



The bridge to possible

Cisco HyperFlex 5.0 for Citrix VDI with VMware ESXi for up to 1000 Users

Cisco Public

Cisco HyperFlex 5.0 for Citrix VDI with VMware ESXi for up to 1000 Users

Deployment Guide for Cisco HyperFlex 5.0 for Citrix VDI using Cisco HyperFlex HX240 M6 Nodes



Document Organization

This document is organized into the following chapters:

- [Executive Summary](#)
- [Solution Overview](#)
- [Technology Overview](#)
- [Solution Design](#)
- [Design Elements](#)
- [Installation](#)
- [Validate](#)
- [Summary](#)
- [Appendices](#)

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>.

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 1000 VDI session workload on a 4-node Cisco HyperFlex system. We provide deployment guidance and performance data for Citrix Virtual Desktops 1912 LTSR virtual desktops running Microsoft Windows 10 with Office 2016 and Windows Server 2019 for HSD. The solution is a pre-integrated, best-practice data center architecture built on the Cisco Unified Computing System (Cisco UCS), the Cisco Nexus® 9000 family of switches and Cisco HyperFlex Data Platform software version 5.0.

The solution payload is 100 percent virtualized on Cisco HyperFlex HXAF240C-M6 hyperconverged nodes booting through on-board M.2 SATA SSD drive running VMware vSphere 7 hypervisor and the Cisco HyperFlex Data Platform storage controller virtual machine. The virtual desktops are configured with Virtual Desktops 1912 LTSR, which incorporates both non-persistent virtual Windows 10 desktops, and hosted applications and remote desktop service (RDS) Microsoft Server 2019 based desktops. The solution provides unparalleled scale and management simplicity. Citrix Virtual Desktops Provisioning Services or Machine Creation Services Windows 10 desktops, full clone desktops or Virtual Apps server-based desktops 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

This chapter is organized into the following subjects:

- [Audience](#)
- [Purpose of this Document](#)
- [Documentation Roadmap](#)
- [Solution Summary](#)

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, Citrix 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 four different Citrix Virtual Desktops/Virtual Apps workloads with Cisco UCS 6400 series Fabric Interconnects and Cisco Nexus 9000 series switches.

Documentation Roadmap

For the comprehensive documentation suite, refer to the Cisco UCS HX-Series Documentation Roadmap: https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/HX_Documentation_Roadmap/HX_Series_Doc_Roadmap.html.

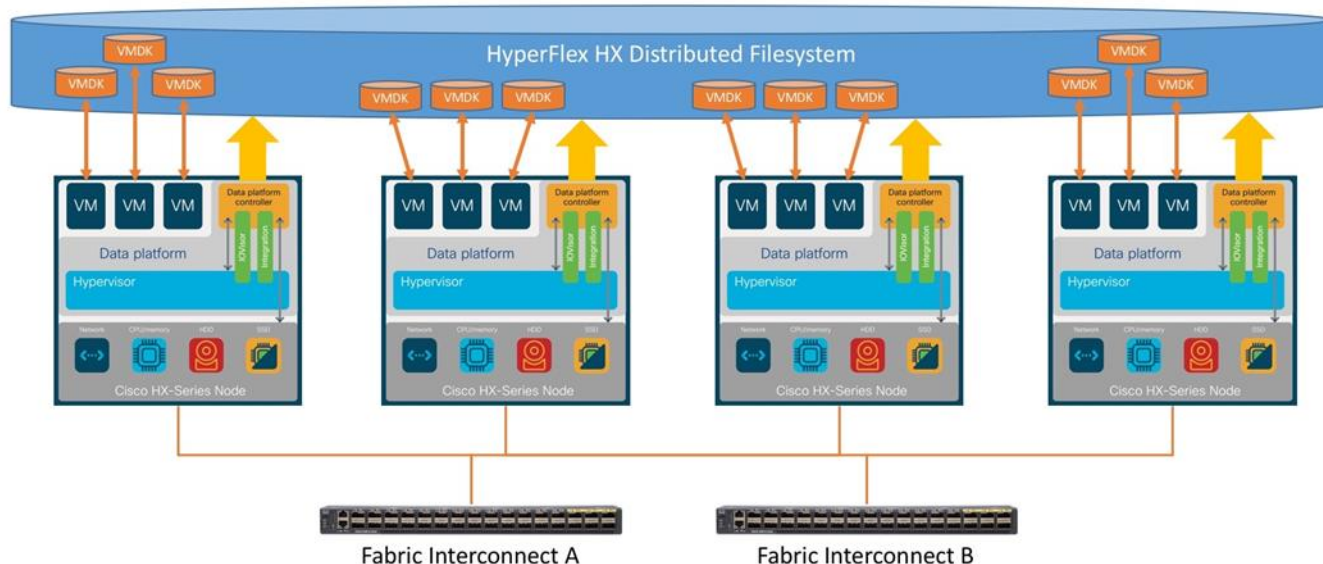
Note: A login is required for the Documentation Roadmap.

Hyperconverged Infrastructure link: <http://hyperflex.io>.

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 1. Cisco HyperFlex 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 HXAF240-M6 All-Flash Rack-Mount Servers
- Cisco HyperFlex Data Platform Software
- VMware vSphere ESXi Hypervisor
- VMware vCenter Server (end-user supplied)

Technology Overview

This chapter is organized into the following subjects:

- [Cisco Unified Computing System](#)
- [Cisco Fabric Interconnect](#)

Cisco Unified Computing System

The 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.

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 backbone for the Cisco UCS C-Series, 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 unified 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 2. Cisco UCS 6454 Fabric Interconnect



Cisco HyperFlex HX240 M6 Node family

The Cisco HyperFlex HX240 M6 Node family delivers high disk capacity (up to 28 drives) in a 2-socket, 2RU package ideal for storage-intensive applications. Physically, the system is delivered as a cluster of three or more Cisco HyperFlex HX240 M6 Nodes, HX240 M6 All NVMe Nodes, HX240 M6 All Flash Nodes, or HX240 M6 LFF Nodes. The nodes are integrated into a single system by a pair of Cisco UCS 6300 or 6400 Series Fabric Interconnects, creating clusters that deliver the performance and storage capacity needed by workloads. All nodes use Intel Xeon Scalable CPUs and next-generation DDR4 memory and offer 12-Gbps SAS throughput.

Figure 3. HXAF240-M6 All-Flash Node



Cisco VIC 1477 MLOM Interface Cards

The Cisco UCS VIC 1477 is a dual-port Small Form-Factor Pluggable (SFP28) mLOM card designed for the M6 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 1477 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 4. Cisco VIC 1477 mLOM Card



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:

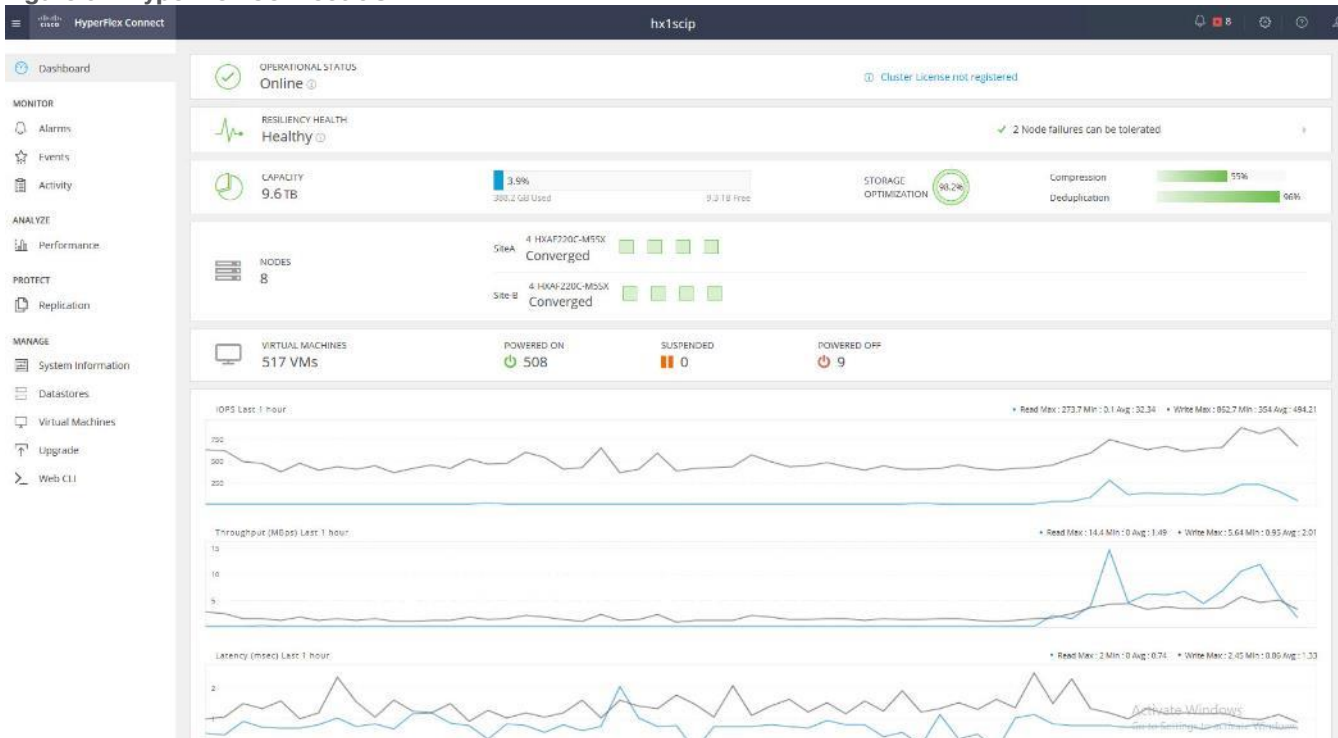
- Data protection creates multiple copies of the data across the cluster so that data availability is not affected if single or multiple components fail (depending on the replication factor configured).
- Stretched clusters allow nodes to be evenly split between two physical locations, keeping a duplicate 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 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>>.

