi 6.7. All of the virtual desktops, user data, profiles, and supporting infrastructure components, includ-ing Active Directory, SQL Servers, VMware Horizon components, Horizon VDI desktops and RDSH servers

were hosted as virtual machines. This provides customers with complete flexibility for maintenance and capacity additions because the entire system runs on the Cisco HyperFlex hyper-converged infrastructure with stateless Cisco UCS HX-series servers. (Infrastructure VMs were hosted on two Cisco UCS C220 M4 Rack Servers outside of the HX cluster to deliver the highest capacity and best economics for the solution.)

- Cisco data center management: Cisco maintains industry leadership with the new Cisco UCS Manager 4.0(1b) software that simplifies scaling, guarantees consistency, and eases maintenance. Cisco's ongoing development efforts with Cisco UCS Manager, Cisco UCS Central, and Cisco UCS Director ensure that customer environments are consistent locally, across Cisco UCS Domains and across the globe. Cisco UCS software suite offers increasingly simplified operational and deployment management, and it continues to widen the span of control for customer organizations' subject matter experts in compute, storage, and network.

- Cisco 40G Fabric: Our 40G unified fabric story gets additional validation on 6400 Series Fabric Interconnects as Cisco runs more challenging workload testing, while maintaining unsurpassed user response times.

- Cisco HyperFlex Connect (HX Connect): An all-new HTML 5 based Web UI Introduced with HyperFlex v2.5 or later is available for use as the primary management tool for Cisco HyperFlex. 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.

- Cisco HyperFlex storage performance: Cisco HyperFlex provides industry-leading hyper converged storage performance that efficiently handles the most demanding I/O bursts (for example, login storms), high write throughput at low latency, delivers simple and flexible business continuity and helps reduce storage cost per desktop.

- Cisco HyperFlex agility: Cisco HyperFlex System enables users to seamlessly add, upgrade or remove storage from the infrastructure to meet the needs of the virtual desktops.

- Cisco HyperFlex vCenter integration: Cisco HyperFlex plugin for VMware vSphere provides easy-button automation for key storage tasks such as storage provisioning and storage resize, cluster health status and performance monitoring directly from the vCenter web client in a single pane of glass. Experienced vCenter administrators have a near zero learning curve when HyperFlex is introduced into the environment.

- VMware Horizon 7 advantage: VMware Horizon 7 follows a new unified product architecture that supports both Virtual Desktops and Remote Desktop Server Hosted server sessions. This new Horizon release simplifies tasks associated with large-scale VDI management. This modular solution supports seamless delivery of Windows apps and desktops as the number of user increase. In addition, PCoIP and Blast extreme enhancements help to optimize performance and improve the user experience across a variety of endpoint device types, from workstations to mobile devices including laptops, tablets, and smartphones.

- Optimized for performance and scale. For hosted shared desktop sessions, the best performance was achieved when the number of vCPUs assigned to the Horizon 7 RDSH virtual machines did not exceed the number of hyper-threaded (logical) cores available on the server. In other words, maximum performance is obtained when not overcommitting the CPU resources for the virtual machines running virtualized RDS systems.

- Provisioning desktop machines made easy: VMware Horizon 7 provisions Remote Desktop Hosted Sessions (RDSH) virtual desktops as well as hosted shared desktop virtual machines for this solution using a single method for both, the "Automated floating assignment desktop pool." "Dedicated user assigned desktop pool" for persistent desktops was provisioned in the same Horizon 7 administrative console. Horizon 7 introduces a new provisioning technique for non-persistent virtual desktops called "Instant-clone." The new method greatly reduces the amount of life-cycle spend and the maintenance windows for the guest OS.

All-Flash Versus Hybrid

The initial HyperFlex product release featured hybrid converged nodes, which use a combination of solid-state disks (SSDs) for the short-term storage caching layer, and hard disk drives (HDDs) for the long-term storage capacity layer. The hybrid HyperFlex system is an excellent choice for entry-level or midrange storage solutions, and hybrid solutions have been successfully deployed in many non-performance sensitive virtual environments. Meanwhile, there is significant growth in deployment of highly performance sensitive and mission critical applications. The primary challenge to the hybrid HyperFlex system from these highly performance sensitive applications, is their increased sensitivity to high storage latency. Due to the characteristics of the spinning hard disks, it is unavoidable that their higher latency becomes the bottleneck in the hybrid system. Ideally, if all of the storage operations were to occur in the caching SSD layer, the hybrid system's performance will be excellent. However, in several scenarios, the amount of data being written and read exceeds the caching layer capacity, placing larger loads on the HDD capacity layer, and the subsequent increases in latency will naturally result in reduced performance.

Cisco All-Flash HyperFlex systems are an excellent option for customers with a requirement to support high performance, latency sensitive workloads. With a purpose built, flash-optimized and high-performance log-based filesystem, the Cisco All-Flash HyperFlex system provides:

- Predictable high performance across all the virtual machines on HyperFlex All-Flash and compute-only nodes in the cluster.

- Highly consistent and low latency, which benefits data-intensive applications and databases such as Microsoft SQL and Oracle.

- Support for NVMe caching SSDs, offering an even higher level of performance.

- Future ready architecture that is well suited for flash-memory configuration:
  ◦ Cluster-wide SSD pooling maximizes performance and balances SSD usage so as to spread the wear.
  ◦ A fully distributed log-structured filesystem optimizes the data path to help reduce write amplification.
  ◦ Large sequential writing reduces flash wear and increases component longevity.
  ◦ Inline space optimization, such as deduplication and compression, minimizes data operations and reduces wear.

- Lower operating cost with the higher density drives for increased capacity of the system.

- Cloud scale solution with easy scale-out and distributed infrastructure and the flexibility of scaling out independent resources separately.

Cisco HyperFlex support for hybrid and all-flash models now allows customers to choose the right platform configuration based on their capacity, applications, performance, and budget requirements. All-flash configurations offer repeatable and sustainable high performance, especially for scenarios with a larger working set of data, in

other words, a large amount of data in motion. Hybrid configurations are a good option for customers who want the simplicity of the Cisco HyperFlex solution, but their needs focus on capacity-sensitive solutions, lower budgets, and fewer performance-sensitive applications.

Cisco HyperFlex Compute-Only Nodes

All current model Cisco UCS M4 and M5 generation servers, except the Cisco UCS C880 M4 and Cisco UCS C880 M5, may be used as compute-only nodes connected to a Cisco HyperFlex cluster, along with a limited number of previous M3 generation servers. Any valid CPU and memory configuration is allowed in the compute-only nodes, and the servers can be configured to boot from SAN, local disks, or internal SD cards. The following servers may be used as compute-only nodes:

- Cisco UCS B200 M3 Blade Server

- Cisco UCS B200 M4 Blade Server

- Cisco UCS B200 M5 Blade Server

- Cisco UCS B260 M4 Blade Server

- Cisco UCS B420 M4 Blade Server

- Cisco UCS B460 M4 Blade Server

- Cisco UCS B480 M5 Blade Server

- Cisco UCS C220 M3 Rack-Mount Servers

- Cisco UCS C220 M4 Rack-Mount Servers

- Cisco UCS C220 M5 Rack-Mount Servers

- Cisco UCS C240 M3 Rack-Mount Servers

- Cisco UCS C240 M4 Rack-Mount Servers

- Cisco UCS C240 M5 Rack-Mount Servers

- Cisco UCS C460 M4 Rack-Mount Servers

- Cisco UCS C480 M5 Rack-Mount Servers

Cisco HyperFlex 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 expected in 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:

- Data protection creates multiple copies of the data across the cluster so that data availability is not affected if single or multiple components fail (depending on the replication factor configured).

- Stretched clusters allow nodes to be evenly split between two physical locations, keeping a duplicate copy of all data in both locations, thereby providing protection in case of an entire site failure.

- Logical availability zones provide multiple logical grouping of nodes and distributes the data across these groups in such a way that no single group has more than one copy of the data. This enables enhanced protection from node failures, allowing for more nodes to fail while the overall cluster remains online.

- Deduplication is always on, helping reduce storage requirements in virtualization clusters in which multiple operating system instances in guest 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.

- Replication copies virtual machine level snapshots from one Cisco HyperFlex cluster to another, to facilitate recovery from a cluster or site failure, via a failover to the secondary site of all VMs.

- Encryption stores all data on the caching and capacity disks in an encrypted format, to prevent accidental data loss or data theft. Key management can be done using local Cisco UCS Manager managed keys, or third-party Key Management Systems (KMS) via the Key Management Interoperability Protocol (KMIP).

- 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 duplicate virtual storage volumes so that virtual machines can be cloned simply through metadata operations, with actual data copied only for write operations.

- Snapshots help facilitate backup and remote-replication operations, which are needed in enterprises that require always-on data availability.

Cisco Desktop Virtualization Solutions: Data Center

The Evolving Workplace

Today's IT departments are facing a rapidly evolving workplace environment. The workforce is becoming increasingly diverse and geographically dispersed, including offshore contractors, distributed call center operations, knowledge and task workers, partners, consultants, and executives connecting from locations around the world at all times.

This workforce is also increasingly mobile, conducting business in traditional offices, conference rooms across the enterprise campus, home offices, on the road, in hotels, and at the local coffee shop. This workforce wants to use a growing array of client computing and mobile devices that they can choose based on personal preference. These trends are increasing pressure on IT to ensure protection of corporate data and prevent data leakage or loss through any combination of user, endpoint device, and desktop access scenarios (Figure 1).

These challenges are compounded by desktop refresh cycles to accommodate aging PCs and bounded local storage and migration to new operating systems, specifically Microsoft Windows 10 and productivity tools, specifically Microsoft Office 2016.

Figure 1　Cisco Data Center Partner Collaboration
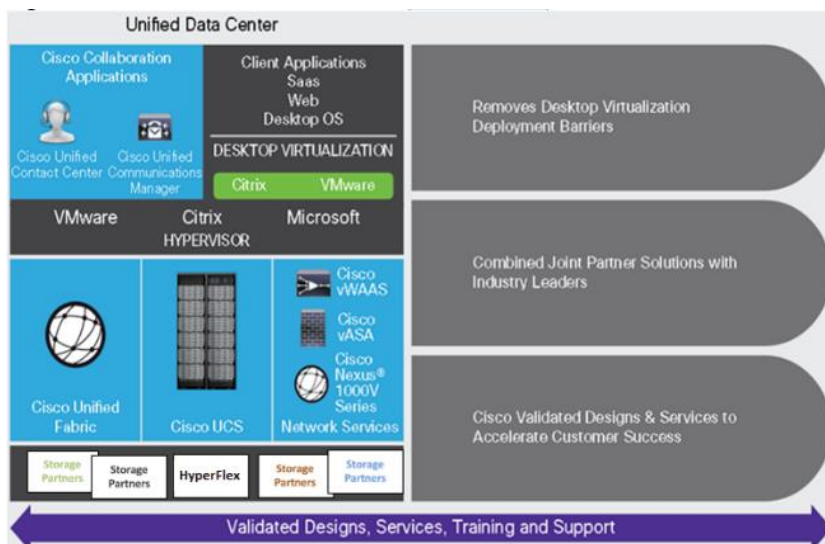
Some of the key drivers for desktop virtualization are increased data security, the ability to expand and contract capacity and reduced TCO through increased control and reduced management costs.

Cisco Desktop Virtualization Focus

Cisco focuses on three key elements to deliver the best desktop virtualization data center infrastructure: simplification, security, and scalability. The software combined with platform modularity provides a simplified, secure, and scalable desktop virtualization platform.

Simplified

Cisco UCS and Cisco HyperFlex provide a radical new approach to industry-standard computing and provides the core of the data center infrastructure for desktop virtualization. Among the many features and benefits of Cisco UCS are the drastic reduction in the number of servers needed, in the number of cables used per server and the capability to rapidly deploy or re-provision servers through Cisco UCS service profiles. With fewer servers and cables to manage and with streamlined server and virtual desktop provisioning, operations are significantly simplified. Thousands of desktops can be provisioned in minutes with Cisco UCS Manager service profiles and Cisco storage partners' storage-based cloning. This approach accelerates the time to productivity for end users, improves business agility, and allows IT resources to be allocated to other tasks.

Cisco UCS Manager automates many mundane, error-prone data center operations such as configuration and provisioning of server, network, and storage access infrastructure. In addition, Cisco UCS B-Series Blade Servers, C-Series and HX-Series Rack Servers with large memory footprints enable high desktop density that helps reduce server infrastructure requirements.

Simplification also leads to more successful desktop virtualization implementation. Cisco and its technology partners like VMware have developed integrated, validated architectures, including predefined hyper-converged architecture infrastructure packages such as HyperFlex. Cisco Desktop Virtualization Solutions have been tested with VMware vSphere.

Secure

Although virtual desktops are inherently more secure than their physical predecessors, they introduce new security challenges. Mission-critical web and application servers using a common infrastructure such as virtual desktops are now at a higher risk for security threats. Inter-virtual machine traffic now poses an important security consideration that IT managers need to address, especially in dynamic environments in which virtual machines, using VMware vMotion, move across the server infrastructure.

Desktop virtualization, therefore, significantly increases the need for virtual machine-level awareness of policy and security, especially given the dynamic and fluid nature of virtual machine mobility across an extended computing infrastructure. The ease with which new virtual desktops can proliferate magnifies the importance of a virtualization-aware network and security infrastructure. Cisco data center infrastructure (Cisco UCS and Cisco Nexus Family solutions) for desktop virtualization provides strong data center, network, and desktop security, with comprehensive security from the desktop to the hypervisor. Security is enhanced with segmentation of virtual desktops, virtual machine-aware policies and administration, and network security across the LAN and WAN infrastructure.

Scalable

Growth of a desktop virtualization solution is accelerating, so a solution must be able to scale, and scale predictably, with that growth. The Cisco Desktop Virtualization Solutions support high virtual-desktop density (desktops per server) and additional servers scale with near-linear performance. Cisco data center infrastructure provides a flexible platform for growth and improves business agility. Cisco UCS Manager service profiles allow on-demand desktop provisioning and make it just as easy to deploy dozens of desktops as it is to deploy thousands of desktops.

Cisco HyperFlex servers provide near-linear performance and scale. Cisco UCS implements the patented Cisco Extended Memory Technology to offer large memory footprints with fewer sockets (with scalability to up to 3.0 terabyte (TB) of memory with 2- and 4-socket servers). Using unified fabric technology as a building block, Cisco UCS server aggregate bandwidth can scale to up to 80 Gbps per server, and the northbound Cisco UCS fabric interconnect can output 2 terabits per second (Tbps) at line rate, helping prevent desktop virtualization I/O and memory bottlenecks. Cisco UCS, with its high-performance, low-latency unified fabric-based networking architecture, supports high volumes of virtual desktop traffic, including high-resolution video and communications traffic. In addition, Cisco HyperFlex helps maintain data availability and optimal performance during boot and login storms as part of the Cisco Desktop Virtualization Solutions. Recent Cisco Validated Designs based on VMware Horizon, Cisco HyperFlex solutions have demonstrated scalability and performance, with up to 5000 Remote Desktop Hosted Sessions (RDSH) and 4000 windows 10 virtual desktops and up and running in ~15 minutes.

Cisco UCS and Cisco Nexus data center infrastructure provides an excellent platform for growth, with transparent scaling of server, network, and storage resources to support desktop virtualization, data center applications, and cloud computing.

Savings and Success

The simplified, secure, scalable Cisco data center infrastructure for desktop virtualization solutions saves time and money compared to alternative approaches. Cisco UCS enables faster payback and ongoing savings (better **ROI and lower TCO) and provides the industry's greatest virtual desktop density per server, reducing both capital** expenditures (CapEx) and operating expenses (OpEx). The Cisco UCS architecture and Cisco Unified Fabric also

enables much lower network infrastructure costs, with fewer cables per server and fewer ports required. In addition, storage tiering and deduplication technologies decrease storage costs, reducing desktop storage needs by up to 50 percent.

The simplified deployment of Cisco HyperFlex for desktop virtualization accelerates the time to productivity and enhances business agility. IT staff and end users are more productive more quickly, and the business can respond to new opportunities quickly by deploying virtual desktops whenever and wherever they are needed. The high-performance Cisco systems and network deliver a near-native end-user experience, allowing users to be productive anytime and anywhere.

The key measure of desktop virtualization for any organization is its efficiency and effectiveness in both the near term and the long term. The Cisco Desktop Virtualization Solutions are very efficient, allowing rapid deployment, requiring fewer devices and cables, and reducing costs. The solutions are also extremely effective, providing the services that end users need on their devices of choice while improving IT operations, control, and data security. **Success is bolstered through Cisco's best**-in-class partnerships with leaders in virtualization and through tested and validated designs and services to help customers throughout the solution lifecycle. Long-term success is **enabled through the use of Cisco's scalable, flexible, and secure architecture as the platform for desktop virtual**-ization.

The ultimate measure of desktop virtualization for any end-user is a great experience. Cisco HyperFlex delivers class-leading performance with sub-second base line response times and index average response times at full load of just under one second.

Use Cases

The following are some typical use cases:

- Healthcare: Mobility between desktops and terminals, compliance, and cost
- Federal government: Teleworking initiatives, business continuance, continuity of operations (COOP), and training centers
- Financial: Retail banks reducing IT costs, insurance agents, compliance, and privacy
- Education: K-12 student access, higher education, and remote learning
- State and local governments: IT and service consolidation across agencies and interagency security
- Retail: Branch-office IT cost reduction and remote vendors
- Manufacturing: Task and knowledge workers and offshore contractors
- Microsoft Windows 10 migration
- Graphic intense applications
- Security and compliance initiatives
- Opening of remote and branch offices or offshore facilities
- Mergers and acquisitions

Figure 2 shows the VMware Horizon 7 on vSphere 6.7 built on Cisco UCS components and the network connections. The reference architecture reinforces the "wire-once" strategy, because as additional storage is added to the architecture, no re-cabling is required from the hosts to the Cisco UCS fabric interconnect.

Solution Summary

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 2   HyperFlex System Overview
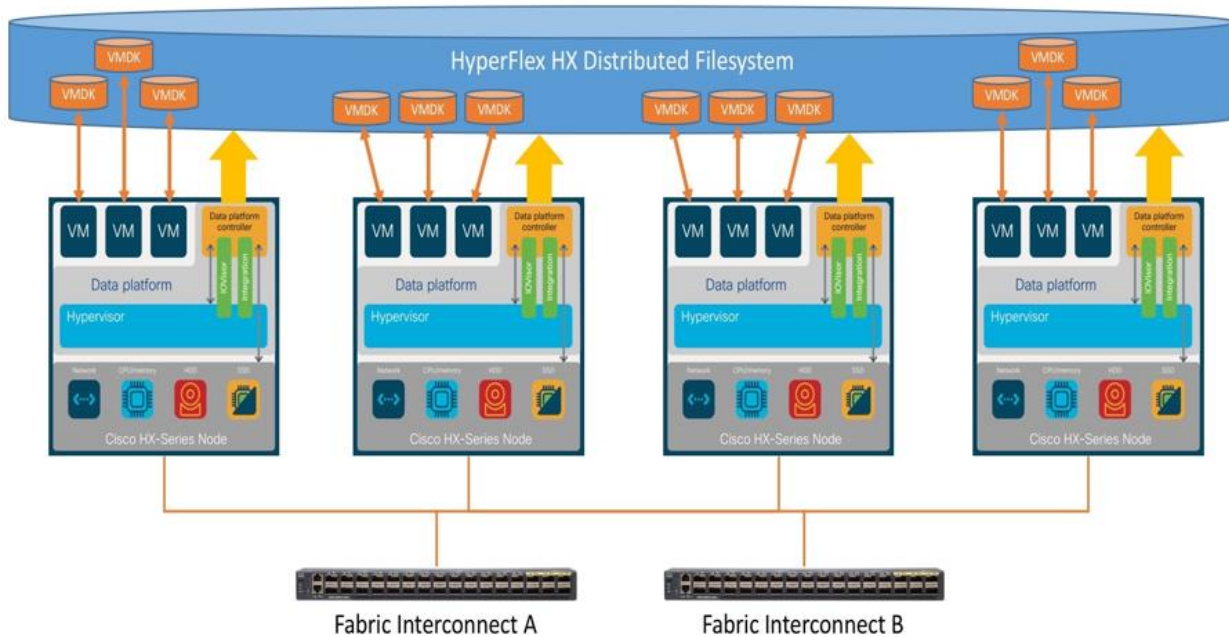
The following are the components of a Cisco HyperFlex system using the VMware ESXi Hypervisor:

- One pair of Cisco UCS Fabric Interconnects, choose from models: Cisco UCS 6454 Fabric Interconnect
- Eight Cisco HyperFlex HX-Series Rack-Mount Servers, choose from models:
  - Cisco HyperFlex HXAF220c-M5SX All-Flash Rack-Mount Servers
- Eight Cisco UCS C220 M5 Rack Servers for Compute-Only nodes
- Eight Cisco UCS B200 M5 Blade Servers for Compute-Only nodes
- Cisco HyperFlex Data Platform Software
- VMware vSphere ESXi Hypervisor
- VMware vCenter Server (end-user supplied)
- VMware Horizon RDSH Server Sessions & Desktops

Technology Overview

Cisco Unified Computing System

Cisco Unified Computing System (Cisco UCS) 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, 25 Gigabit Ethernet or 40 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, 25-Gbps or 40-Gbps unified network fabric, with an option for 100-Gbps uplinks. This network foundation consolidates LANs, SANs, and high-performance computing networks which are often 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 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 their choice of storage protocol and physical architecture, and enhanced 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. Cisco UCS can also be managed by Cisco Intersight, a cloud-based management and monitoring platform which offers a single pane of glass portal for multiple Cisco UCS systems across multiple locations.

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.

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

Cisco UCS Fabric Interconnect

The Cisco UCS Fabric Interconnect (FI) is a core part of the Cisco Unified Computing System, providing both network connectivity and management capabilities for the system. Depending on the model chosen, the Cisco UCS Fabric Interconnect offers line-rate, low-latency, lossless Ethernet, Fibre Channel over Ethernet (FCoE) and Fibre Channel connectivity. Cisco UCS Fabric Interconnects provide the management and communication back-bone for the Cisco UCS C-Series, Cisco UCS S-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 Fabric Interconnects become part of a single, highly available management domain. In addition, by supporting unified fabrics, the Cisco UCS Fabric Interconnects provide both the LAN and SAN connectivity for all servers within its domain. The product family supports Cisco low-latency, lossless Ethernet unified network fabric capabilities, which increase the reliability, efficiency, and scalability of Ethernet networks. The Fabric Interconnect supports multiple traffic classes over the Ethernet fabric from the servers to the uplinks. 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 6454 Fabric Interconnect

The Cisco UCS 6454 54-Port Fabric Interconnect is a One-Rack-Unit (1RU) 10/25/40/100 Gigabit Ethernet, FCoE and Fibre Channel switch offering up to 3.82 Tbps throughput and up to 54 ports. The switch has 28 10/25-Gbps Ethernet ports, 4 1/10/25-Gbps Ethernet ports, 6 40/100-Gbps Ethernet uplink ports and 16 uni-fied ports that can support 10/25-Gbps Ethernet ports or 8/16/32-Gbps Fibre Channel ports. All Ethernet ports are capable of supporting FCoE. Cisco HyperFlex nodes can connect at 10-Gbps or 25-Gbps speeds depending on the model of Cisco VIC card in the nodes and the SFP optics or cables chosen.

Figure 3   Cisco UCS 6454 Fabric Interconnect

Cisco UCS 64108 108-Port Fabric Interconnect

The Cisco UCS 64108 Fabric Interconnect (FI) is a 2-RU top-of-rack switch that mounts in a standard 19-inch rack such as the Cisco R Series rack. The 64108 is a 10/25/40/100 Gigabit Ethernet, FCoE and Fiber Channel switch offering up to 7.42 Tbps throughput and up to 108 ports. The switch has 16 unified ports (port numbers 1-16) that can support 10/25-Gbps SFP28 Ethernet ports or 8/16/32-Gbps Fibre Channel ports, 72 10/25-Gbps Ethernet SFP28 ports (port numbers 17-88), 8 1/10/25-Gbps Ethernet SFP28 ports (port numbers 89-96), and 12 40/100-Gbps Ethernet QSFP28 uplink ports (port numbers 97-108). All Ethernet ports are capable of supporting FCoE.

The Cisco UCS 64108 Fabric Interconnect also has one network management port, one console port for setting the initial configuration, and one USB port for saving or loading configurations. The FI also includes L1/L2 ports for connecting two fabric interconnects for high availability. For more information, see:
https://www.cisco.com/c/en/us/support/servers-unified-computing/ucs-64108-fabric-interconnect/model.html

Figure 4　Cisco UCS 64108 Fabric Interconnect Front and Back View

Front view

Rear view

Figure 3.
Cisco UCS 64108 (2 RU) 108-Port Fabric Interconnect

> The Cisco 64108 108-Port Fabric Interconnect two-rack-unit (2RU) is also available for higher port den-
> sity requirements.

Cisco HyperFlex HX-Series Nodes

A standard HyperFlex cluster requires a minimum of three HX-Series **"converged"** nodes (such as nodes with shared disk storage). Data is replicated across at least two of these nodes, and a third node is required for con-tinuous 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 additional disks, up to the platform's physical limit, for long term storage and capacity.

Cisco HyperFlex HXAF220c-M5SX All-Flash Node

This small footprint Cisco HyperFlex all-flash model contains a 240 GB M.2 form factor solid-state disk (SSD) that acts as the boot drive, a 240 GB housekeeping SSD drive, either a single 375 GB Optane NVMe SSD, a 1.6 TB NVMe SSD or 1.6 TB SAS SSD write-log drive, and six to eight 960 GB or 3.8 TB SATA SSD drives for stor-age capacity. For configurations requiring self-encrypting drives, the caching SSD is replaced with an 800 GB SAS SED SSD, and the capacity disks are also replaced with 960 GB or 3.8 TB SED SSDs. Each HXAF220C M5 node supports 2xT4 GPU cards and supporting GPUs with HXAF240 could be substituted.

Figure 5　HXAF220c-M5SX All-Flash Node

> NVME SSDs are generally 2x of the enterprise caching SSDs. Please chose the appropriate disk sizes for
> Hyperflex all flash-nodes Optane NVMe SSD drives and caching drives.

> While the Optane and NVMe options can provide a higher level of performance, the partitioning of the
> three disk options is the same, therefore the amount of cache available on the system is the same re-

gardless of the model chosen. Caching amounts are not factored in as part of the overall cluster capacity, only the capacity disks contribute to total cluster capacity.

Cisco VIC 1457 MLOM Interface Cards

The Cisco UCS VIC 1387 Card is a dual-port Enhanced Quad Small Form-Factor Pluggable (QSFP+) 40-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. The VIC 1387 is used in conjunction with the Cisco UCS 6332 or 6332-16UP model Fabric Interconnects.

The Cisco UCS VIC 1457 is a quad-port Small Form-Factor Pluggable (SFP28) mLOM card designed for the M5 generation of Cisco UCS C-Series Rack Servers. The card supports 10-Gbps or 25-Gbps Ethernet and FCoE, where the speed of the link is determined by the model of SFP optics or cables used. The card can be configured to use a pair of single links, or optionally to use all four links as a pair of bonded links. The Cisco UCS VIC 1457 is used in conjunction with the Cisco UCS 6454 model Fabric Interconnect.

The mLOM is 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, each dynamically configured as either a network interface card (NICs) or host bus adapter (HBA). The personality of the interfaces is set programmatically using the service profile associated with the server. The number, type (NIC or HBA), identity (MAC address and World Wide Name [WWN]), failover policy, adapter settings, bandwidth, and quality-of-service (QoS) policies of the PCIe interfaces are all specified using the service profile.

Figure 6   Cisco VIC 1457 mLOM Card

Cisco HyperFlex Connect HTML5 Management Web Page

An HTML 5 based Web UI named HyperFlex Connect is available for use as the primary management tool for Cisco HyperFlex. 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. To use the HyperFlex Connect UI, connect using a web browser to the HyperFlex cluster IP address: http://<hx controller cluster ip>.

Figure 7   HyperFlex 24 Node Cluster is Created and Healthy

cisco   **HyperFlex Installer**   4.0(2a)   ESXI

Progress | Summary

Cluster Name HXCLST `ONLINE` `HEALTHY`

Version	4.0.2a-35199	vCenter Server	10.10.50.40
Cluster Management IP Address	10.10.50.200	vCenter Datacenter Name	HXCL
Cluster Data IP Address	10.10.52.200	vCenter Cluster Name	HXCL
Replication Factor	3	DNS Server(s)	10.10.51.21, 10.10.51.22
Available Capacity	12.9 TB	NTP Server(s)	10.10.50.253, 10.10.50.252

Servers

Model	Serial Number	Management Hypervisor	Management Storage Controller	Data Network Hypervisor	Data Network Storage Controller
HXAF220C-M5SX	WZP21490FRS	10.10.50.51	10.10.50.151	10.10.52.51	10.10.52.151
HXAF220C-M5SX	WZP21490FP2	10.10.50.52	10.10.50.152	10.10.52.52	10.10.52.152
HXAF220C-M5SX	WZP22120C8N	10.10.50.53	10.10.50.153	10.10.52.53	10.10.52.153
HXAF220C-M5SX	WZP21480PPA	10.10.50.54	10.10.50.154	10.10.52.54	10.10.52.154
HXAF220C-M5SX	WZP21490FR2	10.10.50.55	10.10.50.155	10.10.52.55	10.10.52.155
HXAF220C-M5SX	WZP22020D8P	10.10.50.56	10.10.50.158	10.10.52.56	10.10.52.158
HXAF220C-M5SX	WZP212416UQ	10.10.50.57	10.10.50.157	10.10.52.57	10.10.52.157
HXAF220C-M5SX	WZP220216VM	10.10.50.58	10.10.50.156	10.10.52.58	10.10.52.156
UCSC-C220-M5SX	WZP2208006S	10.10.50.67		10.10.52.67	
UCSC-C220-M5SX	WZP213717GV	10.10.50.68		10.10.52.68	
UCSC-C220-M5SX	WZP2208006B	10.10.50.69		10.10.52.69	
UCSC-C220-M5SX	WZP2208006Q	10.10.50.70		10.10.52.70	
UCSC-C220-M5SX	WZP2208006V	10.10.50.71		10.10.52.71	
UCSC-C220-M5SX	WZP2208005K	10.10.50.72		10.10.52.72	
UCSC-C220-M5SX	WZP2208074	10.10.50.73		10.10.52.73	
UCSC-C220-M5SX	WZP2208006C	10.10.50.74		10.10.52.74	
UCSB-B200-M5	FOX1813GDVE	10.10.50.75		10.10.52.75	
UCSB-B200-M5	FOX1813GDVE	10.10.50.76		10.10.52.76	
UCSB-B200-M5	FOX1813GDVE	10.10.50.77		10.10.52.77	
UCSB-B200-M5	FOX1813GDVE	10.10.50.78		10.10.52.78	
UCSB-B200-M5	FOX1813GDVE	10.10.50.79		10.10.52.79	
UCSB-B200-M5	FOX1813GDVE	10.10.50.80		10.10.52.80	
UCSB-B200-M5	FOX1813GDVE	10.10.50.81		10.10.52.81	
UCSB-B200-M5	FOX1813GDVE	10.10.50.82		10.10.52.82	

Back to Workflow Selection | Launch HyperFlex Connect

Figure 8   Hyperflex GUI Displays a 24 Node Cluster is Online and Healthy

Cisco Intersight Cloud Based Management
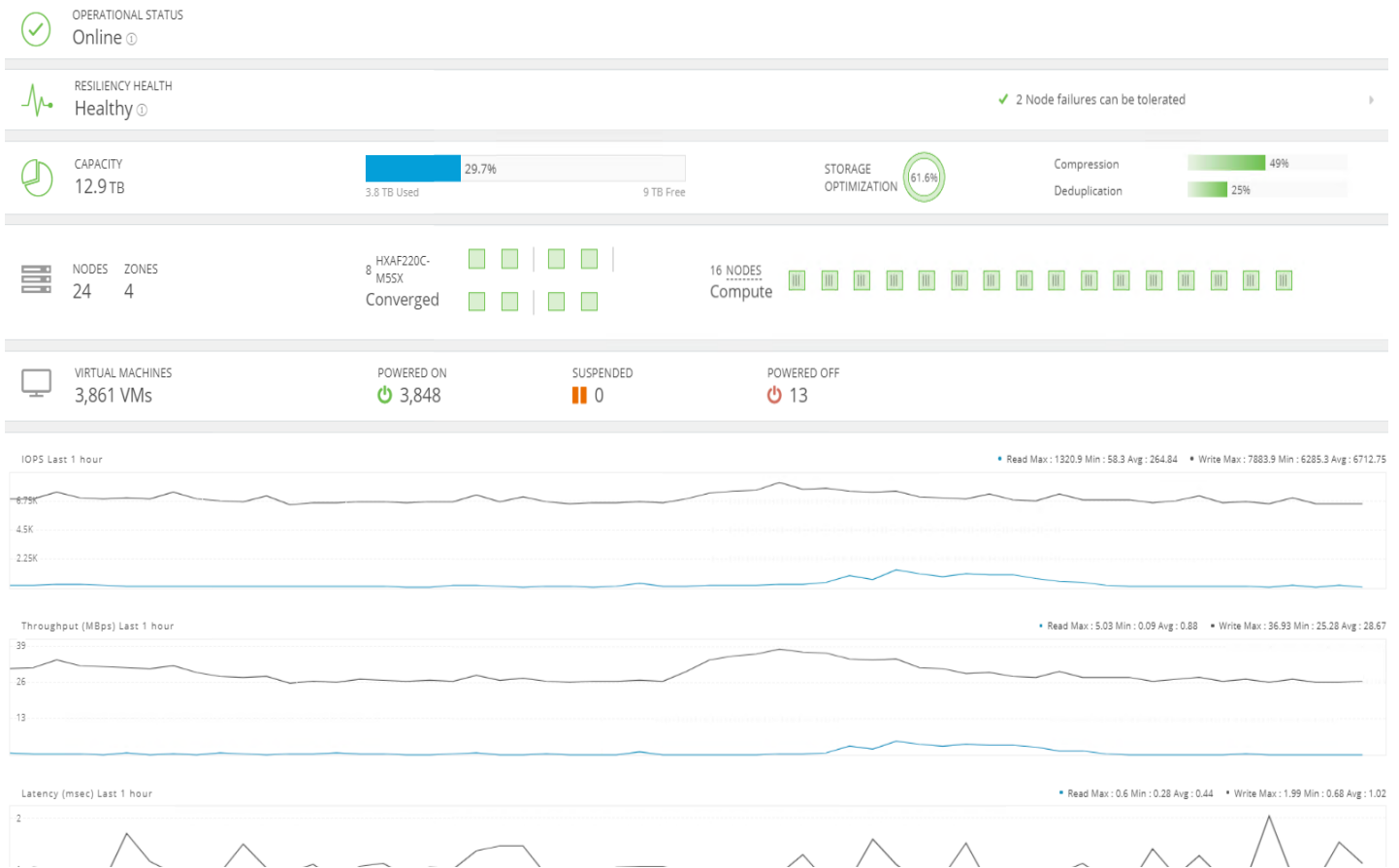
Cisco Intersight (https://intersight.com) is the latest visionary cloud-based management tool, designed to pro-vide a centralized off-site management, monitoring and reporting tool for all of your Cisco UCS based solutions, and can be used to deploy and manage Cisco HyperFlex clusters. Cisco Intersight offers direct links to Cisco UCS Manager and Cisco HyperFlex Connect for systems it is managing and monitoring. The Cisco Intersight website and framework is being constantly upgraded and extended with new and enhanced features inde-pendently of the products that are managed, meaning that many new features and capabilities can come with no downtime or upgrades required by the end users. This unique combination of embedded and online technolo-gies results in a complete cloud-based management solution that can care for Cisco HyperFlex throughout the entire lifecycle, from deployment through retirement.

Figure 9   Cisco Intersight
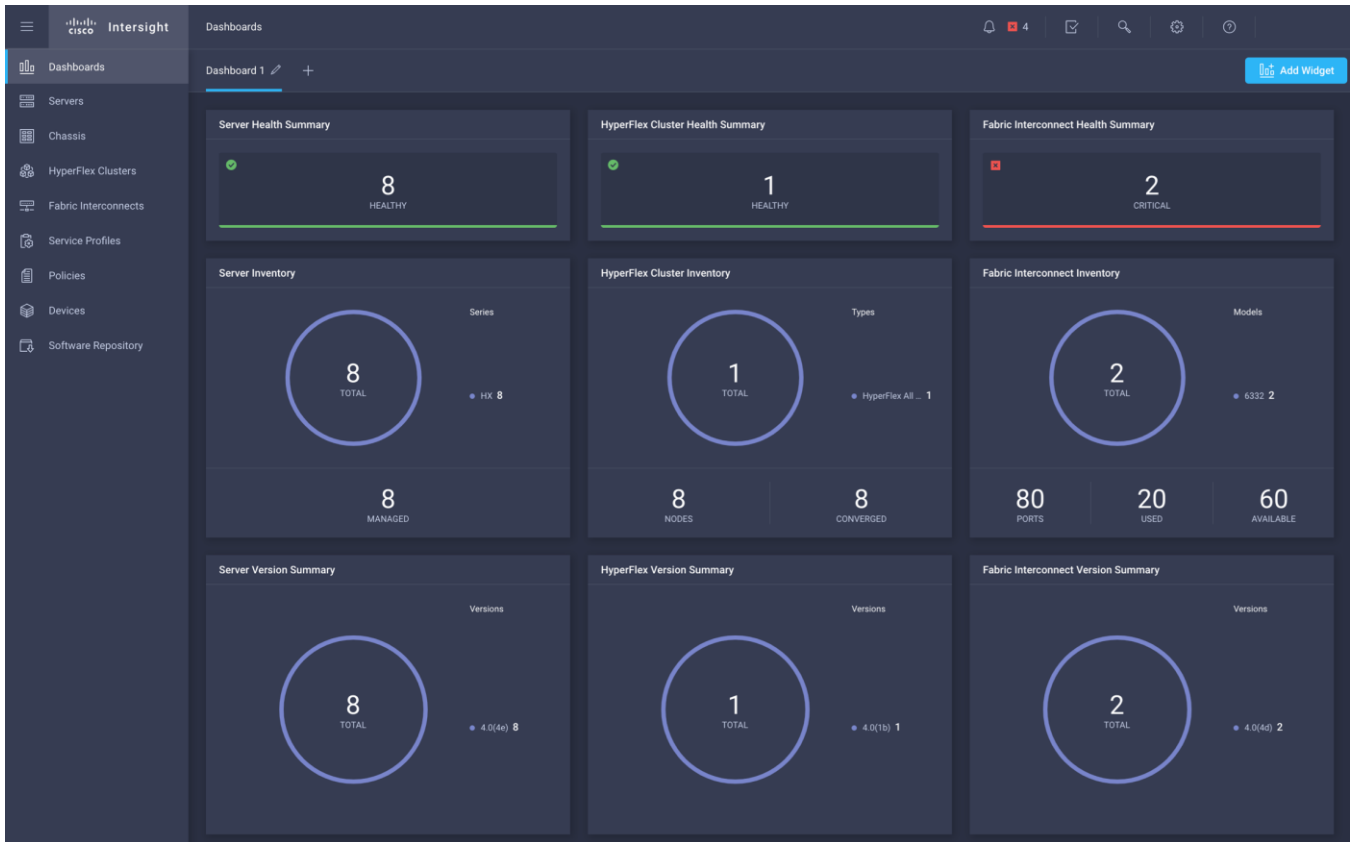
Cisco HyperFlex HX Data Platform Administration Plug-in

The Cisco HyperFlex HX Data Platform is also administered secondarily through a VMware vSphere web client plug-in, which is deployed automatically by the Cisco HyperFlex installer.

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 as software in user space within a virtual machine, and intercepts and handles all I/O from the guest virtual machines. The Storage Controller Virtual Machine (SCVM) uses the VMDirectPath I/O feature to provide direct PCI passthrough **control of the physical server's SAS disk controller**, or direct control of the PCI attached NVMe based SSDs. 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 or SDDs as a capacity layer for distributed storage. The controller integrates the data platform into the VMware vSphere cluster through the use of three preinstalled VMware ESXi vSphere Installation Bundles (VIBs) on each node:

- 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. The IO Visor intercepts guest VM IO traffic, and intelligently redirects it to the HyperFlex SCVMs.

- 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 via ma-

nipulation of the filesystem metadata rather than actual data copying, providing rapid response, and thus rapid deployment of new environments.

- stHypervisorSvc: This VIB adds enhancements and features needed for HyperFlex data protection and VM replication.

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.

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).

- 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 of 2 entire nodes in a cluster of 5 nodes or greater, without losing data and resorting to restore from backup or other recovery processes. RF3 is recommended for all production systems.

- 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 of 1 entire node without losing data and resorting to restore from backup or other recovery processes. RF2 is suitable for non-production systems, or environments where the extra data protection is not needed. HyperFlex stretched clusters use the RF2 setting, however there are 2 copies of the data kept in both halves of the cluster, so effectively there are four copies stored.

Data Write and Compression Operations

Internally, the contents of each virtual disk are subdivided and spread across multiple servers by the HXDP software. For each write operation, the data is intercepted by the IO Visor module on the node where the VM is running, a primary node is determined for that particular operation via a hashing algorithm, and then sent to the primary node via the network. The primary node compresses the data in real time, writes the compressed data to the write log on its caching SSD, and replica copies of that compressed data are sent via the network and written to the write log on the caching SSD of the remote nodes in the cluster, according to the replication factor setting. For example, at RF=3 a write operation will be written to write log of the primary node for that virtual disk address, and two additional writes will be committed in parallel on two other nodes. Because the virtual disk contents have been divided and spread out via the hashing algorithm for each unique operation, this method results in all writes being spread across all nodes, avoiding the problems with data locality and "noisy" VMs consuming all the IO capacity of a single node. 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. This process speeds up read requests when reads are requested of data that has recently been written.

Data Destaging and Deduplication

The Cisco HyperFlex HX Data Platform constructs multiple write log 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 nodes for the Hybrid system or to the SSD capacity layer of the nodes for the All-Flash or All-NVMe systems. During the destaging process, data is deduplicated before being written to the capacity storage layer, and the resulting data can now be written to the HDDs or SDDs of the server. On hybrid systems, the now deduplicated and compressed data is also written to the dedicated read cache area of the caching SSD, which speeds up read requests of data that has recently been written. When the data is destaged to the capacity disks, it is written in a single sequential operation, avoiding disk head seek thrashing on the spinning disks and accomplishing the task in the minimal amount of time. 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, which benefits both the HDD and SDD configurations.

Figure 10 HyperFlex HX Data Platform Data Movement
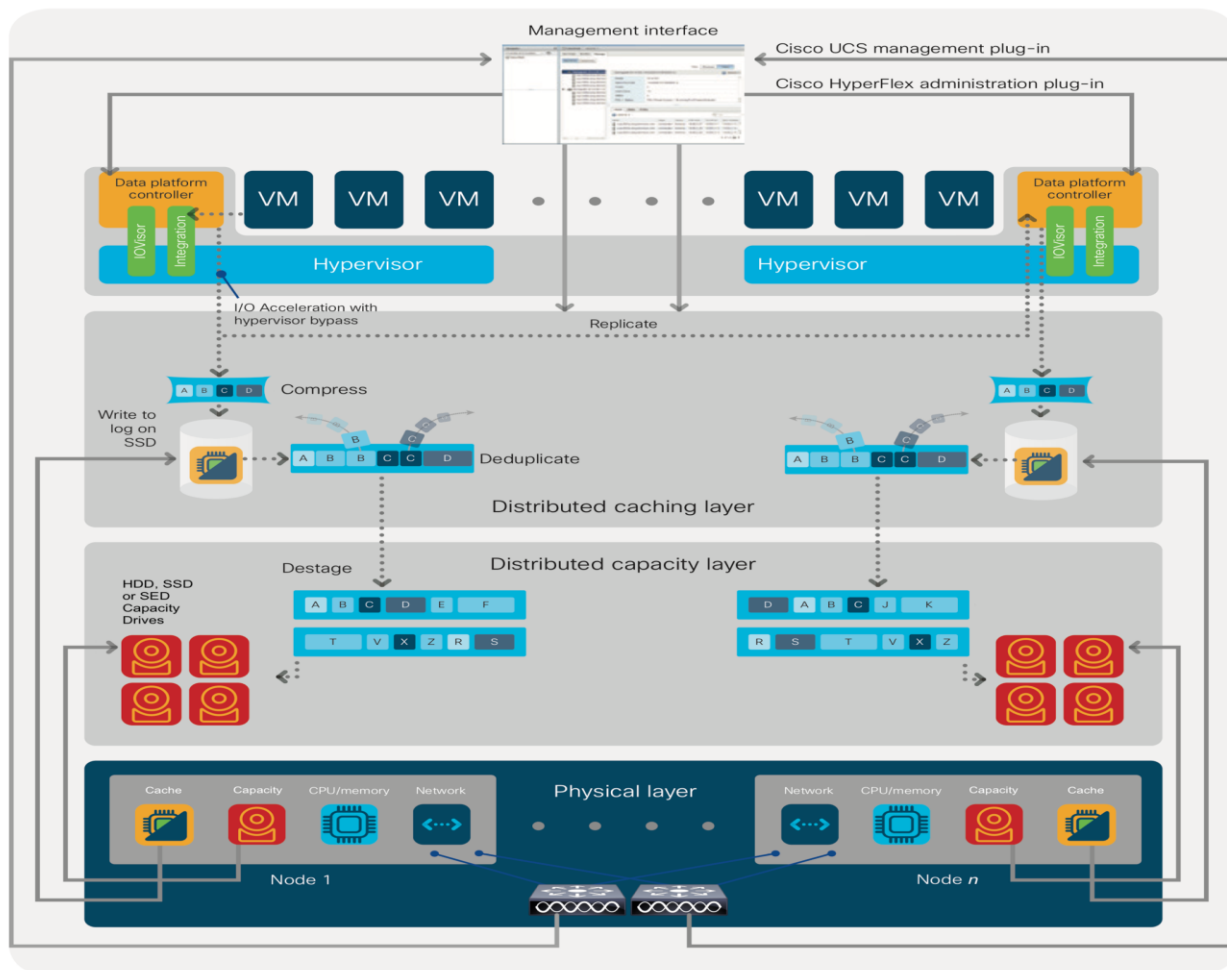
Data Read Operations

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 layer disk. 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 caching SSDs of hybrid nodes. Finally, if the data has not been accessed in a significant amount of time, the filesystem will retrieve the requested data from the distributed capacity layer. As requests for reads are made to the distributed filesystem and the data is retrieved from the capacity layer, the caching SSDs of hybrid nodes 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, ensures that data is served at the highest possible speed, leveraging the caching SSDs of the nodes fully and equally. All-flash and all-NVMe configurations do not employ a dedicated read cache, because such caching does not provide any performance benefit since the persistent data copy already resides on high-performance SSDs.

In summary, the Cisco HyperFlex HX Data Platform implements a distributed, log-structured file system that performs data operations via two configurations:

- In a Hybrid configuration, the data platform provides a caching layer using SSDs to accelerate read requests and write responses, and it implements a storage capacity layer using HDDs.

- In an All-Flash or all-NVMe configuration, the data platform provides a dedicated caching layer using high endurance SSDs to accelerate write responses, and it implements a storage capacity layer also using SSDs. Read requests are fulfilled directly from the capacity SSDs, as a dedicated read cache is not needed to accelerate read operations.

Data Optimization

The Cisco HyperFlex HX Data Platform provides finely detailed inline deduplication and variable block inline compression that is always on for objects in the cache (SSD and memory) and capacity (SSD or HDD) layers. Unlike other solutions, which require you to turn off these features to maintain performance, the deduplication and compression capabilities in the Cisco data platform are designed to sustain and enhance performance and significantly reduce physical storage capacity requirements.

Data Deduplication

Data deduplication is used on all storage in the cluster, including memory and SSD drives. Based on a patent-pending Top-K Majority algorithm, the platform uses conclusions from empirical research that show that most data, when sliced into small data blocks, has significant deduplication potential based on a minority of the data blocks. By fingerprinting and indexing just these frequently used blocks, high rates of deduplication can be achieved with only a small amount of memory, which is a high-value resource in cluster nodes (Figure 11).

Figure 11 Cisco HyperFlex HX Data Platform Optimizes Data Storage with No Performance Impact
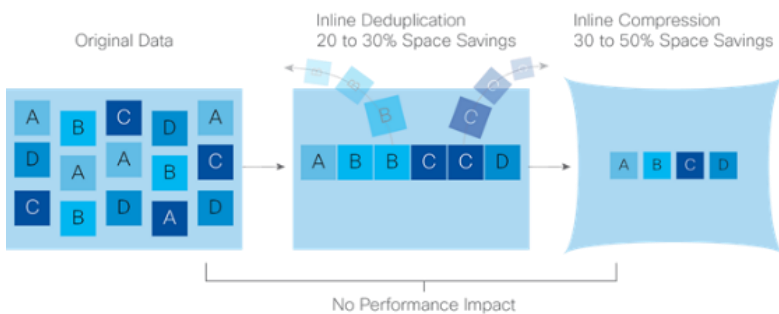
Inline Compression

The Cisco HyperFlex HX Data Platform uses high-performance inline compression on data sets to save storage capacity. Although other products offer compression capabilities, many negatively affect performance. In contrast, the Cisco data platform uses CPU-offload instructions to reduce the performance impact of compression operations. In addition, the log-structured distributed-objects layer has no effect on modifications (write operations) to previously compressed data. Instead, incoming modifications are compressed and written to a new location, and the existing (old) data is marked for deletion, unless the data needs to be retained in a snapshot.

The data that is being modified does not need to be read prior to the write operation. This feature avoids typical read-modify-write penalties and significantly improves write performance.

Log-Structured Distributed Objects

In the Cisco HyperFlex HX Data Platform, the log-structured distributed-object store layer groups and compresses data that filters through the deduplication engine into self-addressable objects. These objects are written to disk in a log-structured, sequential manner. All incoming I/O—including random I/O—is written sequentially to both the caching (SSD and memory) and persistent (SSD or HDD) tiers. The objects are distributed across all nodes in the cluster to make uniform use of storage capacity.

By using a sequential layout, the platform helps increase flash-memory endurance. Because read-modify-write operations are not used, there is little or no performance impact of compression, snapshot operations, and cloning on overall performance.

Data blocks are compressed into objects and sequentially laid out in fixed-size segments, which in turn are sequentially laid out in a log-structured manner ([Figure 12](#)). Each compressed object in the log-structured segment is uniquely addressable using a key, with each key fingerprinted and stored with a checksum to provide high levels of data integrity. In addition, the chronological writing of objects helps the platform quickly recover from media or node failures by rewriting only the data that came into the system after it was truncated due to a failure.

Figure 12   HyperFlex HX Data Platform Optimizes Data Storage with No Performance Impact

Self-Describing Variably Sized Compressed Objects
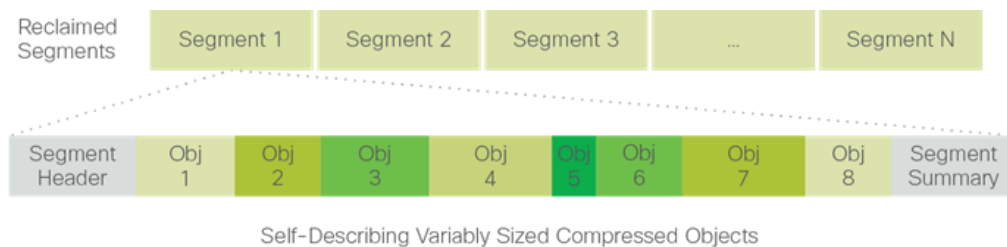
Encryption

Securely encrypted storage optionally encrypts both the caching and persistent layers of the data platform. Integrated with enterprise key management software, or with passphrase-protected keys, encrypting data at rest helps you comply with HIPAA, PCI-DSS, FISMA, and SOX regulations. The platform itself is hardened to Federal Information Processing Standard (FIPS) 140-1 and the encrypted drives with key management comply with the FIPS 140-2 standard.

Data Services

The Cisco HyperFlex HX Data Platform provides a scalable implementation of space-efficient data services, including thin provisioning, space reclamation, pointer-based snapshots, and clones—without affecting performance.

Thin Provisioning

The platform makes efficient use of storage by eliminating the need to forecast, purchase, and install disk capacity that may remain unused for a long time. Virtual data containers can present any amount of logical space to applications, whereas the amount of physical storage space that is needed is determined by the data that is written. You can expand storage on existing nodes and expand your cluster by adding more storage-intensive nodes as your business requirements dictate, eliminating the need to purchase large amounts of storage before you need it.

Snapshots

The Cisco HyperFlex HX Data Platform uses metadata-based, zero-copy snapshots to facilitate backup operations and remote replication: critical capabilities in enterprises that require always-on data availability. Space-efficient snapshots allow you to perform frequent online data backups without worrying about the consumption of physical storage capacity. Data can be moved offline or restored from these snapshots instantaneously.

- Fast snapshot updates: When modified-data is contained in a snapshot, it is written to a new location, and the metadata is updated, without the need for read-modify-write operations.

- Rapid snapshot deletions: You can quickly delete snapshots. The platform simply deletes a small amount of metadata that is located on an SSD, rather than performing a long consolidation process as needed by solutions that use a delta-disk technique.

- Highly specific snapshots: With the Cisco HyperFlex HX Data Platform, you can take snapshots on an individual file basis. In virtual environments, these files map to drives in a virtual machine. This flexible specificity allows you to apply different snapshot policies on different virtual machines.

Many basic backup applications read the entire dataset, or the changed blocks since the last backup at a rate that is usually as fast as the storage, or the operating system can handle. This can cause performance implications since HyperFlex is built on Cisco UCS with 40GbE that could result in multiple gigabytes per second of backup throughput. These basic backup applications, such as Windows Server Backup, should be scheduled during off-peak hours, particularly the initial backup if the application lacks some form of change block tracking.

Full featured backup applications, such as Veeam Backup and Replication v9.5, have the ability to limit the amount of throughput the backup application can consume which can protect latency sensitive applications during the production hours. With the release of v9.5 update 2, Veeam is the first partner to integrate HX native snapshots into the product. HX Native snapshots do not suffer the performance penalty of delta-disk snapshots, and do not require heavy disk IO impacting consolidation during snapshot deletion.

Particularly important for SQL administrators is the Veeam Explorer for SQL which can provide transaction level recovery within the Microsoft VSS framework. The three ways Veeam Explorer for SQL Server works to restore SQL Server databases include; from the backup restore point, from a log replay to a point in time, and from a log replay to a specific transaction – all without taking the VM or SQL Server offline.

Fast, Space-Efficient Clones

In the Cisco HyperFlex HX Data Platform, clones are writable snapshots that can be used to rapidly provision items such as virtual desktops and applications for test and development environments. These fast, space-efficient clones rapidly replicate storage volumes so that virtual machines can be replicated through just metadata operations, with actual data copying performed only for write operations. With this approach, hundreds of clones can be created and deleted in minutes. Compared to full-copy methods, this approach can save a significant amount of time, increase IT agility, and improve IT productivity.

Clones are deduplicated when they are created. When clones start diverging from one another, data that is common between them is shared, with only unique data occupying new storage space. The deduplication engine eliminates data duplicates in the diverged clones to further reduce the clone's storage footprint.

Data Replication and Availability

In the Cisco HyperFlex HX Data Platform, the log-structured distributed-object layer replicates incoming data, improving data availability. Based on policies that you set, data that is written to the write cache is synchronously replicated to one or two other SSD drives located in different nodes before the write operation is acknowledged to the application. This approach allows incoming writes to be acknowledged quickly while protecting data from SSD or node failures. If an SSD or node fails, the replica is quickly re-created on other SSD drives or nodes using the available copies of the data.

The log-structured distributed-object layer also replicates data that is moved from the write cache to the capacity layer. This replicated data is likewise protected from SSD or node failures. With two replicas, or a total of three data copies, the cluster can survive uncorrelated failures of two SSD drives or two nodes without the risk of data loss. Uncorrelated failures are failures that occur on different physical nodes. Failures that occur on the same node affect the same copy of data and are treated as a single failure. For example, if one disk in a node fails and subsequently another disk on the same node fails, these correlated failures count as one failure in the system. In this case, the cluster could withstand another uncorrelated failure on a different node. See the Cisco HyperFlex HX Data Platform system administrator's guide for a complete list of fault-tolerant configurations and settings.

If a problem occurs in the Cisco HyperFlex HX controller software, data requests from the applications residing in that node are automatically routed to other controllers in the cluster. This same capability can be used to upgrade or perform maintenance on the controller software on a rolling basis without affecting the availability of the cluster or data. This self-healing capability is one of the reasons that the Cisco HyperFlex HX Data Platform is well suited for production applications.

In addition, native replication transfers consistent cluster data to local or remote clusters. With native replication, you can snapshot and store point-in-time copies of your environment in local or remote environments for backup and disaster recovery purposes.

Data Rebalancing

A distributed file system requires a robust data rebalancing capability. In the Cisco HyperFlex HX Data Platform, no overhead is associated with metadata access, and rebalancing is extremely efficient. Rebalancing is a non-disruptive online process that occurs in both the caching and persistent layers, and data is moved at a fine level of specificity to improve the use of storage capacity. The platform automatically rebalances existing data when nodes and drives are added or removed or when they fail. When a new node is added to the cluster, its capacity and performance is made available to new and existing data. The rebalancing engine distributes existing data to the new node and helps ensure that all nodes in the cluster are used uniformly from capacity and performance perspectives. If a node fails or is removed from the cluster, the rebalancing engine rebuilds and distributes copies of the data from the failed or removed node to available nodes in the clusters.

Online Upgrades

Cisco HyperFlex HX-Series systems and the HX Data Platform support online upgrades so that you can expand and update your environment without business disruption. You can easily expand your physical resources; add processing capacity; and download and install BIOS, driver, hypervisor, firmware, and Cisco UCS Manager updates, enhancements, and bug fixes.

Cisco Nexus 93180 Switches

The Cisco Nexus 93180YC-FX Switches has 48 10/25-Gbps Small Form Pluggable Plus (SFP+) ports and 6 Quad 40/100-Gbps SFP+ (QSFP+) uplink ports. All the ports are line rate, delivering 3.6 Tbps of throughput in a 1-rack-unit (1RU) form factor. Cisco Nexus 93180-YC-FX benefits are listed below:

- Specifications at-a-Glance
  - 1 rack unit (1RU)
  - 48 x 1/10/25-Gbps fiber ports
  - 6 x 40/100-Gbps QSFP28 ports
  - Up to 3.6 Tbps of bandwidth
- Architectural Flexibility
  - Leaf-node support for Cisco ACI architecture with flexible port configuration
  - Seamless convergence thanks to 48 downlink ports that can work as 1/10/25-Gbps Ethernet or FCoE ports or as 8/16/32-Gbps Fibre Channel ports
  - Easy migration with 6 uplink ports that can be configured as 40/100-Gbps Ethernet or FCoE ports
- Feature Rich
  - Automated policy-based systems management with Cisco ACI
  - Open APIs enable third-party integration with our partners
  - Better management of speed mismatch between access and uplink ports with 40 MB of shared buffer space
  - Support for Fibre Channel interfaces for back-end storage connectivity
- Highly Available and Efficient Design
  - High-performance, non-blocking architecture
  - Easily deployed into either a hot-aisle or a cold-aisle configuration
  - Redundant, hot-swappable power supplies and fan trays
- Simplified Operations
  - Automate IT workflows and shorten app deployment from weeks to minutes
- Top-notch Security
  - Whitelist model, policy enforcement and application security with Cisco ACI micro-segmentation
  - Wire-rate MAC sec encryption on all ports
- Real-time Visibility and Telemetry
  - Built-in Cisco Tetration sensors for rich traffic-flow telemetry and line-rate data collection
  - Get actionable insights in less than 1 second
  - Get visibility into everything in your data center
- Investment Protection
  - Flexible migration options with support for 10-Gbps and 25-Gbps access connectivity and 40-Gbps and 100-Gbps uplinks

- ◦ Cisco's 40-Gbps bidirectional transceiver allows for reuse of an existing 10 Gigabit Ethernet multimode cabling plant for 40 Gigabit Ethernet
- Resources
  - ◦ [Cisco Nexus 9300-EX and 9300-FX Platform Leaf](#)
  - ◦ [Switches for Cisco Application Centric Infrastructure Data Sheet](#)

Figure 13 Cisco Nexus 93180YC-FX Switch

VMware vSphere 6.7

VMware provides virtualization software. VMware's enterprise software hypervisors for servers—VMware vSphere ESX, vSphere ESXi, and vSphere—are bare-metal hypervisors that run directly on server hardware without requiring an additional underlying operating system. VMware vCenter Server for vSphere provides central management and complete control and visibility into clusters, hosts, virtual machines, storage, networking, and other critical elements of your virtual infrastructure.

VMware vSphere 6.5 introduces many enhancements to vSphere Hypervisor, VMware virtual machines, vCenter Server, virtual storage, and virtual networking, further extending the core capabilities of the vSphere platform.

What's New

The new enhancements in this release are as follows:

- ixgben driver enhancements: The ixgben driver adds queue pairing to optimize CPU efficiency.

- VMXNET3 enhancements: ESXi 6.7 Update 3 adds guest encapsulation offload and UDP, and ESP RSS support to the Enhanced Networking Stack (ENS). Checksum calculations are offloaded from encapsulated packets to the virtual device emulation and you can run RSS on UDP and ESP packets on demand. UDP RSS supports both IPv4 and IPv6, while ESP RSS supports only IPv4. The feature requires a corresponding VMXNET3 v4 driver.

- NVIDIA virtual GPU (vGPU) enhancements: With vCenter Server 6.7 Update 3, you can configure virtual machines and templates with up to four vGPU devices to cover use cases requiring multiple GPU accelerators attached to a virtual machine. To use the vMotion vGPU feature, you must set the vgpu.hotmigrate.enabled advanced setting to true and make sure that both your vCenter Server and ESXi hosts are running vSphere 6.7 Update 3.
  vMotion of multi GPU-accelerated virtual machines might fail gracefully under heavy GPU workload due to the maximum switch over time of 100 secs. To avoid this failure, either increase the maximum allowable switch over time or wait until the virtual machine is performing a less intensive GPU workload.

- bnxtnet driver enhancements: ESXi 6.7 Update 3 adds support for Broadcom 100G Network Adapters and multi-RSS feed to the bnxtnet driver.

- QuickBoot support enhancements: ESXi 6.7 Update 3 whitelists Dell EMC vSAN Ready Nodes servers R740XD and R640 for QuickBoot support.

- Configurable shutdown time for the sfcbd service: With ESXi 6.7 Update 3, you can configure the shutdown time of the sfcbd service. You can set a shutdown time depending on the third-party CIM provider that you use, or else use the default setting of 10 sec.

- New SandyBridge microcode: ESXi 6.7 Update 3 adds new SandyBridge microcode to the cpu-microcode VIB to bring SandyBridge security up to par with other CPUs and fix per-VM Enhanced vMotion Compatibility (EVC) support. For more information, see VMware knowledge base article [1003212](#).

VMware vCenter Server

VMware vCenter Server provides the following features:

- Migration Tool

- Improved appliance management

- Native high availability

- Native backup and restore

- There are also general improvements to vCenter Server 6.7, including the vSphere Web Client and the fully supported HTML5-based vSphere Client.

VMware ESXi 6.7 Hypervisor

VMware ESXi 6.7 Hypervisor provides the following enhancements:

- With vSphere 6.7, administrators can find significant improvement in patching, upgrading, and managing configuration of ESXi hosts through vSphere Update Manager that is enabled by default.

- VMware tool and virtual hardware upgrade

- Improvement in Host Profile, as well as in day to day operations

- Improvement in manageability and configuration rules for Auto-Deploy

- Enhanced monitoring added option to monitor GPU usage.

- Dedicated Gateways for VMkernel Network Adapter

- VMware vSphere Storage I/O Control Using Storage Policy Based Management

VMware Horizon

VMware Horizon desktop virtualization solutions built on a unified architecture, so they are simple to manage and flexible enough to meet the needs of all your organization's users. You use the same architecture and management tools to manage public, private, and hybrid cloud deployments as you do for on premises deployments

- VMware Horizon Virtual machines and RDSH known as server-based hosted sessions: These are applications hosted from Microsoft Windows servers to any type of device, including Windows PCs, Macs, smartphones, and tablets. Some VMware editions include technologies that further optimize the experience of using Windows applications on a mobile device by automatically translating native mobile-device display, navigation, and controls to Windows applications; enhancing performance over mobile networks; and enabling developers to optimize any custom Windows application for any mobile environment.

- VMware Horizon RDSH session users also known as server-hosted desktops: These are inexpensive, locked-down Windows virtual desktops hosted from Windows server operating systems. They are well suited for users, such as call center employees, who perform a standard set of tasks.

Advantages of Using VMware Horizon

VMware Horizon 7 version 7.12 provides the following new features and enhancements:

- Instant Clones
  - A new type of desktop virtual machines that can be provisioned significantly faster than the traditional View Composer linked clones.
  - A fully functional desktop can be provisioned in two seconds or less.
  - Recreating a desktop pool with a new OS image can be accomplished in a fraction of the time it takes a View Composer desktop pool because the parent image can be prepared well ahead of the scheduled time of pool recreation.
  - Clones are automatically rebalanced across available datastores.
  - View storage accelerator is automatically enabled.
  - You can use NVIDIA GRID vGPUs with instant-clone desktop pools. Configuring PCoIP as the display protocol with NVIDIA GRID vGPU is a technical preview feature.
  - You can select multiple vLAN networks to create a larger instant-clone desktop pool. Only the static port group is supported.
  - You can use the internal VM debug mode to troubleshoot internal virtual machines in an instant-clone desktop pool or in an instant-clone farm.
  - Administrators can perform a restart or reset of the virtual desktops managed by the vCenter Server.
  - You can perform maintenance on instant-clone virtual machines by putting the ESXi hosts into maintenance mode. Use vSphere Web Client to put the ESXi host into maintenance mode. The ESXi host maintenance operation automatically deletes the parent virtual machines from that ESXi host.

- VMware Blast Extreme
  - VMware Blast Extreme is now fully supported on the Horizon platform.
  - Connections to physical machines that have no monitors attached are supported with NVIDIA graphics cards. This is a technical preview feature for Horizon 7 version 7.12.
  - The Blast Secure Gateway includes Blast Extreme Adaptive Transport (BEAT) networking, which dynamically adjusts to network conditions such as varying speeds and packet loss.
  - Administrators can select the VMware Blast display protocol as the default or available protocol for pools, farms, and entitlements.
  - End users can select the VMware Blast display protocol when connecting to remote desktops and applications.
  - VMware Blast Extreme features include:
    - TCP and UDP transport support
    - H.264 support for the best performance across more devices
    - Reduced device power consumption for longer battery life

- NVIDIA GRID acceleration for more graphical workloads per server, better performance, and a superior remote user experience
- True SSO
  - For VMware Identity Manager Integration, True SSO streamlines the end-to-end login experience. After users log in to VMware Identity Manager using a smart card or an RSA SecurID or RADIUS token, users are not required to also enter Active Directory credentials in order to use a remote desktop or application.
  - Uses a short-lived Horizon virtual certificate to enable a password-free Windows login.
  - Supports using either a native Horizon Client or HTML Access.
  - System health status for True SSO appears in the Horizon Administrator dashboard.
  - Can be used in a single domain, in a single forest with multiple domains, and in a multiple-forest, multiple-domain setup.
- Smart Policies
  - Control of the clipboard cut-and-paste, client drive redirection, USB redirection, and virtual printing desktop features through defined policies.
  - PCoIP session control through PCoIP profiles.
  - Conditional policies based on user location, desktop tagging, pool name, and Horizon Client registry values.
- Configure the Clipboard Memory Size for VMware Blast and PCoIP Sessions
  Horizon administrators can configure the server clipboard memory size by setting GPOs for VMware Blast and PCoIP sessions. Horizon Client 4.1 users on Windows, Linux, and Mac OS X systems can configure the client clipboard memory size. The effective memory size is the lesser of the server and client clipboard memory size values.
- VMware Blast Network Recovery Enhancements
  Network recovery is now supported for VMware Blast sessions initiated from iOS, Android, Mac OS X, Linux, and Chrome OS clients. Previously, network recovery was supported only for Windows client sessions. If you lose your network connection unexpectedly during a VMware Blast session, Horizon Client attempts to reconnect to the network, and you can continue to use your remote desktop or application. The network recovery feature also supports IP roaming, which means you can resume your VMware Blast session after switching to a WiFi network.
- Configure Horizon Administrator to not remember the login name
  Horizon administrators can configure not to display the Remember user name check box and therefore not remember the administrator's login name.
- Allow Mac OS X Users to Save Credentials
  Horizon administrators can configure Connection Server to allow Horizon Client Mac OS X systems to remember a user's user name, password, and domain information. If users choose to have their credentials saved, the credentials are added to the login fields in Horizon Client on subsequent connections.
- Windows 10
  - Windows 10 is supported as a desktop guest operating system
  - Horizon Client runs on Windows 10
  - Smart card is supported on Windows 10

- The Horizon User Profile Migration tool migrates Windows 7, 8/8.1, Server 2008 R2, or Server 2012 R2 user profiles to Windows 10 user profiles.

- RDS Desktops and Hosted Apps

  - Graphics Support. Existing 3D vDGA and GRID vGPU graphics solutions on VDI desktops have been extended to RDS hosts, enabling graphics-intensive applications to run on RDS desktops and Hosted Apps.

  - Enhanced Load Balancing. A new capability provides load balancing of server farm applications based on memory and CPU resources.

  - One-Way AD Trusts
    One-way AD trust domains are now supported. This feature enables environments with limited trust relationships between domains without requiring Horizon Connection Server to be in an external domain.

- Cloud Pod Architecture (CPA) Enhancements

  - Hosted App Support. Support for application remoting allows applications to be launched using global entitlements across a pod federation.

  - HTML Access (Blast) Support. Users can use HTML Access to connect to remote desktops and applications in a Cloud Pod Architecture deployment.

- Access Point Integration

  - Access Point is a hardened Linux-based virtual appliance that protects virtual desktop and application resources to allow secure remote access from the Internet. Access Point provides a new authenticating DMZ gateway to Horizon Connection Server. Smart card support on Access Point is available as a Tech Preview. Security server will continue to be available as an alternative configuration. For more information, see [Deploying and Configuring Access Point](#).

- FIPS

  - Install-time FIPS mode allows customers with high security requirements to deploy Horizon 6.

- Graphics Enhancements

  - AMD vDGA enables vDGA pass-through graphics for AMD graphics hardware.

  - 4K resolution monitors (3840x2160) are supported.

- Horizon Administrator Enhancements

  - Horizon Administrator shows additional licensing information, including license key, named user and concurrent connection user count.

  - Pool creation is streamlined by letting Horizon administrators clone existing pools.

- Horizon 7 for Linux Desktop Enhancements

  - Support for managed virtual machines

  - Support for smart card redirection with SSO

  - Support for Horizon Client for iOS

  - Support for SLES 12 SP1

  - Support for H.264 encoder software

  - True SSO support on the RHEL/CentOS 7 desktops
    True single sign-on (SSO) is now supported on RHEL 7 and CentOS desktops.

- ◦ Additional supported platforms
  Support the Ubuntu 18.04, RHEL/CentOS 6.10, and RHEL/CentOS 7.5 platforms have been added.
- ◦ Instant-clone support for RHEL 7.1 and later versions
  You can now create instant-clone floating desktop pool on systems with RHEL 7.1 or later installed.

- Additional Features
  - ◦ Support for IPv6 with VMware Blast Extreme on security servers.
  - ◦ Horizon Administrator security protection layer. See VMware Knowledge Base (KB) article 2144303 for more information.
  - ◦ Protection against inadvertent pool deletion.
  - ◦ RDS per-device licensing improvements.
  - ◦ Support for Intel vDGA.
  - ◦ Support for AMD Multiuser GPU Using vDGA.
  - ◦ More resilient upgrades.
  - ◦ Display scaling for Windows Horizon Clients.
  - ◦ DPI scaling is supported if it is set at the system level and the scaling level is greater than 100.

- vSphere Support
  - ◦ vSphere 6.7 U2 is supported.

- Virtual Desktops
  - ◦ vMotion support for vGPU-enabled virtual desktops

What are VMware RDS Hosted Sessions?

The following describes the VMware RDS Hosted Sessions:

- An RDS host is a server computer that hosts applications and desktop sessions for remote access. An RDS host can be a virtual machine or a physical server.

- An RDS host has the Microsoft Remote Desktop Services role, the Microsoft Remote Desktop Session Host service, and Horizon Agent installed. Remote Desktop Services was previously known as Terminal Services. The Remote Desktop Session Host service allows a server to host applications and remote desktop sessions. With Horizon Agent installed on an RDS host, users can connect to applications and desktop sessions by using the display protocol PCoIP or Blast Extreme. Both protocols provide an optimized user experience for the delivery of remote content, including images, audio, and video.

- The performance of an RDS host depends on many factors. For information on how to tune the performance of different versions of Windows Server, see http://msdn.microsoft.com/library/windows/hardware/gg463392.aspx.

- Horizon 7 supports at most one desktop session and one application session per user on an RDS host.

- When users submit print jobs concurrently from RDS desktops or applications that are hosted on the same RDS host, the ThinPrint server on the RDS host processes the print requests serially rather than in parallel. This can cause a delay for some users. Note that the print server does not wait for a print job to complete before processing the next one. Print jobs that are sent to different printers will print in parallel.

- If a user launches an application and also an RDS desktop, and both are hosted on the same RDS host, they share the same user profile. If the user launches an application from the desktop, conflicts may result if both applications try to access or modify the same parts of the user profile, and one of the applications may fail to run properly.

- The process of setting up applications or RDS desktops for remote access involves the following tasks:

- Installing Applications
  - If you plan to create application pools, you must install the applications on the RDS hosts. If you want Horizon 7 to automatically display the list of installed applications, you must install the applications so that they are available to all users from the Start menu. You can install an application at any time before you create the application pool. If you plan to manually specify an application, you can install the application at any time, either before or after creating an application pool.

- Important
  - When you install an application, you must install it on all the RDS hosts in a farm and in the same location on each RDS host. If you do not, a health warning will appear on the View Administrator dashboard. In such a situation, if you create an application pool, users might encounter an error when they try to run the application.
  - When you create an application pool, Horizon 7 automatically displays the applications that are available to all users rather than individual users from the Start menu on all of the RDS hosts in a farm. You can choose any applications from that list. In addition, you can manually specify an application that is not available to all users from the Start menu. There is no limit on the number of applications that you can install on an RDS host.

Farms, RDS Hosts, Desktop and Application Pools

With VMware Horizon, you can create desktop and application pools to give users remote access to virtual machine-based desktops, session-based desktops, physical computers, and applications. Horizon takes advantage of Microsoft Remote Desktop Services (RDS) and VMware PC-over-IP (PCoIP) technologies to provide high-quality remote access to users as follows:

- RDS Hosts
  - RDS hosts are server computers that have Windows Remote Desktop Services and View Agent installed. These servers host applications and desktop sessions that users can access remotely. To use RDS desktop pools or applications, your end users must have access to Horizon Client 3.0 or later software.

- Desktop Pools
  - There are three types of desktop pools: automated, manual, and RDS. Automated desktop pools use a vCenter Server virtual machine template or snapshot to create a pool of identical virtual machines. Manual desktop pools are a collection of existing vCenter Server virtual machines, physical computers, or third-party virtual machines. In automated or manual pools, each machine is available for one user to access remotely at a time. RDS desktop pools are not a collection of machines, but instead, provide users with desktop sessions on RDS hosts. Multiple users can have desktop sessions on an RDS host simultaneously.

- Application Pools
  - Application pools let you deliver applications to many users. The applications in application pools run on a farm of RDS hosts.

- Farms
  - Farms are collections of RDS hosts and facilitate the management of those hosts. Farms can have a variable number of RDS hosts and provide a common set of applications or RDS desktops to users. When you create an RDS desktop pool or an application pool, you must specify a farm. The RDS hosts in the farm provide desktop and application sessions to users.

Figure 14   VMware Horizon Architectural Overview
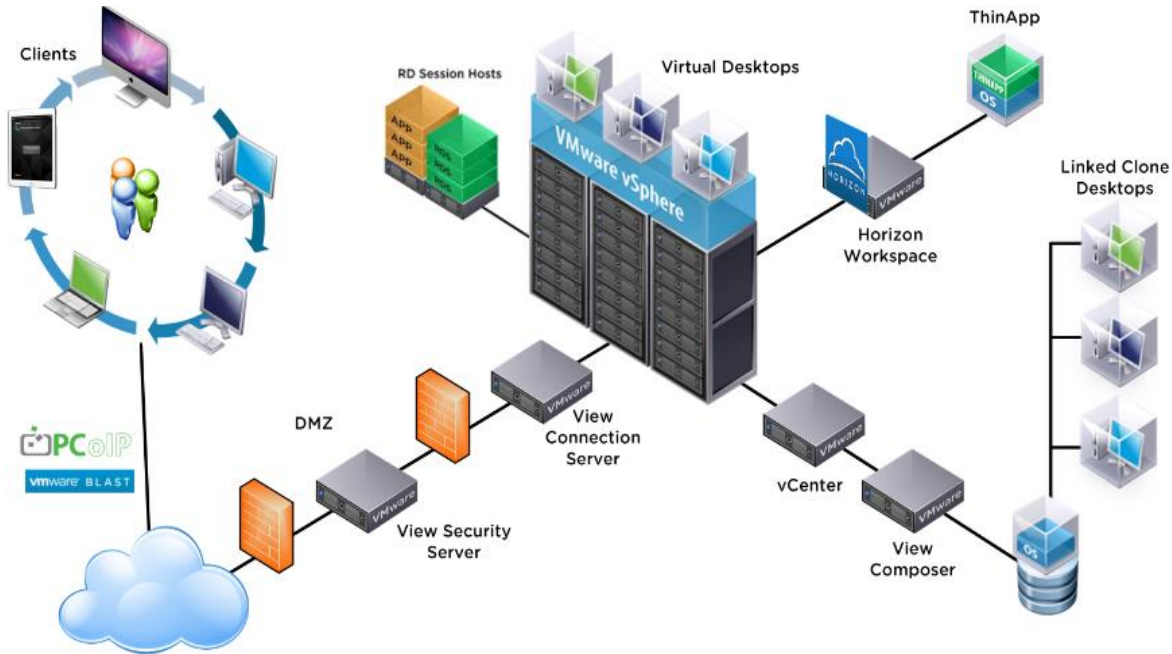
Architecture and Design of VMware Horizon on Cisco Unified Computing System and Cisco HyperFlex System Design Fundamentals

There are many reasons to consider a virtual desktop solution, such as an ever growing and diverse base of user devices, complexity in management of traditional desktops, security, and even Bring Your Own Computer (BYOC) to work programs. The first step in designing a virtual desktop solution is to understand the user community and the type of tasks that are required to successfully execute their role. The following sample user classifications are provided:

- Knowledge Workers today do not just work in their offices all day; they attend meetings, visit branch offices, work from home, and even coffee shops. These anywhere workers expect access to all of their same applications and data wherever they are.

- External Contractors are increasingly part of your everyday business. They need access to certain portions of your applications and data, yet administrators still have little control over the devices they use and the locations they work from. Consequently, IT is stuck making trade-offs on the cost of providing these workers a device vs. the security risk of allowing them access from their own devices.

- Task Workers perform a set of well-defined tasks. These workers access a small set of applications and have limited requirements from their PCs. However, since these workers are interacting with your customers, partners, and employees, they have access to your most critical data.

- Mobile Workers need access to their virtual desktop from everywhere, regardless of their ability to connect to a network. In addition, these workers expect the ability to personalize their PCs, by installing their own applications and storing their own data, such as photos and music, on these devices.

- Shared Workstation users are often found in state-of-the-art university and business computer labs, conference rooms or training centers. Shared workstation environments have the constant requirement to re-provision desktops with the latest operating systems and applications as the needs of the organization change, tops the list.

After the user classifications have been identified and the business requirements for each user classification have been defined, it becomes essential to evaluate the types of virtual desktops that are needed based on user requirements. There are essentially five potential desktops environments for each user:

- Traditional PC: A traditional PC is what –typically‖ constituted a desktop environment: physical device with a locally installed operating system.

- Remote Desktop Server Session (RDSH) Hosted Shared Desktop: A hosted, server-based desktop is a desktop where the user interacts through a delivery protocol. With hosted, server-based desktops, a single installed instance of a server operating system, such as Microsoft Windows Server 2012,2016 or 2019, is shared by multiple users simultaneously. Each user receives a desktop "session" and works in an isolated memory space. Changes made by one user could impact the other users.

- Remote Desktop Hosted Sessions (RDSH): A Remote Desktop Hosted Sessions (RDSH) is a virtual desktop running either on virtualization layer (ESX) or on bare metal hardware. The user does not work with and sit in front of the desktop, but instead the user interacts through a delivery protocol.

- Published Applications: Published applications run entirely on the VMware RDSH Session Hosts and the user interacts through a delivery protocol. With published applications, a single installed instance of an application, such as Microsoft, is shared by multiple users simultaneously. Each user receives an application "session" and works in an isolated memory space.

- Streamed Applications: Streamed desktops and applications run entirely on the user's local client device and are sent from a server on demand. The user interacts with the application or desktop directly, but the resources may only available while they are connected to the network.

- Local Virtual Desktop: A local virtual desktop is a desktop running entirely on the user's local device and continues to operate when disconnected from the network. In this case, the user's local device is used as a type 1 hypervisor and is synchronized with the data center when the device is connected to the network.

For the purposes of the validation represented in this document, both Horizon Virtual Desktops and Remote Desktop sever Hosted Sessions were validated. Each of the sections provides some fundamental design decisions for this environment.

Understanding Applications and Data

When the desktop user groups and sub-groups have been identified, the next task is to catalog group application and data requirements. This can be one of the most time-consuming processes in the VDI planning exercise but is essential for the VDI project's success. If the applications and data are not identified and co-located, performance will be negatively affected.

The process of analyzing the variety of application and data pairs for an organization will likely be complicated by the inclusion cloud applications, like SalesForce.com. This application and data analysis is beyond the scope of

this Cisco Validated Design but should not be omitted from the planning process. There are a variety of third-party tools available to assist organizations with this crucial exercise.

Project Planning and Solution Sizing Sample Questions

Now that user groups, their applications and their data requirements are understood, some key project and solution sizing questions may be considered.

General project questions should be addressed at the outset, including:

- Has a VDI pilot plan been created based on the business analysis of the desktop groups, applications, and data?
- Is there infrastructure and budget in place to run the pilot program?
- Are the required skill sets to execute the VDI project available? Can we hire or contract for them?
- Do we have end user experience performance metrics identified for each desktop sub-group?
- How will we measure success or failure?
- What is the future implication of success or failure?

Below is a short, non-exhaustive list of sizing questions that should be addressed for each user sub-group:

- What is the desktop OS planned? Windows 7, Windows 8, or Windows 10?
- 32-bit or 64-bit desktop OS?
- How many virtual desktops will be deployed in the pilot? In production? All Windows 7/8/10?
- How much memory per target desktop group desktop?
- Are there any rich media, Flash, or graphics-intensive workloads?
- What is the end point graphics processing capability?
- Will VMware RDSH for Remote Desktop Server Hosted Sessions used?
- If RDSH is used what is the desktop OS planned? Server 2008, Server 2012, 2016 or Server 2019?
- How many RDSH sessions will be deployed in the pilot? In production?
- What is the hypervisor for the solution?
- What is the storage configuration in the existing environment?
- Are there sufficient IOPS available for the write-intensive VDI workload?
- Will there be storage dedicated and tuned for VDI service?
- Is there a voice component to the desktop?
- Is anti-virus a part of the image?
- Is user profile management (such as non-roaming profile based) part of the solution?
- What is the fault tolerance, failover, disaster recovery plan?
- Are there additional desktop sub-group specific questions?

Desktop Virtualization Design Fundamentals

An ever growing and diverse base of user devices, complexity in management of traditional desktops, security, and even Bring Your Own Device (BYOD) to work programs are prime reasons for moving to a virtual desktop solution.

VMware Horizon Design Fundamentals

VMware Horizon 7 integrates Remote Desktop Server Hosted sessions users and VDI desktop virtualization technologies into a unified architecture that enables a scalable, simple, efficient, mixed users and manageable solution for delivering Windows applications and desktops as a service.

Users can select applications from an easy-to-use "store" that is accessible from tablets, smartphones, PCs, Macs, and thin clients. VMware Horizon delivers a native touch-optimized experience via PCoIP or Blast Extreme high-definition performance, even over mobile networks.

Horizon VDI Pool and RDSH Servers Pool

Collections of identical Virtual Machines (VMs) or physical computers are managed as a single entity called a Desktop Pool. In this CVD, VM provisioning relies on VMware View Composer aligning with VMware Horizon View Connection Server and vCenter Server components. Machines in these Pools are configured to run either a Windows Server 2019 OS (for RDSH hosted shared sessions) or a Windows 10 Desktop OS (for, instant clone and persistent VDI desktops).

> ⚠ Server OS and Desktop OS Machines were configured in this CVD to support RDSH hosted shared desktops and a variety of VDI Remote Desktop Hosted Sessions (RDSH) virtual desktops.

Figure 15 VMware Horizon Design Overview
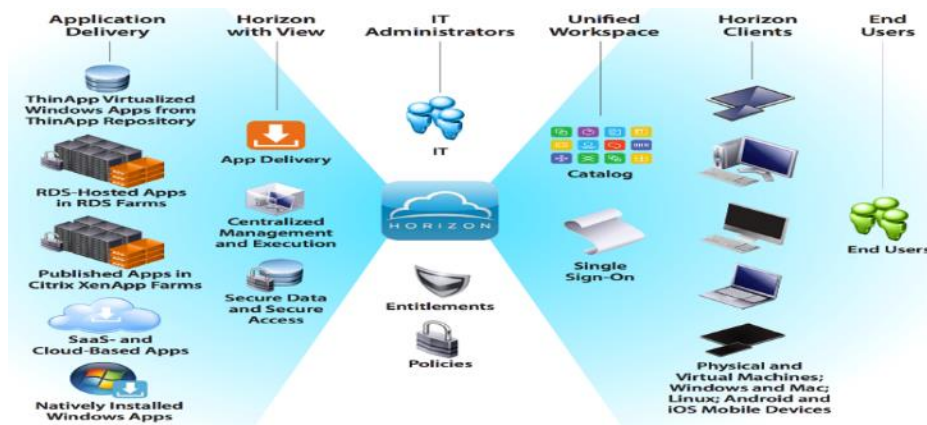
Figure 16 Horizon VDI and RDSH Desktop Delivery Based on Display Protocol (PCoIP/Blast/RDP)
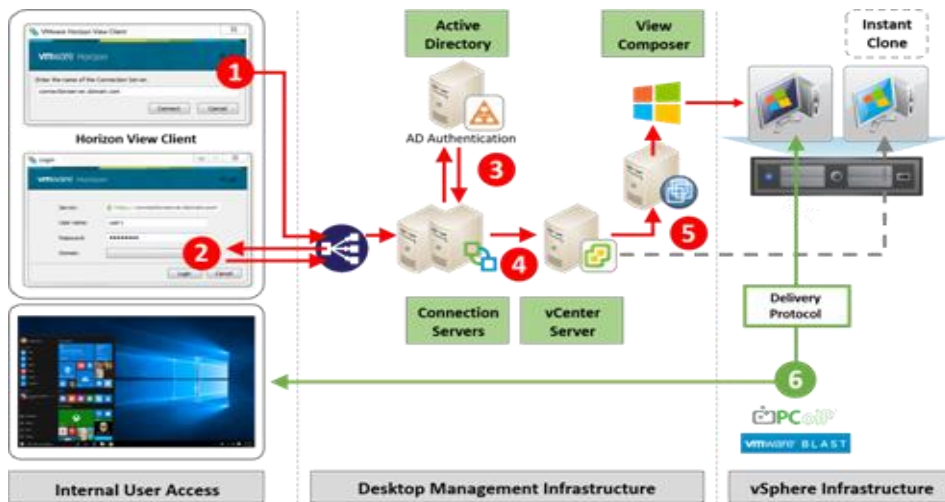
Multiple Site Configuration

If you have multiple regional sites, you can use any of the Load Balances Tools to direct the user connections to the most appropriate site to deliver the desktops and application to users.

Figure 17 illustrating sites, shows a site created in two data centers. Having two sites globally, rather than just one, minimizes the amount of unnecessary WAN traffic. Two Cisco blade servers host the required infrastructure services (Domain Controllers, DNS, DHCP, Profile, SQL, VMware Horizon View Connection Servers, View Composer server and web servers).

Figure 17 Multisite Configuration Overview
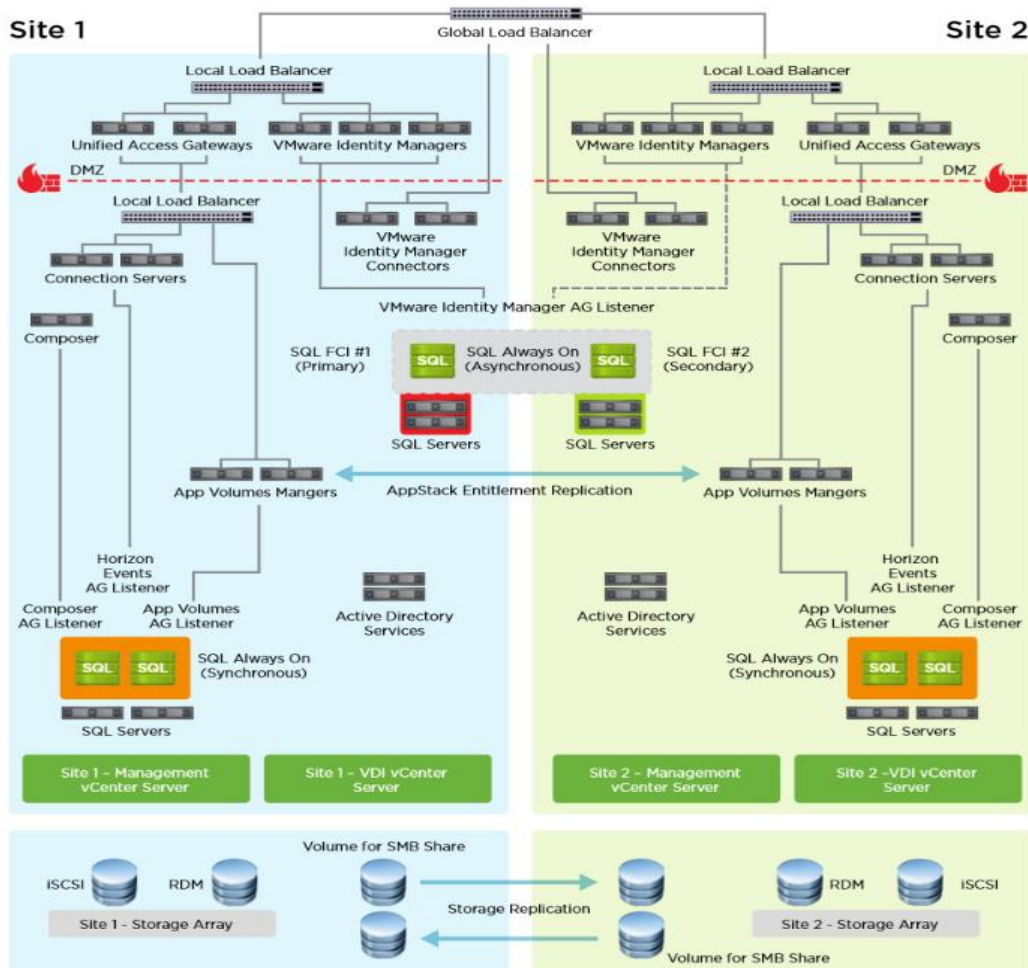
Based on the requirement and no of data centers or remote location, we can choose any of the available Load balancing software or tools accelerates the application performance, load balances servers, increases security, and optimizes the user experience.

Multi-Site configuration is shown as the example.

Designing a VMware Horizon Environment for Various Workload Types

With VMware Horizon 7, the method you choose to provide applications or desktops to users depends on the types of applications and desktops you are hosting and available system resources, as well as the types of users and user experience you want to provide.

Machine	User Type and Experience
Server OS machines	You want: Inexpensive server-based deliver to minimize the cost of delivering applications to a large number of users while providing a secure, high-definition user experience.
	Your users: Perform well-defined tasks and do not require personalization or offline

Machine	User Type and Experience
	access to applications. Users may include task workers such as call center operators and retail workers, or users that share workstations. Application types: Any application.
Desktop OS machines	You want: A client-base application delivery solution that is secure, provides centralized management, and supports a large number of users per host server (or hypervisor), while providing users with applications that display seamlessly in high-definition. Your users: Are internal, external contractors, third-party collaborators, and other provisional team members. Users do not require offline access to hosted applications. Application types: Applications that might not work well with other applications or might interact with the operating system, such as .NET framework. These types of applications are ideal for hosting on virtual machines. Applications running on older operating systems such as Windows XP or Windows Vista, and older architectures such as 32-bit or 16-bit. By isolating each application on its own virtual machine, if one machine fails, it does not impact other users.
Remote PC Access	You want: Employees with secure remote access to a physical computer without using a VPN. For example, the user may be accessing their physical desktop PC from home or through a public WIFI hotspot. Depending upon the location, you may want to restrict the ability to print or copy and paste outside of the desktop. This method enables BYO device support without migrating desktop images into the datacenter. Your users: Employees or contractors that have the option to work from home but need access to specific software or data on their corporate desktops to perform their jobs remotely. Host: The same as Desktop OS machines. Application types: Applications that are delivered from an office computer and display seamlessly in high definition on the remote user's device.

For the Cisco Validated Design described in this document, individual configuration of Remote Desktop Server Hosted sessions (RDSH) using RDS-based Server OS machines and Virtual Desktops using Desktop OS machines via Instant-clone, Linked- clone automated pool and Full clone persistent desktops were configured and tested. The following sections discuss design decisions relative to the VMware Horizon deployment, including this CVD test environment.

Solution Design

Requirements

The following sections detail the physical hardware, software revisions, and firmware versions required to install a single cluster of the Cisco HyperFlex system. This solution cluster will have a two-to-one ratio for Compute Only nodes to Converged Nodes. There are 16 Compute Only Nodes and 8 Converged Nodes for a total of 24 nodes.

Physical Components

Table 2　HyperFlex System Components

Component	Hardware Required
Fabric Interconnects	Two Cisco UCS 6454 Fabric Interconnects
Servers	Eight Cisco HyperFlex HXAF220c-M5SX All-Flash rack servers

For complete server specifications and more information, please refer to the Cisco HyperFlex HXAF220c M5 Node Spec Sheet

Table 3 lists the hardware component options for the HXAF220c-M5SX server model:

Table 3　Table 4　HXAF220c-M5SX Server Options

HXAF220c-M5SX Options	Hardware Required
Processors	Chose a matching pair of 2nd Generation Intel Xeon 6230 Processor Scalable Family CPUs
Memory	786 GB total memory using 64 GB DDR4 2933 MHz 1.2v modules depending on CPU type
Disk Controller	Cisco 12Gbps Modular SAS HBA
SSDs	Standard: One 240 GB 2.5 Inch Enterprise Value 6G SATA SSD 1.6 TB 2.5 Inch Extreme Performance SAS SSD Six to eight 3.8 TB 2.5 Inch Enterprise Value 6G SATA SSDs, or six to eight 960 GB 2.5 Inch Enterprise Value 6G SATA SSDs SED: One 240 GB 2.5 Inch Enterprise Value 6G SATA SSD One 800 GB 2.5 Inch Enterprise Performance 12G SAS SED SSD Six to eight 3.8 TB 2.5 Inch Enterprise Value 6G SATA SED SSDs, or six to eight 960 GB 2.5 Inch Enterprise Value 6G SATA SED SSDs
Network	Cisco UCS VIC 1457 VIC MLOM
Boot Device	One 240 GB M.2 form factor SATA SSD
microSD Card	One 32GB microSD card for local host utilities storage (Not used in this study)
Optional	

Table 4 lists the hardware component options for the Cisco UCS 220c-M5SX server model.

Table 4   Cisco UCS 229C-M5SX Server Options

UCS220c-M5SX Options	Hardware Required
Processors	Chose a matching pair of 2nd Generation Intel Xeon 6230 Processor Scalable Family CPUs
Memory	786 GB total memory using 64 GB DDR4 2933 MHz 1.2v modules depending on CPU type
Disk Controller	Cisco 12Gbps Modular SAS HBA
SSDs	Standard SED
Network	Cisco UCS VIC 1457 VIC MLOM
Boot Device	One 240 GB M.2 form factor SATA SSD
microSD Card	One 32GB microSD card for local host utilities storage (Not used in this study)
Optional	

Table 5 lists the hardware component options for the UCS240c-M5L server model:

Table 5   Cisco UCS B200-M5 Server Options

B200-M5 Options	Hardware Required
Processors	Chose a matching pair of 2nd Generation Intel Xeon Processor Scalable Family CPUs
Memory	786 GB total memory using 64 GB DDR4 2933 MHz 1.2v modules depending on CPU type
Disk Controller	Cisco 12Gbps Modular SAS HBA
SSDs	One 240 GB 2.5 Inch Enterprise Value 6G SATA SSD
One 3.2 TB 2.5 Inch Enterprise Performance 12G SAS SSD	
HDDs	Six to twelve 12 TB, 8 TB or 6 TB SAS 7.2K RPM LFF HDD
Network	Cisco UCS VIC1440 VIC

B200-M5 Options	Hardware Required
Boot Device	One 240 GB M.2 form factor SATA SSD
microSD Card	One 32GB microSD card for local host utilities storage
Optional	Cisco HyperFlex Acceleration Engine card

Software Components

The software components of the Cisco HyperFlex system must meet minimum requirements for the Cisco UCS firmware, hypervisor version, and the Cisco HyperFlex Data Platform software in order to interoperate properly.

For additional hardware and software combinations, refer to the public Cisco UCS Hardware Compatibility webpage: https://ucshcltool.cloudapps.cisco.com/public/

Table 6 lists the software components and the versions required for the Cisco HyperFlex 4.0 system:

Table 6   Software Components

Component	Software Required
Hypervisor	VMware ESXi 6.7 Update 3 CISCO Custom Image for ESXi 6.7 Update 3 for HyperFlex: HX-ESXi-6.7U3-15160138-Cisco-Custom-6.7.3.3-install-only.iso NOTE: Using a published Cisco custom ESXi ISO installer file is required when in-stalling/reinstalling ESXi or upgrading to a newer version prior to installing Hy-perFlex. An offline bundle file is also provided to upgrade ESXi on running clusters. NOTE: ESXi 6.0 is not supported on servers equipped with the Cisco VIC1457 card, or the HXAF220c-M5N model servers. Each of these requires ESXi 6.5 Update 3 or higher. NOTE: VMware vSphere Standard, Essentials Plus, ROBO, Enterprise or Enterprise Plus licensing is required from VMware.
Management Server	VMware vCenter Server for Windows or vCenter Server Appliance 6.0 U3c or later. Refer to http://www.vmware.com/resources/compatibility/sim/interop_matrix.php for interoperability of your ESXi version and vCenter Server.
Cisco HyperFlex Data Plat-form	Cisco HyperFlex HX Data Platform Software 4.0(2a)
Cisco UCS Firmware	Cisco UCS Infrastructure software, B-Series and C-Series bundles, revision 4.0(4g) or later.

Licensing

Cisco HyperFlex systems must be properly licensed using Cisco Smart Licensing, which is a cloud-based software licensing management solution used to automate many manual, time consuming and error prone licensing tasks. Cisco HyperFlex 2.5 and later communicate with the Cisco Smart Software Manager (CSSM) online service via a Cisco Smart Account, to check out or assign available licenses from the account to the Cisco Hyper-Flex cluster resources. Communications can be direct via the internet, they can be configured to communicate via a proxy server, or they can communicate with an internal Cisco Smart Software Manager satellite server, which caches and periodically synchronizes licensing data. In a small number of highly secure environments, systems can be provisioned with a Permanent License Reservation (PLR) which does not need to communicate with CSSM. Contact your Cisco sales representative or partner to discuss if your security requirements will necessitate use of these permanent licenses. New HyperFlex cluster installations will operate for 90 days without licensing as an evaluation period, thereafter the system will generate alarms and operate in a non-compliant mode. Systems without compliant licensing will not be entitled to technical support.

For more information on the Cisco Smart Software Manager satellite server, visit this website: https://www.cisco.com/c/en/us/buy/smart-accounts/software-manager-satellite.html

Beginning with Cisco HyperFlex 3.0, licensing of the system requires one license per node from one of three different licensing editions; Edge licenses, Standard licenses, or Enterprise licenses. Depending on the type of cluster being installed, and the desired features to be activated and used in the system, licenses must be purchased from the appropriate licensing tier. Additional features in the future will be added to the different licensing editions as they are released, the features listed below are current only as of the publication of this document.

Table 7 lists an overview of the licensing editions, and the features available with each type of license:

Table 7   HyperFlex System License Editions

HyperFlex Licensing Edition	Edge	Standard (in addition to Edge)	Enterprise (in addition to Standard)
Features Available	HyperFlex Edge clusters without Fabric Interconnects 220 SFF model servers only Hybrid or All-Flash ESXi Hypervisor only Replication Factor 2 only 1 Gb or 10 Gb Ethernet only Compression Deduplication HyperFlex native snap-	HyperFlex standard clusters with Fabric Interconnects 220 and 240 SFF server models and 240 LFF server models Replication Factor 3 Hyper-V and Kubernetes platforms Cluster expansions Compute-only nodes up to 1:1 ratio 10 Gb, 25 Gb or 40 Gb Ethernet	Stretched clusters 220 all-NVMe server models Cisco HyperFlex Acceleration Engine cards Compute-only nodes up to 2:1 ratio

HyperFlex Licensing Edition	Edge	Standard (in addition to Edge)	Enterprise (in addition to Standard)
	shots	Data-at-rest encryption using self-encrypting disks	
	Rapid Clones		
	HyperFlex native replication	Logical Availability Zones	
	Management via vCenter plugin, HyperFlex Connect, or Cisco Intersight		

Licensing required to replicate 2:1 CO: HCI

For replicating this CVD with 2:1 CO:HCI ratio, the Enterprise edition License is required.

For a comprehensive guide to licensing and all the features in each edition, consult the Cisco HyperFlex Licensing Guide here:
https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/b_Cisco_HyperFlex_Systems_Ordering_and_Licensing_Guide/b_Cisco_HyperFlex_Systems_Ordering_and_Licensing_Guide_chapter_01001.html

Physical Topology

Topology Overview

The Cisco HyperFlex system is composed of a pair of Cisco UCS Fabric Interconnects along with up to thirty-two HX-Series rack-mount servers per cluster. Up to thirty-two compute-only servers can also be added per HyperFlex cluster. Adding Cisco UCS rack-mount servers and/or Cisco UCS 5108 Blade chassis, which house Cisco UCS blade servers, allows for additional compute resources in an extended cluster design. The two Fabric Interconnects both connect to every HX-Series rack-mount server, and both connect to every Cisco UCS 5108 blade chassis, and Cisco UCS rack-mount server. Upstream network connections, also referred to as "north-bound" network connections are made from the fabric interconnects to the customer datacenter network at the time of installation.

Figure 18 HyperFlex Standard Cluster Topology
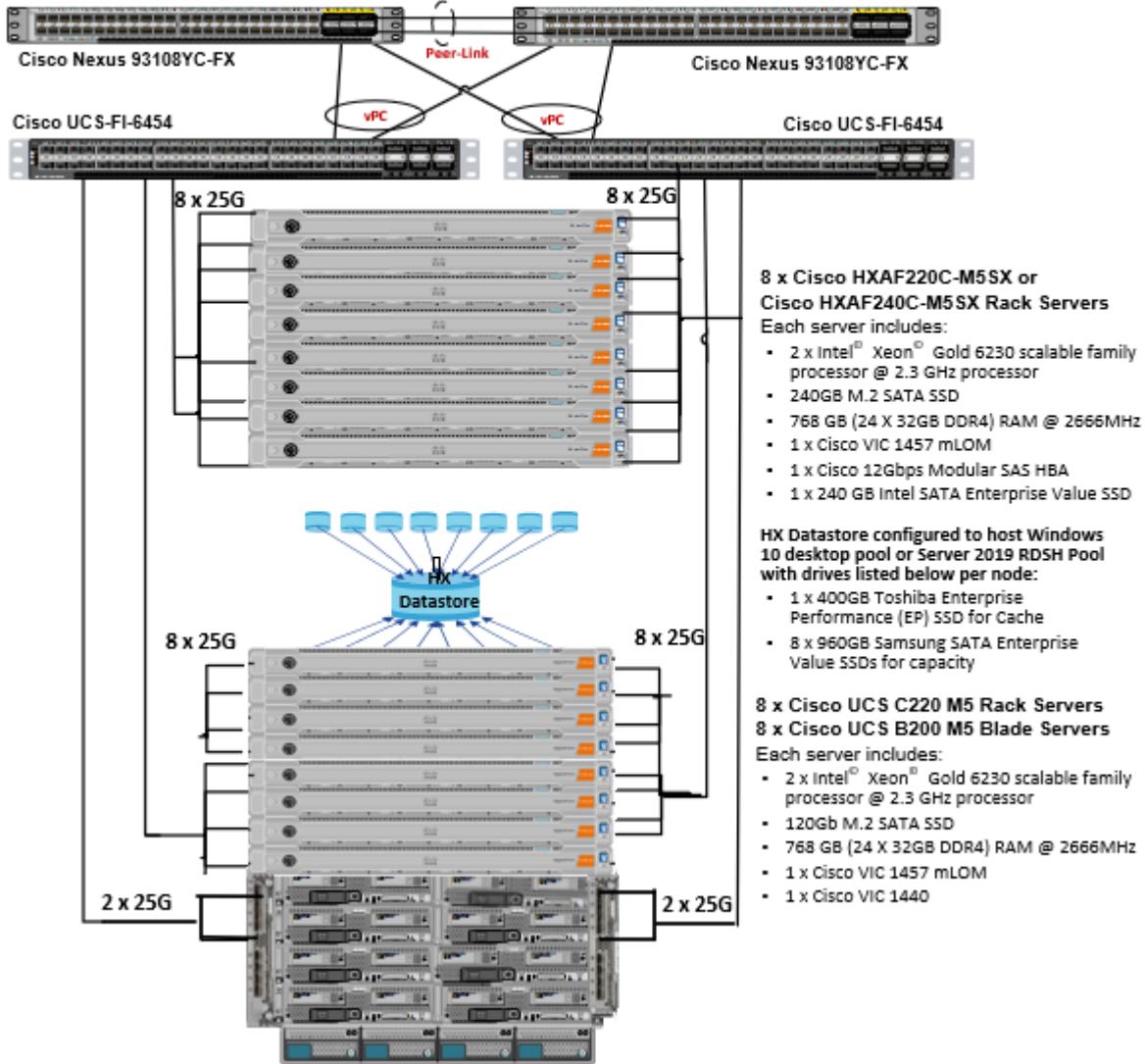
Hardware Deployed

The solution contains the following hardware as shown in [Figure 17](#):

- Two Cisco Nexus 93180YC-FX Layer 2 Access Switches

- Two Cisco Fabric Interconnects 6454

- Two Cisco UCS C220 M4 Rack servers with dual socket Intel Xeon E5-2620v4 2.1-GHz 8-core processors, 128GB RAM 2133-MHz and VIC1227 mLOM card for the hosted infrastructure with N+1 server fault tolerance. (Not show in the diagram).

- Eight Cisco HXAF220C M5 Rack servers with Intel Xeon Gold 6230 scalable family 2.3-GHz 18-core processors, 768GB RAM 2666-MHz and VIC1457 mLOM cards running Cisco HyperFlex data platform v4.0(1a) for the virtual desktop workloads with N+1 server fault tolerance.

- Eight Cisco UCS C220 M5 Rack servers with Intel Xeon Gold 6230 scalable family 2.3-GHz 18-core processors, 768GB RAM 2666-MHz and VIC1457 mLOM cards running Cisco HyperFlex data platform v4.0(1a) for the virtual desktop workloads with N+1 server fault tolerance

- Eight Cisco UCS B200 M5 Blade servers with Intel Xeon Gold 6230 scalable family 2.3-GHz 18-core processors, 768GB RAM 2666-MHz and VIC1440 cards running Cisco HyperFlex data platform v4.0(1a) for the virtual desktop workloads with N+1 server fault tolerance.

Software Deployed

Table 8 lists the software and firmware version used in the study.

Table 8   Software and Firmware Versions

Vendor	Product	Version
Cisco	UCS Component Firmware	4.0(4g) bundle release
Cisco	UCS Manager	4.0(4g) bundle release
Cisco	UCS HXAF220c-M5S rack server	4.0(4g) bundle release
Cisco	VIC 1457	4.2(2b)
Cisco	UCS B200 M5 blade server	4.0(1b) bundle release
Cisco	VIC 1440	4.2(2d)
Cisco	HyperFlex Data Platform	4.0.(2a)
Cisco	Cisco eNIC	2.1.2.71
Cisco	Cisco fNIC	1.6.0.37
Network	Cisco Nexus 9000 NX-OS	7.0(3)I7(2)
VMware	Horizon Connection Server	7.12.0-15770369
VMware	Horizon Agent	7.12.0-15805436
VMware	Horizon Client	5.4.3-16346110
VMware	vCenter Server Appliance	6.7-0 -42000

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 always active, 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.

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 via GUI and CLI tools. This port is 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 port 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 port 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. This port is 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. Cisco HyperFlex M5 generation servers can be configured with the Cisco VIC 1387 or VIC 1457 cards. The standard and redundant connection practice for the VIC 1387 is to connect port 1 of the VIC card (the right-hand port) to a port on FI A, and port 2 of the VIC card (the left-hand port) to a port on FI B (Figure 19). For the VIC 1457 card, the standard and redundant practice is to connect port 1 of the VIC card (the left-hand most port) to a port on FI A, and connect port 3 (the right-center port) to a port on FI B (Figure 20). An optional configuration method for servers containing the Cisco VIC 1457 card is to cable the servers with 2 links to each FI, using ports 1 and 2 to FI A, and ports 3 and 4 to FI B. The HyperFlex installer checks for these configurations, and that all **servers' cabling matches.** Failure to follow this cabling best practice can lead to errors, discovery failures, and loss of redundant connectivity.

All nodes within a Cisco HyperFlex cluster must be connected at the same communication speed, for example, mixing 10 Gb with 25 Gb interfaces is not allowed. In addition, for clusters that contain only M5 generation nodes, all of the nodes within a cluster must contain the same model of Cisco VIC cards.

Various combinations of physical connectivity between the Cisco HX-series servers and the Fabric Interconnects are possible, but only specific combinations are supported. Table 9 lists the possible connections, and which of these methods is supported.

Table 9   Supported Physical Connectivity

Fabric Interconnect Model	6248	6296	6332		6332-16UP			6454	
Port Type	10GbE	10GbE	40GbE	10GbE Breakout	40GbE	10GbE Breakout	10GbE onboard	10GbE	25GbE
M4 with VIC 1227	✓	✓	✕	✕	✕	✕	✕	✓	✕
M4 with VIC 1387	✕	✕	✓	✕	✓	✕	✕	✕	✕
M4 with VIC 1387 + QSA	✓	✓	✕	✕	✕	✕	✕	✓	✕
M5 with VIC 1387	✕	✕	✓	✕	✓	✕	✕	✕	✕
M5 with VIC 1387 + QSA	✓	✓	✕	✕	✕	✕	✕	✓	✕
M5 with VIC 1457	✓	✓	✕	✕	✕	✕	✕	✓	✓

Figure 19 HX-Series Server with Cisco VIC 1457 Connectivity
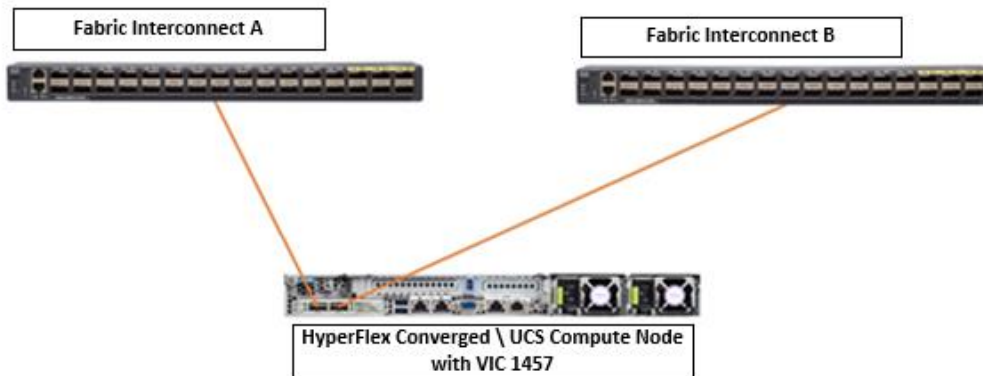
Figure 20 Cisco UCS B200 M5 Compute Only Node with Cisco VIC 1440 Connectivity
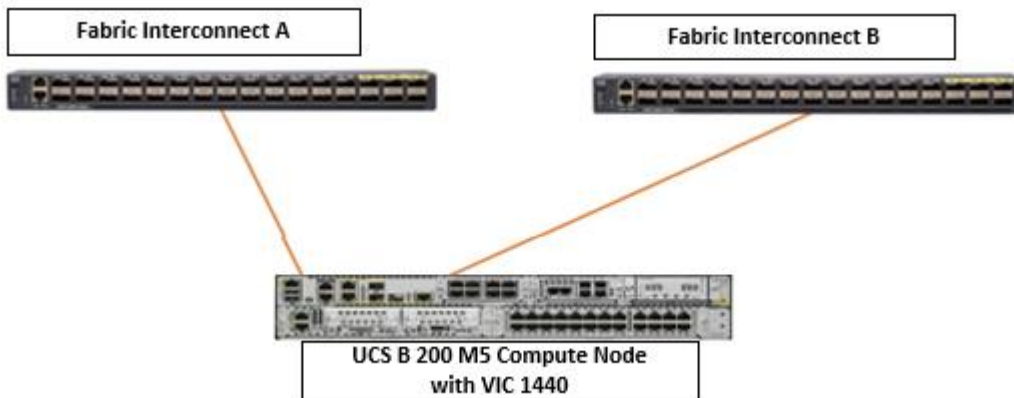
Cisco UCS B-Series Blade Servers

HyperFlex extended clusters can also incorporate 1-32 Cisco UCS blade servers for additional compute capacity. 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 card installed in each blade server across the chassis backplane. The standard connection practice is to connect 1-8 10 GbE links, or 1-4 40 GbE links (depending on the IOMs and FIs purchased) from the left-side IOM, or IOM 1, to FI A, and to connect the same number of 10 GbE or 40 GbE links from the right-side IOM, or IOM 2, to FI B (Figure 21). All other cabling configurations are invalid, and can lead to errors, discovery failures, and loss of redundant connectivity.

Figure 21 Cisco UCS 5108 Chassis Connectivity
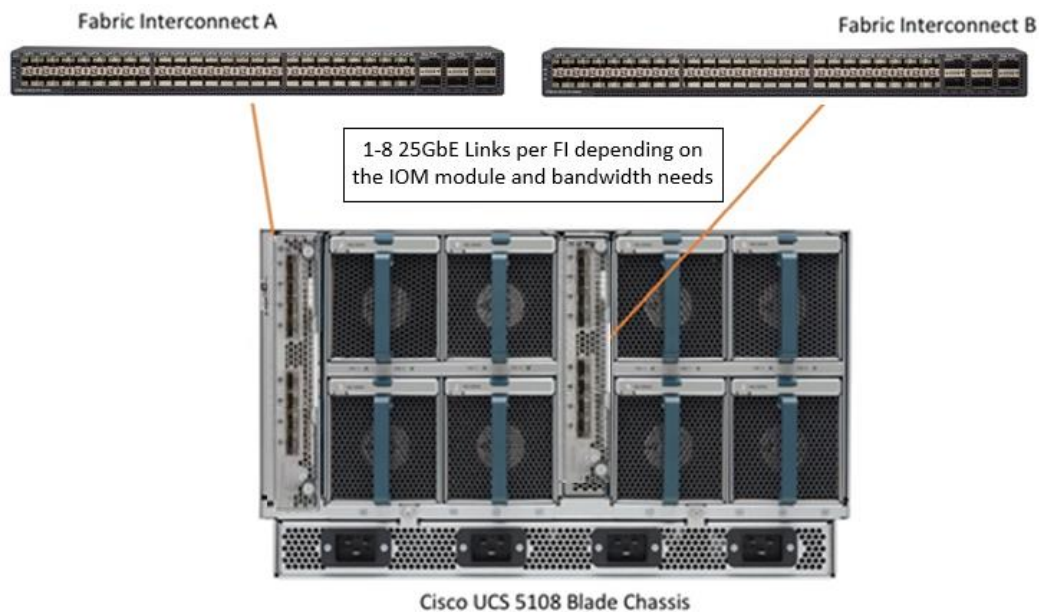
Cisco UCS C-Series Rack-Mount Servers

HyperFlex extended clusters can also incorporate 1-32 Cisco UCS Rack-Mount Servers for additional compute capacity. The Cisco UCS C-Series Rack-Mount Servers are connected directly to the Cisco UCS Fabric Interconnects in Direct Connect mode. Internally the Cisco UCS C-Series servers are configured with the Cisco VIC 1227, 1387 or 1457 network interface card (NIC) installed in a modular LAN on motherboard (MLOM) slot, which have dual 10 Gigabit Ethernet (GbE), quad 10/25 Gigabit Ethernet (GbE) ports or dual 40 Gigabit Ethernet (GbE) ports. The standard and redundant connection practice for connecting standard Cisco UCS C-Series servers to the Fabric Interconnects is identical to the method described earlier for the HX-Series servers. Failure to follow this cabling practice can lead to errors, discovery failures, and loss of redundant connectivity.

Figure 22 Cisco UCS C-Series Server Connectivity
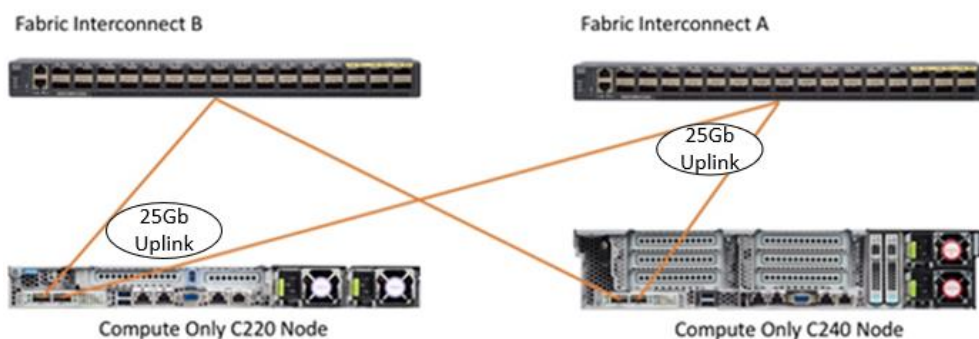
Logical Topology

Logical Network Design

The Cisco HyperFlex system has communication pathways that fall into four defined zones (Figure 23):

- 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 also 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 via the FI management ports.
  - ESXi host management interfaces.
  - Storage Controller VM management interfaces.
  - A roaming HX cluster management interface.
  - Storage Controller VM replication interfaces.
  - A roaming HX cluster replication 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, which are trunked to the Cisco UCS Fabric Interconnects via 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 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. This zone is primarily jumbo frame traffic therefore jumbo frames must be enabled on the Cisco UCS uplinks. In this zone are multiple components:

- A VMkernel interface used for storage traffic on 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 north-bound of the Cisco UCS domain. For that reason, the VLAN used for HX vMotion traffic must be able to traverse the network uplinks from the Cisco UCS domain, reaching FI A from FI B, and vice-versa.

Figure 23 Logical Network Design
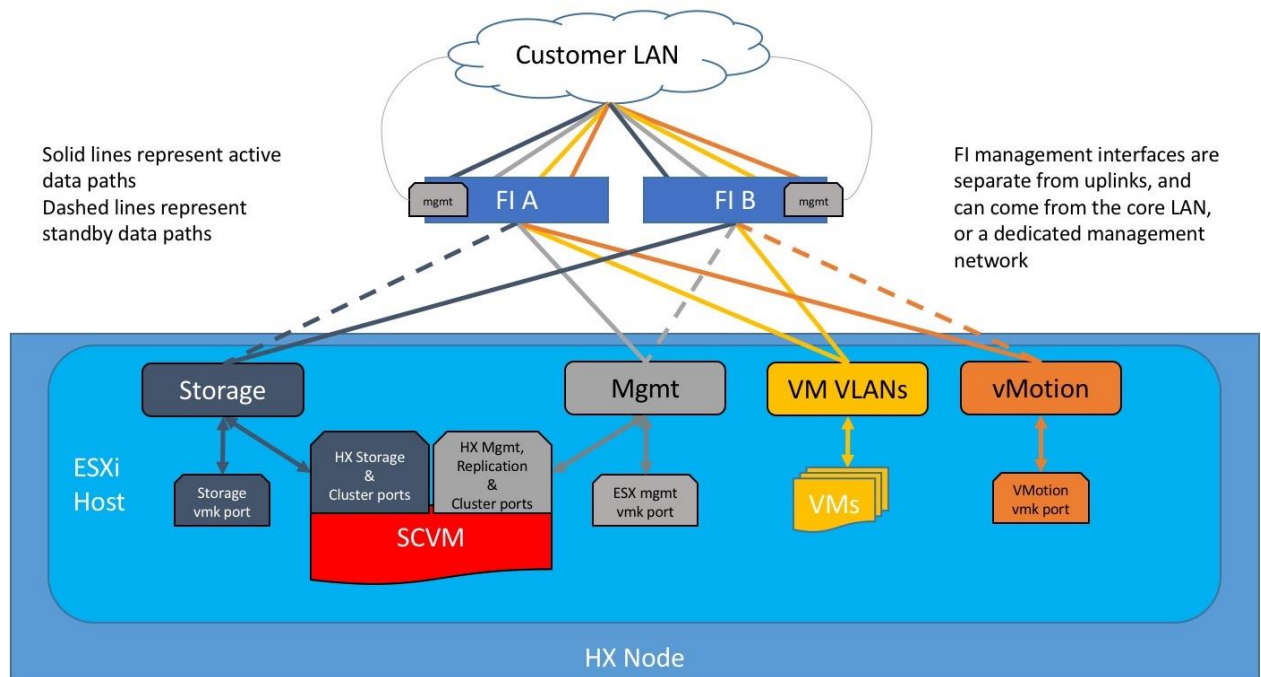
Logical Availability Zones

Larger scale HyperFlex clusters are subject to higher failure risks, simply due to the number of nodes in the cluster. While any individual node's risk of failure is the same no matter how many nodes there are, with clusters up to 32 converged nodes in size, there is a logically higher probability that a single node could fail, when compared to a cluster with fewer nodes. To mitigate these risks in larger scale clusters, a HyperFlex cluster of eight nodes or more can be configured with a feature called Logical Availability Zones (LAZ). The Logical Availability Zones feature groups 2 or more HyperFlex nodes together into a logically defined zone, a minimum of 3 zones are created, and the data in the cluster is distributed in such a way that no blocks are written to the nodes within a single zone more than once. Due to this enhanced distribution pattern of data across zones, wherein each zone has multiple servers, clusters with LAZ enabled can typically withstand more failures than clusters which operate without it. The number of failures that can tolerated varies depending on the number of zones in the cluster, and the number of servers in each of the zones. Generally speaking, multiple node failures across one or two zones will be tolerated better, and with less risk than multiple nodes failing across three or more zones. Note that the failure tolerance shown in the HyperFlex Connect dashboard will always present a "worst case scenario"

view, meaning that even though the dashboard may state that two failures can be tolerated, in fact two servers could fail and the cluster can remain online, and the failure tolerance may still remain at two.

Logical availability zones should not be confused with the concept of fault domains. An example of a fault domain would be a subset of the nodes in a single HyperFlex cluster being powered by one uninterruptable power supply (UPS) or connected to one power distribution unit (PDU), meanwhile the remaining nodes would be connected to another UPS or PDU. If one of the UPS' or PDUs were to fail, then there would be a simultaneous failure of multiple nodes. While LAZ may actually prevent the cluster from failing in this scenario, to guarantee it would require that the zone membership be manually controlled, so that a failure of all of the servers protected by a single UPS or PDU, would be distributed in such a way that it would not cause an outage. The LAZ feature is not designed to be manually configured in this way, instead the zone membership is determined automatically by the system. If a HyperFlex cluster needs to be physically split in half due to a physical limitation, such as the UPS example above, or a distance requirement for fault tolerance, then the cluster should be built as a stretched cluster instead of using LAZ.

Figure 24 illustrates an example of the data distribution method for clusters with Logical Availability Zones enabled, set to replication factor 3, where each zone only contains one of the three copies of the data in the cluster. This cluster consists of eight nodes, which the system configures into four zones.

Figure 24 Logical Availability Zone Data Distribution
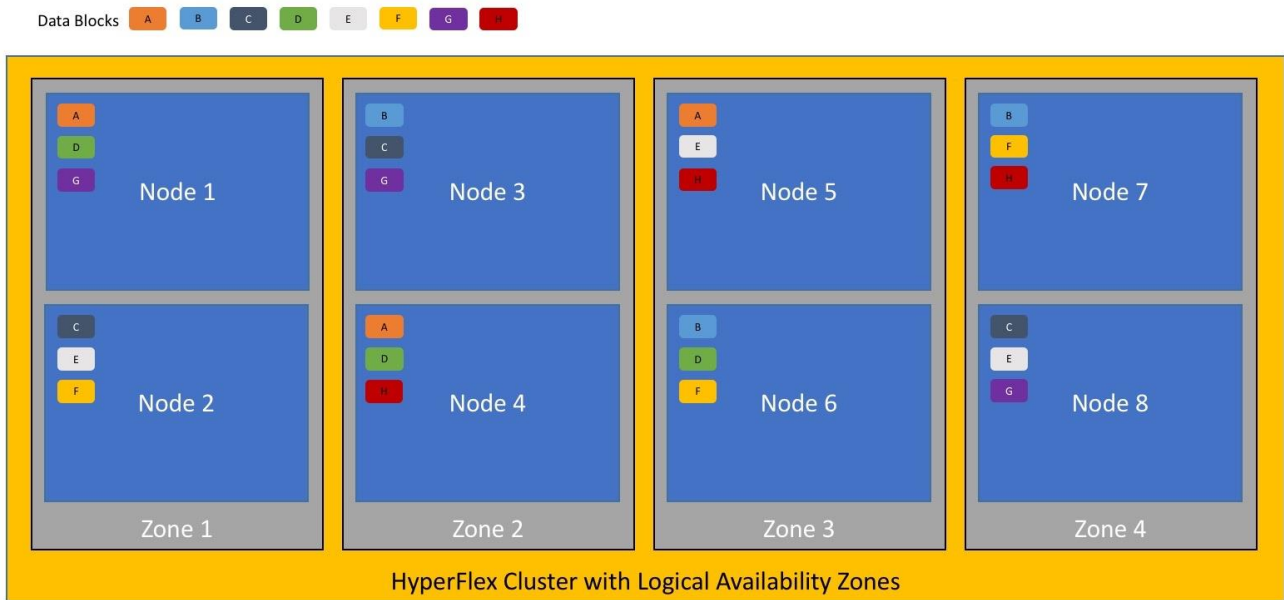
Logical availability zones are subject to the following requirements and limitations:

- Only HyperFlex clusters with 8 nodes or more can be configured with logical availability zones during the installation process.

- Logical Availability Zones can be enabled during the HyperFlex cluster installation, or it can be enabled via the command line at a later time. It is recommended to enable this feature during installation, in order to avoid a large migration and reorganization of data across the cluster, which would be necessary to comply with the data distribution rules if LAZ is turned on in a cluster already containing data.

- The number of zones can be manually specified as 3, 4, 5, or you can allow the installer to automatically choose, which is the recommended setting.

- The HyperFlex cluster determines which nodes participate in each zone, and this configuration cannot be modified.

- To maintain the most balanced consumption of space and data distribution, it is recommended that the number of nodes in a cluster are whole multiples of 3, 4, 5, or 7. For example, 8 nodes would evenly divide into 4 zones of 2 servers each, and 9 nodes would divide evenly into 3 zones of 3 servers each. Eleven nodes would create an unbalanced number of nodes across the zones, leading to unbalanced space consumption on the nodes.

- In addition to the previous point, expansion of a cluster should be done in multiples of the number of zones, when the cluster is operating with LAZ enabled. Expanding in such a way preserves a matched number of nodes in each zone and prevents any unbalance of space consumption. For example, a cluster with 3 zones should be expanded by adding 3 more nodes, because adding only 1 or 2 nodes would lead to an imbalance, as would adding 4 nodes.

Considerations

Version Control

The software revisions 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

VMware vCenter Server 6.0 Update 3c or later is required due to the requirement for TLS 1.2 with Cisco HyperFlex 4.0. The following best practice guidance applies to installations of HyperFlex 4.0:

- 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 multiple methods of deployment if no external vCenter server is already available:
http://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/TechNotes/Nested_vcenter_on_hyperflex.html

This document does not explain the installation and configuration of VMware vCenter Server for Windows, or the vCenter Server Appliance.

Scale

Cisco HyperFlex standard clusters currently scale from a minimum of 3 to a maximum of 32 Cisco HX-series converged nodes with small form factor (SFF) disks per cluster. A converged node is a member of the cluster which provides storage resources to the HX Distributed Filesystem. **For the compute intensive "**extended**"** cluster design, a configuration with 3 to 32 Cisco HX-series converged nodes can be combined with up to 32 com-

pute nodes. It is required that the number of compute-only nodes should always be less than or equal to number of converged nodes when using the HyperFlex Standard licenses. If using HyperFlex Enterprise licenses, the number of compute-only nodes can grow to as much as twice the number of converged nodes. Regardless of the licensing used, the combined maximum size of any HyperFlex cluster cannot exceed 64 nodes. Once the maximum size of a single **cluster has been reached, the environment can be "scaled out" by adding additional** HX model servers to the Cisco UCS domain, installing an additional HyperFlex cluster on them, and controlling them via the same vCenter server. There is no longer any limit to the number of clusters that can be created in a single UCS domain, the practical limits will instead be reached due to the number of ports available on the Fabric Interconnects. Up to 100 HyperFlex clusters can be managed by a single vCenter server. When using Cisco Intersight for management and monitoring of Cisco HyperFlex clusters, there are no practical limits to the number of clusters being managed.

Cisco HyperFlex All-NVMe HXAF220c-M5N model servers are limited to a maximum of sixteen nodes per cluster and are not allowed to deploy more compute-only nodes than converged nodes, regardless of licensing.

Cisco HyperFlex HX240c-M5L model servers with large form factor (LFF) disks are limited to a maximum of sixteen nodes per cluster and cannot be mixed within the same cluster as models with small form factor (SFF) disks. In the case where the HX240c-M5L nodes use the 12 TB capacity disks, the maximum number of converged nodes is limited to 8.

Cisco HyperFlex systems deployed in a stretched cluster configuration require a minimum of two Cisco HX-series converged nodes per physical site and support a maximum of sixteen converged nodes per physical site when using small-form-factor (SFF) disks. When using large-form-factor (LFF) disks, the maximum number of converged nodes allowed in a stretched cluster is 8. Each site requires a pair of Cisco UCS Fabric Interconnects, to form an individual UCS domain in both sites.

lists the minimum and maximum scale for various installations of the Cisco HyperFlex system.

Table 10   HyperFlex Cluster Scale

Cluster Type	Minimum Converged Nodes Required	Maximum Converged Nodes	Maximum Compute-only Nodes Allowed	Maximum Total Cluster Size
Standard with SFF disks	3	32	32	64
Standard with LFF disks	3	16	32	48
Standard with 12 TB LFF disks	3	8	16	24
Standard with all-NVMe disks	3	16	16	32
Stretched with SFF disks	2 per site	16 per site	21 per site	32 per site

Cluster Type	Minimum Con-verged Nodes Required	Maximum Con-verged Nodes	Maximum Compute-only Nodes Al-lowed	Maximum Total Cluster Size
Stretched with LFF disks	2 per site	8 per site	16 per site	64 per cluster

Capacity

Overall usable cluster capacity is based on a number of factors. The number of nodes in the cluster, the number and size of the capacity layer disks, and the replication factor of the HyperFlex HX Data Platform, all affect the cluster capacity. In addition, configuring a cluster as a stretched cluster across two sites modifies the data dis-tribution method, which reduces capacity in favor of data availability. Caching disk sizes are not calculated as part of the cluster capacity.

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 120×10^9 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 percent.

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

Table 11   SI Unit Values (Decimal Prefix)

Value	Symbol	Name
1000 bytes	kB	Kilobyte
1000 kB	MB	Megabyte
1000 MB	GB	Gigabyte
1000 GB	TB	Terabyte

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 12   IEC Unit Values (binary prefix)

Value	Symbol	Name
1024 bytes	KiB	Kibibyte
1024 KiB	MiB	Mebibyte

Value	Symbol	Name
1024 MiB	GiB	Gibibyte
1024 GiB	TiB	Tebibyte

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 and HyperFlex Connect GUI when viewing cluster capacity, allocation, and consumption, and also within most operating systems.

Table 13 lists a set of HyperFlex HX Data Platform cluster usable capacity values, using binary prefix, for an array of cluster configurations. These values provide an example of the capacity calculations, for determining the appropriate size of HX cluster to initially purchase, and how much capacity can be gained by adding capacity disks.

Table 13   Cluster Usable Capacities

HX-Series Server Model	Node Quantity	Capacity Disk Size (each)	Capacity Disk Quantity (per node)	Cluster Usable Capacity at RF=2	Cluster Usable Capacity at RF=3
HXAF220c-M5SX	8	3.8 TB	8	102.8 TiB	68.6 TiB
		960 GB	8	25.7 TiB	17.1 TiB
		800 GB	8	21.4 TiB	14.3 TiB
HXAF240c-M5SX	8	3.8 TB	6	77.1 TiB	51.4 TiB
			15	192.8 TiB	128.5 TiB
			23	295.7 TiB	197.1 TiB
		960 GB	6	19.3 TiB	12.9 TiB
			15	48.2 TiB	32.1 TiB
			23	73.9 TiB	49.3 TiB
		800 GB	6	16.1 TiB	10.7 TiB
			15	40.2 TiB	26.8 TiB
			22	58.9 TiB	39.3 TiB

HX-Series Server Model	Node Quantity	Capacity Disk Size (each)	Capacity Disk Quantity (per node)	Cluster Usable Capacity at RF=2	Cluster Usable Capacity at RF=3
HX240c-M5L	8	6 TB	6	120.5 TiB	80.3 TiB
			12	241.0 TiB	160.7 TiB
		8 TB	6	160.7 TiB	107.1 TiB
			12	321.3 TiB	214.2 TiB

Capacity calculations methods for all servers are identical regardless of model. Calculations are based upon the number of nodes, the number of capacity disks per node, and the size of the capacity disks. Table 13 is not a comprehensive list of all capacities and models available.

Design Elements

Installing the HyperFlex system is done via the Cisco Intersight online management portal, or through a deployable HyperFlex installer virtual machine, available for download at cisco.com as an OVA file. The installer performs most of the Cisco UCS configuration work, and also performs significant portions of the ESXi configuration. Finally, the installer will 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 prerequisite steps needed for installation, and how to then utilize the HyperFlex Installer for the remaining configuration steps. This document focuses on the use of Cisco Intersight for the initial deployment of a Cisco HyperFlex cluster.

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 via 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 up-

links 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, however 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 instead be directed over the Cisco UCS uplinks because that traffic must travel from fabric A to fabric B, or vice-versa. 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. Cisco recommends that the uplink bandwidth configured is greater than or equal to double the bandwidth available to each Hyperflex converged node. For example, if the nodes are connected at 10 Gigabit speeds, then each Fabric Interconnect should have at least 20 Gigabit of uplink bandwidth available. 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 25 Connectivity with Single Uplink to Single Switch
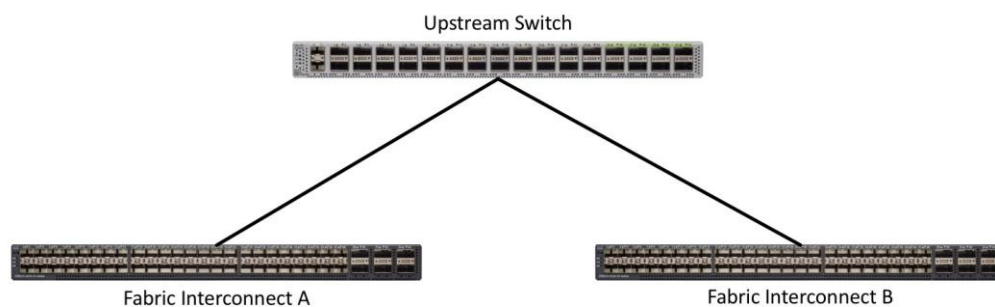
Port Channels to Single Switch

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

Figure 26 Connectivity with Port-Channels to Single Switch
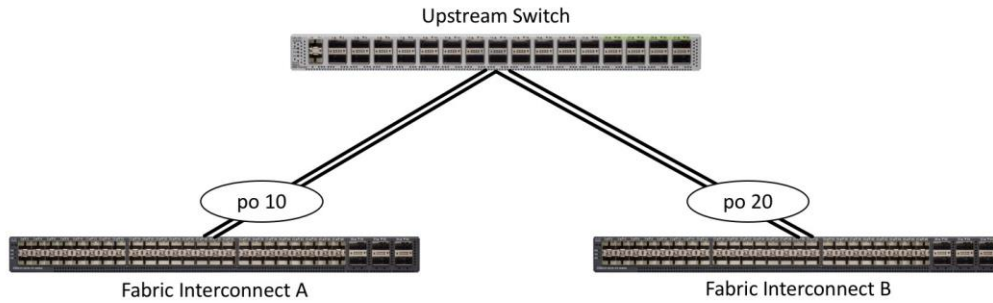
Single Uplinks or Port Channels to Multiple Switches

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 via 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 Figure 27 could also be port-channels.

Figure 27 Connectivity with Multiple Uplink Switches
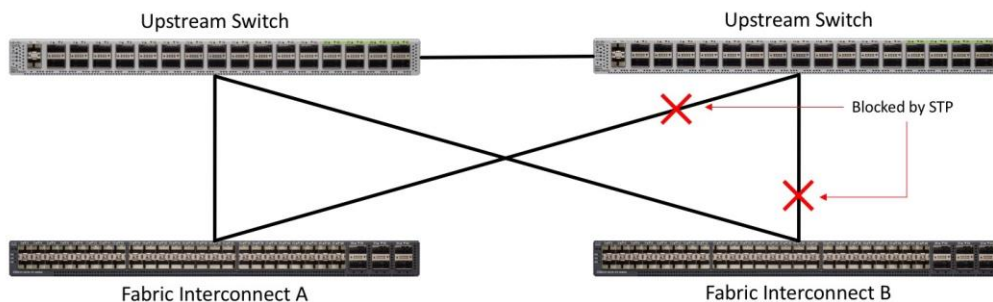
vPC to Multiple Switches

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 Cisco 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.

Figure 28 Connectivity with vPC
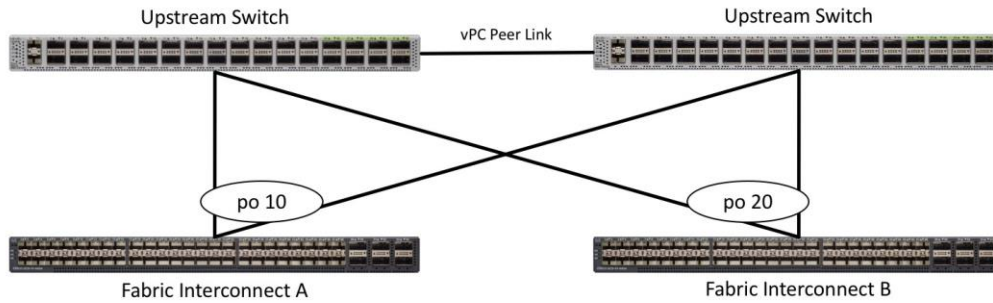
VLANs and Subnets

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 hx-storage-data VLAN must be a separate VLAN ID from the remaining VLANs. Table 14 lists the VLANs created by the Hyper-Flex installer in Cisco UCS, and their functions:

Table 14   VLANs

VLAN Name	VLAN ID	Purpose
hx-inband-mgmt	Customer supplied	ESXi host management interfaces HX Storage Controller VM management interfaces HX Storage Cluster roaming management interface
hx-inband-repl	Customer supplied	HX Storage Controller VM Replication interfaces HX Storage Cluster roaming replication interface
hx-storage-data	Customer supplied	ESXi host storage VMkernel interfaces HX Storage Controller storage network interfaces HX Storage Cluster roaming storage interface
vm-network	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 data centers. 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 by default 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. In addition, the default MTU for the hx-vmotion VLAN is also set to use jumbo frames. 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 configuration 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, including Cisco UCS firmware upgrades, or when a cable or port failure would cause storage traffic to traverse the northbound Cisco UCS uplink switches.

HyperFlex clusters can be configured to use standard size frames of 1500 bytes, however Cisco recommends that this configuration only be used in environments where the Cisco UCS uplink switches are not capable of passing jumbo frames, and that jumbo frames be enabled in all other situations.

Cisco UCS Design

This section describes 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, external management 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 new Cisco UCS sub-organization is created. The sub-organization is created underneath the root level of the Cisco UCS hierarchy, and is used to contain all policies, pools, templates, and service profiles used by HyperFlex, which prevents problems from overlapping settings across policies and pools. This arrangement also allows for organizational control using Role-Based Access Control (RBAC) and administrative locales within Cisco UCS Manager 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 29 Cisco UCS HyperFlex Sub-Organization
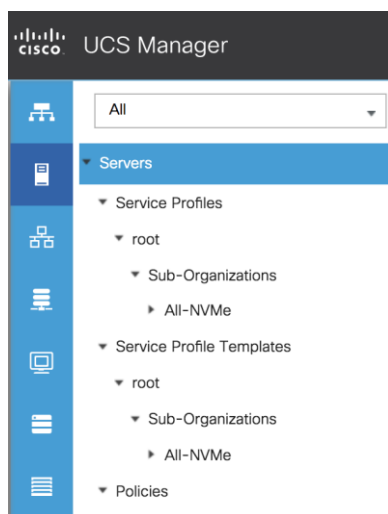
Cisco UCS LAN Policies

QoS System Classes

Specific Cisco 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. Table 15 and Figure 30 details the QoS System Class settings configured for HyperFlex:

Table 15   QoS System Classes

Priority	Enabled	CoS	Packet Drop	Weight	MTU	Multicast Optimized
Platinum	Yes	5	No	4	9216	No
Gold	Yes	4	Yes	4	Normal	No
Silver	Yes	2	Yes	Best-effort	Normal	Yes
Bronze	Yes	1	Yes	Best-effort	9216	No
Best Effort	Yes	Any	Yes	Best-effort	Normal	No
Fibre Channel	Yes	3	No	5	FC	N/A

Figure 30 QoS System Classes

Priority	Enabled	CoS	Packet Drop	Weight	Weight (%)	MTU	Multicast Optimized
Platinum	☑	5	☐	4	25	9216	☐
Gold	☑	4	☑	4	25	normal	☐
Silver	☑	2	☑	best-effort	6	normal	☑
Bronze	☑	1	☑	best-effort	6	9216	☐
Best Effort	☑	Any	☑	best-effort	6	normal	☐
Fibre Channel	☑	3	☐	5	32	fc	N/A

⚠ Changing the QoS system classes on a Cisco UCS 6332 or 6332-16UP model Fabric Interconnect requires both FIs to reboot in order to take effect.

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. Table 16 lists the QoS Policies configured for HyperFlex and their default assignment to the vNIC templates created.

Table 16   HyperFlex QoS Policies

Policy	Priority	Burst	Rate	Host Control	Used by vNIC Template
Platinum	Platinum	10240	Line-rate	None	storage-data-a storage-data-b
Gold	Gold	10240	Line-rate	None	vm-network-a vm-network-b
Silver	Silver	10240	Line-rate	None	hv-mgmt-a hv-mgmt-b
Bronze	Bronze	10240	Line-rate	None	hv-vmotion-a hv-vmotion-b
Best Effort	Best Effort	10240	Line-rate	None	N/A

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. Table 17 and Figure 31 detail the Multicast Policy configured for HyperFlex:

Table 17   Multicast Policy

Name	IGMP Snooping State	IGMP Snooping Querier State
HyperFlex	Enabled	Disabled

Figure 31 Multicast Policy

VLANs

VLANs are created by the HyperFlex installer to support a base HyperFlex system, with a VLAN for vMotion, and a single or multiple VLANs defined for guest VM traffic. 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). Table 18 lists the VLANs configured for HyperFlex.

Table 18   Cisco UCS VLANs

Name	ID	Type	Transport	Native	VLAN Sharing	Multicast Policy
<<hx-inband-mgmt>>	<<us-er_defi ned>>	LAN	Ether	No	None	HyperFlex
<<hx-inband-repl>>	<<us-er_defi ned>>	LAN	Ether	No	None	HyperFlex
<<hx-storage-data>>	<<us-er_defi ned>>	LAN	Ether	No	None	HyperFlex
<<vm-network>>	<<us-er_defi ned>>	LAN	Ether	No	None	HyperFlex
<<hx-vmotion>>	<<us-er_defi ned>>	LAN	Ether	No	None	HyperFlex

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 Cisco 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. A new IP pool, named "hx-ext-mgmt" is created in the HyperFlex sub-organization, and populated with a block of IP addresses, a subnet mask, and a default gateway by the HyperFlex installer. The default IP pool named "ext-mgmt", in the root organization is no longer used as of HyperFlex 2.5 for new installations.

Figure 32 Management IP Address Pool

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 via Cisco UCS Service Profiles. The number of virtual NIC (vNIC) interfaces, their VLAN associations, MAC addresses, QoS policies and more are all applied dynamically as part of the service profile 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 Cisco UCS fabric and the vNIC placement order. Finally, the last byte is incremented according to the number of MAC addresses created in the pool, which by default is 100. To avoid overlaps, when you define the values in the HyperFlex installer you must ensure that the first four bytes of the MAC

address pools are unique for each HyperFlex cluster installed in the same layer 2 network, and also different from MAC address pools in other Cisco UCS domains which may exist.

Table 19 lists the MAC Address Pools configured for HyperFlex and their default assignment to the vNIC templates created:

Table 19   MAC Address Pools

Name	Block Start	Size	Assignment Order	Used by vNIC Template
hv-mgmt-a	00:25:B5:<xx>:A1:01	100	Sequential	hv-mgmt-a
hv-mgmt-b	00:25:B5:<xx>:B2:01	100	Sequential	hv-mgmt-b
hv-vmotion-a	00:25:B5:<xx>:A7:01	100	Sequential	hv-vmotion-a
hv-vmotion-b	00:25:B5:<xx>:B8:01	100	Sequential	hv-vmotion-b
storage-data-a	00:25:B5:<xx>:A3:01	100	Sequential	storage-data-a
storage-data-b	00:25:B5:<xx>:B4:01	100	Sequential	storage-data-b
vm-network-a	00:25:B5:<xx>:A5:01	100	Sequential	vm-network-a
vm-network-b	00:25:B5:<xx>:B6:01	100	Sequential	vm-network-b

Figure 33 MAC Address Pools

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 regis-tration, MAC address forging, and the action taken on the vNIC status if the Cisco 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 carry-ing guest VM traffic. This allows for more flexibility, even though the policies are currently configured with the same settings. Table 20 lists the Network Control Policies configured for HyperFlex, and their default assignment to the vNIC templates created:

Table 20   Network Control Policy

Name	CDP	MAC Register Mode	Action on Uplink Fail	MAC Security	Used by vNIC Template
HyperFlex-infra	Enabled	Only Native VLAN	Link-down	Forged: Allow	hv-mgmt-a hv-mgmt-b hv-vmotion-a hv-vmotion-b storage-data-a storage-data-b

Name	CDP	MAC Register Mode	Action on Uplink Fail	MAC Security	Used by vNIC Template
HyperFlex-vm	Enabled	Only Native VLAN	Link-down	Forged: Allow	vm-network-a vm-network-b

Figure 34 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. An additional feature named "vNIC Redundancy" allows vNICs to be configured in pairs, so that the settings of one vNIC template, designated as a primary template, will automatically be applied to a configured secondary template. For all HyperFlex vNIC templates, the "A" side vNIC template is configured as a primary template, and the related "B" side vNIC template is a secondary. In each case, the only configuration difference between the two templates is which fabric they are configured to connect through. The following tables detail the initial settings in each of the vNIC templates created by the HyperFlex installer.

Table 21   vNIC Template hv-mgmt-a

vNIC Template Name:	hv-mgmt-a
Setting	Value
Fabric ID	A
Fabric Failover	Disabled

vNIC Template Name:	hv-mgmt-a	
Target	Adapter	
Type	Updating Template	
MTU	1500	
MAC Pool	hv-mgmt-a	
QoS Policy	silver	
Network Control Policy	HyperFlex-infra	
VLANs	<<hx-inband-mgmt>>	Native: No

Table 22   vNIC Template hv-mgmt-b

vNIC Template Name:	hv-mgmt-b	
Setting	Value	
Fabric ID	B	
Fabric Failover	Disabled	
Target	Adapter	
Type	Updating Template	
MTU	1500	
MAC Pool	hv-mgmt-b	
QoS Policy	silver	
Network Control Policy	HyperFlex-infra	
VLANs	<<hx-inband-mgmt>>	Native: No

Table 23   vNIC Template hv-vmotion-a

vNIC Template Name:	hv-vmotion-a
Setting	Value
Fabric ID	A

vNIC Template Name:	hv-vmotion-a	
Fabric Failover	Disabled	
Target	Adapter	
Type	Updating Template	
MTU	9000	
MAC Pool	hv-vmotion-a	
QoS Policy	bronze	
Network Control Policy	HyperFlex-infra	
VLANs	<<hx-vmotion>>	Native: No

Table 24   vNIC Template hx-vmotion-b

Drive Function	Drive PID	Applicable Platforms	Version
7.12TB SED SSD Capacity drive	HX-SD76TBEM2NK9	All existing HX M5 servers except All NVMe	4.0(2c)
SKUs for existing 3.8TB and 960GB Capacity drive capacity	HX-SD38TBEM2NK9, HX-SD960GBM2NK9	All existing HX M5 servers except All NVMe	4.0(2c)
Alternate system (or house-keeping) drive	HX-SD480G6I1X-EV	All existing HX M5 servers except All NVMe	4.0(2b)
Alternate system (or house-keeping) drive	HX-SD480GM1X-EV	All existing HX M5 servers except All NVMe	4.0(2b)
New 960G FIPS compliant SED SSD data drives	HX-SD960G2HTNK9	HXAF220C-M5SX, HXAF240C-M5SX	4.0(2b)
Alternate boot drive	HX-M2-960GB	All existing HX M5 servers	4.0(2a)
All NVMe 4TB Capacity drive	HX-NVME2H-I4000	All NVMe: HXAF220C-M5SN	4.0(2a)
New high density All NVMe 8TB Capacity drive	HX-NVMEHW-I8000	All NVMe: HXAF220C-M5SN	4.0(2a)
New high density All Flash			

Drive Function	Drive PID	Applicable Platforms	Version
7.12TB Capacity drive			
Maximum number of drives per node is 12 drives on HX AF C240.	HX-SD76T61X-EV	All Flash Configuration – namely: HXAF220C-M5SX, HXAF240C-M5SX, HXAF-E-220M5SX	
ESX support only.	4.0(2a)		
New 3.8TB FIPS compliant SED SSD data drives	HX-SD38T2HTNK9	HXAF220C-M5SX, HXAF240C-M5SX	4.0(2a)
Alternate drive for 8TB LFF capacity	HX-HD8T7K4KAN	HX240C-M5L	4.0(2a)

vNIC Template Name:	hv-vmotion-b	
Setting	Value	
Fabric ID	B	
Fabric Failover	Disabled	
Target	Adapter	
Type	Updating Template	
MTU	9000	
MAC Pool	hv-vmotion-b	
QoS Policy	bronze	
Network Control Policy	HyperFlex-infra	
VLANs	<<hx-vmotion>>	Native: No

Table 25   vNIC Template storage-data-a

Drive Function	Drive PID	Applicable Platforms	Version
7.12TB SED SSD Capacity drive	HX-SD76TBEM2NK9	All existing HX M5 servers except All NVMe	4.0(2c)
SKUs for existing 3.8TB and 960GB Capacity drive capacity	HX-SD38TBEM2NK9, HX-SD960GBM2NK9	All existing HX M5 servers except All NVMe	4.0(2c)
Alternate system (or house-	HX-SD480G6I1X-EV	All existing HX M5 servers except All NVMe	4.0(2b)

Drive Function	Drive PID	Applicable Platforms	Version
keeping) drive			
Alternate system (or house-keeping) drive	HX-SD480GM1X-EV	All existing HX M5 servers except All NVMe	4.0(2b)
New 960G FIPS compliant SED SSD data drives	HX-SD960G2HTNK9	HXAF220C-M5SX, HXAF240C-M5SX	4.0(2b)
Alternate boot drive	HX-M2-960GB	All existing HX M5 servers	4.0(2a)
All NVMe 4TB Capacity drive	HX-NVME2H-I4000	All NVMe: HXAF220C-M5SN	4.0(2a)
New high density All NVMe 8TB Capacity drive	HX-NVMEHW-I8000	All NVMe: HXAF220C-M5SN	4.0(2a)
New high density All Flash 7.12TB Capacity drive			
Maximum number of drives per node is 12 drives on HX AF C240.	HX-SD76T61X-EV	All Flash Configuration – namely: HXAF220C-M5SX, HXAF240C-M5SX, HXAF-E-220M5SX	
ESX support only.	4.0(2a)		
New 3.8TB FIPS compliant SED SSD data drives	HX-SD38T2HTNK9	HXAF220C-M5SX, HXAF240C-M5SX	4.0(2a)
Alternate drive for 8TB LFF capacity	HX-HD8T7K4KAN	HX240C-M5L	4.0(2a)

vNIC Template Name:	storage-data-a
Setting	Value
Fabric ID	A
Fabric Failover	Disabled
Target	Adapter
Type	Updating Template
MTU	9000

MAC Pool	storage-data-a
QoS Policy	platinum
Network Control Policy	HyperFlex-infra
VLANs	<<hx-storage-data>>	Native: No

Table 26   vNIC Template storage-data-b

vNIC Template Name:	storage-data-b	
Setting	Value	
Fabric ID	B	
Fabric Failover	Disabled	
Target	Adapter	
Type	Updating Template	
MTU	9000	
MAC Pool	storage-data-b	
QoS Policy	platinum	
Network Control Policy	HyperFlex-infra	
VLANs	<<hx-storage-data>>	Native: No

Table 27   vNIC Template vm-network-a

vNIC Template Name:	vm-network-a
Setting	Value
Fabric ID	A
Fabric Failover	Disabled
Target	Adapter
Type	Updating Template
MTU	1500

vNIC Template Name:	vm-network-a	
MAC Pool	vm-network-a	
QoS Policy	gold	
Network Control Policy	HyperFlex-vm	
VLANs	<<vm-network>>	Native: no

Table 28   vNIC Template vm-network-b

vNIC Template Name:	vm-network-b	
Setting	Value	
Fabric ID	B	
Fabric Failover	Disabled	
Target	Adapter	
Type	Updating Template	
MTU	1500	
MAC Pool	vm-network-b	
QoS Policy	gold	
Network Control Policy	HyperFlex-vm	
VLANs	<<vm-network>>	Native: no

LAN Connectivity Policies

Cisco 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, then 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. Table 29 lists the LAN Connectivity Policy configured for HyperFlex:

Table 29   LAN Connectivity Policy

Policy Name	Use vNIC Template	vNIC Name	vNIC Template Used	Adapter Policy
HyperFlex	Yes	hv-mgmt-a	hv-mgmt-a	HyperFlex
		hv-mgmt-b	hv-mgmt-b	
		hv-vmotion-a	hv-vmotion-a	
		hv-vmotion-b	hv-vmotion-b	
		storage-data-a	storage-data-a	
		storage-data-b	storage-data-b	
		vm-network-a	vm-network-a	
		vm-network-b	vm-network-b	

Cisco UCS Servers 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 Adapter Policy named "HyperFlex", configured for HyperFlex:

Figure 35 Cisco UCS Adapter Policy Resources

Figure 36 Cisco UCS Adapter Policy Options

BIOS Policies

Cisco UCS Manager utilizes policies applied via the service profiles, in order to modify settings in the BIOS of the associated server. Cisco HX-Series M5 generation servers no longer use predefined BIOS setting defaults derived from Cisco UCS Manager, instead the servers have default BIOS tokens set from the factory. The current default token settings can be viewed at the following website:
https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/Reference-Docs/Server-BIOS-Tokens/4-0/b_UCS_BIOS_Tokens_Guide_4_0.html

A BIOS policy named "HyperFlex-m5" is created by the HyperFlex installer to modify the settings of the M5 generation servers. The modified settings are as follows:

- System altitude is set to "Auto"

- CPU performance is set to "HPC"

- CPU direct cache access is set to "Enabled"

- Intel Virtualization Technology is set to "Enabled"

- IMC Interleave is set to "Auto"

- Sub NUMA clustering is set to "Disabled"

- Processor C states are all set to "Disabled"

- Power Technology is set to "Performance"

- Energy Performance is set to "Performance"

- LLC Prefetch is set to "Disabled"

- XPT Prefetch is set to "Disabled"

- Intel VTD coherency support is set to "Disabled"

- Intel VT for Directed IO is set to "Enabled"

- Intel VTD interrupt Remapping is set to "Enabled"

- Serial Port A is enabled

- PCI Memory mapped IO above 4GB is set to "Enabled"

- Console Redirection is set to "Serial Port A"

- Out of band management is set to "Enabled"

A third BIOS policy named "HyperFlex-nvme" is also created with the same settings as found in the "HyperFlex-m5" policy above.

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 M5 generation rack-mount servers have their VMware ESXi hypervisors installed to an internal M.2 SSD boot drive; therefore, they require a unique boot policy defining that the servers should boot from that location. The HyperFlex installer configures a boot policy named "HyperFlex-m5" specifying boot from the M.2 SSDs, referred to as "Embedded Disk", which is used by the HyperFlex M5 converged nodes, and should not be modified. The HyperFlex installer configures a boot policy named "hx-compute-m5", which can be modified as needed for the boot method used by the M5 generation compute-only nodes. Figure 37 details the HyperFlex Boot Policy.

Figure 37 Cisco UCS M5 Boot Policy

Host Firmware Packages

Cisco UCS Host Firmware Packages represent one of the most powerful features of the Cisco UCS platform; the ability to control the firmware revision of all the managed blades and rack-mount servers via 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-m5"** which uses the simple package definition method, applying firmware revisions to all components that matches a specific Cisco UCS firmware bundle, versus defining the firmware revisions part by part. Figure 38 details the Host Firmware Package configured by the HyperFlex installer:

Figure 38 Cisco UCS M5 Host Firmware Package

Actions	Properties		
Delete	Name	: **HyperFlex-m5**	
Show Policy Usage	Description	: Recommended Host Firmware Packages for M5 Hyp	
Use Global	Owner	: **Local**	
Modify Package Versions	Blade Package : **4.0(4d)B**		Blade Backup Package :
Modify Backup Package Versions	Rack Package : **4.0(4d)C**		Rack Backup Package :
	Service Pack :		

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 four Local Disk Configuration Policies which allows any local disk configuration. **The policy named "HyperFlex-m5" is used by the service profile tem-plate named "hx-nodes-m5", which is for the HyperFlex M5 generation converged servers, and should not** be modified.

Meanwhile, the policies named "hx-compute" and "hx-compute-m5" are used by the service profile templates named "compute-nodes" and "compute-nodes-m5", which are used by compute-only nodes. The two com-pute-only node policies can be modified as needed to suit the local disk configuration that will be used in com-pute-only nodes.

Figure 39 details the Local Disk Configuration Policy configured by the HyperFlex installer.

Figure 39 Cisco UCS M5 Local Disk Configuration Policy

Additional policies are created for use by Cisco UCS M4 generation HX-series servers, including additional BIOS policies, Boot Policies, Host Firmware Packages and Local Disk Configuration Policies. Because this document no longer describes the installation and configuration of M4 generation hardware, the settings in these policies are not outlined here. Please refer to previous editions of this Cisco Validated Design document as a reference for these policies targeted at M4 generation hardware.

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 ser**vice 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 Manager, 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".** In addition, the On Next Boot setting is enabled, which will automatically apply changes the next time the server is rebooted, without any secondary acknowledgement. Figure 40 details the Maintenance Policy configured by the HyperFlex installer:

Figure 40 Cisco UCS Maintenance Policy

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. Figure 41 details the Power Control Policy configured by the HyperFlex installer:

Figure 41 Cisco UCS Power Control Policy

Properties

Name   : **HyperFlex**

Description   : Recommended Power control policy for HyperFlex s

Owner   : **Local**

Fan Speed Policy : Any ▼

Power Capping

If you choose **cap**, the server is allocated a certain amount of power based on its priority within its being the highest priority. If you choose **no-cap**, the server is exempt from all power capping.

⦿ No Cap ○ cap

Cisco UCS Manager only enforces power capping when the servers in a power group require more servers run at full capacity regardless of their priority.

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. Figure 42 details the Scrub Policy configured by the HyperFlex installer:

Figure 42 Cisco UCS Scrub Policy

Properties

Name   : **HyperFlex**

Description   : Recommended Scrub policy for HyperFlex servers

Owner   : **Local**

Disk Scrub   : ⦿ No ○ Yes

BIOS Settings Scrub   : ⦿ No ○ Yes

FlexFlash Scrub   : ⦿ No ○ Yes

Persistent Memory Scrub : ⦿ No ○ Yes

Serial over LAN Policies

Cisco UCS Serial over LAN (SoL) Policies enable console output which is sent to the serial port of the server, to be accessible via 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 ad-

ministrators are working remotely, the ability to send and receive that traffic via the LAN is very helpful. Connections to a SoL session can be initiated from Cisco UCS Manager. The HyperFlex installer creates a SoL policy named "HyperFlex" to enable SoL sessions and uses this feature to configure the ESXi hosts' management networking configuration. Figure 43 details the SoL Policy configured by the HyperFlex installer:

Figure 43 Cisco UCS Serial over LAN Policy

Properties

Name	:	**HyperFlex**
Description	:	Recommended Serial over LAN policy for HyperFlex
Owner	:	**Local**
Serial over LAN State :		○ Disable ● Enable
Speed	:	115200 ▼

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.

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 a group of 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 service profile templates named "hx-nodes-m5" and "compute-nodes-m5", each with nearly the same configuration, except for the BIOS, firmware, local disk configuration and boot policies. 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 tables detail the service profile templates configured by the HyperFlex installer:

Table 30 Cisco UCS Service Profile Template Settings and Values

Service Profile Template Name:	hx-nodes-m5
Setting	**Value**
UUID Pool	Hardware Default
Associated Server Pool	None

Service Profile Template Name:	hx-nodes-m5
Maintenance Policy	HyperFlex
Management IP Address Policy	hx-ext-mgmt
Local Disk Configuration Policy	HyperFlex-m5
LAN Connectivity Policy	HyperFlex
Boot Policy	HyperFlex-m5
BIOS Policy	HyperFlex-m5
Firmware Policy	HyperFlex-m5
Power Control Policy	HyperFlex
Scrub Policy	HyperFlex
Serial over LAN Policy	HyperFlex
vMedia Policy	Not defined

Service Profile Template Name:	compute-nodes-m5
Setting	Value
UUID Pool	Hardware Default
Associated Server Pool	None
Maintenance Policy	HyperFlex
Management IP Address Policy	hx-ext-mgmt
Local Disk Configuration Policy	hx-compute-m5
LAN Connectivity Policy	HyperFlex
Boot Policy	hx-compute-m5
BIOS Policy	HyperFlex-m5

Service Profile Template Name:	compute-nodes-m5
Firmware Policy	HyperFlex-m5
Power Control Policy	HyperFlex
Scrub Policy	HyperFlex
Serial over LAN Policy	HyperFlex
vMedia Policy	Not defined

⚠ Additional templates are created for use by Cisco UCS M4 generation HX-series servers. Because this document no longer covers the installation and configuration of M4 generation hardware, the settings in these templates are not outlined here. Please refer to previous editions of this Cisco Validated Design document as a reference for these templates targeted at M4 generation hardware.

vNIC/vHBA Placement

In order to control the order of detection of the vNICs and vHBAs defined in service profiles, Cisco 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. In certain hardware configurations, the physical mapping of the installed 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 placement and detection order of the defined vNICs and vHBAs does not refer to physical cards, but instead refers to a vCon. HX-series servers are most often configured with a single Cisco UCS VIC mLOM card. An optional configuration does allow for two VIC cards to be used for an extra layer of physical re-dundancy. To accommodate this option, the vCon placement policy alternates between vCon 1 and vCon 2. If two cards were present, then the 8 vNICs would be evenly distributed across both cards. With a single Cisco VIC card installed, the only available placement is on vCon 1. In this scenario, all the vNICs defined in the service profile templates for HX-series servers will be placed on vCon 1, despite some of them being set to be placed on vCon 2. In either case, the resulting detection order is the same, giving a consistent enumeration of the inter-faces as seen by the VMware ESXi hypervisor.

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 same network inter-faces in a known and identical order, and they will always be connected to the same VLANs via the same net-work fabrics. Table 31 lists 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 31   vNIC Placement

vNIC	Placement	Order	Fabric	VLAN	ESXi interface enumeration
hv-mgmt-a	1	1	A	<<hx-inband-mgmt>>	vmnic0
hv-mgmt-b	2	5	B	<<hx-inband-	vmnic4

vNIC	Placement	Order	Fabric	VLAN	ESXi interface enumeration
				mgmt>>	
storage-data-a	1	2	A	<<hx-storage-data>>	vmnic1
storage-data-b	2	6	B	<<hx-storage-data>>	vmnic5
vm-network-a	1	3	A	<<vm-network>>	vmnic2
vm-network-b	2	7	B	<<vm-network>>	vmnic6
hv-vmotion-a	1	4	A	<<hx-vmotion>>	vmnic3
hv-vmotion-b	2	8	B	<<hx-vmotion>>	vmnic7

> ESXi VMDirectPath relies on a fixed PCI address for the passthrough devices. If the configuration is changed by adding or removing vNICs or vHBAs, then the order of the devices seen in the PCI tree will change. The ESXi hosts will subsequently need to reboot one additional time in order to repair the configuration, which they will do automatically.

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 Cisco 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. A third port group is created for cluster to cluster VM snapshot replication traffic. The VLANs are not Native VLANs as assigned to the vNIC templates, and therefore they are defined 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 via NFS. The switch has two uplinks, active on fabric B and standby on fabric A, with jumbo frames highly recommended. 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 templates, and therefore they are defined 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 VLANs are not Native VLANs as assigned to the vNIC templates, and therefore they are defined 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 highly recommended. The IP addresses of the VMkernel ports (vmk2) are configured during the post_install script execution. The VLAN is not a Native VLAN as assigned to the vNIC templates, and therefore they are defined in ESXi/vSphere.

The following table and figures provide more details about the ESXi virtual networking design as built by the HyperFlex installer by default:

Table 32   Virtual Switches

Virtual Switch	Port Groups	Active vmnic(s)	Passive vmnic(s)	VLAN IDs	Jumbo
vswitch-hx-inband-mgmt	Management Network Storage Controller Management Network	vmnic0	vmnic4	<<hx-in-band-mgmt>>	no
	Storage Controller Replication Network	vmnic0	vmnic4	<<hx-in-band-repl>>	no
vswitch-hx-storage-data	Storage Controller Data Network Storage Hypervisor Data Network	vmnic5	vmnic1	<<hx-storage-data>>	yes
vswitch-hx-vm-network	vm-network-<<VLAN ID>>	vmnic2 vmnic6		<<vm-network>>	no
vmotion	vmotion-<<VLAN ID>>	vmnic3	vmnic7	<<hx-vmotion>>	yes

Figure 44 ESXi Network Design
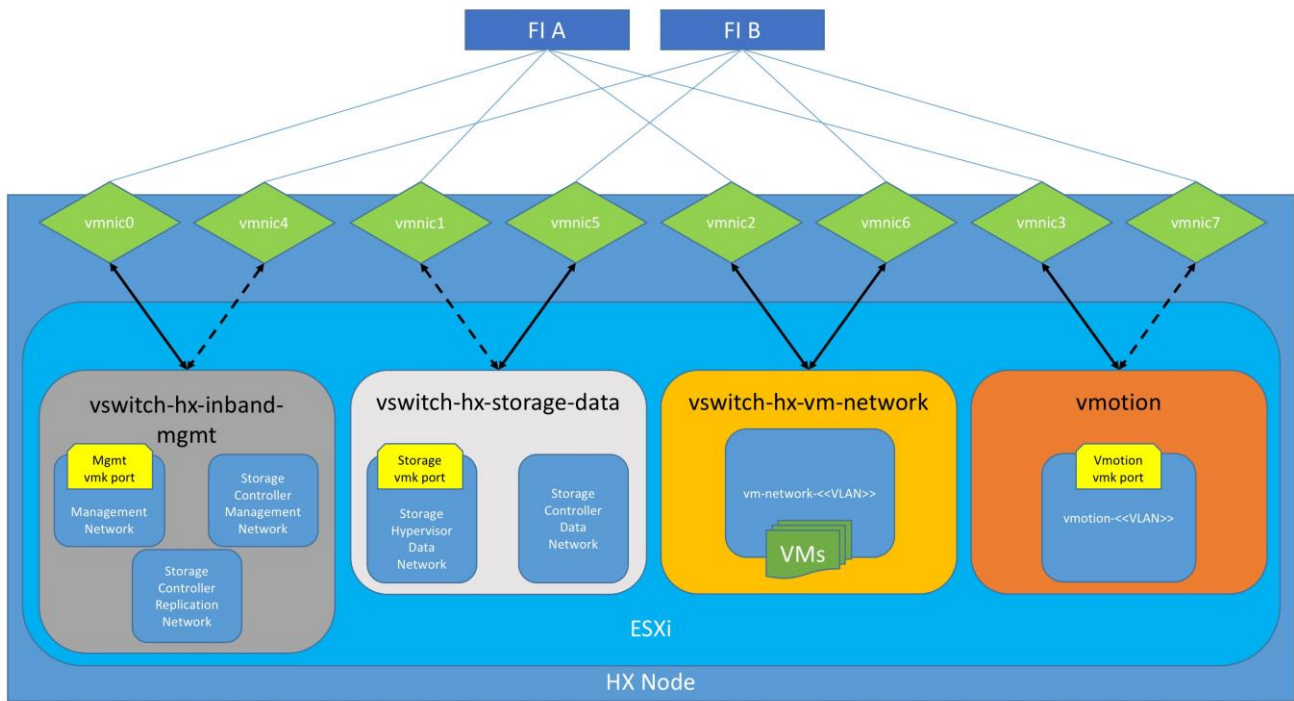
VMDirectPath I/O Passthrough

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 passthrough. With the appropriate driv-er 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 con-trol 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. In all-flash model servers equipped with an NVMe caching SSD, VMDirectPath is also configured for the caching disk, since it is not connected to an HBA card. In all-NVMe model servers there is no SAS HBA at all, and all of the NVMe caching and capacity SSDs are configured via VMDirectPath I/O so that the controller VMs have direct access to all of the disks. Other disks, connected to dif-ferent controllers, such as the M.2 boot SSDs, remain under the control of the ESXi hypervisor. Lastly, when the Cisco HyperFlex Acceleration Engine card is installed, VMDirectPath I/O is also configured to give the controller VMs direct access to the cards as well. The configuration of the VMDirectPath I/O feature is done by the Cisco HyperFlex installer and requires no manual steps.

Storage Platform Controller Virtual Machines

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 storage controller VM runs custom soft-ware and services that manage and maintain the Cisco HX Distributed Filesystem. The services and processes that run within the controller VMs are not exposed directly to the ESXi hosts, although the controller VMs are configured to automatically start and stop with the ESXi hosts and protected from accidental deletion. Manage-ment and visibility into the function of the controller VMs, and the Cisco HX Distributed Filesystem is done via the

HyperFlex Connect HTML management webpage, or 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 and vCenter plugins are all done by the Cisco HyperFlex installer and requires no manual steps.

Controller Virtual Machine Locations

The physical storage location of the controller VMs differs among the Cisco HX-Series rack servers, due to differences with the physical disk location and connections on those server models. 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 via VMDirectPath I/O. The configuration details of the models are as follows:

- HX220c M5, HXAF220c M5, HX240c M5L, HX240c M5 and HXAF240c M5: The server boots the ESXi hypervisor from the internal M.2 form factor SSD. The M.2 SSD is partitioned by the ESXi installer, and the remaining 216 GB of space is used as a VMFS datastore. The controller VM's root filesystem is stored on a 2.5 GB virtual disk, /dev/sda, which is placed on this VMFS datastore. The controller VM has full control of all the front and rear facing SAS based hot-swappable disks via PCI passthrough control of the SAS HBA. The controller VM operating system sees the 240 GB SSD, also commonly called the "housekeeping" disk as /dev/sdb, and places HyperFlex binaries and logs on this disk. The remaining disks seen by the controller VM OS are used by the HX Distributed filesystem for caching and capacity layers.