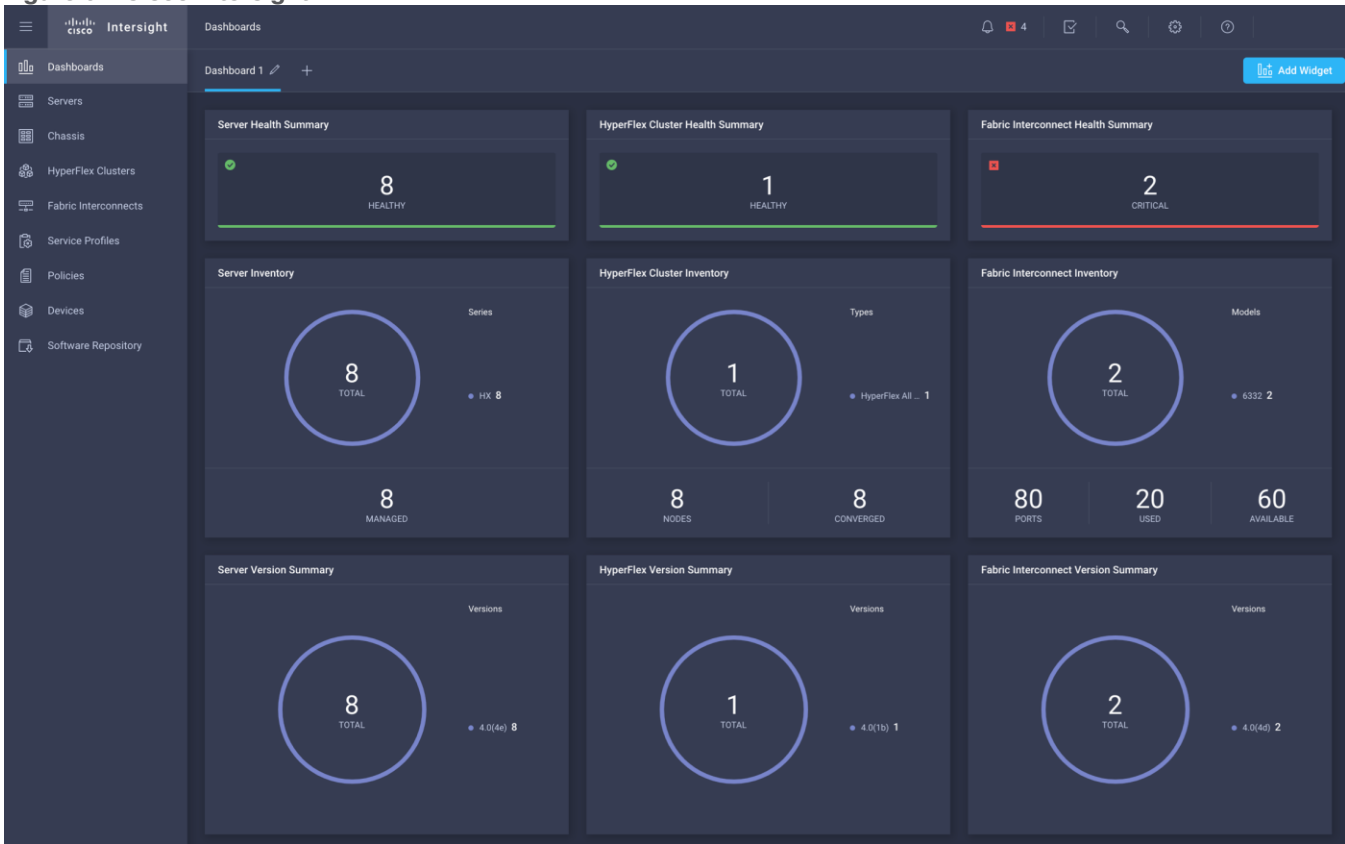
Figure 5. HyperFlex Connect GUI



Cisco Intersight Cloud Based Management

Cisco Intersight (<https://intersight.com>) is the latest visionary cloud-based management tool, designed to provide 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 independently 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 technologies results in a complete cloud-based management solution that can care for Cisco HyperFlex throughout the entire lifecycle, from deployment through retirement.

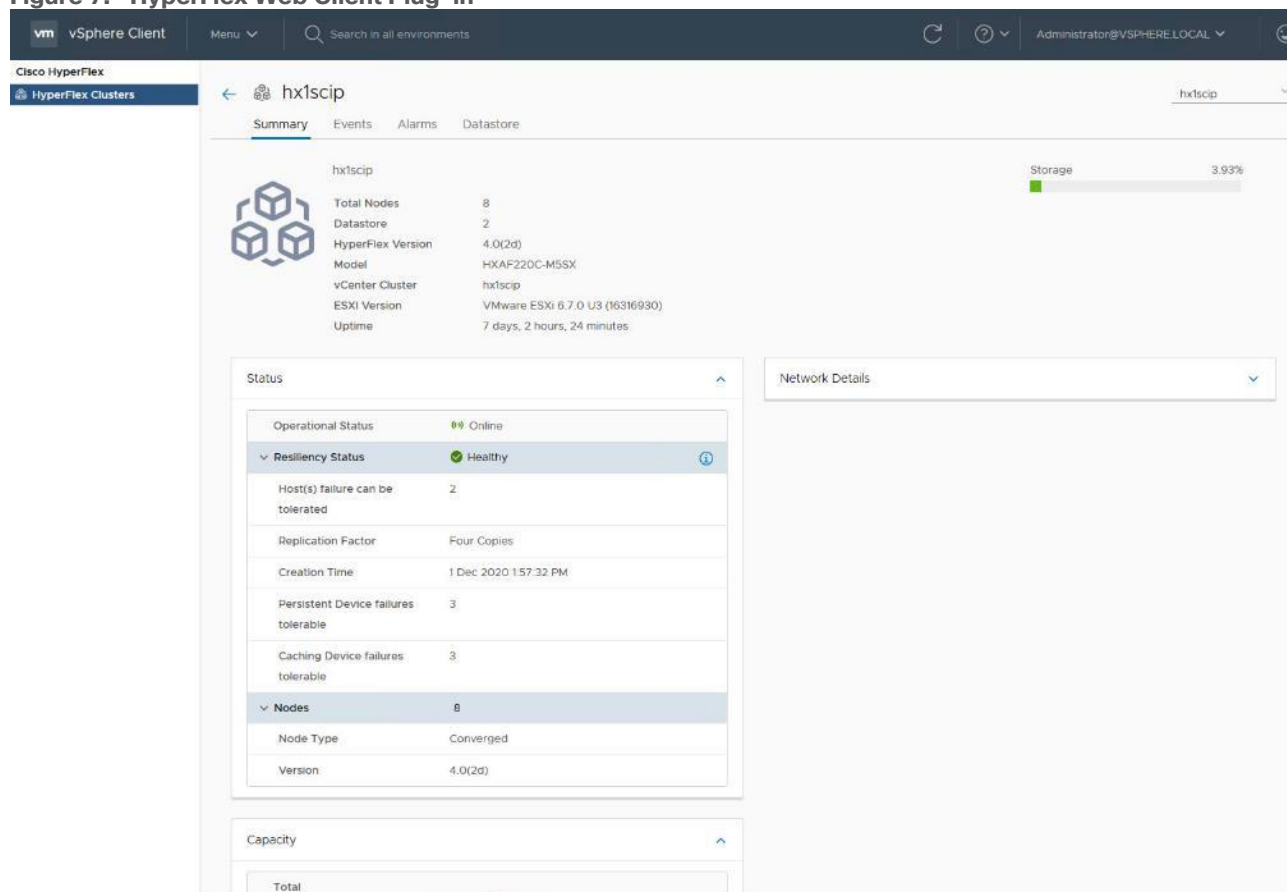
Figure 6. Cisco Intersight



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.

Figure 7. HyperFlex Web Client Plug-in



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 manipulation 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

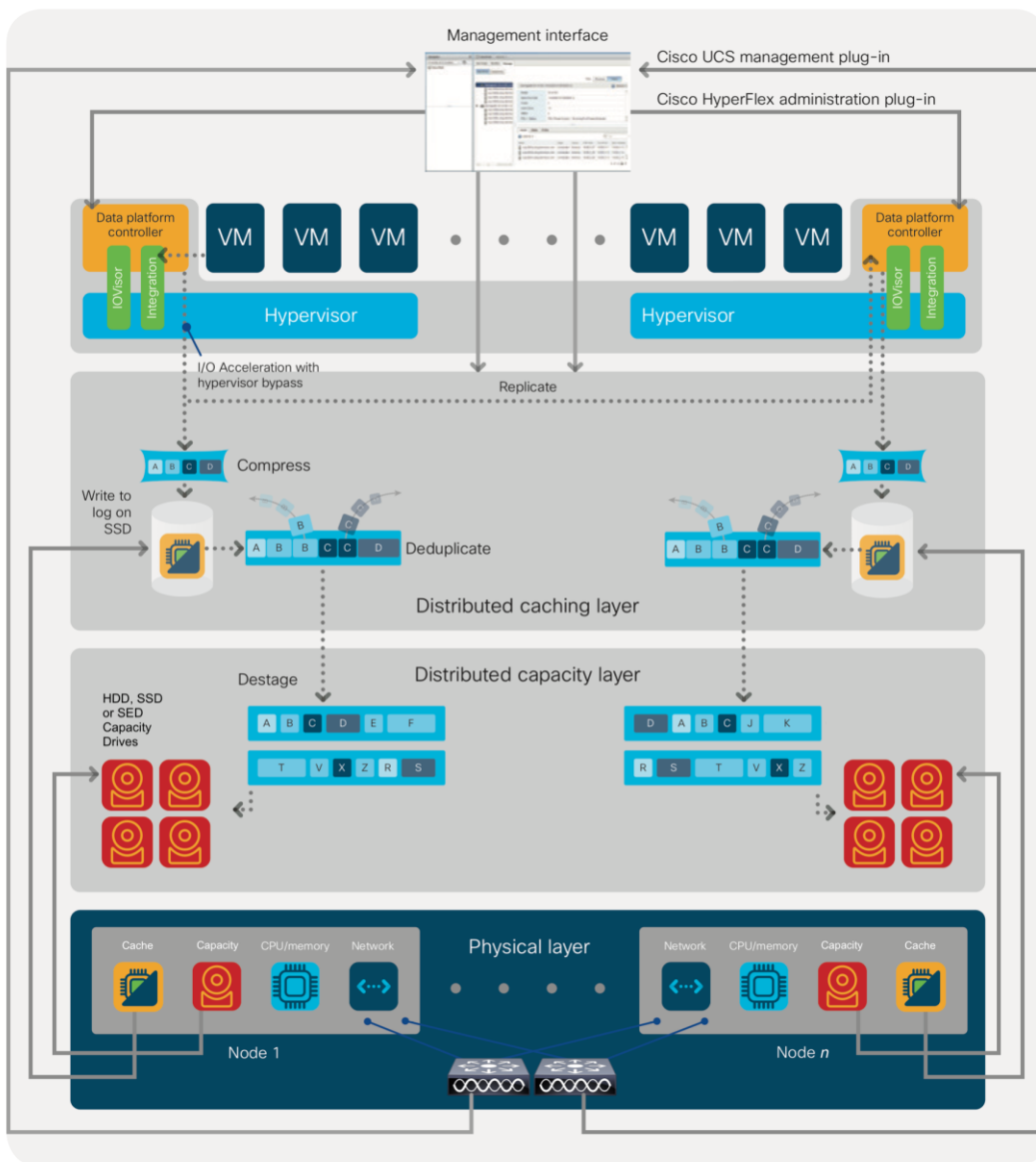
Internally, the contents of each virtual disk are subdivided and spread across multiple servers by the HXDP software. For each write operation, the data is intercepted by the IO Visor module on the node where the VM is running, a primary node is determined for that particular operation via a hashing 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 8. 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.

Solution Design

This chapter is organized into the following subjects:

- [Requirements](#)
- [Considerations](#)

Requirements

This subject is organized into the following sections:

- [Physical Components](#)
- [Software Components](#)
- [Licensing](#)
- [Physical Topology](#)
- [Fabric Interconnects](#)
- [HX-Series Rack-Mount Servers](#)
- [Cisco UCS C-Series Rack-Mount Servers](#)
- [Logical Topology](#)
- [Logical Availability Zones](#)

Physical Components

[Table 1](#) lists the required physical components and hardware.

Table 1. HyperFlex System Components

Component	Hardware
Fabric Interconnects	Four Cisco UCS 6454 Fabric Interconnects
Servers	Four Cisco HyperFlex HXAF240-M6 All-Flash rack servers

For complete server specifications and more information, please refer to the HXAF240-M6 Spec Sheet: <https://www.cisco.com/c/dam/en/us/products/collateral/hyperconverged-infrastructure/hyperflex-hx-series/hyperflex-hx240c-m6-nvme-spec-sheet.pdf>

[Table 2](#) lists the hardware component options for the HXAF240-M6 server model.

Table 2. HXAF240-M6 Server Options

Component	Hardware
Processors	Chose a matching pair of 3rd Generation Intel Xeon 8000 Processor Scalable Family CPUs
Memory	1 TB total memory using 64 GB DDR4 3200 MHz 1.2v modules depending on CPU type
Disk Controller	Cisco 12Gbps Modular SAS HBA

Component		Hardware
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 VIC1477 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		

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 Cisco UCS Hardware Compatibility webpage: <https://ucshcltool.cloudapps.cisco.com/public/>

[Table 3](#) lists the software components and the versions required for the Cisco HyperFlex 4.0 system.

Table 3. Software Components

Component	Software
Hypervisor	VMware ESXi 7 CISCO Custom Image for ESXi 7 for HyperFlex Note: Use of a published Cisco custom ESXi ISO installer file is required when installing/reinstalling ESXi or upgrading to a newer version prior to installing HyperFlex. 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 VIC1477 card, or the HXAF220c-M6N 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 For interoperability of your ESXi version and vCenter Server, refer to http://www.vmware.com/resources/compatibility/sim/interop_matrix.php .

Component	Software
	Using ESXi 6.5 on the HyperFlex nodes also requires using vCenter Server 6.5. Accordingly, using ESXi 6.7 hosts requires using vCenter Server 6.7.
Cisco HyperFlex Data Platform	Cisco HyperFlex HX Data Platform Software 5.0
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 HyperFlex 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, go to:

<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 4](#) lists an overview of the licensing editions, and the features available with each type of license.