- HX220c M5N: The server boots the ESXi hypervisor from the internal M.2 form factor SSD. The M.2 SSD is partitioned by the ESXi installer, and the remaining 216 GB of space is used as a VMFS datastore. The controller VM's root filesystem is stored on a 2.5 GB virtual disk, /dev/sda, which is placed on this VMFS datastore. The controller VM has full control of all the front facing NVMe based hot-swappable SSDs directly connected through the PCIe bus via PCI Passthrough. The controller VM operating system sees the 240 GB SSD, also commonly called the "housekeeping" disk as /dev/sdb, and places HyperFlex binaries and logs on this disk. The remaining disks seen by the controller VM OS 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 45 All M5 Generation Servers Controller VM Placement Except All-NVMe
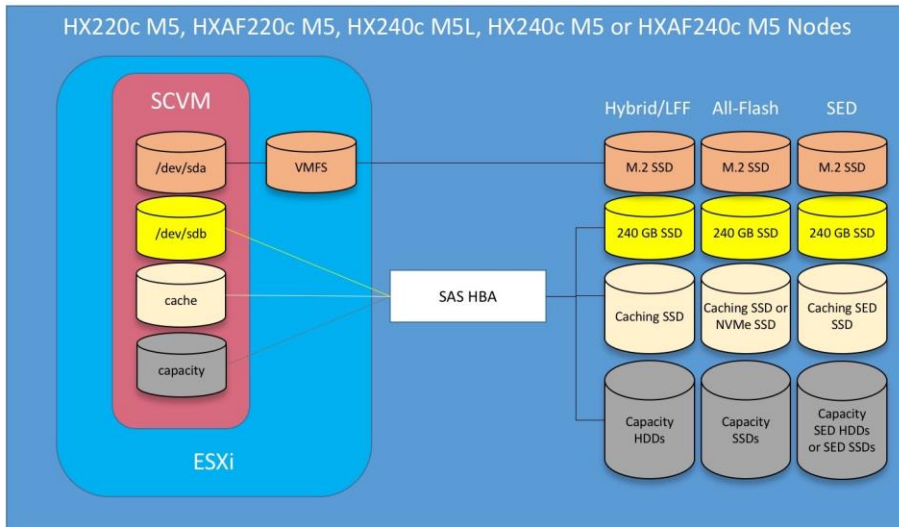
Figure 46 All-NVMe M5 Controller VM Placement
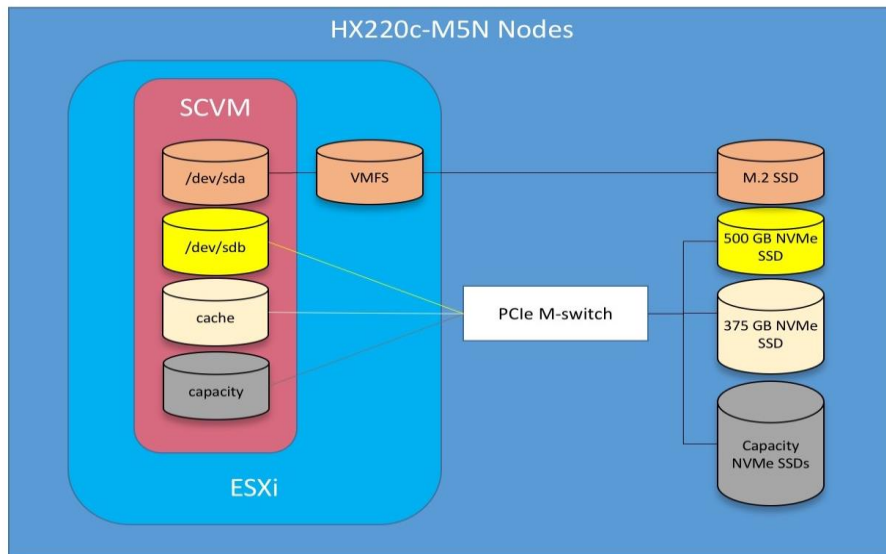
HyperFlex compute-only nodes install a lightweight controller VM in the VMFS datastore automatically created during the installation of ESXi. This VM performs no storage functions and is only used for node coordination.

HyperFlex Datastores

A new HyperFlex cluster has no default datastores configured for virtual machine storage, therefore the datastores must be created using the vCenter Web Client plugin or the HyperFlex Connect GUI. 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 HyperFlex Con-

nect or the vCenter plugin when actual space consumption results in low amounts of free space, and alerts will be sent via auto support email alerts. Overall space consumption in the HyperFlex clustered filesystem is optimized by the default deduplication and compression features.

Figure 39 Datastore Example
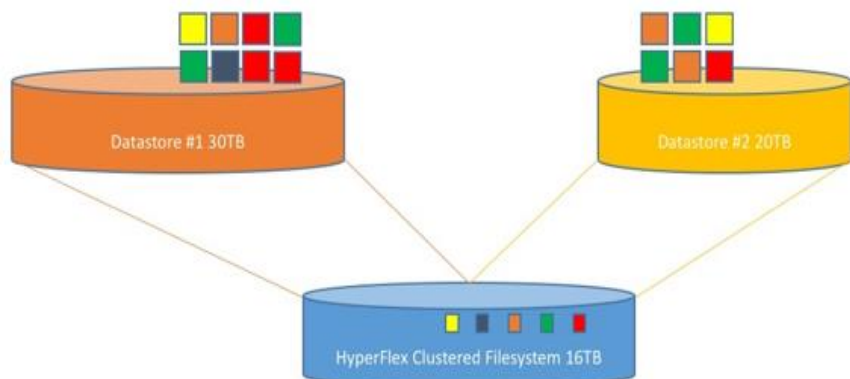
CPU Resource Reservations

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. This is a soft guarantee, meaning in most situations the SCVMs are not using all of the CPU resources reserved, therefore allowing the guest VMs to use them. Table 33 lists the CPU resource reservation of the storage controller VMs:

Table 33 Controller VM CPU Reservations

Server Models	Number of vCPU	Shares	Reservation	Limit
All hybrid and all-flash models	8	Low	10800 MHz	unlimited
All-NVMe models	12	Low	10800 MHz	unlimited

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. Table 34 lists the memory resource reservation of the storage controller VMs:

Table 34 Controller VM Memory Reservations

Server Models	Amount of Guest Memory	Reserve All Guest Memory
HX220c-M5SX	48 GB	Yes

Server Models	Amount of Guest Memory	Reserve All Guest Memory
HXAF220c-M5SX		
HXAF220c-M5N HX240c-M5SX HXAF240c-M5SX	72 GB	Yes
HX240c-M5L	78 GB	Yes

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 are reduced and simplified due to the previous factory work. For the purpose of this document, the setup process is described presuming that this factory pre-installation work was done, thereby leveraging the tools and processes developed by Cisco to simplify the process and dramatically reduce the deployment time. The following sections will guide you through the prerequisites and manual steps needed prior to using the HyperFlex installer via Cisco Intersight, how to configure the HyperFlex profiles in Cisco Intersight and perform the installation, then 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 management of the Cisco UCS cluster. In addition, at least one IP address per Cisco UCS blade or HX-series rack-mount server is required for the hx-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 Replication: These addresses are used by the HyperFlex Storage Platform Controller VMs for clusters that are configured to replicate VMs to one another. One IP address per HX node is required, plus

one additional IP address as a roaming clustered replication interface. These addresses are assigned to a pool as part of a post-installation activity described later in this document and are not needed to complete the initial installation of a HyperFlex cluster. These addresses can be from the same subnet as the Hyper-Flex and ESXi management addresses, but it is recommended that the VLAN ID and subnet be unique.

- 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. These addresses are automatically provisioned to the nodes from the link-local IPv4 subnet of 169.254.0.0/16 and do not need to be manually assigned prior to installation. Two IP addresses per node in the HyperFlex cluster are assigned from the subnet, and a single additional IP address is assigned as the roaming HyperFlex cluster storage interface. The third octet of the IP addresses is derived from the MAC address pool prefix by converting that value to a decimal number, thereby creating a unique subnet for each cluster, as the subnet mask set on the hosts for these VMkernel ports is actually 255.255.255.0. The value for the fourth octet is sequentially set, starting with .1 for the overall cluster, then proceeding to .2 for the vmk1(Hypervisor) port of the first server, then .3 for the Storage Controller VM of the first server. The second server would be assigned .4 for its vmk1 port, and .5 for its Storage Controller VM, and this pattern continues for each subsequent server. It is recommended to provision a VLAN ID that is not used in the network for other purposes. Finally, if the Cisco UCS domain is going to contain multiple HyperFlex clusters, it is recommended to use a different VLAN ID for the HyperFlex storage traffic for each cluster, as this is a safer method, guaranteeing that storage traffic from multiple clusters 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, although this configuration would require additional manual steps.

The following tables will assist in gathering the required IP addresses for the installation of an 8-node standard HyperFlex cluster, or a 4+4 extended cluster by listing the addresses required, plus an example IP configuration:

Table cells shaded in black do not require an IP address.

Table 35   HyperFlex Standard Cluster IP Addressing

Address Group:	UCS	HyperFlex and ESXi Management			HyperFlex Storage		VMotion
VLAN ID:							
Subnet:							
Subnet Mask:							
Gateway:							
Device	UCS Management Addresses	ESXi Management Interfaces	Storage Controller VM Management Interfaces	Storage Controller VM Replication Interfaces	ESXi Hypervisor Storage VMkernel Interfaces	Storage Controller VM Storage Interfaces	VMotion VMkernel Interfaces
Fabric Interconnect							
Fabric Interconnect							
UCS Manager							
HyperFlex Cluster							
HyperFlex Node #1							
HyperFlex Node #2							
HyperFlex Node #3							
HyperFlex Node #4							
HyperFlex Node #5							
HyperFlex Node #6							
HyperFlex Node #7							
HyperFlex Node #8							

If the on-premises HyperFlex installer VM is used instead of Cisco Intersight for the installation, then IP addresses for the HyperFlex storage components must be manually assigned and provided during the installation process.

HyperFlex extended clusters are also addressed similarly to a standard cluster, they require additional IP addresses for Cisco UCS management and ESXi management, as listed below:

Table 36   HyperFlex Standard Cluster Example IP Addressing

Address Group:	UCS	HyperFlex and ESXi Management			HyperFlex Storage		VMotion
VLAN ID:	132	30		133	101		35
Subnet:	10.29.132.0	10.29.132.0		192.168.101.0	169.254.0.0		192.168.201.0
Subnet Mask:	255.255.255.0	255.255.255.0		255.255.255.0	255.255.255.0		255.255.255.0
Gateway:	10.29.132.1	10.29.132.1		192.168.101.1			
Device	UCS Management Addresses	ESXi Management Interfaces	Storage Controller VM Management Interfaces	Storage Controller VM Replication Interfaces	ESXi Hypervisor Storage VMkernel Interfaces	Storage Controller VM Storage Interfaces	VMotion VMkernel Interfaces
Fabric Interconnect	10.29.132.104						
Fabric Interconnect	10.29.132.105						
UCS Manager	10.29.132.106						
HyperFlex Cluster			10.29.132.182	192.168.101.40			
HyperFlex Node #1	10.29.132.166	10.29.132.174	10.29.132.183	192.168.101.41			192.168.201.61
HyperFlex Node #2	10.29.132.167	10.29.132.175	10.29.132.184	192.168.101.42			192.168.201.62
HyperFlex Node #3	10.29.132.168	10.29.132.176	10.29.132.185	192.168.101.43			192.168.201.63
HyperFlex Node #4	10.29.132.169	10.29.132.177	10.29.132.186	192.168.101.44			192.168.201.64
HyperFlex Node #5	10.29.132.170	10.29.132.178	10.29.132.187	192.168.101.45			192.168.201.65
HyperFlex Node #6	10.29.132.171	10.29.132.179	10.29.132.188	192.168.101.46			192.168.201.66
HyperFlex Node #7	10.29.132.172	10.29.132.180	10.29.132.189	192.168.101.47			192.168.201.67
HyperFlex Node #8	10.29.132.173	10.29.132.181	10.29.132.190	192.168.101.48			192.168.201.68

IP addresses for Cisco UCS Management, plus HyperFlex and ESXi Management can come from the same subnet, or can be separate subnets, as long as the HyperFlex installer can reach them both.

DHCP versus Static IP

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

DNS

DNS servers are highly recommended to be configured for querying Fully Qualified Domain Names (FQDN) 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 and reverse PTR records **for the ESXi hypervisor hosts'** management interfaces. Additional DNS A records can be created for the Storage Controller Management inter-faces, 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, by listing the in-formation required, and an example configuration:

Table 37   DNS Server Information

Item	Value
DNS Server #1	
DNS Server #2	

Item	Value
DNS Domain	
vCenter Server Name	
SMTP Server Name	
UCS Domain Name	
HX Server #1 Name	
HX Server #2 Name	
HX Server #3 Name	
HX Server #4 Name	
HX Server #5 Name	
HX Server #6 Name	
HX Server #7 Name	
HX Server #8 Name	

Table 38   DNS Server Example Information

Item	Value
DNS Server #1	10.29.132.110
DNS Server #2	
DNS Domain	hxdom.local
vCenter Server Name	vcenter.hxdom.local
SMTP Server Name	outbound.cisco.com
UCS Domain Name	HX-FI
HX Server #1 Name	hxaf220m5n-01.hxdom.local
HX Server #2 Name	hxaf220m5n-02.hxdom.local

Item	Value
HX Server #3 Name	hxaf220m5n-03.hxdom.local
HX Server #4 Name	hxaf220m5n-04.hxdom.local
HX Server #5 Name	hxaf220m5n-05.hxdom.local
HX Server #6 Name	hxaf220m5n-06.hxdom.local
HX Server #7 Name	hxaf220m5n-07.hxdom.local
HX Server #8 Name	hxaf220m5n-08.hxdom.local

NTP

Consistent time clock synchronization is required across the components of the HyperFlex system, provided by reliable NTP servers, accessible for querying in the Cisco UCS Management network group, and the HyperFlex and ESXi Management group. NTP is used by Cisco UCS Manager, vCenter, the ESXi hypervisor hosts, and the HyperFlex Storage Platform Controller VMs. The use of public NTP servers is highly discouraged, instead a reliable internal NTP server should be used.

The following tables will assist with gathering the required NTP information for the installation by listing the information required, and an example configuration:

Table 39   NTP Server Information

Item	Value
NTP Server #1	
NTP Server #2	
Timezone	

Table 40   NTP Server Example Information

Item	Value
NTP Server #1	ntp1.hxdom.local
NTP Server #2	ntp2.hxdom.local
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 Cisco UCS Fabric Interconnects that comprise the HyperFlex system; a VLAN for the HyperFlex and ESXi Management group, a VLAN for the Hyper-Flex Storage group, a VLAN for the VMotion group, and at least one VLAN for the guest VM traffic. If HyperFlex Replication is to be used, another VLAN must be created and trunked for the replication traffic. The VLAN names and IDs must be supplied during the HyperFlex installation wizard.

The following tables will assist with gathering the required VLAN information for the installation by listing the information required, and an example configuration:

Table 41   VLAN Information

Name	ID
<<hx-inband-mgmt>>	
<<hx-inband-repl>>	
<<hx-storage-data>>	
<<hx-vm-data>>	

Table 42   VLAN Example Information

Name	ID
hx-mgmt	30
hx-repl	35
hx-storage	101
vm-network-100	34
vmotion-200	201

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 Cisco 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. Cisco recommends that the uplink bandwidth configured is greater than or equal to double the bandwidth available to each Hyperflex converged node. For example, if the nodes are connected at 10 Gigabit speeds, then each Fabric Interconnect should have at least 20 Gigabit of uplink bandwidth available.

The following tables will assist with gathering the required network uplink information for the installation by listing the information required, and an example configuration:

Table 43 Network Uplink Configuration

Fabric Interconnect Port	Port Channel	Port Channel Type	Port Channel ID	Port Channel Name
A	☐ Yes ☐ No	☐ LACP		
	☐ Yes ☐ No	☐ vPC		
	☐ Yes ☐ No			
	☐ Yes ☐ No			
B	☐ Yes ☐ No	☐ LACP		
	☐ Yes ☐ No	☐ vPC		
	☐ Yes ☐ No			
	☐ Yes ☐ No			

Table 44 Network Uplink Example Configuration

Fabric Interconnect Port	Port Channel	Port Channel Type	Port Channel ID	Port Channel Name
A	1/49 ☒ Yes ☐ No	☐ LACP		
	1/50 ☒ Yes ☐ No	☒ vPC	10	vpc-10
	☐ Yes ☐ No			
	☐ Yes ☐ No			
B	1/49 ☒ Yes ☐ No	☐ LACP		
	1/50 ☒ Yes ☐ No	☒ vPC	20	vpc-20
	☐ Yes ☐ No			
	☐ Yes ☐ No			

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 by listing the information required and an example configuration:

Table 45 Usernames and Passwords

Account	Username	Password
HX Installer Administrator	root	<<hx_install_root_pw>>
UCS Administrator	admin	<<ucs_admin_pw>>
ESXi Administrator	root	<<esxi_root_pw>>
HyperFlex Administrator	admin	<<hx_admin_pw>>
vCenter Administrator	<<vcenter_administrator>>	<<vcenter_admin_pw>>

Table 46 Example Usernames and Passwords

Account	Username	Password
HX Installer Administrator	root	Cixxxxxx
UCS Administrator	admin	Cixxxxxx
ESXi Administrator	root	Clxxxxxx
HyperFlex Administrator	admin	Clxxxxxx
vCenter Administrator	administrator@vsphere.local	Cixxxxxxx

Physical Installation

Install the Fabric Interconnects, the HX-Series rack-mount servers, standard Cisco UCS C-series rack-mount servers, the Cisco UCS 5108 chassis, the Cisco UCS Fabric Extenders, and the Cisco UCS blades according to their corresponding hardware installation guides listed below. For a stretched cluster deployment, the physical installation is identical to a standard cluster, only it is duplicated in two different physical locations.

Cisco UCS 6400 Series Fabric Interconnect:

https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/hw/6454-install-guide/6454.html

HX220c M5 Server:

https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HX_series/HX220c_M5/HX220c_M5.html

Cisco UCS 5108 Chassis, Servers, and Fabric Extenders:

http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/hw/chassis-install-guide/ucs5108_install.pdf

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.

Table 47 provides an example cabling map for installation of a Cisco HyperFlex system, with eight HyperFlex converged servers, and one Cisco UCS 5108 chassis.

Table 47   Example Cabling Map

Device	Port	Connected To	Port	Type	Length	Note
UCS6454-A	L1	UCS6454-B	L1	CAT5	1FT	
UCS6454-A	L2	UCS6454-B	L2	CAT5	1FT	
UCS6454-A	mgmt0	Customer LAN				
UCS6454-A	1/1	HX Server #1	mLOM port 1	Twinax	3M	Server 1
UCS6454-A	1/2	HX Server #2	mLOM port 1	Twinax	3M	Server 2
UCS6454-A	1/3	HX Server #3	mLOM port 1	Twinax	3M	Server 3
UCS6454-A	1/4	HX Server #4	mLOM port 1	Twinax	3M	Server 4
UCS6454-A	1/5	HX Server #5	mLOM port 1	Twinax	3M	Server 5
UCS6454-A	1/6	HX Server #6	mLOM port 1	Twinax	3M	Server 6
UCS6454-A	1/7	HX Server #7	mLOM port 1	Twinax	3M	Server 7
UCS6454-A	1/8	HX Server #8	mLOM port 1	Twinax	3M	Server 8
UCS6454-A	1/9	C220 Server #9	mLOM port 1	Twinax	3M	Server 9
UCS6454-A	1/10	C220 Server #10	mLOM port 1	Twinax	3M	Server 10
UCS6454-A	1/11	C220 Server #11	mLOM port 1	Twinax	3M	Server 11
UCS6454-A	1/12	C220 Server #12	mLOM port 1	Twinax	3M	Server 12
UCS6454-A	1/13	C220 Server #13	mLOM port 1	Twinax	3M	Server 13
UCS6454-A	1/14	C220 Server #14	mLOM port 1	Twinax	3M	Server 14
UCS6454-A	1/15	C220 Server #15	mLOM port 1	Twinax	3M	Server 15
UCS6454-A	1/16	C220 Server #16	mLOM port 1	Twinax	3M	Server 16
UCS6454-A	1/17	2408XP #1	IOM1 port 1	Twinax	3M	Chassis 1
UCS6454-A	1/18	2408XP #1	IOM1 port 2	Twinax	3M	Chassis 1
UCS6454-A	1/19					

Device	Port	Connected To	Port	Type	Length	Note
UCS6454-A	1/20					
UCS6454-A	1/21					
UCS6454-A	1/22					
UCS6454-A	1/23					
UCS6454-A	1/24					
UCS6454-A	1/25					
UCS6454-A	1/26					
UCS6454-A	1/27					
UCS6454-A	1/28					
UCS6454-A	1/29					
UCS6454-A	1/30					
UCS6454-A	1/31					
UCS6454-A	1/32					
UCS6454-A	1/33					
UCS6454-A	1/34					
UCS6454-A	1/35					
UCS6454-A	1/36					
UCS6454-A	1/37					
UCS6454-A	1/38					
UCS6454-A	1/39					
UCS6454-A	1/40					
UCS6454-A	1/41					

Device	Port	Connected To	Port	Type	Length	Note
UCS6454-A	1/42					
UCS6454-A	1/43					
UCS6454-A	1/44					
UCS6454-A	1/45					
UCS6454-A	1/46					
UCS6454-A	1/47					
UCS6454-A	1/48					
UCS6454-A	1/49	Customer LAN				uplink
UCS6454-A	1/50	Customer LAN				uplink
UCS6454-A	1/51					
UCS6454-A	1/52					
UCS6454-A	1/53					
UCS6454-A	1/54					

Device	Port	Connected To	Port	Type	Length	Note
UCS6454-B	L1	UCS6454-A	L1	CAT5	1FT	
UCS6454-B	L2	UCS6454-A	L2	CAT5	1FT	
UCS6454-B	mgmt0	Customer LAN				
UCS6454-B	1/1	HX Server #1	mLOM port 3	Twinax	3M	Server 1
UCS6454-B	1/2	HX Server #2	mLOM port 3	Twinax	3M	Server 2
UCS6454-B	1/3	HX Server #3	mLOM port 3	Twinax	3M	Server 3
UCS6454-B	1/4	HX Server #4	mLOM port 3	Twinax	3M	Server 4
UCS6454-B	1/5	HX Server #5	mLOM port 3	Twinax	3M	Server 5
UCS6454-B	1/6	HX Server #6	mLOM port 3	Twinax	3M	Server 6
UCS6454-B	1/7	HX Server #7	mLOM port 3	Twinax	3M	Server 7
UCS6454-B	1/8	HX Server #8	mLOM port 3	Twinax	3M	Server 8
UCS6454-B	1/9	C220 Server #9	mLOM port 3	Twinax	3M	Server 9
UCS6454-B	1/10	C220 Server #10	mLOM port 3	Twinax	3M	Server 10
UCS6454-B	1/11	C220 Server #11	mLOM port 3	Twinax	3M	Server 11
UCS6454-B	1/12	C220 Server #12	mLOM port 3	Twinax	3M	Server 12
UCS6454-B	1/13	C220 Server #13	mLOM port 3	Twinax	3M	Server 13
UCS6454-B	1/14	C220 Server #14	mLOM port 3	Twinax	3M	Server 14
UCS6454-B	1/15	C220 Server #15	mLOM port 3	Twinax	3M	Server 15
UCS6454-B	1/16	C220 Server #16	mLOM port 3	Twinax	3M	Server 16
UCS6454-B	1/17	2408XP #2	IOM2 port 1	Twinax	3M	Chassis 1
UCS6454-B	1/18	2408XP #2	IOM2 port 2	Twinax	3M	Chassis 1
UCS6454-B	1/19					

Device	Port	Connected To	Port	Type	Length	Note
UCS6454-B	1/20					
UCS6454-B	1/21					
UCS6454-B	1/22					
UCS6454-B	1/23					
UCS6454-B	1/24					
UCS6454-B	1/25					
UCS6454-B	1/26					
UCS6454-B	1/27					
UCS6454-B	1/28					
UCS6454-B	1/29					
UCS6454-B	1/30					
UCS6454-B	1/31					
UCS6454-B	1/32					
UCS6454-B	1/33					
UCS6454-B	1/34					
UCS6454-B	1/35					
UCS6454-B	1/36					
UCS6454-B	1/37					
UCS6454-B	1/38					
UCS6454-B	1/39					
UCS6454-B	1/40					
UCS6454-B	1/41					

Device	Port	Connected To	Port	Type	Length	Note
UCS6454-B	1/42					
UCS6454-B	1/43					
UCS6454-B	1/44					
UCS6454-B	1/45					
UCS6454-B	1/46					
UCS6454-B	1/47					
UCS6454-B	1/48					
UCS6454-B	1/49	Customer LAN				uplink
UCS6454-B	1/50	Customer LAN				uplink
UCS6454-B	1/51					
UCS6454-B	1/52					
UCS6454-B	1/53					
UCS6454-B	1/54					

Cisco UCS Installation

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

Cisco UCS Fabric Interconnect A

To configure Fabric Interconnect A, follow these steps:

1. Make sure 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, and 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:  HX1-FI

Physical Switch Mgmt0 IP address : 10.29.132.104

Physical Switch Mgmt0 IPv4 netmask : 255.255.255.0

IPv4 address of the default gateway : 10.29.132.1

Cluster IPv4 address : 10.29.132.106

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

  DNS IP address : 10.29.132.110

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

  Default domain name : hxdom.local

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

Following configurations will be applied:

  Switch Fabric=A
  System Name=HX1-FI
  Enforced Strong Password=no
  Physical Switch Mgmt0 IP Address=10.29.132.104
  Physical Switch Mgmt0 IP Netmask=255.255.255.0
  Default Gateway=10.29.132.1
  Ipv6 value=0
  DNS Server=10.29.132.110
  Domain Name=hxdom.local

  Cluster Enabled=yes
  Cluster IP Address=10.29.132.106
  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 Fabric Interconnect B

To configure Fabric Interconnect B, follow these 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, and 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: 10.29.132.104
    Peer Fabric interconnect Mgmt0 IPv4 Netmask: 255.255.255.0
    Cluster IPv4 address          : 10.29.132.106

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

  Physical Switch Mgmt0 IP address : 10.29.132.105


  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

To log into the Cisco UCS Manager environment, follow these steps:

1. Open a web browser and navigate to the Cisco UCS Manager Cluster IP address.
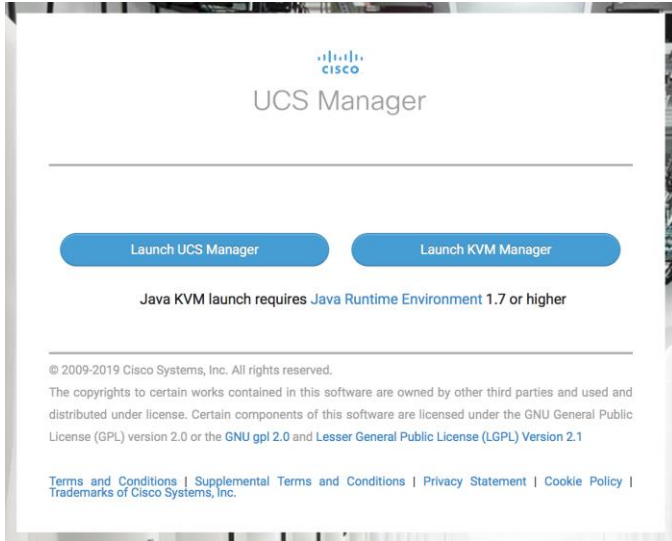
2. Click the "Launch UCS Manager" HTML link to open the Cisco UCS Manager web client.

3. At the login prompt, enter "admin" as the username, and enter the administrative password that was set during the initial console configuration.

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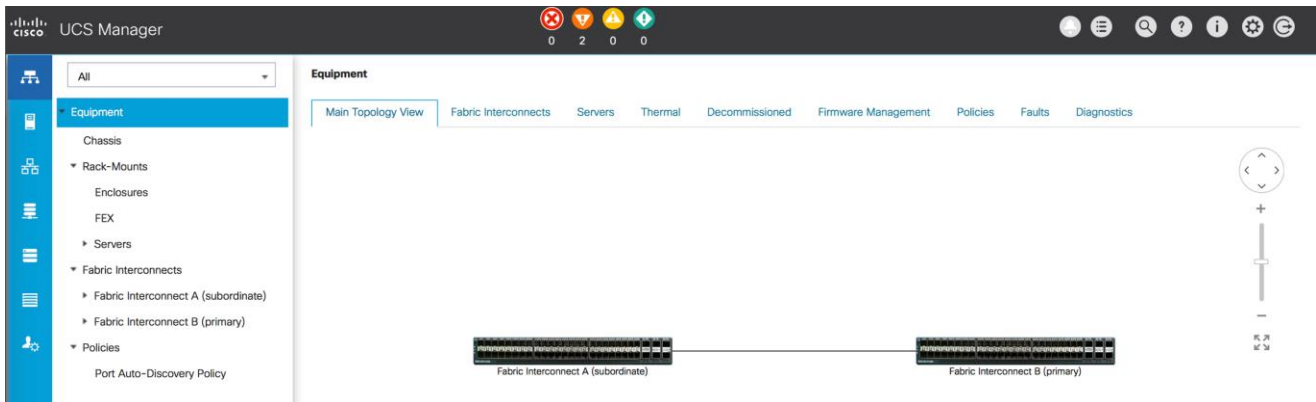
Cisco UCS Configuration

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

Cisco UCS Firmware

Your Cisco UCS 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 4.0(4d). If the firmware version of the Fabric Interconnects is older

than this version, the firmware 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 these instructions:

https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Firmware-Mgmt/4-0/b_UCSM_GUI_Firmware_Management_Guide_4-0.html

NTP

To synchronize the Cisco UCS environment time to the NTP server, follow these steps:

1. In Cisco UCS Manager, click Admin.

2. In the navigation pane, choose All > Time Zone Management, and click the carat next to Time Zone Management to expand it.

3. Click Timezone.

4. In the Properties pane, choose the appropriate time zone in the Time Zone menu.

5. Click Add NTP Server.

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

7. Click OK.

8. Click Save Changes and then click OK.

Uplink Ports

The Ethernet ports of a Cisco UCS 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, follow these steps:

1. In Cisco UCS Manager, click Equipment.

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

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

4. Click Yes to confirm the configuration then click OK.

5. Choose Fabric Interconnects > Fabric Interconnect B > Fixed Module or Expansion Module as appropriate > Ethernet Ports.

6. Choose 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, where their role is listed as "Network".

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, which will use the previously configured uplink ports. To configure the necessary port channels in the Cisco UCS environment, follow these steps:

1. In Cisco UCS Manager, click LAN.

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

3. Right-click Port Channels underneath Fabric A, then click Create Port Channel.

4. Enter the port channel ID number as the unique ID of the port channel (this does not have to match the port-channel ID on the upstream switch).

5. Enter the name of the port channel.

6. Click Next.

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

8. Click Finish.

9. Click OK.

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

11. Right-click Port Channels underneath Fabric B, then click Create Port Channel.

12. Enter the port channel ID number as the unique ID of the port channel (this does not have to match the port-channel ID on the upstream switch).

13. Enter the name of the port channel.

14. Click Next.

15. Click each port from Fabric Interconnect B that will participate in the port channel, then 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 blades as compute-only nodes in an extended cluster design, additional settings must be configured for connecting the Cisco UCS 5108 blade chassis. The Chassis Discovery policy defines the number of links between the Fabric Interconnect and the Cisco UCS 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 Cisco UCS Manager will automatically combine them into port-channels. Cisco best practices recommends using link grouping, and the number of links per side is dependent on the hardware used

in Cisco UCS 5108 chassis, and the model of Fabric Interconnects. For 10 GbE connections Cisco recommends 4 links per side, and for 40 GbE connections Cisco recommends 2 links per side.

To configure the necessary policy and setting, follow these steps:

1. In Cisco UCS Manager, click Equipment, then click Equipment.

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 uplink ports that are cabled per side, between the chassis and the Fabric Interconnects.

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

5. Set the backplane speed preference to 4x10 Gigabit or 40 Gigabit.

6. Click Save Changes.

7. Click OK.

Main Topology View	Fabric Interconnects	Servers	Thermal	Decommissioned	Firmware Management	Policies

Global Policies	Autoconfig Policies	Server Inheritance Policies	Server Discovery Policies	SEL Policy	Power Groups

Chassis/FEX Discovery Policy

Action : 1 Link ▼

Link Grouping Preference : ◉ None ○ Port Channel

Backplane Speed Preference : ◉ 40G ○ 4x10G

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. When 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 Cisco UCS Manager 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 Cisco 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 Cisco UCS Manager before configuring the ports for the next server. The same numbering procedure applies to blade server chassis, although chassis and rack-mount server numbers are separate from each other.

Auto Configuration

A new feature in Cisco UCS Manager 3.1(3a) and later is Server Port Auto-Discovery, which automates the configuration of ports on the Fabric Interconnects as server ports when a Cisco UCS rack-mount server or blade chassis is connected to them. The firmware on the rack-mount servers or blade chassis Fabric Extenders must already be at version 3.1(3a) or later in order for this feature to function properly. Enabling this policy eliminates the manual steps of configuring each server port, however it can configure the servers in a somewhat random order depending upon the circumstances. An example of how to use this feature in an orderly manner would be

to have the policy already set, then to mount, cable and apply power to each new server one-by-one. In this scenario the servers should be automatically discovered in the order you racked them and applied power.

An example of how the policy can result in unexpected ordering would be when the policy has not been enabled, then all of the new servers are racked, cabled, and have power applied to them. If the policy is enabled afterwards, it will likely not discover the servers in a logical order. For example, the rack-mount server at the bottom of the stack, which you may refer to as server #1, and you may have plugged into port 1 of both Fabric Interconnects, could be discovered as server 2, or server 5, and so on. In order to have fine control of the rack-mount server or chassis numbering and order in this scenario, the manual configuration steps listed in the next section must be followed.

To configure automatic server port definition and discovery, follow these steps:

1. In Cisco UCS Manager, click the Equipment button on the left-hand side.

2. In the navigation tree, under Policies, click Port Auto-Discovery Policy.

3. In the properties pane, set Auto Configure Server Port option to Enabled.

4. Click Save Changes.

5. Click OK.

6. Wait for a brief period, until the rack-mount servers appear in the Equipment tab underneath Equipment > Rack Mounts > Servers, or the chassis appears underneath Equipment > Chassis.
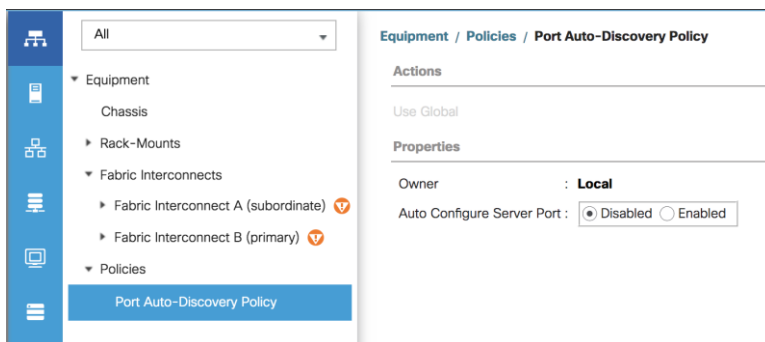
Manual Configuration

To manually define the specified ports to be used as server ports, and have control over the numbering of the servers, follow these steps:

1. In Cisco UCS Manager, click Equipment.

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

3. Choose 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. Choose Fabric Interconnects > Fabric Interconnect B > Fixed Module or Expansion Module as appropriate > Ethernet Ports.

6. Choose 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 pair of server ports, until all rack-mount servers and chassis appear in the order desired in the Equipment tab.

Server Discovery

As previously described, when 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, follow these steps:

1. In Cisco UCS Manager, click Equipment and then click Equipment at the top of the navigation tree.

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

3. Click the Blade Servers or Rack-Mount Servers sub-tab as appropriate, then view the servers' status in the Overall Status column.

HyperFlex Installer VM 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/download/home/286305544/type/286305994/release/4.0(1b)

This document is based on the Cisco HyperFlex 4.0(1b) release filename: Cisco-HX-Data-Platform-Installer-v4.0.2a-33133-esx.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 importing of OVA format files. For the purpose of this document, the process described uses an existing ESXi server managed by vCenter to run the HyperFlex installer OVA and deploying it via the VMware vSphere Web 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 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 via SSH.

For detailed information about all ports required for the installation of Cisco HyperFlex, refer to Appendix A of the HyperFlex 4.0 Hardening Guide: https://www.cisco.com/c/dam/en/us/support/docs/hyperconverged-infrastructure/hyperflex-hx-data-platform/HX-Hardening_Guide.pdf

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 Table 48 to document the settings to be used for the HyperFlex installer VM:

Table 48   HyperFlex Installer Settings

Setting	Value
IP Address	
Subnet Mask	
Default Gateway	
DNS Server	
NTP Server(s)	
Root Password	

Deploy Installer OVA

To deploy the HyperFlex installer OVA, follow these steps:

1. Open the vSphere HTML5 Web Client webpage of a vCenter server where the installer OVA will be deployed and log in with admin privileges.

2. In the vSphere Web Client, from the Home view, click Hosts and Clusters.

3. From the Actions menu, click Deploy OVF Template.

4. Choose the Local file option, then click the Choose Files button and locate the *Cisco-HX-Data-Platform-Installer-v4.0.2a-33133-esx.ova* file, click the file and click Open.

5. Click Next.

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

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

8. After the file validation, review the details and click Next.

9. Choose a Thin provision virtual disk format, and the datastore to store the new virtual machine, then click Next.

10. 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.

11. If DHCP is to be used for the installer VM, leave the fields blank, except for the NTP server value and click Next. If static address settings are to be used, fill in the fields for the DNS server, Default Gateway, NTP Servers, IP address, and subnet mask.

12. Enter and confirm a new password used to log in to the installer VM after it is deployed, then click Next.

13. Review the final configuration and click Finish.

14. The installer VM will take a few minutes to deploy, once it has deployed, power on the new VM and proceed to the next step.

HyperFlex Installer Web Page

The HyperFlex installer is accessed via 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:

Figure 47 HyperFlex Installer VM IP Address

To access the HyperFlex installer webpage, follow these steps:

1. Open a web browser on the local computer and navigate to the IP address of the installer VM. For example, open 10.xx.xx.xx

2. Click accept or continue to bypass any SSL certificate errors.

3. At the login screen, enter the username: root

4. At the login screen, enter the password which was set during the OVA deployment.

5. Verify the version of the installer in the lower right-hand corner of the Welcome page is the correct version.

6. Check the box for "I accept the terms and conditions" and click Login.
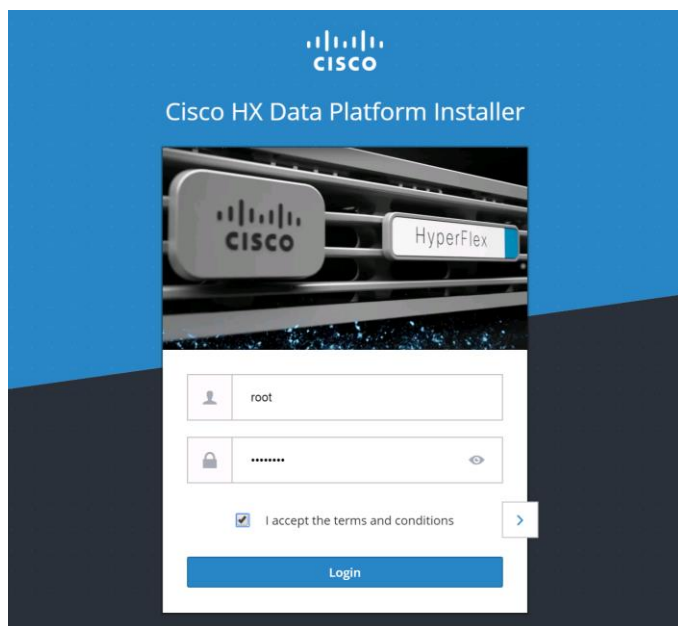
Cisco HyperFlex Cluster Configuration

To configuring the Cisco HyperFlex Cluster, follow this step:

1. Log into the HX Installer virtual machine through a web browser
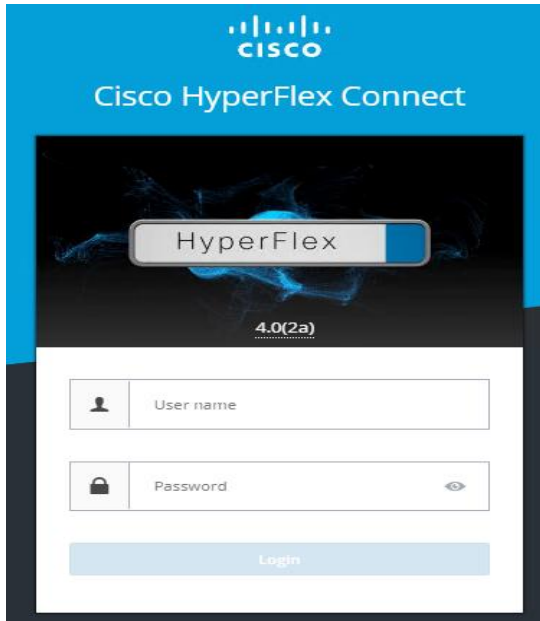
Create a HyperFlex Cluster

To create a HyperFlex cluster, follow these steps:

1. Choose the workflow for cluster creation to deploy a new HyperFlex cluster on eight Cisco HXAF220c-M5S nodes.
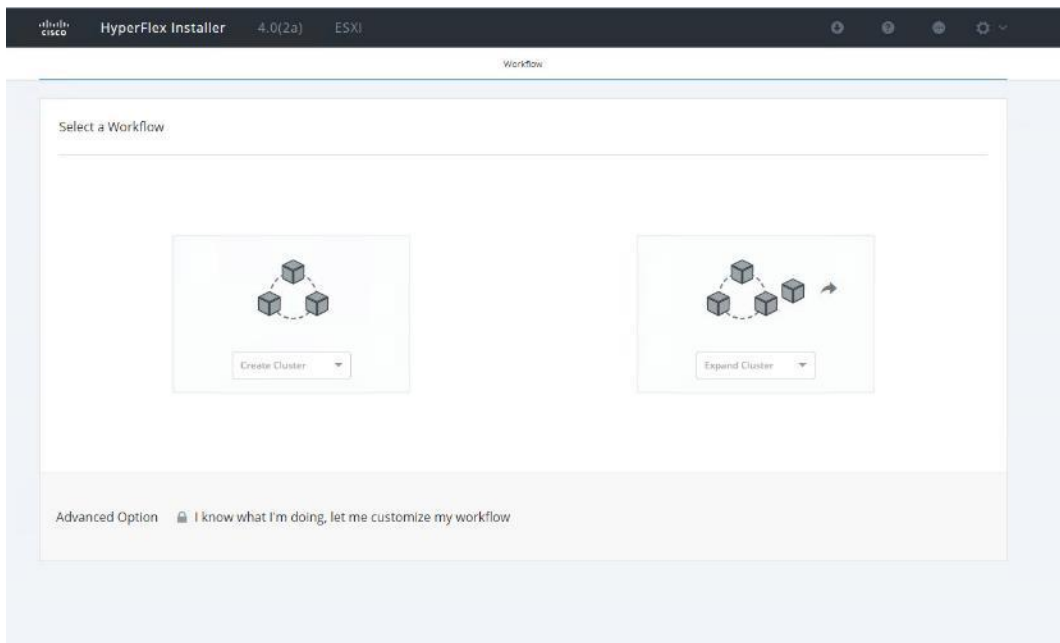
2. On the credentials page, enter the access details for Cisco UCS Manager, vCenter server, and Hypervisor. Click Continue.

3. Choose the top-most check box at the top right corner of the HyperFlex installer to select all unassociated servers. (To configure a subset of available of the HyperFlex servers, manually click the check box for individual servers.)

4. Click Continue after completing server selection.

| ⚠ | The required server ports can be configured from Installer workflow but it will extend the time to complete server discovery. Therefore, we recommend configuring the server ports and complete HX node discovery in Cisco UCS Manager as described in the Pre-requisites section above prior starting workflow for HyperFlex installer. |

Configure Server Ports (Optional)

If you choose to allow the installer to configure the server ports, follow these steps:

1. Click Configure Server Ports at the top right corner of the Server Selection window.

2. Provide the port numbers for each Fabric Interconnect in the form:

   A1/x-y,B1/x-y   where A1 and B1 designate Fabric Interconnect A and B and where x=starting port number and y=ending port number on each Fabric Interconnect.

3. Click Configure.
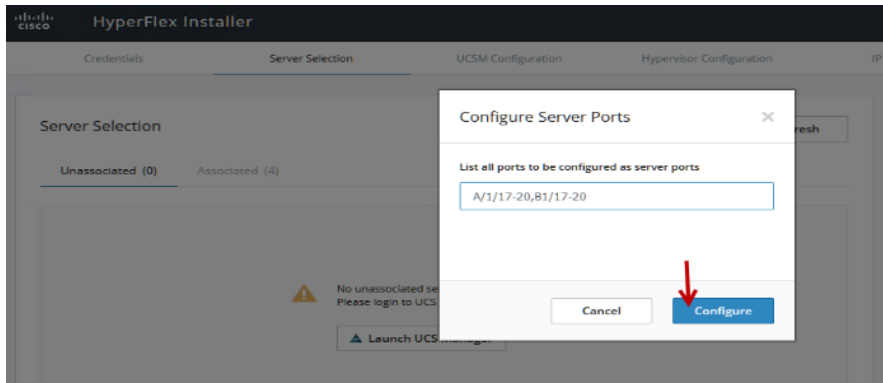
4. Enter the Details for the Cisco UCS Manager Configuration:

   a. Enter the VLAN ID for hx-inband-mgmt, hx-storage-data, hx-vmotion, vm-network.

   b. MAC Pool Prefix: The prefix to use for each HX MAC address pool. Please select a prefix that does not conflict with any other MAC address pool across all Cisco UCS domains.

   c. The blocks in the MAC address pool will have the following format:

   ◦ ${prefix}:${fabric_id}${vnic_id}:{service_profile_id}

   ◦ The first three bytes should always be "00:25:B5".

   ⚠ The first three bytes should always be "00:25:B5."

5. Enter range of IP address to create a block of IP addresses for external management and access to CIMC/KVM.

6. Cisco UCS firmware version is set to 4.0 (1b) which is the required Cisco UCS Manager release for HyperFlex v4.0(1a) installation.

7. Enter HyperFlex cluster name.

8. Enter Org name to be created in Cisco UCS Manager.

9. Click Continue.

Configure Hypervisor Settings

To configure the Hypervisor settings, follow these steps:

1. In the Configure common Hypervisor Settings section, enter:

    a. Subnet Mask

    b. Gateway

    c. DNS server(s)

2. In the Hypervisor Settings section:

    a. Choose check box Make IP Address and Hostnames Sequential if they are following in sequence.

    b. Provide the starting IP Address.

    c. Provide the starting Host Name or enter Static IP address and Host Names manually for each node

3. Click Continue.
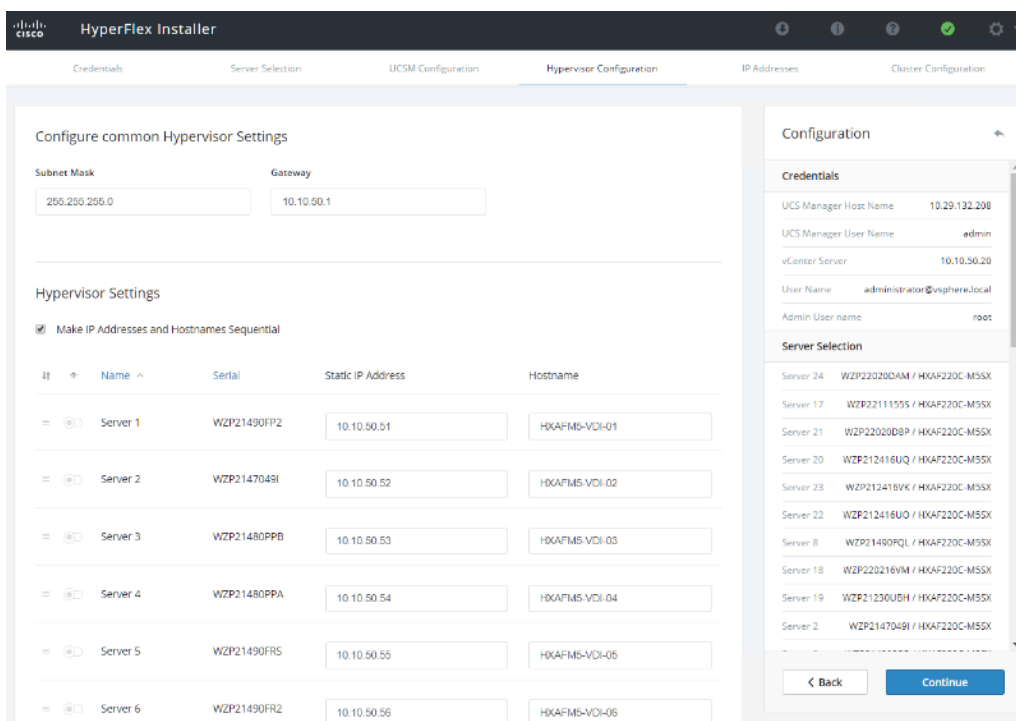
IP Addresses

To add the IP addresses, follow these steps:

> ⚠ When the IP Addresses page appears, the hypervisor IP address for each node that was configured in the Hypervisor Configuration tab, appears under the Management Hypervisor column.

Three additional columns appear on this page:

- Storage Controller/Management

- Hypervisor/Data

- Storage Controller/Data

> ⚠️ The Data network IP addresses are for vmkernel addresses for storage access by the hypervisor and storage controller virtual machine.

1. On the IP Addresses page, check the box Make IP Addresses Sequential or enter the IP address manually for each node for the following requested values:

   a. Storage Controller/Management

   b. Hypervisor/Data

   c. Storage Controller/Data

2. Enter subnet and gateway details for the Management and Data subnets configured.

3. Click Continue to proceed.

4. On the Cluster Configuration page, enter the following:

   a. Cluster Name

   b. Cluster management IP address

   c. Cluster data IP Address

   d. Set Replication Factor: 2 or 3

   e. Controller virtual machine password

   f. vCenter configuration

      ▪ vCenter Datacenter name

- vCenter Cluster name
g. System Services
    - DNS Server(s)
    - NTP Server(s)
    - Time Zone
h. Auto Support
    - Click the check box for Enable Auto Support
    - Mail Server
    - Mail Sender
    - ASUP Recipient(s)
i. Advanced Networking
    - Management vSwitch
    - Data vSwitch
j. Advanced Configuration
    - Click the check box to Optimize for VDI only deployment
    - Enable jumbo Frames on Data Network
    - Clean up disk partitions (optional)
        - vCenter Single-Sign-On server

○ vCenter Single-Sign-On server

5. The configuration details can be exported to a JSON file by clicking the down arrow icon in the top right corner of the Web browser page as shown in the screenshot below.

6. Configuration details can be reviewed on Configuration page on right side section. Verify entered details for IP address entered in Credentials page, server selection for cluster deployment and creation workflow, Cisco UCS Manager configuration, Hypervisor Configuration, IP addresses.

7. Click Start after verifying details.

When the installation workflow begins, it will go through the Cisco UCS Manager validation.

Configuration	
IP Blocks	10.29.132.41-77
Subnet Mask	255.255.255.0
Gateway	10.29.132.1
UCS Server Firmware Version	3.2(1d)
HyperFlex Cluster Name	HXAF-M5-HZVDI
Org Name	HXAF-M5-HZVDI
iSCSI Storage	false
VLAN A Name	hx-ext-storage-iscsi-a
VLAN B Name	hx-ext-storage-iscsi-b
FC Storage	false
WWxN Pool	20:00:00:25:B5:
VSAN A Name	hx-ext-storage-fc-a
VSAN B Name	hx-ext-storage-fc-b
Hypervisor Configuration	
Subnet Mask	255.255.255.0
Gateway	10.10.50.1
DNS Server(s)	10.10.51.21,10.10.51.22
Server 1	
Static IP Address	10.10.50.51
Hostname	HXAFM5-HZVDI-01
Server 2	
Static IP Address	10.10.50.52
Hostname	HXAFM5-HZVDI-02

> ⚠ If QoS system class is not defined as per the requirement HyperFlex installer will go ahead and make required changes. There will be a warning generated accordingly in HyperFlex Installer workflow. For 6300 series Fabric Interconnect change in QoS system class requires reboot of FIs.

8. After a successful validation, the workflow continues with the Cisco UCS Manager configuration.

Progress

Configuration

Start   Validations   UCSM Configuration   Hypervisor Configuration   Deploy Validation   Deploy   Create Validation   Cluster Creation

UCSM Configuration in Progress

UCSM Configuration

UCSM Configuration - Overall

In Progress

✓	Login to UCS API
✓	Inventory physical servers
✓	Validate UCS firmware version
✓	Setting flags for firmware validation
✓	Get inventory of firmware bundles
✓	Download firmware bundle
✓	Configure UCS Fabric Interconnect
✓	Configure FI Server Ports
✓	Configure QoS classes
✓	Configure org for the hx cluster
✓	Configure VLANs
✓	Configure Host Firmware policy
✓	Configure MAC address pools
✓	Configure QoS policies
✓	Configure Network Control policies
·	Configure HyperFlex cluster
⟳	Configure Adapter policies

Configuration

Credentials

UCS Manager Host Name	10.29.132.40
User Name	admin
vCenter Server	10.10.50.20
User Name	administrator@vsphere.local
Admin User Name	root

Server Selection

Server 2	WZP212416UO / HXAF220C-M5SX
Server 3	WZP212416VK / HXAF220C-M5SX
Server 1	WZP21230UBH / HXAF220C-M5SX
Server 4	WZP212416UQ / HXAF220C-M5SX

UCSM Configuration

VLAN Name	hx-inband-mgmt
VLAN ID	50
VLAN Name	hx-storage-data
VLAN ID	52
VLAN Name	hx-vmotion
VLAN ID	53
VLAN Name	vm-network
VLAN ID(s)	54
MAC Pool Prefix	00:25:B5:23
IP Blocks	10.29.132.41-77
Subnet Mask	255.255.255.0
Gateway	10.29.132.1
UCS Server Firmware Version	3.2(2b)
HyperFlex Cluster Name	HXAF-M5-HZVDI

9. After a successful Cisco UCS Manager configuration, the installer proceeds with the Hypervisor configuration.

10. After a successful Hypervisor configuration, the deploy validation task is performed which checks for the required component and accessibility prior Deploy task is performed on Storage Controller virtual machine.

11. Installer performs deployment task after successfully validating Hypervisor configuration.

12. After a successful deployment of the ESXi hosts configuration, the Controller virtual machine software components for HyperFlex installer checks for validation prior to creating the cluster.

13. After a successful validation, the installer creates and starts the HyperFlex cluster service.

14. After a successful HyperFlex Installer virtual machine workflow completion, the installer GUI provides a summary of the cluster that has been created.

Cisco HyperFlex Cluster Expansion

For this exercise, you will add the compute node workflow as part of the cluster expansion.

Prerequisites

You need to configure the service profile for compute-only nodes and install ESXi hypervisor.

To add the compute node workflow, follow these steps:

1. Log into Cisco UCS Manger.

2. Under "hx-cluster" sub-organization:

   a. In the existing vMedia policy "HyperFlex" add vMedia mount details to boot ESXi image from data platform installer virtual machine.

   b. For Hostname/IP Address – Add IP address of data-platform installer virtual machine which can also communicate with Cisco UCS Manager.

Create vMedia Mount

Name	:	COMPNODE-VMedia
Description	:	COMPNODE-M5-VMedia
Device Type	:	⦿ CDD ◯ HDD
Protocol	:	◯ NFS ◯ CIFS ⦿ HTTP ◯ HTTPS
Hostname/IP Address	:	10.29.132.18
Image Name Variable	:	⦿ None ◯ Service Profile Name
Remote File	:	e-ESXi-6.5U2-8935087-Cisco-Custom-6.5.2.2.iso
Remote Path	:	images
Username	:	root
Password	:	••••••••
Remap on Eject	:	☐

OK   Cancel

3. Change the existing service profile template to accommodate the new changes; install ESXi through vMedia policy.

4. In the existing service profile template "compute-nodes" choose vMedia Policy tab.

5. Click Modify vMedia Policy.

6. From the drop-down list of vMedia Policy, choose HyperFlex.

7. In the existing service profile template "compute-nodes" click Boot Order tab.

8. Click Modify Boot Policy.

9. From the drop-down list of Boot Policies, choose HyperFlexInstall.

10. Save changes.
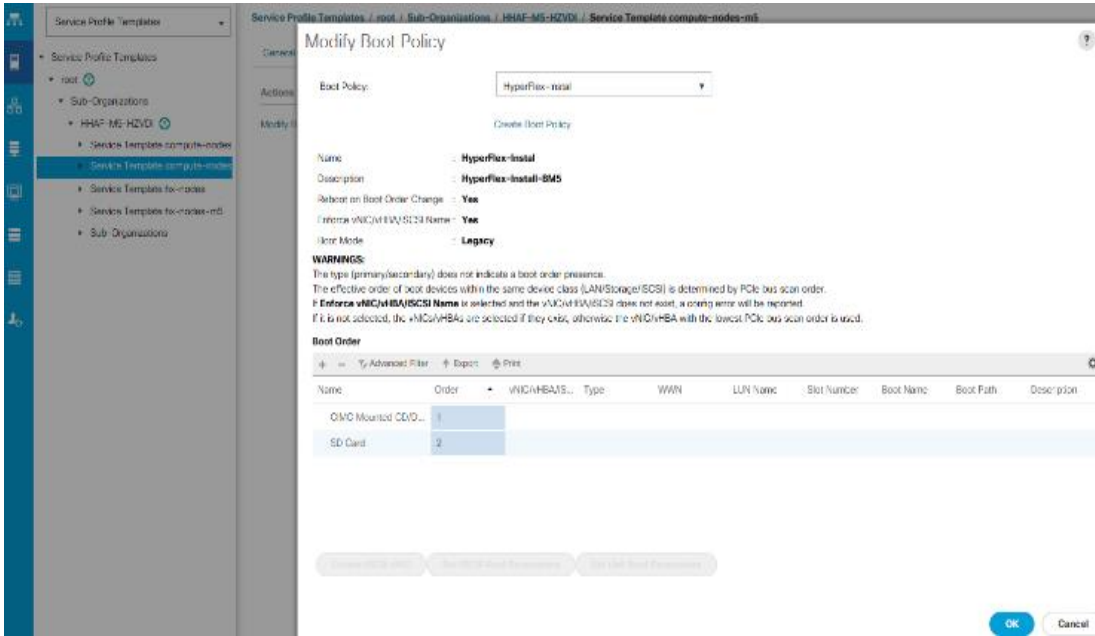
11. Create the service profile from the "compute-nodes" updating service profile template located in the Hyper-Flex cluster sub organization.
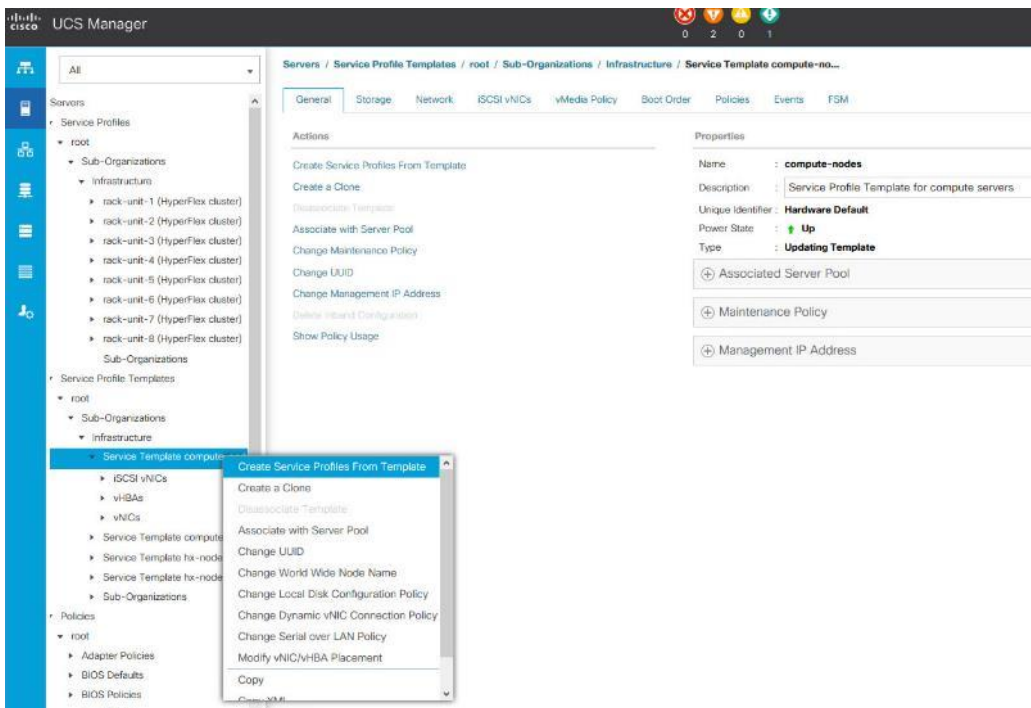
12. Add the Naming Prefix and Number of Instances to be created.
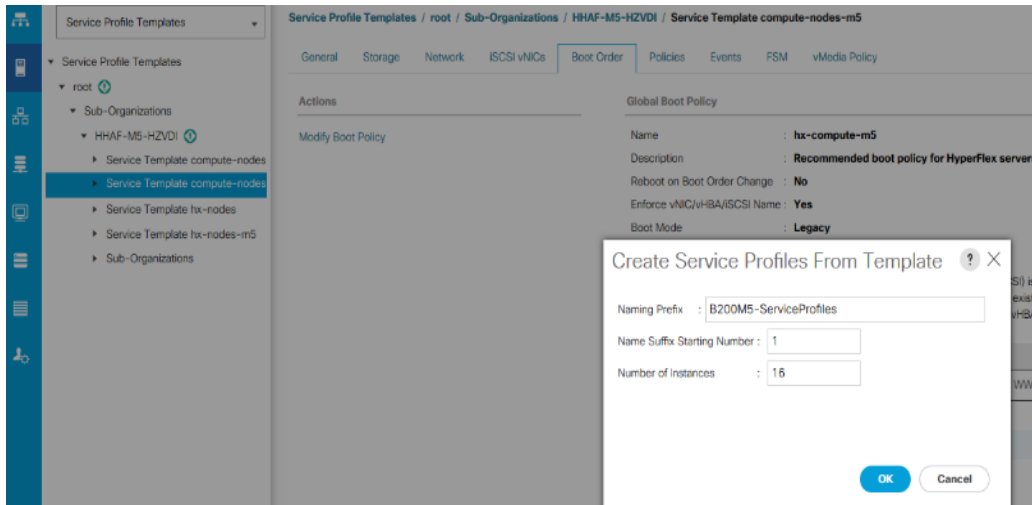
13. Click OK.
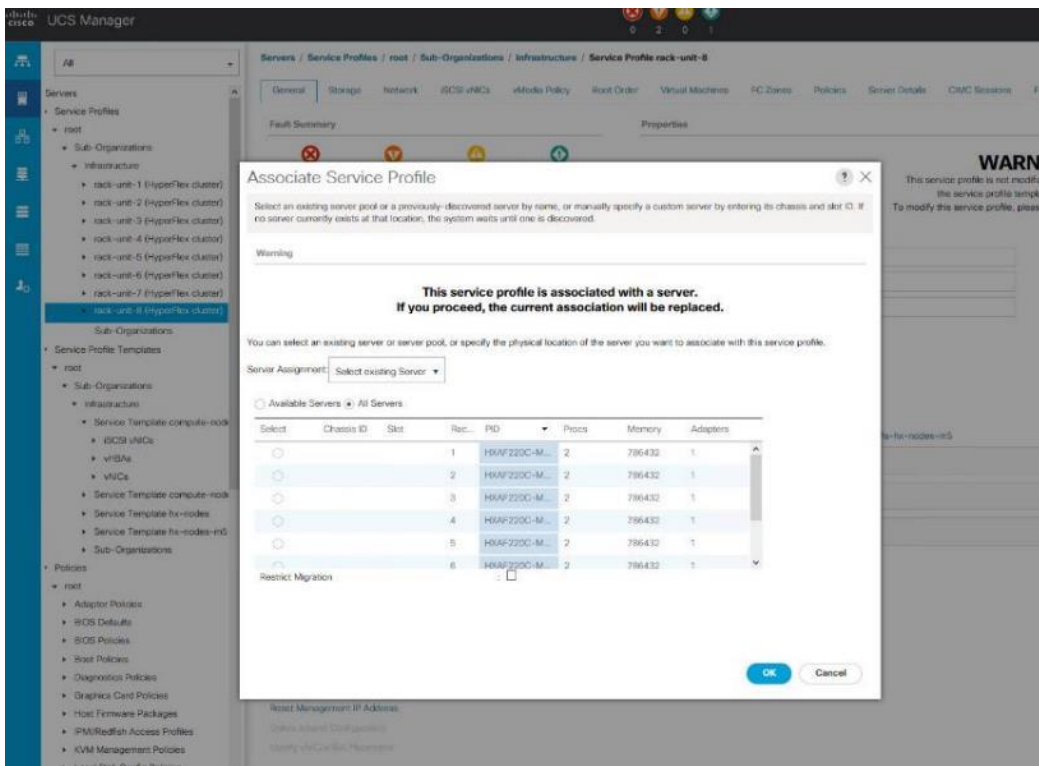
14. After the of ESXi install, assign the VLAN tag on the ESXi host; the static IP address configuration is located in the Configure Management Network section.
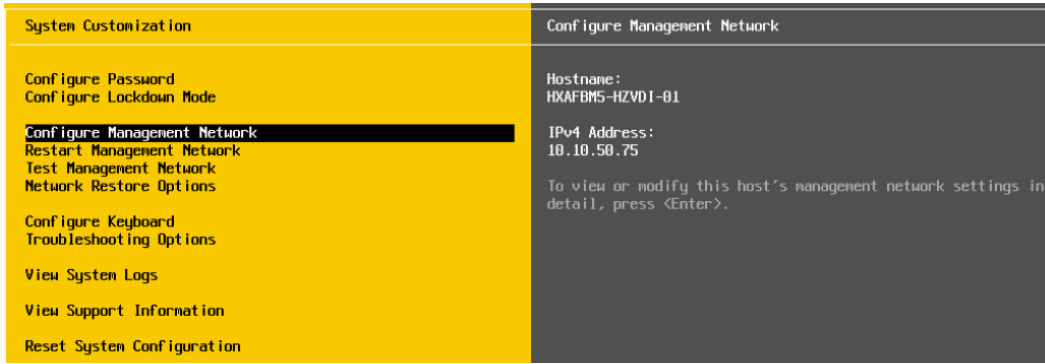
15. Log into the HyperFlex data platform installer WebUI. Click "I know what I'm doing, let me customize my work-flow."
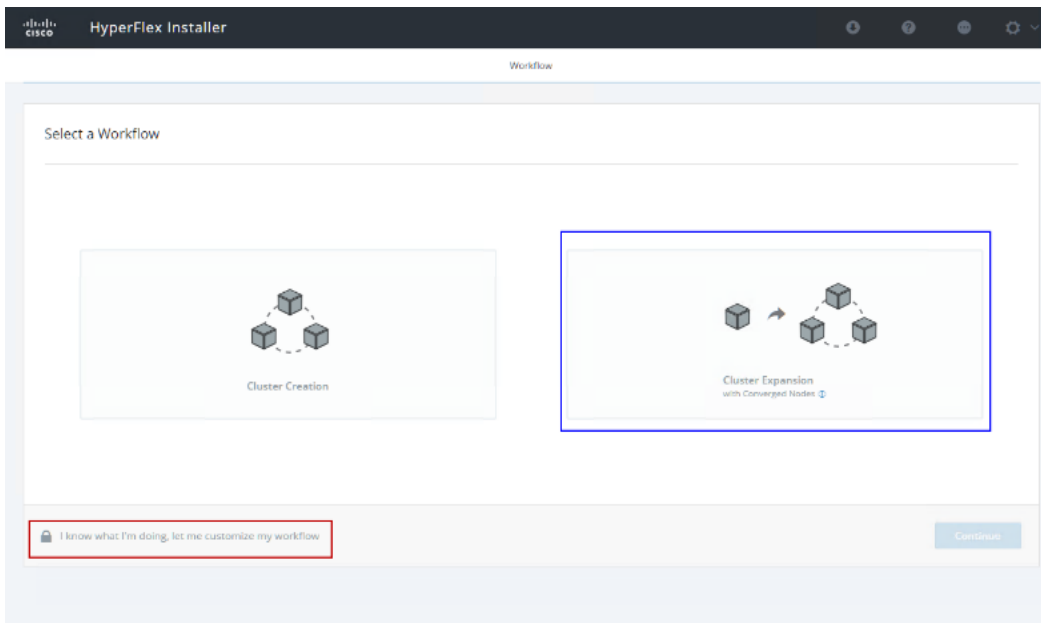
16. Choose Deploy HX Software, Expand HX Cluster. Click Continue.
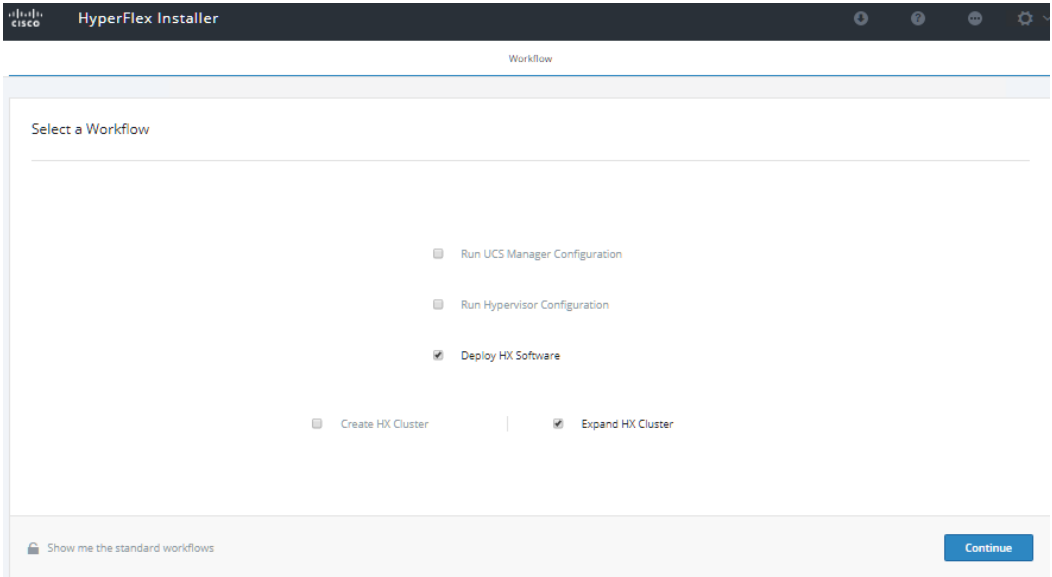
17. Enter the credentials for vCenter server, and ESXi. Click Continue.
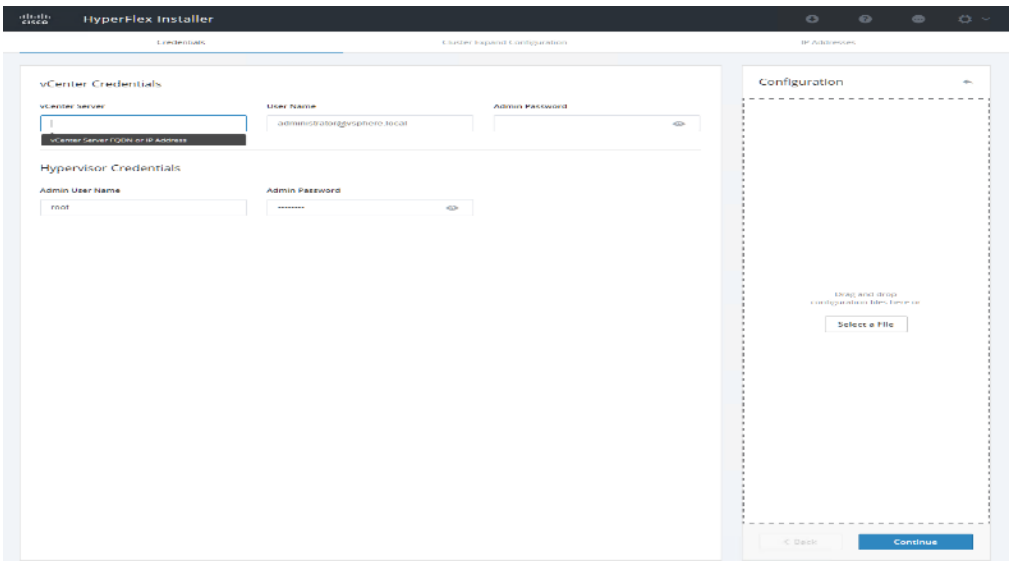
18. Choose Cluster to expand, click Continue.
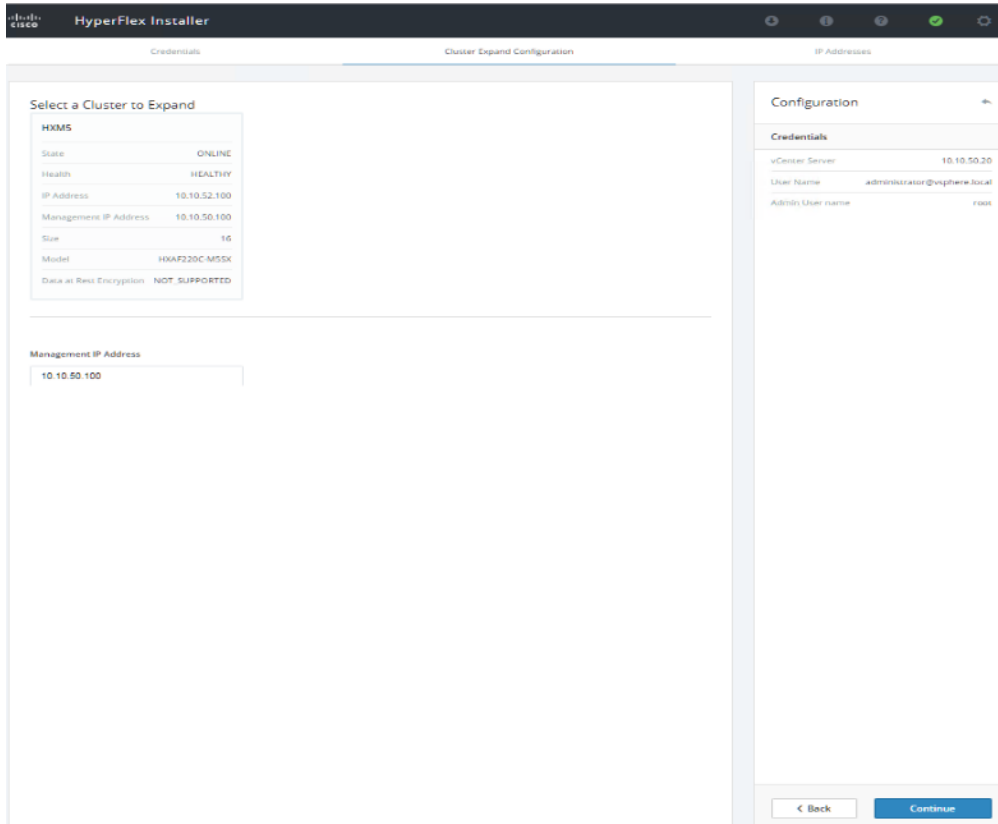
⚠️ Since you are performing a compute-node only expansion, no servers report into the Cisco UCS Manager configuration tab.

---

19. Click Add Compute Server tab for N number of compute-only node expansion to existing HyperFlex cluster. Provide Hypervisor Management IP address and vmkernel IP address to access storage cluster. Click Continue.

IP Addresses

Add Compute Server   Add Converged Server

☑ Make IP Addresses Sequential

		Management - VLAN			Data - VLAN (FQDN or IP Address)			
↕ ◆	Name⌃	Hypervisor	①	Storage Controller ①	Hypervisor	①	Storage Controller ①	
=	Server 1 compute	10.10.50.75			10.10.52.75			✕
=	Server 2 compute	10.10.50.76			10.10.52.76			✕
=	Server 3 compute	10.10.50.77			10.10.52.77			✕
=	Server 4 compute	10.10.50.78			10.10.52.78			✕
=	Server 5 compute	10.10.50.79			10.10.52.79			✕
=	Server 6 compute	10.10.50.80			10.10.52.80			✕
=	Server 7 compute	10.10.50.81			10.10.52.81			✕
=	Server 8 compute	10.10.50.82			10.10.52.82			✕
=	Server 9 compute	10.10.50.83			10.10.52.83			✕
=	Server 10 compute	10.10.50.84			10.10.52.84			✕
=	Server 11 compute	10.10.50.85			10.10.52.85			✕
=	Server 12 compute	10.10.50.86			10.10.52.86			✕
=	Server 13 compute	10.10.50.87			10.10.52.87			✕
=	Server 14 compute	10.10.50.88			10.10.52.88			✕
=	Server 15 compute	10.10.50.89			10.10.52.89			✕
=	Server 16 compute	10.10.50.90			10.10.52.90			✕

Controller VM Password

Configuration ↰

Credentials

vCenter Server	10.10.50.20
User Name	administrator@vsphere.local
Admin User name	root

Cluster Expand Configuration

Management Cluster	10.10.50.100

‹ Back   **Start**

The cluster expansion workflow starts and performs the deploy validation task.

The deployment of HyperFlex controller virtual machine create and deployment begins.

The Expansion Validation begins.

The Summary displays.

As part of the cluster creation operations, the HyperFlex Installer adds HyperFlex functionality to the vSphere vCenter identified in earlier steps. This functionality allows vCenter administrators to manage the HyperFlex cluster entirely from their vSphere Web Client.

20. Click Launch vSphere Web Client.

Cisco HyperFlex installer creates and configures a controller virtual machine on each converged or compute-only node. The naming **convention used is "stctlvm-<Serial Number for Cisco UCS Node>"** shown in .

Do **not** change the name or any resource configuration for the controller virtual machine.

Figure 48 Cisco UCS Node Naming Convention
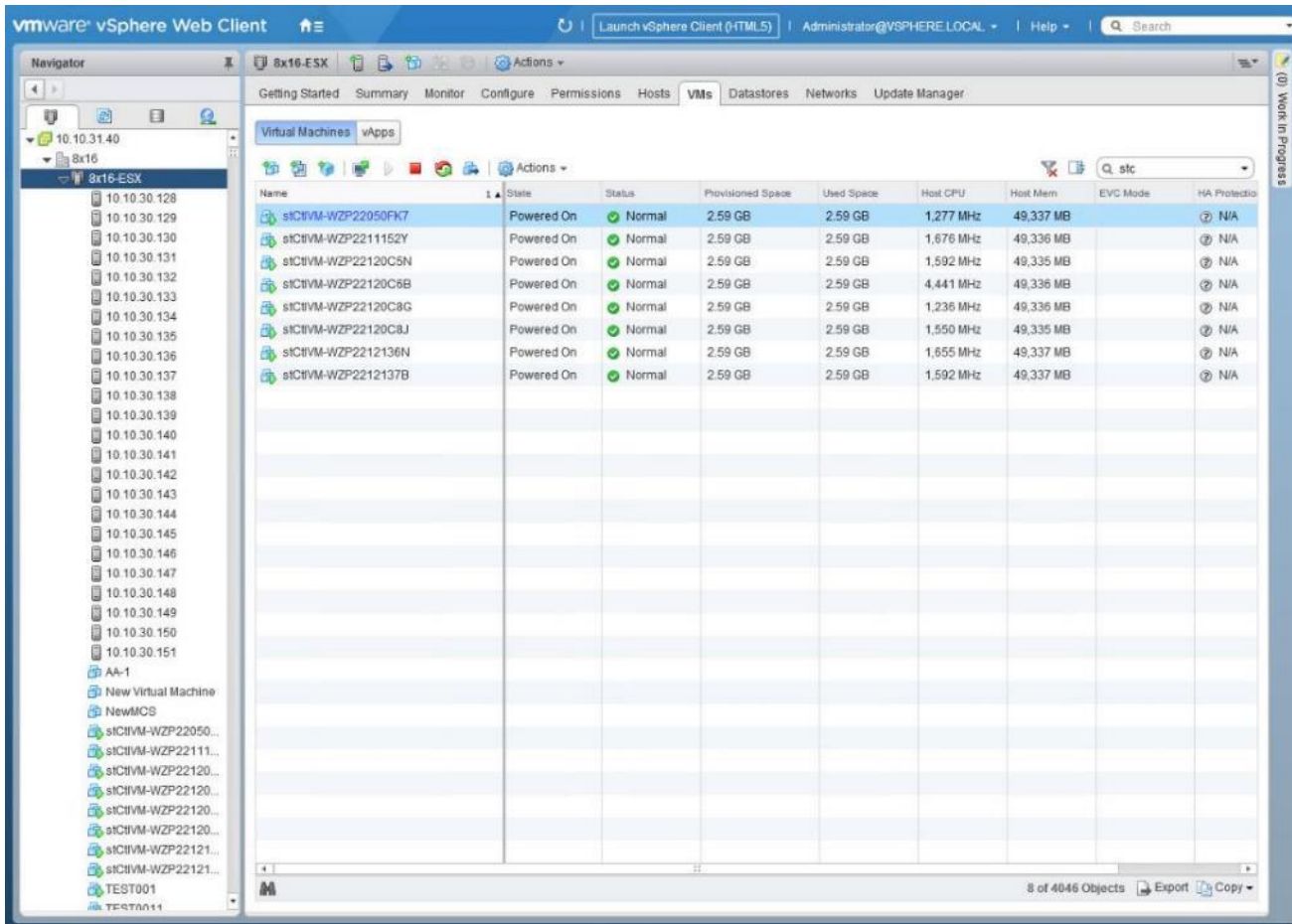
Claim Devices in Intersight (optional)

The Cisco UCS Manager device connector allows Cisco Intersight to manage the Cisco UCS domain and all of the connected HyperFlex servers and claim them for cloud management.

To configure the claim devices, follow these steps:

1. Log into the Cisco UCS Manager web interface of the Cisco Fabric Interconnects which are connected to the Cisco HX-series servers that will comprise the new Cisco HyperFlex cluster being installed.

2. From the left-hand navigation pane click Admin, then click Device Connector.

3. Note that the Cisco UCS domain shows a status of "Not Claimed". Copy the Device ID and the Claim Code by clicking on the small clipboard icons.
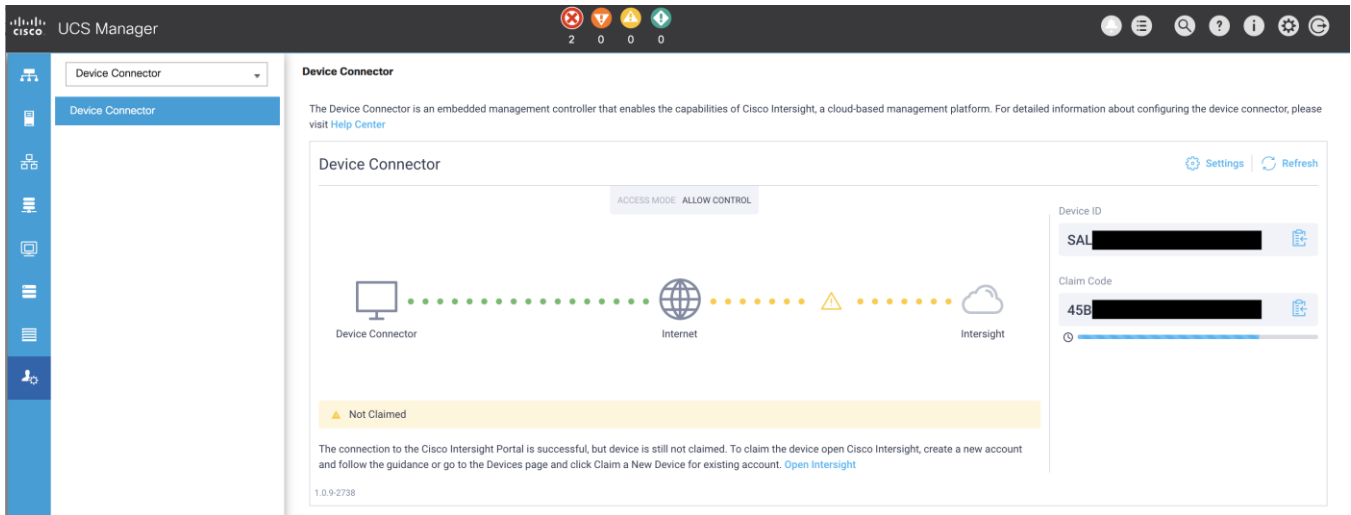
4. Open a web browser and navigate to the Cisco Intersight Cloud Management platform https://intersight.com/.

5. Login with your Cisco ID and password. If this is the first time using Intersight, it is recommended you take a site tour to be guided through some main features.

6. To Claim a new device, from the left-hand Navigation pane, click Devices, in the Device window, choose Claim a New Device at the right top corner.
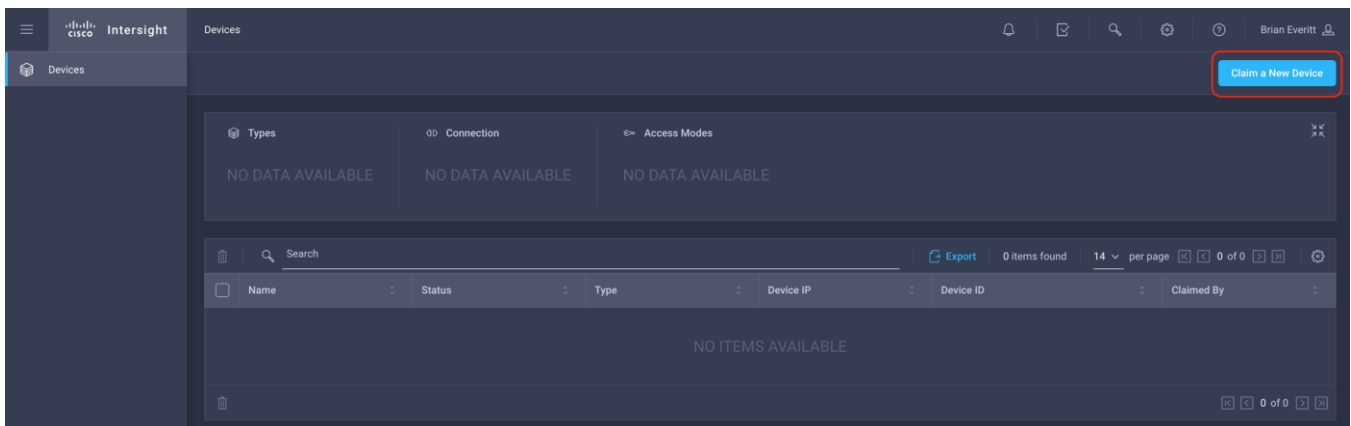
7. Ensure the option for Direct Claim is chosen, then input the Device ID and Claim Code obtained from Cisco UCS management GUI. Use copy and paste for accuracy. Click Claim.
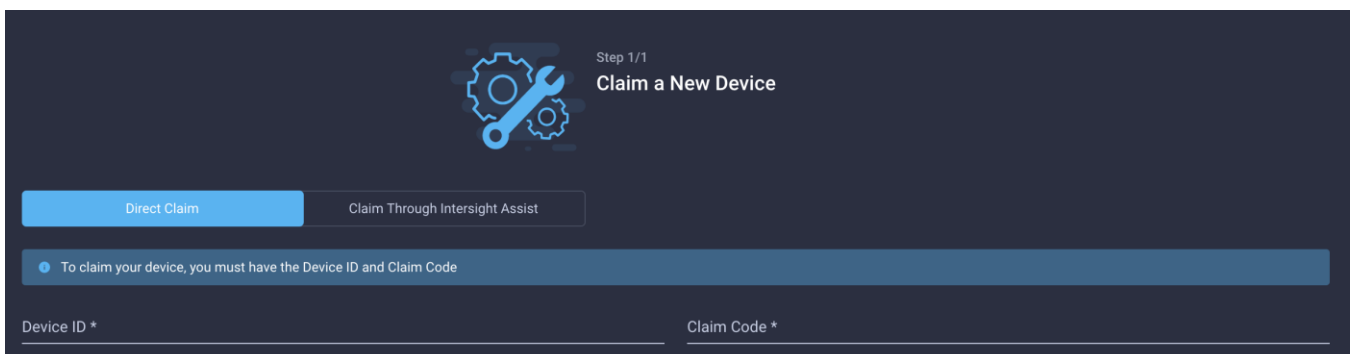
8. Wait until the device is claimed successfully and to appear in the list of devices.

9. Click the Refresh link in the Cisco UCS Manager Device Connector screen. The Device Connector now shows this device is claimed.
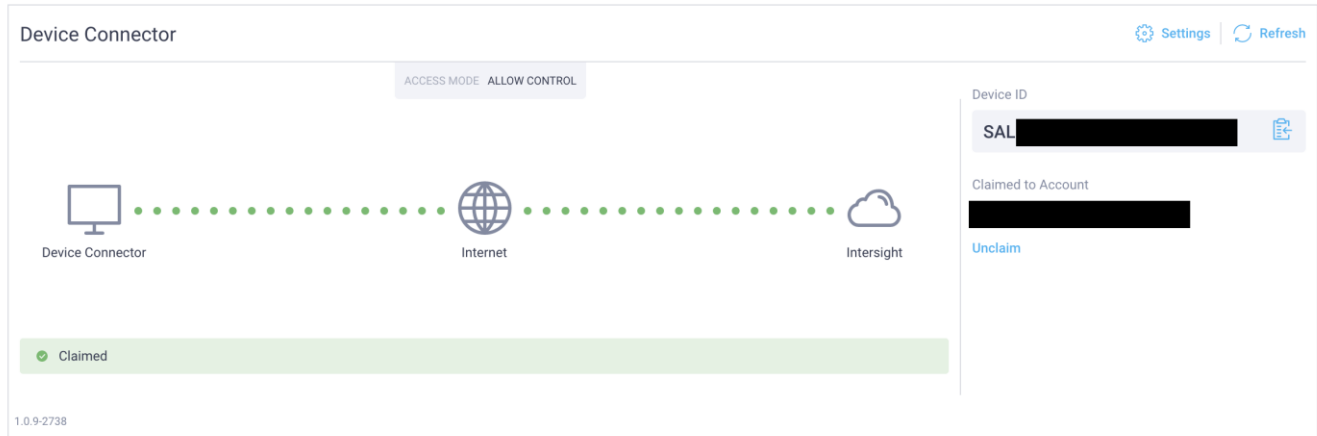
10. In the Device window, the Cisco UCS Fabric Interconnect domain should now show as connected devices.
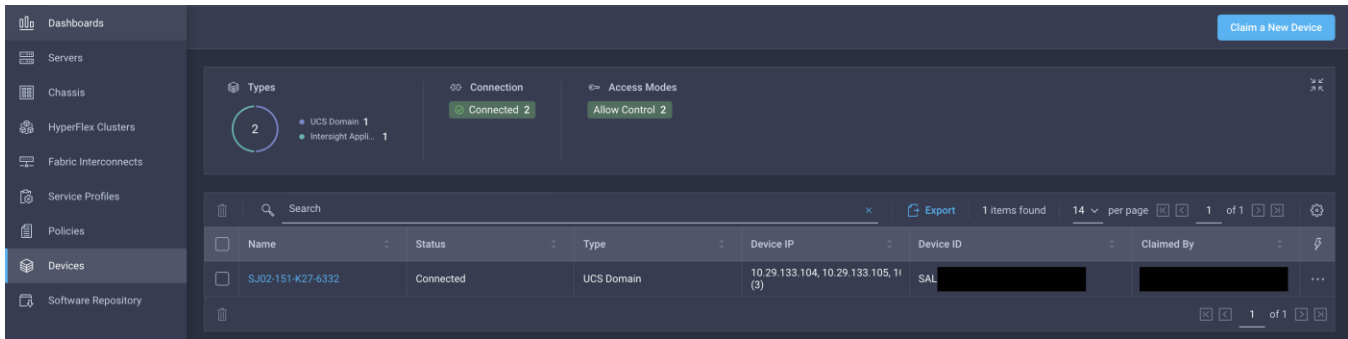
Post-install Configuration

Prior to putting a new HyperFlex cluster into production, a few post-install tasks must be completed. To auto-mate the post installation procedures and verify the HyperFlex cluster configuration, a post_install script has been provided on the HyperFlex Controller VMs. To run this script, follow these steps:

1. SSH to the cluster management IP address and login using <root> username and the controller VM password provided during installation. Verify the cluster is online and healthy using "stcli cluster info" or "stcli cluster storage-summary".

```
root@SpringpathControllerT7DB8MDX0A:~# stcli cluster storage-summary
address: 169.254.37.1
name: All-NVMe
state: online
uptime: 0 days 2 hours 27 minutes 43 seconds
activeNodes: 8 of 8
compressionSavings: 76.99%
deduplicationSavings: 0.0%
freeCapacity: 13.2T
healingInfo:
    inProgress: False
resiliencyInfo:
    messages:
        Storage cluster is healthy.
    state: 1
    nodeFailuresTolerable: 2
    cachingDeviceFailuresTolerable: 2
    persistentDeviceFailuresTolerable: 2
    zoneResInfoList: None
spaceStatus: normal
totalCapacity: 13.4T
totalSavings: 76.99%
usedCapacity: 148.6G
zkHealth: online
clusterAccessPolicy: lenient
dataReplicationCompliance: compliant
dataReplicationFactor: 3
```

2. Run the following command in the shell, and press enter:

    `/usr/share/springpath/storfs-misc/hx-scripts/post_install.py`

3. Choose the first post_install workflow type – New/Existing Cluster.

4. Enter the HX Storage Controller VM root password for the HX cluster (use the one entered during the HX Cluster installation).

5. Enter the vCenter server username and password.

```
root@SpringpathControllerT7DB8MDX0A:~# /usr/share/springpath/storfs-misc/hx-scripts/post_install.py

Select post_install workflow-

 1. New/Existing Cluster
 2. Expanded Cluster
 3. Generate Certificate

 Note:  Workflow No.3 is mandatory to have unique SSL certificate in the cluster.
        By Generating this certificate, it will replace your current certificate.
        If you're performing cluster expansion, then this option is not required.

 Selection: 1
Logging in to controller localhost
HX CVM admin password:
Getting ESX hosts from HX cluster...
vCenter URL: 10.29.133.120
Enter vCenter username (user@domain): administrator@vsphere.local
vCenter Password:
Found datacenter Datacenter
Found cluster All-NVMe

post_install to be run for the following hosts:
 hxaf220m5n-01.hx.lab.cisco.com
 hxaf220m5n-02.hx.lab.cisco.com
 hxaf220m5n-03.hx.lab.cisco.com
 hxaf220m5n-04.hx.lab.cisco.com
 hxaf220m5n-05.hx.lab.cisco.com
 hxaf220m5n-06.hx.lab.cisco.com
 hxaf220m5n-07.hx.lab.cisco.com
 hxaf220m5n-08.hx.lab.cisco.com
```

6. Enter ESXi host root password (use the one entered during the HX Cluster installation).

7. You must license the vSphere hosts through the script or complete this task in vCenter before continuing. Failure to apply a license will result in an error when enabling HA or DRS in subsequent steps. Enter "n" if you have already registered the license information in vCenter.

8. Enter "y" to enable HA/DRS.

9. Enter "y" to disable the ESXi hosts' SSH warning. SSH running in ESXi is required in HXDP 2.6.

10. Add the vMotion VMkernel interfaces to each node by entering "y". Input the netmask, the vMotion VLAN ID, and the vMotion IP addresses for each of the hosts as prompted.

```
 Enter ESX root password:

Enter vSphere license key?  (y/n) n

Enable HA/DRS on cluster? (y/n) y
Successfully completed configuring cluster HA.
Successfully completed configuring cluster DRS.

Disable SSH warning? (y/n) y

Add vmotion interfaces? (y/n) y
 Netmask for vMotion: 255.255.255.0
 VLAN ID: (0-4096) 200
 vMotion MTU is set to use jumbo frames (9000 bytes). Do you want to change to 1500 bytes? (y/n) n
 vMotion IP for hxaf220m5n-01.hx.lab.cisco.com: 192.168.200.61
 Adding vmotion-200 to hxaf220m5n-01.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-01.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-02.hx.lab.cisco.com: 192.168.200.62
 Adding vmotion-200 to hxaf220m5n-02.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-02.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-03.hx.lab.cisco.com: 192.168.200.63
 Adding vmotion-200 to hxaf220m5n-03.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-03.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-04.hx.lab.cisco.com: 192.168.200.64
 Adding vmotion-200 to hxaf220m5n-04.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-04.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-05.hx.lab.cisco.com: 192.168.200.65
 Adding vmotion-200 to hxaf220m5n-05.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-05.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-06.hx.lab.cisco.com: 192.168.200.66
 Adding vmotion-200 to hxaf220m5n-06.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-06.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-07.hx.lab.cisco.com: 192.168.200.67
 Adding vmotion-200 to hxaf220m5n-07.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-07.hx.lab.cisco.com
 vMotion IP for hxaf220m5n-08.hx.lab.cisco.com: 192.168.200.68
 Adding vmotion-200 to hxaf220m5n-08.hx.lab.cisco.com
 Adding vmkernel to hxaf220m5n-08.hx.lab.cisco.com
```

11. You may add VM network portgroups for guest VM traffic. Enter "n" to skip this step and create the portgroups manually in vCenter. Or if desired, VM network portgroups can be created and added to the vm-network vSwitch. This step will add identical network configuration to all nodes in the cluster.

12. Enter "y" to run the health check on the cluster.

13. A summary of the cluster will be displayed upon completion of the script. Make sure the cluster is healthy.

Initial Tasks and Testing

Datastores

Create a datastore for storing the virtual machines. This task can be completed by using the vSphere Web Client HX plugin, or by using the HyperFlex Connect HTML management webpage. To configure a new datastore via the HyperFlex Connect webpage, follow these steps:

1. Use a web browser to open the HX cluster IP management URL.

2. Enter a local credential, or a vCenter RBAC credential with administrative rights for the username, and the corresponding password.

3. Click Login.

4. Click Datastores and then click Create Datastore.

5. In the popup, enter the Datastore Name and size. For most applications, leave the Block Size at the default of 8K. Only dedicated Virtual Desktop Infrastructure (VDI) environments should choose the 4K Block Size option.

6. Click Create Datastore.
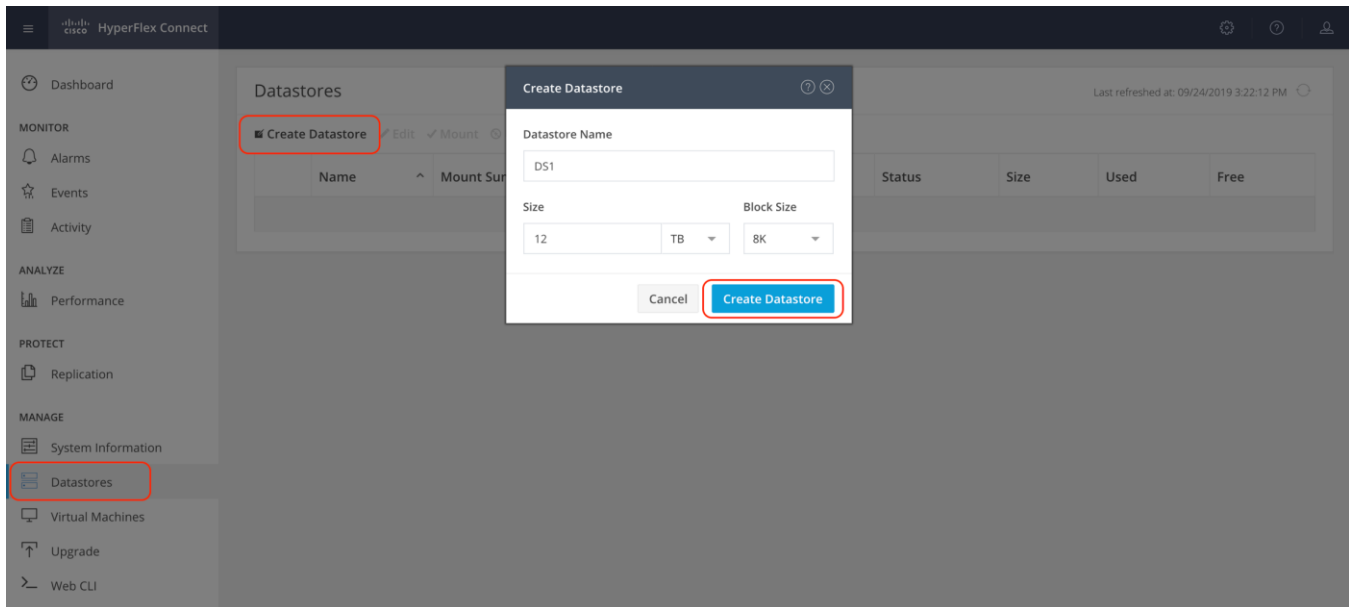
Create VM

In order to perform the initial testing and to learn about the features in the HyperFlex cluster, create a test virtual machine stored on your new HX datastore in order to take a snapshot and perform a cloning operation.
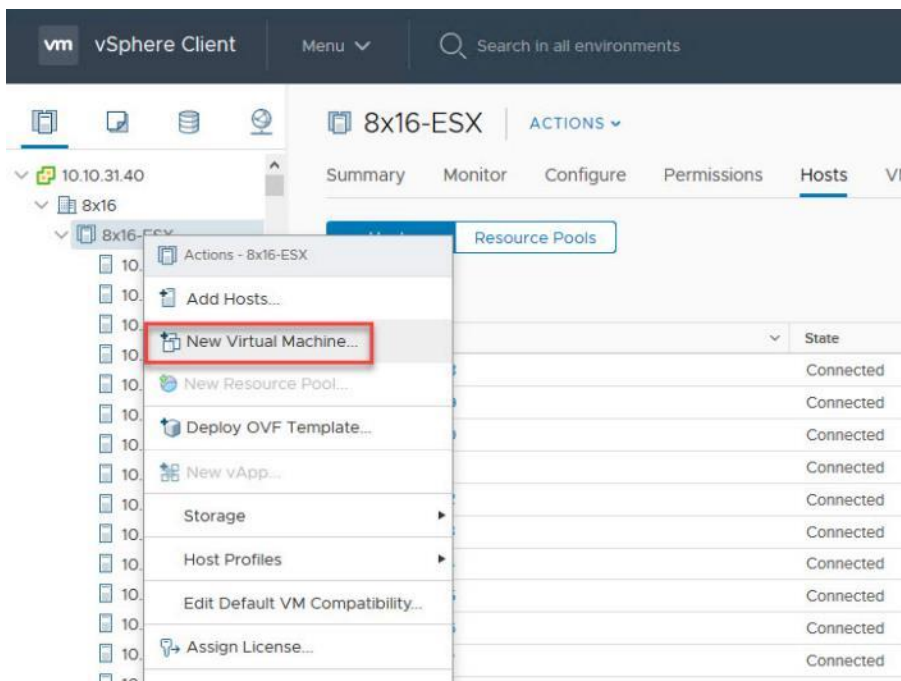
HX Data Platform Native Snapshots Overview

HX Data Platform Native Snapshots are a backup feature that saves versions (states) of working VMs. VMs can be reverted to native snapshots.

Use the HX Data Platform plug-in to take native snapshots of your VMs. HX Data Platform native snapshot options include create a native snapshot, revert to any native snapshot, and delete a native snapshot. Timing options include: Hourly, Daily, and Weekly, all in 15-minute increments.

A native snapshot is a reproduction of a VM that includes the state of the data on all VM disks and the VM power state (on, off, or suspended) at the time the native snapshot is taken. Take a native snapshot to save the current state of the VM, so that you have the option to revert to the saved state.

You can take a native snapshot when a VM is powered on, powered off, or suspended. For additional information about VMware snapshots, see the VMware KB, Understanding VM snapshots in ESXi (1015180)

Benefits of HX Data Platform Native Snapshots

HX Data Platform native snapshots use native technology. Native snapshots provide the following benefits:

- Reverting registered VMs - If a VM is registered, whether powered-on or powered-off, native snapshots, same as VM snapshots, can be used to revert to an earlier point in time at which the snapshot was created.

- High performance - The HX Data Platform native snapshot process is fast because it does not incur I/O overhead.

- VM I/O independent - The HX Data Platform native snapshot creation time is independent of the I/O on the VM.

- VM performance - HX Data Platform native snapshots do not degrade VM performance.

- Crash-consistent. HX Data Platform native snapshots are crash-consistent by default - I/O crash consistency is defined as maintaining the correct order of write operations to enable an application to restart properly from a crash.

- Application-consistent - You can select the quiesce option of the stcli vm snapshot command through the HX Data Platform CLI to enable HX Data Platform native snapshots to be application-consistent. The applications in the guest VM run transparently exactly like they do in the host VM. For details, see the Cisco HX Data Platform Command Line Interface Reference CLI Reference.

Quiescing a file system is a process of bringing the on-disk data of a physical or virtual computer into a state suitable for backups. This process might include operations such as flushing dirty buffers from the operating system's in-memory cache to disk, or other higher-level application-specific tasks.

If your system displays quiesce errors, review the VMware KB article Troubleshooting Volume Shadow Copy (VSS) quiesce related issues (1007696).

- Scheduled snapshots tolerant to node failures - Scheduled snapshots are tolerant to administrative operations that require a node shutdown, such as HX maintenance mode and HX online upgrade.

- Scheduled Snapshots are tolerant to failures in other HX clusters in multi cluster environments.

- Unified interface – You can manage native snapshots created through the HX Data Platform plug-in using VMware snapshot managerTM.

- Individual or grouped – You can take native snapshots on a VM level, VM folder level, or resource pool level.

- Granular progress and error reporting – These monitoring tasks performed at Task level for Resource Pool, Folder and VM level snapshot.

- Instantaneous snapshot delete – Deletion of a snapshot and consolidation is always instantaneous.

- Parallel batch snapshots – Support for up to 255 VMs in a Resource Pool or Folder for parallel batched snapshots.

- VDI deployment support – HX scheduled snapshots are supported for desktop VMs on VDI deployments which are using VMware native technology.

- Recoverable VM – The VM is always recoverable when there are snapshot failures.

- Datastore access – Snapshots work on partially mounted/accessible datastores as long as the VM being snapshotted is on an accessible mountpoint.

Snapshots

Take a snapshot of the new virtual machine prior to powering it on.

> When you create the first snapshot of a VM using Snapshot Now, the HX DataPlatform plug-in creates a base snapshot called a SENTINEL snapshot. The SENTINEL snapshot ensures follow-on snapshots are all native snapshots. SENTINEL snapshots prevent reverted VMs from having VMware redo log-based virtual disks. Redo log-based virtual disks occur when an original snapshot is deleted and the VM is reverted to the second oldest snapshot.

To take an instant snapshot of a VM, follow these steps:

1. In the HyperFlex Connect webpage, click the Virtual Machines menu, then click the name of the VM to snapshot.
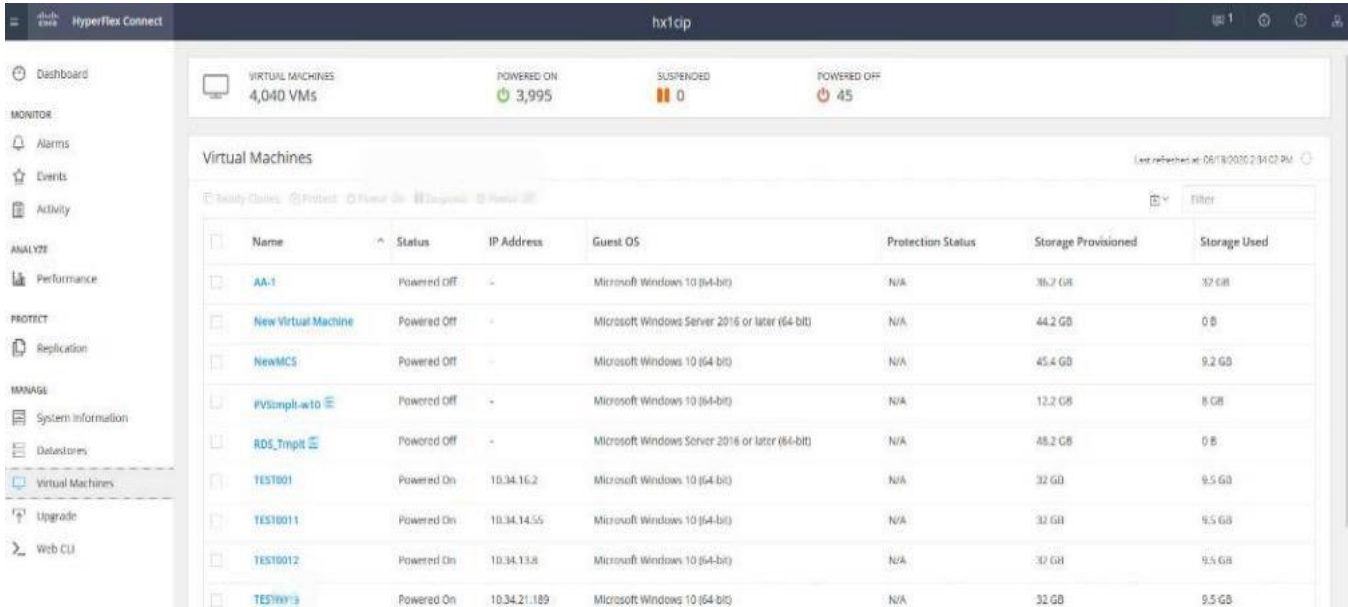
2. Click the Actions drop-down list, then choose Snapshot Now.
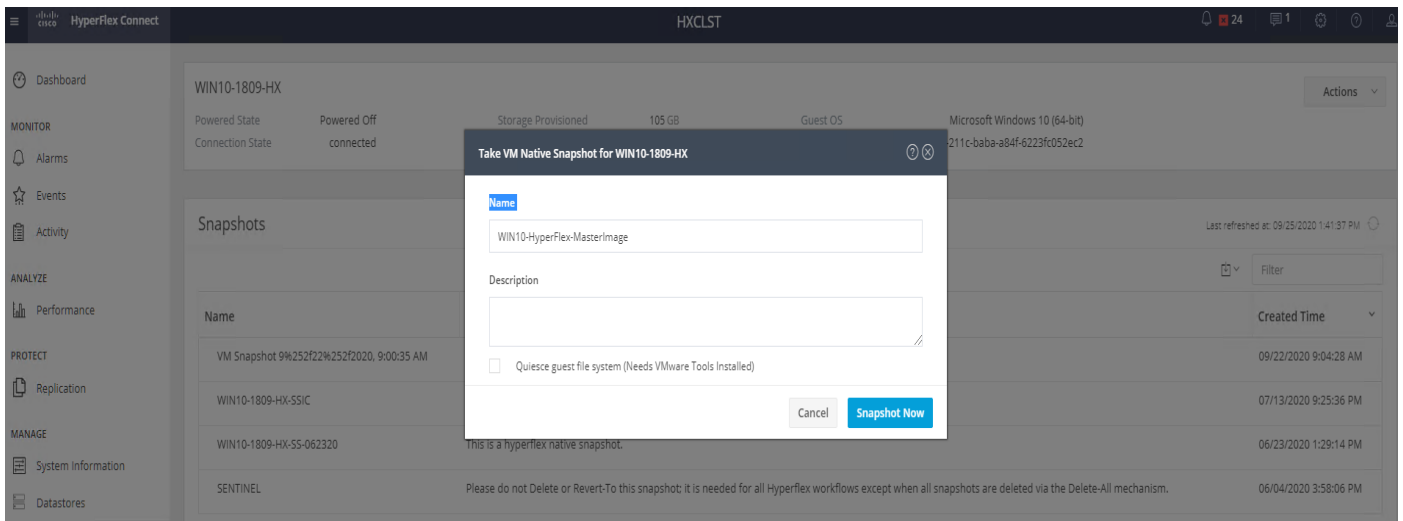
A Snapshot is created as shown below:

3. Input the snapshot name, a description if desired, and choose whether to quiesce the VM, then click Snapshot Now.

Hyperflex Ready Clones

HX Data Platform Ready Clones is a pioneer storage technology that enables you to rapidly create and customize multiple cloned VMs from a host VM. It enables you to create multiple copies of VMs that can then be used as standalone VMs.

A Ready Clone, similar to a standard clone, is a copy of an existing VM. The existing VM is called the host VM. When the cloning operation is complete, the Ready Clone is a separate guest VM.

Changes made to a Ready Clone do not affect the host VM. A Ready Clone's MAC address and UUID are different from that of the host VM.

Installing a guest operating system and applications can be time consuming. With Ready Clone, you can make many copies of a VM from a single installation and configuration process.

Clones are useful when you deploy many identical VMs to a group.

To create the Ready Clones, follow these steps:

1. In the HyperFlex Connect webpage, click the Virtual Machines menu, click the checkbox to choose the VM to clone, then click Ready Clones.

2. Input the Number of clones to create, a customization specification if needed, a resource pool if needed, and a naming prefix, then click Clone to start the operation. The clones will be created in seconds.
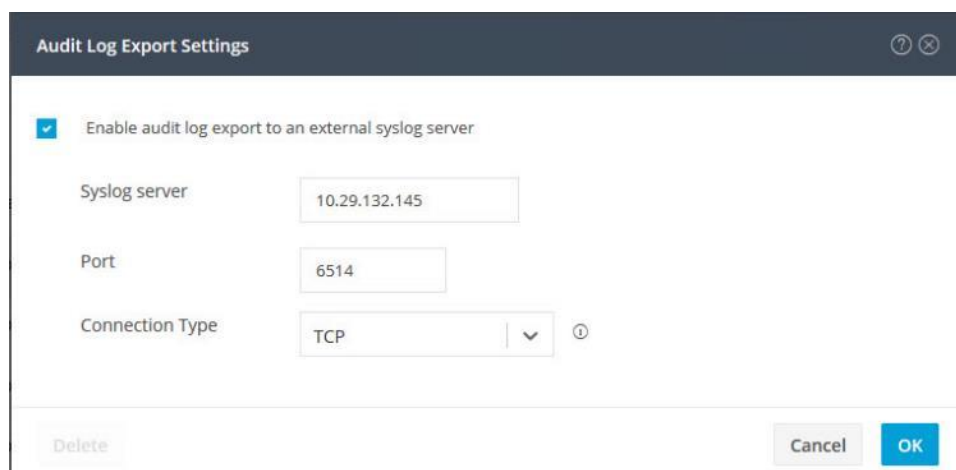
⚠ Ready clones were not used to clone virtual machines in this testing.

Audit Logging

By default, the HyperFlex controller VMs store logs locally for many functions, including the filesystem logs, security auditing, CLI commands and shell access, single sign-on logs, and more. These logs are rotated periodically and could be lost if there were a total failure of a controller VM. In order to store these logs externally from the HyperFlex cluster, audit logging can be enabled in HX Connect to send copies of these logs to an external syslog server. From this external location, logs can be monitored, generate alerts, and stored long term. HX

Connect will not monitor the available disk space on the syslog destination, nor will it generate an alarm if the destination server is full. To enable audit logging, follow these steps:

1. Use a web browser to open the HX cluster IP management URL.

2. From the HyperFlex Connect webpage, click the gear shaped icon in the upper right-hand corner, and click Audit Log Export Settings.

3. Click to check the box to Enable audit log export to an external syslog server.

4. Enter the syslog server IP address and TCP port.

5. Choose TCP or TLS as the connection type. If using TLS, client certificate and private key pair files must be provided. Alternatively, a self-signed certificate can be used. Click browse to choose the appropriate files.

6. Click OK.

Audit log exports can be temporarily disabled or completely deleted at a later time from the same location.

To store ESXi diagnostic logs in a central location in case they are needed to help diagnose a host failure, it is recommended to enable a syslog destination for permanent storage of the ESXi host logs for all Cisco HyperFlex hosts. It is possible to use the vCenter server as the log destination in this case, or another syslog receiver of your choice.

To configure syslog for ESXi, follow these steps:

1. Log into the ESXi host via SSH as the root user.

2. Enter the following commands, replacing the IP address in the first command with the IP address of the vCenter server that will receive the syslog logs:

```
[root@hx220-01:~] esxcli system syslog config set --loghost='udp://10.29.132.120'
[root@hx220-01:~] esxcli system syslog reload
[root@hx220-01:~] esxcli network firewall ruleset set -r syslog -e true
[root@hx220-01:~] esxcli network firewall refresh
```

3. Repeat for each ESXi host.

Auto-Support and Notifications

Auto-Support should be enabled for all clusters during the initial HyperFlex installation. Auto-Support enables Call Home to automatically send support information to Cisco TAC, and notifications of tickets to the email address specified. If the settings need to be modified, they can be changed in the HyperFlex Connect HTML management webpage.

The events that will automatically open a support ticket with Cisco TAC are as follows:

- Cluster Capacity Changed
- Cluster Unhealthy
- Cluster Health Critical
- Cluster Read Only
- Cluster Shutdown
- Space Warning
- Space Alert
- Space Critical
- Disk Blacklisted
- Infrastructure Component Critical
- Storage Timeout

To change Auto-Support settings, follow these steps:

1. From the HyperFlex Connect webpage, click the gear shaped icon in the upper right-hand corner, and click Auto-Support Settings.

2. Enable or disable Auto-Support as needed.

3. Enter the email address to receive alerts when Auto-Support events are generated.

4. Enable or disable Remote Support as needed. Remote support allows Cisco TAC to connect to the HX cluster and accelerate troubleshooting efforts.

5. Enter in the information for a web proxy if needed.

6. Click OK.
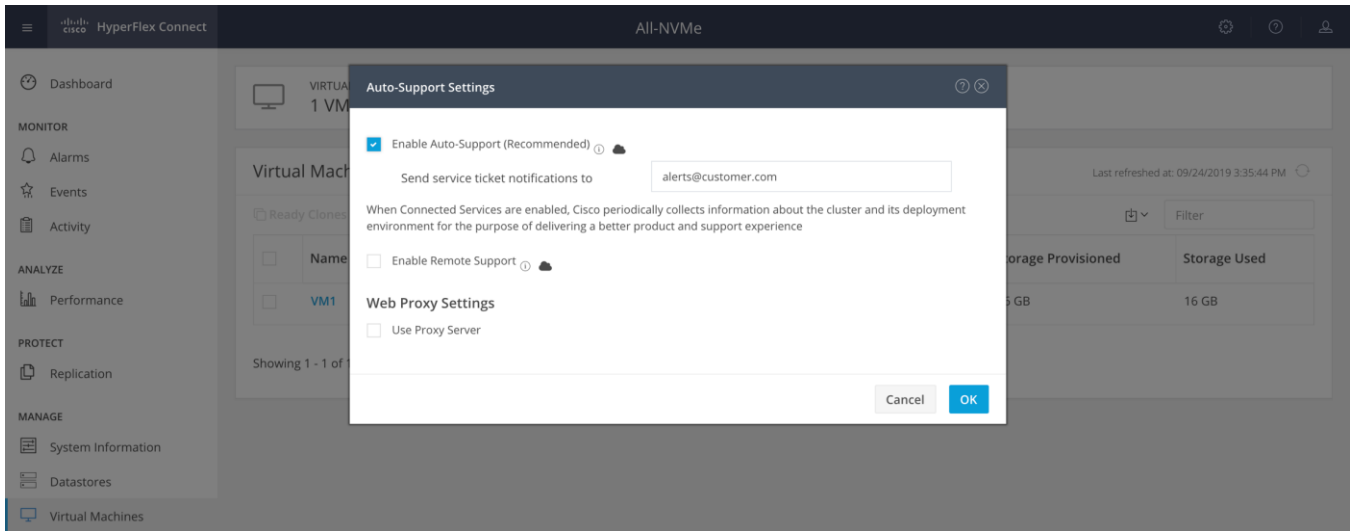
![warning icon] Email notifications that come directly from the HyperFlex cluster can also be enabled.

To enable direct email notifications, follow these steps:

1. From the HyperFlex Connect webpage, click the gear shaped icon in the upper right-hand corner, and click Notifications Settings.

2. Enter the DNS name or IP address of the outgoing email server or relay, the email address the notifications will come from, and the recipients.

3. Click OK.
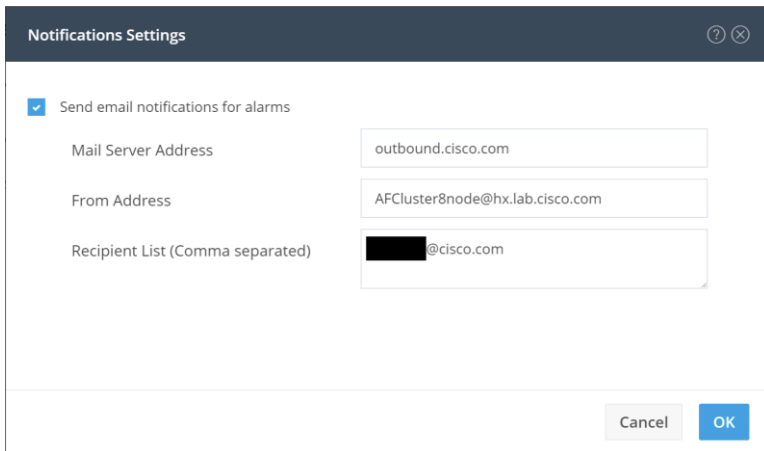
Smart Licensing

HyperFlex 2.5 and later utilizes Cisco Smart Licensing, which communicates with a Cisco Smart Account to validate and check out HyperFlex licenses to the nodes, from the pool of available licenses in the account. At the beginning, Smart Licensing is enabled but the HX storage cluster is unregistered and in a 90-day evaluation period or EVAL MODE. For the HX storage cluster to start reporting license consumption, it must be registered with the Cisco Smart Software Manager (SSM) through a valid Cisco Smart Account. Before beginning, verify that you have a Cisco Smart account, and valid HyperFlex licenses are available to be checked out by your HX cluster.

To create a Smart Account, click
https://webapps.cisco.com/software/company/smartaccounts/home?route=module/accountcreation then go to
Cisco Software Central > Request a Smart Account

To activate and configure smart licensing, follow these steps:

1. Log into a controller VM. Confirm that your HX storage cluster is in Smart Licensing mode.

   ```
   # stcli license show status

   Smart Licensing is ENABLED

   Registration:
     Status: UNREGISTERED
     Export-Controlled Functionality: Not Allowed

   License Authorization:
     Status: EVAL MODE
     Evaluation Period Remaining: 88 days, 1 hr, 33 min, 41 sec
     Last Communication Attempt: NONE

   License Conversion:
    Automatic Conversion Enabled: true
    Status: NOT STARTED

   Utility:
     Status: DISABLED

   Transport:
     Type: TransportCallHome
   ```

2. Feedback will show Smart Licensing is ENABLED, Status: UNREGISTERED, and the amount of time left during the 90-day evaluation period (in days, hours, minutes, and seconds).

3. Navigate to Cisco Software Central (https://software.cisco.com/) and log in to your Smart Account.

4. From Cisco Smart Software Manager, generate a registration token.

5. In the License pane, click Smart Software Licensing to open Cisco Smart Software Manager.

6. Click Inventory.

7. From the virtual account where you want to register your HX storage cluster, click General, and then click New Token.

8. In the Create Registration Token dialog box, add a short Description for the token, enter the number of days you want the token to be active and available to use on other products, and check Allow export-controlled functionality on the products registered with this token.

9. Click Create Token.

10. From the New ID Token row, click the Actions drop-down list, and click Copy.

11. Log into a controller VM.

12. Register your HX storage cluster, where *idtoken-string* is the New ID Token from Cisco Smart Software Manager.

```
# stcli license register --idtoken idtoken-string
```

13. Confirm that your HX storage cluster is registered.

```
# stcli license show summary
```

The cluster is now ready. You may run any other preproduction tests that you wish to run at this point.

Virtual Machine Snapshot and Cloning

To take a VM snapshot and clone, follow these steps:

1. Right-click on master image and click Snapshots.
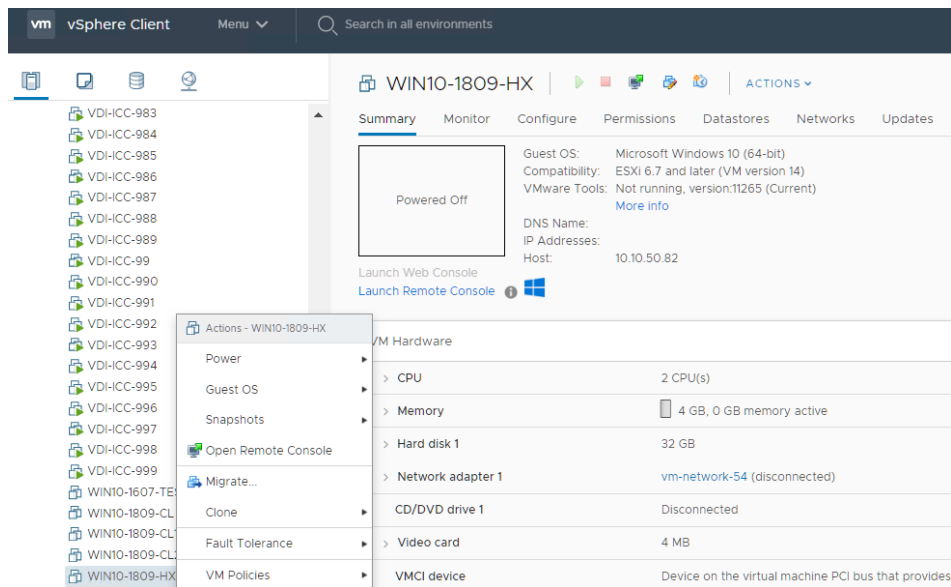
2. Provide name of master image snapshot.
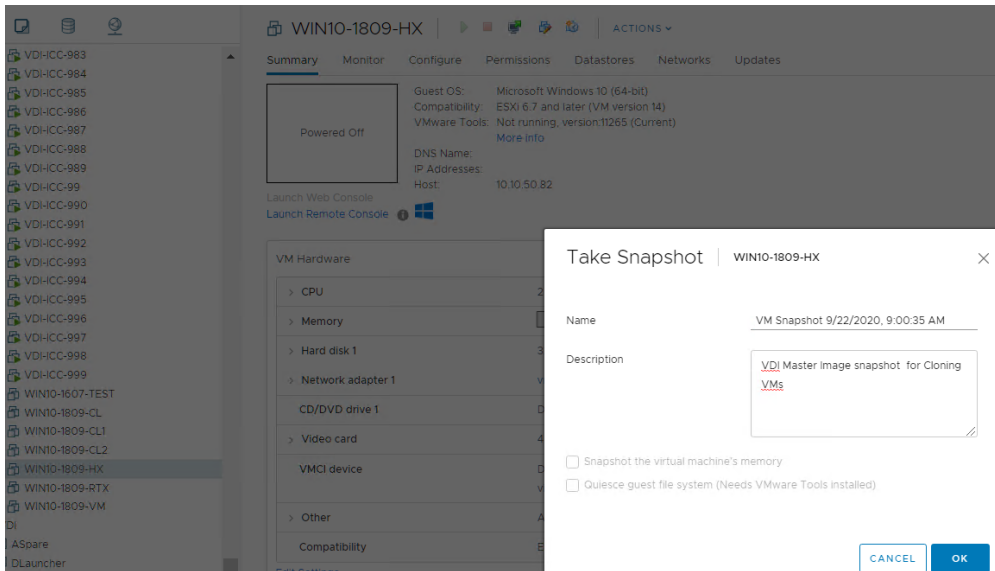
ESXi Hypervisor Installation

HX converge nodes come from the factory with a copy of the ESXi hypervisor pre-installed, however, the compute only nodes do not come from factory with ESXi pre-installed. However, there are scenarios where it may be necessary to redeploy or reinstall ESXi on an HX node.

In addition, this process can be used to deploy ESXi on rack mount or blade servers that will function as HX compute-only nodes. The HyperFlex system requires a Cisco custom ESXi ISO file to be used, which has Cisco hardware specific drivers pre-installed, and customized settings configured to ease the installation process. The Cisco custom ESXi ISO file is available to download at cisco.com.

ESXi Kickstart ISO

The HX custom ISO is based on the Cisco custom ESXi 6.7 Update 3 ISO release with the filename: *HX-ESXi-6.7U3-15160138-Cisco-Custom-6.7.3.3-install-only.iso* and is available from the Cisco web site:

https://software.cisco.com/download/home/286305544/type/286305994/release/4.0(2a)

The custom Cisco HyperFlex ESXi ISO automatically performs the following tasks with no user interaction required:

- Accept the End User License Agreement

- Configure the root password to: Cxxxxxxx

- Install ESXi to the internal mirrored Cisco FlexFlash SD cards, or the internal M.2 SSD

- Set the default management network to use vmnic0, and obtain an IP address via 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

Cisco UCS vMedia and Boot Policies

By using a Cisco UCS vMedia policy, the custom Cisco HyperFlex ESXi installation ISO file can be mounted to all of the HX servers automatically. 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.

⚠ **WARNING!** While vMedia policies are very efficient for installing multiple servers, using vMedia policies as described could lead to an accidental reinstall of ESXi on any existing server that is rebooted with this policy applied. Please be certain that the servers being rebooted while the policy is in effect are the servers you wish to reinstall. Even though the custom ISO will not continue without a secondary confirmation, extreme caution is recommended. This procedure needs to be carefully monitored and the boot policy should be changed back to original settings immediately after the intended servers are rebooted, and the ESXi installation begins. Using this policy is only recommended for new installs or rebuilds. Alternatively, you can manually choose the boot device using the KVM console during boot, and pressing F6, instead of making the vMedia device the default boot selection.

To configure the Cisco UCS vMedia and Boot Policies, follow these steps:

1. Copy the *HX-ESXi-6.7U3-15160138-Cisco-Custom-6.7.3.3-install-only.iso* file to an available web server folder, NFS share or CIFS share. In this example, an open internal web server folder is used.

2. In Cisco UCS Manager, click the Servers button on the left-hand side of the screen.

3. Expand Servers > Policies > root > Sub-Organizations > <<HX_ORG>> > vMedia Policies and click vMedia Policy HyperFlex.

4. In the configuration pane, click Create vMedia Mount.

5. Enter a name for the mount, for example: ESXi.

6. Choose the CDD option.

7. Choose CIFS as the protocol.

8. Enter the IP address of the CIFS server where the file was copied, for example: 10.29.132.120

9. Choose None as the Image Variable Name.

10. Enter HX-ESXi-6.7U3-15160138-Cisco-Custom-6.7.3.3-install-only.iso as the Remote File.

11. Enter the Remote Path to the installation file.

Create vMedia Mount ? ✕

Name	:	ESXi
Description	:	
Device Type	:	⦿ CDD ◯ HDD
Protocol	:	◯ NFS ⦿ CIFS ◯ HTTP ◯ HTTPS
Authentication Protocol :		Ntlm ▼
Hostname/IP Address	:	10.29.132.145
Image Name Variable	:	⦿ None ◯ Service Profile Name
Remote File	:	HX-ESXi-6.7U3-15160138-Cisco-Custom-6.7.3.3-
Remote Path	:	iso
Username	:		
Password	:	
Reman on Eiect	·	☐

OK   **Cancel**

12. Click OK.

13. Choose Servers > Service Profile Templates > root > Sub-Organizations > <<HX_ORG>> > Service Template hx-nodes.

14. In the configuration pane, click the vMedia Policy tab.

15. Click Modify vMedia Policy.

16. Chose the HyperFlex vMedia Policy from the drop-down selection and click OK twice.

17. For Compute-Only nodes (if necessary), choose Servers > Service Profile Templates > root > Sub-Organizations > <<HX_ORG>> > Service Template compute-nodes.

18. In the configuration pane, click the vMedia Policy tab.

19. Click Modify vMedia Policy.

20. Chose the HyperFlex vMedia Policy from the drop-down selection and click OK twice.

21. Choose Servers > Policies > root > Sub-Organizations > <<HX_ORG>> > Boot Policy HyperFlex.

22. In the navigation pane, expand the section titled CIMC Mounted vMedia.

23. Click the entry labeled Add CIMC Mounted CD/DVD.

24. Choose 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.

25. Click Save Changes and click OK.

Local Devices

CIMC Mounted vMedia

Add CIMC Mounted CD/DVD

Add CIMC Mounted HDD

vNICs

vHBAs

iSCSI vNICs

EFI Shell

Boot Order

+ − Advanced Filter ↑ Export 🖶 Print

Name	Or... ▲	vNIC...	Type	WWN	LUN ...	Slo
CIMC Mounted CD/DVD	1					
CD/DVD	2					
SD Card	3					

↑ Move Up   ↓ Move Down   🗑 Delete

Set Uefi Boot Parameters

Install ESXi

To begin the installation after modifying the vMedia policy, Boot policy and service profile template, the servers need to be rebooted. To complete the reinstallation, it is necessary to open a remote KVM console session to each server being worked on. To open the KVM console and reboot the servers, follow these steps:

1. In Cisco UCS Manager, click Equipment.

2. Expand Equipment > Rack mounts > Servers > Server 1.

3. In the configuration pane, click KVM Console.

4. The remote KVM Console window will open in a new browser tab. Click Continue to any security alerts that appear and click the hyperlink to start the remote KVM session.

5. Repeat Steps 2-4 for all additional servers whose console you need to monitor during the installation.

6. In Cisco UCS Manager, click Equipment.

7. Expand Equipment > Rack-Mount Servers > Servers.

8. In the configuration pane, click the first server to be rebooted, then shift+click the last server to be rebooted, selecting all of the servers.

9. Right-click the mouse and click Reset.

Equipment

Main Topology View | Fabric Interconnects | Servers

Blade Servers | **Rack-Mount Servers**

⚙ Advanced Filter | ⬆ Export | 🖶 Print

Name	Overall ...	PID	Mo...	Serial	Pro...	Use...
Server 1	⬆ OK	HX...	Cis...	FC...	org...	
Server 2	⬆ OK	HX...	Cis...	FC...	org...	
Server 3	⬆ OK	HX...	Cis...	FC...	org...	
Server 4				C...	org...	
Server 5				C...		
Server 6				C...		
Server 7				C...		
Server 8				C...		

Boot Server
Shutdown Server
Reset
Start Fault Suppression
Stop Fault Suppression
Copy
Copy XML
Delete

10. Click OK.

11. Choose 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. Alternatively, the individual KVM consoles can be used to perform a power cycle one-by-one.

13. When the server boots from the installation ISO file, you will see a customized Cisco boot menu. In the Cisco customized installation boot menu, choose "HyperFlex Converged Node – HX PIDs Only" and press Enter.

14. Enter "ERASE" in all uppercase letters, and press Enter to confirm and install ESXi.

15. (Optional) When installing Compute-Only nodes, the appropriate Compute-Only Node option for the boot location to be used should be selected. The "Fully Interactive Install" option should only be used for debugging purposes.

The ESXi installer continues the installation process automatically, there may be error messages seen on screen temporarily, but they can be safely ignored. When the process is complete, the standard ESXi console screen will display as shown below:
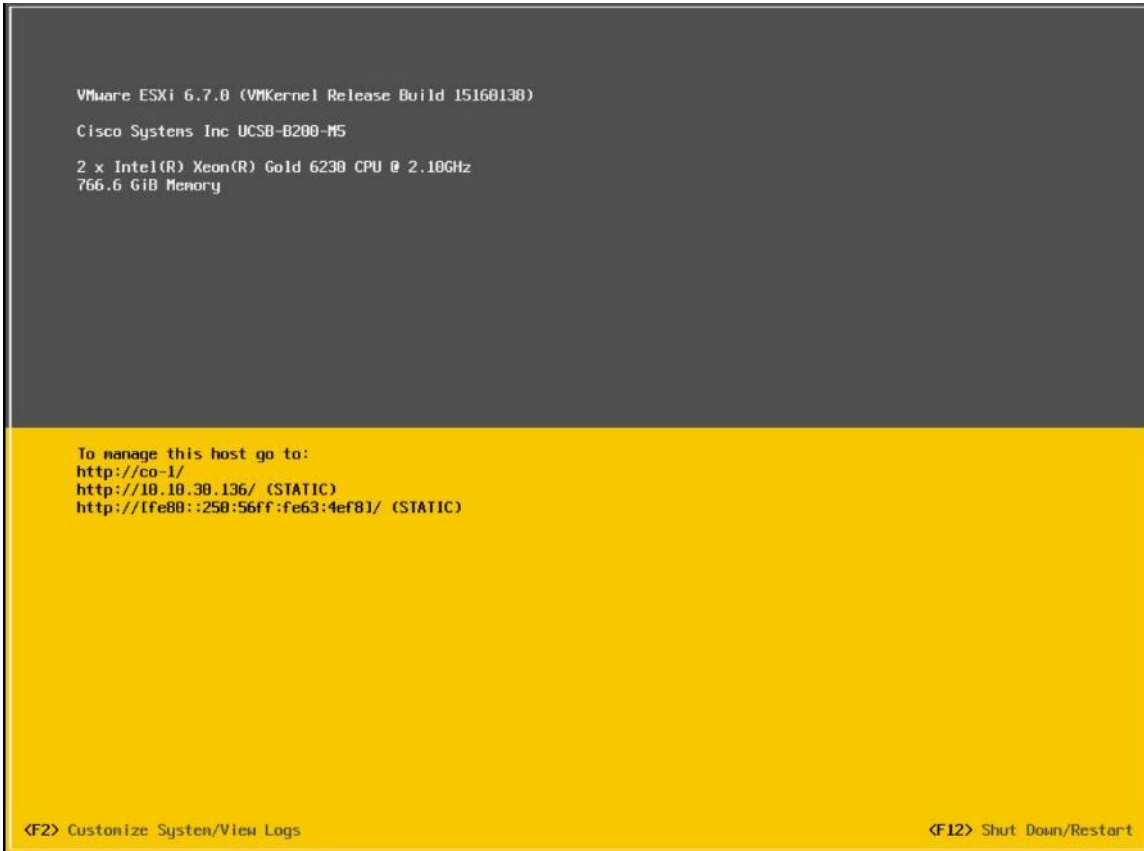
Undo vMedia and Boot Policy Changes

When 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, follow these steps:

1. Choose Servers > Policies > root > Sub-Organizations > <<HX_ORG>> > Boot Policy HyperFlex.

2. Choose the CIMC Mounted CD/DVD entry in the Boot Order list and click Delete.

3. Click Save Changes and click OK.

The changes made to the vMedia policy and service profile template may also be undone once the ESXi installations have all completed fully, or they may be left in place for future installation work.

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 an extended cluster. At the time of this document, Cisco Intersight cannot perform HyperFlex cluster expansions, therefore the Cisco HyperFlex installer VM must be used. The Cisco HyperFlex installer VM is deployed via a downloadable ISO image from cisco.com.

Expansion with Compute-Only Nodes

The following technical guidelines must be followed when adding compute-only nodes to a Cisco HyperFlex cluster:

- The number of compute-only nodes cannot exceed the number of HyperFlex converged nodes within a single HyperFlex cluster unless the appropriate HyperFlex Enterprise licenses have been purchased, allowing up to a 2:1 ratio of compute-only nodes to converged nodes.

- The Cisco UCS infrastructure firmware revision, which provides the firmware for Cisco UCS Manager and the Fabric Interconnects, must be maintained at the minimum version required for the HyperFlex converged nodes, or higher, at all times.

- The version of VMware ESXi installed on the compute-only nodes must be compatible with the Cisco HyperFlex version in use, and it must match the version installed on the HyperFlex converged nodes.

- While the CPU models and memory capacities between the compute-only nodes and the HyperFlex converged nodes do not have to match, configuring the nodes to have similar capacities is recommended.

- Mixing different models of compute-only nodes is allowed within the same cluster. Example: using Cisco UCS C220 M3 and Cisco UCS C240 M4 servers as compute-only nodes is allowed.

- Mixing CPU generations will require configuring VMware Enhanced vMotion Compatibility (EVC) in order to allow vMotion to work between the compute-only nodes and the converged nodes. Enabling EVC typically requires all VMs to be powered off including the HyperFlex Storage Controller VMs, therefore the HyperFlex cluster must be shut down for an outage. If it is known ahead of time that EVC will be needed, then it is easier to create the vCenter cluster object and enable EVC prior to installing HyperFlex.

- Connectivity between compute-only nodes and the HyperFlex cluster must be within the same Cisco UCS domain, and networking speeds of the additional compute-only nodes should match the speeds of the existing converged nodes. Connecting compute-only nodes from a different Cisco UCS domain is not allowed, nor is connecting standalone rack-mount servers from outside of the Cisco UCS domain allowed.

- Blade servers installed in the Cisco UCS 5108 Blade Chassis can connect through 10 GbE,25GbE or 40 GbE chassis links, using the Cisco UCS 2204XP, 2208XP, or 2304 model Fabric Extenders. The Fabric Extenders, also called I/O Modules (IOMs), are typically installed in pairs, and connect the 5108 chassis to the Fabric Interconnects, which provide all the networking and management for the blades. Care must be taken not to oversubscribe and saturate the chassis links.

- Compute-only nodes can be configured to boot from SAN, local disks, or internal SD cards. No other internal storage should be present in a compute-only node. Manual configuration of the appropriate boot policy will be necessary if booting from any device other than SD cards.

- Compute-only nodes can be configured with additional vNICs or vHBAs in order to connect to supported external storage arrays via NFS, iSCSI or Fibre Channel, in the same way as HyperFlex converged nodes are allowed to do.

- Care must be taken that the addition of the compute-only nodes will not significantly impact the HyperFlex cluster by creating additional load, or by consuming too much space. Pay close attention to the space consumption and performance requirements of any net-new VMs that will run on the additional compute-only nodes, and also note the current cluster performance and space utilization. If no new VMs will be created, then the current cluster performance will not be impacted.

The Cisco HyperFlex installer VM has a wizard for Cluster Expansion with converged nodes and compute-only nodes, however the compute-only node process requires some additional manual steps to install the ESXi hypervisor on the nodes. To expand an existing HyperFlex cluster with compute-only nodes, creating an extended HyperFlex cluster, follow these steps:

1. On the HyperFlex installer webpage click the drop-down list for Expand Cluster, then click Standard Cluster.
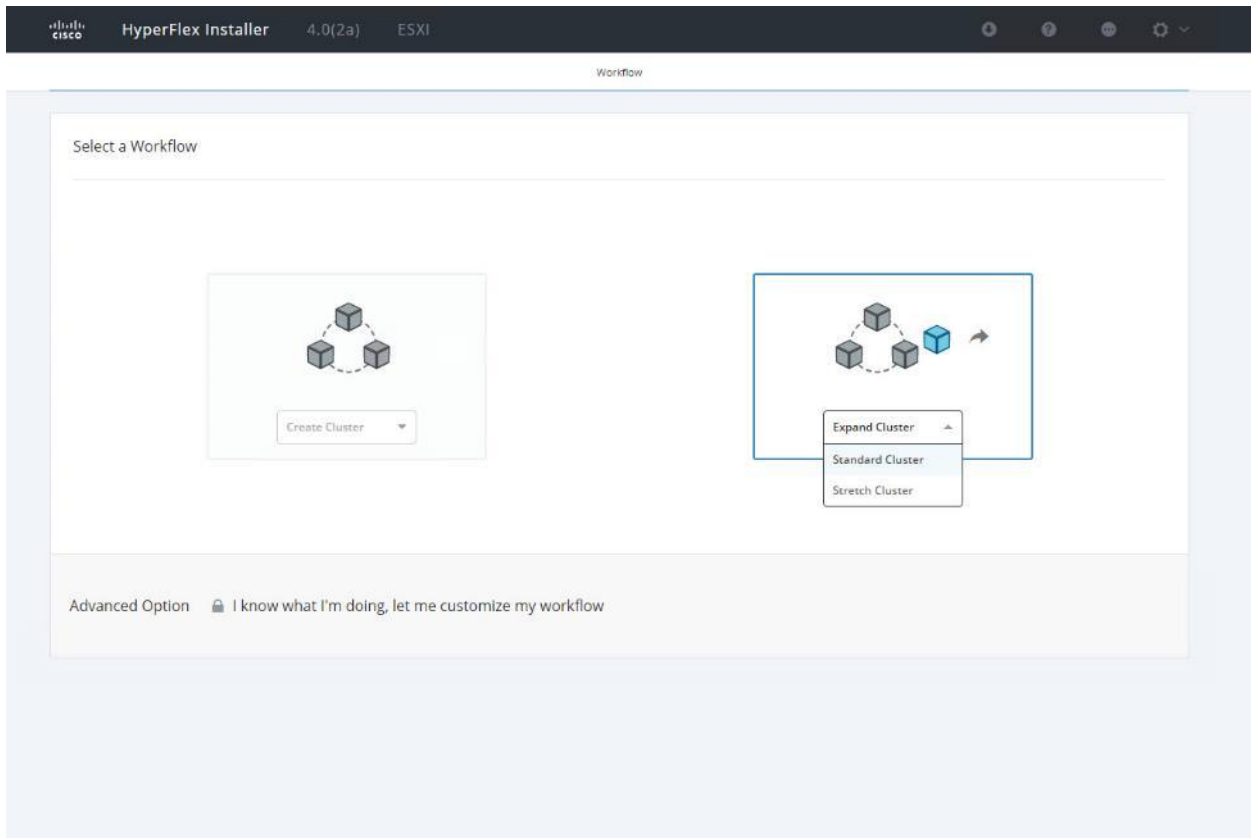
2. Enter the Cisco UCS Manager and vCenter DNS hostname or IP address, the admin usernames, and the passwords for the UCS domain where the existing and new nodes are, and the managing vCenter server of the cluster to be expanded. Optionally, you can import a JSON file that has the configuration information, except for the appropriate passwords. Click Continue.

3. Choose the HX cluster to expand and enter the cluster management password, then click Continue. If the installer has been reset, or otherwise does not detect a cluster to expand, enter the HX cluster management IP address, username, and password of a different cluster instead.
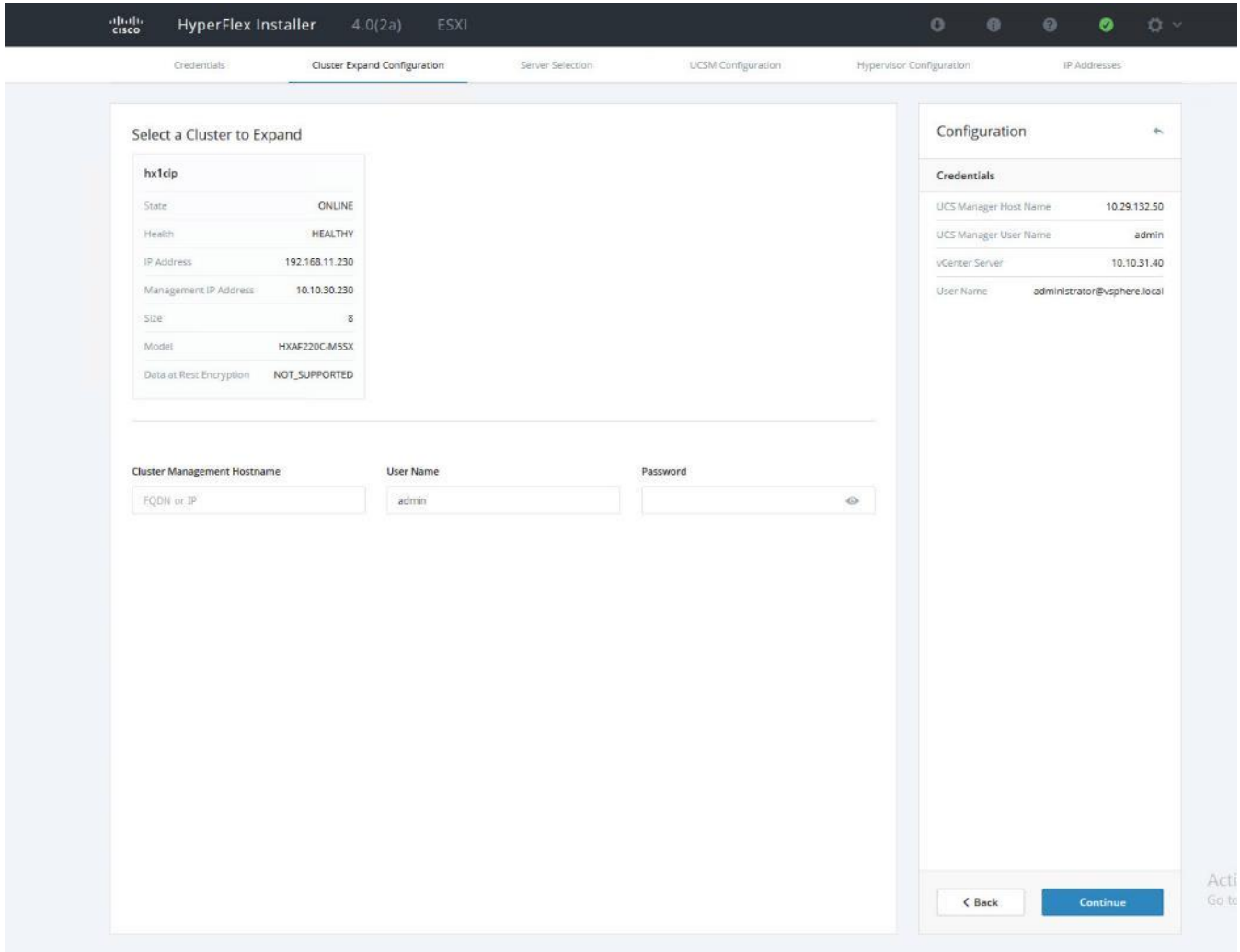
4. Choose the unassociated compute-only servers you want to add to the existing HX cluster, then click Continue.

5. On the UCSM Configuration page, all the settings should be pre-populated with the correct values for the existing cluster. The only value that is required is to create an additional IP address block for the hx-ext-mgmt IP address pool. Enter a range of IP addresses sufficient to assign to the new server(s), along with the subnet mask and gateway address, then click Continue.
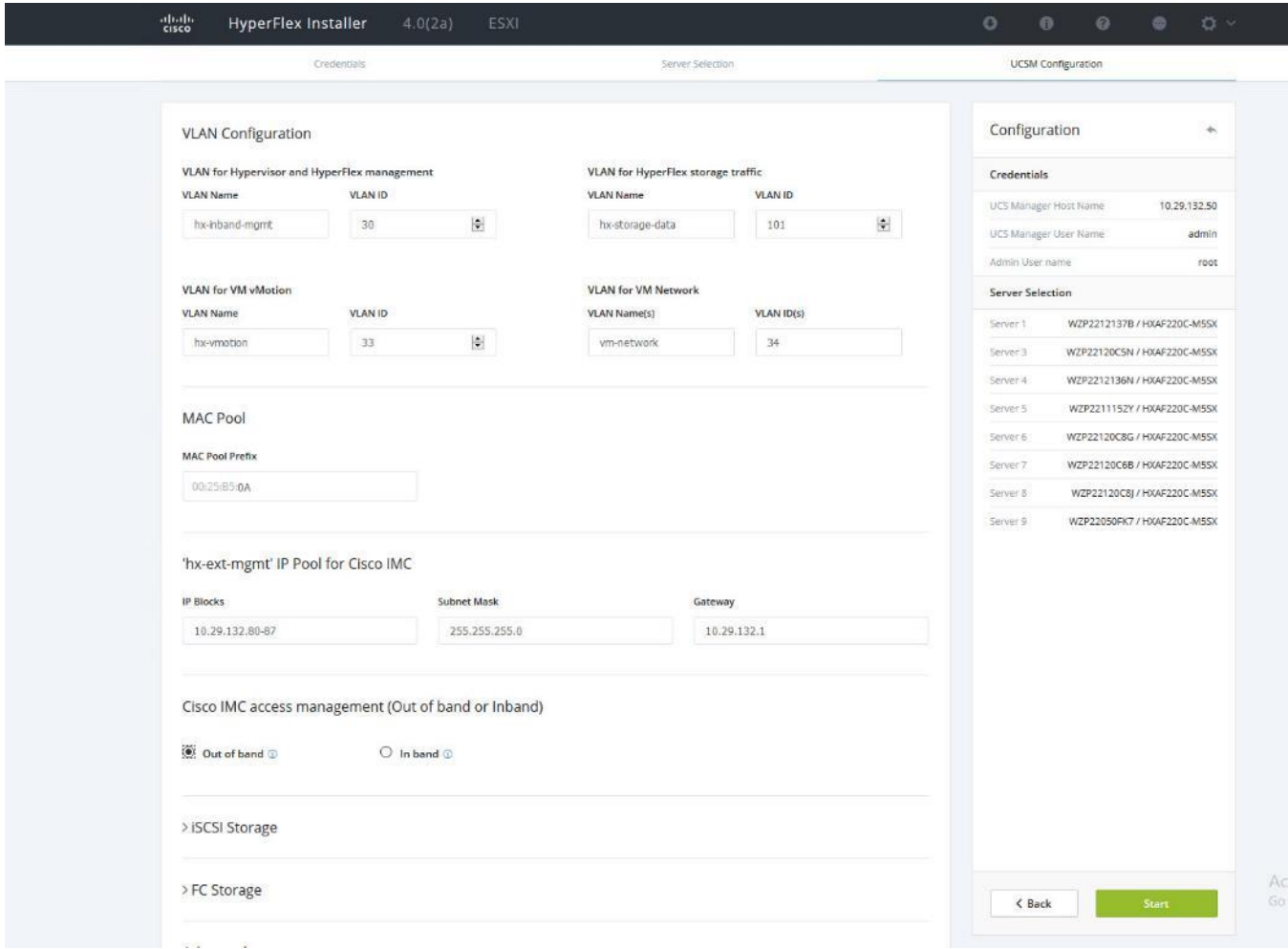
6. Enter the subnet mask, gateway, DNS, and IP addresses for the Hypervisors (ESXi hosts) as well as host names, then click Continue. The IPs will be assigned through Cisco UCS Manager to the ESXi systems.

Credentials   Cluster Expand Configuration   Server Selection   UCSM Configuration   Hypervisor Configuration   IP Addresses

CIMC IP Address Allocation

Number of Servers to be expanded: 16

Number of IP Addresses required: 16

Number of IP Addresses available: 0

Add CIMC Block to Pool   ✕

Starting IP Block   **Number of Addresses**

10.29.132.102   16

Subnet Mask   **Gateway**

255.255.255.0   10.29.132.1

Cancel   **Add**

Add IP Address to CIMC IP Pool

> Show Existing IP Pool Inventory

Configuration   ↩

Credentials

UCS Manager Host Name   10.29.132.50

UCS Manager User Name   admin

vCenter Server   10.10.31.40

User Name   administrator@vsphere.local

Cluster Expand Configuration

Management Cluster   10.10.30.230

Server Selection

Server 2/2   FCH22247MMC / UCSB-B200-M5
Server 2/3   FCH22247KXN / UCSB-B200-M5
Server 2/1   FCH22247MXM / UCSB-B200-M5
Server 2/6   FLM2331010R / UCSB-B200-M5
Server 2/7   FCH22247MZ7 / UCSB-B200-M5
Server 2/4   FCH22247N42 / UCSB-B200-M5
Server 2/5   FCH22247LJF / UCSB-B200-M5
Server 2/8   FCH22067RJP / UCSB-B200-M5
Server 1/5   FCH21447GNV / UCSB-B200-M5
Server 1/4   FCH22057T1X / UCSB-B200-M5
Server 1/7   FCH22067QZJ / UCSB-B200-M5
Server 1/6   FCH22067QYS / UCSB-B200-M5
Server 1/1   FCH22247M7Q / UCSB-B200-M5
Server 1/3   FCH22057SZM / UCSB-B200-M5
Server 1/2   FCH22057RWS / UCSB-B200-M5
Server 1/8   FCH22067QNZ / UCSB-B200-M5

< Back   Continue

Activate W
Go to Setting

Success: CIMC Block successfully created.   ✕

CIMC IP Address Allocation

Number of Servers to be expanded: 16

Number of IP Addresses required: 16

Number of IP Addresses available: 16

✓ Sufficient IP Addresses available.

Add IP Address to CIMC IP Pool

> Show Existing IP Pool Inventory

Configuration ↰

Credentials

UCS Manager Host Name	10.29.132.50
UCS Manager User Name	admin
vCenter Server	10.10.31.40
User Name	administrator@vsphere.local

Cluster Expand Configuration

Management Cluster	10.10.30.230

Server Selection

Server 2/2	FCH22247MMC / UCSB-B200-M5
Server 2/3	FCH22247KXN / UCSB-B200-M5
Server 2/1	FCH22247MXM / UCSB-B200-M5
Server 2/6	FLM2331010R / UCSB-B200-M5
Server 2/7	FCH22247MZ7 / UCSB-B200-M5
Server 2/4	FCH22247N42 / UCSB-B200-M5
Server 2/5	FCH22247LJF / UCSB-B200-M5
Server 2/8	FCH22067RJP / UCSB-B200-M5
Server 1/5	FCH21447GNV / UCSB-B200-M5
Server 1/4	FCH22057T1X / UCSB-B200-M5
Server 1/7	FCH22067QZJ / UCSB-B200-M5
Server 1/6	FCH22067QYS / UCSB-B200-M5
Server 1/1	FCH22247M7Q / UCSB-B200-M5
Server 1/3	FCH22057SZM / UCSB-B200-M5
Server 1/2	FCH22057RWS / UCSB-B200-M5
Server 1/8	FCH22067QNZ / UCSB-B200-M5

< Back   Continue

7. Enter the additional IP addresses for the Management and Data networks of the new nodes.

The HyperFlex Data VLAN IP addresses are automatically assigned during an installation via Cisco Intersight, however when expanding a cluster this step must be done manually. All addresses in the Data VLAN come from the link-local subnet of 169.254.0.0/16. The third octet is derived from converting the MAC address pool prefix into a binary number. It is critical to examine the existing addresses and take note of the existing value of the third octet for the vmk1 ports of the existing servers, as the subnet mask set on the hosts is actually 255.255.255.0. Therefore, if the third octet for the new values entered is not matched to the existing servers then there will be failures and errors. The value for the fourth octet is sequentially set, starting with .1 for the overall cluster, then proceeding to .2 for the vmk1(Hypervisor) port of the first server, then .3 for the Storage Controller VM of the first server. The second server would be assigned .4 for its vmk1 port, and .5 for its Storage Controller VM. It is important to note the ending values for these assignments among the existing servers, then continue this same addressing pattern for the new servers being added. In this example, a cluster with 4 con-verged nodes is being expanded with a 5th compute-only node, so the vmk1 (Hypervisor) port for the new server is .10, and there is no Storage Controller VM, so no IP addresses are required for that.

8. Enter the current password that is set on the Controller VMs.

9. (Optional) At this step more servers can be added for expansion if these servers already have service profiles associated and the hypervisor is ready, by clicking on Add Compute Server or Add Converged Server and then entering the IP addresses for the storage controller management and data networks.

10. Click Start.

11. Click Continue to accept the warning that by default new compute-only nodes are configured to boot from redundant Flex-flash cards. If necessary, follow the instructions referenced to create a new local disk configuration policy and boot policy for the new compute-only nodes.

12. Validation of the configuration will now start. After validation, the installer will create the compute-only node service profiles and associate them with the selected servers. Once the service profiles are associated, the installer will move on to the UCSM Configuration step. If the hypervisor is already installed, then move ahead to step 36. If the ESXi hypervisor has not been previously installed on the compute-only nodes, the installer will stop with errors as shown below due to the missing hypervisor. Continue to step 13 and do not click Retry UCSM Configuration until the hypervisor has been installed.

To install ESXi onto the new compute-only nodes, follow these steps:

1. In Cisco UCS Manager, click Servers.

2. Expand Servers > Service Profiles > root > Sub-Organizations > <<HX_ORG>>.

3. Each new compute-only node will have a new service profile, for example: chassis-1_blade-1. Right-click the new service profile and click KVM Console. The remote KVM console will open in a new browser tab. Accept any SSL errors or alerts, then click the link to launch the KVM console.
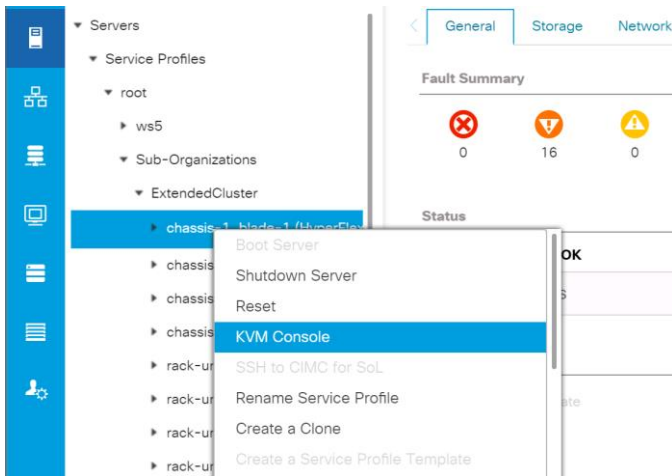
4. Repeat step 3 for each new service profile, that is associated with the new compute-only nodes.

5. In the remote KVM tab, click the Virtual Media button in the top right-hand corner of the screen, then click Activate Virtual Devices.
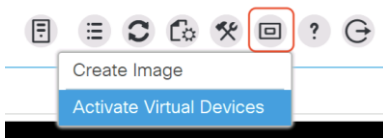
6. In the remote KVM tab, click Virtual Media, then click the CD/DVD option.
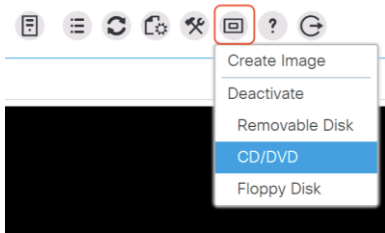
7. Click Choose File, browse for the Cisco custom ESXi ISO installer file for HyperFlex nodes, and click Open.

8. Click Map Drive.
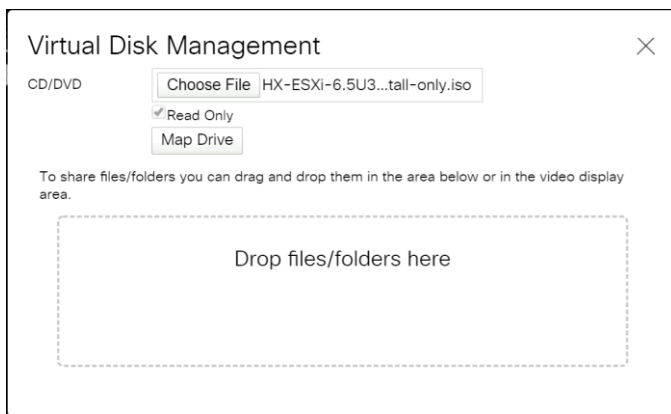
9. Repeat steps 1-8 for all the new compute-only nodes.

10. In the remote KVM tab, click the Server Actions button in the top right-hand corner of the screen, the click Reset.

11. Click OK.

12. Choose the Power Cycle option, then click OK.

13. Click OK.

14. Observe the server going through the POST process until the following screen is seen. When it appears, press the F6 key to enter into the boot device selection menu.
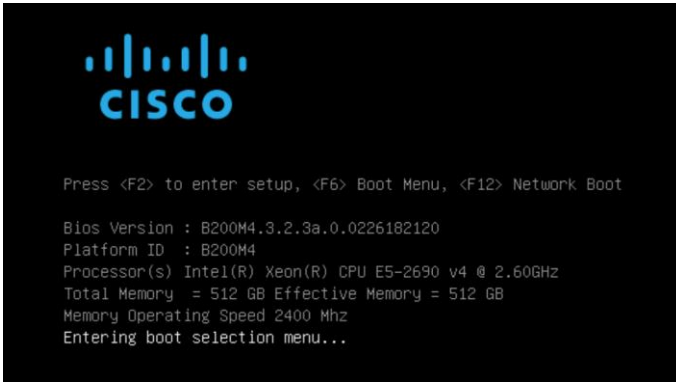
15. Choose Cisco vKVM-mapped vDVD1.22, then press Enter.

16. The server will boot from the remote KVM mapped ESXi ISO installer and display the following screen:

17. Choose the appropriate installation option for the compute-only node you are installing, either installing to SD cards, local disks, or booting from SAN, then press Enter.

```
              HyperFlex ESXi Installer - 6.7 U3 (Build 15160138)

    Select an Install Option:
          HyperFlex Converged Node - HX PIDs Only (DO NOT USE FOR UPGRADE)
          Compute-Only Node - Install to SD Cards/M.2 (DO NOT USE FOR UPGRADE)
          Compute-Only Node - Install to Local Disk (DO NOT USE FOR UPGRADE)
          Compute-Only Node - Install to Remote Disk (SAN) (DO NOT USE FOR UPGR
          Fully Interactive Install (DO NOT USE FOR UPGRADE)

    Exit and boot according to BIOS

    Press F12 for Full Help

This option SHOULD NOT be used for upgrades, see help section for details.

Select this option for compute-only nodes and only if you wish to install ESXi to SD card
(M3/M4/M5) or M.2 SSD (M5 Only).

Enter "ERASE" in all CAPS to confirm and agree to start the installation.
This is a DESTRUCTIVE process and cannot be reversed. Ensure a re-image is required.
```
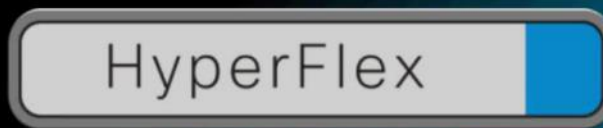
18. Type "ERASE" in all capital letters and press Enter to accept the warning and continue the installation.

19. The ESXi installer will now automatically perform the installation to the boot media. As you watch the process, some errors may be seen, but they can be ignored. Once the new server has completed the ESXi installation, it will be waiting at the console status screen shown below.

```
VMware ESXi 6.7.0 (VMKernel Release Build 15160138)

Cisco Systems Inc UCSB-B200-M5

2 x Intel(R) Xeon(R) Gold 6230 CPU @ 2.10GHz
766.6 GiB Memory




To manage this host go to:
http://co-1/
http://10.10.30.136/ (STATIC)
http://[fe80::250:56ff:fe63:4ef8]/ (STATIC)




<F2> Customize System/View Logs                          <F12> Shut Down/Restart
```
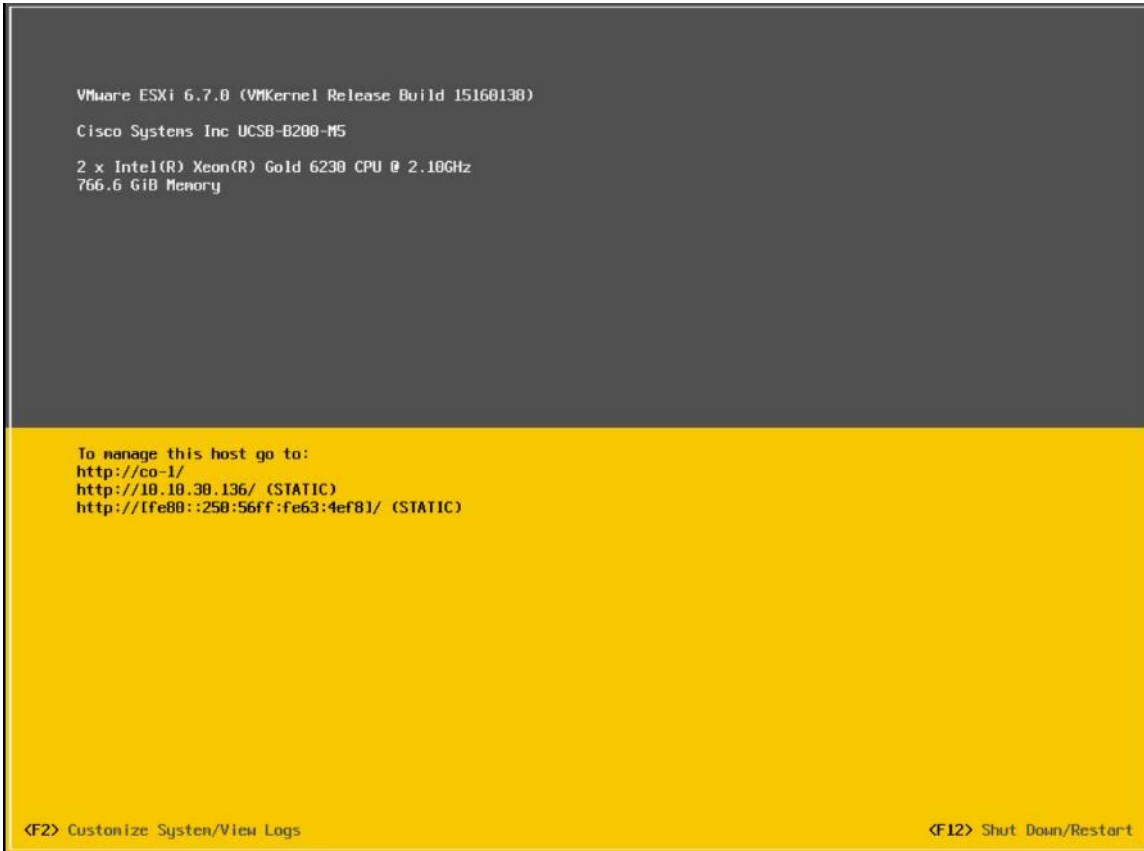
20. Repeat steps 1-19 for all the additional new compute-only nodes being added to the HX cluster.

21. When all the new nodes have finished their fresh ESXi installations, return to the HX installer, where the error in step 12 was seen. Click Retry UCSM Configuration.

22. The HX installer will now proceed to complete the deployment and perform all the steps listed at the top of the screen along with their status.

23. When the expansion is completed, a summary screen showing the status of the expanded cluster and the expansion operation is shown.

Model	Serial Number	Management Hypervisor	Management Storage Controller	Data Network Hypervisor	Data Network Storage Controller
HXAF220C-M5SX	WZP2212137B	10.10.30.128	10.10.30.200	192.168.11.128	192.168.11.200
HXAF220C-M5SX	WZP22120C5N	10.10.30.129	10.10.30.201	192.168.11.129	192.168.11.201
HXAF220C-M5SX	WZP2212136N	10.10.30.130	10.10.30.202	192.168.11.130	192.168.11.202
HXAF220C-M5SX	WZP2211152Y	10.10.30.131	10.10.30.203	192.168.11.131	192.168.11.203
HXAF220C-M5SX	WZP22120C8G	10.10.30.132	10.10.30.204	192.168.11.132	192.168.11.204
HXAF220C-M5SX	WZP22120C6B	10.10.30.133	10.10.30.205	192.168.11.133	192.168.11.205
HXAF220C-M5SX	WZP22120C8J	10.10.30.134	10.10.30.206	192.168.11.134	192.168.11.206
HXAF220C-M5SX	WZP22050FK7	10.10.30.135	10.10.30.207	192.168.11.135	192.168.11.207
UCSB-B200-M5	FOX1926G3BA	10.10.30.136		192.168.11.136	
UCSB-B200-M5	FOX1926G3BA	10.10.30.137		192.168.11.137	
UCSB-B200-M5	FOX1926G3BA	10.10.30.138		192.168.11.138	
UCSB-B200-M5	FOX1926G3BA	10.10.30.139		192.168.11.139	
UCSB-B200-M5	FOX1926G3BA	10.10.30.140		192.168.11.140	

After the install has completed, the new compute-only node is added to the cluster and it will have mounted the existing HyperFlex datastores, however the new node still requires some post installation steps in order to be consistent with the configuration of the existing nodes. For example, the new compute-only node will not have a vMotion vmkernel interface, and it may not have all of the guest VM networks configured. The easiest method to make the changes is to use the post_install script, choosing option 2 to configure an Expanded cluster, or the configuration can be done manually.

A list of additional configuration steps necessary includes:

- Disable SSH warning
- Creation of the guest VM port groups
- Creation of the vMotion vmkernel port
- Syslog Server Configuration

If at a later time the post_install script needs to be run against a specific HX cluster, the cluster can be specified by using the --cluster-ip switch and entering the cluster's management IP address.

To validate the configuration, vMotion a VM to the new compute-only node. You can validate that your VM is now running on the compute-only node through the Summary tab of the VM.

Management

Customers have two choices for managing their HyperFlex System: HyperFlex Connect or cloud-based Intersight. Both methods are described in the following sections.

HyperFlex Connect

HyperFlex Connect is the easy to use, and powerful primary management tool for HyperFlex clusters. HyperFlex Connect is an HTML5 web-based GUI tool which runs on all of the HX nodes and is accessible via the cluster management IP address.

Local Access

Logging into HyperFlex Connect can be done using pre-defined local accounts. The default predefined administrative account is named "admin". The password for the default admin account is set during the cluster creation as the cluster password. Using local access is only recommended when vCenter direct or SSO credentials are not available.

Role-Based Access Control

HyperFlex Connect provides Role-Based Access Control (RBAC) via integrated authentication with the vCenter Server managing the HyperFlex cluster. You can have two levels of rights and permissions within the HyperFlex cluster:

- Administrator: Users with administrator rights in the managing vCenter server will have read and modify rights within HyperFlex Connect. These users can make changes to the cluster settings and configuration.

- Read-Only: Users with read-only rights in the managing vCenter server will have read rights within HyperFlex Connect. These users cannot make changes to the cluster settings and configuration.

Users can log into HyperFlex Connect using direct vCenter credentials, for example, administrator@vsphere.local, or using vCenter Single Sign-On (SSO) credentials such as an Active Directory user, for example, domain\user. Creation and management of RBAC users and rights must be done via the vCenter Web Client or vCenter 6.5 HTML5 vSphere Client.

To manage the HyperFlex cluster using HyperFlex Connect, follow these steps:

1. Using a web browser, open the HyperFlex cluster's management IP address via HTTPS.

2. Enter a local credential, such as local/root, or a vCenter RBAC credential for the username, and the corresponding password.