Table 4. HyperFlex System License Editions

HyperFlex Licensing Edition	Edge	Standard (in addition to Edge)	Enterprise (in addition to Standard)
Available Features	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	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	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)
	only Compression Deduplication HyperFlex native snapshots Rapid Clones HyperFlex native replication Management via vCenter plugin, HyperFlex Connect, or Cisco Intersight	1:1 ratio 10 Gb, 25 Gb or 40 Gb Ethernet Data-at-rest encryption using self-encrypting disks Logical Availability Zones	

For a comprehensive guide to licensing and all the features in each edition, consult the Cisco HyperFlex Licensing Guide here:

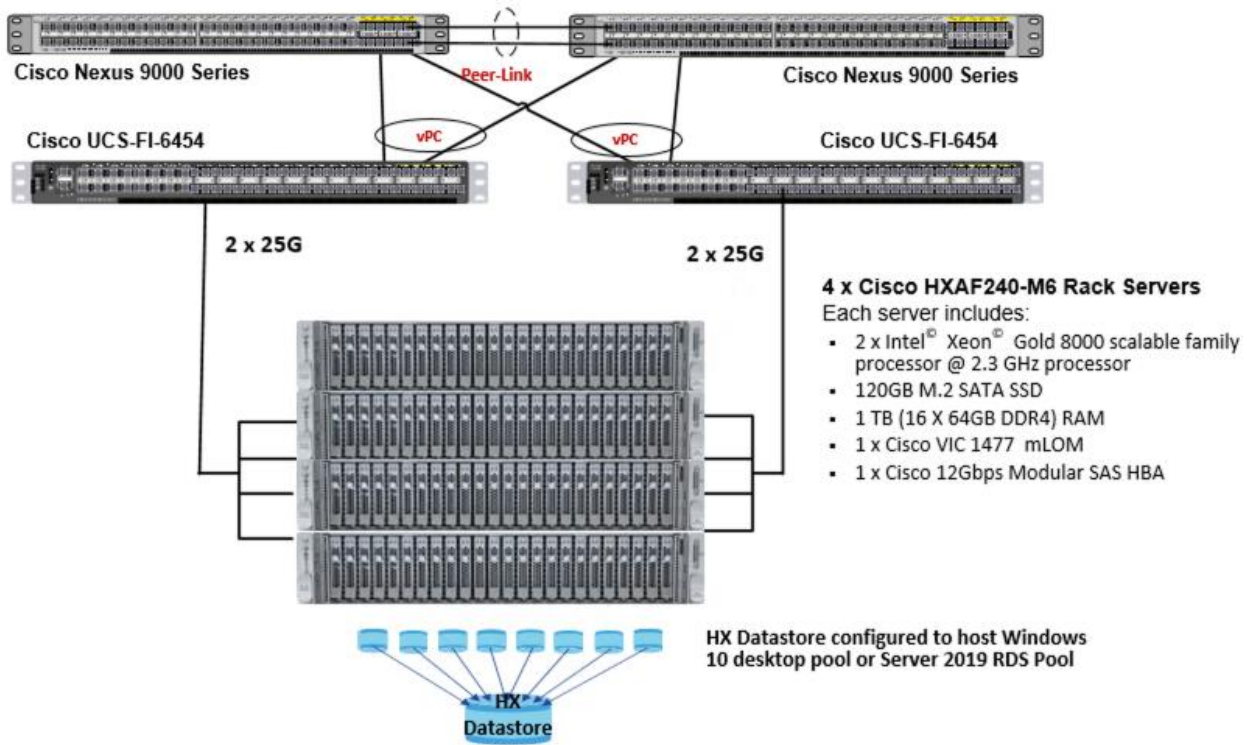
https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/b_Cisco_HyperFlex_Systems_Ordering_and_Licensing_Guide/b_Cisco_HyperFlex_Systems_Ordering_and_Licensing_Guide_chapter_01001.html

Physical Topology

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 “northbound” network connections are made from the Fabric Interconnects to the customer datacenter network at the time of installation.

Figure 9. HyperFlex Cluster Topology

Cisco HyperFlex VDI, Reference Architecture



Fabric Interconnects

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

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 M6 generation servers are configured with the Cisco VIC 1477 cards. The standard and redundant connection practice for the VIC 1477 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 10](#)). For the VIC 1477 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. An optional configuration method for servers containing the Cisco VIC 1477 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.

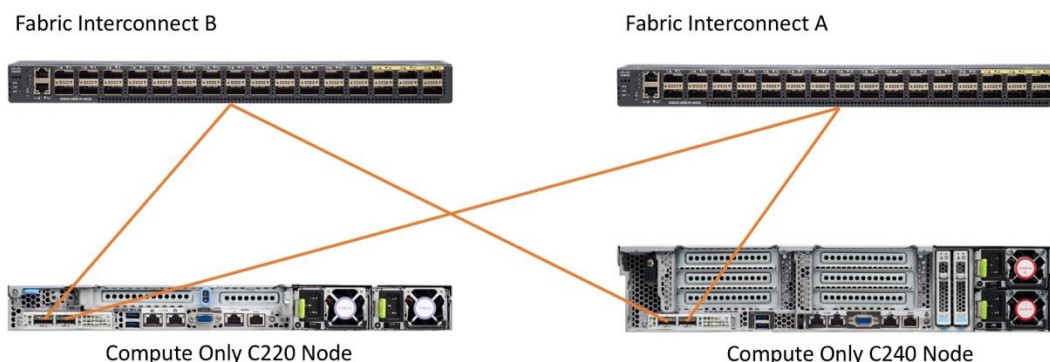
All nodes within a Cisco HyperFlex cluster must be connected at the same communication speed, for example, mixing 10 Gb with 25 Gb interfaces is not allowed. In addition, for clusters that contain only M6 generation nodes, all of the nodes within a cluster must contain the same model of Cisco VIC cards.

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 1477 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.

Note: Failure to follow this cabling practice can lead to errors, discovery failures, and loss of redundant connectivity.

Figure 10. Cisco UCS C-Series Server Connectivity



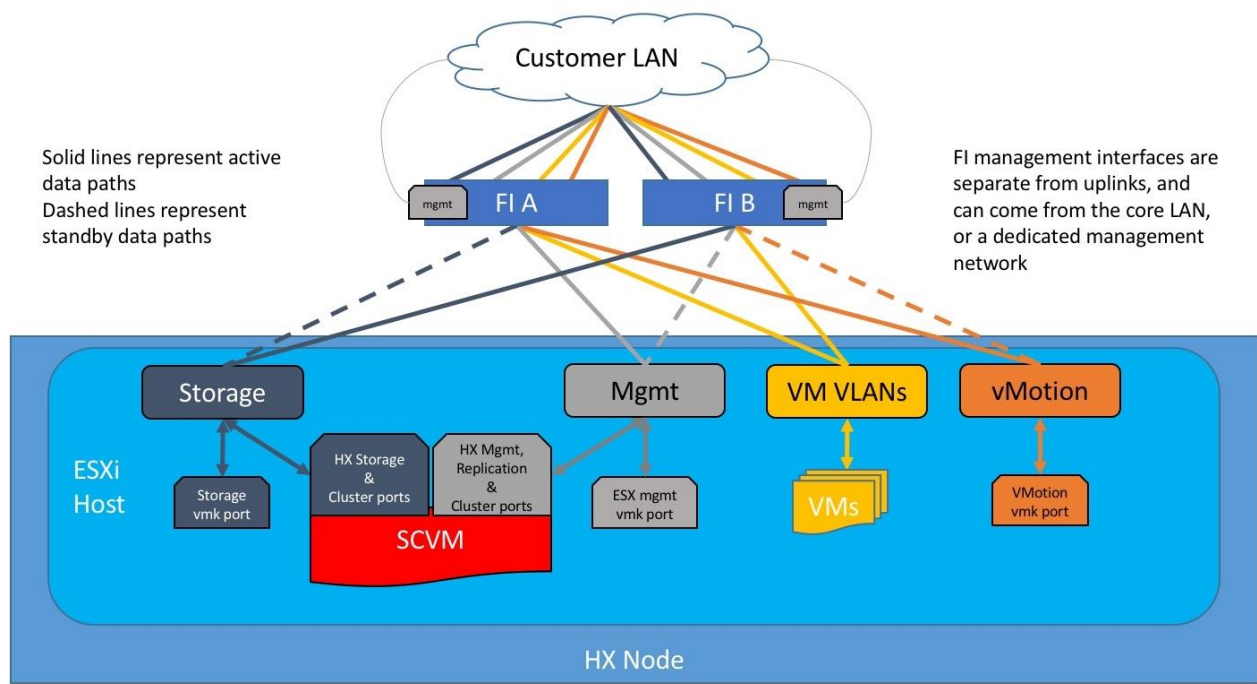
Logical Topology

Logical Network Design

The Cisco HyperFlex system has communication pathways that fall into four defined zones ([Figure 11](#)):

- **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 northbound 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 11. Logical Network Design



Logical Availability Zones

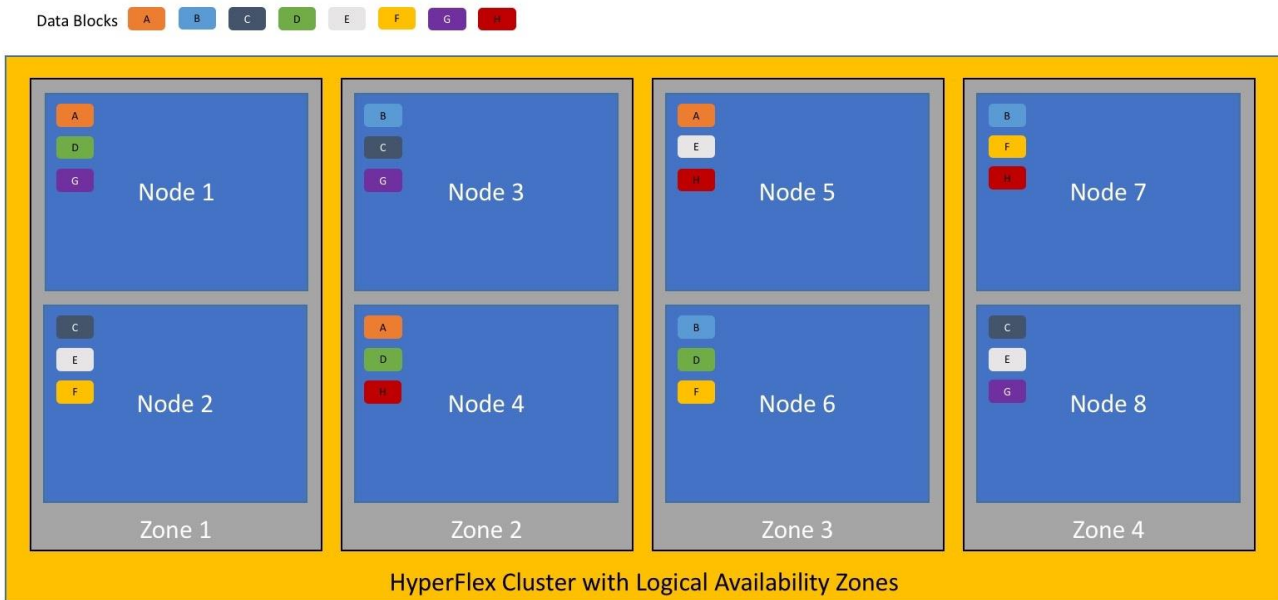
Larger scale HyperFlex clusters are subject to higher failure risks, simply due to the number of nodes in the cluster. While any individual node’s risk of failure is the same no matter how many nodes there are, with clusters up to 32 converged nodes in size, there is a logically higher probability that a single v 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 12](#) 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 12. 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 8 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 7.0 or later is required due to the requirement for TLS 1.2 with Cisco HyperFlex 5.0. The following best practice guidance applies to installations of HyperFlex 5.0:

- 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

Note: This document does not cover 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 compute 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-M6N 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-M6L 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-M6L 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.

[Table 5](#) lists the minimum and maximum scale for various installations of the Cisco HyperFlex system:

Table 5. 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 64 per cluster
Stretched with LFF disks	2 per site	8 per site	16 per site	24 per site 48 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 distribution 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 6. 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 7. IEC Unit Values (binary prefix)

Value	Symbol	Name
1024 bytes	KiB	Kibibyte
1024 KiB	MiB	Mebibyte
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 8](#) 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 8. 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
HXAF240-M6	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-M6SX	8	3.8 TB	6	77.1 TiB	51.4 TiB
			15	192.8 TiB	128.5 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	
		960 GB	23	295.7 TiB	197.1 TiB	
			6	19.3 TiB	12.9 TiB	
			15	48.2 TiB	32.1 TiB	
		800 GB		23	73.9 TiB	49.3 TiB
				6	16.1 TiB	10.7 TiB
				15	40.2 TiB	26.8 TiB
				22	58.9 TiB	39.3 TiB
HX240c-M6L	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	

Note: 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. The above table is not a comprehensive list of all capacities and models available.

Design Elements

This chapter is organized into the following subjects:

- [Network Design](#)
- [Cisco UCS Design](#)
- [ESXi Host Design](#)

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](https://www.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.

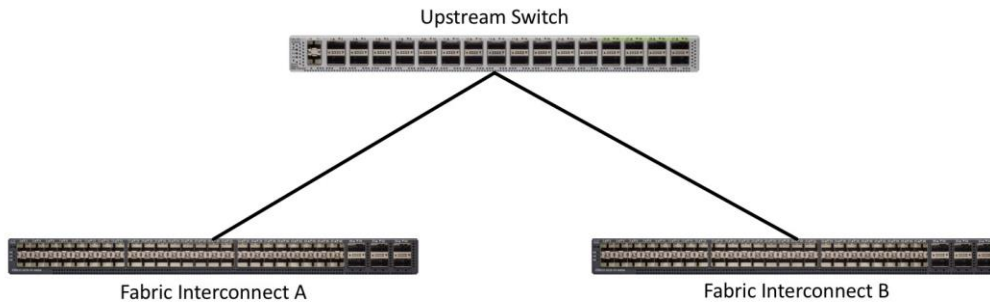
Uplinks need to be connected and active from both Fabric Interconnects. For redundancy, multiple uplinks can be used on each FI, either as 802.3ad Link Aggregation Control Protocol (LACP) port-channels or using individual links. For the best level of performance and redundancy, uplinks can be made as LACP port-channels to multiple upstream Cisco switches using the virtual port channel (vPC) feature. Using vPC uplinks allows all uplinks to be active passing data, plus protects against any individual link failure, and the failure of an upstream switch. Other uplink configurations can be redundant, 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.

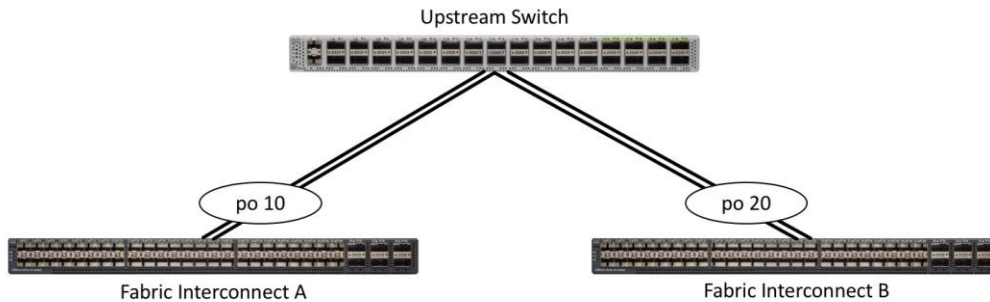
Figure 13. Connectivity with Single Uplink to Single Switch



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.

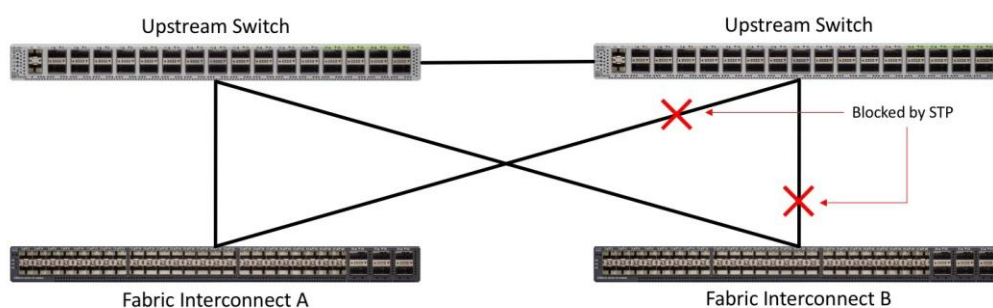
Figure 14. Connectivity with Port-Channels to Single Switch



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 the figure below could also be port-channels.

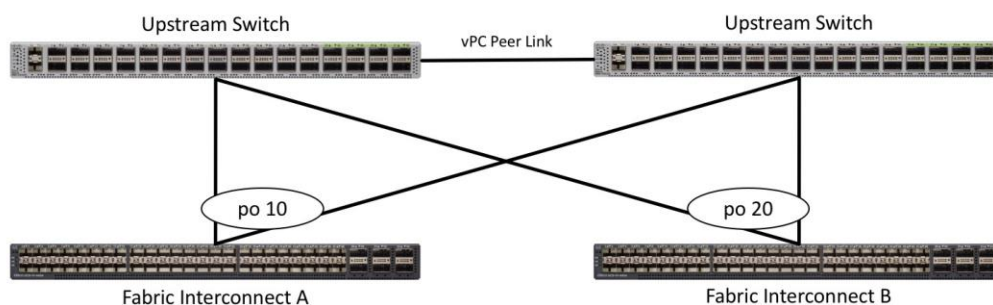
Figure 15. Connectivity with Multiple Uplink Switches



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 16. 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 9](#) lists the VLANs created by the HyperFlex installer in Cisco UCS, and their functions:

Table 9. 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

VLAN Name	VLAN ID	Purpose
		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

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

Jumbo Frames

All HyperFlex storage traffic traversing the hx-storage-data VLAN, and subnet is configured 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

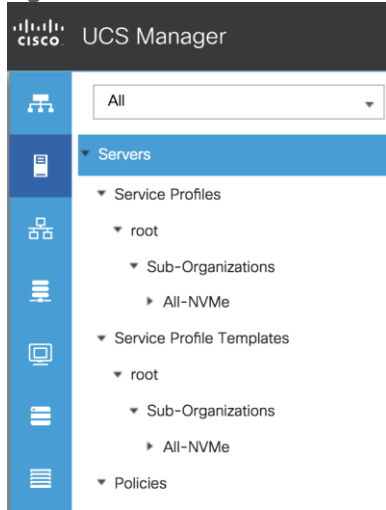
This section describes the elements within Cisco UCS Manager that are configured by the Cisco HyperFlex installer. Many of the configuration elements are fixed in nature, meanwhile the HyperFlex installer does allow for some items to be specified at the time of creation, for example VLAN names and IDs, 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 17. 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 10](#) and [Figure 18](#) details the QoS System Class settings configured for HyperFlex.

Table 10. 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 18. QoS System Classes

Priority	Enabled	CoS	Packet Drop	Weight	Weight (%)	MTU	Multicast Optimized
Platinum	<input checked="" type="checkbox"/>	5	<input type="checkbox"/>	4	25	9216	<input type="checkbox"/>
Gold	<input checked="" type="checkbox"/>	4	<input checked="" type="checkbox"/>	4	25	normal	<input type="checkbox"/>
Silver	<input checked="" type="checkbox"/>	2	<input checked="" type="checkbox"/>	best-effort	6	normal	<input checked="" type="checkbox"/>
Bronze	<input checked="" type="checkbox"/>	1	<input checked="" type="checkbox"/>	best-effort	6	9216	<input type="checkbox"/>
Best Effort	<input checked="" type="checkbox"/>	Any	<input checked="" type="checkbox"/>	best-effort	6	normal	<input type="checkbox"/>
Fibre Channel	<input checked="" type="checkbox"/>	3	<input type="checkbox"/>	5	32	fc	N/A

Note: 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 11](#) lists the QoS Policies configured for HyperFlex, and their default assignment to the vNIC templates created.

Table 11. 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 12](#) and [Figure 19](#) details the Multicast Policy configured for HyperFlex.

Table 12. Multicast Policy

Name	IGMP Snooping State	IGMP Snooping Querier State
HyperFlex	Enabled	Disabled

Figure 19. Multicast Policy

Properties

Name : **HyperFlex**

IGMP Snooping State : Enabled Disabled

IGMP Snooping Querier State : Enabled Disabled

Owner : **Local**

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 13](#) details the VLANs configured for HyperFlex.

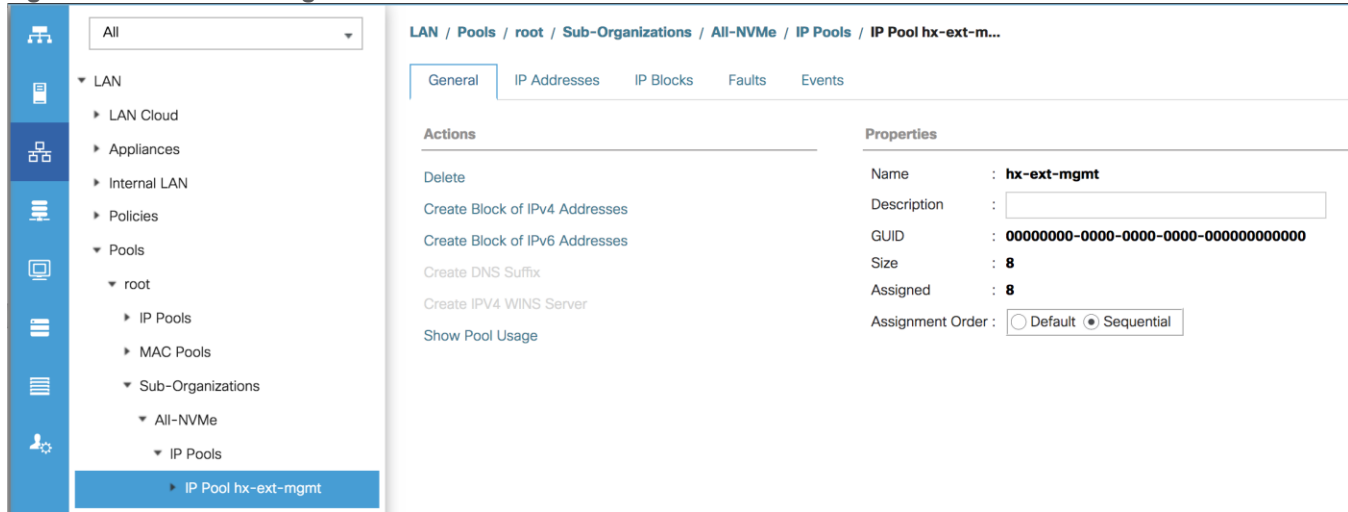
Table 13. Cisco UCS VLANs

Name	ID	Type	Transport	Native	VLAN Sharing	Multicast Policy
<<hx-inband-mgmt>>	<<user_defined>>	LAN	Ether	No	None	HyperFlex
<<hx-inband-repl>>	<<user_defined>>	LAN	Ether	No	None	HyperFlex
<<hx-storage-data>>	<<user_defined>>	LAN	Ether	No	None	HyperFlex
<<vm-network>>	<<user_defined>>	LAN	Ether	No	None	HyperFlex
<<hx-vmotion>>	<<user_defined>>	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 20. 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 (for example, 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 14](#) details the MAC Address Pools configured for HyperFlex and their default assignment to the vNIC templates created.