3. Click Login.

4. The Dashboard view will be shown after a successful login.
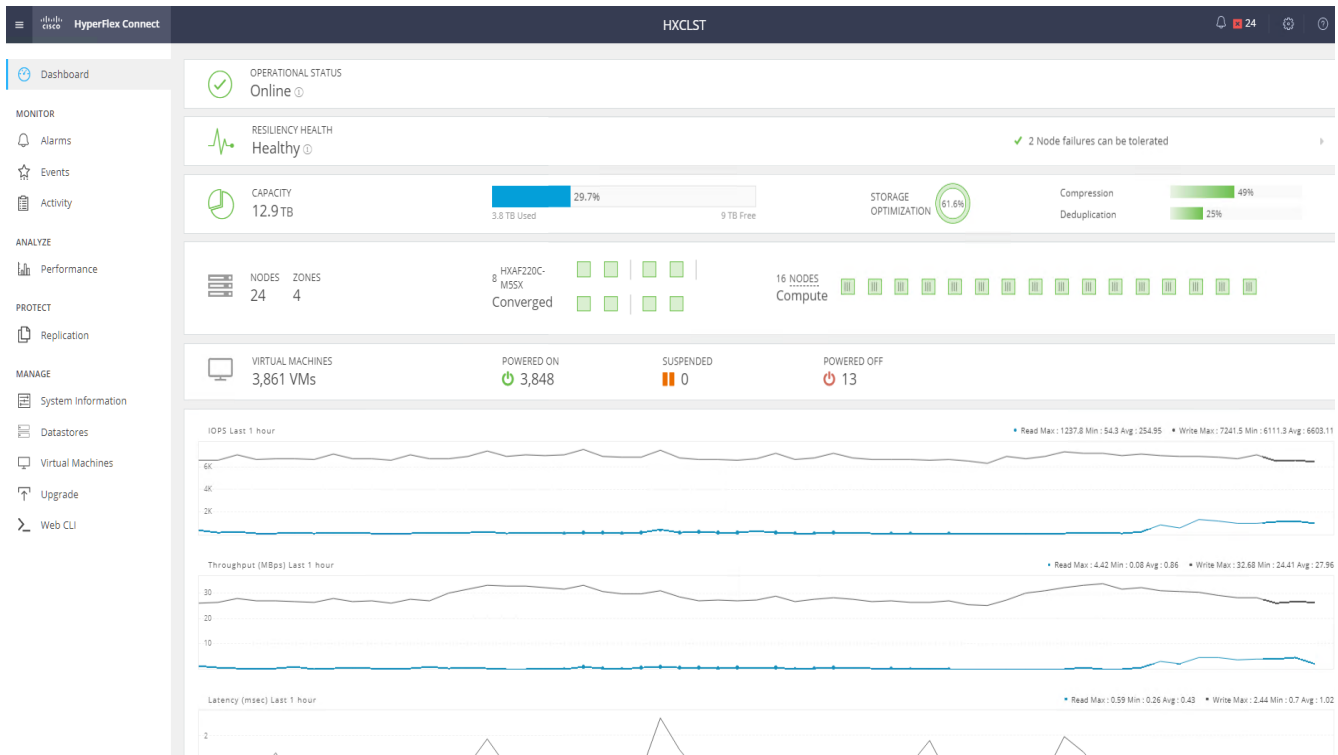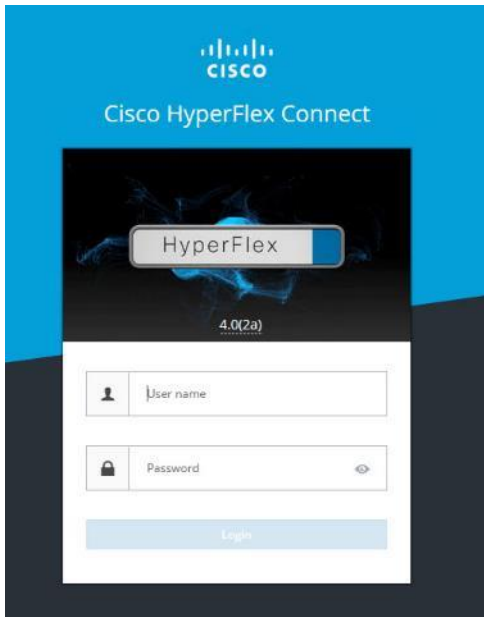
Dashboard

From the Dashboard view, several elements are presented:

- Cluster operational status, overall cluster health, and the cluster's current node failure tolerance.

- Cluster storage capacity used and free space, compression and deduplication savings, and overall cluster storage optimization statistics.

- Cluster size and individual node health.

- Cluster IOPs, storage throughput, and latency for the past 1 hour.

Monitor

HyperFlex Connect provides for additional monitoring capabilities, including:

- Alarms: Cluster alarms can be viewed, acknowledged, and reset.
- Event Log: The cluster event log can be viewed, specific events can be filtered for, and the log can be exported.
- Activity Log: Recent job activity, such as ReadyClones can be viewed and the status can be monitored.
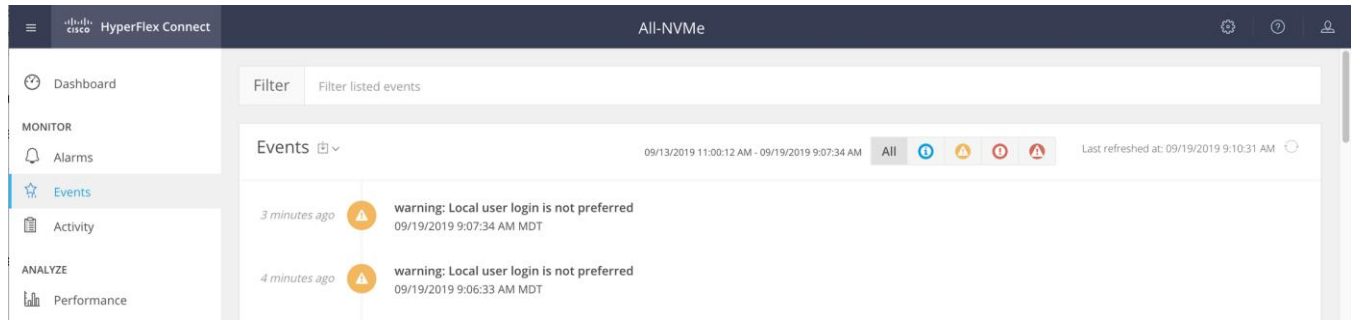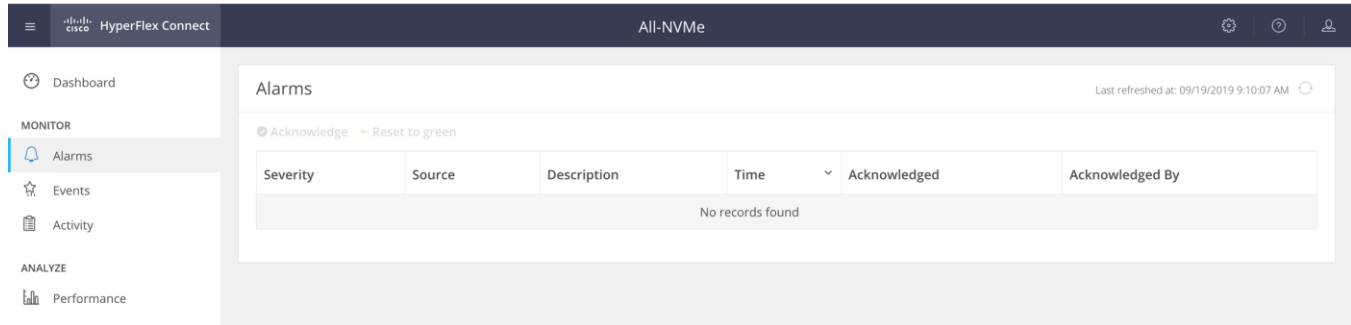
Analyze

The historical and current performance of the HyperFlex cluster can be analyzed via the built-in performance charts. The default view shows read and write IOPs, bandwidth, and latency over the past 1 hour for the entire cluster. Views can be customized to see individual nodes or datastores, and change the timeframe shown in the charts.

All-NVMe

Cluster ← Change Views   Change Intervals →   1h 1d 1w 1m 6m | Sep 19, 2019 - Sep 19, 2019   Last refreshed at: 09/19/2019 9:11:41 AM

8:19 AM   8:28 AM   8:36 AM   8:45 AM   8:53 AM   9:02 AM

IOPS 09/19/2019 8:11:00 AM - 09/19/2019 9:09:00 AM   09/19/2019 8:29:00 AM MDT   • Read Max: 0 Min:0 Avg: 0   • Write Max: 14.4 Min:6.6 Avg: 13.02   • Aggregated Max: 14.4 Min:6.6 Avg: 13.02

Throughput (MBps) 09/19/2019 8:11:00 AM - 09/19/2019 9:09:00 AM   • Read Max: 0 Min:0 Avg: 0   • Write Max: 0.05 Min:0.02 Avg: 0.04   • Aggregated Max: 0.05 Min:0.02 Avg: 0.04

Latency (msec) 09/19/2019 8:11:00 AM - 09/19/2019 9:09:00 AM   • Read Max: 0 Min:0 Avg: 0   • Write Max: 0.54 Min:0.34 Avg: 0.37   • Average Max: 0.54 Min:0.34 Avg: 0.37

	• Write	• Read
IOPS	13.8	No Data Available
Throughput (MBps)	0.047	No Data Available
Latency (msec)	0.355	No Data Available

Protect

HyperFlex Connect is used as the management tool for all configuration of HyperFlex Data Protection features, including VM replication and data-at-rest encryption.

Manage

HyperFlex Connect presents several views and elements for managing the HyperFlex cluster:

- System Information: Presents a detailed view of the cluster configuration, software revisions, hosts, disks, and cluster uptime. Support bundles can be generated to be shared with Cisco TAC when technical support is needed. Views of the individual nodes and the individual disks are available. In these views, nodes can be placed into HX Maintenance Mode, and self-encrypting disks can be securely erased.

- Datastores: Presents the datastores present in the cluster, and allows for datastores to be created, mounted, unmounted, edited or deleted, as described earlier in this document as part of the cluster setup.

- Virtual Machines: Presents the VMs present in the cluster and allows for the VMs to be powered on or off, cloned via HX ReadyClone, Snapshots taken, and protected via native replication.

- Upgrade: One-click upgrades to the HXDP software, ESXi host software and Cisco UCS firmware can be initiated from this view.

- Web CLI: A web-based interface, from which CLI commands can be issued and their output seen, as opposed to directly logging into the SCVMs via SSH.

Cisco Intersight Cloud-Based Management

Cisco Intersight management is enabled via embedded code running on the Cisco UCS Fabric Interconnects, and in the Cisco HyperFlex software, known as device connectors. To enable Intersight management, the device connectors are registered online at the Cisco Intersight website, https://intersight.com when logged into the website with a valid cisco.com account used to manage your environments. Cisco Intersight can be used to manage and monitor HyperFlex clusters and UCS domains with the following software revisions:

- Cisco UCS Manager and Infrastructure Firmware version 3.2 and later
- Cisco HyperFlex software version 2.5(1a) or later

The Cisco UCS Fabric Interconnects, and the Cisco HyperFlex nodes must have DNS lookup capabilities and access to the internet. If direct access to the internet is not available, the systems can be configured to connect via an HTTPS proxy server.

Cisco Intersight Licensing

Cisco Intersight is offered in two editions; a Base license which is free to use, and offers a large variety of monitoring, inventory and reporting features, and an added cost Essentials license, which adds advanced monitoring, server policy and profile configuration, firmware management, virtual KVM features, and more. New features and capabilities will be added to the different licensing tiers over time. A 90-day trial of the Essentials license is available for use as an evaluation period.

To configure Cisco Intersight licensing, follow these steps:

1. Using a web browser, log into the Cisco Intersight webpage at https://intersight.com (you must have a valid cisco.com CCO account).

2. In the Dashboards view, click the gear shaped icon in the upper right-hand corner, then click Licensing.
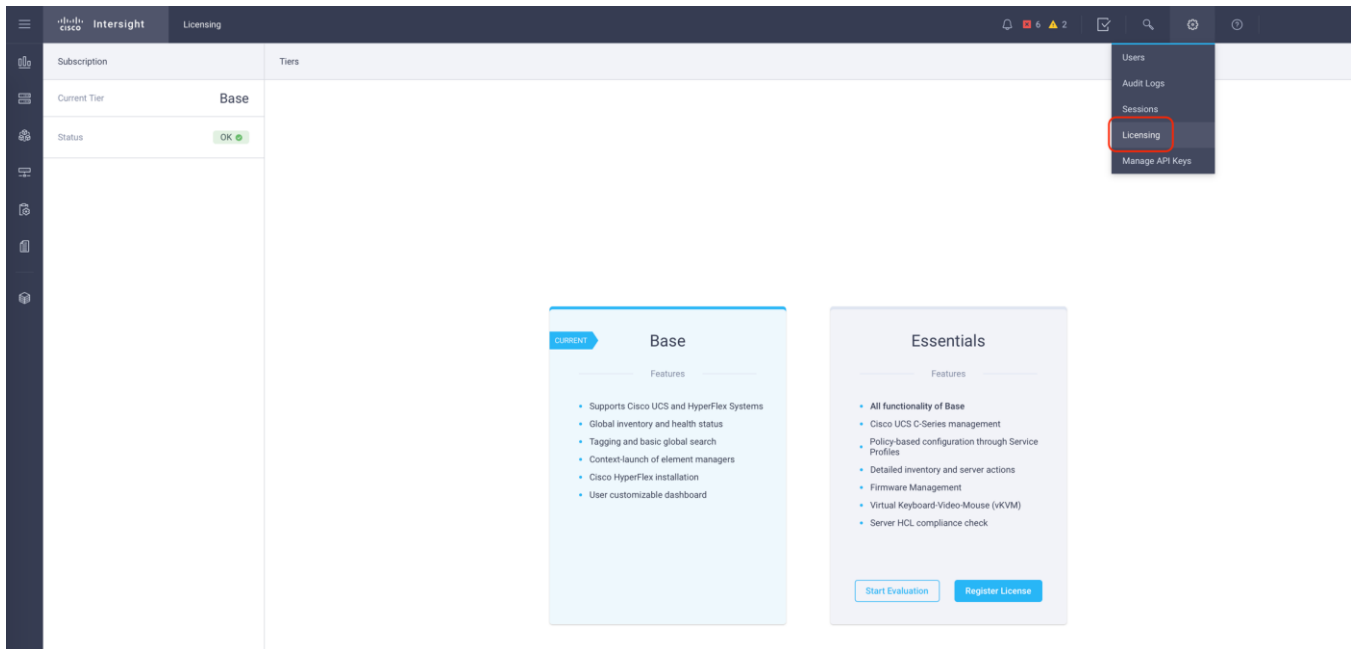
3. Click Start Evaluation to begin a 90-day Essentials license trial or click Register License.

Cisco Intersight HyperFlex Management

To connect Cisco Intersight to the Cisco HyperFlex cluster(s), and the Cisco UCS Domain(s) in your environments, follow the steps in this section.

Connect Cisco UCS Manager

To connect to Cisco UCS Manager (UCSM), follow these steps:

1. Using a web browser, log into the Cisco UCS Manager webpage.

2. From a second browser window or tab, log into the Cisco Intersight webpage at https://intersight.com (you must have a valid cisco.com CCO account).
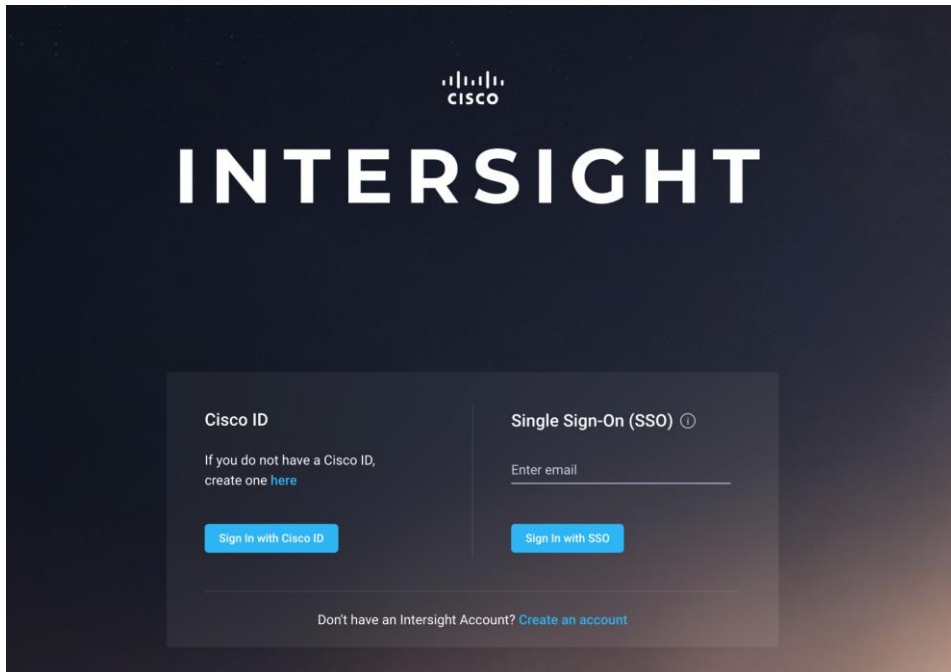
3. Click Devices.
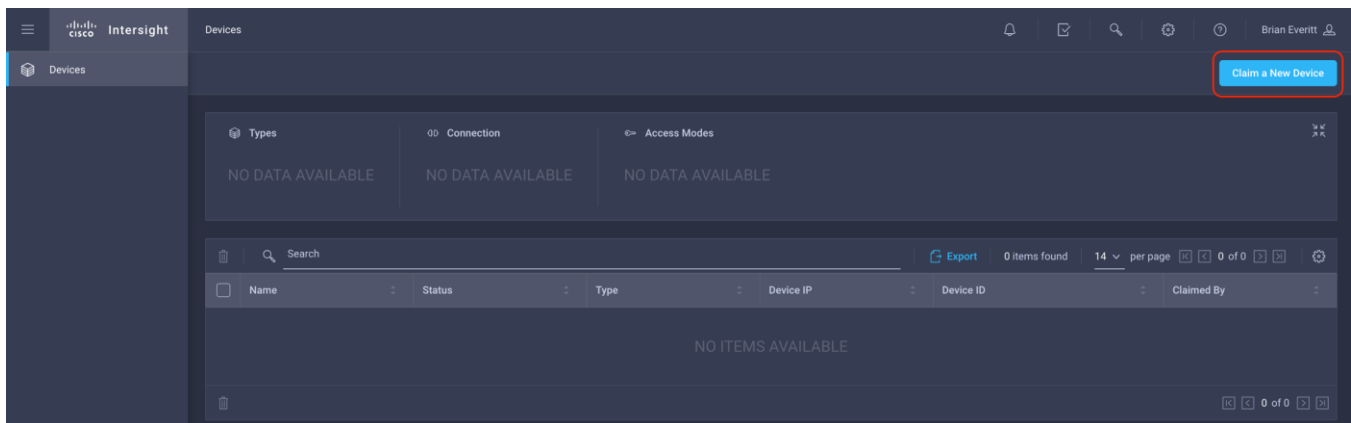
4. Click Claim A New Device.

5. In Cisco UCS Manager, click Admin.

6. In the Admin tree, click Device Connector.

7. If necessary, click HTTPS Proxy Settings and then click Manual. Enter the Proxy server IP address or DNS hostname, the TCP port, enable authentication then enter a username and password if necessary, then click Save.

8. If desired, the Access Mode for Cisco Intersight can be set to Read-only, or management can be disabled from this screen.

9. In the main Cisco UCS Manager screen, you will see a Device ID and a Claim Code for this Cisco UCS Domain. Copy these two codes by clicking on the clipboard icons and pasting them to the Device ID and Claim Code fields in the Cisco Intersight "Claim A New Device" window, then click Claim.
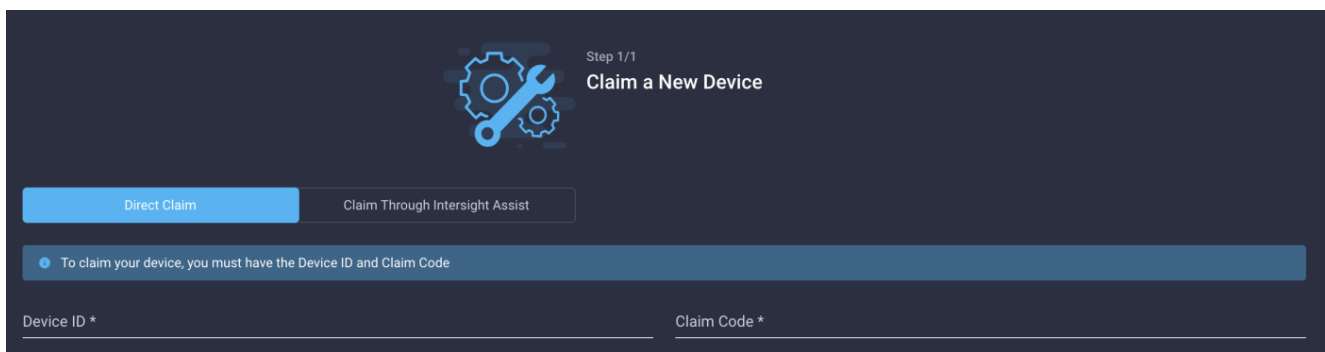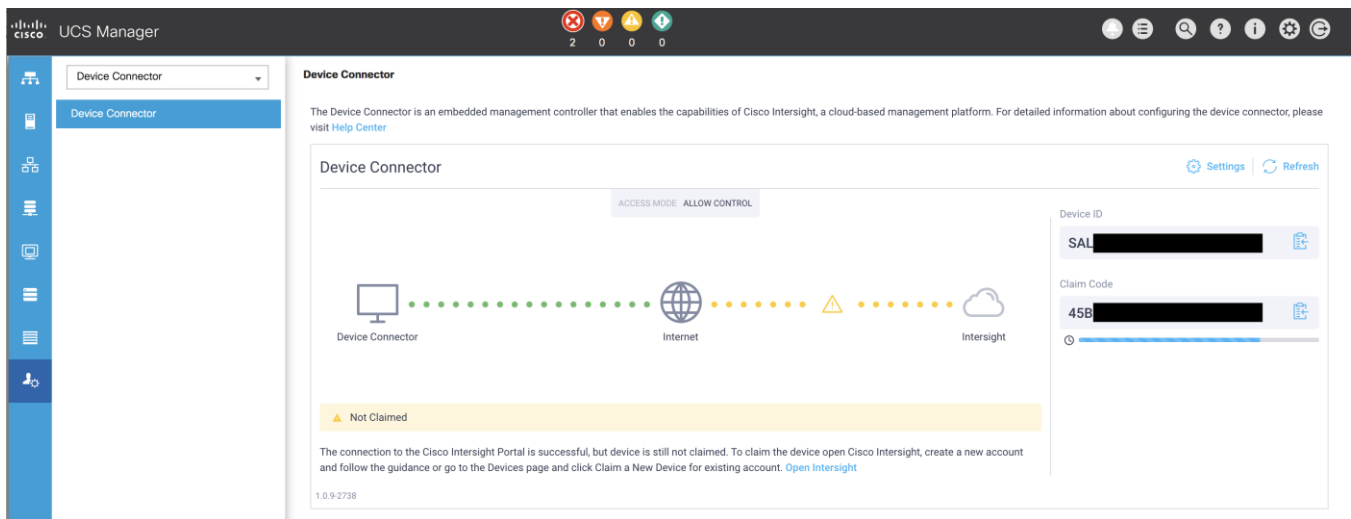
The Cisco UCS Domain will now show the system as Claimed in the Device Connector screen.

Connect Cisco HyperFlex Clusters

To connect Cisco HyperFlex Clusters, follow these steps:

1. Use a web browser to open the HX Connect webpage at the cluster's management IP address, for example: https://10.29.132.182/

2. Enter a local credential or a vCenter RBAC credential for the username and the corresponding password.

3. Click Login.

4. From a second browser window or tab, log into the Cisco Intersight webpage at https://intersight.com (you must have a valid cisco.com CCO account).

5. In the left-hand navigation buttons, click Devices.

6. Click Claim A New Device.

7. In the HyperFlex Connect Dashboard page, click Edit Settings in the top right-hand corner, then click Device Connector.
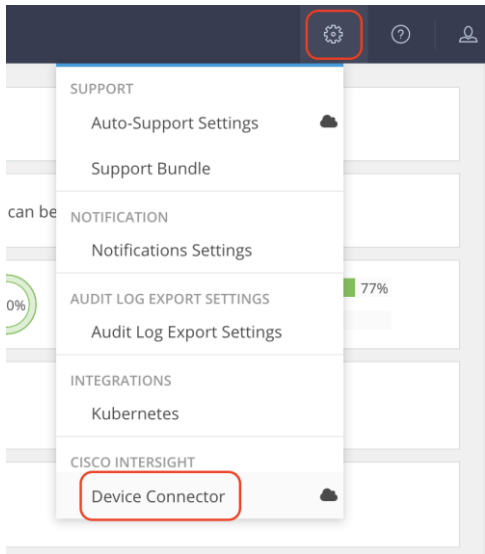
8. If necessary, to modify the Proxy settings, click the Settings button, and click the Proxy Settings link on the left-hand side. Enable the Proxy configuration button, then enter the Proxy server IP address or DNS host-name, the TCP port, enable authentication then enter a username and password if necessary, then click Save.

9. If desired, the Access Mode for Cisco Intersight can be set to Read-only, or management can be disabled from this screen.

10. In the HyperFlex Connect screen, a Device ID, and a Claim Code for this HyperFlex cluster will be shown. Copy these two codes by clicking on the clipboard icons and pasting them to the Device ID and Claim Code fields in the Cisco Intersight "Claim A New Device" window, then click Claim.
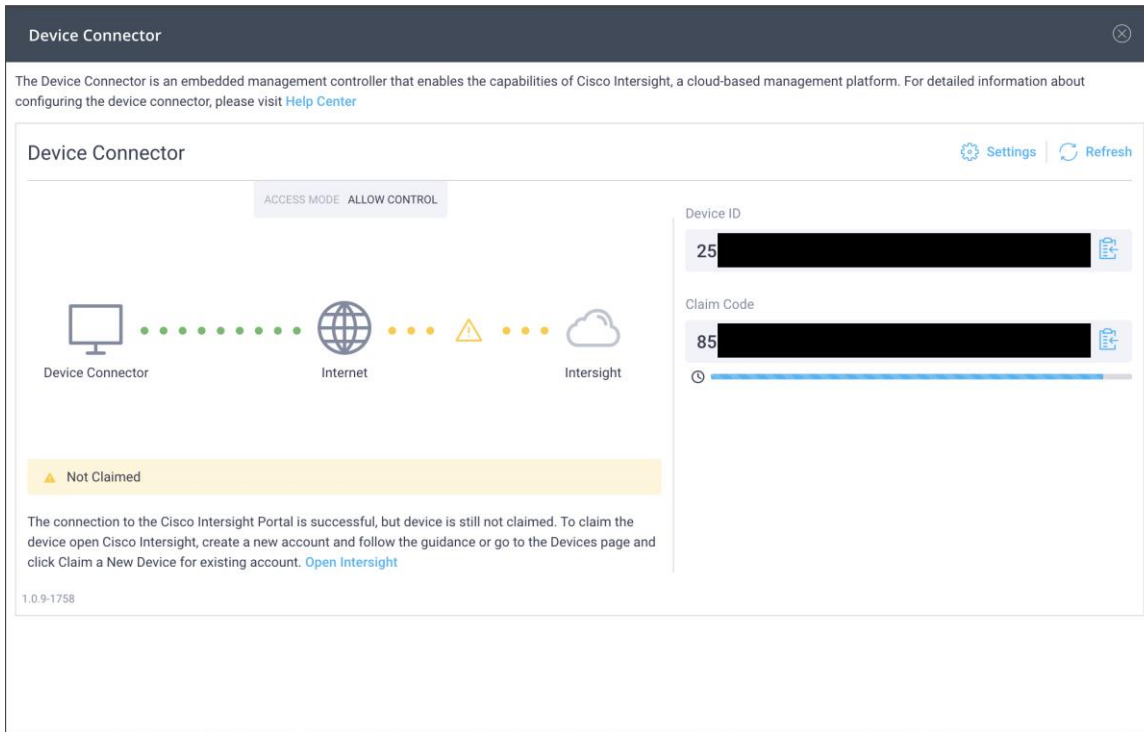
11. The Cisco HyperFlex Cluster will now show the system as Claimed in the Device Connector screen.

Dashboard

The Cisco Intersight Dashboard provides a single screen overview of all connected Cisco UCS Domains, the servers within those domains, the HyperFlex Clusters running in the domains, along with their health statuses, storage utilization, port counts, and more. Elements on the screen are clickable and will drill down into other sections of the page to view further details.

Servers

The Servers screen provides details of all the individual servers within the connected and managed UCS domains.

HyperFlex Clusters

The HyperFlex Clusters screen provides details of all the HyperFlex clusters that are connected and managed by Cisco Intersight. By clicking the ellipses (...) the HyperFlex Connect GUI for the clusters can be directly connected to in another browser window or tab.
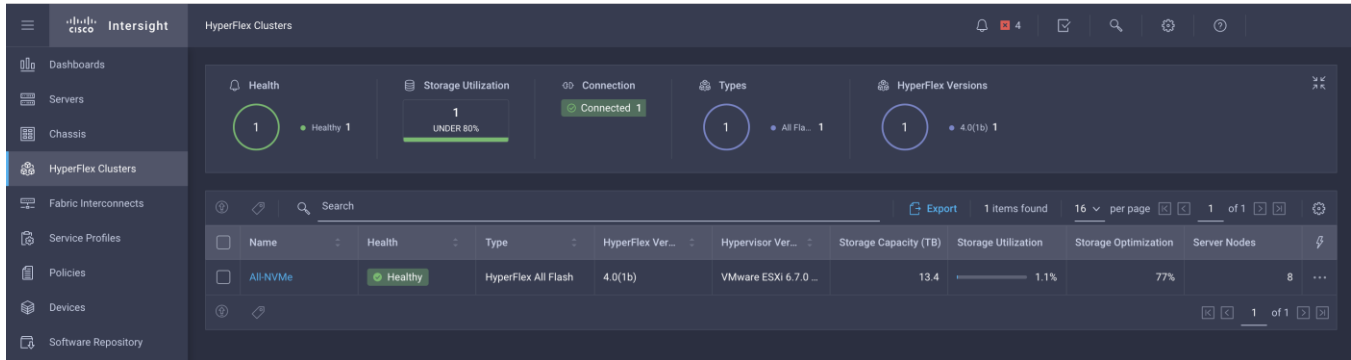
Fabric Interconnects

The Fabric Interconnects screen provides details of all the UCS domains that are connected and managed by Cisco Intersight. By clicking the ellipses (...) the Cisco UCS Manager webpage for the domain can be directly connected to in another browser window or tab, or a session can be opened to the CLI of the Fabric Interconnect.
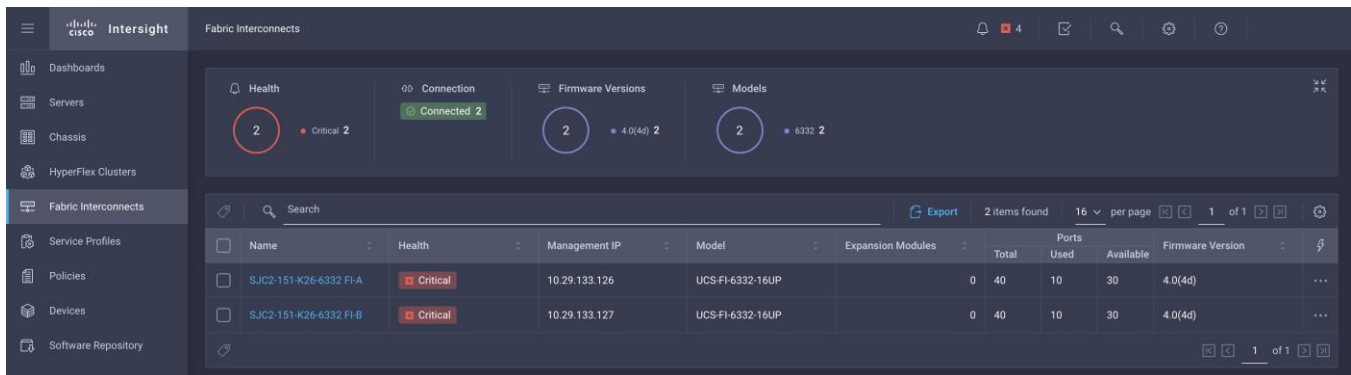
Profiles and Policies

Cisco Intersight Service Profiles and Policies pages are only available with Intersight Essentials licensing, except for configuring a Cisco HyperFlex Cluster Profile as outlined earlier in this document.

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 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 via 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.

Figure 49 HyperFlex Management – SENTINEL Snapshot
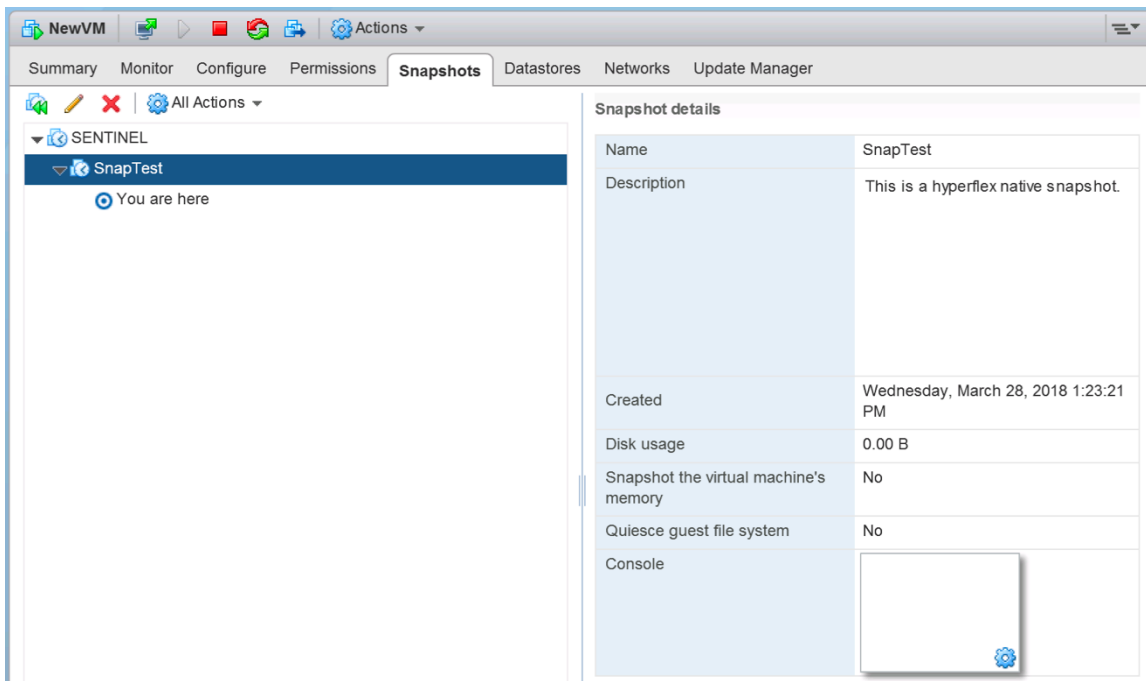
- If large numbers of scheduled snapshots need to be taken, distribute the time of the snapshots taken by placing the VMs into multiple folders or resource pools. For example, schedule two resource groups, each with several VMs, to take snapshots separated by 15-minute intervals in the scheduler window. Snapshots will be processed in batches of 8 at a time, until the scheduled task is completed.

Figure 50 HyperFlex Management – Schedule Snapshots
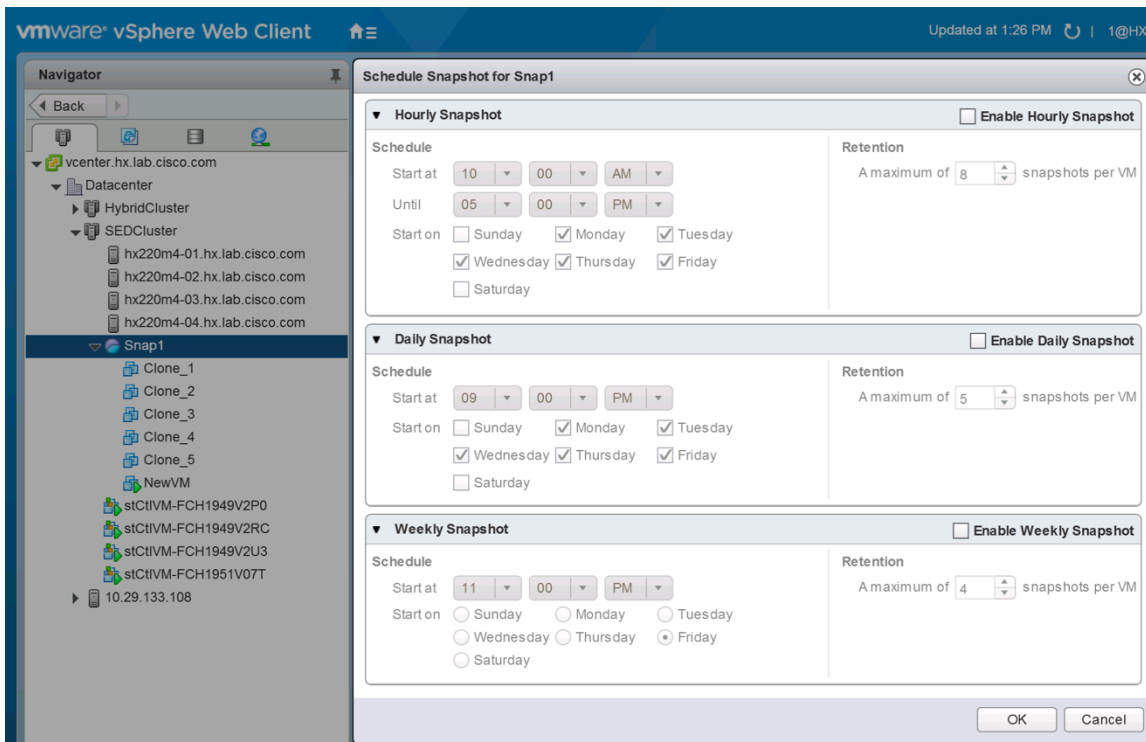
Storage vMotion

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 a storage vMotion of virtual machine disk files has little value in the HyperFlex system. Furthermore, storage vMotion can create additional filesystem consumption and generate additional unnecessary metadata within the filesystem, which must later be cleaned up via the filesystem's internal cleaner process.

> ⚠ It is recommended not to perform a storage vMotion of a guest VM between datastores within the same HyperFlex cluster. Storage vMotion 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, can cause ReadyClone and Snapshot errors, and lead to degraded performance in stretched clusters.

Maintenance Mode

Cisco HyperFlex clusters which have been originally installed using HXDP version 4.0(1b) or later no longer require the use of "HX Maintenance Mode" in order to evacuate the converged nodes for reboots, patches, or other work. Use of the standard enter/exit maintenance mode available in the vCenter web client or HTML5 web client is sufficient. Clusters which are upgraded from earlier revisions to version 4.0(1b) or later can also use standard vSphere maintenance mode, after undergoing a process to remove vSphere ESX Agent Manager (EAM) components and settings that are no longer required. These instructions are available upon request from your Cisco sales team or technical support contacts.

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 Installation Checklist

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

1. 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.

2. Verify the expected cluster capacity is seen in the HX Connect Dashboard summary screen.

3. Create a test virtual machine that accesses the HyperFlex datastore and is able to perform read/write operations.

4. Perform a virtual machine migration (vMotion) of the test virtual machine to a different host on the cluster.

5. During the vMotion of the virtual machine, make sure the test virtual machine can perform a continuous ping to its 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 via 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 in the HX Connect Dashboard.

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 in the HX Connect Dashboard 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 Cisco UCS Manager, but all will be cleared after the FI comes back online.

Reinstall HX Cluster

If a Cisco HyperFlex cluster needs to be reinstalled, contact your local Cisco account or support team in order to be provided with a cluster cleanup guide. Note that the process will be destructive and result in the loss of all the VMs and all the data stored in the HyperFlex distributed filesystem.

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

1. Clean up the existing environment by:

   a. Delete the existing HX virtual machines and HX datastores.

   b. Destroy the HX cluster.

   c. Remove the HX cluster from vCenter.

   d. Remove the vCenter MOB entries for the HX extension.

   e. Delete the HX sub-organization and HX VLANs in Cisco UCS Manager.

2. Using the HX OVA-based installer VM, use the customized version of the installation workflow by selecting the "I know what I am doing" link.
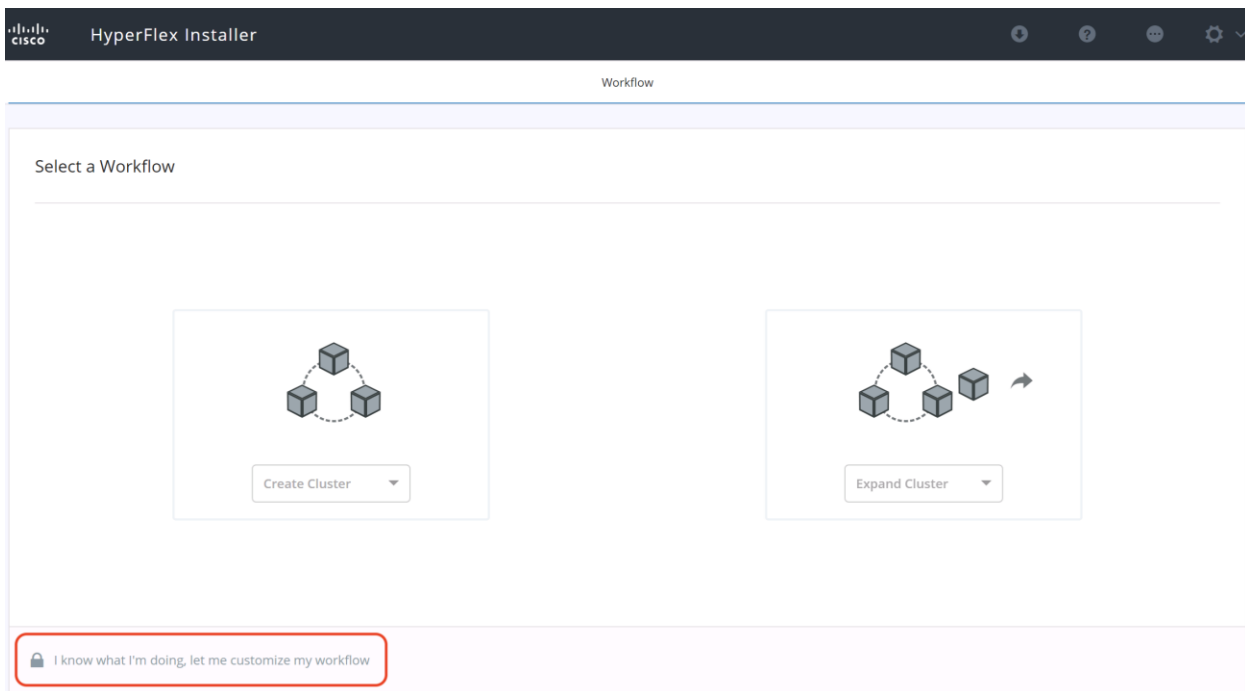
3. Use customized workflow and only choose the "Run UCS Manager Configuration" option, click Continue.
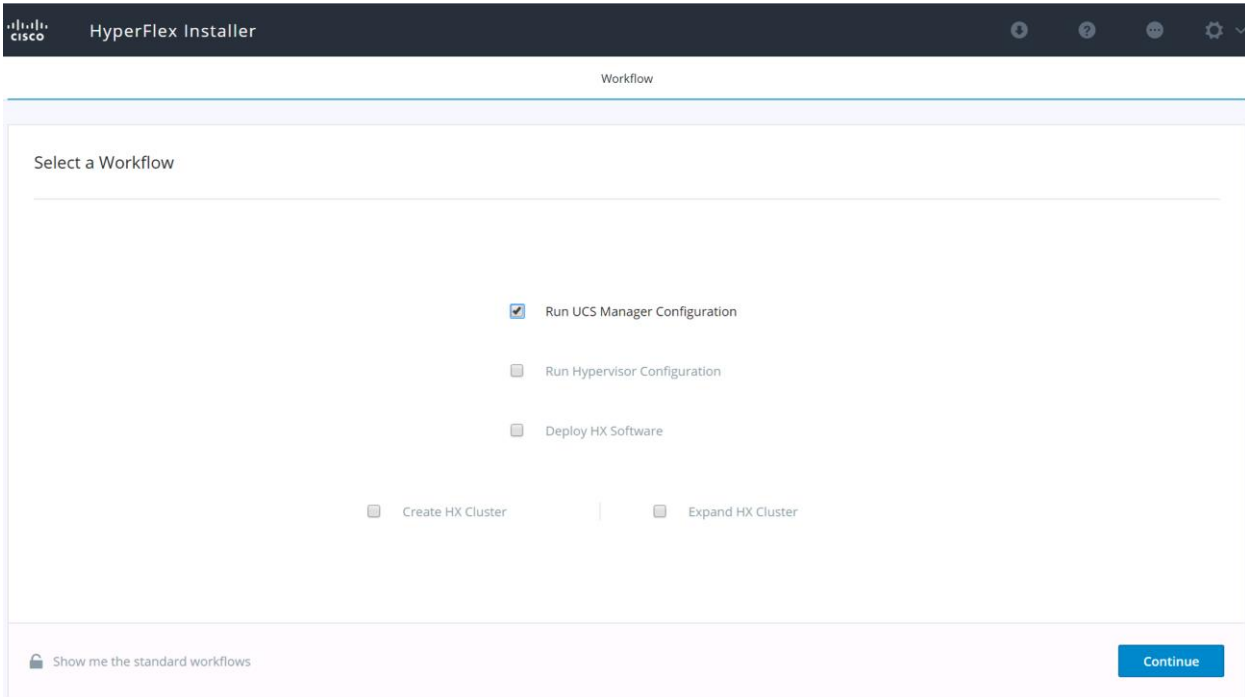
4. When the Cisco UCS Manager configuration is complete, HX hosts are associated with HX service profiles and powered on. Now perform a fresh ESXi installation using the custom ISO image and following the steps in section Cisco UCS vMedia and Boot Policies.

5. When the ESXi fresh installations are all finished, use the customized workflow, and choose the remaining 3 options; ESXi Configuration, Deploy HX Software, and Create HX Cluster, to continue and complete the HyperFlex cluster installation.
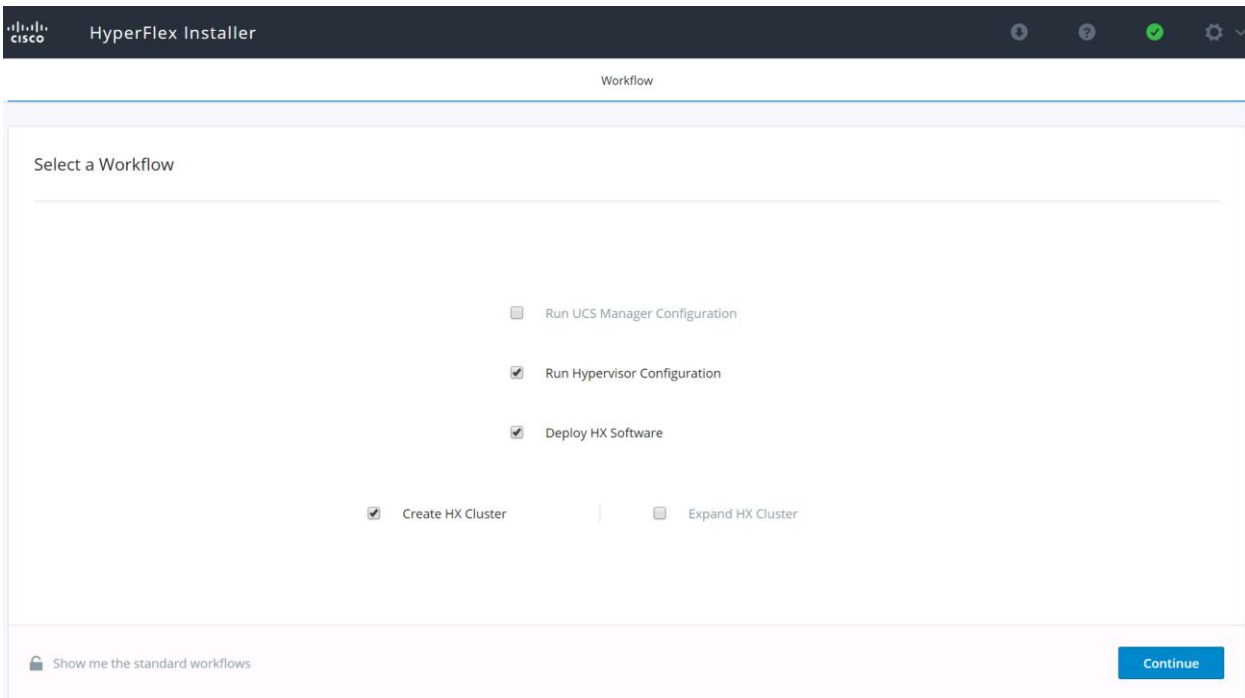
Build the Virtual Machines and Environment for Workload Testing

Software Infrastructure Configuration

This section details how to configure the software infrastructure components that comprise this solution. Install and configure the infrastructure virtual machines by following the process provided in .

Table 49   Test Infrastructure Virtual Machine Configuration

Configuration	VMware Virtual Desktops Horizon Administrator Virtual Machines	VMware Provisioning Servers Virtual Machines
Operating System	Microsoft Windows Server 2019	Microsoft Windows Server 2019
Virtual CPU amount	6	8
Memory amount	8 GB	12 GB
Network	VMNIC	Network
Disk-1 (OS) size and location	40 GB	Disk-1 (OS) size and location
Configuration	**Microsoft Active Directory DC's Virtual Machines**	**VMware Profile Servers Virtual Machines**
Operating system	Microsoft Windows Server 2019	Operating system
Virtual CPU amount	4	
Memory amount	4 GB	
Network	VMNIC	
Disk size and location	40 GB	
Configuration	**Microsoft SQL Server Virtual Machine**	**Additional Supporting Virtual Machine**
Operating system	Microsoft Windows Server 2019	Microsoft Windows Server 2019
Virtual CPU amount	8	4
Memory amount	16 GB	8 GB
Network	VMNIC	Network

Configuration	VMware Virtual Desktops Horizon Administrator Virtual Machines	VMware Provisioning Servers Virtual Machines
Disk-1 (OS) size and location	40 GB	Disk-1 (OS) size and location

Prepare the Master Images

This section details how to create the golden (or master) images for the environment. virtual machines for the master images must first be installed with the software components needed to build the golden images. For this CVD, the images contain the basics needed to run the Login VSI workload.

To prepare the master virtual machines for the Remote Desktop Hosted Sessions (RDSH) VDI Desktops there are three major steps to complete when the base virtual machine has been created:

- Installing OS
- Installing application software
- Installing the Virtual Delivery Agents (VDAs)

The master image HVD and RDSH virtual machines were configured as listed in Table 50.

Table 50   HVD and RDSH Configurations

Configuration	VDI	Configuration
Operating system	Microsoft Windows 10 64-bit	Microsoft Windows Server 2019
Virtual CPU amount	2	8
Memory amount	4.0 GB (reserved)	32 GB (reserved)
Network	VMNIC vm-network	VMNIC vm-network
VMware PVS vDisk size and location	40 GB	80 GB
Additional software used for testing	Microsoft Office 2016 Login VSI 4.1.32 (Knowledge Worker Workload)	Microsoft Office 2016 Login VSI 4.1.32 (Knowledge Worker Workload)

Horizon 7 Infrastructure Components Installation

This section details how to configure the software infrastructure components that comprise this solution.

The prerequisites for installing the view connection server, replica server(s) and composer server is to have Windows 2012, 2016 or 2019 virtual machines ready.

⚠ **In this study, we used Windows Server 2019 virtual machines for all Horizon infrastructure servers.**

Download the VMware Horizon 7 installation package from this link:

https://my.vmware.com/web/vmware/downloads/info/slug/desktop_end_user_computing/vmware_horizon/7_12

This following section provides a detailed, systematic installation process for Horizon 7 v7.12

Install Horizon Connection / Replica Servers

To install the Horizon Connection/Replica Servers, follow these steps:

1. Open view connection server installation, VMware-viewconnectionserver-x86_64-7.12.0-15770369.exe.

2. Click Next.
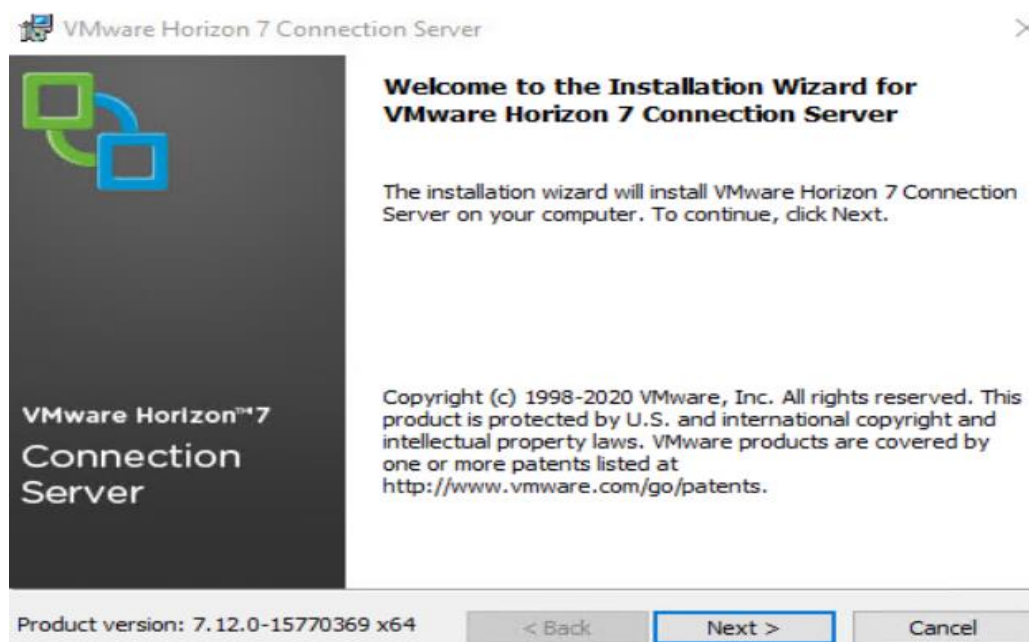
3. Accept the EULA then click Next.

4. Keep the default destination folder and click Next.
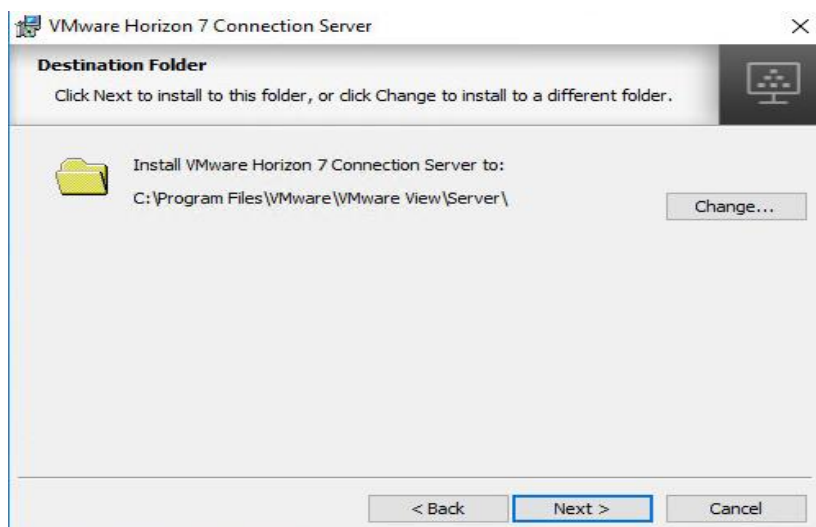
5. Select type of instance intended to install.

6. Select Standard Server instance for primary connection server installation.
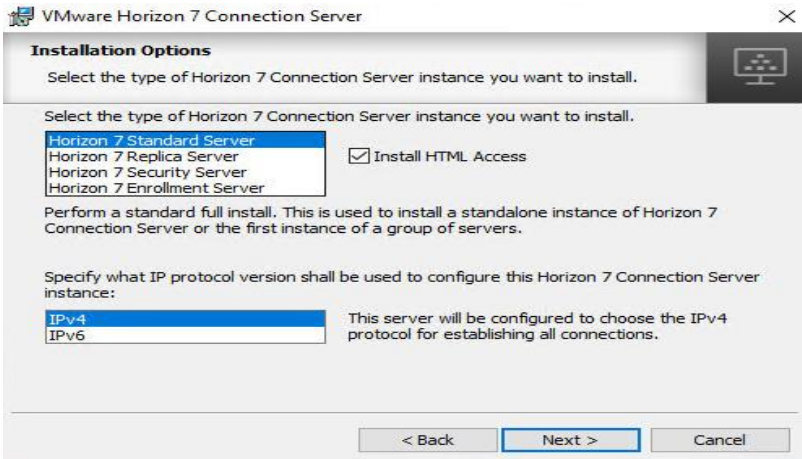
7. Select Replica server instance for fault tolerant connection server configuration after completion of Standard Server instance installation.
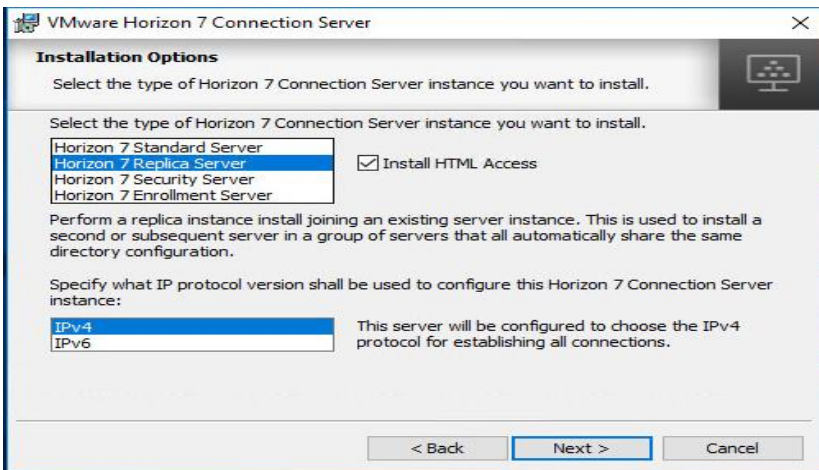
8. Enter the Data Recovery Password.
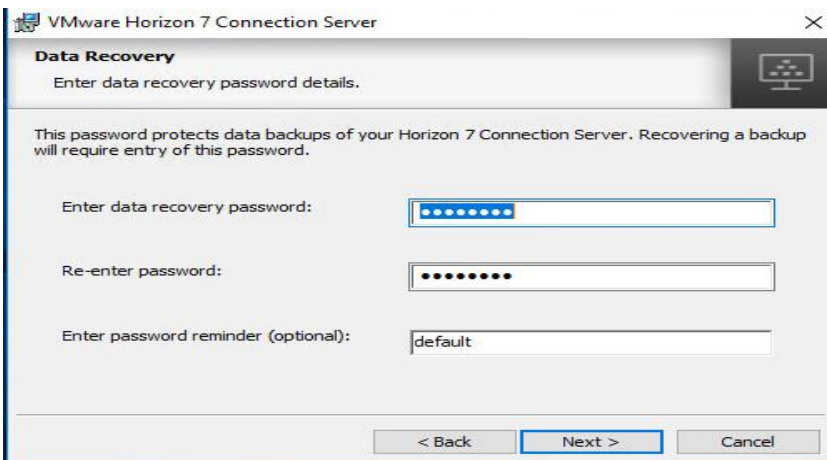
9. Click Next.
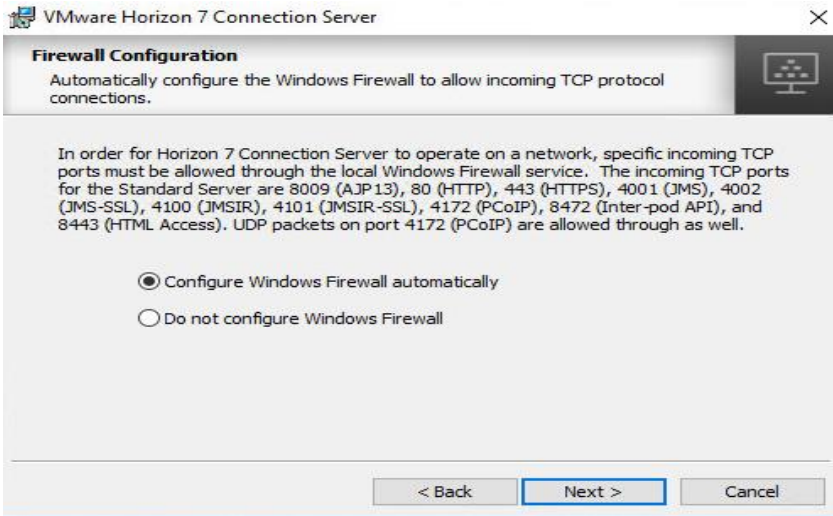
10. Select authorized users and group, click Next.
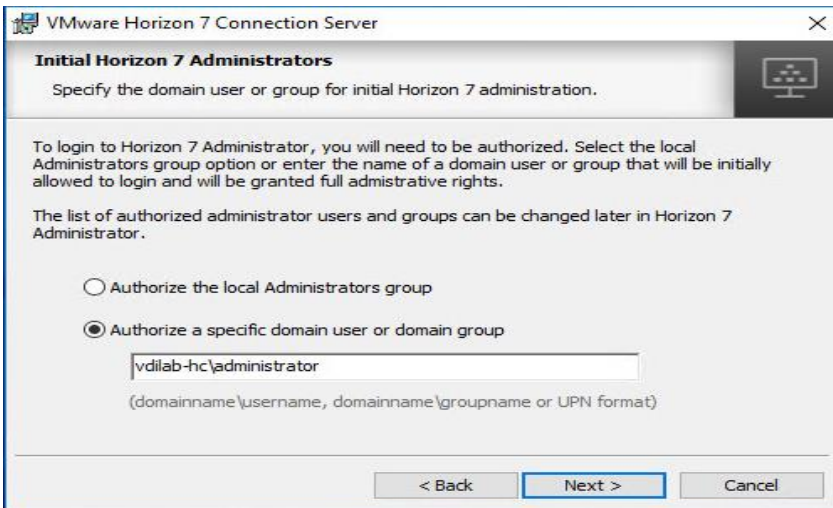
11. Enter domain credentials for previously specified domain user/group.
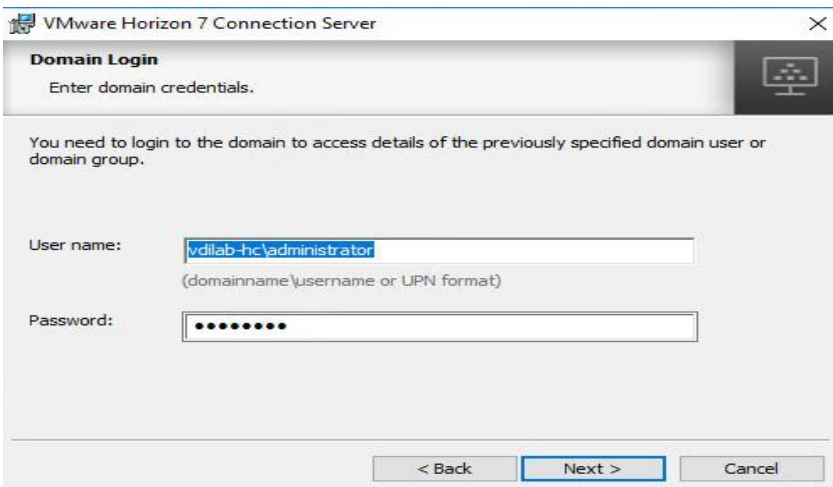
12. Opt-in or Opt-out of User Experience Improvement Program. Click Next.
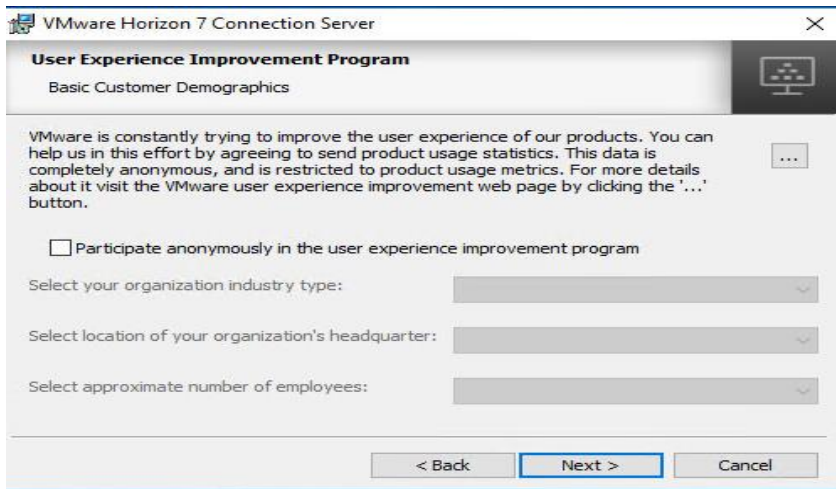
13. Click Install.
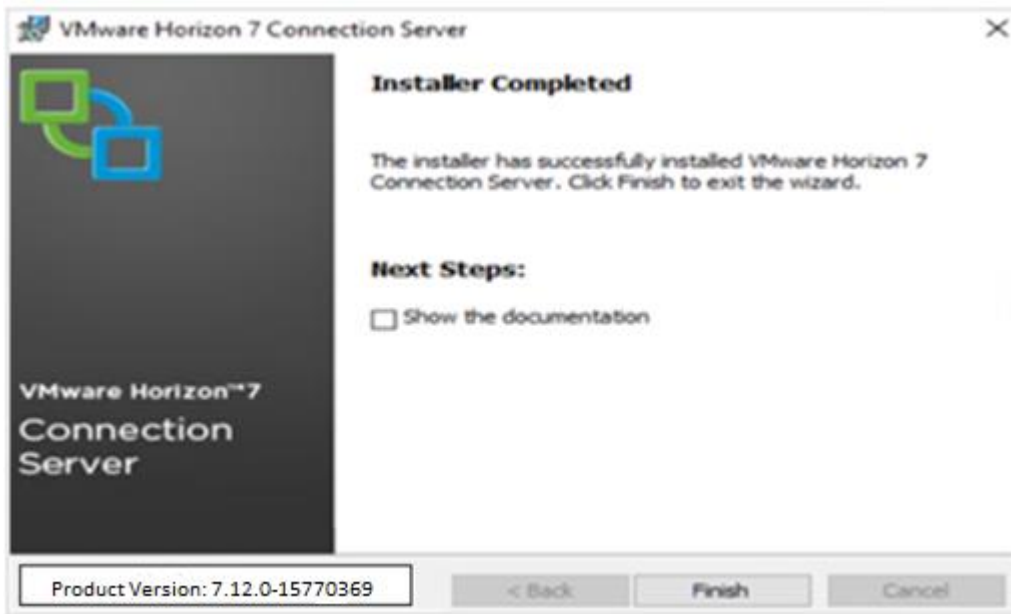
14. Click Finish.

Create a Microsoft Management Console Certificate Request

To generate a Horizon View SSL certificate request, use the Microsoft Management Console (MMC) Certificates snap-in:

https://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2068666

Configure the Horizon 7 Environment

To configure the Horizon 7 environment, follow this step:

1. Open WebUI, Login to https://<Horizon_Connection_ server_Management_IP_Address>/admin.
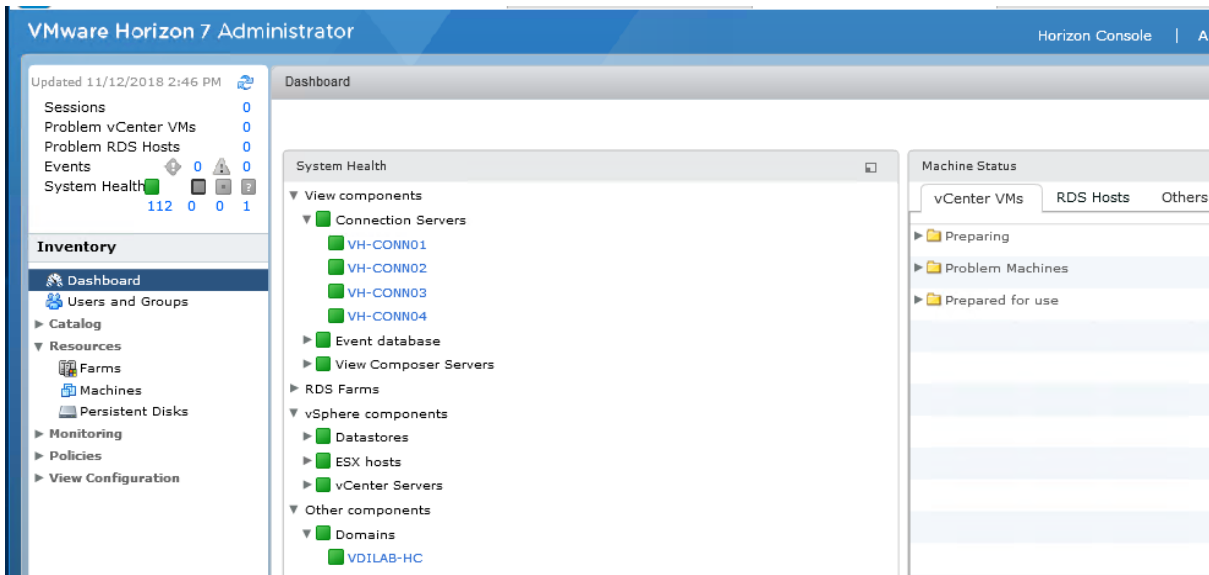
Configure Event Database

To configure the Event Database, follow these steps:

1. Configure the Event Database by adding Database Server, Database name, login credentials and prefix for the table from the Horizon 7 Administrator, View Configuration, Event Configuration node of the Inventory pane.

2. Click Edit in the action pane.
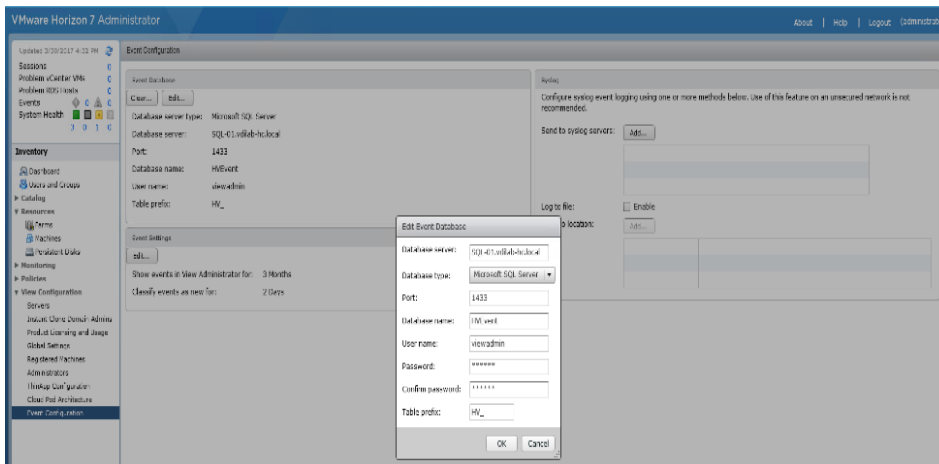
The details are shown below:

Configure Horizon 7 Licenses

To configure the Horizon 7 licenses, follow these steps:

1. Click View Configuration.

2. Select Product Licensing and Usage.

3. Click Edit License in the action pane.

4. Add the License Serial Number.

5. Click OK.

Configure vCenter

To configure the vCenter, follow these steps:

1. In View Configuration, Select Servers. Click Add vCenter Server tab.

2. Enter vCenter Server IP Address or FQDN, login credentials.

3. Advanced Settings options can be modified to change existing operations limit. Keep the advanced settings options as default.
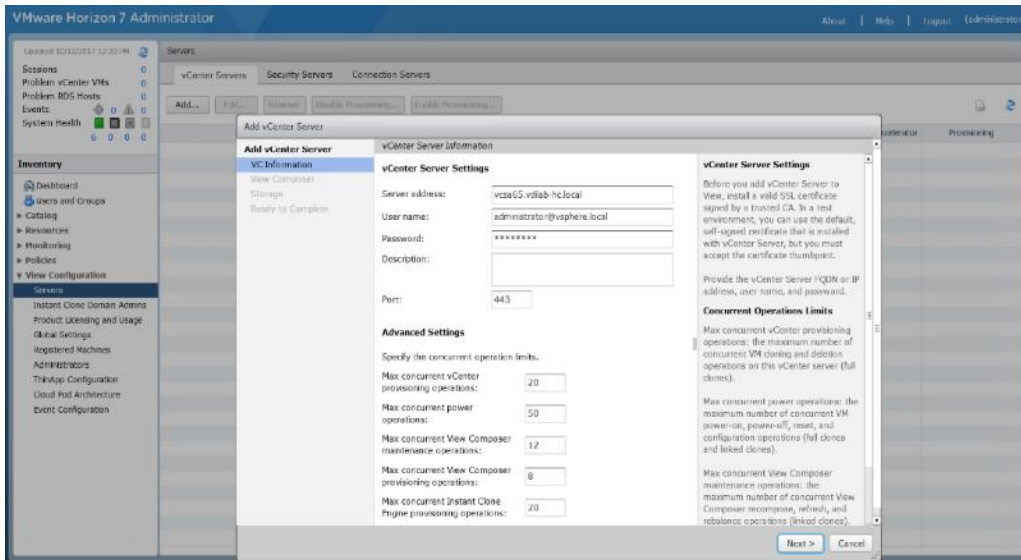
4. Click View certificate. Accept the certificate.
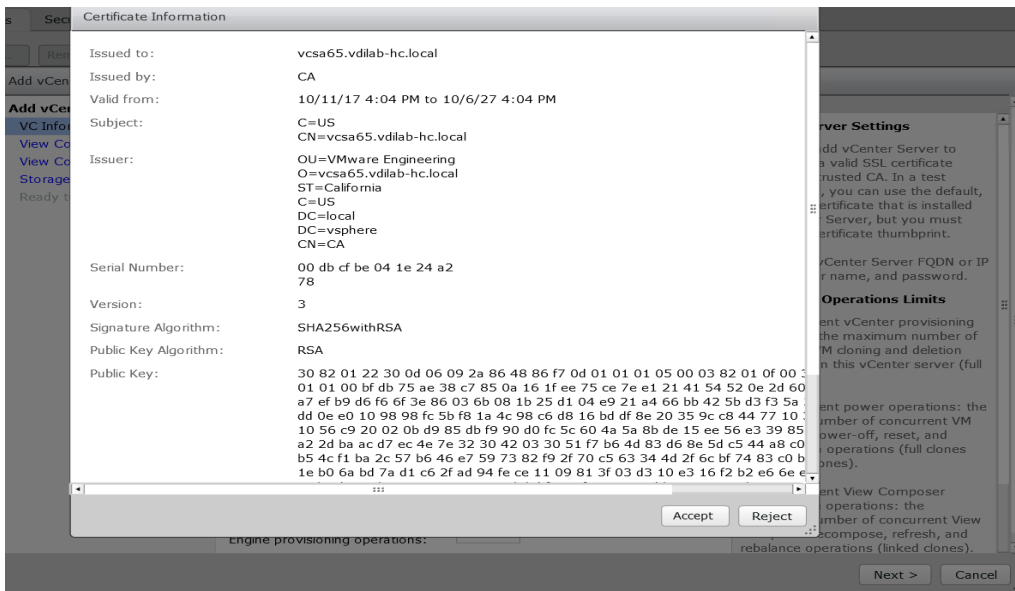
5. Add View composer settings, View composer server FQDN or Management IP address, login credentials.

6. Click Next.
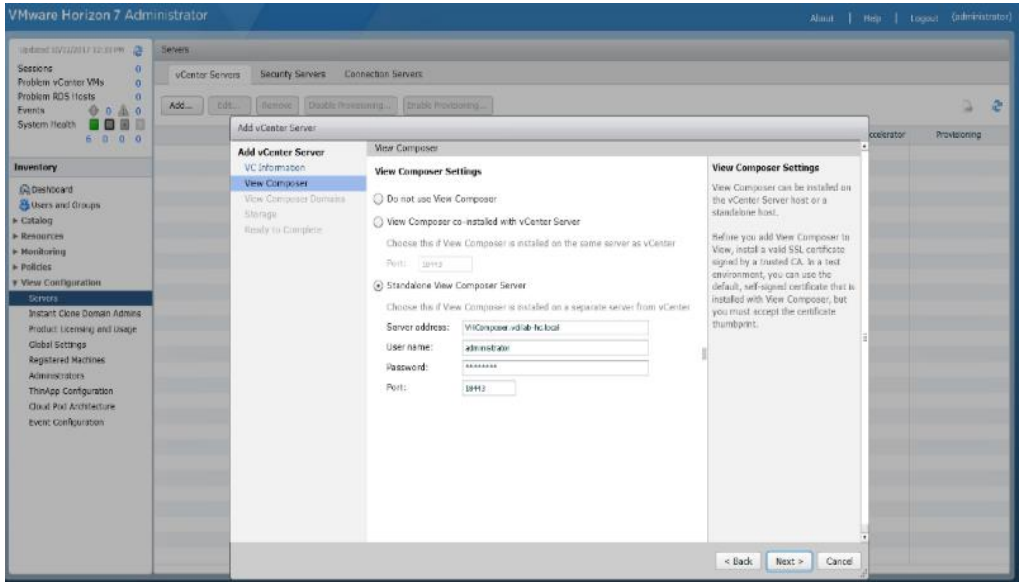
7. View and accept the certificate.
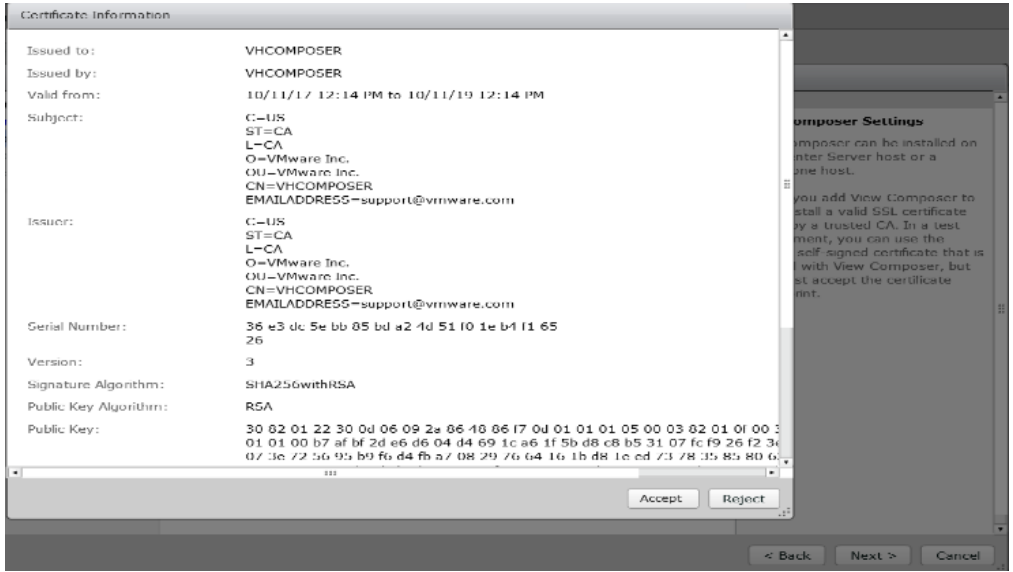
8. Click Add a new domain or Edit the existing domain.

9. Click Next.
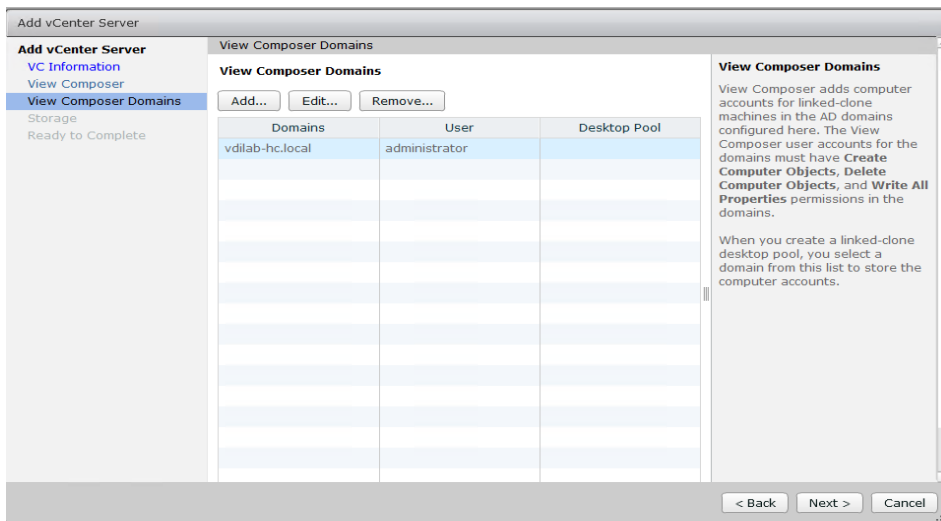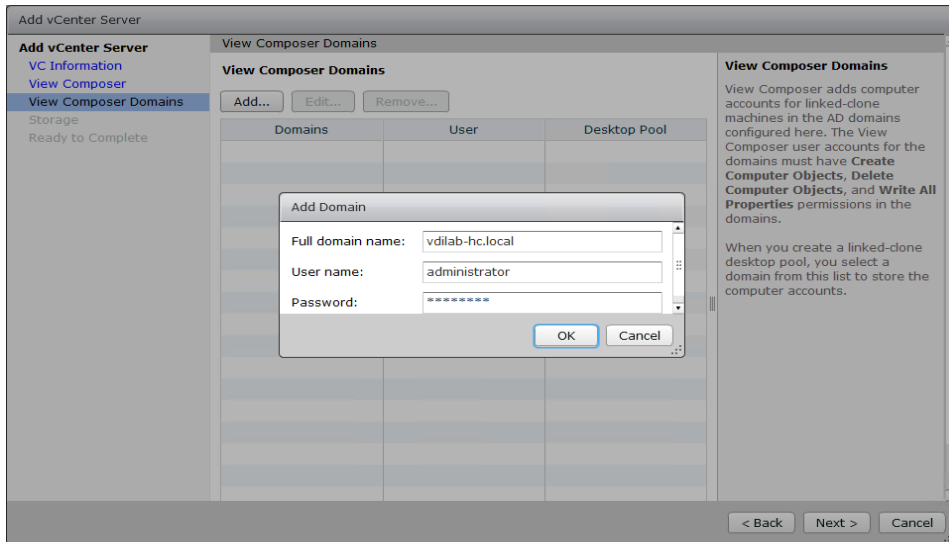
10. In Storage settings, select Reclaim VM disk space and View Storage Accelerator.