Table 14. 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

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

Figure 22. Network Control Policy

Properties

Name : **HyperFlex-infra**

Description : Network Control policy for infrastructure vNICs Hype

Owner : **Local**

CDP : Disabled Enabled

MAC Register Mode : Only Native Vlan All Host Vlans

Action on Uplink Fail : Link Down Warning

MAC Security

Forge : Allow Deny

LLDP

Transmit : Disabled Enabled

Receive : Disabled Enabled

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 16. vNIC Template hv-mgmt-a

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

vNIC Template Name	hv-mgmt-a	
Fabric Failover	Disabled	
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 17. 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 18. vNIC Template hv-vmotion-a

vNIC Template Name	hv-vmotion-a	
Setting	Value	
Fabric ID	A	
Fabric Failover	Disabled	
Target	Adapter	
Type	Updating Template	
MTU	9000	

vNIC Template Name	hv-vmotion-a	
MAC Pool	hv-vmotion-a	
QoS Policy	bronze	
Network Control Policy	HyperFlex-infra	
VLANs	<<hx-vmotion>>	Native: No

Table 19. vNIC Template hx-vmotion-b

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 20. vNIC Template storage-data-a

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 21. 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 22. 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	
MAC Pool	vm-network-a	
QoS Policy	gold	
Network Control Policy	HyperFlex-vm	
VLANs	<<vm-network>>	Native: no

Table 23. vNIC Template vm-network-b

vNIC Template Name		vm-network-b
Setting	Value	
Fabric ID	B	
Fabric Failover	Disabled	

vNIC Template Name	vm-network-b	
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 24](#) details the LAN Connectivity Policy configured for HyperFlex.

Table 24. 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 23. Cisco UCS Adapter Policy Resources

⊖ Resources

Pooled	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Transmit Queues	:	<input type="text" value="1"/>	[1-1000]
Ring Size	:	<input type="text" value="256"/>	[64-4096]
Receive Queues	:	<input type="text" value="1"/>	[1-1000]
Ring Size	:	<input type="text" value="512"/>	[64-4096]
Completion Queues	:	<input type="text" value="2"/>	[1-2000]
Interrupts	:	<input type="text" value="4"/>	[1-1024]

Figure 24. Cisco UCS Adapter Policy Options

⊖ Options

Transmit Checksum Offload	:	<input type="radio"/> Disabled <input checked="" type="radio"/> Enabled	
Receive Checksum Offload	:	<input type="radio"/> Disabled <input checked="" type="radio"/> Enabled	
TCP Segmentation Offload	:	<input type="radio"/> Disabled <input checked="" type="radio"/> Enabled	
TCP Large Receive Offload	:	<input type="radio"/> Disabled <input checked="" type="radio"/> Enabled	
Receive Side Scaling (RSS)	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Accelerated Receive Flow Steering	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Network Virtualization using Generic Routing Encapsulation	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Virtual Extensible LAN	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Failback Timeout (Seconds)	:	<input type="text" value="5"/>	[0-600]
Interrupt Mode	:	<input checked="" type="radio"/> MSI X <input type="radio"/> MSI <input type="radio"/> IN Tx	
Interrupt Coalescing Type	:	<input checked="" type="radio"/> Min <input type="radio"/> Idle	
Interrupt Timer (us)	:	<input type="text" value="125"/>	[0-65535]
RoCE	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Advance Filter	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	
Interrupt Scaling	:	<input checked="" type="radio"/> Disabled <input type="radio"/> Enabled	

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 M6 generation servers no longer use pre-defined 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 here:

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-M6” is created by the HyperFlex installer to modify the setting of M6 generation servers. The settings modified 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-M6” 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 M6 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-M6” specifying boot from the M.2 SSDs, referred to as “Embedded Disk”, which is used by the HyperFlex M6 converged nodes, and should not be modified. The HyperFlex installer configures a boot policy named “hx-compute-M6”, which can be modified as needed for the boot method used by the M6 generation compute-only nodes. [Figure 25](#) details the HyperFlex Boot Policy.

Figure 25. Cisco UCS M6 Boot Policy

Actions	Properties
Delete	Name : HyperFlex-m5
Show Policy Usage	Description : Recommended boot policy for HyperFlex servers
Use Global	Owner : Local
	Reboot on Boot Order Change : <input type="checkbox"/>
	Enforce vNIC/vHBA/iSCSI Name : <input checked="" type="checkbox"/>
	Boot Mode : <input checked="" type="radio"/> Legacy <input type="radio"/> Uefi

Warning

The type (primary/secondary) does not indicate a boot order presence.
 The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determined by PCIe bus scan order.
 If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a config error will be reported.
 If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the lowest PCIe bus scan order is used.

Local Devices	Boot Order																		
<input type="checkbox"/> Local Devices	<table border="1"> <thead> <tr> <th>Name</th> <th>Order</th> <th>vNIC/vHB...</th> <th>Type</th> <th>LUN Name</th> <th>WWN</th> </tr> </thead> <tbody> <tr> <td>CD/DVD</td> <td>1</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Embedded Disk</td> <td>2</td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	Name	Order	vNIC/vHB...	Type	LUN Name	WWN	CD/DVD	1					Embedded Disk	2				
Name	Order	vNIC/vHB...	Type	LUN Name	WWN														
CD/DVD	1																		
Embedded Disk	2																		
<input type="checkbox"/> CIMC Mounted vMedia																			
<input type="checkbox"/> vNICs																			
<input type="checkbox"/> vHBAs																			
<input type="checkbox"/> iSCSI vNICs																			
<input type="checkbox"/> EFI Shell																			

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-M6” 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 26](#) details the Host Firmware Package configured by the HyperFlex installer:

Figure 26. Cisco UCS M6 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-M6” is used by the service profile

template named “hx-nodes-M6”, which is for the HyperFlex M6 generation converged servers and should not be modified.

Meanwhile, the policies named “hx-compute” and “hx-compute-M6” are used by the service profile templates named “compute-nodes” and “compute-nodes-M6”, which are used by compute-only nodes. The two compute-only node policies can be modified as needed to suit the local disk configuration that will be used in compute-only nodes.

[Figure 27](#) details the Local Disk Configuration Policy configured by the HyperFlex installer.

Figure 27. Cisco UCS M6 Local Disk Configuration Policy

Actions	Properties
Delete Show Policy Usage Use Global	<p>Name : HyperFlex-m5</p> <p>Description : Recommended Local Disk policy for M5 HyperFlex s</p> <p>Owner : Local</p> <p>Mode : Any Configuration ▾</p> <p>Protect Configuration : <input checked="" type="checkbox"/></p> <p><small>If Protect Configuration is set, the local disk configuration is preserved if the service profile is disassociated with the server. In that case, a configuration error will be raised when a new service profile is associated with that server if the local disk configuration in that profile is different.</small></p> <hr/> <p>FlexFlash</p> <p>FlexFlash State : <input checked="" type="radio"/> Disable <input type="radio"/> Enable</p> <p><small>If FlexFlash State is disabled, SD cards will become unavailable immediately. Please ensure SD cards are not in use before disabling the FlexFlash State.</small></p> <p>FlexFlash RAID Reporting State : <input checked="" type="radio"/> Disable <input type="radio"/> Enable</p> <p>FlexFlash Removable State : <input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> No Change</p> <p><small>If FlexFlash Removable State is changed, SD cards will become unavailable temporarily. Please ensure SD cards are not in use before changing the FlexFlash Removable State.</small></p>

Note: Additional policies are created for use by Cisco M4 generation HX-series servers, including additional BIOS policies, Boot Policies, Host Firmware Packages and Local Disk Configuration Policies. Because this document no longer covers 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 service profile that requires a reboot of the physical server will result in an immediate reboot of that server. The Cisco best practice is to use a Maintenance Policy set to “user-ack,” which requires a secondary acknowledgement by a user with the appropriate rights within Cisco UCS 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 28](#) details the Maintenance Policy configured by the HyperFlex installer.

Figure 28. Cisco UCS Maintenance Policy

Properties

Name : **HyperFlex**

Description : Recommended maintenance policy for HyperFlex ser

Owner : **Local**

Soft Shutdown Timer : 150 Secs

Storage Config. Deployment Policy : Immediate User Ack

Reboot Policy : Immediate User Ack Timer Automatic

On Next Boot (Apply pending changes at next reboot.)

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 29](#) details the Power Control Policy configured by the HyperFlex installer.

Figure 29. Cisco UCS Power Control Policy

Properties

Name : **HyperFlex**

Description : Recommended Power control policy for HyperFlex se

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 30](#) details the Scrub Policy configured by the HyperFlex installer.

Figure 30. 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 administrators 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 31](#) details the SoL Policy configured by the HyperFlex installer.

Figure 31. 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-M6” and “compute-nodes-M6”, 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 25. Cisco UCS Service Profile Template Settings and Values

Service Profile Template Name	hx-nodes-M6
Setting	Value
UUID Pool	Hardware Default
Associated Server Pool	None
Maintenance Policy	HyperFlex
Management IP Address Policy	hx-ext-mgmt
Local Disk Configuration Policy	HyperFlex-M6
LAN Connectivity Policy	HyperFlex
Boot Policy	HyperFlex-M6
BIOS Policy	HyperFlex-M6
Firmware Policy	HyperFlex-M6
Power Control Policy	HyperFlex
Scrub Policy	HyperFlex
Serial over LAN Policy	HyperFlex
vMedia Policy	Not defined

Service Profile Template Name	compute-nodes-M6
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-M6
LAN Connectivity Policy	HyperFlex
Boot Policy	hx-compute-M6

Service Profile Template Name	compute-nodes-M6
BIOS Policy	HyperFlex-M6
Firmware Policy	HyperFlex-M6
Power Control Policy	HyperFlex
Scrub Policy	HyperFlex
Serial over LAN Policy	HyperFlex
vMedia Policy	Not defined

Note: Additional templates are created for use by Cisco UCS M4 generation HX-series servers. Because this document doesn't explain 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 redundancy. 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 interfaces 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 interfaces in a known and identical order, and they will always be connected to the same VLANs via the same network fabrics. [Table 26](#) 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 26. 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-mgmt>>	vmnic4
storage-data-a	1	2	A	<<hx-storage-	vmnic1

vNIC	Placement	Order	Fabric	VLAN	ESXi Interface Enumeration
				data>>	
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

Note: 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

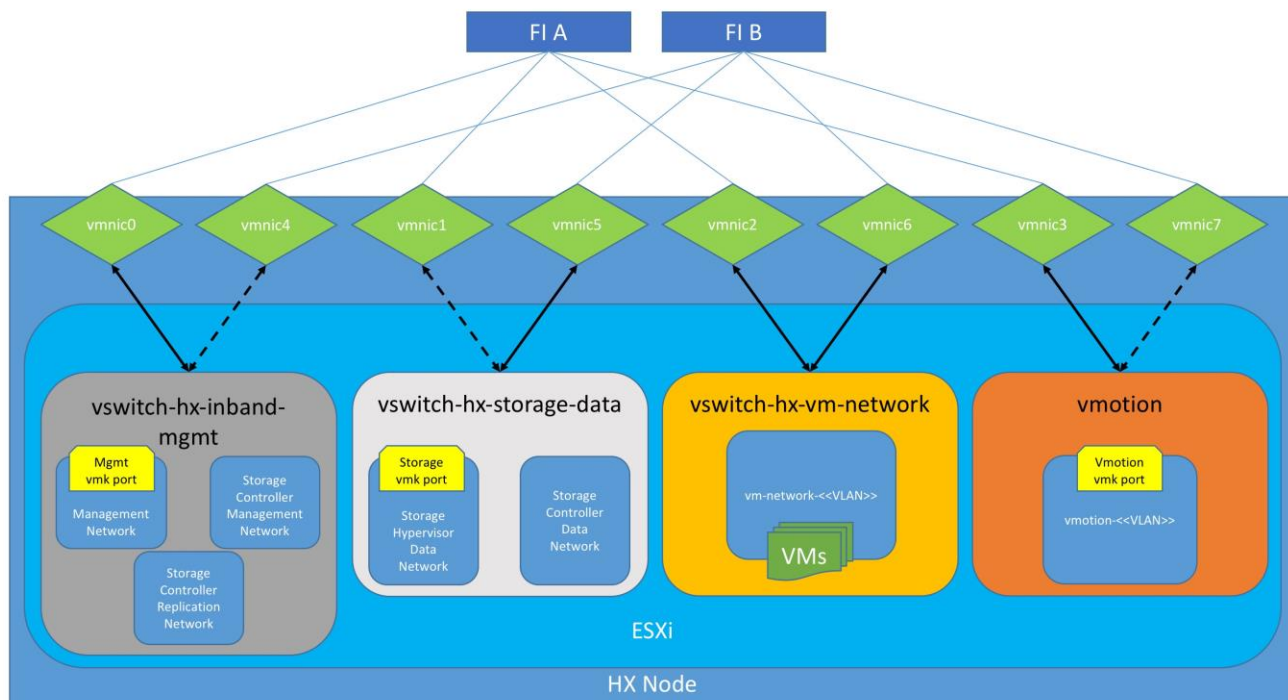
VMkernel ports (vmk2) are configured during the post_install script execution. The VLAN is not a Native VLAN as assigned to the vNIC templates, and therefore they are defined in ESXi/vSphere.