11. Configure default host cache size between 100MB and 2048MB. We configured the maximum, which is 2048MB.

12. Click Next.
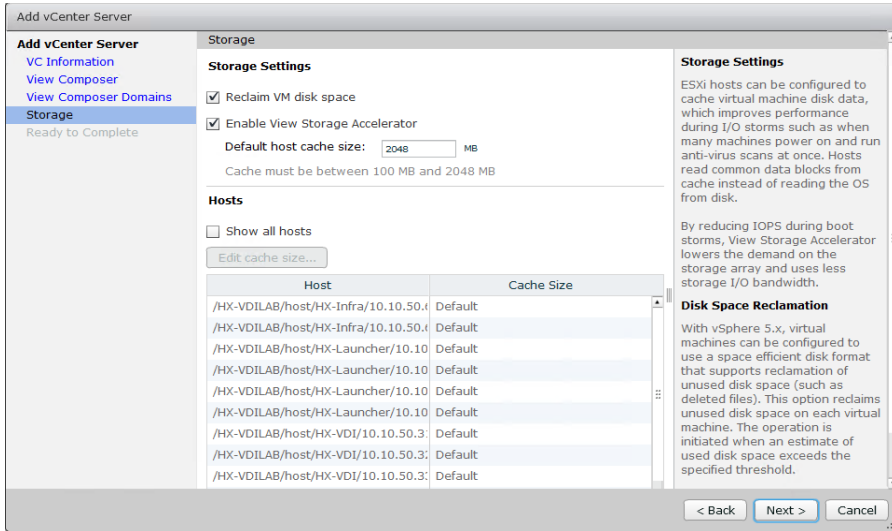
13. Review Add vCenter Server settings and click Finish.
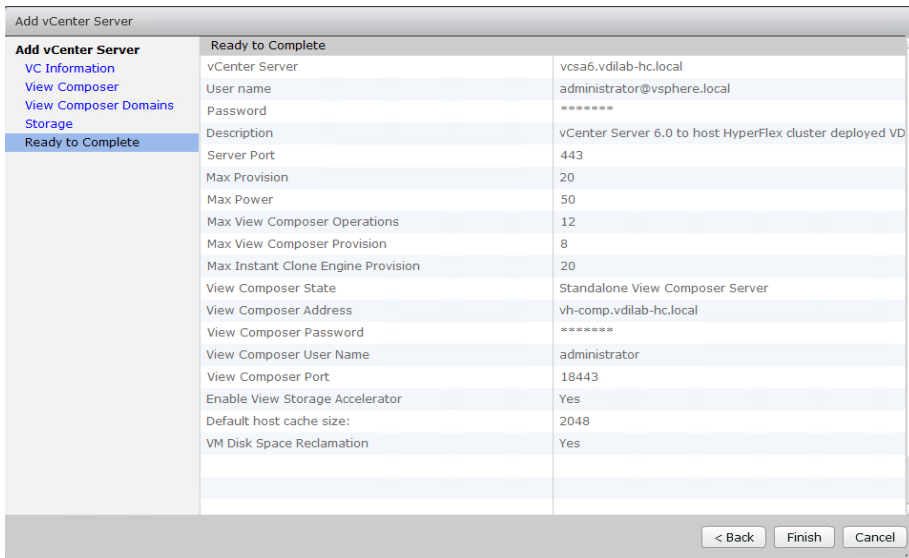
Configure Instant Clone Domain Admins

To configure the instant clone domain admins, follow these steps:

1. Under View Configuration, Click on Instant Clone Domain Admins.

2. Click Add button. Enter credentials for domain user/group.
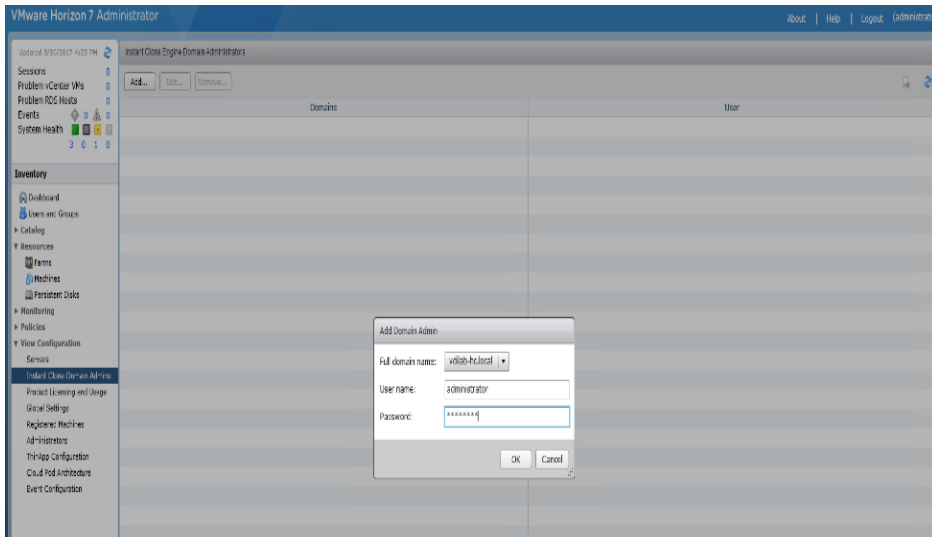
Horizon Persona Manager Installation

To install Horizon Persona Manager, follow these steps:

1. Open Horizon Persona Manager Installer, VMware-personamanagement-x86_64-7.12.0-15085436.exe.

2. Click Next.
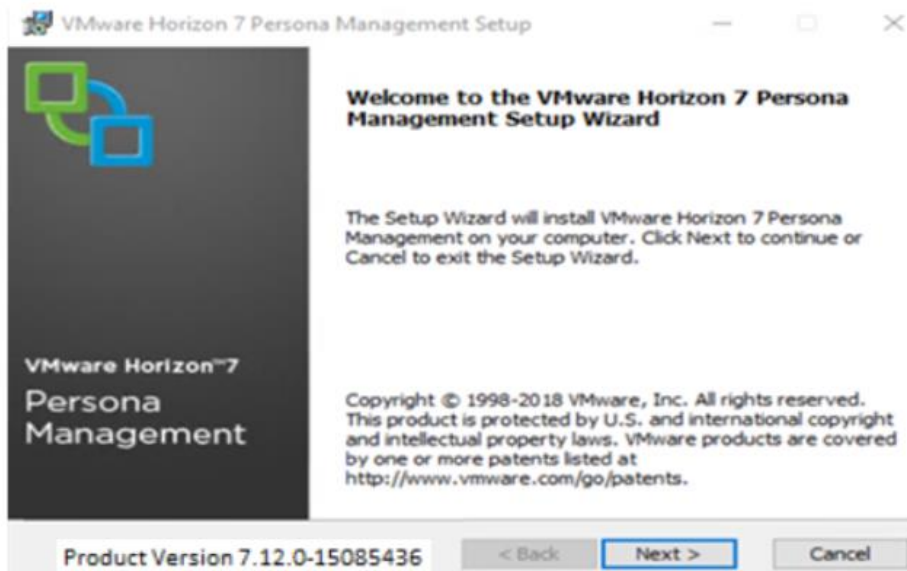
3. Accept the EULA and click Next.
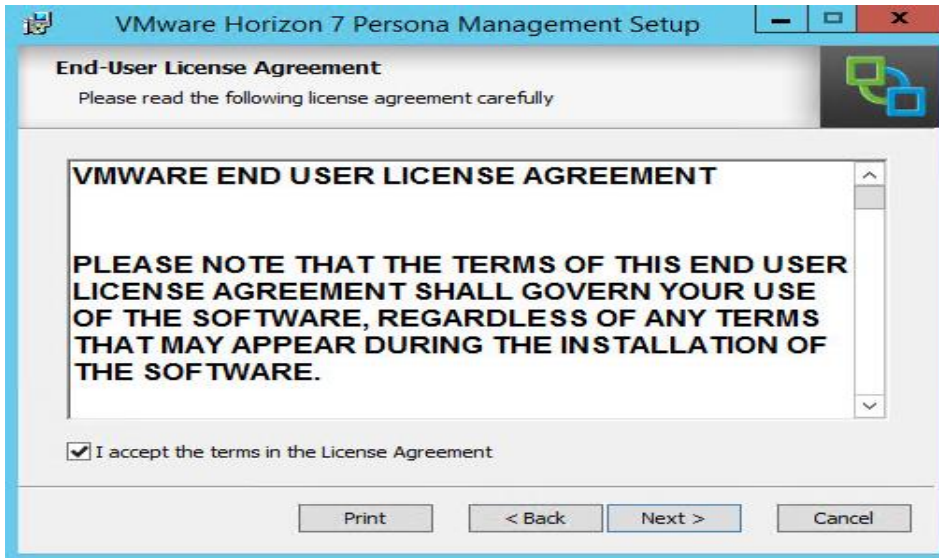
4. Click Install.
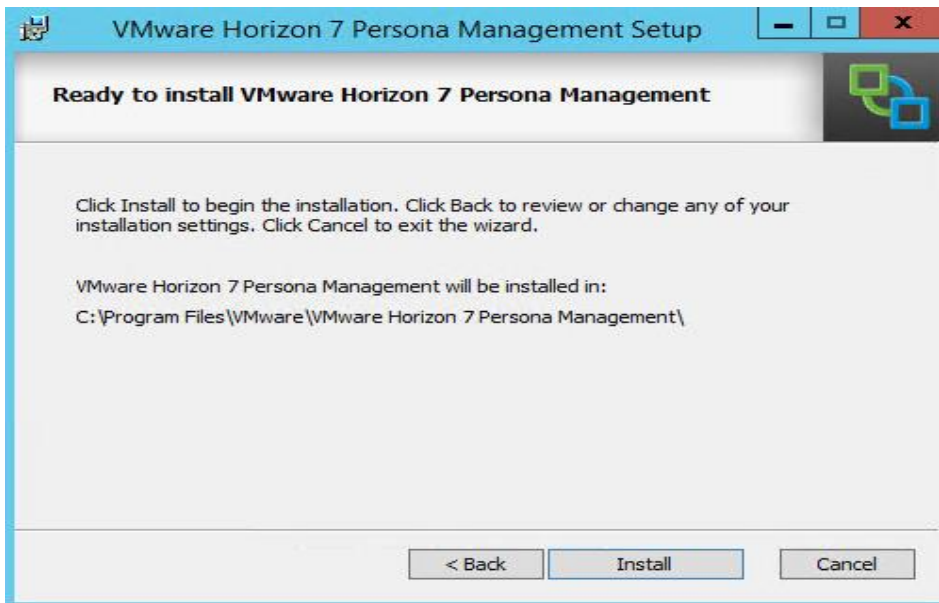
5. Click Finish.
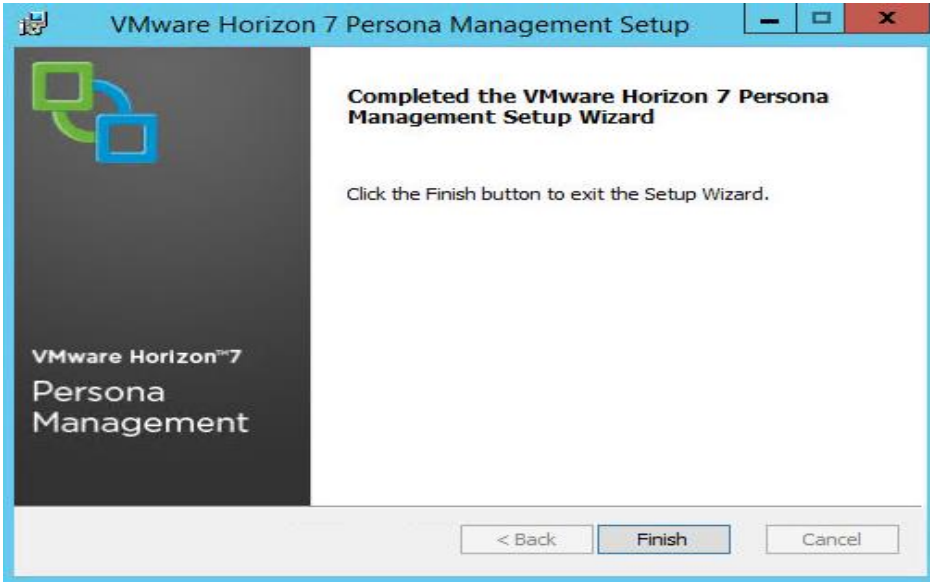
Master Image Creation for Tested Horizon Deployment Types

To create the Master Image for the tested Horizon deployment types, follow this step:

1. Select an ESXi host in an existing infrastructure cluster and create the virtual machines to use as Golden Images with Windows 10 and Office 2016 for Instant Clone, and Full Clone desktops.

We used a 64-bit version of Microsoft Operating System and Office for our testing.

A third master image has been created using Microsoft Windows Server 2019 for Remote Desktop Server Sessions (RDSH) server session virtual machines.

For the Master Image virtual machines, the following parameters were used (Table 51).

Table 51   Golden Image Virtual Machine Parameters

Attribute	Instant / Non-Persistent Clone	Persistent / Full Clone	RDSH server
Desktop operating system	Microsoft Windows 10 Enterprise (64-bit)	Microsoft Windows 10 Enterprise (64-bit)	Microsoft Windows Server 2019 standard (64-bit)
Hardware	VMware Virtual Hardware Version 13	VMware Virtual Hardware Version 13	VMware Virtual Hardware Version 13
vCPU	2	2	8

Attribute	Instant / Non-Persistent Clone	Persistent / Full Clone	RDSH server
Memory	4096 MB	4096 MB*	32768MB
Memory reserved	4096 MB	4096 MB*	32768MB
Video RAM	35 MB	35 MB	4MB
3D graphics	Off	Off	Off
NIC	1	1	1
Virtual network adapter 1	VMXNet3 adapter	VMXNet3 adapter	VMXNet3 adapter
Virtual SCSI controller 0	Paravirtual	Paravirtual	Paravirtual
Virtual disk: VMDK 1	40 GB	100 GB	80 GB
Virtual floppy drive 1	Removed	Removed	Removed
Virtual CD/DVD drive 1	-	-	-
Applications	Login VSI 4.1.32 application installation Adobe Acrobat 11 Adobe Flash Player 16 Doro PDF 1.82 FreeMind Microsoft Internet Explorer Microsoft Office 2016	Login VSI 4.1.32 application installation Adobe Acrobat 11 Adobe Flash Player 16 Doro PDF 1.82 FreeMind Microsoft Internet Explorer Microsoft Office 2016	Login VSI 4.1.32 application installation Adobe Acrobat 11 Adobe Flash Player 16 Doro PDF 1.82 FreeMind Microsoft Internet Explorer Microsoft Office 2016
VMware tools	Release 11.1.0	Release 11.1.0	Release 11.1.0
VMware View Agent	Release 7.12.-15805436	Release 7.12.0-15805436	Release 7.12.0-15805436
Attribute	Instant-clone	Persistent/Full Clone	RDSH Server

* For Persistent Desktops, we configured 4GB of RAM as amount of memory allocated is enough to run LoginVSI Knowledge Worker workload. HyperFlex nodes and compute-only node were configured with 768GB of total memory for this performance study.

Prepare Microsoft Windows 10 and Server 2019 R2 with Microsoft Office 2016

Prepare your master image for one or more of the following use cases:

- VMware Horizon 7 Remote Desktop Server Sessions (RDSH) Server 2019 Virtual Machines

- VMware Horizon 7 Instant Clones non- persistent virtual machines

- VMware Horizon 7 persistent virtual machines

Include Microsoft Office 2016 and other applications used by all pool users in your organization into your master image.

Apply the required Microsoft updates and patches to your master images.

For this study, we added Login VSI target software to enable the use the Login VSI Knowledge Worker workload to benchmark end user experience for each use case.

Optimization of Base Windows 10 or Server 2019 Guest OS

Click the links below for information about how to optimize windows 10 for VDI deployment:

VMware Windows Operating System Optimization Tool Guide: http://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/whitepaper/vmware-view-optimizationguidewindows7-en-white-paper.pdf

VMware Optimization Tool for HVD or HSD Deployment: https://labs.vmware.com/flings/vmware-os-optimization-tool

Virtual Desktop Agent Software Installation for Horizon

To install the Virtual Desktop Agent software for Horizon, follow these steps:

1. For each master image created, open the Horizon View Agent Installer, VMware-viewagent-7.12.0-15805436. Click Next to install.
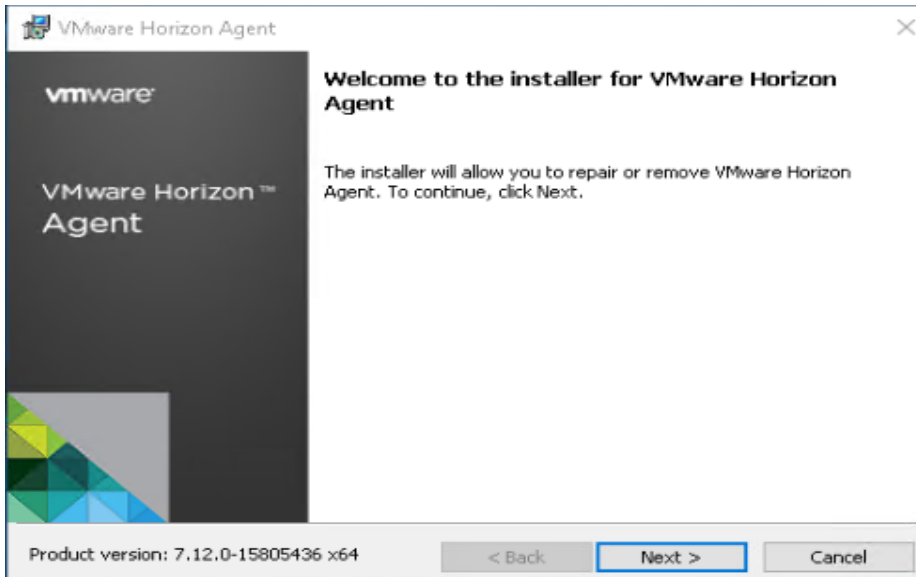
2. Review and accept the EULA Agreement. Click Next.
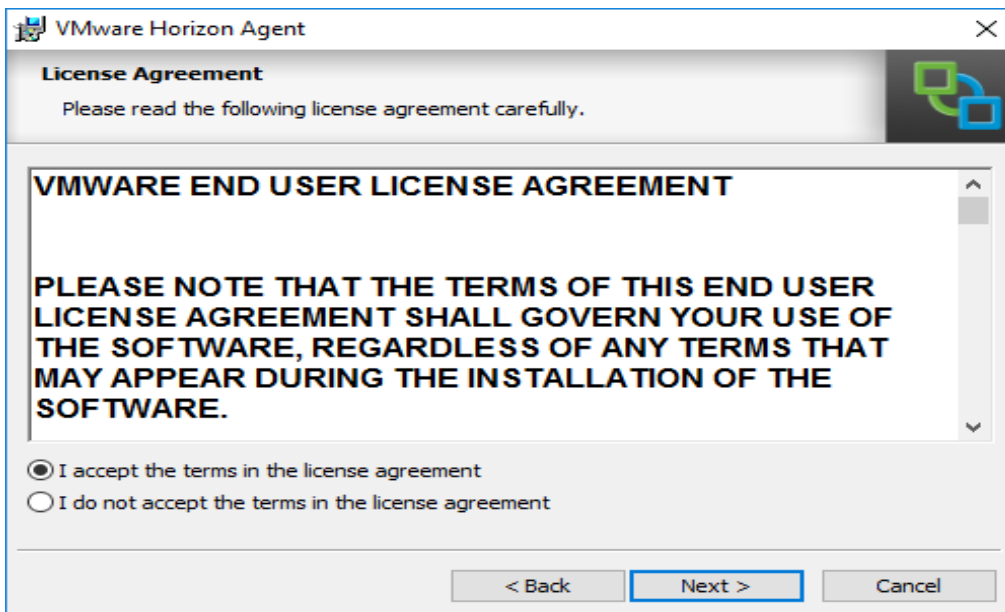
3. Select Network protocol configuration, click Next.
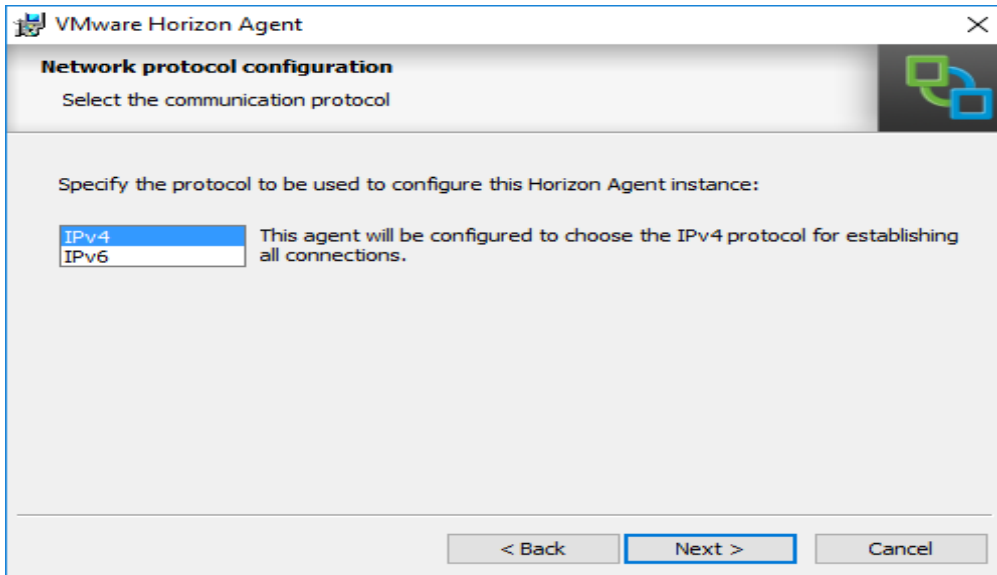
4. Based on the Desktop pool you want to create, select either View Composer Agent or Instant Clone Agent in-stallation. Do not install both features on the same master image.

5. Enable installation of the VMware Horizon View Instant Agent for Instant -clone VDI virtual machines.
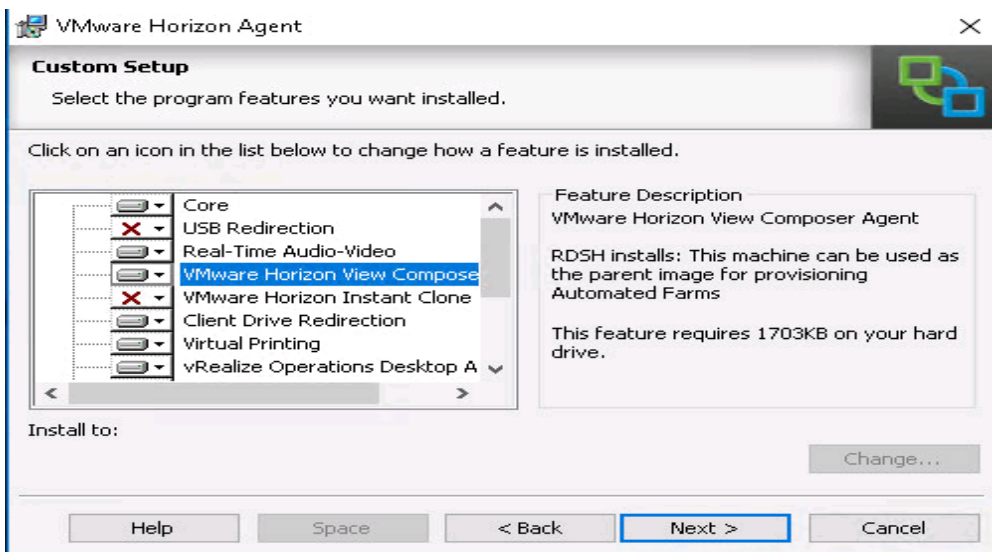
6. Disable the Horizon View Composer Agent and enable the Horizon Instant Clone Agent for Instant Clone float-ing assigned desktop pool creation.

⚠️ The VMware Horizon Composer was not tested for this CVD.
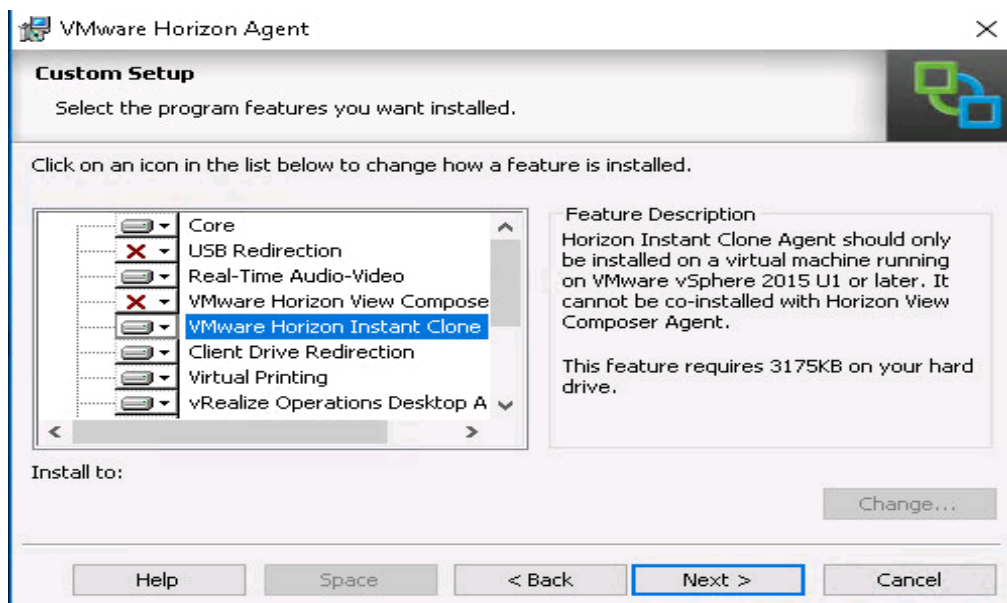
> ⚠ Prior to installing the Horizon View Agent on a Microsoft Server 2019 virtual machine, you must add the Remote Desktop Services role and the Remote Desktop Session Host role service.

7. To add Remote Desktop Services role on Windows Server OS from the Server Manager, use the Add Roles and Features wizard:
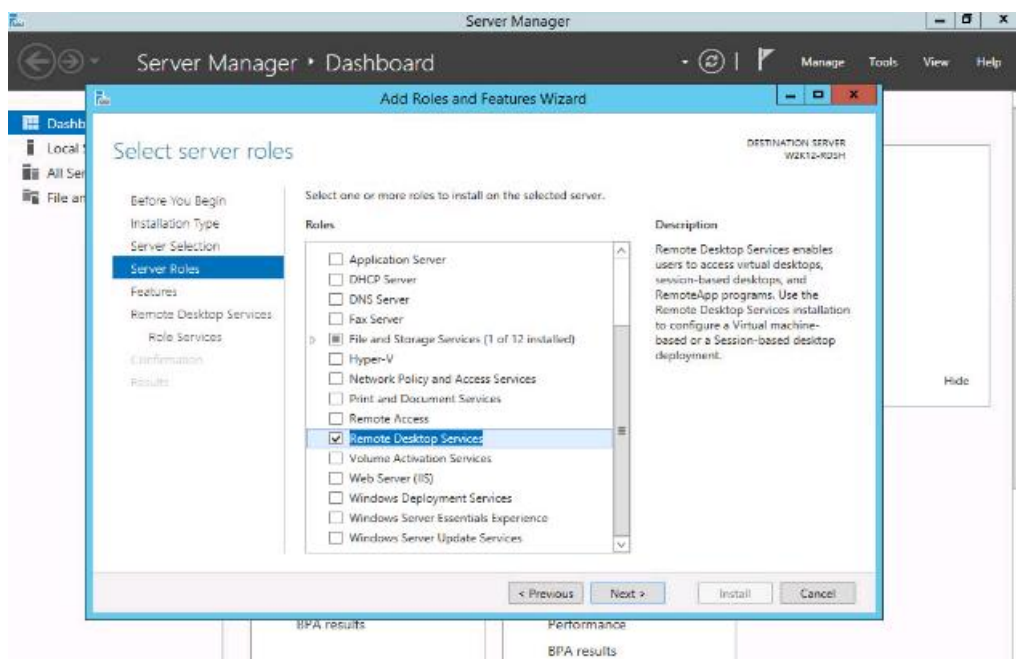
8. Add Remote Desktop Session Host services.
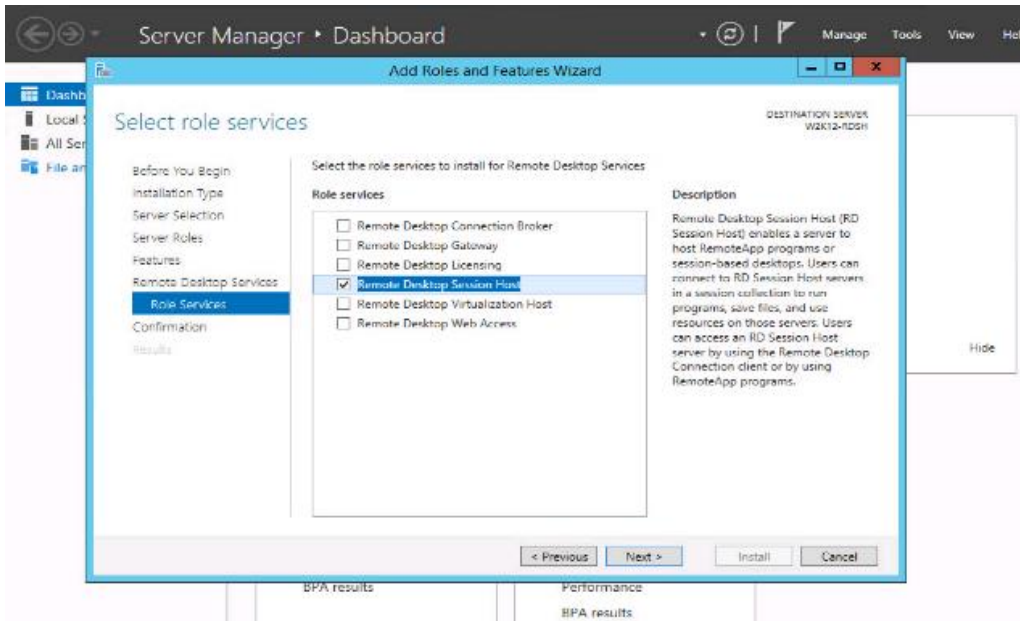
9. Click Install.
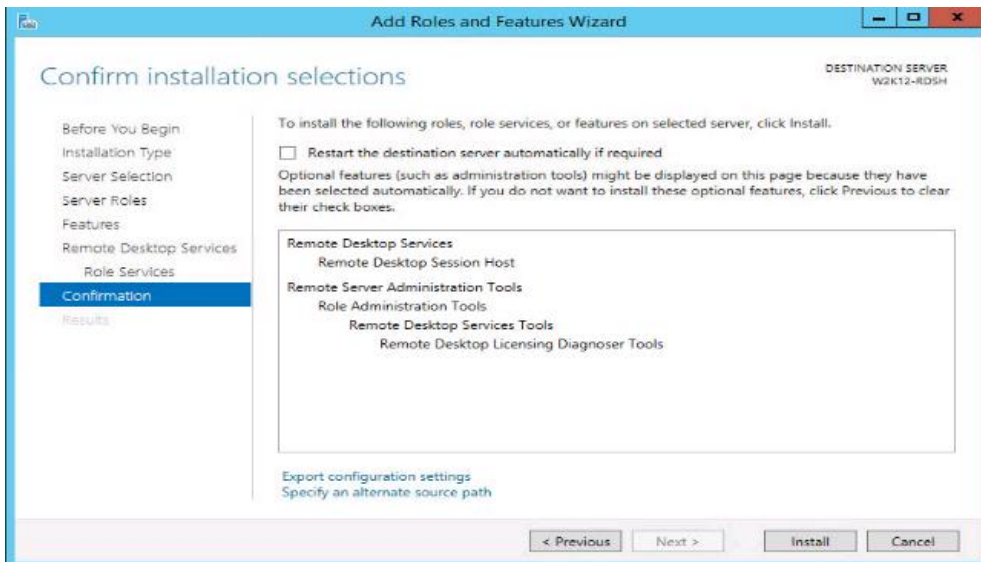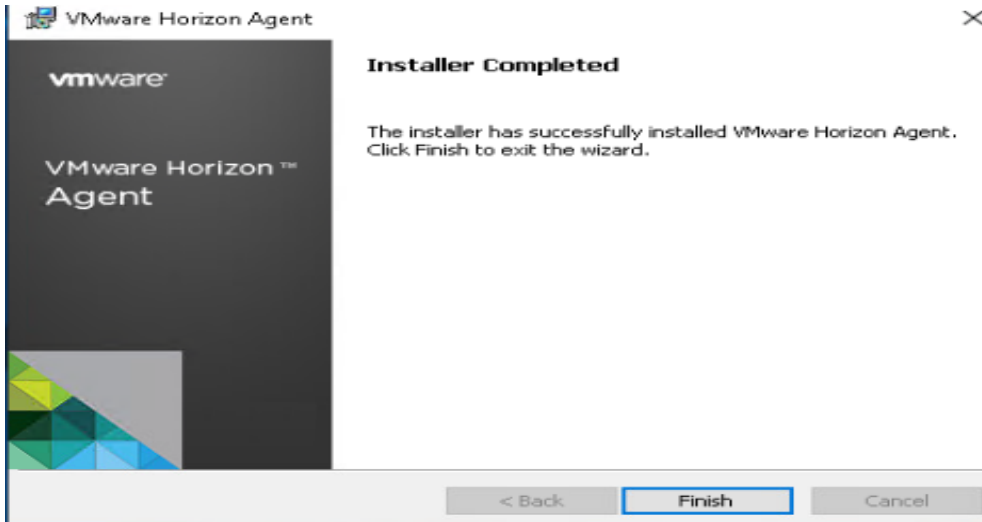
10. View Agent will report as installed in "Desktop Mode" if the Remote Desktop Services not installed.
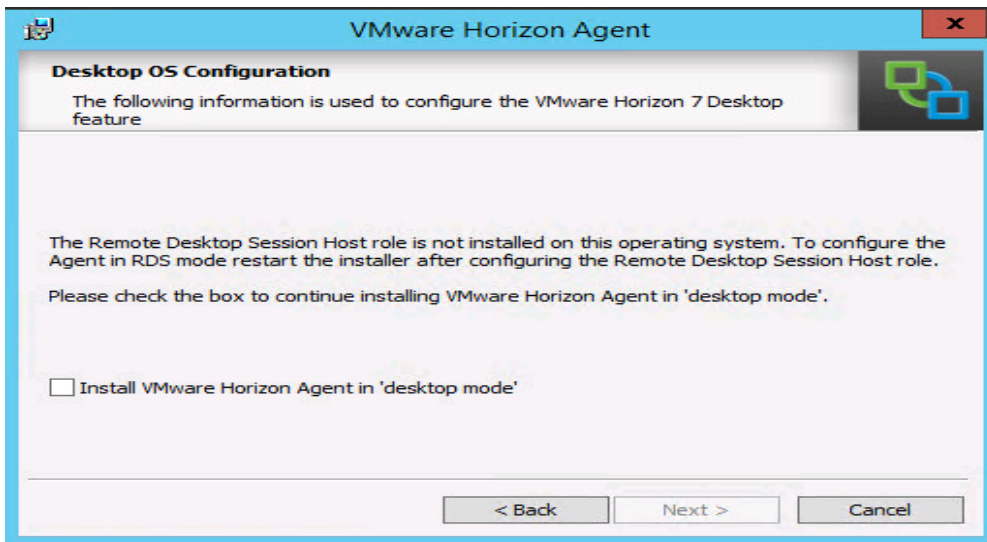
11. Add FQDN or IP address for Connection Server Instance to register the RDSH server.
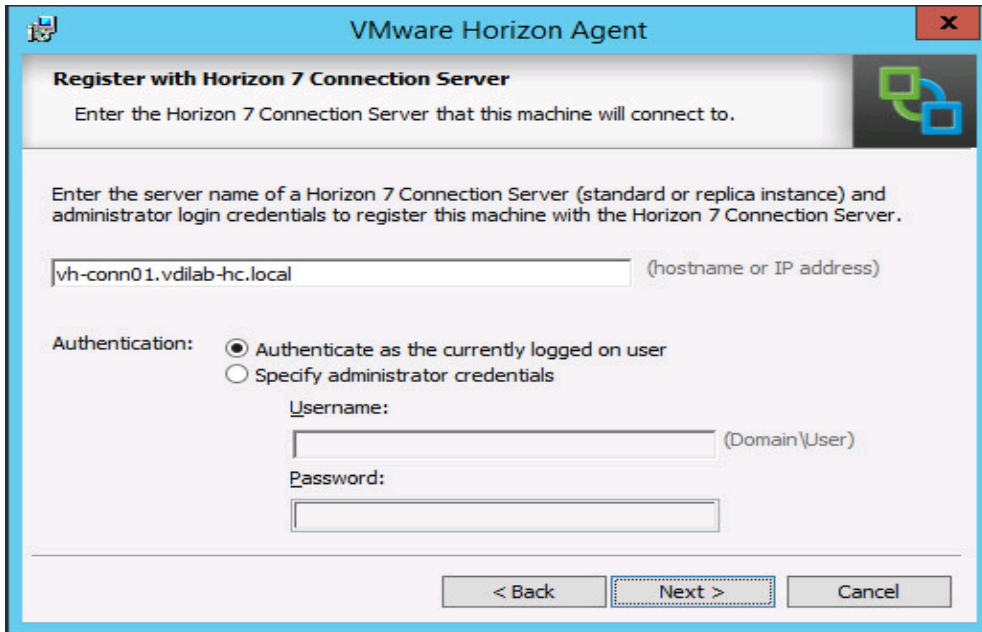
12. Click Install.
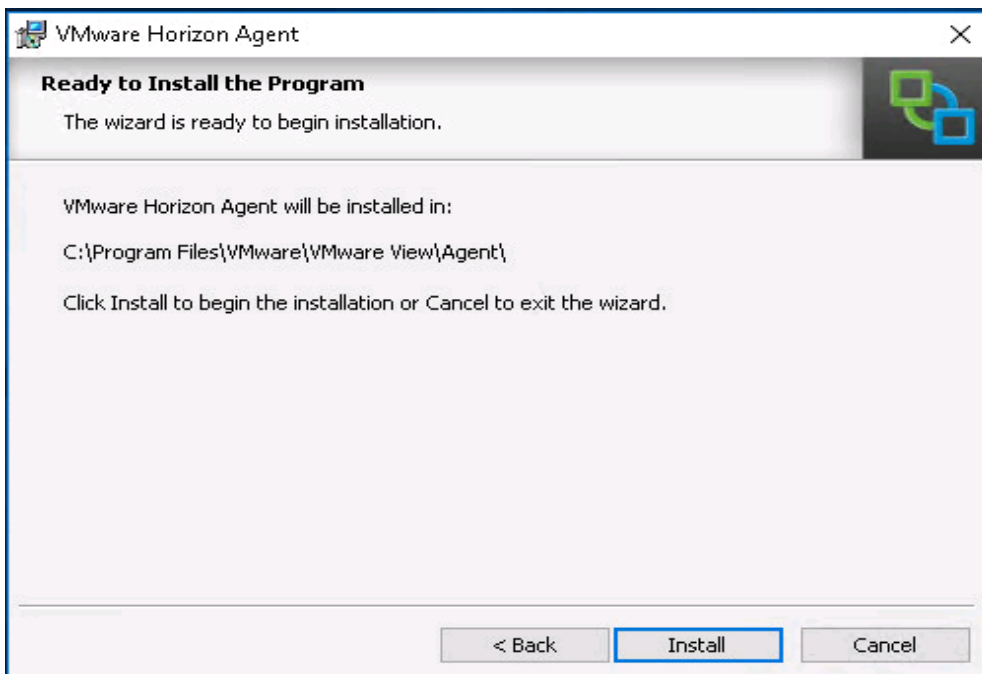
13. Click Finish and restart the VM.
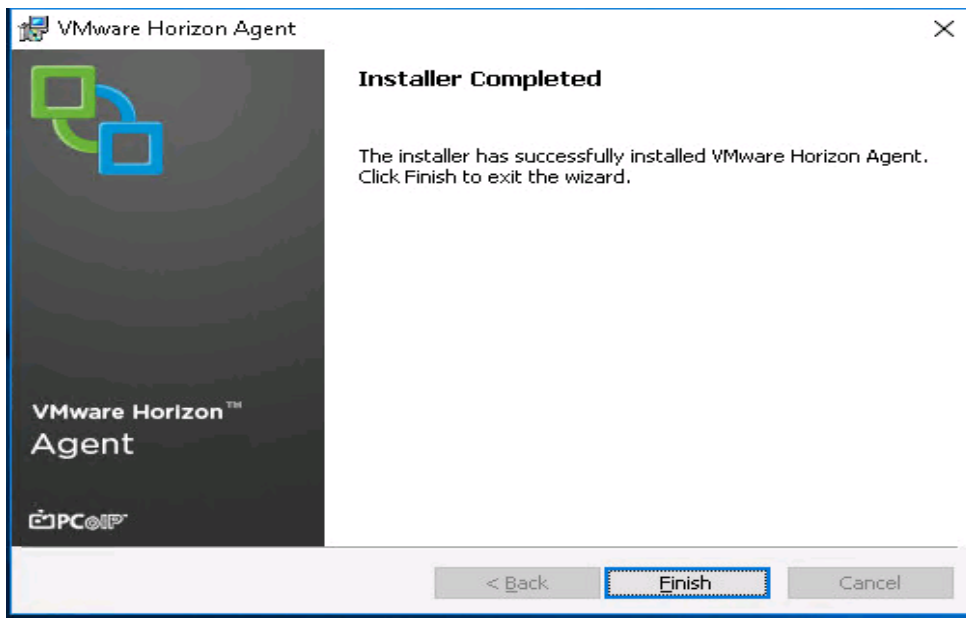
Install Additional Software

To install additional software required for your base windows image, follow these steps:

1. For testing, we installed Microsoft Office 2016 64-bit version.

2. Log into the VSI Target software package to facilitate workload testing.

3. Install service packs and hot fixes required for the additional software components that are being added.

4. Reboot or shut down the VM as required.

Create a Native Snapshot for Automated Desktop Pool Creation

To create a native snapshot for the automated desktop pool, follow these steps:

1. Log into vCenter WebUI.

2. Select the master image for the automated desktop pool creation.

3. Right-click and select Cisco HX Data Platform > Snapshot Now.
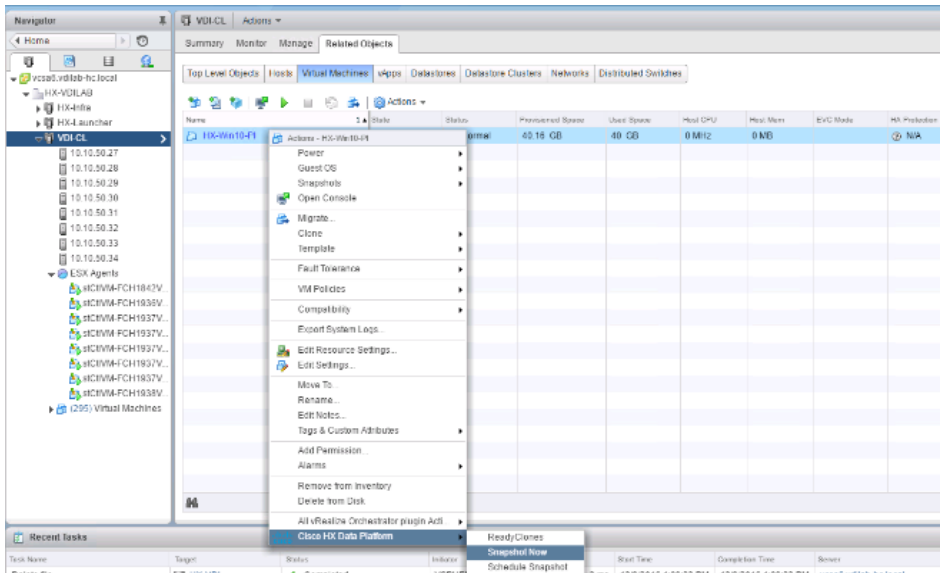
4. Enter a name for the HX native snapshot.
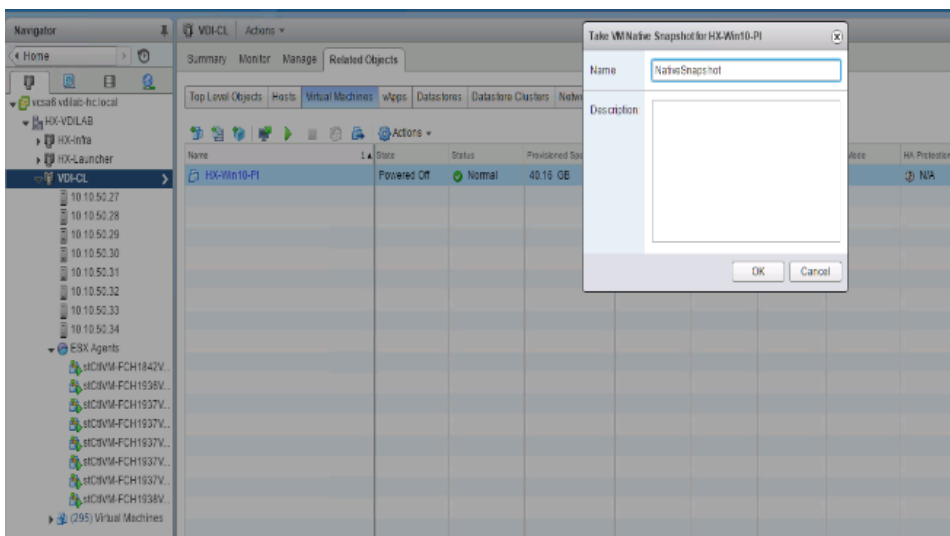
Create Customization Specification for Virtual Desktops

To create Customization Specification for virtual desktops, follow these steps:

1. On vCenter WebUI, select Customization Specification Manager.
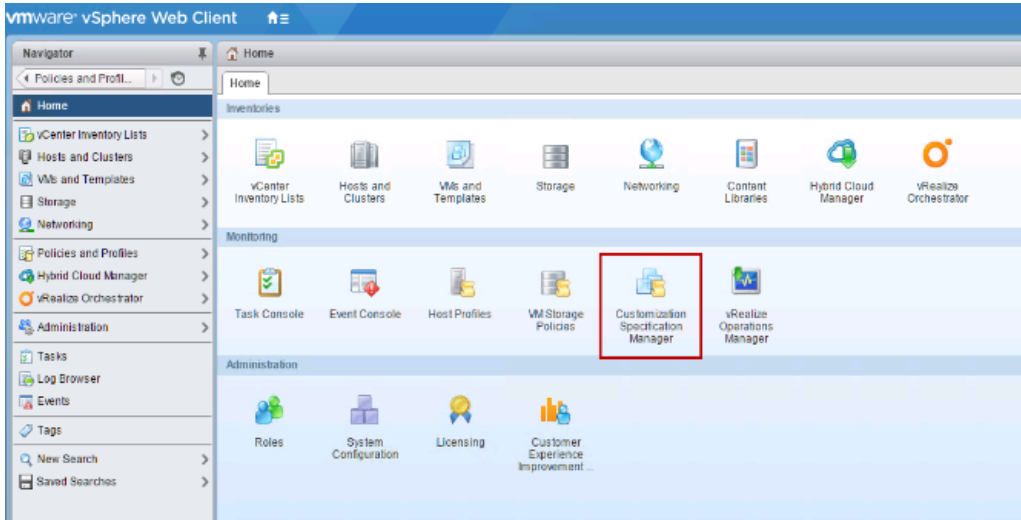
2. Select VM Operating System as Windows for Windows based guest OS optimization. Enter a name.
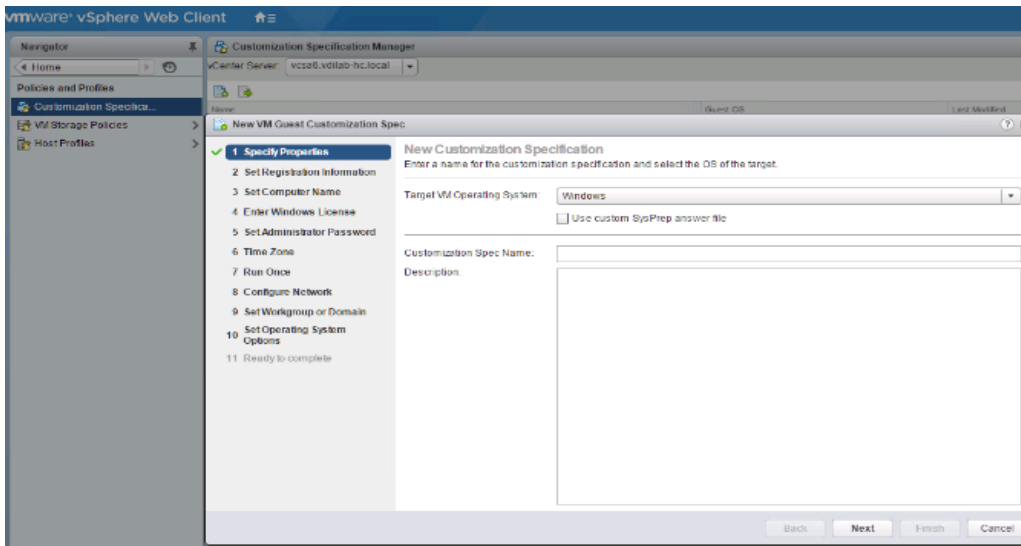
3. Provide name and organization details.
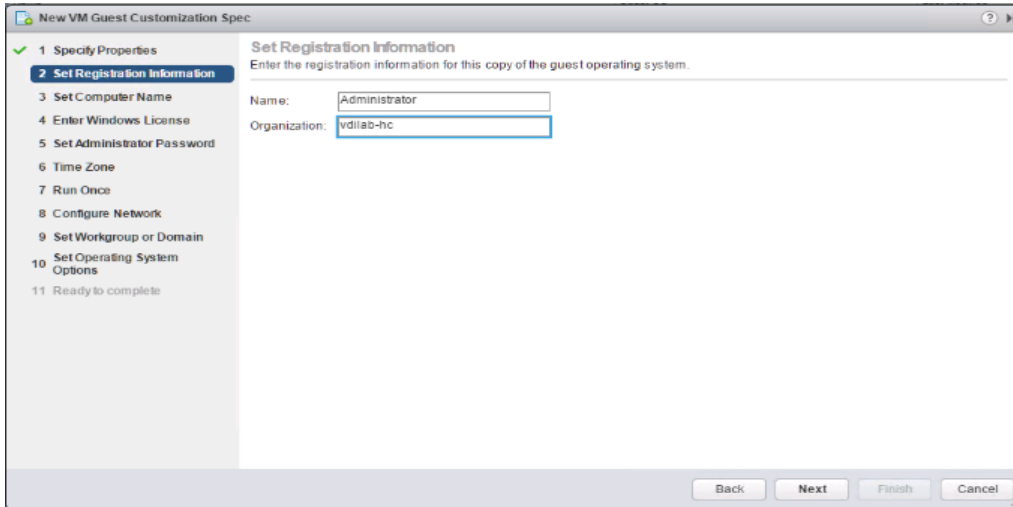
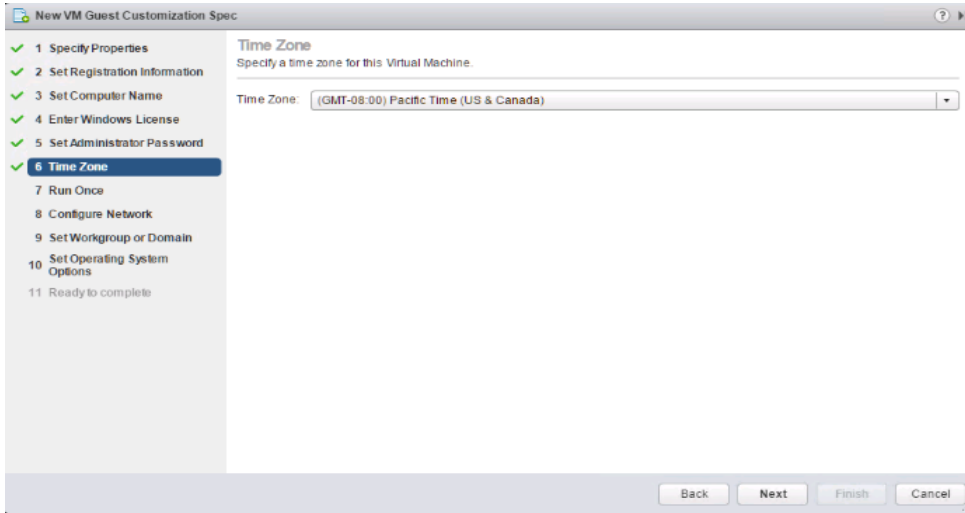4. Provide a computer name. For this solution, we selected Use the virtual machine name.
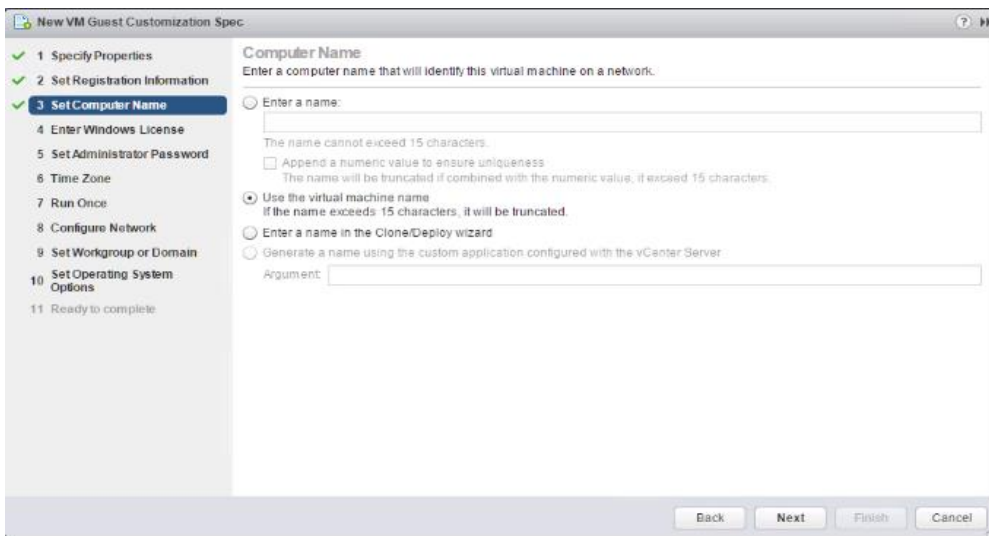
5. Provide the product License key if required.
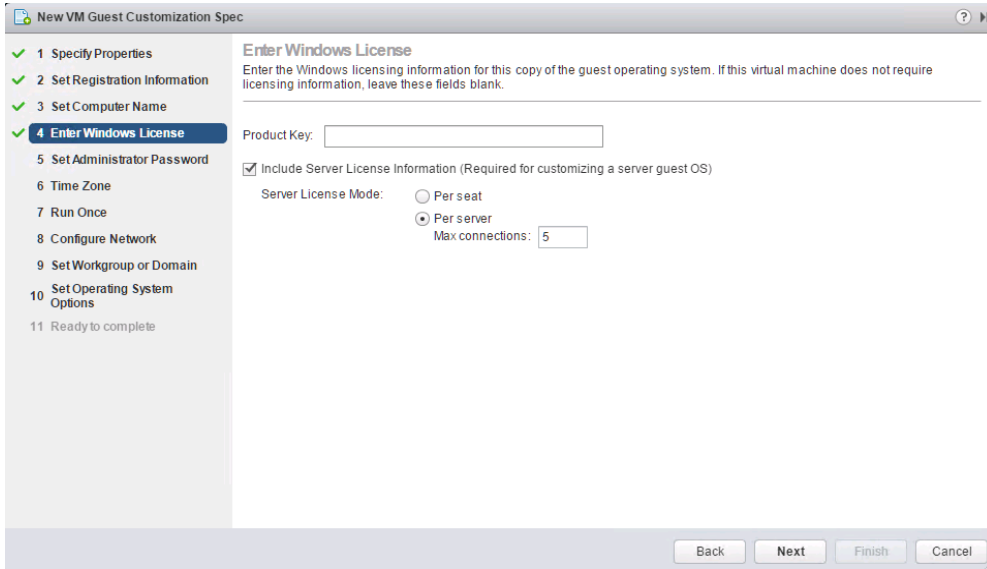
6. Provide Password credentials.
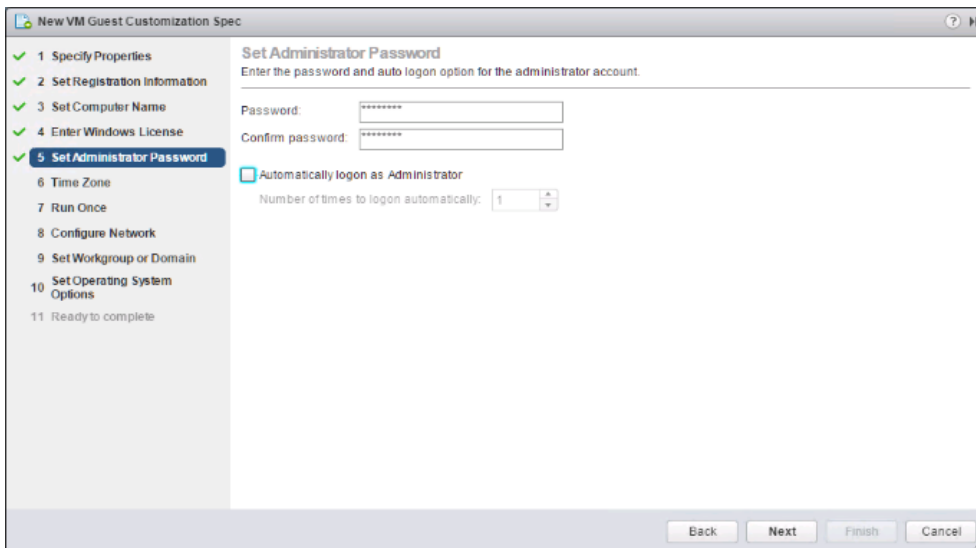
7. Select the Timezone.

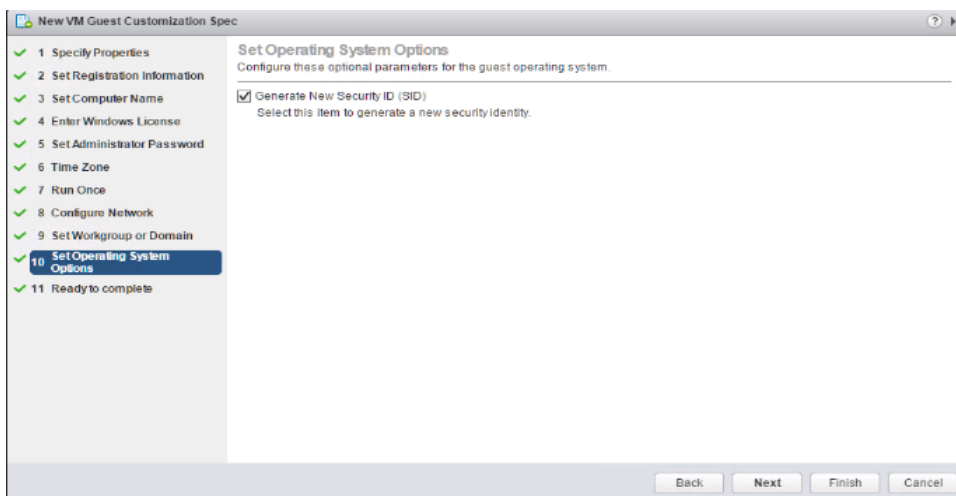8. Add the commands to run when the first-time user logs in, if there are any.
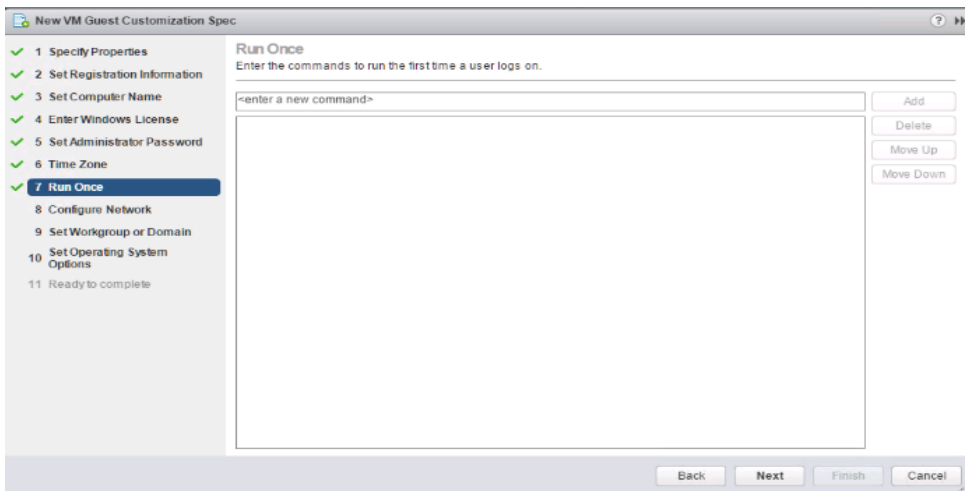
9. Provide the network information whether to use the DHCP server to assign IP address, or manual configuration.
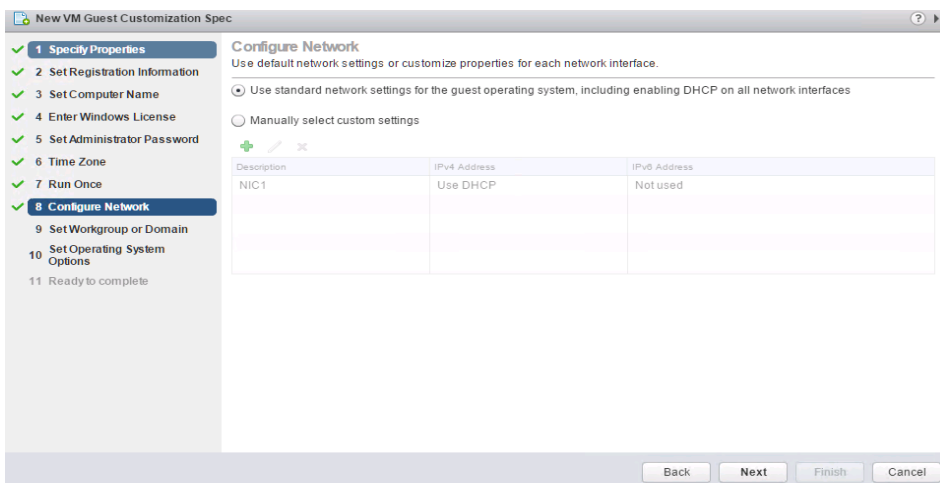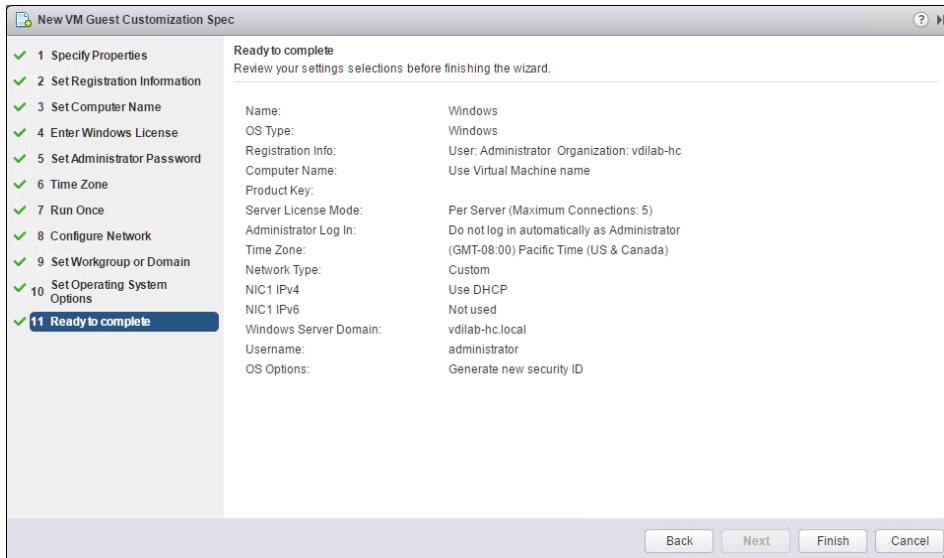
10. Provide the domain name and user credentials.
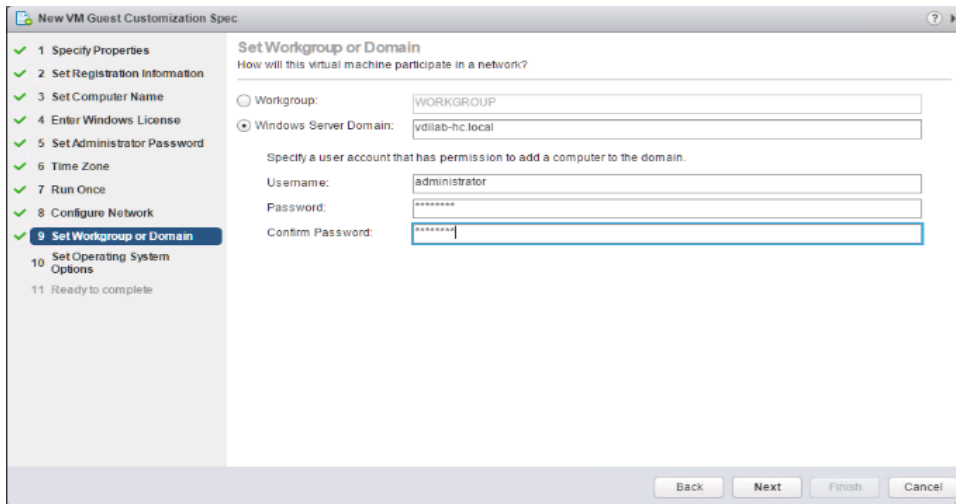
11. Select the checkbox Generate New Security ID (SID).

12. Review and click Next to complete creating the Customization Specs.
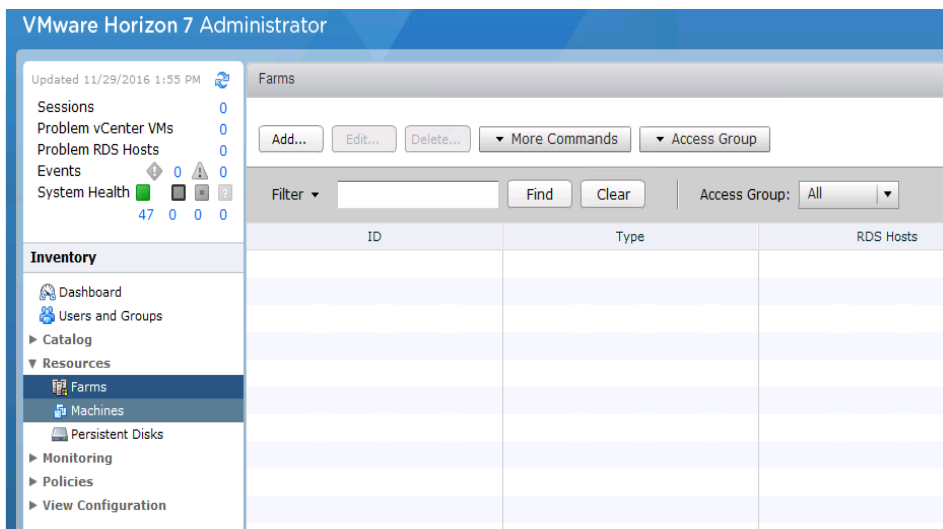
13. Click Finish.

RDSH Farm Creation

Before you can create an RDSH desktop pool, you must first create a RDSH Farm. To create a RDSH Farm, fol-low these steps:

1. In the VMware Horizon Administration console, select Farms under the Resource node of the Inventory pane.

2. Click Add in the action pane to create a new RDSH Farm.

3. Select either to create an Automated or Manual Farm. In this solution, we selected Automated Farm.

⚠️   A Manual Farm requires a manual registration of each RDSH server to Horizon Connection or Replica Server instance.
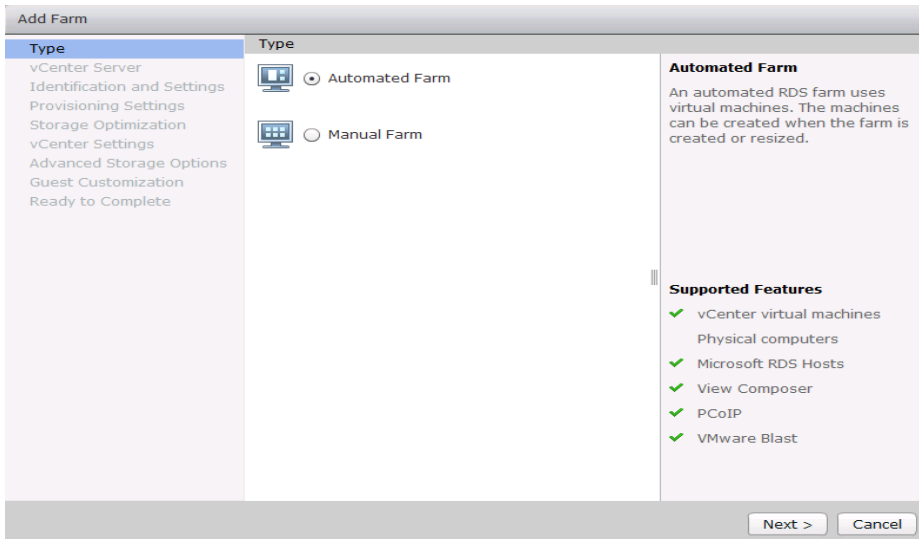
4. Select the vCenter Server and Horizon Composer server that you will use to deploy the Horizon RDSH Farm.

5. Click Next.
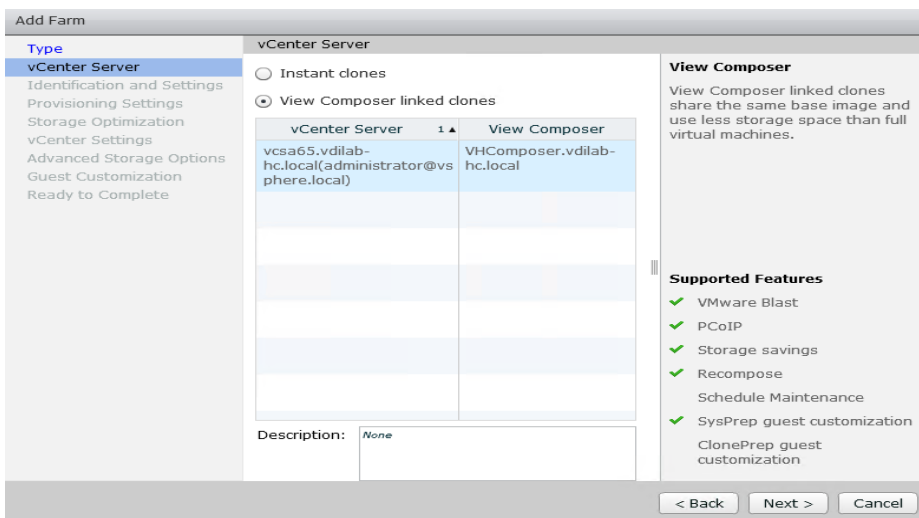
You can choose to create either Instant clones or View Composer linked clones for the RDSH server FARM server VMs. Both have benefits and limitations, but detailing these differences are beyond the scope of this CVD. Please refer to your VMware documentation for more information.

6. Enter the RDSH Farm ID, Access group, Default Display Protocol (Blast/PCoIP/RDP).

7. Select if users are allowed to change the default display protocol, Session timeout, Logoff Disconnected users, and select the checkbox to Enable HTML access.

8. Click Next.

9. Select the provisioning settings, naming convention for RDSH server VM to deploy, and the number of VMs to deploy.

In this study, we deployed 1032 RDSH virtual machines across our 16 node HyperFlex Cluster.

10. Click Next.

11. Click Next.

12. Select vCenter settings, for example; Master Image, snapshot, folder, Host or Cluster, resource pool, storage selection.

13. Click Next.
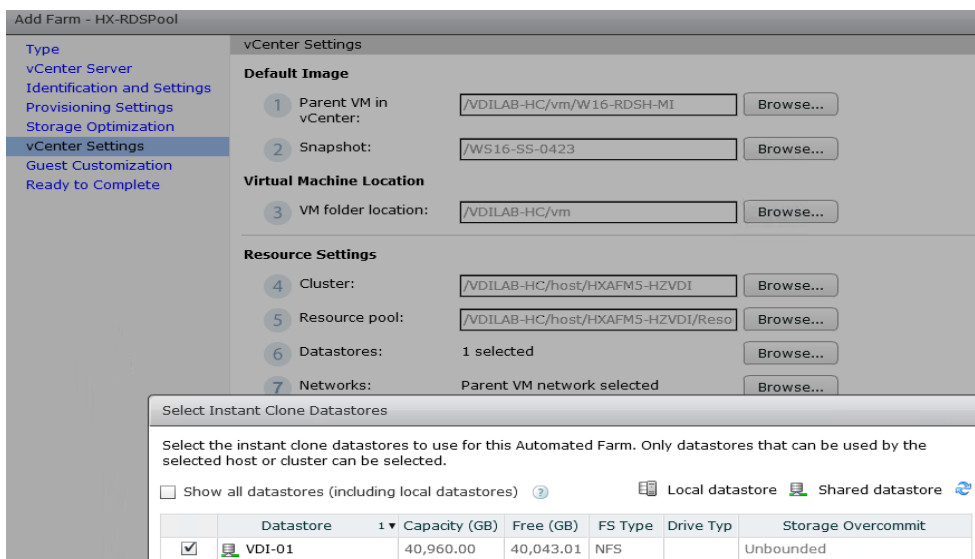
14. For Step 6 Datastores: Browse and choose Unbounded for the Storage Overcommit field.

15. Click OK.
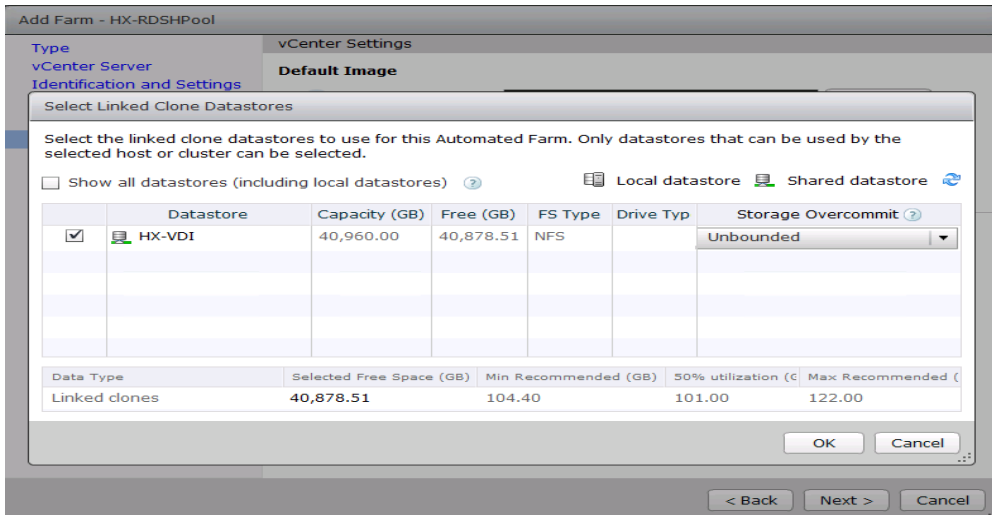
16. Click Next.

17. Select the Active Directory Domain, the Active Directory OU into which the RDSH machines will be provisioned, and the Sysprep file created as part of the customization specific configuration performed earlier.

> ⚠️ If you choose the instant clone pool for the RDSH FARM creation, you may not see the Sys prep guest customization step shown in the screenshot shown below.

18. Click Next.
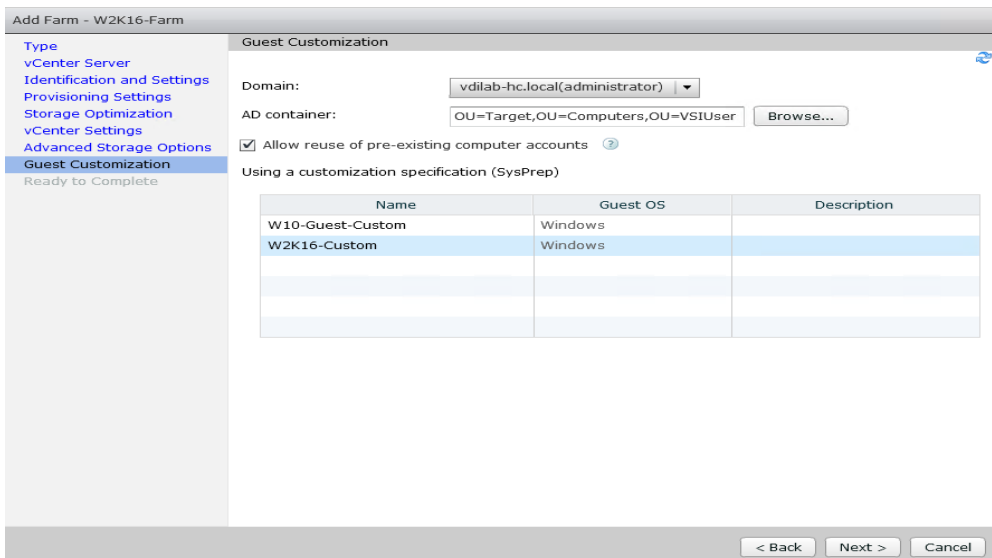
19. Review the pool creation information.

20. Click Finish.
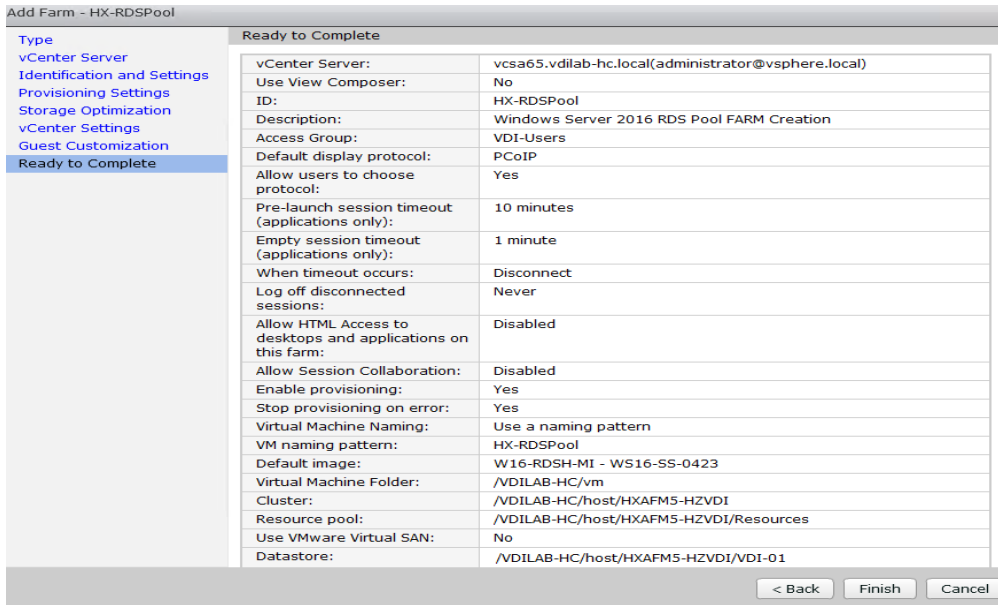
The VMware Horizon Administration console displays the status of the provisioning task and pool settings:

Create the Horizon 7 RDS Published Desktop Pool

To create the Horizon 7 RDS Published Desktop Pool, follow these steps:

1. In the Horizon Administrator console, select Desktop Pools in the Catalog node of the Inventory pane.

2. Click Add in the action pane.
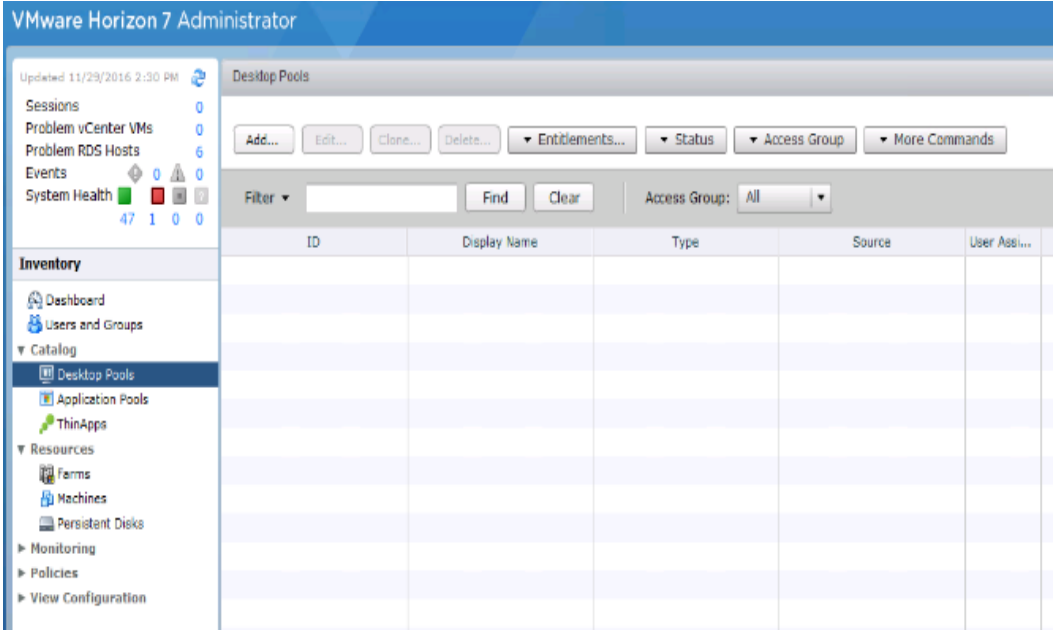
3. Select RDS Desktop pool.

4. Click Next.
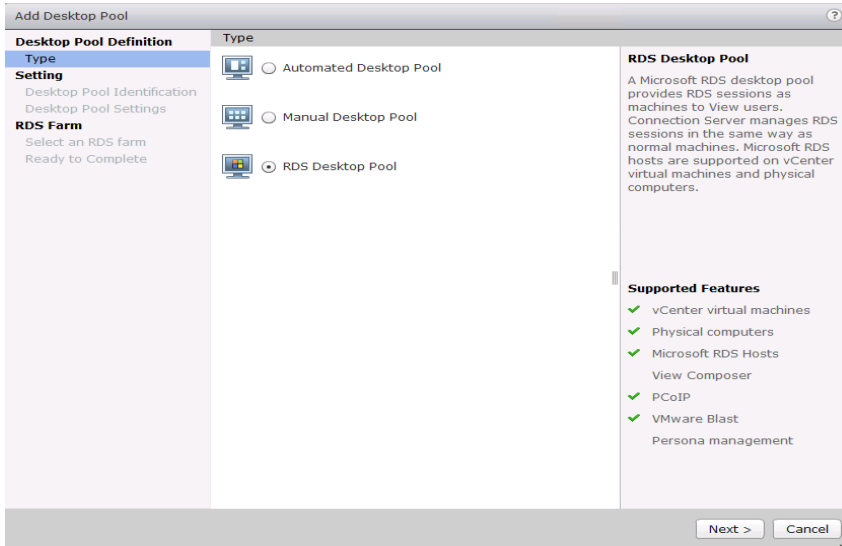
5. Enter Pool ID and Display name.

6. Click Next.
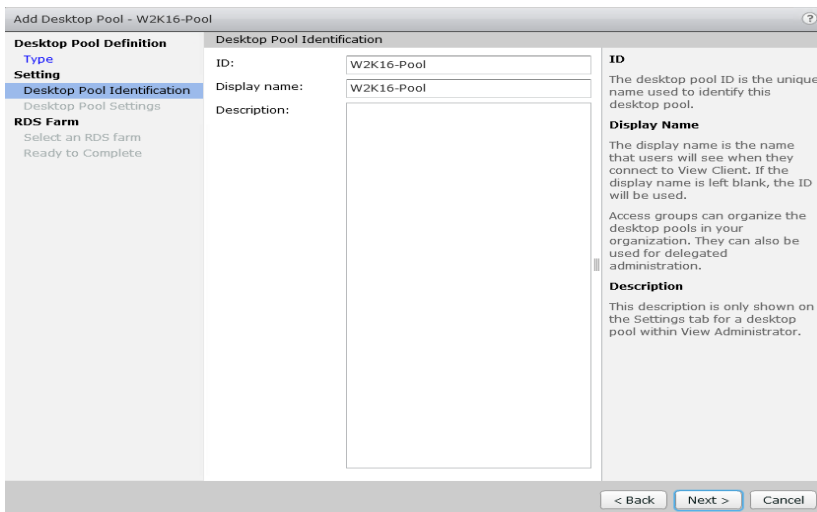
7. Accept the default settings on Desktop Pool Settings page.

8. Click Next.
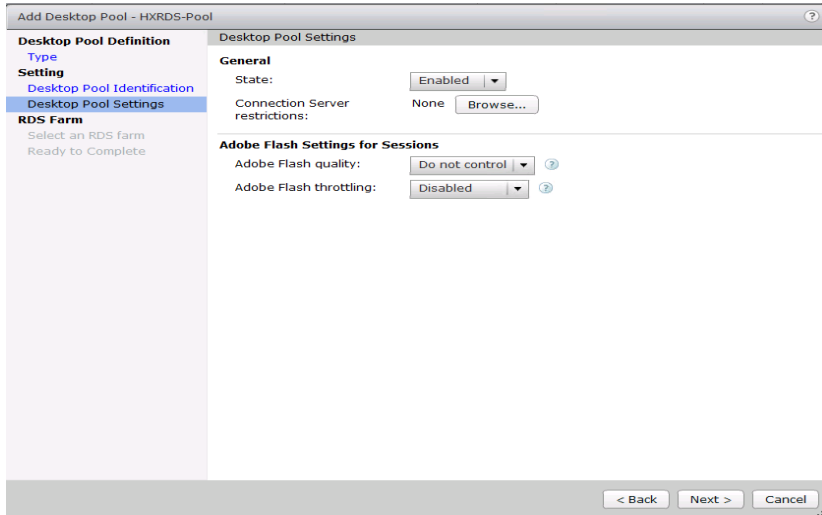
9. Click the "Select an RDS farm for this desktop pool" radio button.

10. Click the farm created in the previous section.

11. Click Next.
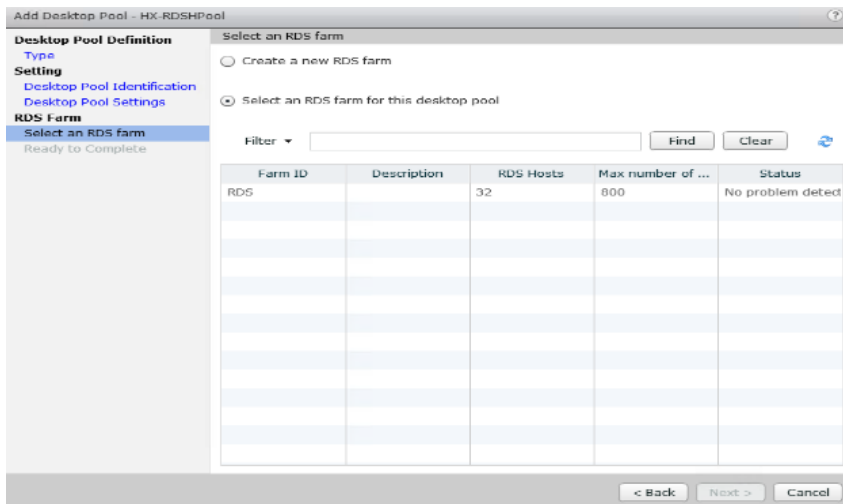
12. Review the pool settings.

13. Select the checkbox "Entitle users after this wizard finishes" to authorize users for the newly create RDSH desktop pool.

14. Click Finish.
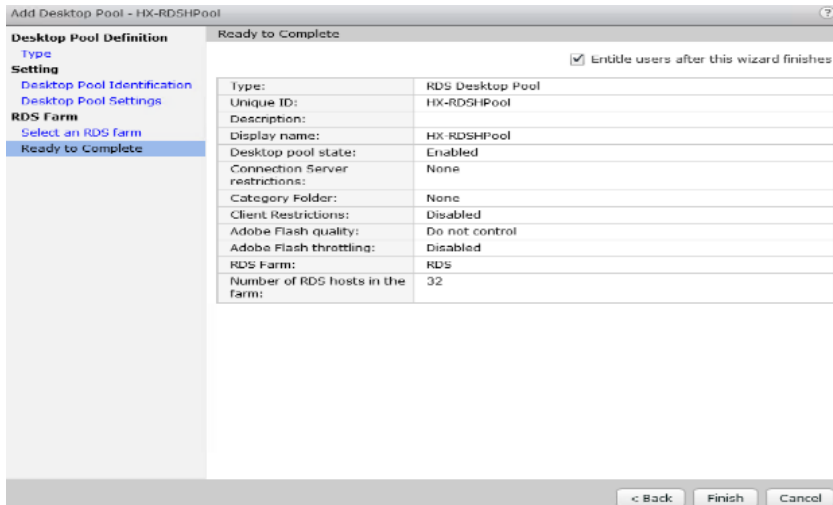
15. Select the Users or Groups checkbox, use the search tools to locate the user or group to be authorized, high-light the user or group in the results box.

16. Click OK.
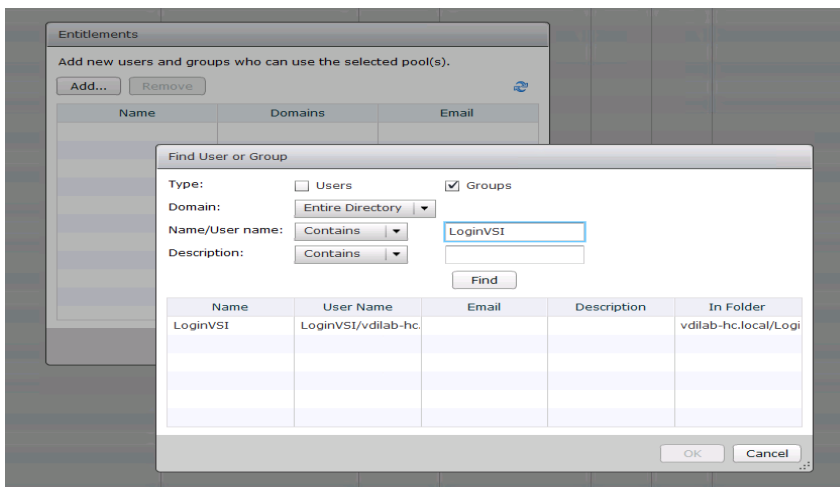
17. You now have a functional RDSH Farm and Desktop Pool with users identified who are authorized to utilize Horizon RDSH sessions.