[Table 27](#) and [Figure 32](#) provide more details about the ESXi virtual networking design as built by the HyperFlex installer by default.

Table 27. Virtual Switches

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

Figure 32. 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 driver for the hardware device, the guest VM sends all I/O requests directly to the physical device, bypassing the hypervisor. In the Cisco HyperFlex system, the Storage Platform Controller VMs use this feature to gain full control of the Cisco 12Gbps SAS HBA cards in the Cisco HX-series rack-mount servers. This gives the controller VMs direct hardware level access to the physical disks installed in the servers, which they consume to construct the Cisco HX Distributed Filesystem. 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 different 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 software and services that manage and maintain the Cisco HX Distributed Filesystem. The services and processes that run within the controller VMs are not exposed 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. Management 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 M6, HXAF220c M6, HX240c M6L, HX240c M6 and HXAF240c M6: 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 M6N: 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 33. All M6 Generation Servers Controller VM Placement Except All-NVMe

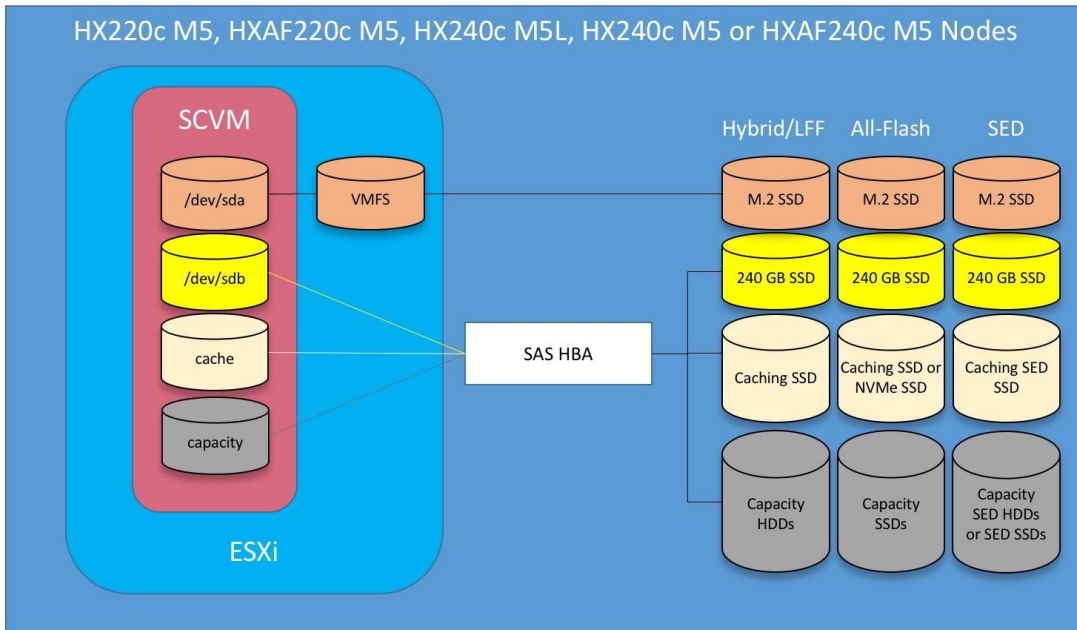
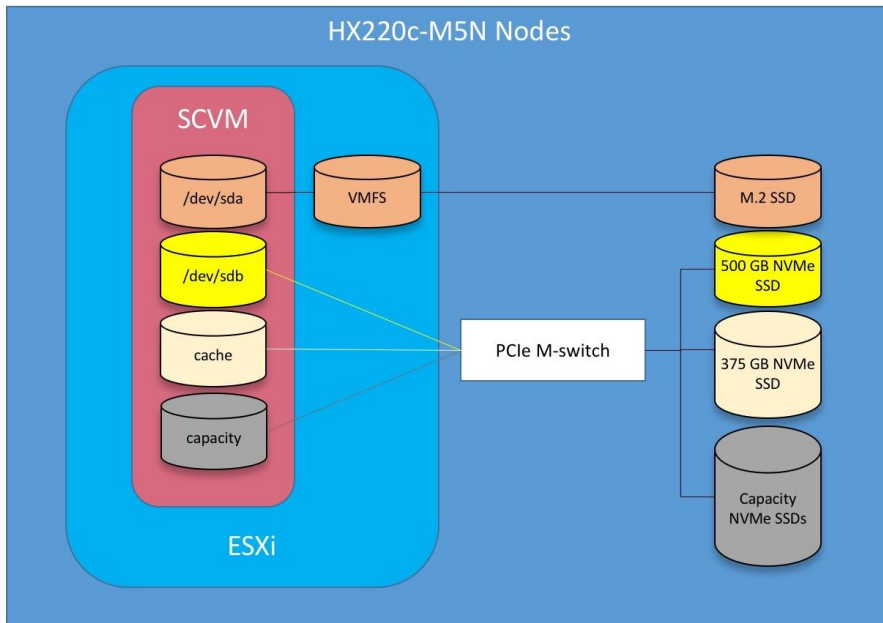


Figure 34. All-NVMe M6 Controller VM Placement



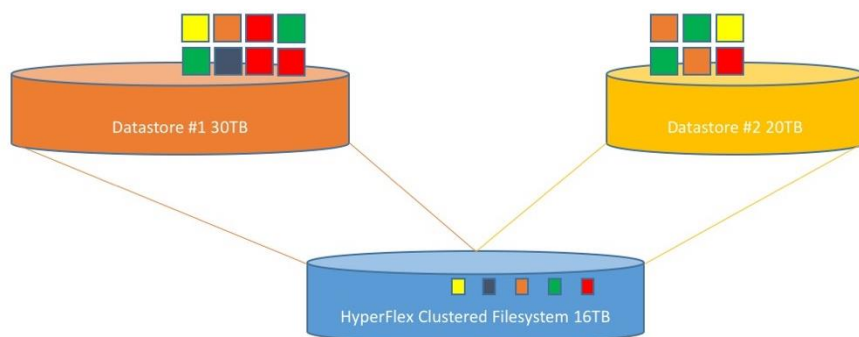
Note: HyperFlex compute-only nodes install a lightweight controller VM in the VMFS datastore automatically created during the installation of ESXi. This VM performs no storage functions and is only used for node coordination.

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 Connect 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 35. 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 28](#) lists the CPU resource reservation of the storage controller VMs.

Table 28. 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 29](#) lists the memory resource reservation of the storage controller VMs.

Table 29. Controller VM Memory Reservations

Server Models	Amount of Guest Memory	Reserve All Guest Memory
HX220c-M6SX	48 GB	Yes

Server Models	Amount of Guest Memory	Reserve All Guest Memory
HXAF240-M6		
HXAF220c-M6N HX240c-M6SX HXAF240c-M6SX	72 GB	Yes
HX240c-M6L	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 preset, 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.

To install a HyperFlex cluster for Virtual Desktop Infrastructure, follow the detailed HX cluster installation found here:

https://www.cisco.com/c/en/us/td/docs/hyperconverged_systems/HyperFlex_HX_DataPlatformSoftware/Installation_VMWare_ESXi/5-0/b-hx-install-guide-for-vmware-esxi-5-0.html

Validate

This chapter is organized into the following subjects:

- [Post-installation Checklist](#)
- [Verify Redundancy](#)

Post-installation Checklist

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

- Verify the expected number of converged storage nodes and compute-only nodes are members of the HyperFlex cluster in the vSphere Web Client plugin manage cluster screen.
- Verify the expected cluster capacity is seen in the HX Connect Dashboard summary screen.
- Create a test virtual machine that accesses the HyperFlex datastore and is able to perform read/write operations.
- Perform a virtual machine migration (vMotion) of the test virtual machine to a different host on the cluster.
- During the vMotion of the virtual machine, make sure the test virtual machine can perform a continuous ping to 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.

- 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.
- 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.
- 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.
- 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.

Note: 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.

- 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.

Build the Virtual Machines and Environment for Workload Testing

This chapter is organized into the following subjects:

- [Software Infrastructure Configuration](#)
- [Prepare the Master Images](#)
- [Install and Configure Citrix Desktop Delivery Controller, Citrix Licensing, and StoreFront](#)
- [Create Delivery Groups](#)
- [Citrix Virtual Desktops Policies and Profile Management](#)

Software Infrastructure Configuration

This subject 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 the following tables.

Configuration	Citrix Virtual Desktops Controllers Virtual Machines	Citrix Provisioning Services Servers Virtual Machines
Operating System	Microsoft Windows Server 2022	Microsoft Windows Server 2022
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
Disk-2 size and location	500GB	Disk-2 (Data) Paravirtual SCSI adapter with ReFS format

Configuration	Microsoft Active Directory DC's Virtual Machines	Citrix Profile Servers Virtual Machines
Operating system	Microsoft Windows Server 2022	Operating system
Virtual CPU amount	4	
Memory amount	4 GB	
Network	VMNIC	
Disk size and location	40 GB	

Configuration	Microsoft SQL Server Virtual Machine	Citrix StoreFront Virtual Machine
Operating system	Microsoft Windows Server 2022	Microsoft Windows Server 2022
Virtual CPU amount	8	4

Configuration	Microsoft SQL Server Virtual Machine	Citrix StoreFront Virtual Machine
Memory amount	16 GB	8 GB
Network	VMNIC	Network
Disk-1 (OS) size and location	40 GB	Disk-1 (OS) size and location
Disk-2 size and location	200 GB Infra-DS volume	Disk-2 size and location

Configuration	Citrix License Server Virtual Machine	NetScaler VPX Appliance Virtual Machine
Operating system	Microsoft Windows Server 2019	NS11.1 52.13.nc
Virtual CPU amount	4	2
Memory amount	4 GB	2 GB
Network	VMNIC	Network
Disk size and location	40 GB	20 GB

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 Hosted Virtual Desktops (HVDs) and Hosted Shared Desktops (HSDs), 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)

[Table 30](#) lists the configuration of the master image HVD and HSD virtual machines.