VMware Horizon Linked-Clone Windows 10 Desktop Pool Creation

To create a VMware Horizon linked-clone Windows 10 Desktop Pool, follow these steps:

1. In Horizon Administrator console, select Desktop Pools in the Catalog node of the Inventory pane.

2. Click Add in the action pane.

3. Select assignment type for pool.
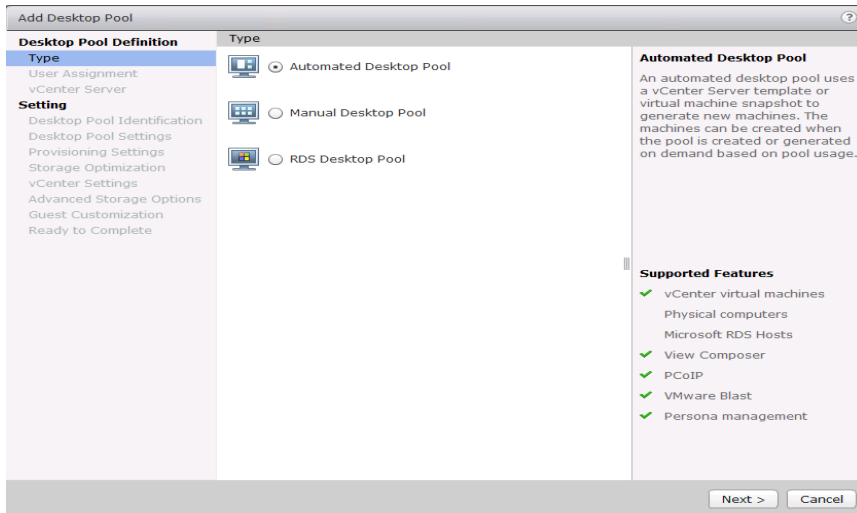
4. Click Next.

5. Select Floating or Dedicated user assignment.

6. Click Next.
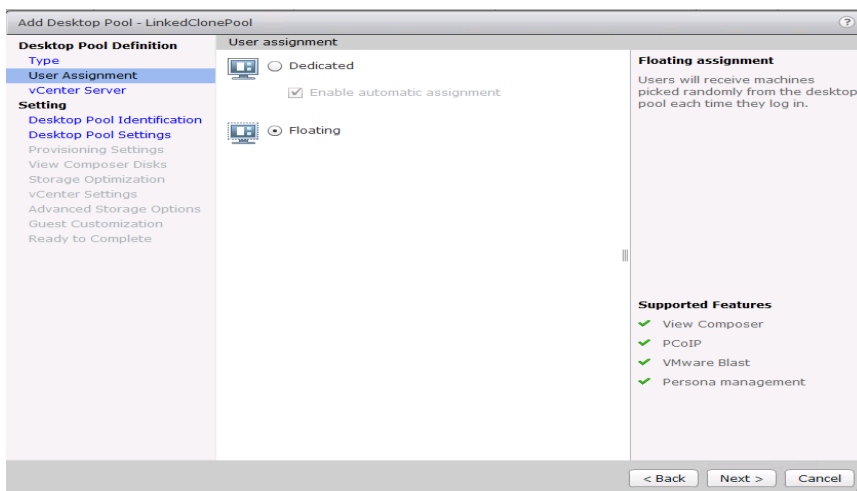
7. Select View Composer Linked Clones, highlight your vCenter and View Composer virtual machine.

8. Click Next.
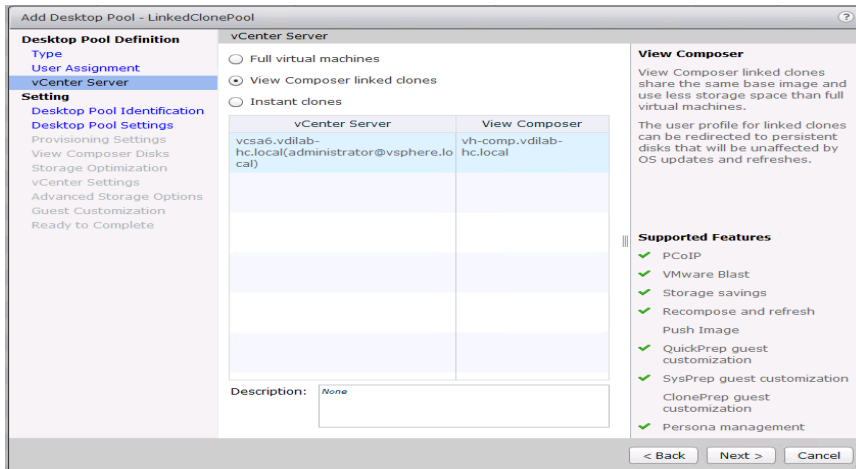
9. Enter pool identification details.

10. Click Next.
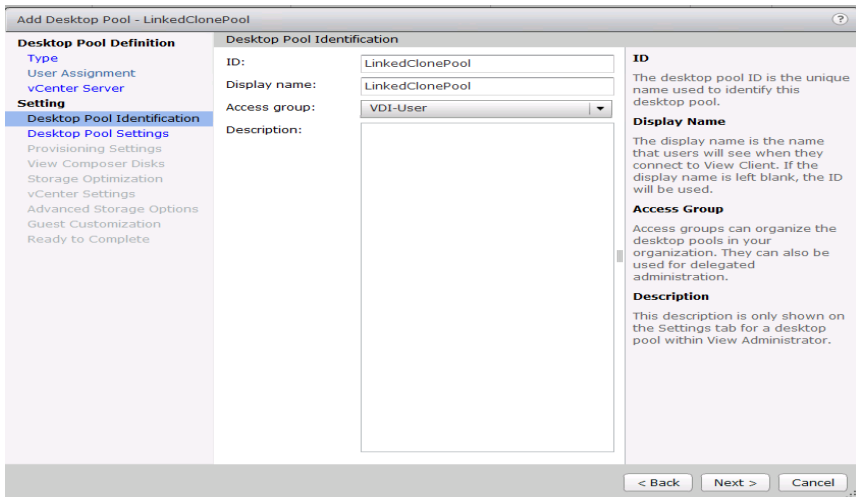
11. Select Desktop Pool settings.

> ⚠ Be sure to scroll down in this dialogue to configure all options.

12. Click Next.
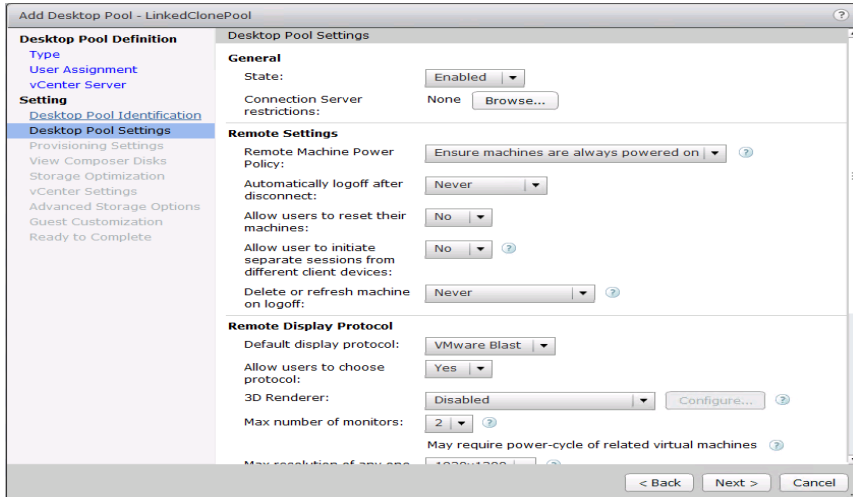
13. Select Provisioning Settings.

14. Click Next.
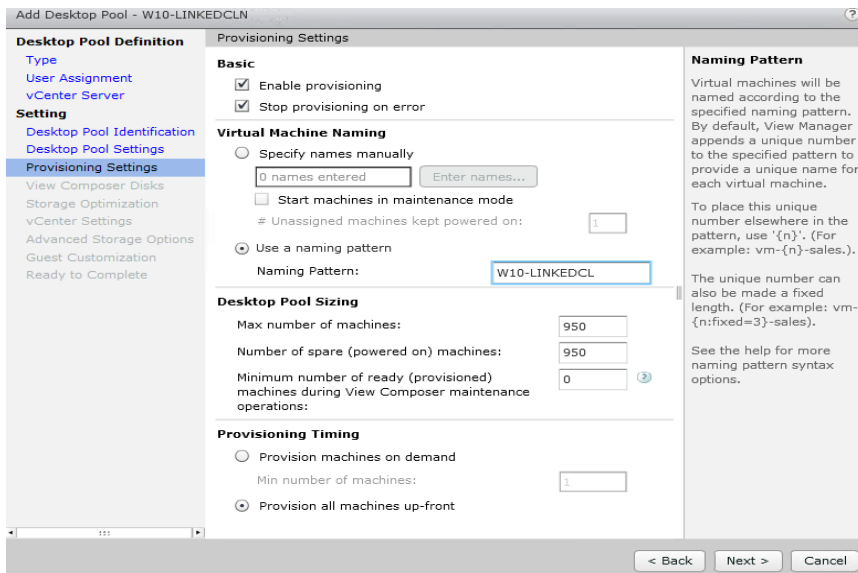
15. Select View Composer disk configuration.

16. Click Next.
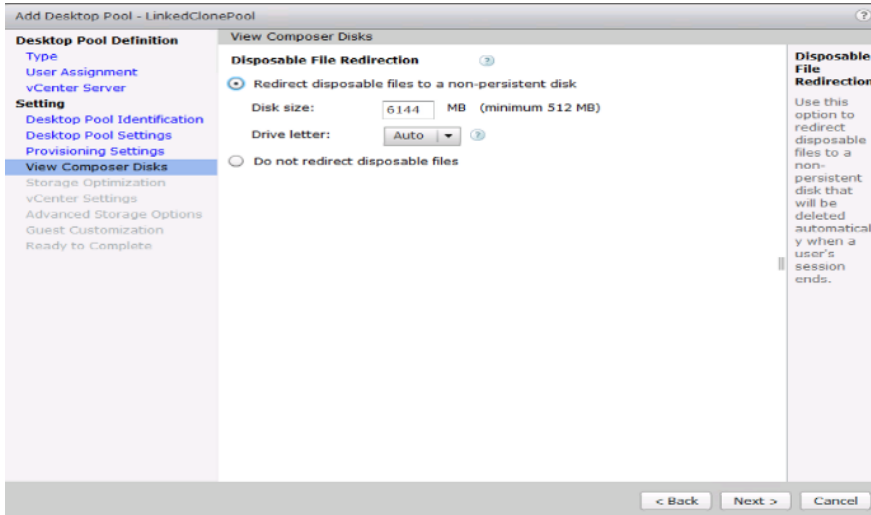
17. Click Next.
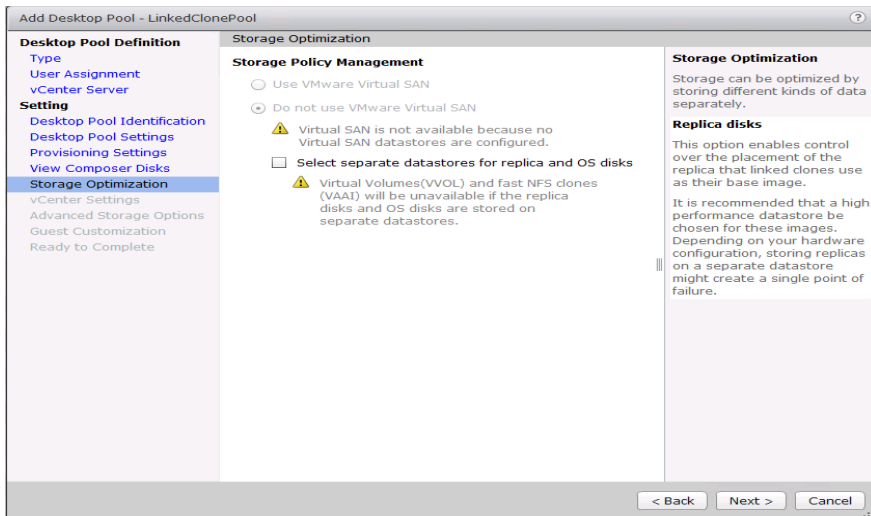
18. Select each of the six required vCenter Settings by using the Browse button next to each field.
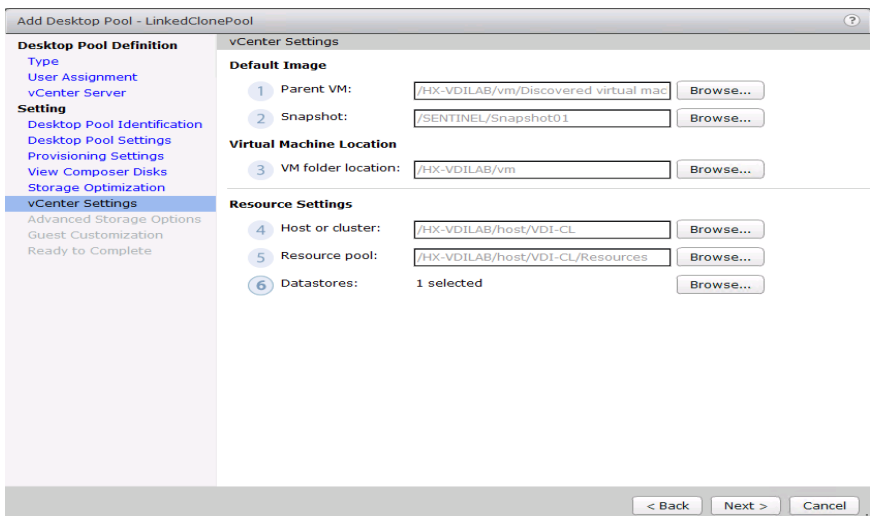
19. For Datastore selection, select the correct datastore and set the Storage Overcommit as "Unbounded."

20. Click OK.

21. Click Next.
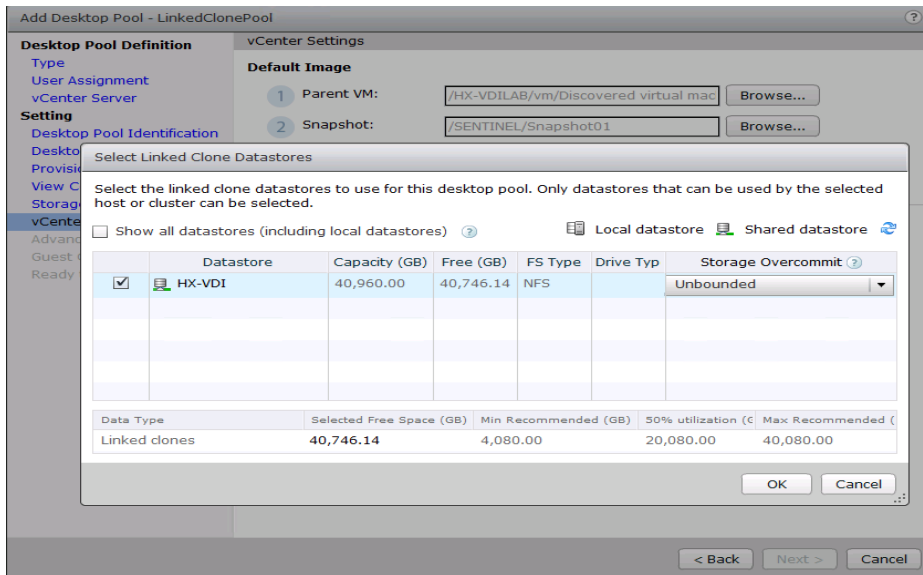
22. Set the Advanced Storage Options using the settings shown in the following screenshot.

23. Click Next.
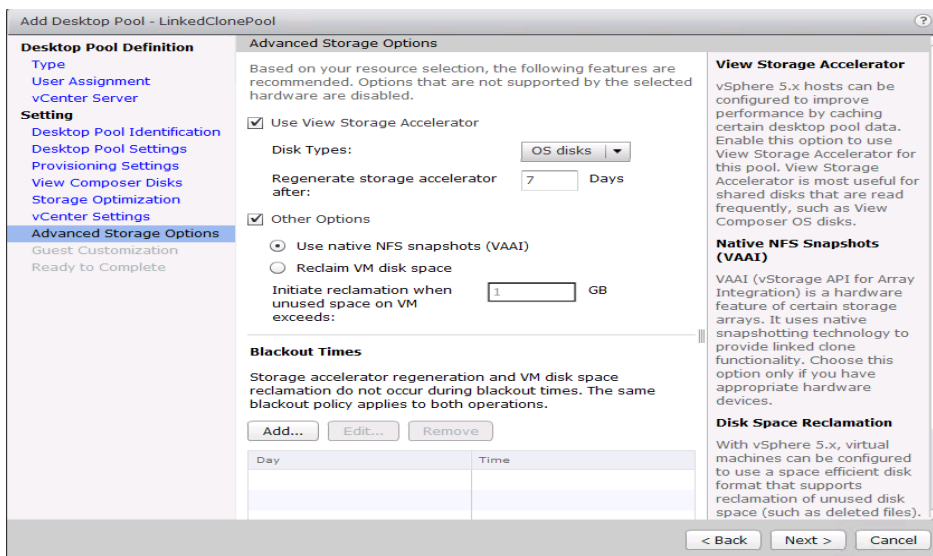
24. Select Guest optimization settings.

25. Select the Active Directory domain, browse to the Active Directory Container where the virtual machines will be provisioned and then choose either the QuickPrep or Sysprep option you would like to use. Highlight the Customization Spec previously prepared.

26. Click Next.
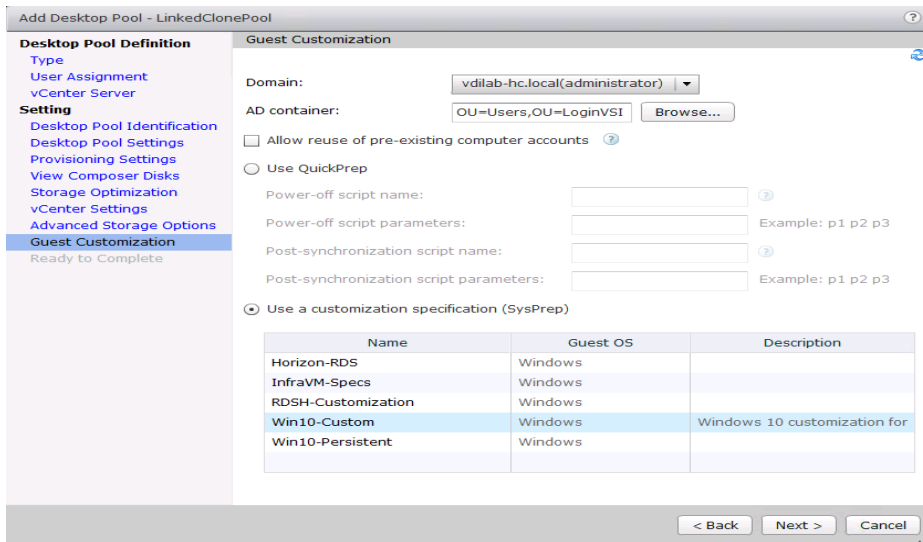
27. Select the checkbox "Entitle users after pool creation wizard completion" if you would like to authorize users as part of this process. Follow instructions provided in the Create Horizon 7 RDS Desktop Pool to authorize users for the Linked Clone Pool.

28. Click Finish to complete the Linked Clone Pool creation process.
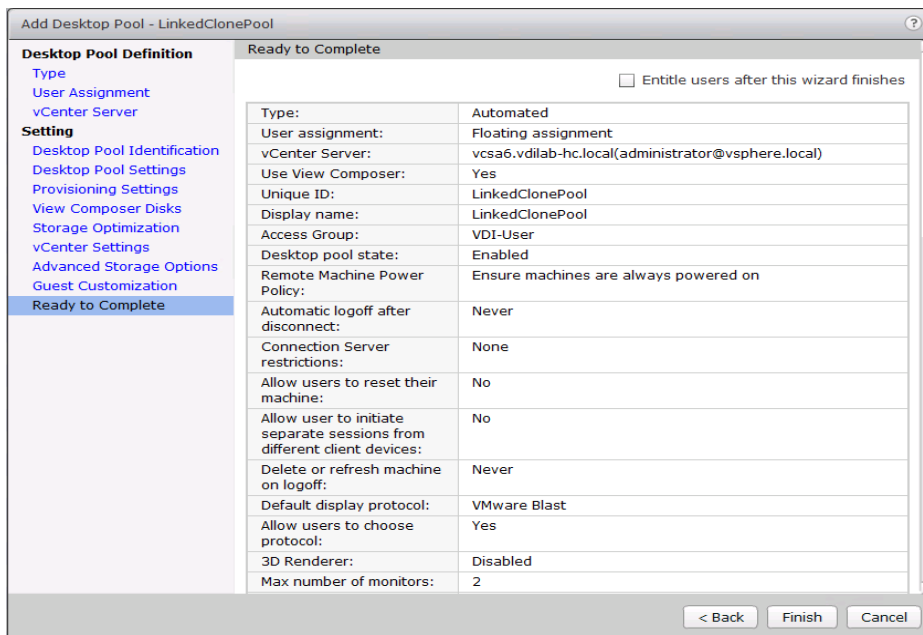
VMware Horizon Instant-Clone Windows 10 Desktop Pool Creation

To create the VMware Horizon Instant-Clone Windows 10 Desktop Pool, follow these steps:

1. In Horizon Administrator console, select Desktop Pools in the Catalog node of the Inventory pane.

2. Click Add in the action pane.

3. Select Automated assignment type for pool.
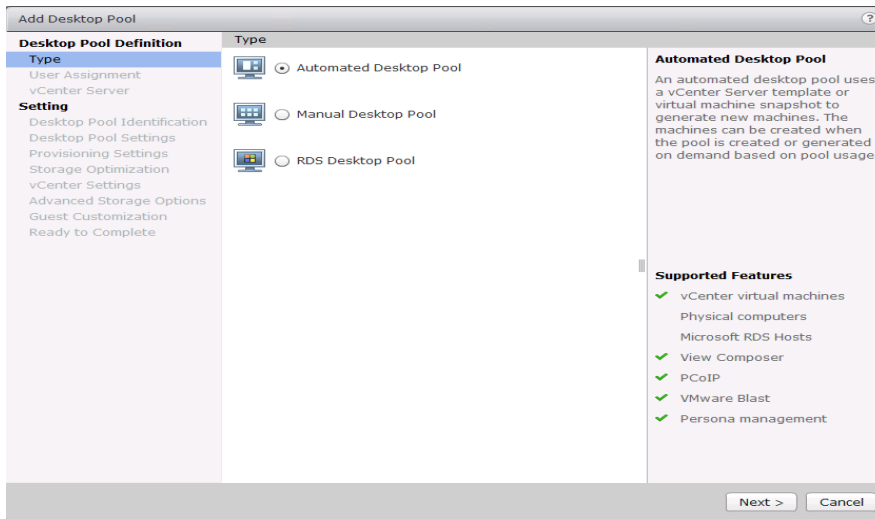
4. Click Next.

5. Select Floating or Dedicate user assignment.

6. Click Next.
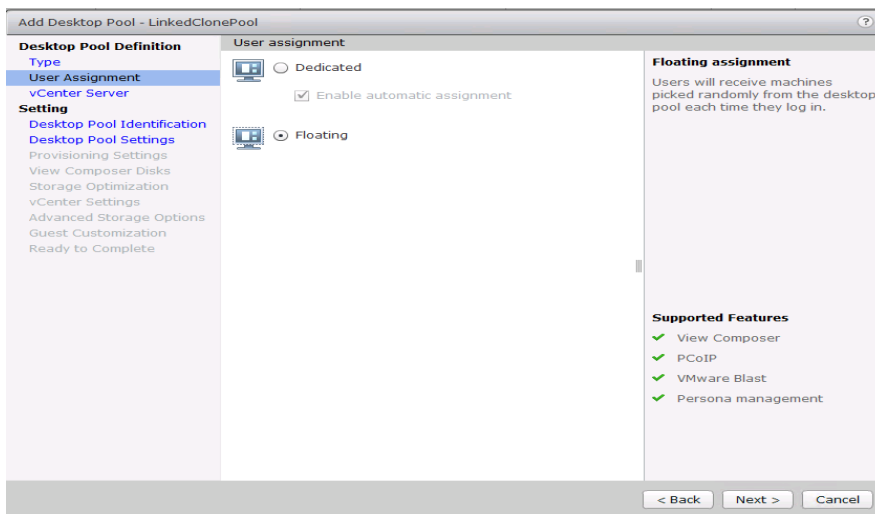
7. Select Instant Clones, highlight your vCenter server, then click Next.
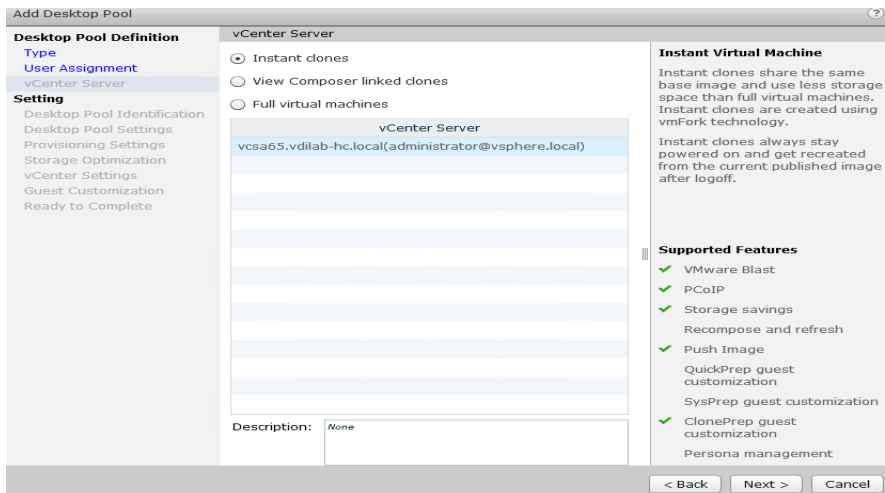
8. Enter pool identification details.

9. Click Next.
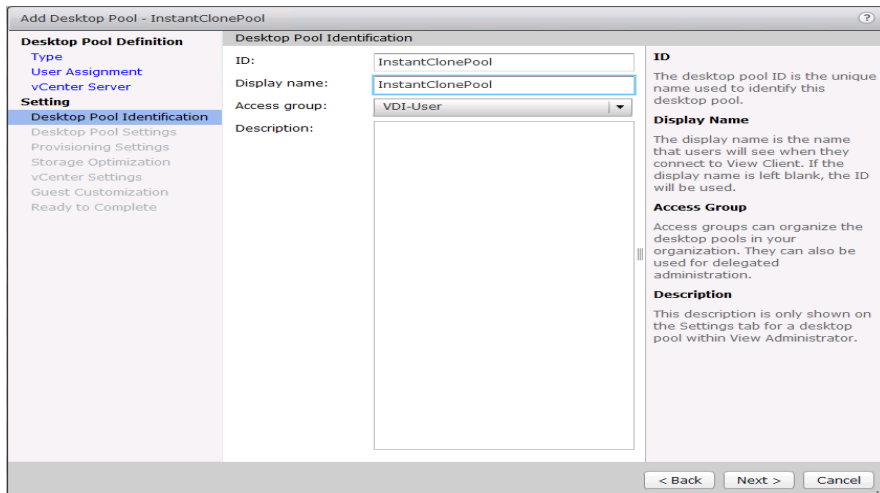
10. Select Desktop Pool settings.

> ⚠️ Be sure to scroll down to choose the Acrobat Flash settings.

11. Click Next.
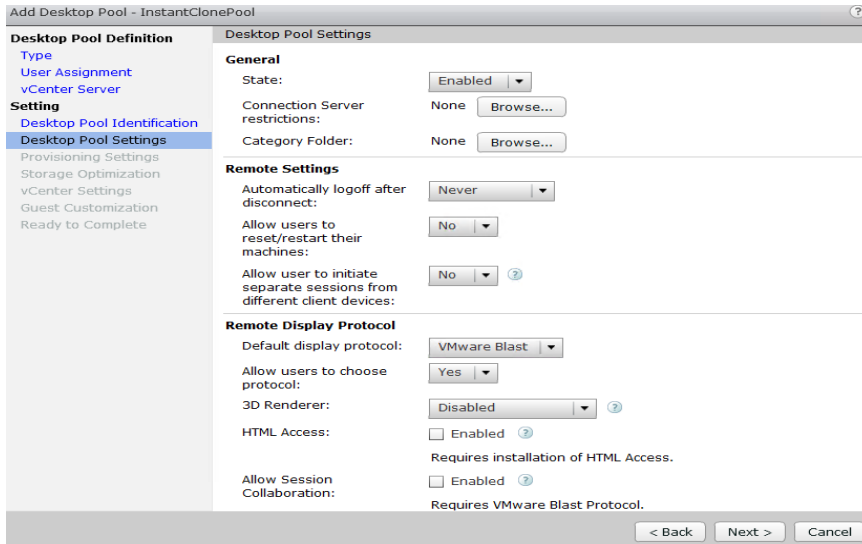
12. Select provisioning settings.

13. Click Next.
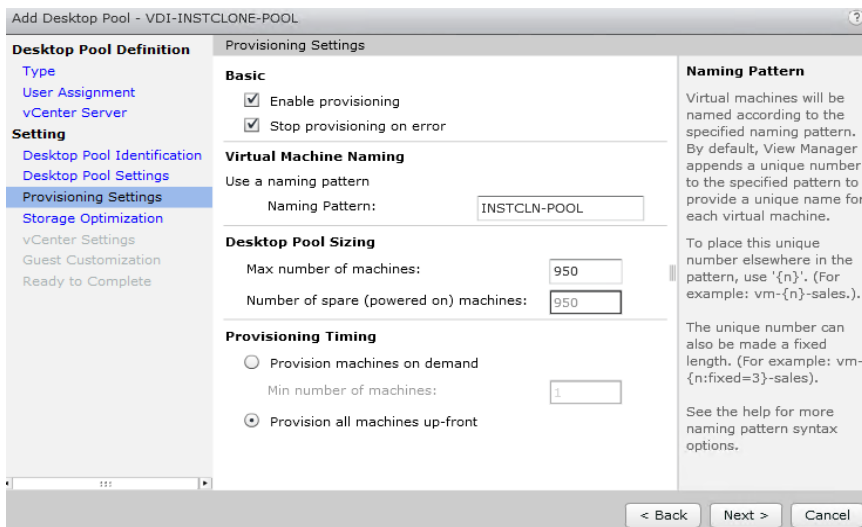
14. Click Next.
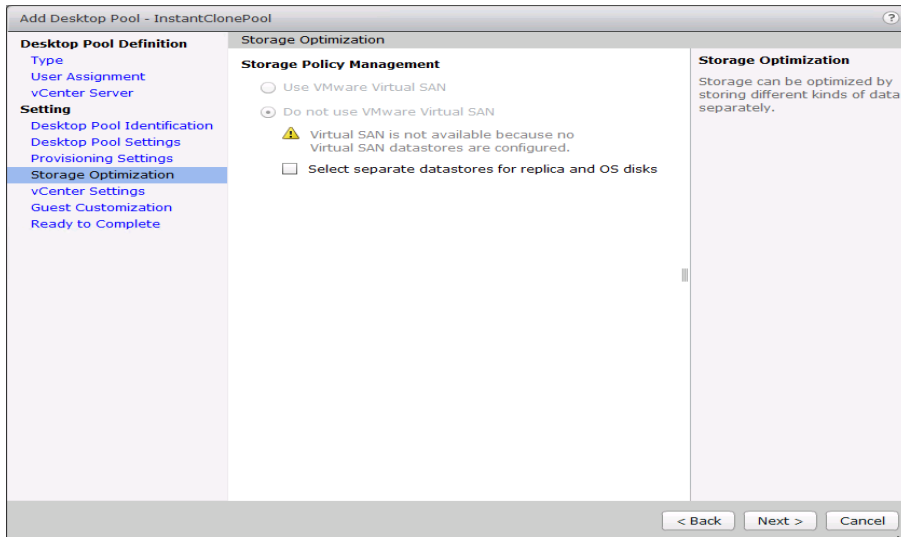
15. Select the vCenter Settings and browse for each of the six required inputs.

16. Click Next.
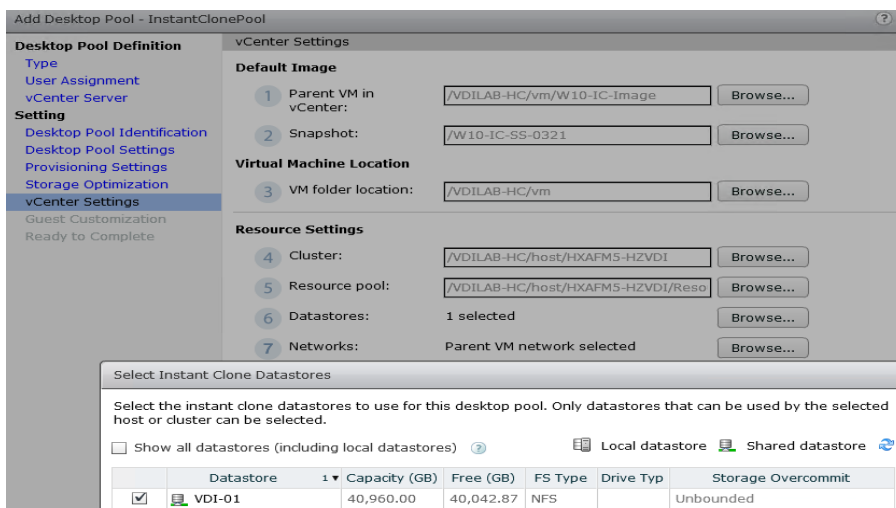
17. For Datastore selection, select the datastore with the storage overcommit as "Unbounded."

18. Click OK.

19. Select Guest Customization.

20. Browse to your Active Directory Domain and to the AD container into which you want your Instant Clone machines provisioned.

21. Click Next.
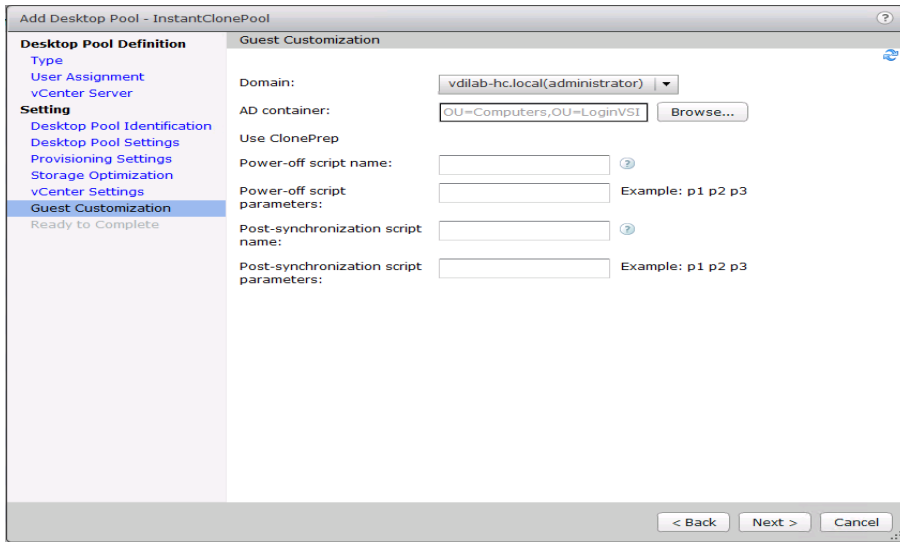
22. Review the summary of the pool configuration.

23. Select the checkbox "Entitle users after pool creation wizard completion" to authorize users or groups for the new pool.

24. Click Finish.
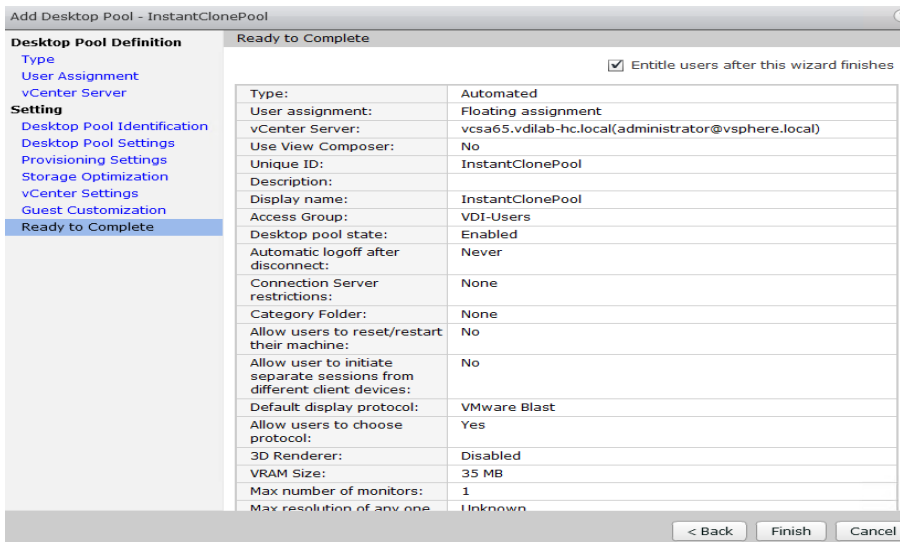
25. Follow the instructions provided in the Create Horizon 7 RDS Desktop Pool to authorize users for the Instant Clone Pool.

VMware Horizon Persistent Windows 10 Desktop Pool Creation

To create the VMware Horizon Persistent Windows 10 Desktop Pool, follow these steps:

1. In Horizon Administrator console, select Desktop Pools in the Catalog node of the Inventory pane.

2. Click Add in the action pane.

3. Select assignment type for pool.
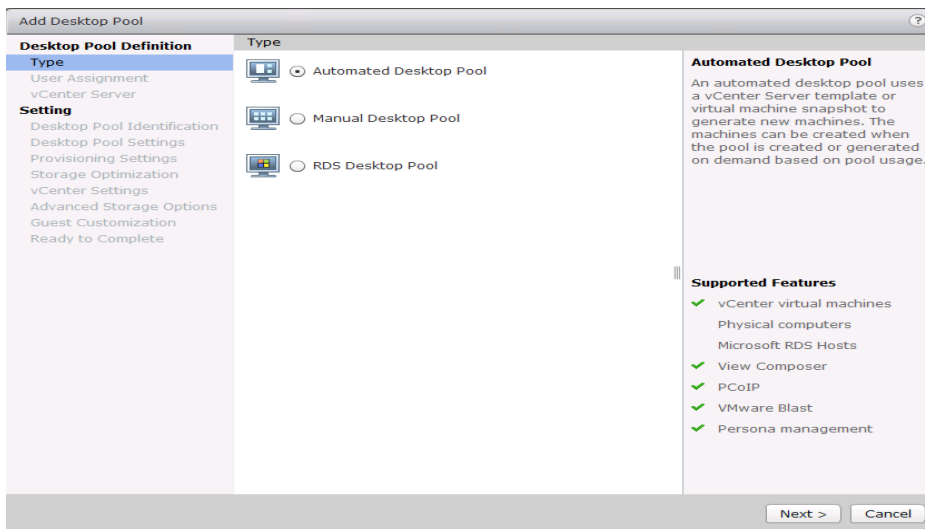
4. Click Next.

5. Select the Dedicated radio button.

6. Select the Enable automatic assignment checkbox, if desired.

7. Click Next.
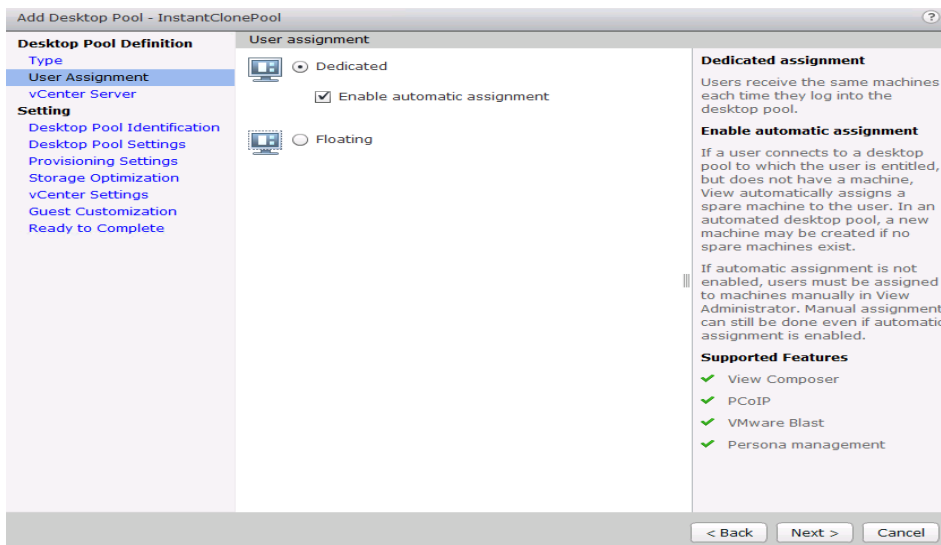
8. Select the Full Virtual Machines radio button and highlight your vCenter and Composer.

9. Click Next.
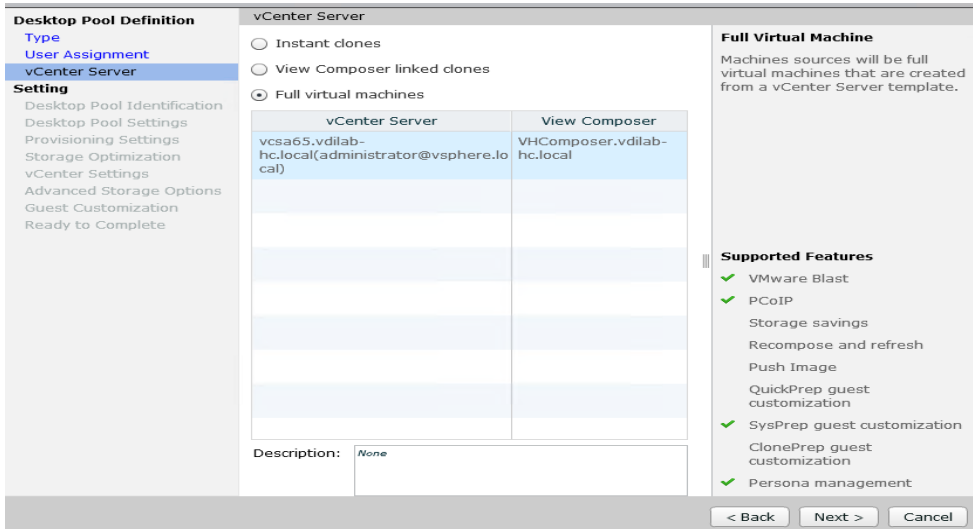
10. Enter the pool identification details.

11. Click Next.
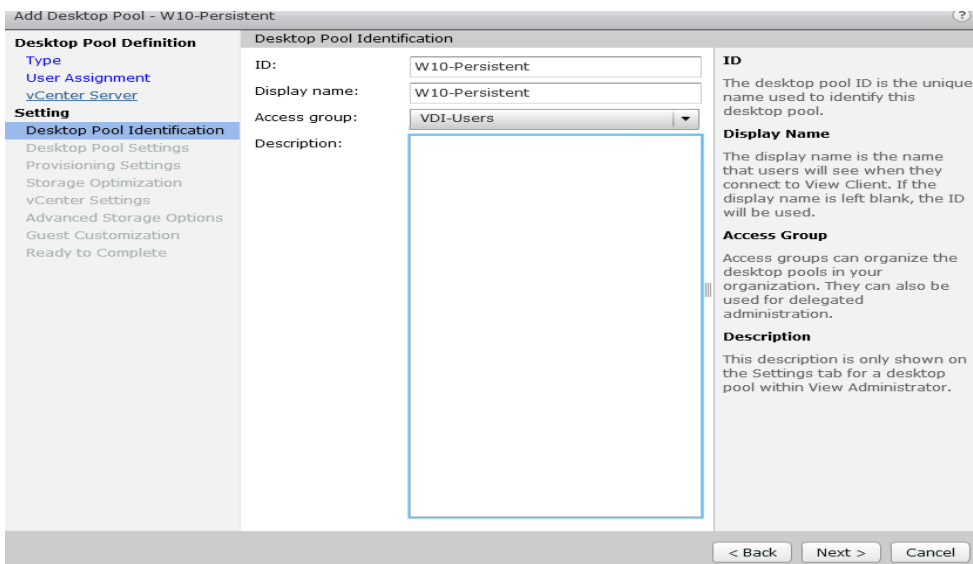
12. Select Desktop Pool settings.

13. Click Next.
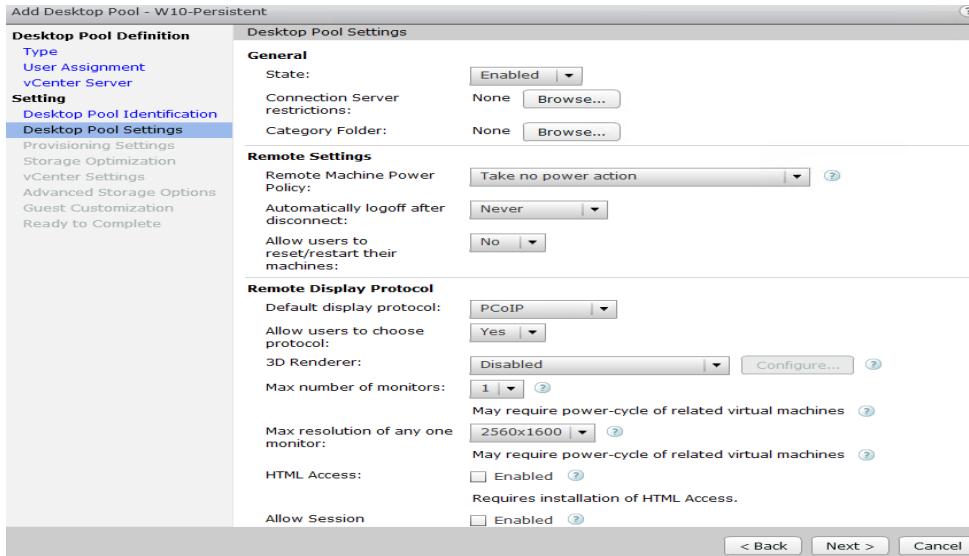
14. Select the provisioning settings to meet your requirements.

15. Click Next.
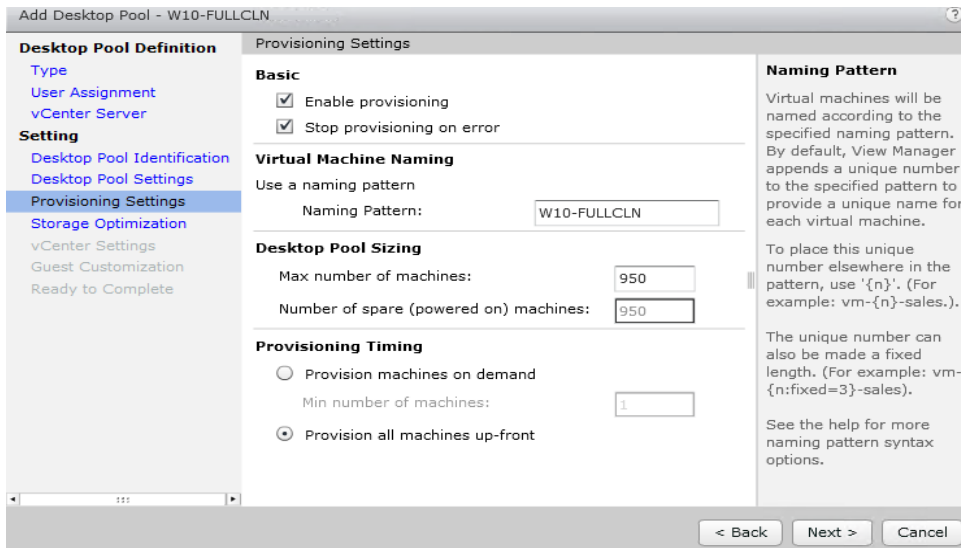
16. Click Next.
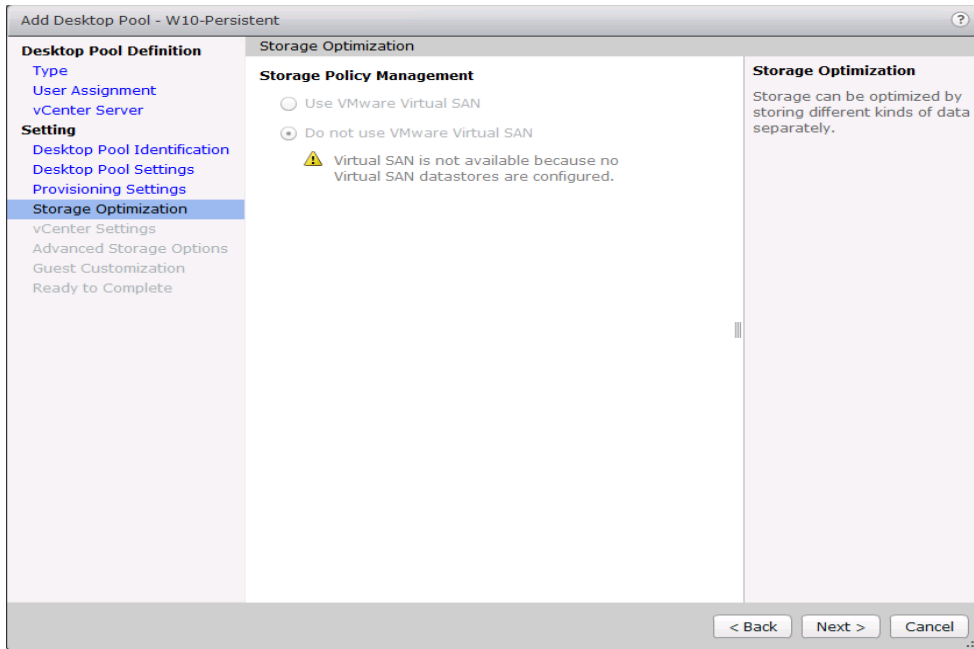
17. Select each of the five vCenter Settings.

18. Click Next.
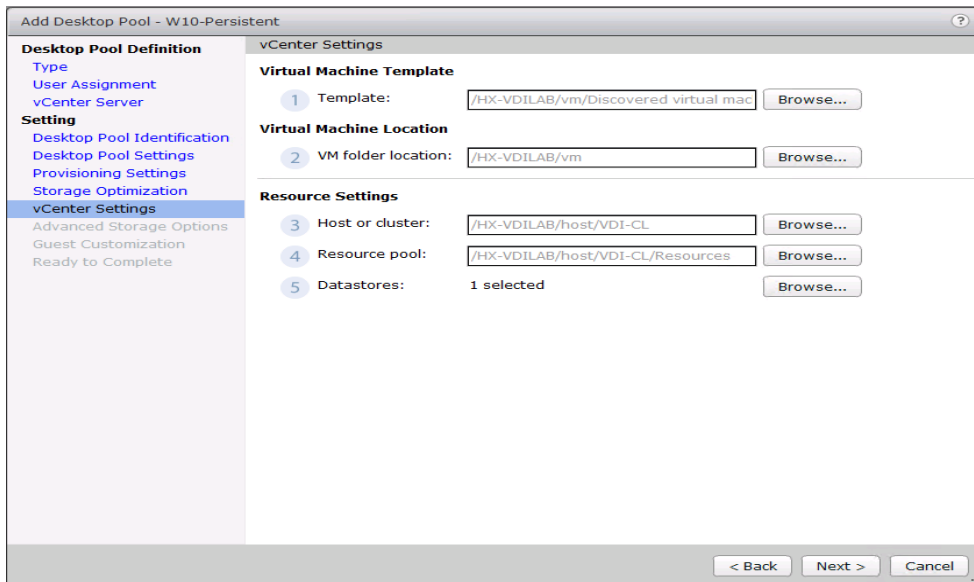
19. For Datastore selection, select the datastore with storage overcommit as "Unbounded."

20. Click OK.
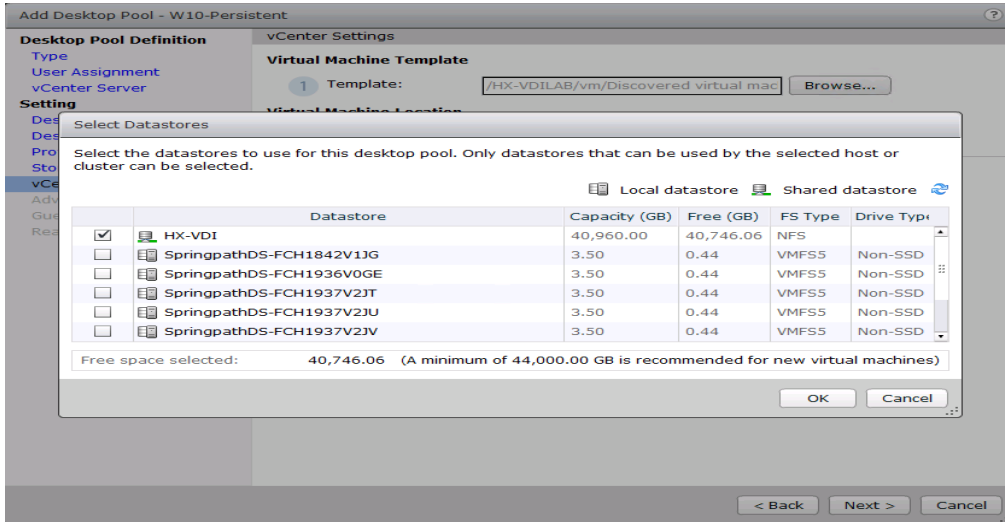
21. Select Advance Storage Options and enable the View Storage Accelerator.

22. Click Next.
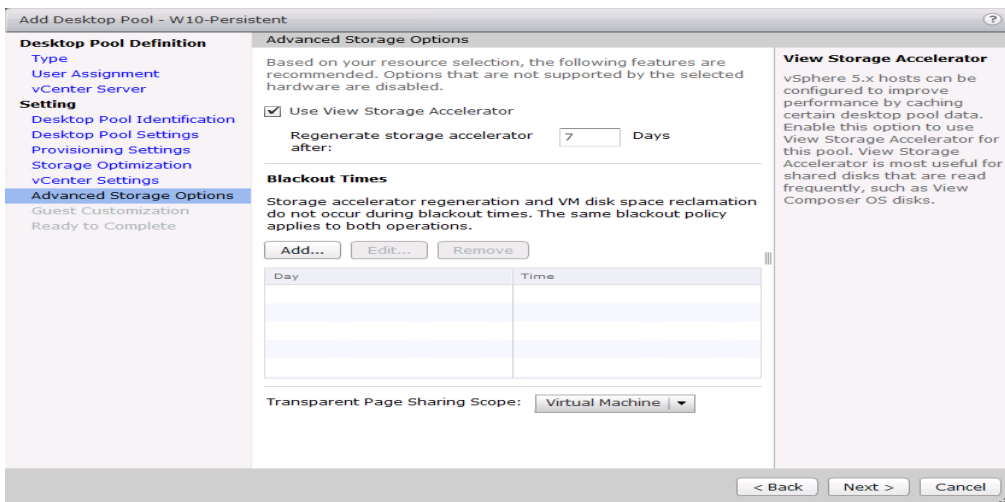
23. Select Guest optimization settings.

24. Click Next.
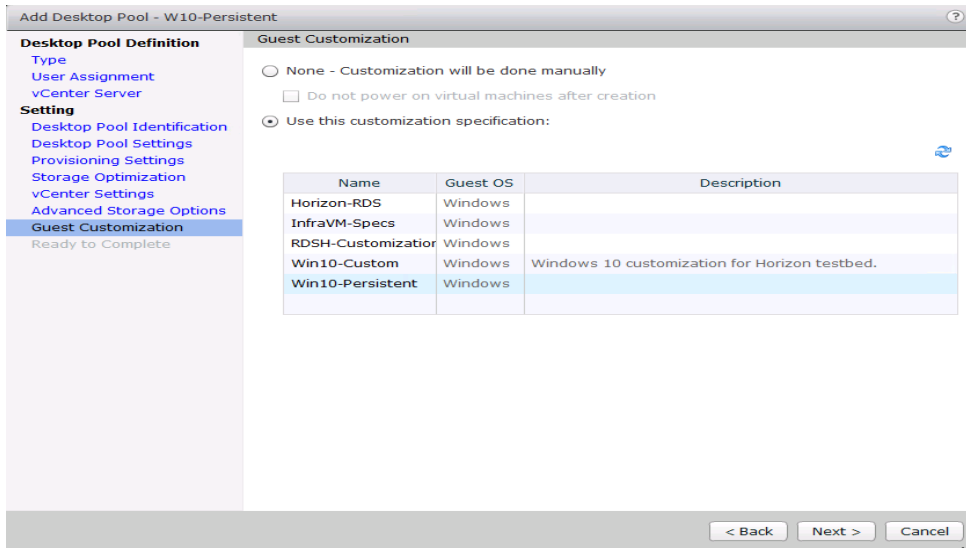
25. Review the summary of the pool you are creating.

26. Select the checkbox "Entitle users after pool creation wizard completion" to authorize users for the pool.

27. Click Finish.
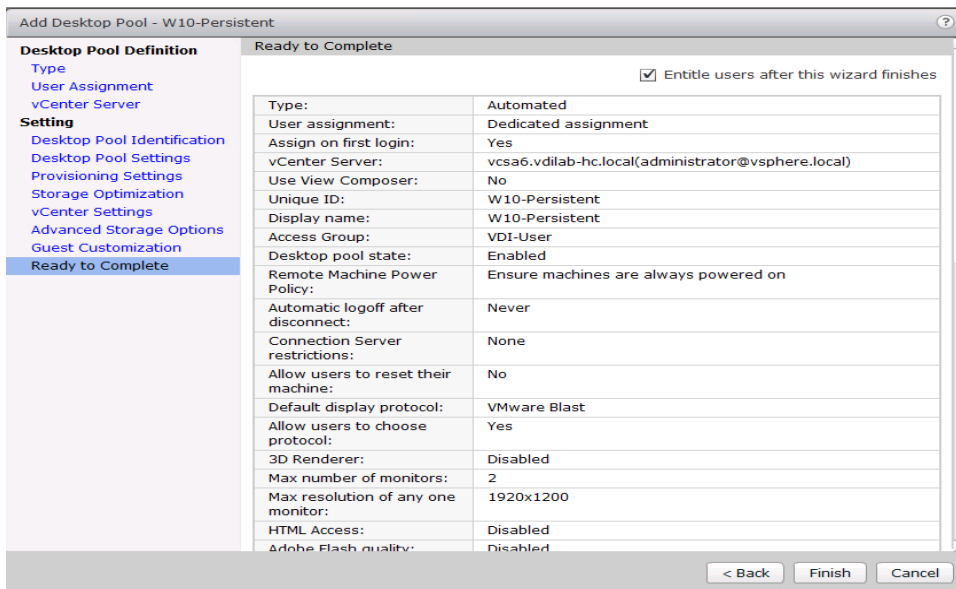
28. Follow the instructions provided in the Create Horizon 7 RDS Desktop Pool to authorize users for the Linked Clone Pool.

Test Setup and Configurations

In this project, we tested a single Cisco HyperFlex cluster running 8 Cisco UCS HXAF220C-M5SX with 8 compute-only Cisco UCS C220 M5 Rack Servers and Cisco UCS 8 B200 M5 Blade Servers.

24 Node Cisco HyperFlex and VMware Horizon RDSH & Desktops,
Two to One (2:1) Compute to Converge Node Ratio, UCS Domain Reference
Architecture

Cisco Nexus 93108YC-FX   Peer-Link   Cisco Nexus 93108YC-FX

Cisco UCS-FI-6454   vPC   vPC   Cisco UCS-FI-6454

8 x 25G   8 x 25G

8 x Cisco HXAF220C-M5SX or
Cisco HXAF240C-M5SX Rack Servers
Each server includes:
- 2 x Intel® Xeon® Gold 6230 scalable family
  processor @ 2.3 GHz processor
- 240GB M.2 SATA SSD
- 768 GB (24 X 32GB DDR4) RAM @ 2666MHz
- 1 x Cisco VIC 1457 mLOM
- 1 x Cisco 12Gbps Modular SAS HBA
- 1 x 240 GB Intel SATA Enterprise Value SSD

HX Datastore configured to host Windows
10 desktop pool or Server 2019 RDSH Pool
with drives listed below per node:
- 1 x 400GB Toshiba Enterprise
  Performance (EP) SSD for Cache
- 8 x 960GB Samsung SATA Enterprise
  Value SSDs for capacity

HX
Datastore

8 x 25G   8 x 25G

8 x Cisco UCS C220 M5 Rack Servers
8 x Cisco UCS B200 M5 Blade Servers
Each server includes:
- 2 x Intel® Xeon® Gold 6230 scalable family
  processor @ 2.3 GHz processor
- 120Gb M.2 SATA SSD
- 768 GB (24 X 32GB DDR4) RAM @ 2666MHz
- 1 x Cisco VIC 1457 mLOM
- 1 x Cisco VIC 1440

2 x 25G   2 x 25G

Hardware Components:

- 2 x Cisco UCS 6454 Fabric Interconnects

- 2 x Cisco Nexus 93108YCPX Access Switches

- 8 x Cisco UCS HXAF220c-M5SX Rack Servers (2 Intel Xeon Gold 6230 scalable family processor at 2.1 GHz, with 768 GB of memory per server [64 GB x 12 DIMMs at 2933 MHz])

- 8 x Cisco UCS C220 M5 Rack Servers (2 Intel Xeon Gold 6230 scalable family processor at 2.1 GHz, with 768 GB of memory per server [64 GB x 12 DIMMs at 2933 MHz])

- 8 x Cisco UCS B200 M5 Blade Servers (2 Intel Xeon Gold 6230 scalable family processor at 2.1 GHz, with 768 GB of memory per server [64 GB x 12 DIMMs at 2933 MHz])

- Cisco VIC 1457 mLOM

- 12G modular SAS HBA Controller

- 240GB M.2 SATA SSD drive (Boot and HyperFlex Data Platform controller virtual machine)

- 240GB 2.5" 6G SATA SSD drive (Housekeeping)

- 400GB 2.5" 6G SAS SSD drive (Cache)

- 6 x 960GB 2.5" SATA SSD drive (Capacity)
- 1 x 32GB mSD card (Upgrades temporary cache)

Software Components:

- Cisco UCS firmware 4.0(4g)
- Cisco HyperFlex Data Platform 4.0.2a
- VMWare ESXi 6.7.0 15160138
- VMware Virtual Desktops 7.12 VMware Horizon 7.12
- VMware User Profile Management
- Microsoft SQL Server 2019
- Microsoft Windows 10, Build 1809
- Microsoft Windows Server 2019
- Microsoft Office 2016
- Login VSI 4.1.32

Test Methodology and Success Criteria

All validation testing was conducted on-site within the Cisco labs in San Jose, California.

The testing results focused on the entire process of the virtual desktop lifecycle by capturing metrics during the desktop boot-up, user logon and virtual desktop acquisition (also referred to as ramp-up,) user workload execution (also referred to as steady state), and user logoff for the Hosted Shared Desktop Session under test.

Test metrics were gathered from the virtual desktop, storage, and load generation software to assess the overall success of an individual test cycle. Each test cycle was not considered passing unless all of the planned test users completed the ramp-up and steady state phases (described below) and unless all metrics were within the permissible thresholds as noted as success criteria.

Three successfully completed test cycles were conducted for each hardware configuration and results were found to be relatively consistent from one test to the next.

You can obtain additional information and a free test license from http://www.loginvsi.com.

Test Procedure

The following protocol was used for each test cycle in this study to insure consistent results.

Pre-Test Setup for Testing

All machines were shut down utilizing the VMware Horizon Administrator Console.

All Launchers for the test were shut down. They were then restarted in groups of 10 each minute until the required number of launchers was running with the Login VSI Agent at a "waiting for test to start" state.

Test Run Protocol

To simulate severe, real-world environments, Cisco requires the log-on and start-work sequence, known as Ramp Up, to complete in 48 minutes. Additionally, we require all sessions started, whether 60 single server users or 4000 full scale test users to become active within two minutes after the last session is launched.

In addition, Cisco requires that the Login VSI Benchmark method is used for all single server and scale testing. This assures that our tests represent real-world scenarios. For each of the three consecutive runs on single server tests, the same process was followed. To run the test protocol, follow these steps:

1. Time 0:00:00 Start esxtop Logging on the following systems:

    ◦ Infrastructure and VDI Host Blades used in test run
    ◦ All Infrastructure virtual machines used in test run (AD, SQL, View Connection brokers, image mgmt., and so on)

2. Time 0:00:10 Start Storage Partner Performance Logging on Storage System.

3. Time 0:05: Boot RDS Machines using VMware Horizon Administrator Console.

4. Time 0:06 First machines boot.

5. Time 0:35 Single Server or Scale target number of RDS Servers registered on XD.

> **⚠** No more than 60 minutes of rest time is allowed after the last desktop is registered and available on VMware Horizon Administrator Console dashboard. Typically, a 20-30-minute rest period for Windows 10 desktops and 10 minutes for RDS virtual machines is sufficient.

6. Time 1:35 Start Login VSI 4.1.32 Knowledge Worker Benchmark Mode Test, setting auto-logoff time at 900 seconds, with Single Server or Scale target number of desktop virtual machines utilizing sufficient number of Launchers (at 20-25 sessions/Launcher).

7. Time 2:23 Single Server or Scale target number of desktop virtual machines desktops launched (48-minute benchmark launch rate).

8. Time 2:25 All launched sessions must become active.

> **⚠** All sessions launched must become active for a valid test run within this window.

9. Time 2:40 Login VSI Test Ends (based on Auto Logoff 900 Second period designated above).

10. Time 2:55 All active sessions logged off.

11. All sessions launched and active must be logged off for a valid test run. The VMware Administrator Dashboard must show that all desktops have been returned to the registered/available state as evidence of this condition being met.

12. Time 2:57 All logging terminated; Test complete.

13. Time 3:15 Copy all log files off to archive; Set virtual desktops to maintenance mode through broker; Shutdown all Windows 7 machines.

14. Time 3:30 Reboot all hypervisors.

15. Time 3:45 Ready for new test sequence.

Success Criteria

Our "pass" criteria for this testing is as follows: Cisco will run tests at a session count levels that effectively utilize the server capacity measured by CPU, memory, storage, and network utilization. We use Login VSI version 4.1.25 to launch Knowledge Worker workload sessions. The number of launched sessions must equal active sessions within two minutes of the last session launched in a test as observed on the VSI Management console.

The VMware Virtual Desktops Studio will be monitored throughout the steady state to make sure of the following:

- All running sessions report In Use throughout the steady state

- No sessions move to unregistered, unavailable, or available state at any time during steady state

Within 20 minutes of the end of the test, all sessions on all launchers must have logged out automatically and the Login VSI Agent must have shut down. Cisco's tolerance for Stuck Sessions is 0.5 percent (half of one percent.) If the Stuck Session count exceeds that value, we identify it as a test failure condition.

Cisco requires three consecutive runs with results within +/-1 percent variability to pass the Cisco Validated Design performance criteria. For white papers written by partners, two consecutive runs within +/-1 percent variability are accepted. (All test data from partner run testing must be supplied along with proposed white paper.)

We will publish Cisco Validated Designs with our recommended workload following the process described and will note that we did not reach a VSImax dynamic in our testing.

The purpose of this testing is to provide the data needed to validate VMware Remote Desktop Server Hosted (RDSH) server using Microsoft Server 2019 and Virtual Desktops using Microsoft Windows 10 with VMware Horizon provisioning on Cisco UCS HXAF220c-M4S, Cisco UCS 220 M4, and Cisco UCS B200 M4 servers.

The information in this section provides data points that a customer may reference in designing their own implementations. These validation results are an example of what is possible under the specific environment conditions outlined here and do not represent the full characterization of VMware and Microsoft products.

Four test sequences, each containing three consecutive test runs generating the same result, were performed to establish system performance and linear scalability.

VSImax 4.1.x Description

The philosophy behind Login VSI is different to conventional benchmarks. In general, most system benchmarks are steady state benchmarks. These benchmarks execute one or multiple processes, and the measured execution time is the outcome of the test. Simply put: the faster the execution time or the bigger the throughput, the faster the system is according to the benchmark.

Login VSI is different in approach. Login VSI is not primarily designed to be a steady state benchmark (however, if needed, Login VSI can act like one). Login VSI was designed to perform benchmarks for SBC or VDI workloads through system saturation. Login VSI loads the system with simulated user workloads using well known desktop applications like Microsoft Office, Internet Explorer, and Adobe PDF reader. By gradually increasing the number of simulated users, the system will eventually be saturated. Once the system is saturated, the response time of the applications will increase significantly. This latency in application response times show a clear indication whether the system is (close to being) overloaded. As a result, by nearly overloading a system it is possible to find out what is its true maximum user capacity.

After a test is performed, the response times can be analyzed to calculate the maximum active session/desktop capacity. Within Login VSI this is calculated as VSImax. When the system is coming closer to its saturation point, response times will rise. When reviewing the average response time, it will be clear the response times escalate at saturation point.

This VSImax is the "Virtual Session Index (VSI)". With Virtual Desktop Infrastructure (VDI) and Terminal Services (RDS) workloads this is valid and useful information. This index simplifies comparisons and makes it possible to understand the true impact of configuration changes on hypervisor host or guest level.

Server-Side Response Time Measurements

It is important to understand why specific Login VSI design choices have been made. An important design choice is to execute the workload directly on the target system within the session instead of using remote sessions. The scripts simulating the workloads are performed by an engine that executes workload scripts on every **target system and are initiated at logon within the simulated user's desktop session context.**

An alternative to the Login VSI method would be to generate user actions client side through the remoting protocol. These methods are always specific to a product and vendor dependent. More importantly, some protocols simply do not have a method to script user actions client side.

For Login VSI the choice has been made to execute the scripts completely server side. This is the only practical and platform independent solutions, for a benchmark like Login VSI.

Calculating VSImax v4.1.x

The simulated desktop workload is scripted in a 48-minute loop when a simulated Login VSI user is logged on, performing generic Office worker activities. After the loop is finished it will restart automatically. Within each loop the response times of sixteen specific operations are measured in a regular interval: sixteen times in within each loop. The response times of these five operations are used to determine VSImax.

The five operations from which the response times are measured are:

- **Notepad File Open (NFO)**
  Loading and initiating VSINotepad.exe and opening the openfile dialog. This operation is handled by the OS and by the VSINotepad.exe itself through execution. This operation seems almost instant from an end-user's point of view.

- **Notepad Start Load (NSLD)**
  Loading and initiating VSINotepad.exe and opening a file. This operation is also handled by the OS and by the VSINotepad.exe itself through execution. This operation seems almost instant from an end-user's point of view.

- **Zip High Compression (ZHC)**
  This action copy's a random file and compresses it (with 7zip) with high compression enabled. The compression will very briefly spike CPU and disk IO.

- **Zip Low Compression (ZLC)**
  This action copy's a random file and compresses it (with 7zip) with low compression enabled. The compression will very briefly disk IO and creates some load on the CPU.

- **CPU**
  Calculates a large array of random data and spikes the CPU for a short period of time.

  These measured operations within Login VSI do hit considerably different subsystems such as CPU (user and kernel), Memory, Disk, the OS in general, the application itself, print, GDI, and so on. These operations are specifically short by nature. When such operations become consistently long: the system is saturated because of excessive queuing on any kind of resource. As a result, the average response times will then escalate. This effect is clearly visible to end-users. If such operations consistently consume multiple seconds the user will regard the system as slow and unresponsive.

Figure 51 Sample of a VSI Max Response Time Graph, Representing a Normal Test
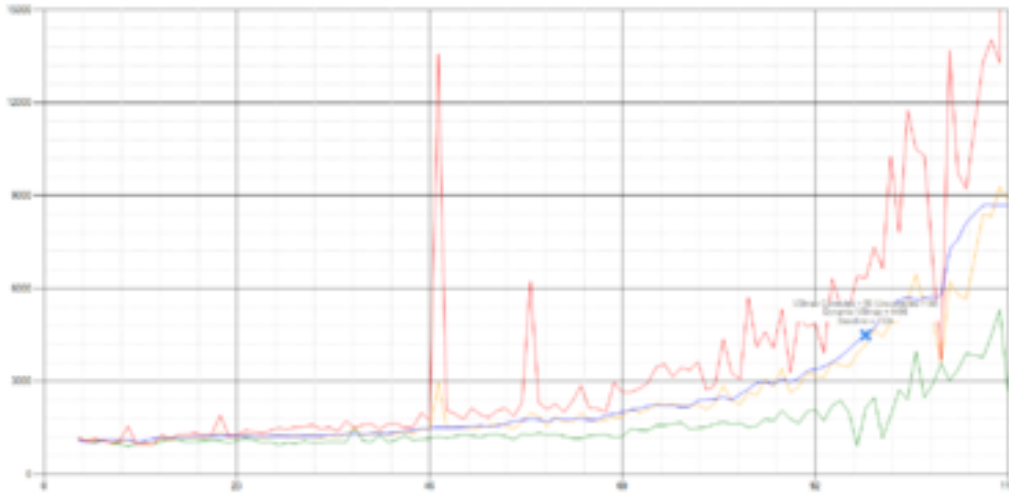
Figure 52 Sample of a VSI Test Response Time Graph with a Clear Performance Issue
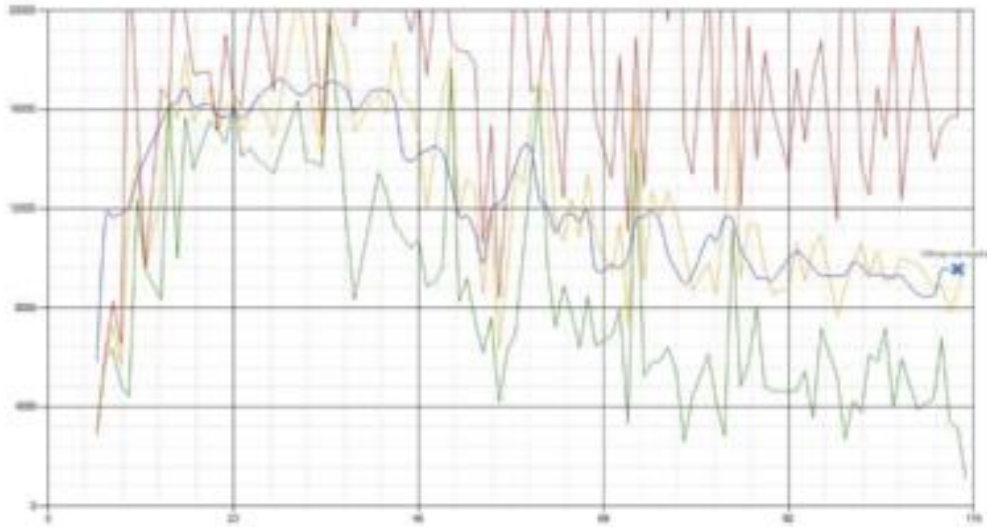
When the test is finished, VSImax can be calculated. When the system is not saturated, and it could complete the full test without exceeding the average response time latency threshold, VSImax is not reached and the amount of sessions ran successfully.

The response times are very different per measurement type, for instance Zip with compression can be around 2800 ms, while the Zip action without compression can only take 75ms. This response time of these actions are weighted before they are added to the total. This ensures that each activity has an equal impact on the total response time.

In comparison to previous VSImax models, this weighting much better represents system performance. All actions have very similar weight in the VSImax total. The following weighting of the response times are applied.

The following actions are part of the VSImax v4.1 calculation and are weighted as follows (US notation):

- Notepad File Open (NFO): 0.75
- Notepad Start Load (NSLD): 0.2
- Zip High Compression (ZHC): 0.125
- Zip Low Compression (ZLC): 0.2
- CPU: 0.75

This weighting is applied on the baseline and normal Login VSI response times.

With the introduction of Login VSI 4.1 we also created a new method to calculate the base phase of an environment. With the new workloads (Taskworker, Powerworker, etc.) enabling 'base phase' for a more reliable baseline has become obsolete. The calculation is explained below. In total 15 lowest VSI response time samples are taken from the entire test, the lowest 2 samples are removed, and the 13 remaining samples are averaged. The result is the Baseline. The calculation is as follows:

- Take the lowest 15 samples of the complete test

- From those 15 samples remove the lowest 2

- Average the 13 results that are left is the baseline

The VSImax average response time in Login VSI 4.1.x is calculated on the number of active users that are logged on the system.

Always a 5 Login VSI response time samples are averaged + 40 percent **of the amount of "active" sessions. For** example, if the active sessions are 60, then latest 5 + 24 (=40 percent of 60) = 31 response time measurement are used for the average calculation.

To remove noise (accidental spikes) from the calculation, the top 5 percent and bottom 5 percent of the VSI response time samples are removed from the average calculation, with a minimum of 1 top and 1 bottom sample. As a result, with 60 active users, the last 31 VSI response time sample are taken. From those 31 samples the top 2 samples are removed and lowest 2 results are removed (5 percent of 31 = 1.55, rounded to 2). At 60 users the average is then calculated over the 27 remaining results.

VSImax v4.1.x is reached when the VSIbase + a 1000 ms latency threshold is not reached by the average VSI response time result. Depending on the tested system, VSImax response time can grow 2 - 3x the baseline average. In end-user computing, a 3x increase in response time in comparison to the baseline is typically regarded as the maximum performance degradation to be considered acceptable.

In VSImax v4.1.x this latency threshold is fixed to 1000ms, this allows better and fairer comparisons between two different systems, especially when they have different baseline results. Ultimately, in VSImax v4.1.x, the performance of the system is not decided by the total average response time, but by the latency is has under load. For all systems, this is now 1000ms (weighted).

The threshold for the total response time is average weighted baseline response time + 1000ms.

When the system has a weighted baseline response time average of 1500ms, the maximum average response time may not be greater than 2500ms (1500+1000). If the average baseline is 3000 the maximum average response time may not be greater than 4000ms (3000+1000).

When the threshold is not exceeded by the average VSI response time during the test, VSImax is not hit and the amount of sessions ran successfully. This approach is fundamentally different in comparison to previous VSImax methods, as it was always required to saturate the system beyond VSImax threshold.

Lastly, VSImax v4.1.x is now always reported with the average baseline VSI response time result. For example: **"The VSImax v4.1 was 125 with a baseline of 1526ms". This helps considerably in the comparison of systems** and gives a more complete understanding of the system. The baseline performance helps to understand the best performance the system can give to an individual user. VSImax indicates what the total user capacity is for the system. These two are not automatically connected and related:

When a server with a very fast dual core CPU, running at 3.6 GHZ, is compared to a 10 core CPU, running at 2,26 GHZ, the dual core machine will give and individual user better performance than the 10-core machine. This is indicated by the baseline VSI response time. The lower this score is, the better performance an individual user can expect.

However, the server with the slower 10 core CPU will easily have a larger capacity than the faster dual core system. This is indicated by VSImax v4.1.x, and the higher VSImax is, the larger overall user capacity can be expected.

With Login VSI 4.1.x a new VSImax method is introduced: VSImax v4.1. This methodology gives much better insight in system performance and scales to extremely large systems.

Test Results

Boot Storms

A key performance metric for desktop virtualization environments is the ability to boot the virtual machines quickly and efficiently to minimize user wait time for their desktop.

As part of Cisco's virtual desktop test protocol, we shut down each virtual machine at the conclusion of a benchmark test. When we run a new test, we cold boot all 4000 desktops and measure the time it takes for the 4000th virtual machine to register as available in the Virtual Desktops Administrator console.

The Cisco HyperFlex HXAF220c-M5SX based All-Flash cluster running Data Platform version 4.0.2a software can accomplish this task in 15 minutes.

Recommended Maximum Workload and Configuration Guidelines

Eight Node Cisco HXAF220c-M5S Rack Server Converge Nodes, Eight Node Cisco UCS C220 M5 Rack Servers, Eight Node Cisco UCS B200 M5 Blade Servers Compute Only Nodes, and HyperFlex All-Flash Cluster

For VMware Remote Desktop Hosted Sessions (RDSH) server sessions Desktop use case, the recommended maximum workload was determined based on both Login VSI Knowledge Worker workload end user experience measures and HXAF220c-M5SX server operating parameters.

This recommended maximum workload approach allows you to determine the server N+1 fault tolerance load the blade can successfully support in the event of a server outage for maintenance or upgrade.

Our recommendation is that the Login VSI Average Response and VSI Index Average should not exceed the Baseline plus 2000 milliseconds to ensure that end-user experience is outstanding. Additionally, during steady state, the processor utilization should average no more than 90-95 percent.

⚠ Memory should never be oversubscribed for Desktop Virtualization workloads.

⚠ Callouts have been added throughout the data charts to indicate each phase of testing.

Test Phase	Description
Boot	Start all RDS and/or VDI virtual machines at the same time.
Login	The Login VSI phase of test is where sessions are launched and start executing the

Test Phase	Description
	workload over a 48 minutes duration.
Steady state	The steady state phase is where all users are logged in and performing various work-load tasks such as using Microsoft Office, Web browsing, PDF printing, playing videos, and compressing files.
Logoff	Sessions finish executing the Login VSI workload and logoff.

Eight Node Cisco HXAF220c-M5S Rack-Mount Server, Eight Cisco UCS C220M5, Cisco UCS B200 M5 Compute Nodes, and HyperFlex All-Flash Cluster

For VMware Remote Desktop Hosted Sessions (RDSH) and Windows 10 Desktop use case, the recommended maximum workload was determined based on both Login VSI Knowledge Worker workload end user experience measures and HXAF220c-M5SX server operating parameters.

This recommended maximum workload approach allows you to determine the server N+1 fault tolerance load the blade can successfully support in the event of a server outage for maintenance or upgrade.

Our recommendation is that the Login VSI Average Response and VSI Index Average should not exceed the Baseline plus 2000 milliseconds to ensure that end-user experience is outstanding. Additionally, during steady state, the processor utilization should average no more than 90-95 percent.

⚠️   Memory should never be oversubscribed for Desktop Virtualization workloads.

Test Phase	Description
Boot	Start all RDS and/or VDI virtual machines at the same time.
Login	The Login VSI phase of test is where sessions are launched and start executing the workload over a 48 minutes duration.
Steady state	The steady state phase is where all users are logged in and performing various work-load tasks such as using Microsoft Office, Web browsing, PDF printing, playing videos, and compressing files.
Logoff	Sessions finish executing the Login VSI workload and logoff.

⚠️   The recommended maximum workload for a Cisco HyperFlex cluster configured on Cisco HXAF220c-M5SX with Intel Xeon Gold 6230 scalable family processors and 768GB of RAM for Windows 10 desktops with Office 2016 is 4000 virtual desktops.

⚠️   The recommended maximum workload for a Cisco HyperFlex cluster configured on Cisco HXAF220c-M5SX with Intel Xeon Gold 6230 scalable family processors and 768GB of RAM for Windows Server 2019 RDS desktop sessions with Office 2016 is 5000 virtual desktops.

5000 RDS Sessions, 3800 Windows 10 VMware Non-Persistent Instant Clones, and 4000 Windows 10 VMware Full Clone Persistent Desktops Testing on 24 Node Cisco HyperFlex Cluster

Hosted Shared desktops with 5000 user sessions on 192 Windows Server 2019 virtual machines on a 24-node HyperFlex cluster.

Test results for 5000 user sessions on VMware RDS highlights include:

- 1.101 second baseline response time

- 1.769 second average response time with 5000 desktops running

- Average CPU utilization of 80 percent during steady state

- Average of 300 GB of RAM used out of 768 GB available

- 50,000 peak I/O operations per second (IOPS) per cluster at steady state

- 500–750MBps peak throughput per cluster at steady state

Figure 53 Login VSI Analyzer Chart for 5000 Windows 2019 VMware RDSH Shared Desktops

NFO NSLD AppStart LogonTimer ShellExecuteWait VSImax per computer Stuck sessions over time Active sessions over time VSImax v4 Data

Raw Data Stuck sessions

Summary Settings VSImax v4 VSImax v4 Detailed VSImax v4 Detailed Weighted VSImax v4 Scatter UMEM IO CPU ZLC ZHC NFP
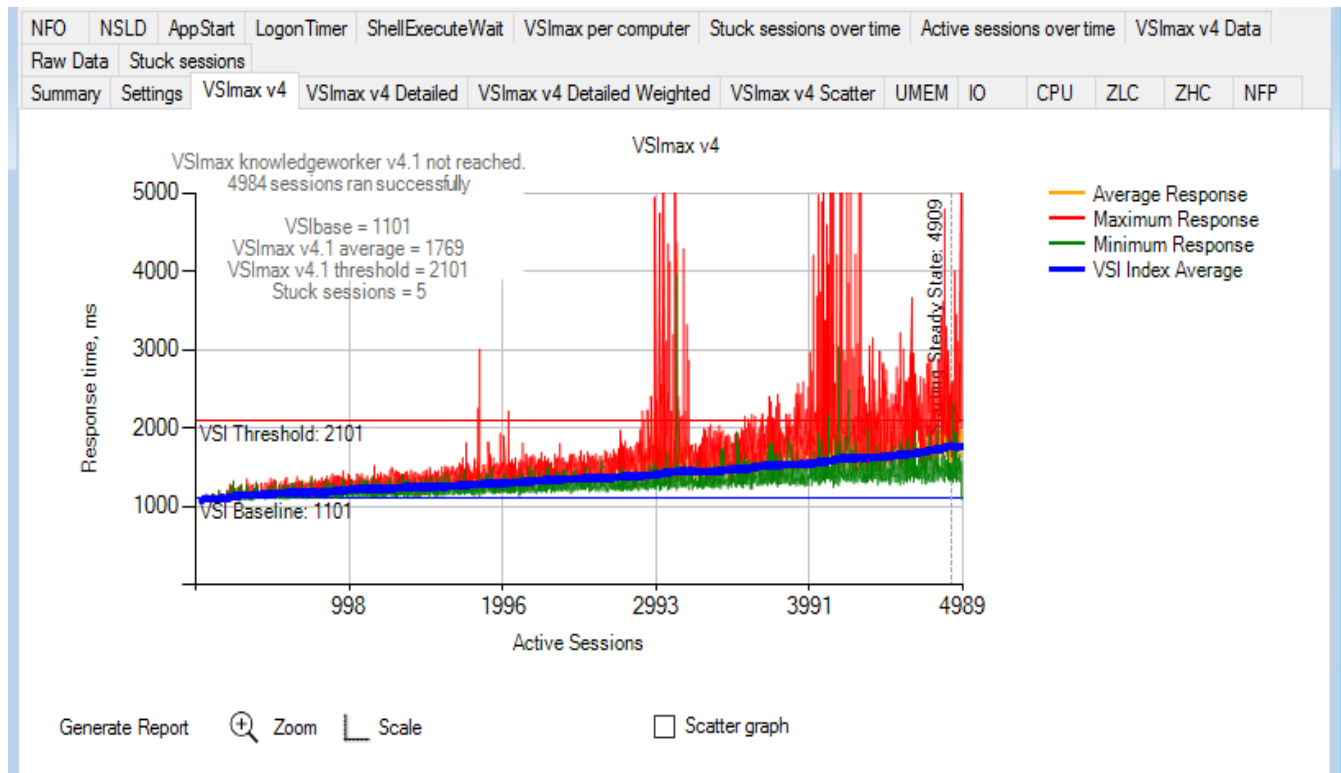
08042020-5000-RDS-64GB-8vCPU-HX-SCALE-RDS-TEST-1340

Successfully completed Login VSI test with **4984** **knowledgeworker** sessions. VSImax (system saturation) was not reached.

Test result review

5000 sessions were configured to be launched in **2880** seconds.

In total **16** sessions failed during the test:

- **0** sessions was/were not successfully launched

- **11** launched sessions failed to become active

- **4989** sessions were active during the test

- **5** sessions got stuck during the test (before VSImax threshold) **> Click Here**

With **4984** sessions the maximum capacity VSImax (v4.1) **knowledgeworker** was not reached with a Login VSI baseline performance score of **1101**

Login VSI index average score is **468** lower than threshold. It might be possible to launch more sessions in this configuration.

Baseline performance of **1101** is: **Good**

Figure 54 Three Consecutive Login VSI Analyzer Chart for 5000 Windows 2019 VMware Remote Desktop Server Sessions
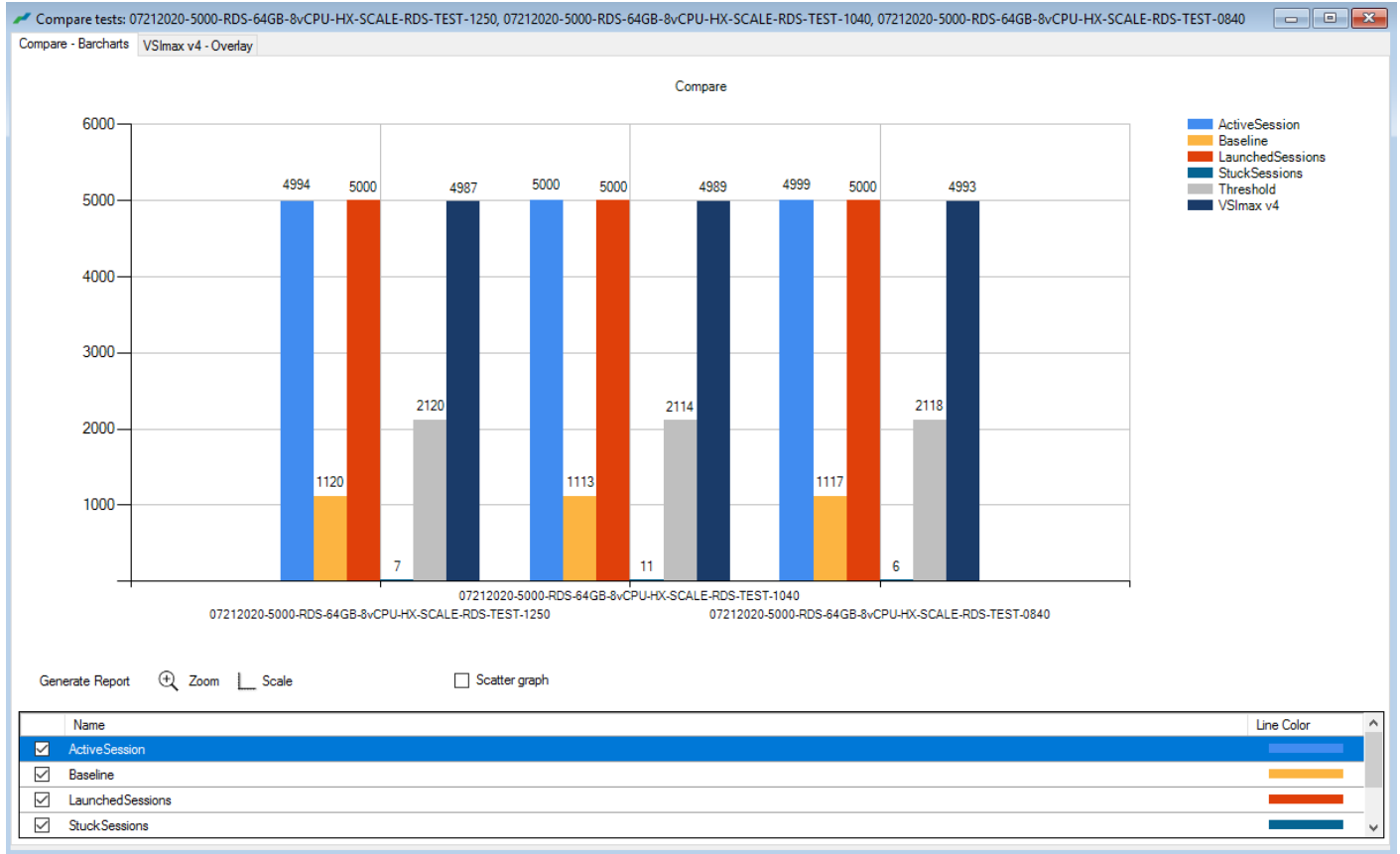
ESX Host Performance Counters

When running a VMware ESXi environment for our VMware Virtual Desktop workloads, it's important to monitor a few key performance counters to ensure the best end-user experience. We typically look for CPU utilization, memory availability, network throughput and Storage performance:

- CPU Performance: With VMware ESXi, using esxtop, our main counter is % Core Utilization.

- Memory Availability: We measure the memory available in megabytes to ensure that memory is not being consumed at a high level.

- Network throughput: We measure the bytes sent and received by the VM Network and Storage vSwitches on each ESXi HX Host.

- Storage performance: We use HyperFlex Connect to monitor and review storage performance during VDI.

The following figures show the results of our workload testing:

Figure 55   8x HyperFlex Converged ESXi Hosts CPU Core Utilization Running 5000 Windows Server 2019 VMware Hosted Shared Desktops (Total % Core Utilization)

Figure 56 16 x HyperFlex Compute-Only ESXi Hosts CPU Core Utilization Running 5000 Windows Server 2019 VMware Hosted Shared Desktops (Total % Core Utilization)

5000 RDS Server on Server 2019 using VMware Horizon Provisioning: Hyperfplex 16x Compute Nodes: ESXTOP Core CPU Util%

Legend:
- \\VDI-COMP09\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP10\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP11\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP12\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST13\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST14\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST15\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST08\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-01\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-02\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-03\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-04\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-05\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-06\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-07\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-08\Physical Cpu(_Total)\% Core Util Time

5000 RDS Server Sessions on Server 2019 using VMwae Horizon Proviosning: 8 Node Hyperflex NonKernal MBytes

Legend:
- \\VDI-HOST01.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST02.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST03.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST04.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST-05.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST-06.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST-07.VDILAB-HC.LOCAL\Memory\NonKernel MBytes
- \\VDI-HOST-08.VDILAB-HC.LOCAL\Memory\NonKernel MBytes

5000 RDS Server Sessions on Server 2019 using VMwae Horizon Proviosning: 16 Node Hyperflex NonKernal MBytes

— \\VDI-COMP09\Memory\NonKernel MBytes
— \\VDI-COMP10\Memory\NonKernel MBytes
— \\VDI-COMP11\Memory\NonKernel MBytes
— \\VDI-COMP12\Memory\NonKernel MBytes
— \\VDI-HOST13\Memory\NonKernel MBytes
— \\VDI-HOST14\Memory\NonKernel MBytes
— \\VDI-HOST15\Memory\NonKernel MBytes
— \\VDI-HOST08\Memory\NonKernel MBytes
— \\VDI-HOST-01\Memory\NonKernel MBytes
— \\VDI-COMP-02\Memory\NonKernel MBytes
— \\VDI-COMP-03\Memory\NonKernel MBytes
— \\VDI-COMP-04\Memory\NonKernel MBytes
— \\VDI-COMP-05\Memory\NonKernel MBytes
— \\VDI-COMP-06\Memory\NonKernel MBytes
— \\VDI-COMP-07\Memory\NonKernel MBytes

MBits Sent / Received for 5000 RDS Server Sessions using VMware Horizon Provisioning for vSwitch-hx-vm-network Switch for 8 Hyperflex Node Converged Nodes
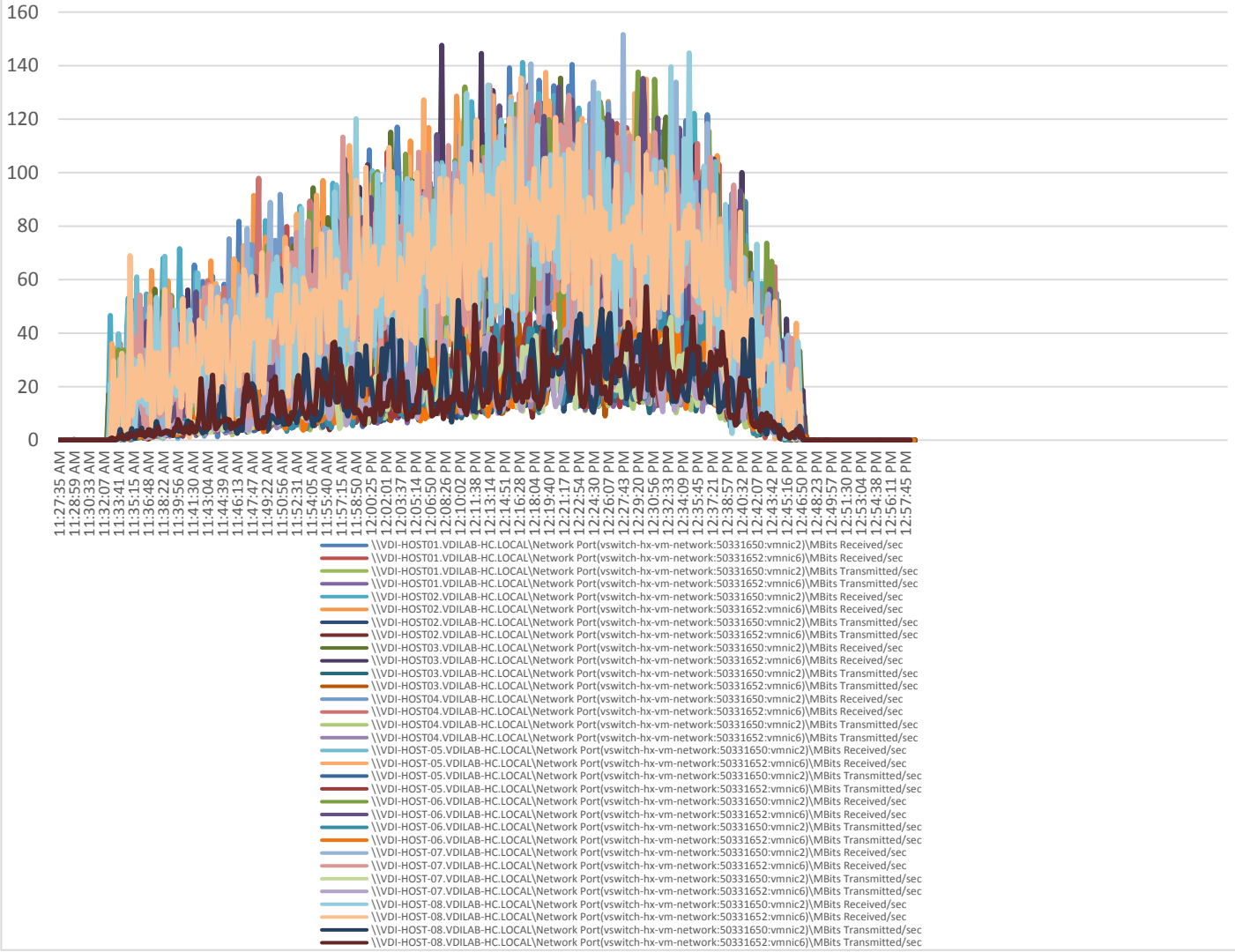
MBits Sent / Received for 5000 RDS Server Sessions using VMware Horizon Provisioning for vSwitch-hx-vm-network Switch for 16 Hyperflex Node Compute Only Nodes
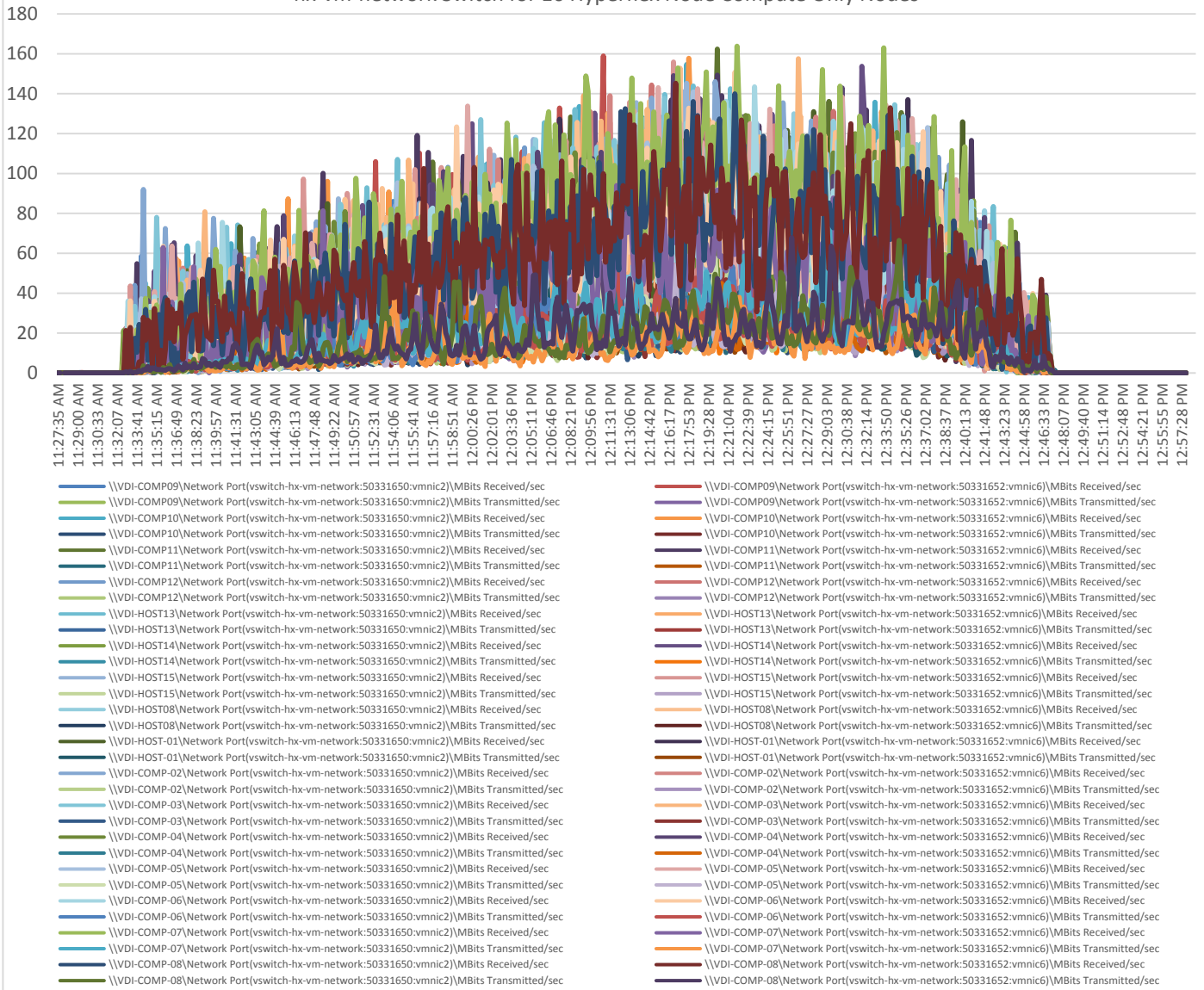
Figure 57 HyperFlex Cluster UI Performance Chart for Knowledge Worker Workload Running 5000 User Test on VMware Remote Desktop Server Sessions (RDSH) Windows Server 2019 User Sessions
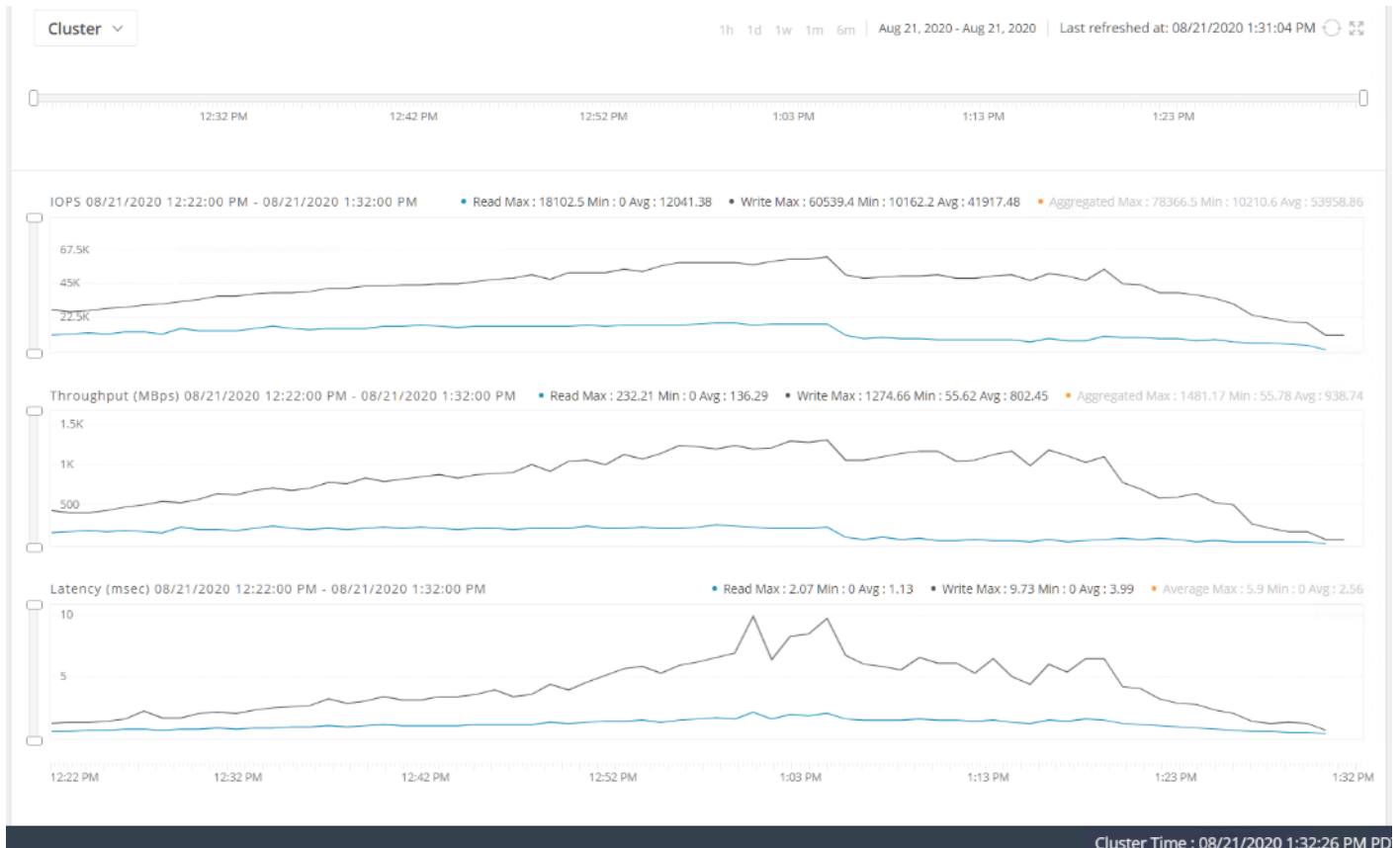
Test results for 3800 VMware VDI non-persistent Desktops using VMware Provisioning Services highlights include:

- 0.815 second baseline response time
- 1.385 second average response time with 4000 desktops running
- Average CPU utilization of 90 percent during steady state
- Average of 550 GB of RAM used out of 768 GB available
- 50000 peak I/O operations per second (IOPS) per cluster at steady state
- 500-750MBps peak throughput per cluster at steady state

Figure 58 Login VSI Analyzer Chart for 3800 Instant Clone Non-Persistent Windows 10 VMware Virtual
Desktop

aug222020_vsiServices_3800users_IC_PCPtest-_0

NFO   NSLD   AppStart   LogonTimer   ShellExecuteWait   VSImax per computer   Active sessions over time   VSImax v4 Data   Raw Data   Failed to become active
Summary   Settings   VSImax v4   VSImax v4 Detailed   VSImax v4 Detailed Weighted   VSImax v4 Scatter   UMEM   IO   CPU   ZLC   ZHC   NFP

aug222020_vsiServices_3800users_IC_PCPtest-_0

Successfully completed Login VSI test with **3796** **knowledgeworker** sessions. VSImax (system saturation) was not reached.

Test result review

3800 sessions were configured to be launched in **2880** seconds.

In total **4** sessions failed during the test:

- **0** sessions was/were not successfully launched
- **4** launched sessions failed to become active
- **3796** sessions were active during the test
- **0** sessions got stuck during the test (before VSImax threshold)

With **3796** sessions the maximum capacity VSImax (v4.1) **knowledgeworker** was not reached with a Login VSI baseline performance score of **815**

Login VSI index average score is **624** lower than threshold. It might be possible to launch more sessions in this configuration.

Baseline performance of **815** is: **Good**

---

aug222020_vsiServices_3800users_IC_PCPtest-_0

NFO   NSLD   AppStart   LogonTimer   ShellExecuteWait   VSImax per computer   Active sessions over time   VSImax v4 Data   Raw Data   Failed to become active
Summary   Settings   VSImax v4   VSImax v4 Detailed   VSImax v4 Detailed Weighted   VSImax v4 Scatter   UMEM   IO   CPU   ZLC   ZHC   NFP

VSImax v4

VSImax knowledgeworker v4.1 not reached.
3796 sessions ran successfully

VSIbase = 815
VSImax v4.1 average = 1335
VSImax v4.1 threshold = 1816
Stuck sessions = 0

VSI Threshold: 1816
VSI Baseline: 815

Response time, ms
Active Sessions

Legend:
- Average Response
- Maximum Response
- Minimum Response
- VSI Index Average

Generate Report   Zoom   Scale   Scatter graph

Name	Line Color
☑ Average Response	
☑ Maximum Response	
☑ Minimum Response	
☑ VSI Index Average	

Figure 59 Three Consecutive Login VSI Analyzer Chart for 3800 Windows 10 VMware Instant Clone Non-
Persistent Virtual Desktops

ESX Host Performance Counters

When running a VMware ESXi environment for our VMware Virtual Desktop workloads, it's important to monitor a few key performance counters to ensure the best end-user experience. We typically look for CPU utilization, memory availability, network throughput and Storage performance:
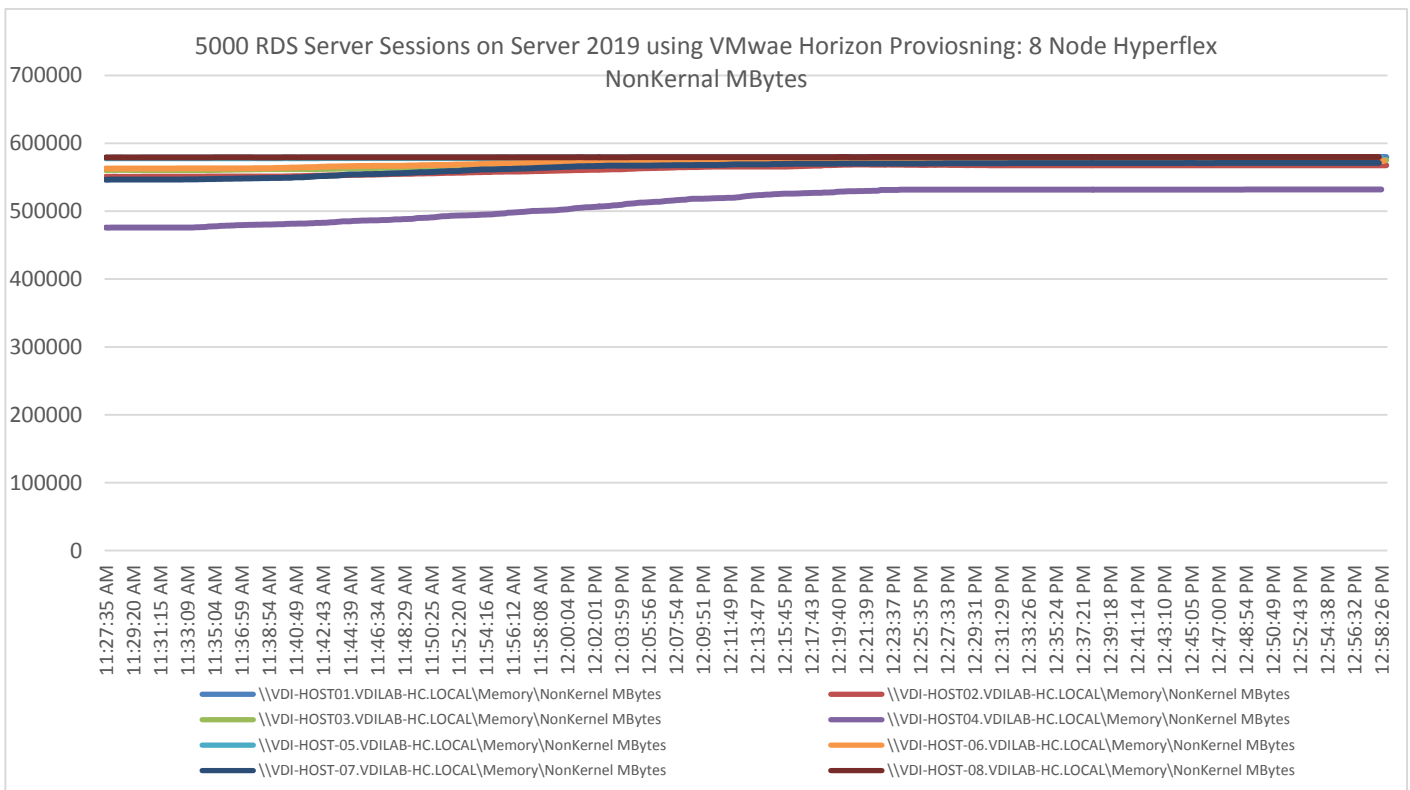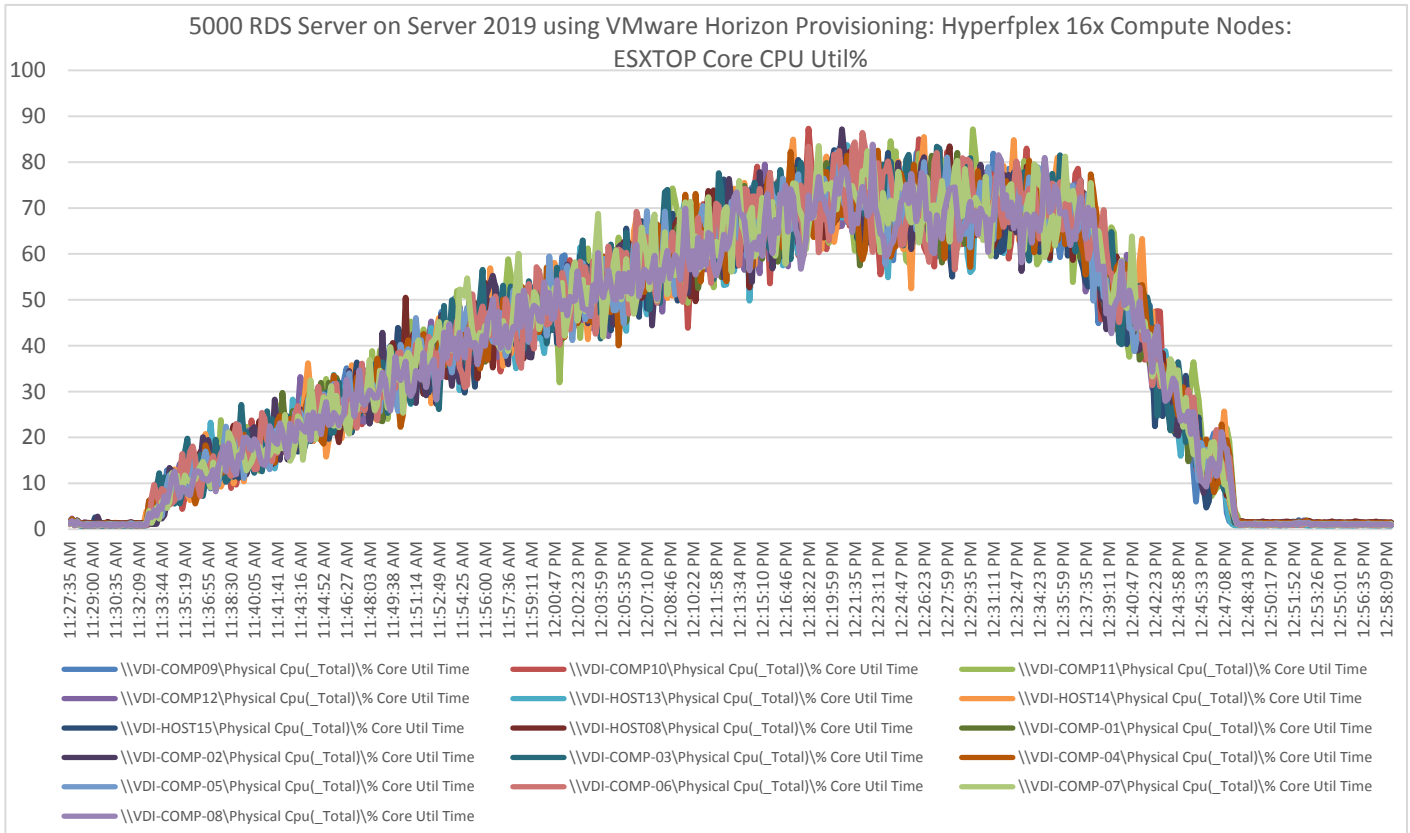
- CPU Performance: With VMware ESXi, using esxtop, our main counter is % Core Utilization.

- Memory Availability: We measure the memory available in megabytes to ensure that memory is not being consumed at a high level.

- We measure the bytes sent and received by the VM Network and Storage vSwitches on each ESXi HX Host. Storage performance: We use HyperFlex Connect to monitor and review storage performance during VDI.

The following figures show the results of our workload testing:

Figure 60 8x HyperFlex Converged ESXi Hosts CPU Core Utilization Running 3800 Windows 10 VMware Non-Persistent Desktops (Total % Core Utilization)

3800 VDI Non-Persistent on WIndows 10 using VMware Horizon Provioning: 8x Hyperflex Converge Nodes: ESXTOP Core CPU%

- \\VDI-HOST01.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST02.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST03.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST04.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST-05.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST-06.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST-07.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST-08.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time

3800 VDI Non-Persistent WIndows 10 using VMware Horizon Provioning: 16x Hyperflex Compute Nodes: ESXTOP CPU Core Util%

- \\VDI-COMP09\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP10\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP11\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP12\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST13\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST14\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST15\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST16\Physical Cpu(_Total)\% Core Util Time
- \\VDI-HOST-01\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-02\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-03\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-04\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-05\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-06\Physical Cpu(_Total)\% Core Util Time
- \\VDI-COMP-07\Physical Cpu(_Total)\% Core Util Time
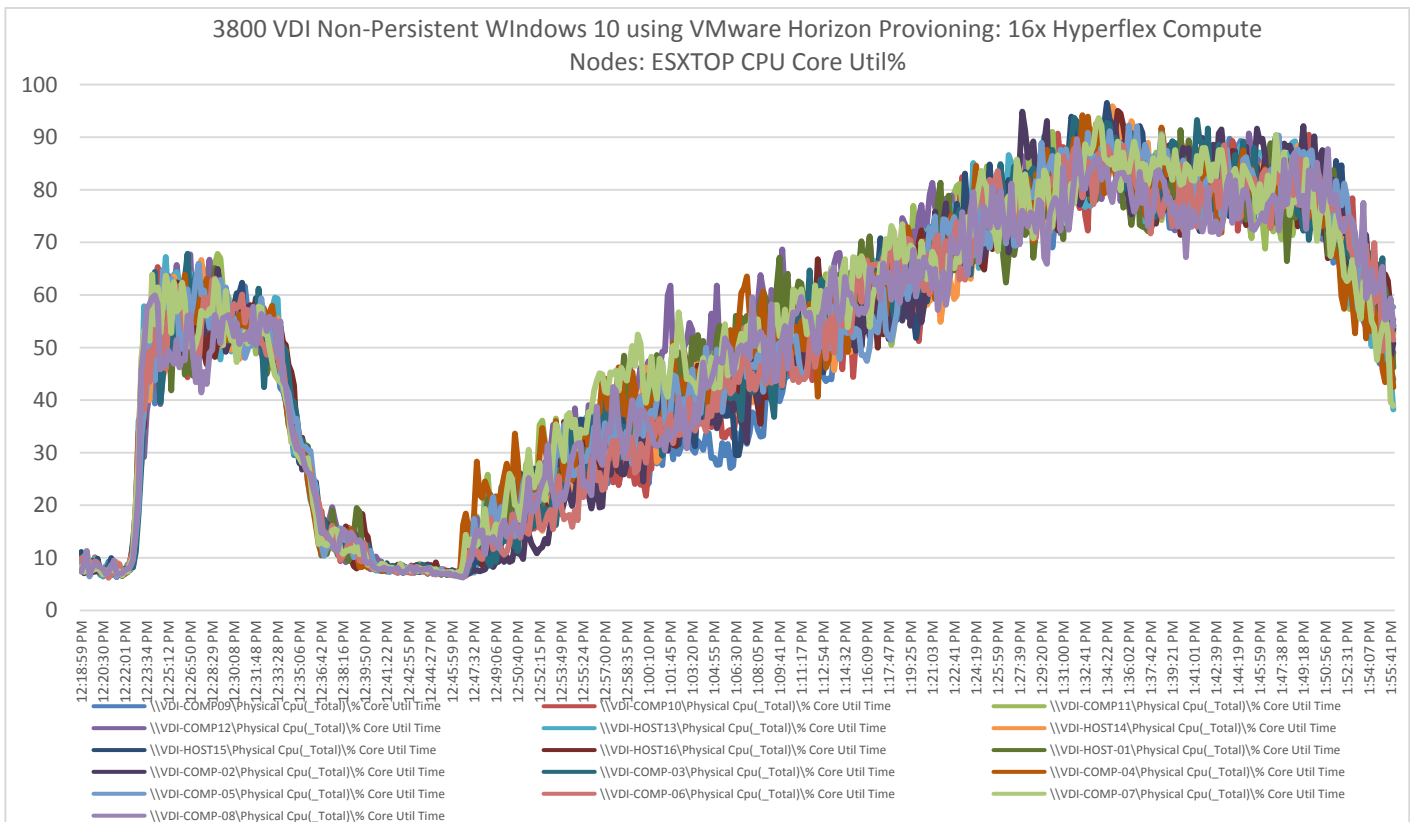- \\VDI-COMP-08\Physical Cpu(_Total)\% Core Util Time

Figure 61 8x HyperFlex Compute-Only ESXi Hosts CPU Core Utilization Running 3800 Windows 10 VMware Non-Persistent Desktops (Total % Core Utilization)

3800 VDI non-persistent on Windows 10 using VMware Horizon Provisioning: 8x Hyperflex Converge Nodes CPU Core Util%
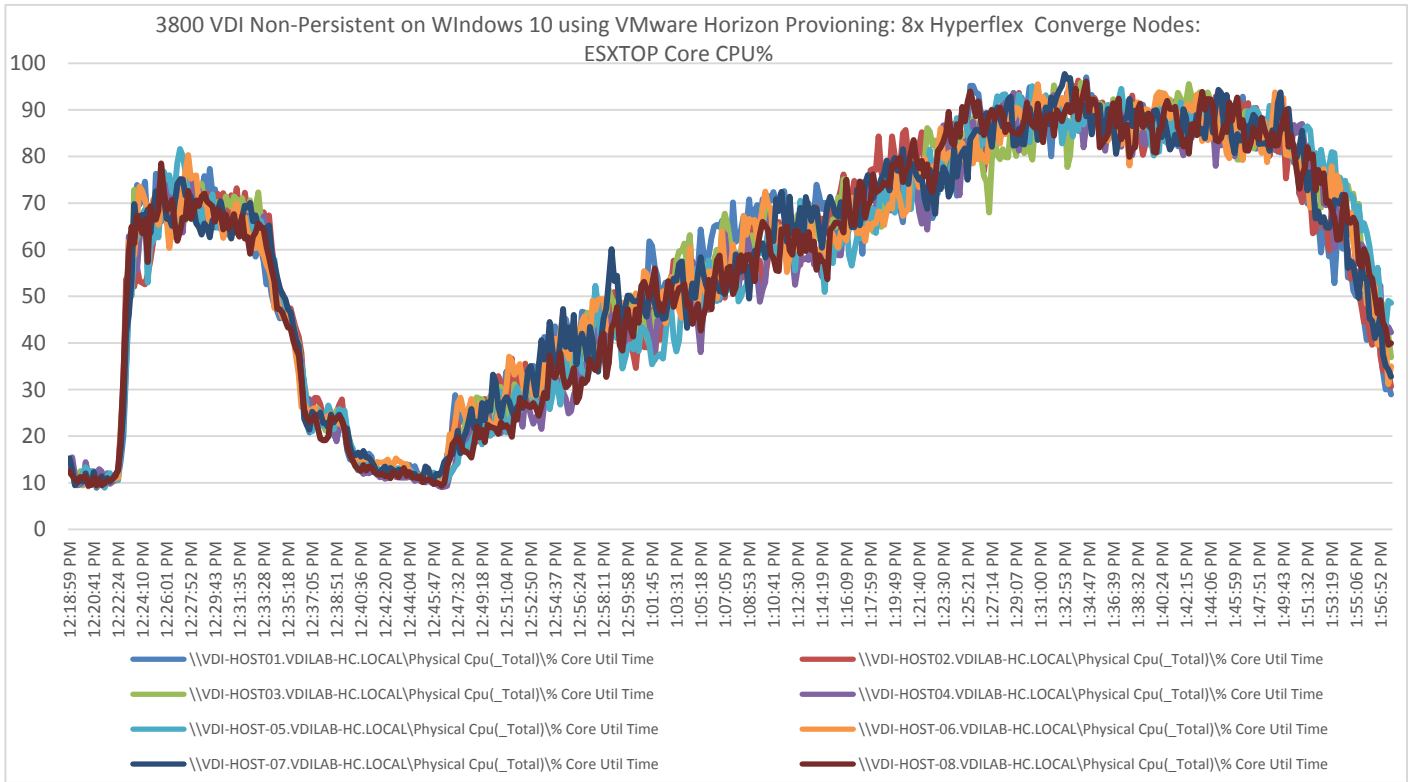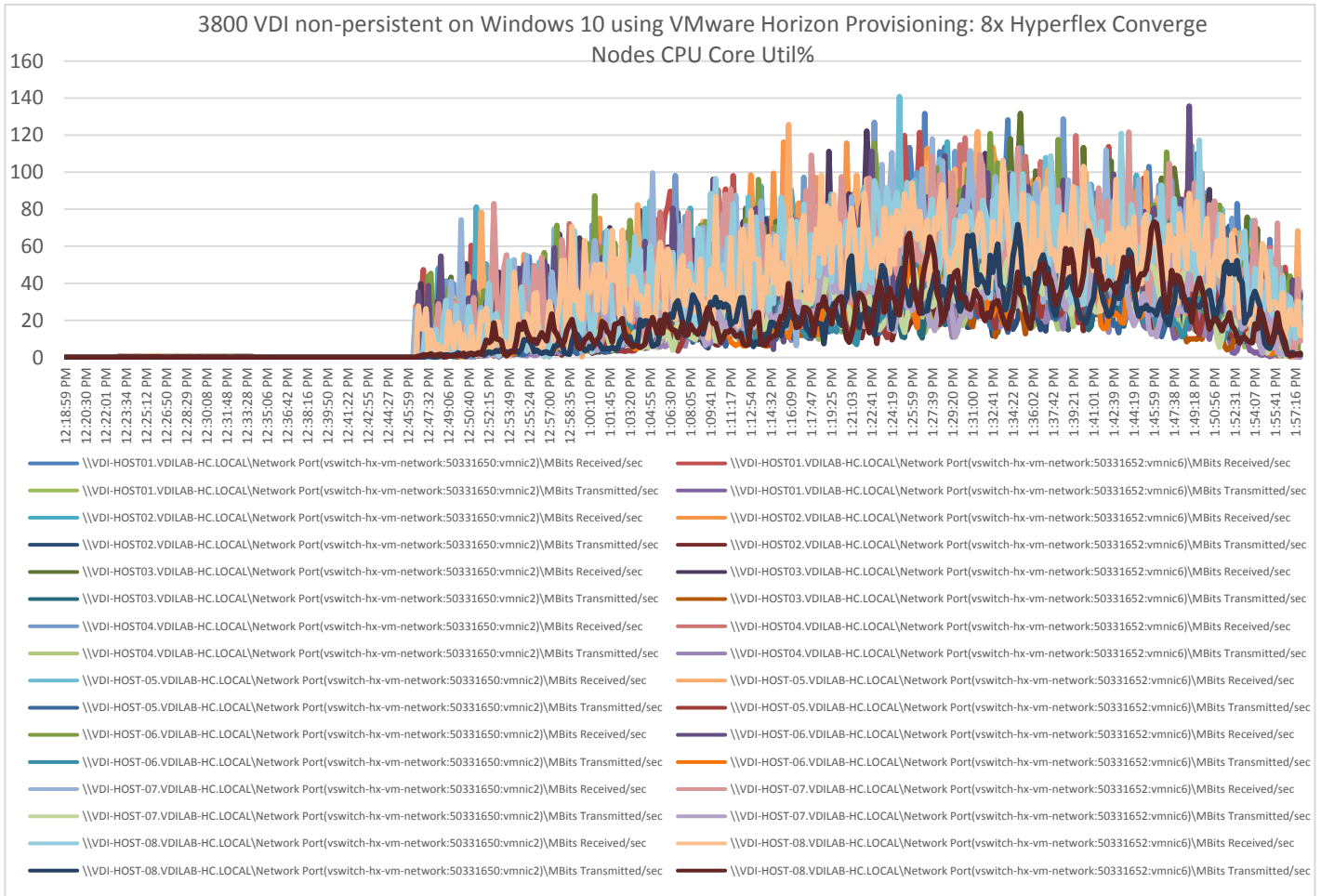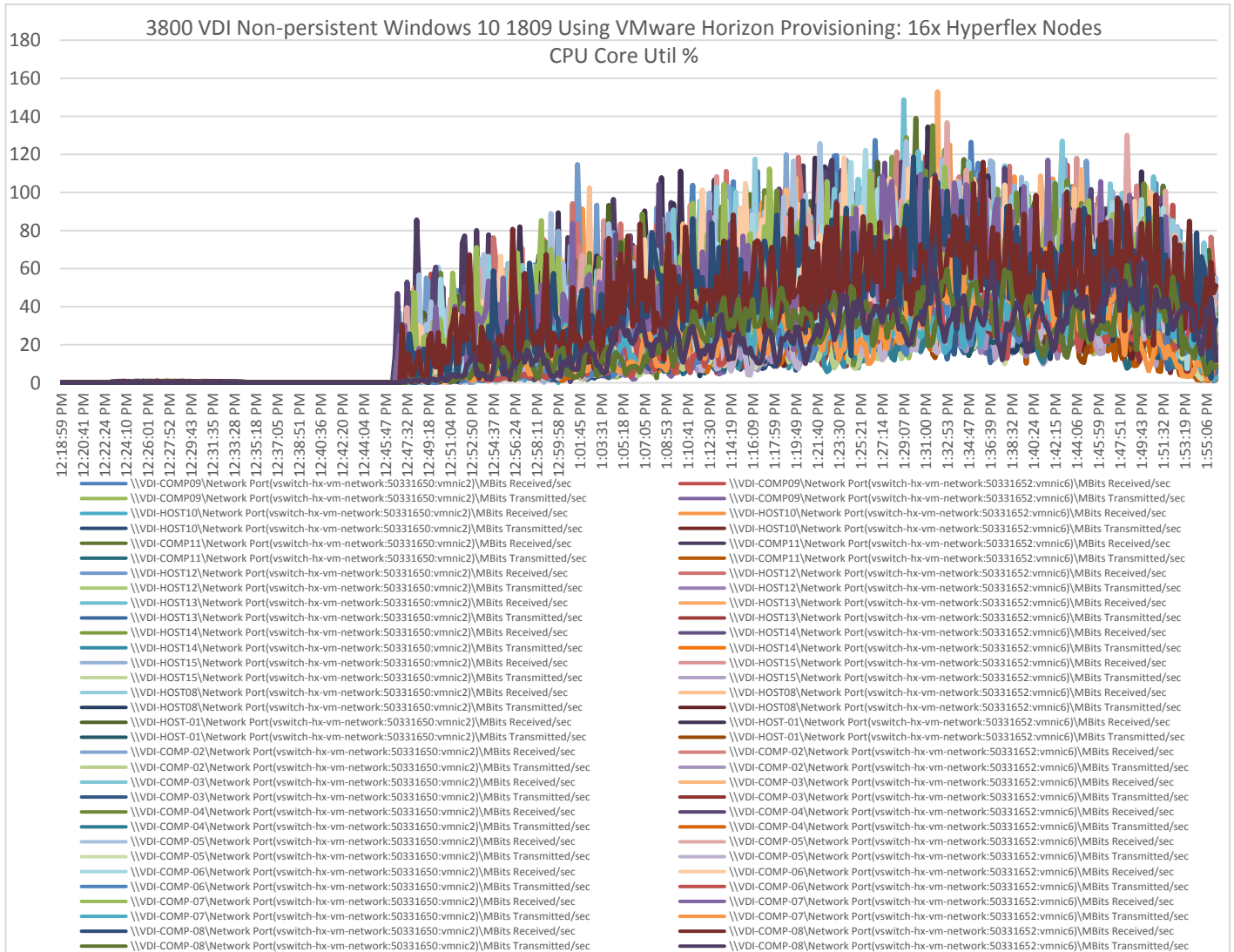
Figure 62 16x HyperFlex Compute-Only ESXi Hosts CPU Core Utilization Running 3800 Windows 10 VMware Non-Persistent Desktops (Total % Core Utilization)

3800 VDI Non-persistent Windows 10 1809 Using VMware Horizon Provisioning: 16x Hyperflex Nodes CPU Core Util %

3800 VDI non-persistent desktops on WIndows 10 using VMware Horizon Provisioning: 8x Hyperflex Converge nodes Non Kernel MBytes
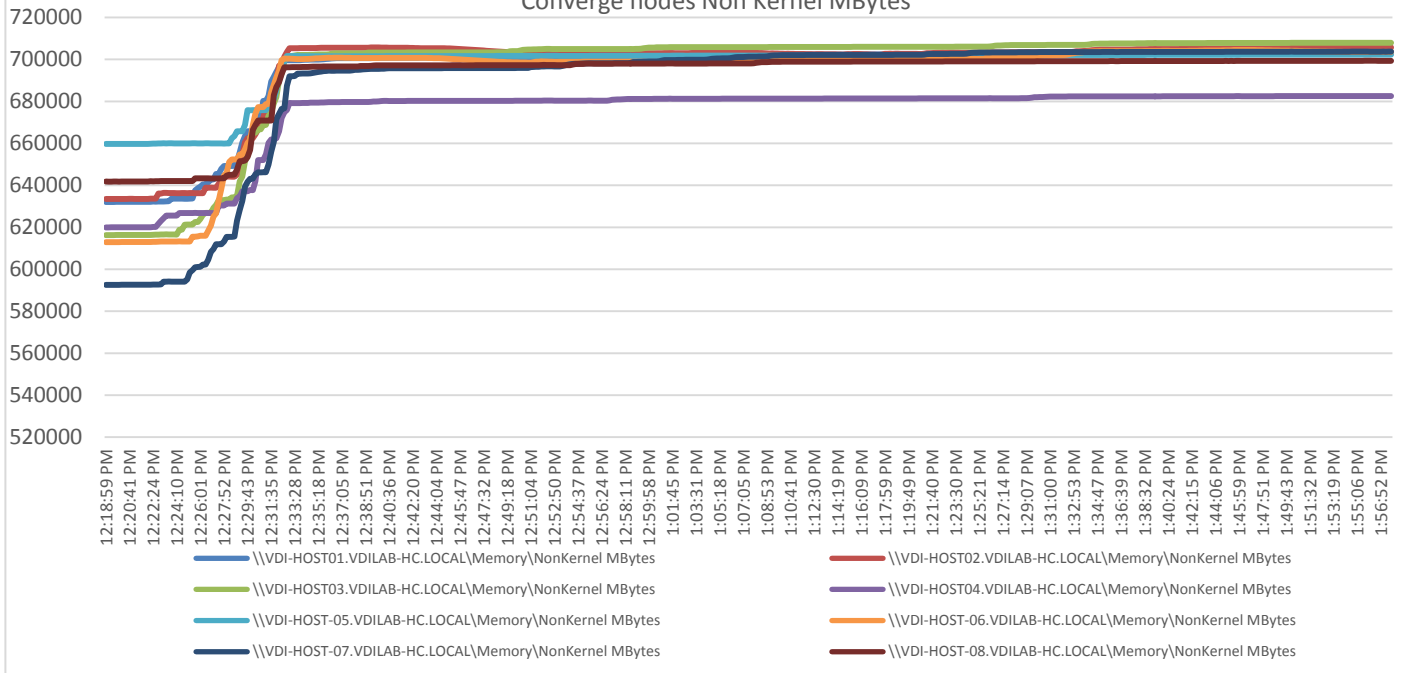
3800 VDI non-persistent desktops on WIndows 10 using VMware Horizon Provisioning: 16x Hyperflex Compute nodes Non Kernel MBytes

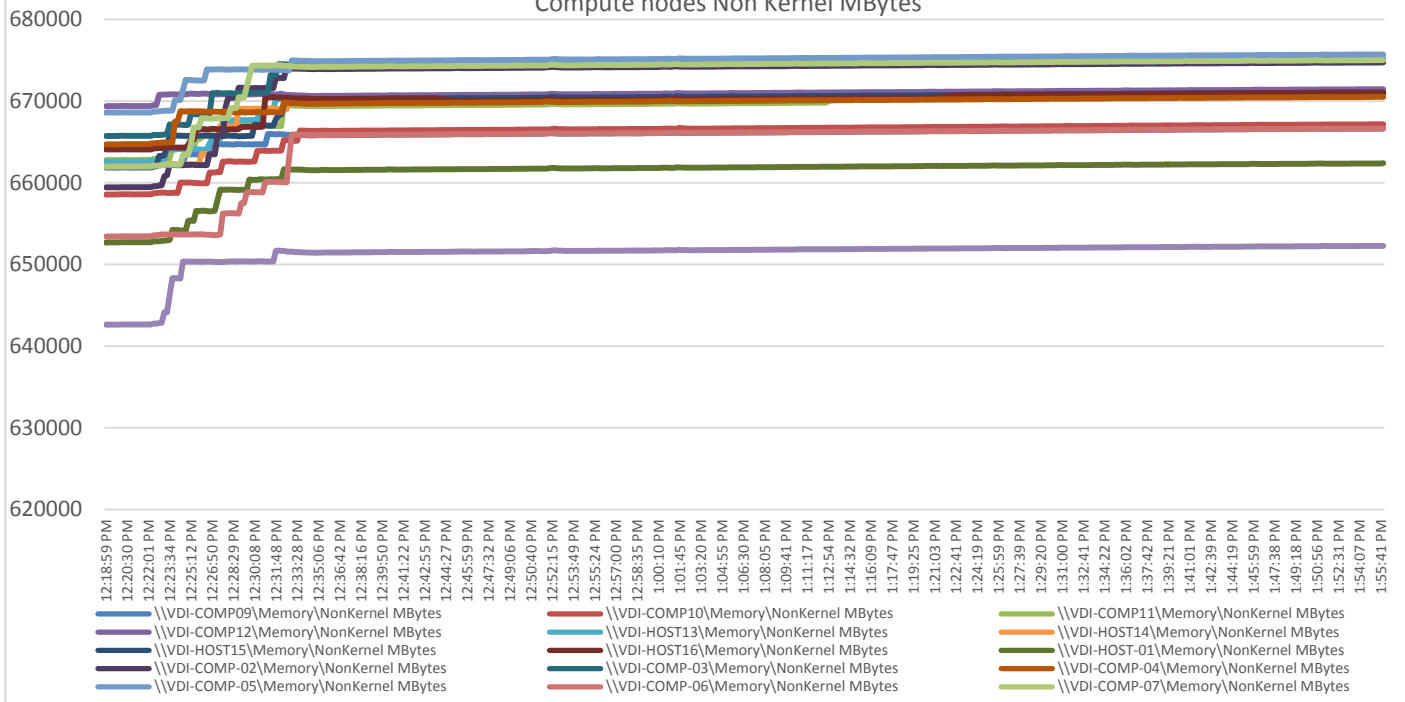Figure 63 HyperFlex Cluster UI Performance Chart for Knowledge Worker Workload Running 3800 User Test on VMware Windows 10 Virtual Desktops
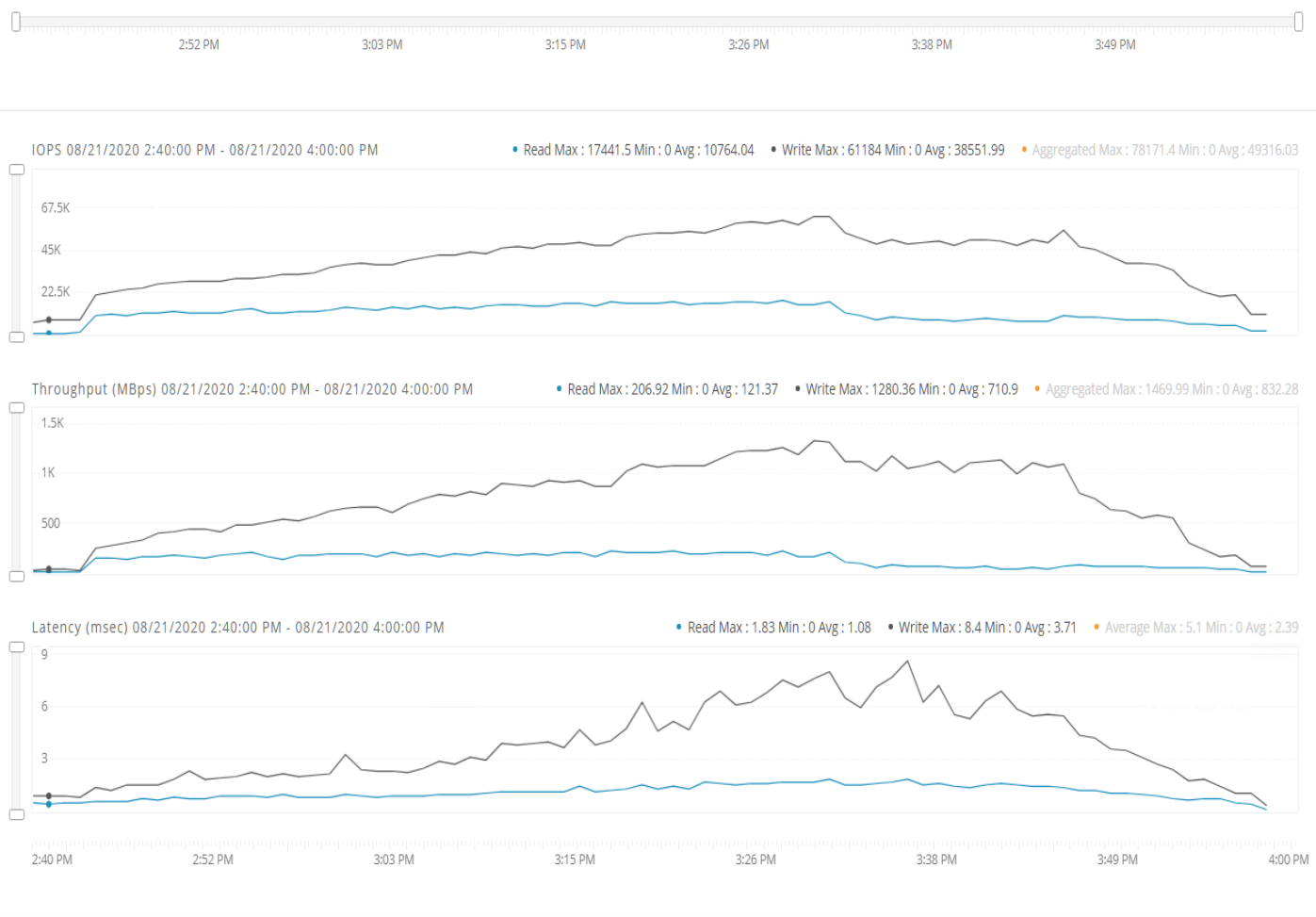
Test results for 4000 VMware VDI Persistent Desktops using Machine Creation Services highlights include:

- 0.826 second baseline response time

- 1.385 second average response time with 4000 desktops running

- Average CPU utilization of 90 percent during steady state

- Average of 650 GB of RAM used out of 768 GB available

- 50000 peak I/O operations per second (IOPS) per cluster at steady state

- 500-750MBps peak throughput per cluster at steady state

Figure 64 Login VSI Analyzer Chart for 4000 Windows 10 Horizon Persistent Virtual Desktops

Figure 65   Three Consecutive Login VSI Analyzer Chart for 4000 Windows 10 VMware persistent Virtual
Desktops
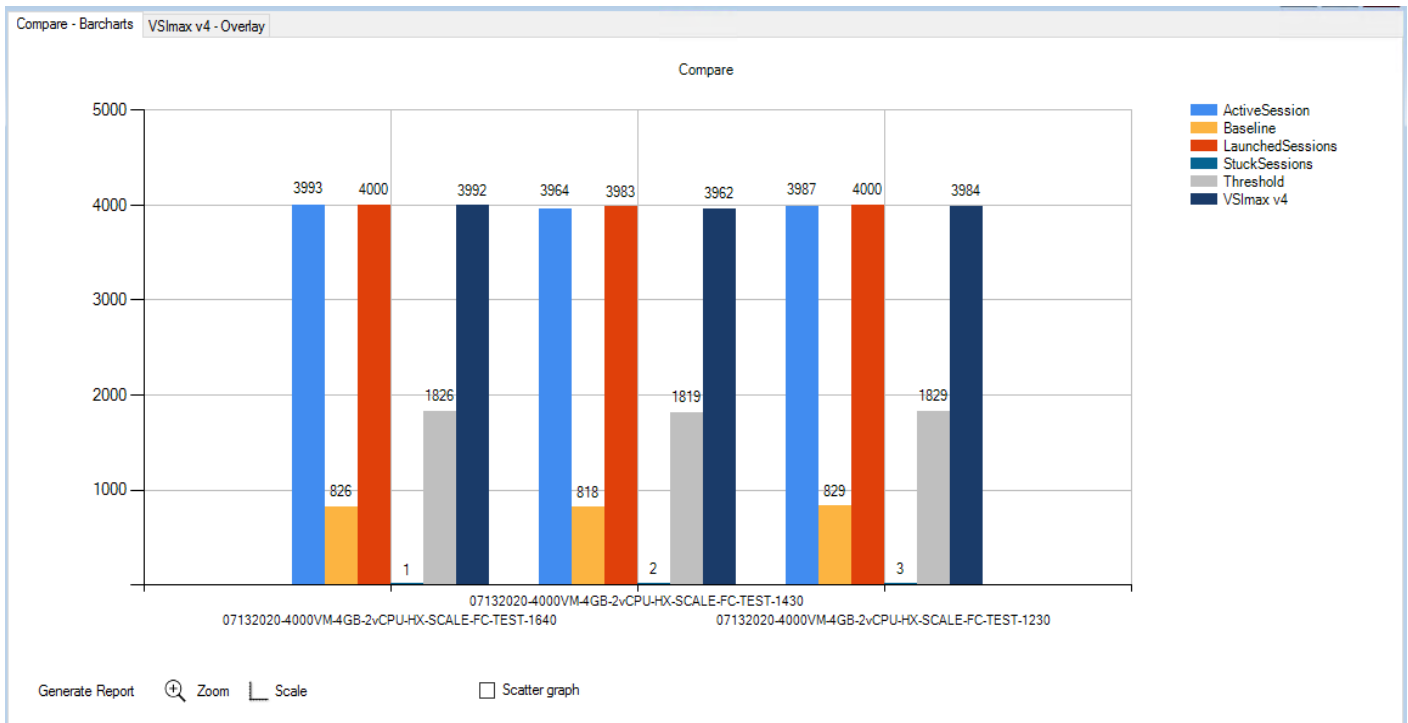
ESX Host Performance Counters

When running a VMware ESXi environment for our VMware **Virtual Desktop workloads, it's important to monitor a** few key performance counters to ensure the best end-user experience. We look for CPU utilization, memory availability, network throughput and Storage performance:

- CPU Performance: With VMware ESXi, using esxtop, our main counter is % Core Utilization.

- Memory Availability: We measure the memory available in megabytes to ensure that memory is not being consumed at a high level.

- We measure the bytes sent and received by the VM Network and Storage vSwitches on each ESXi HX Host.

- Storage performance: We use HyperFlex Connect to monitor and review storage performance during VDI.

The following figures show the results of our workload testing:

Figure 66 8x HyperFlex Converged ESXi Hosts CPU Core Utilization Running 4000 Windows 10 VMware Persistent Desktops (Total % Core Utilization)

ESXTOP CPU Core Util% for 8x Hyperflex Converged Nodes for 4000 VDI Persistent Desktops using VMware Horizon Provisoning

\\VDI-HOST01.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST02.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST03.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST04.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST-05.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST-06.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST-07.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time
\\VDI-HOST-08.VDILAB-HC.LOCAL\Physical Cpu(_Total)\% Core Util Time

Figure 67 16x HyperFlex Converged ESXi Hosts CPU Core Utilization Running 4000 Windows 10 VMware Persistent Desktops (Total % Core Utilization)

ESXTOP CPU Core Util% for 16x Hyperflex Converged Nodes for 4000 VDI Persistent Desktops using VMware Horizon Provisioning

NonKernal Memory used in Mbytes for 8x Converge Nodes Hyperflex Converged for 4000 Users Using VMware Horizon Provisioning
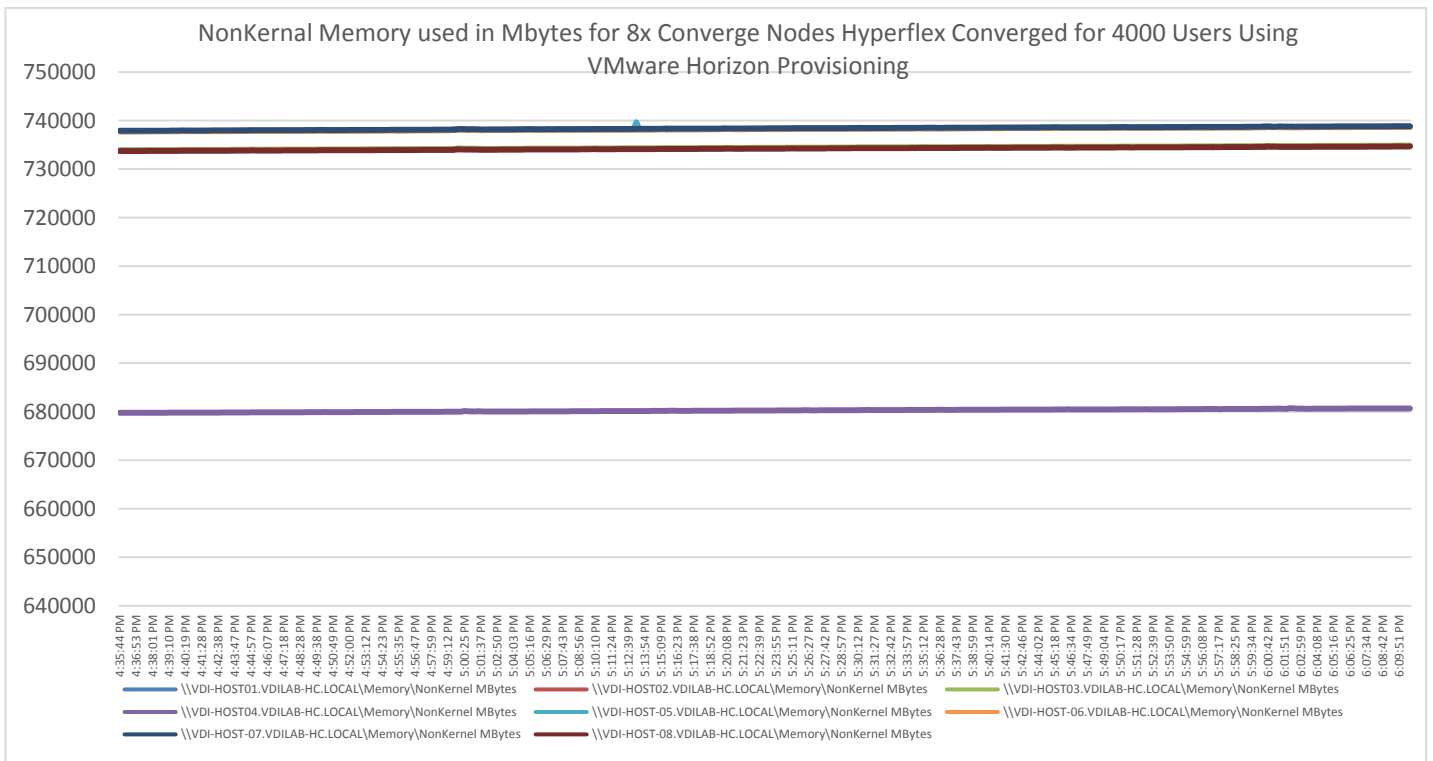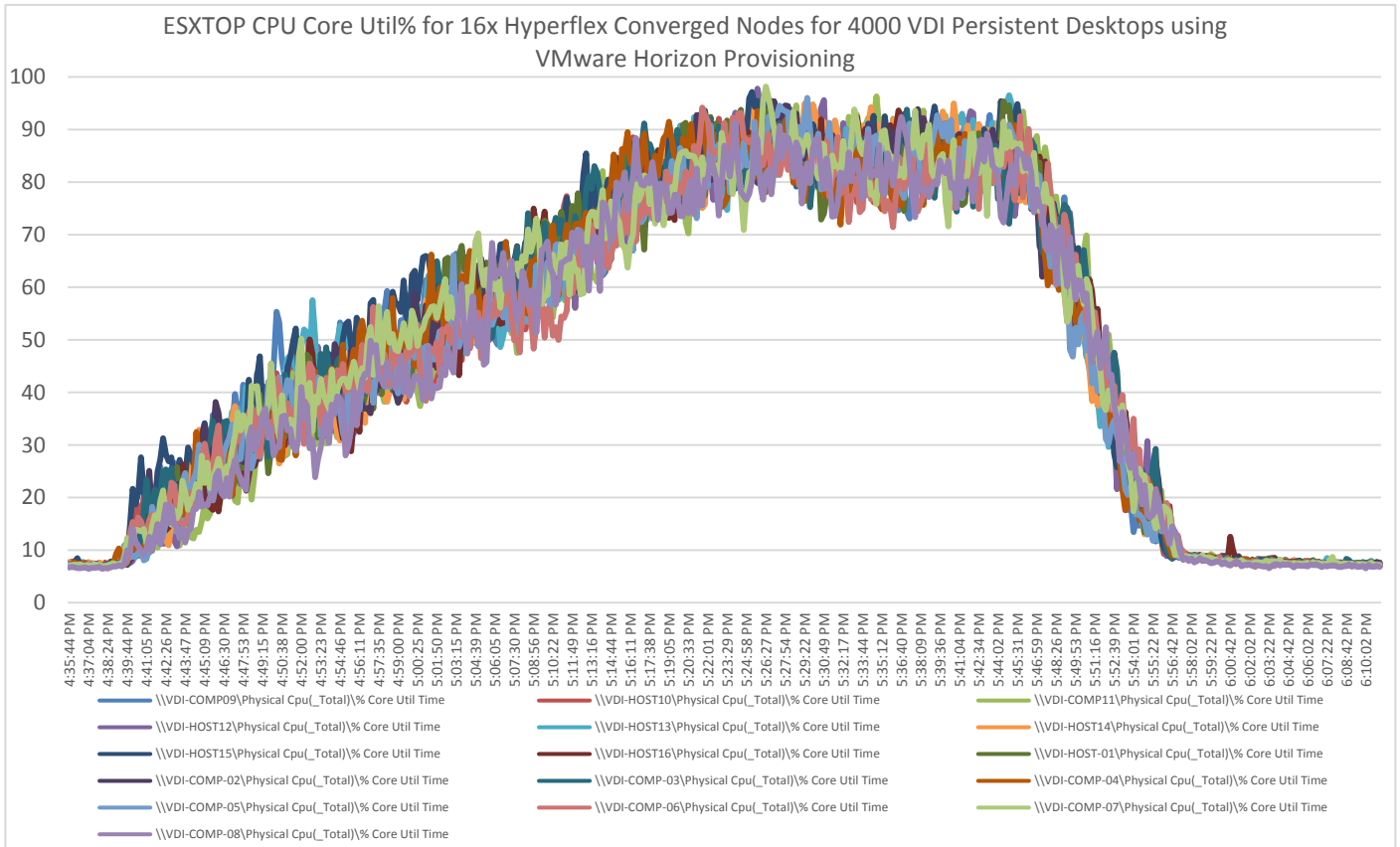
NonKernal Memory used in Mbytes for 16x Hyperflex Compute Only Nodes for 4000 Usres Using VMware Horizon Provisioning
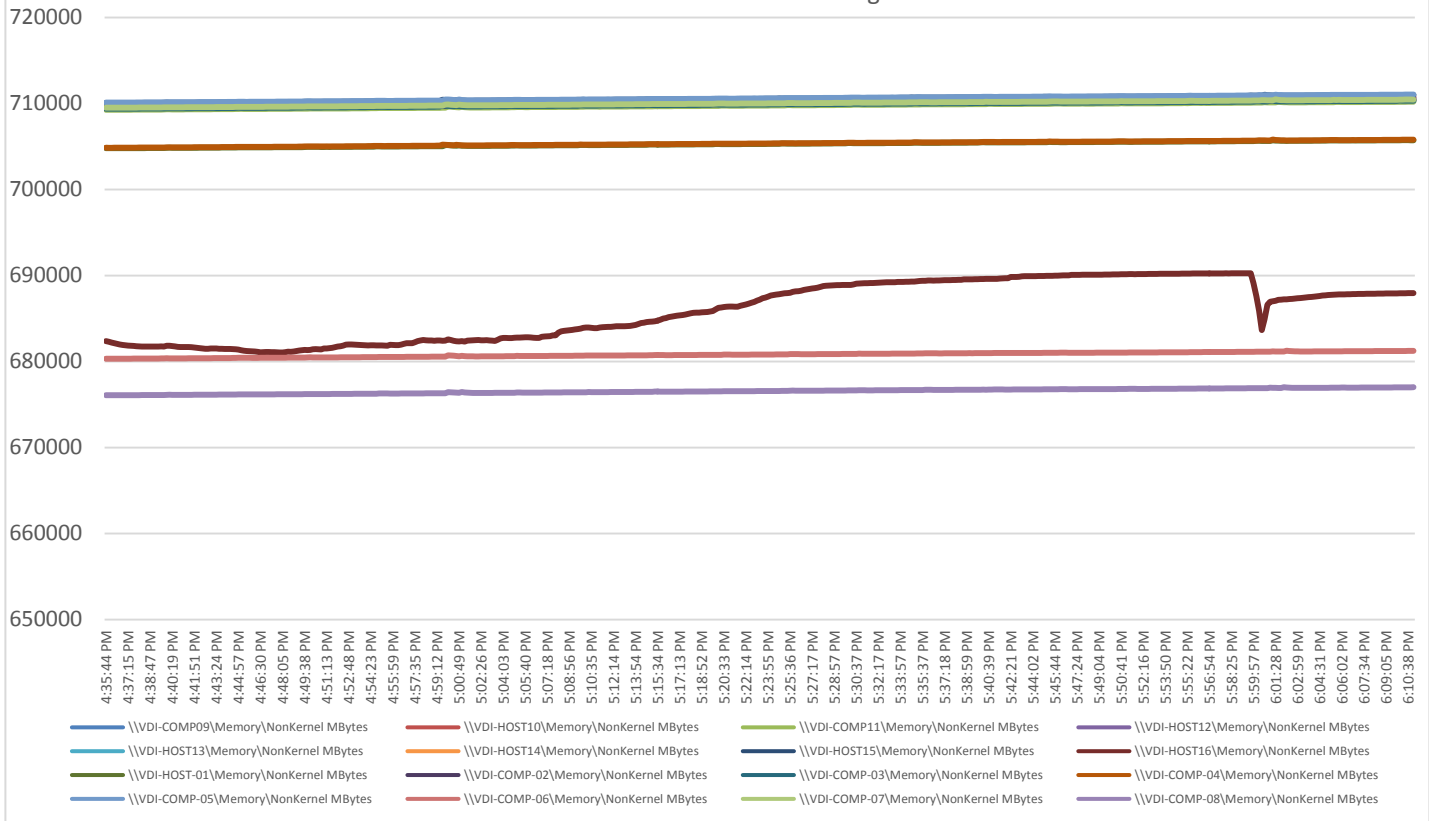
Mbytes Received/Sent for 4000 Full Clone Persistent using VMware Horizon for vSwitch vm-hx-vmnetwork
8x HX Converged Nodes
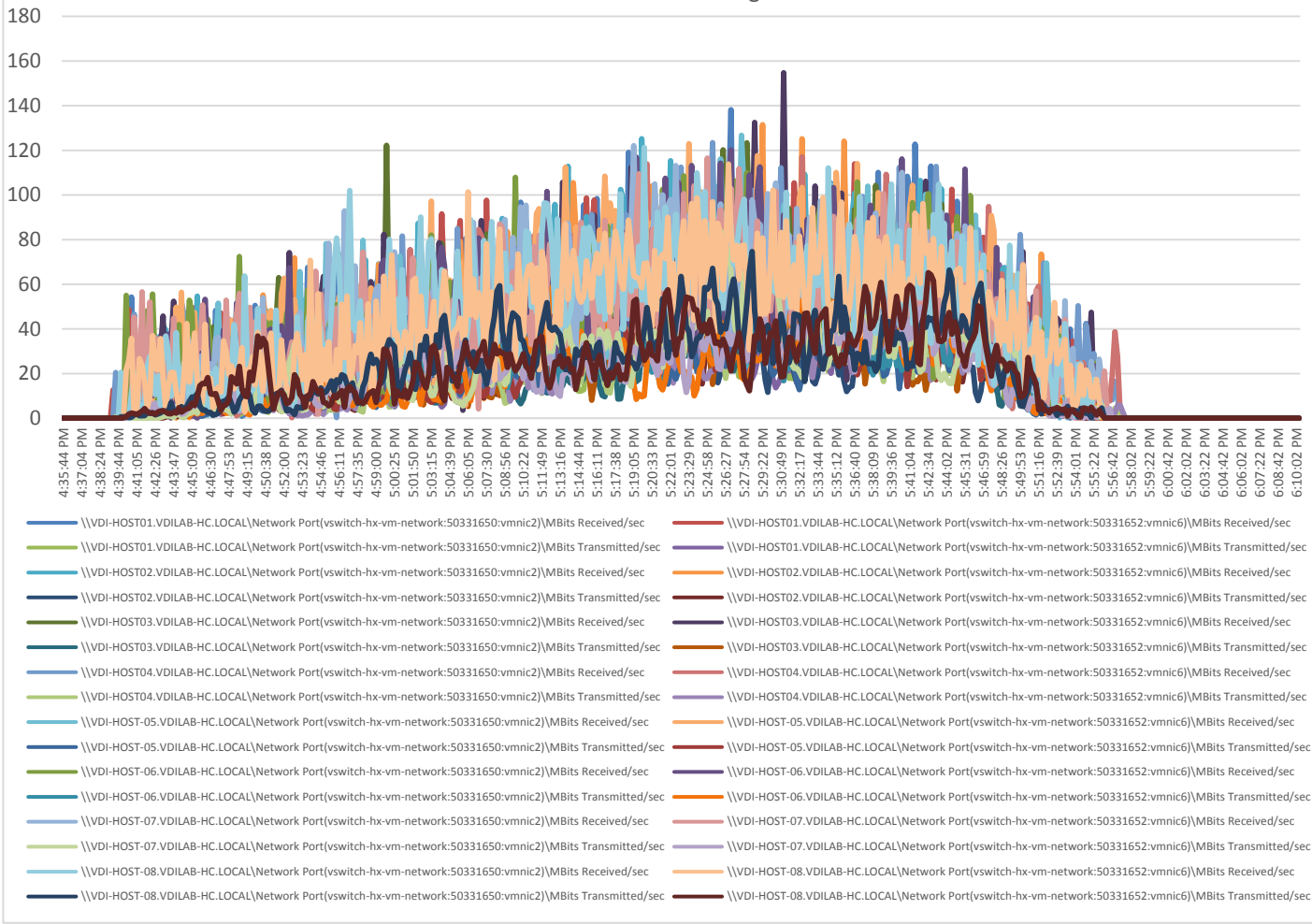
Mbytes Received/Sent for 4000 Full Clone Persistent using VMware Horizon for vSwitch vm-hx-vmnetwork 8x HX Converged Nodes
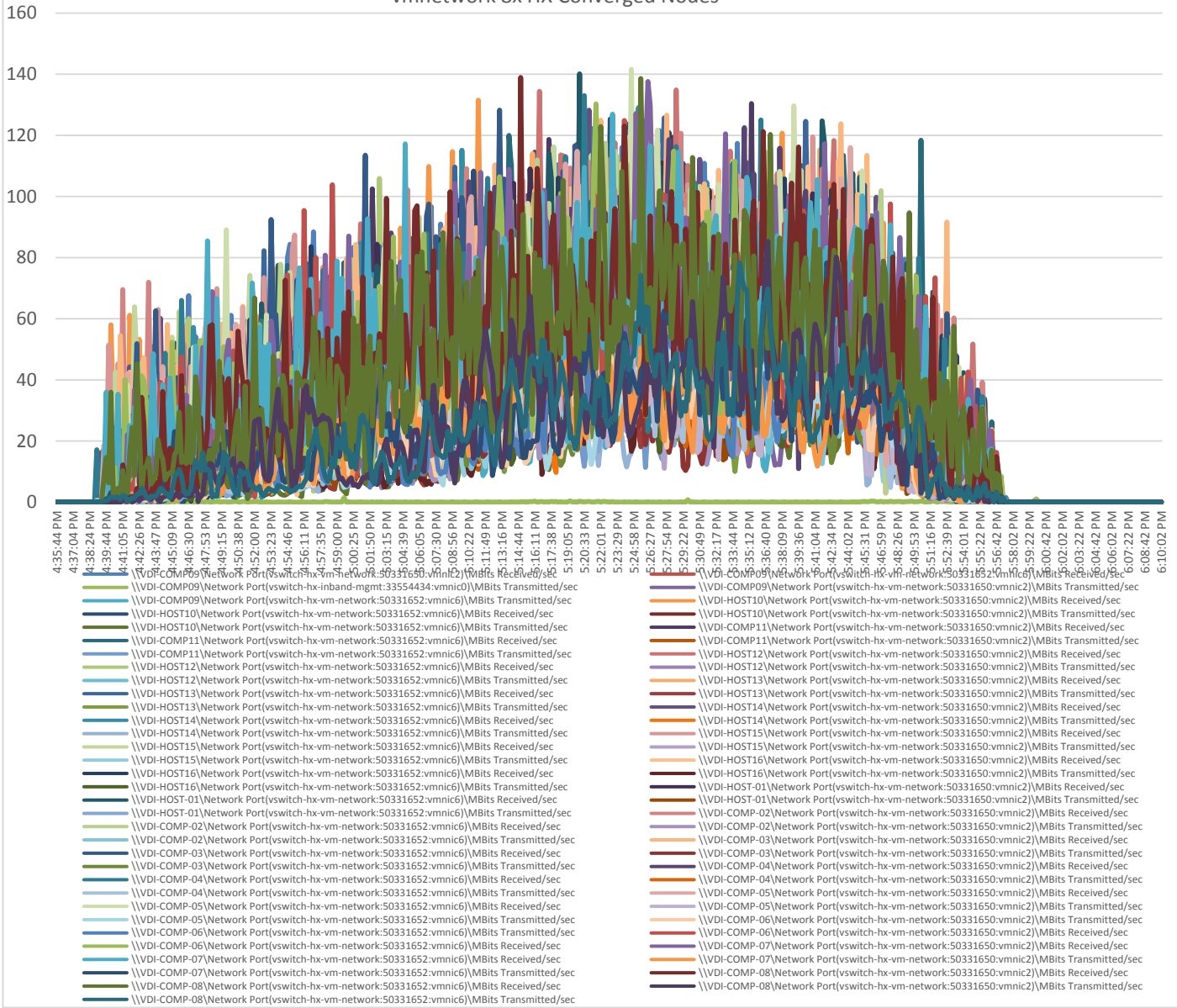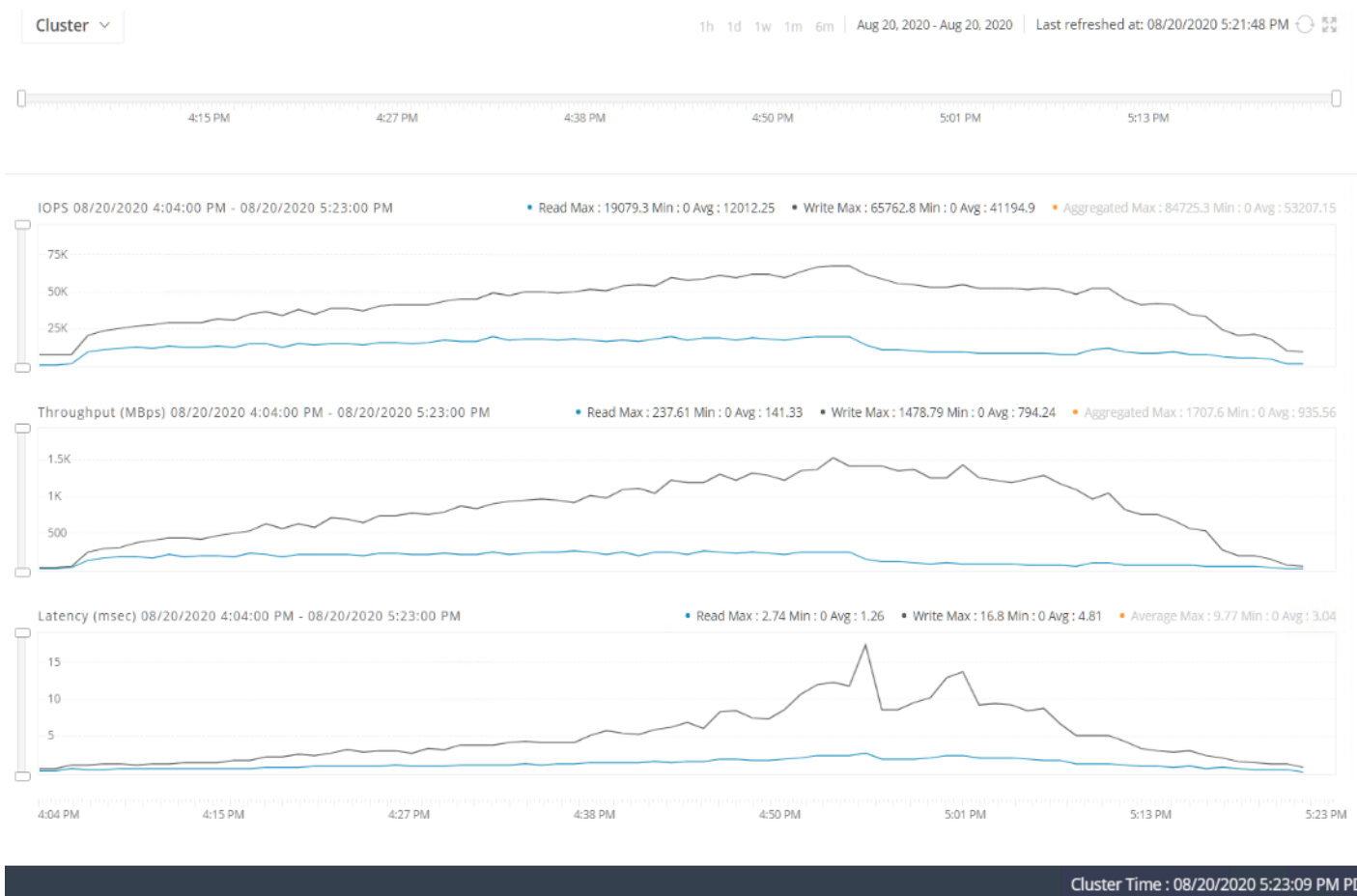
Figure 68 HyperFlex Cluster UI Performance Chart for Knowledge Worker Workload Running 4000 User Test on VMware Windows 10 persistent Virtual Desktops

Summary

This Cisco HyperFlex solution addresses the urgent requirements of IT by delivering a platform that is cost effective and simple to deploy and manage. The architecture and approach used provides for a flexible and high-performance system with a familiar and consistent management model from Cisco. In addition, the solution offers numerous enterprise-class data management features to deliver the next-generation hyper-converged system.

Only Cisco offers the flexibility to add compute only nodes to a true hyper-converged cluster for compute intensive workloads like desktop virtualization. This translates to lower cost for the customer since no hyper-convergence licensing is required for those nodes. This CVD demonstrates a 2:1 compute only to hyperconverged node configuration that would save customers very significant licensing costs.

Delivering responsive, resilient, high-performance VMware Horizon provisioned Microsoft Windows 10 Virtual Machines and Microsoft Windows Server sessions for hosted Apps or desktops has many advantages for desktop virtualization administrators.

The solution is fully capable of supporting graphics accelerated workloads. Each Cisco HyperFlex server can support up to two NVIDIA T4 cards. The Cisco UCS B200 M5 server supports up to two NVIDIA P6 cards for high

density, high performance graphics workload support. For customers who need higher graphics density, our HXAF240c M5 and HX-C240 M5 server can accommodate up to six NVIDIA T4 cards, up to two RTX6000 cards or up to two RTX8000 cards. See our Cisco Graphics White Paper for our fifth generation servers with NVIDIA GPUs and software for details on how to integrate this capability with VMware Virtual Desktops.

The Virtual desktop end-user experience, as measured by the Login VSI tool in benchmark mode, is outstanding with Intel Xeon 2nd Generation Scalable Family processors and Cisco 2933Mhz memory. In fact, we have set a new industry standard in performance for Desktop Virtualization on a hyper-converged platform.

About the Author

Ramesh Guduru, Technical Marketing Engineer, Desktop Virtualization and Graphics Solutions, Cisco Systems, Inc.

Ramesh Guduru is a Cisco Unified Computing System architect, focusing on Virtual Desktop and Application solutions with VMware ESX, Virtual Desktops, Remote Desktop Server Sessions and Microsoft Remote Desktop Services. He has expert product knowledge in application, desktop, and server virtualization across all the major hypervisor platforms and supporting infrastructures including but not limited to Windows Active Directory and Group Policies, User Profiles, DNS, DHCP, Cisco Hyperflex System and major storage platforms.

Mike Brennan, Sr. Product Manager, Desktop Virtualization and Graphics Solutions, Cisco Systems, Inc.

Mike Brennan is Technical and Product Marketing Manager for Virtual Desktop and Graphics Solutions with Cisco Cloud and Compute Business Unit.