Table 30. HVD and HSD Configurations

Configuration	HVDI Virtual Machines	HSD Virtual Machines
Operating system	Microsoft Windows 10 64-bit	Microsoft Windows Server 2019
Virtual CPU amount	2	8
Memory amount	4.0 GB	32 GB
Network	VMNIC vm-network	VMNIC vm-network
Citrix PVS vDisk size and location	24 GB	40 GB

Configuration	HVDI Virtual Machines	HSD Virtual Machines
Citrix PVS write cache Disk size	6 GB	24 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)

Install and Configure Citrix Desktop Delivery Controller, Citrix Licensing, and StoreFront

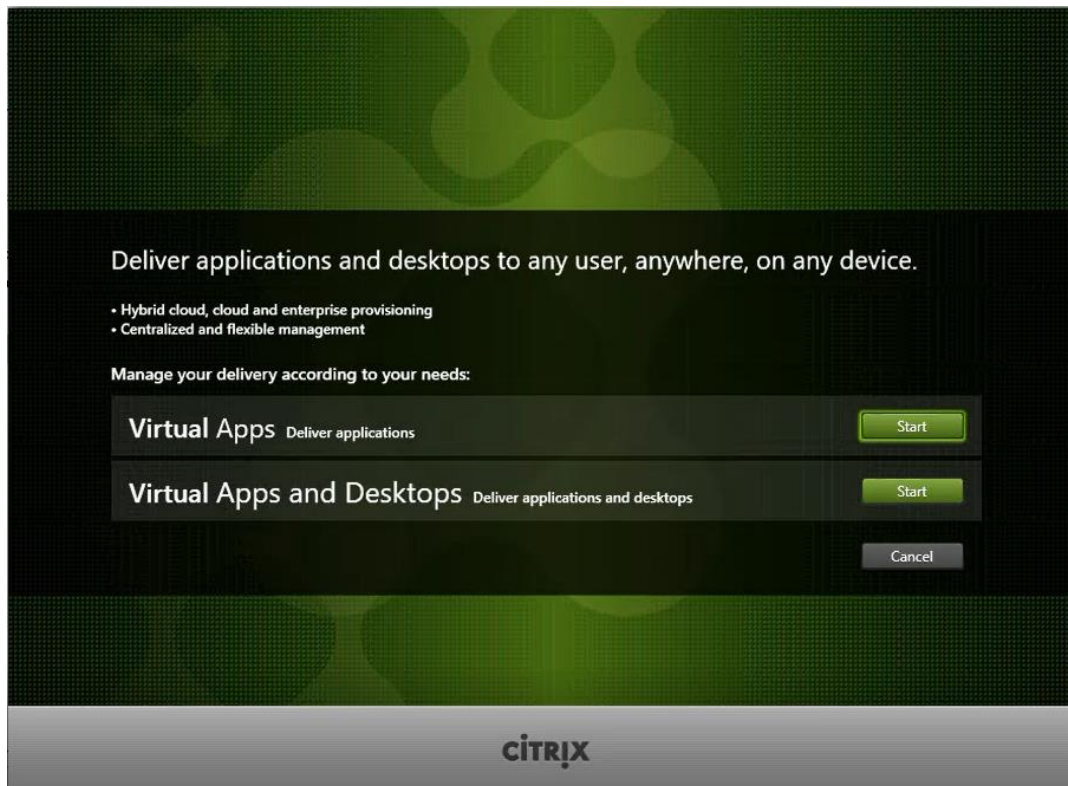
This subject details the installation of the core components of the Citrix Virtual Apps and Desktops 1912 LTSR system. This CVD provides the process to install two Desktop Delivery Controllers to support hosted shared desktops (HSD), non-persistent virtual desktops (VDI), and persistent virtual desktops (VDI).

The process of installing the Desktop Delivery Controller also installs other key Citrix Desktop software components, including Studio, which is used to create and manage infrastructure components, and Director, which is used to monitor performance and troubleshoot problems.

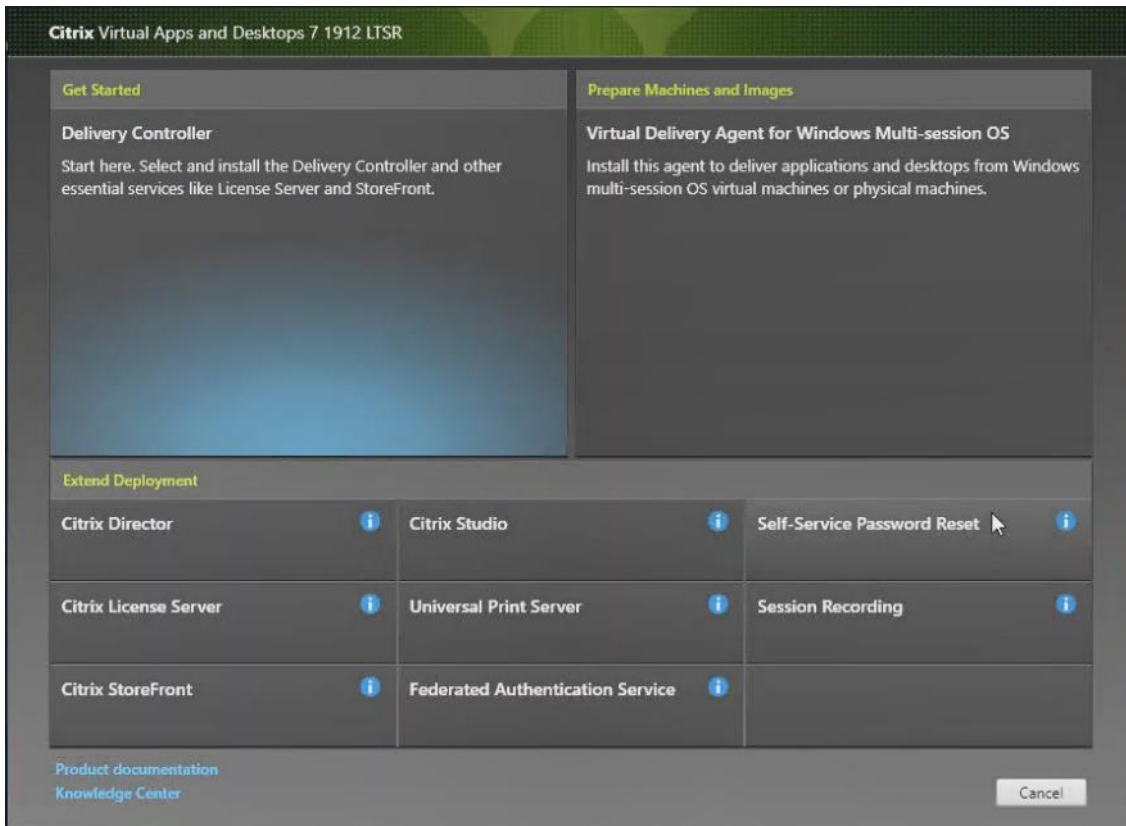
Procedure 1. Install Citrix License Server

Step 1. Connect to the first Citrix License server and launch the installer from the Citrix Virtual Apps and Desktops 1912 LTSR ISO.

Step 2. Click Start.



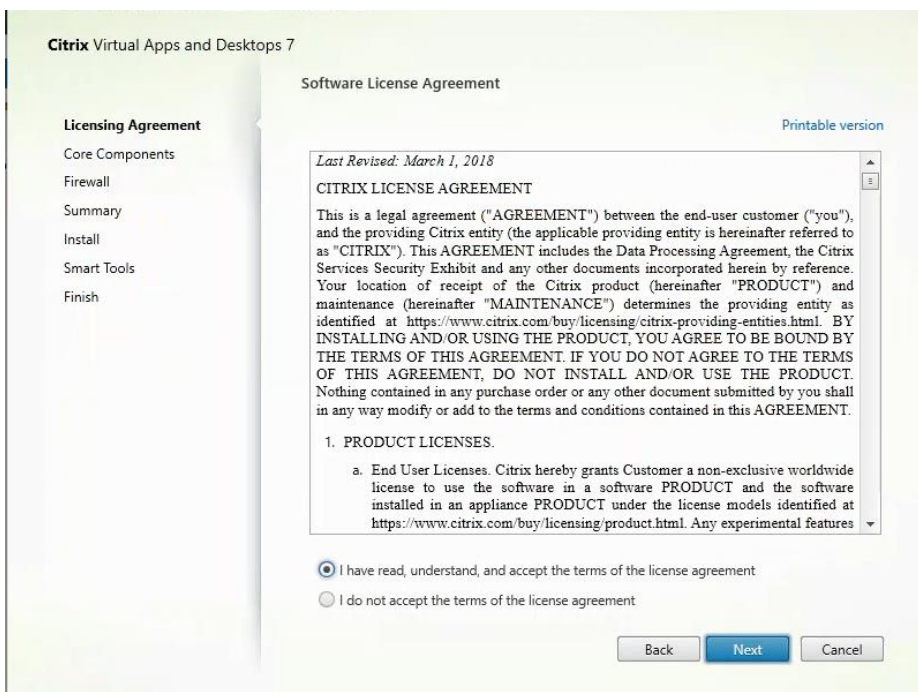
Step 3. Click “Extend Deployment – Citrix License Server.”



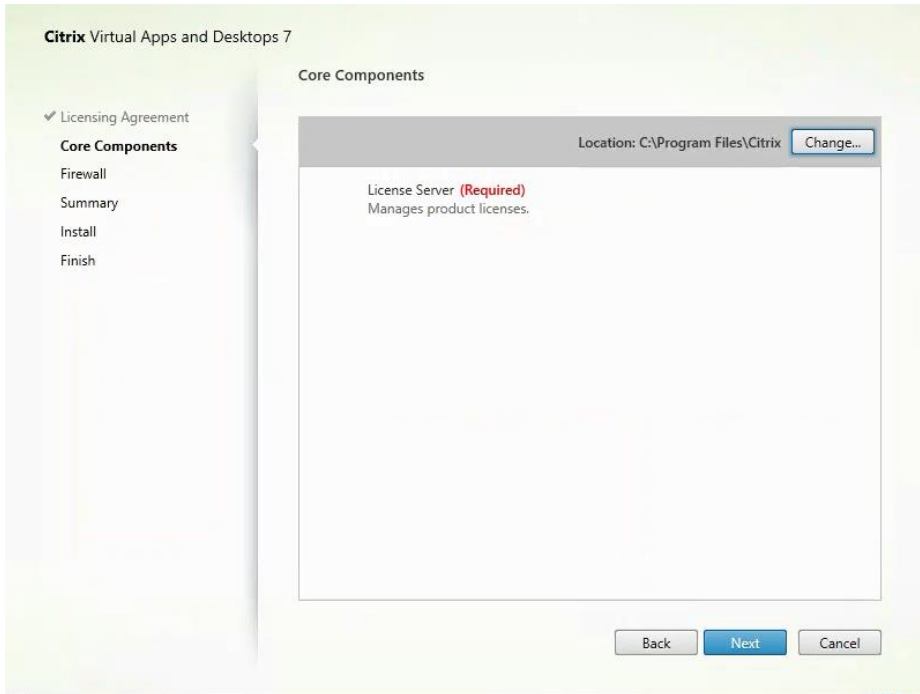
Step 4. Read the Citrix License Agreement.

Step 5. If acceptable, indicate your acceptance of the license by selecting the “I have read, understand, and accept the terms of the license agreement” radio button.

Step 6. Click Next.

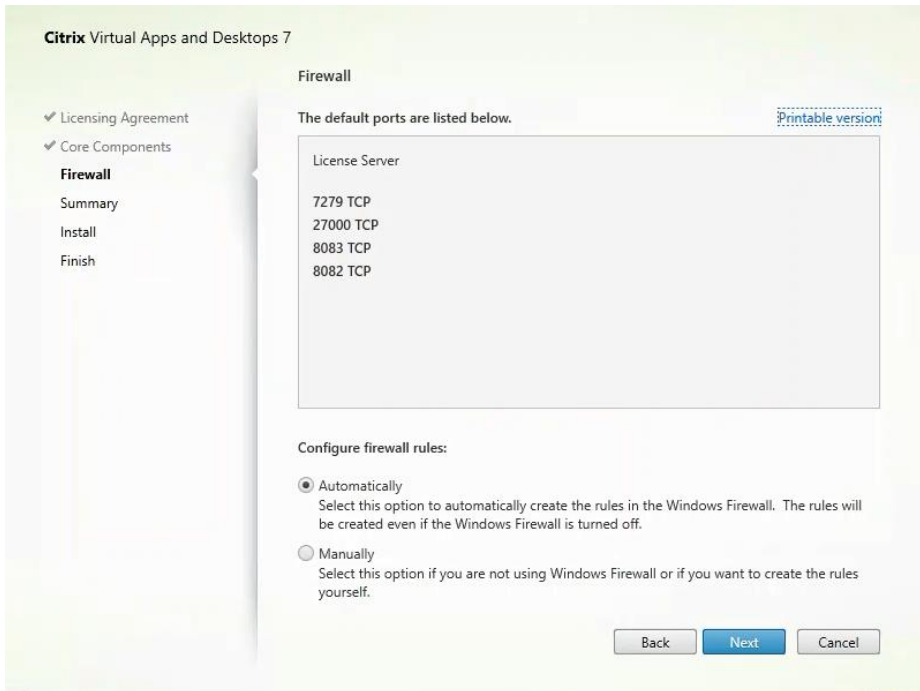


Step 7. Click Next.

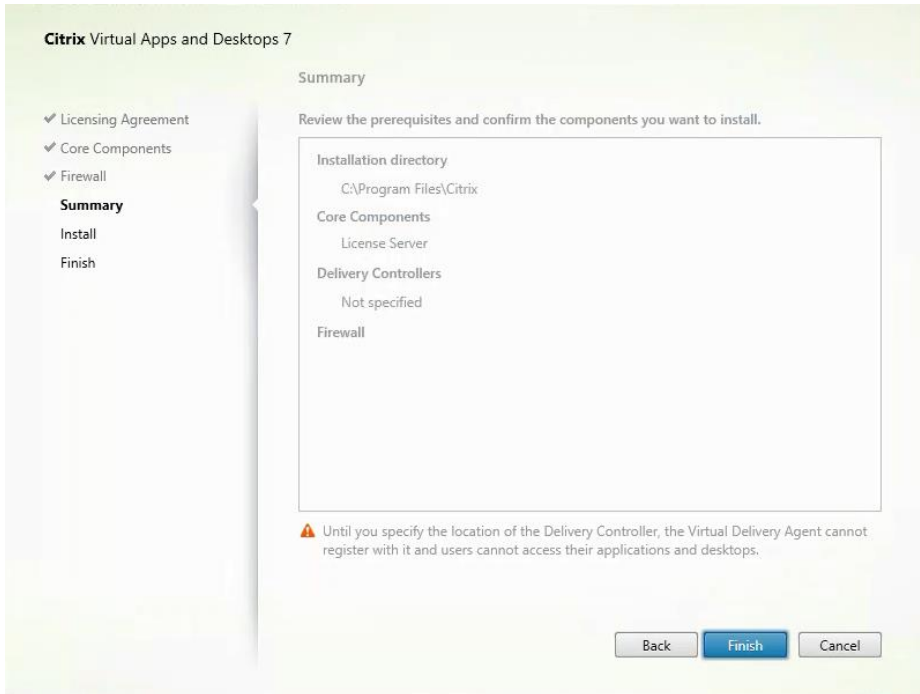


Step 8. Select the default ports and automatically configured firewall rules.

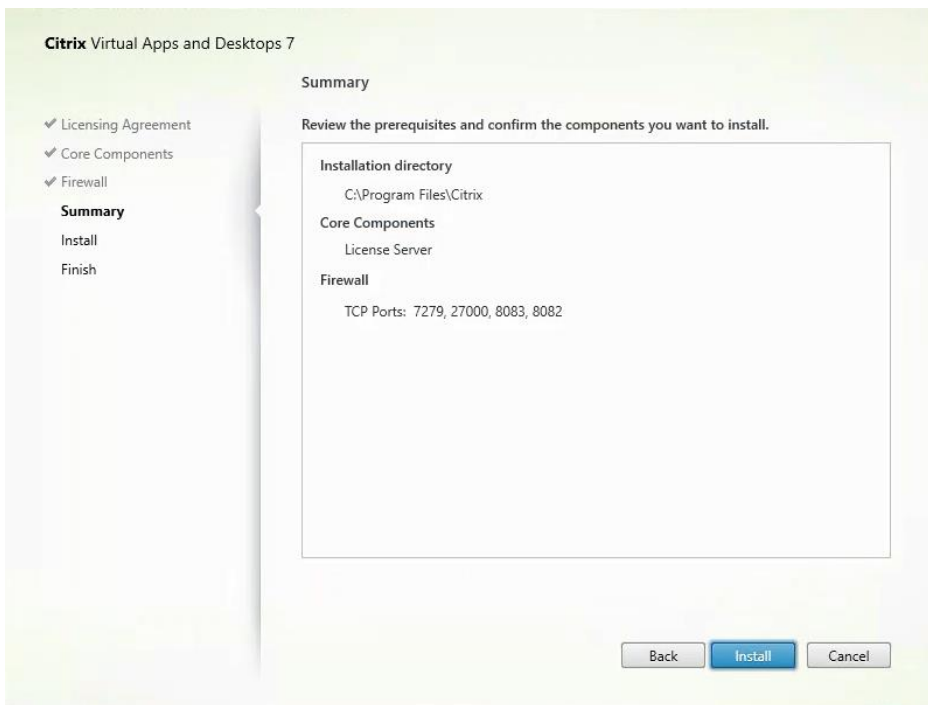
Step 9. Click Next.



Step 10. Click Install.

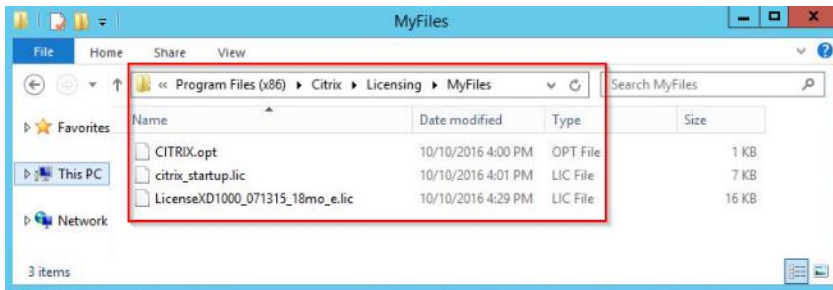


Step 11. Click Finish to complete the installation.



Procedure 2. Install Citrix Licenses

Step 1. Copy the license files to the default location (C:\Program Files (x86)\Citrix\Licensing\ MyFiles) on the license server.

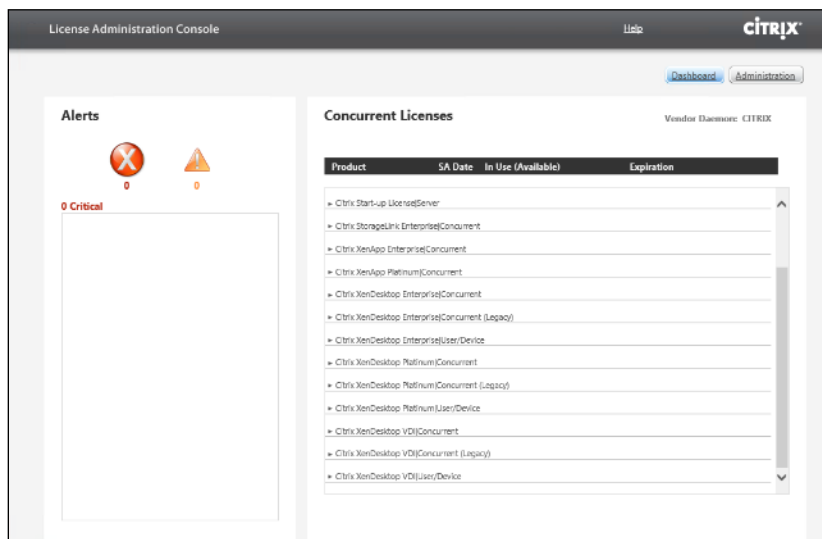


Step 2. Restart the server or Citrix licensing services so that the licenses are activated.

Step 3. Run the application Citrix License Administration Console.



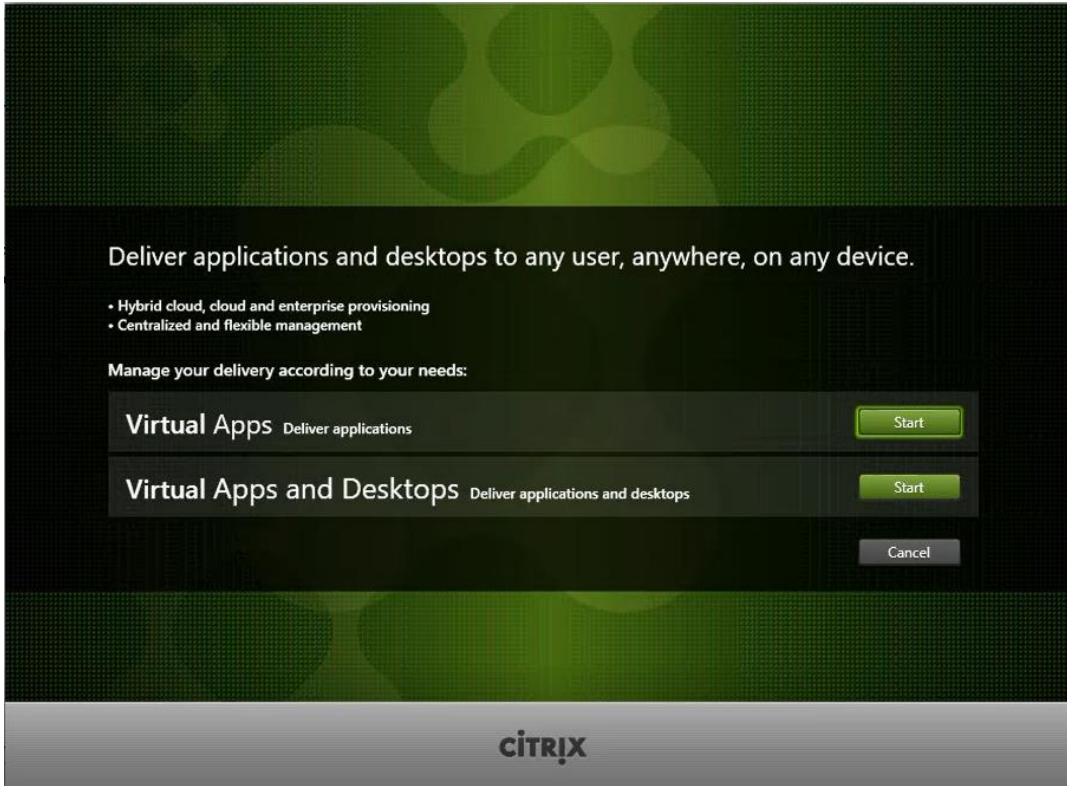
Step 4. Confirm that the license files have been read and enabled correctly.



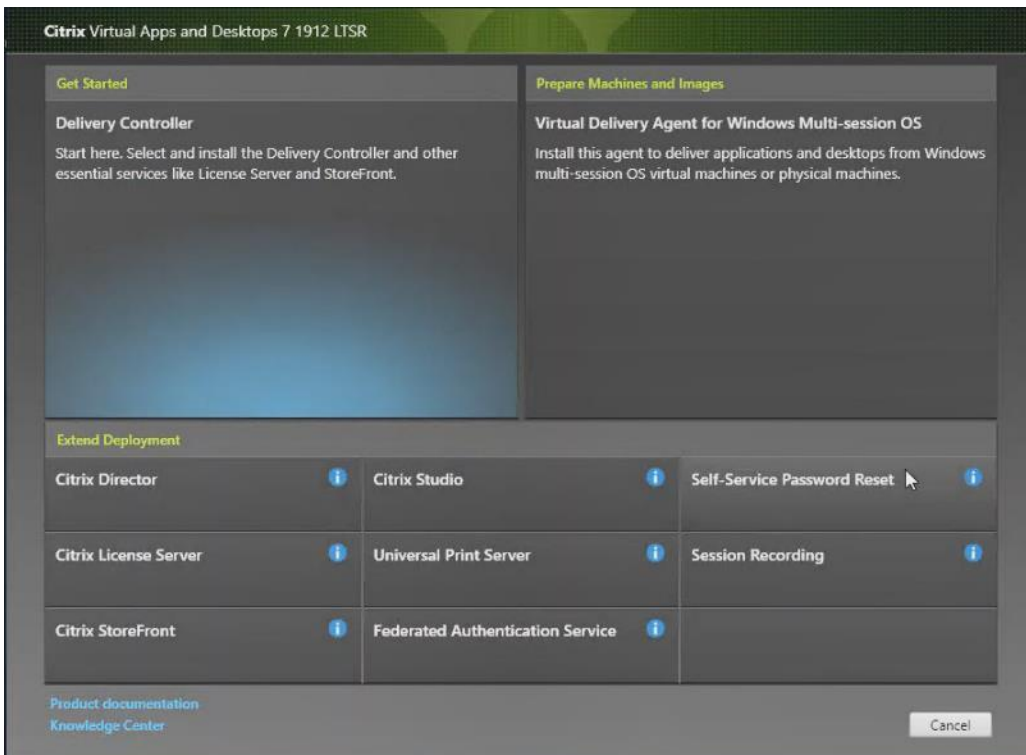
Procedure 3. Install Citrix Desktop Broker/Studio

Step 1. Connect to the first Citrix VDI server and launch the installer from the Citrix Desktop 1912 LTSR ISO.

Step 2. Click Start.

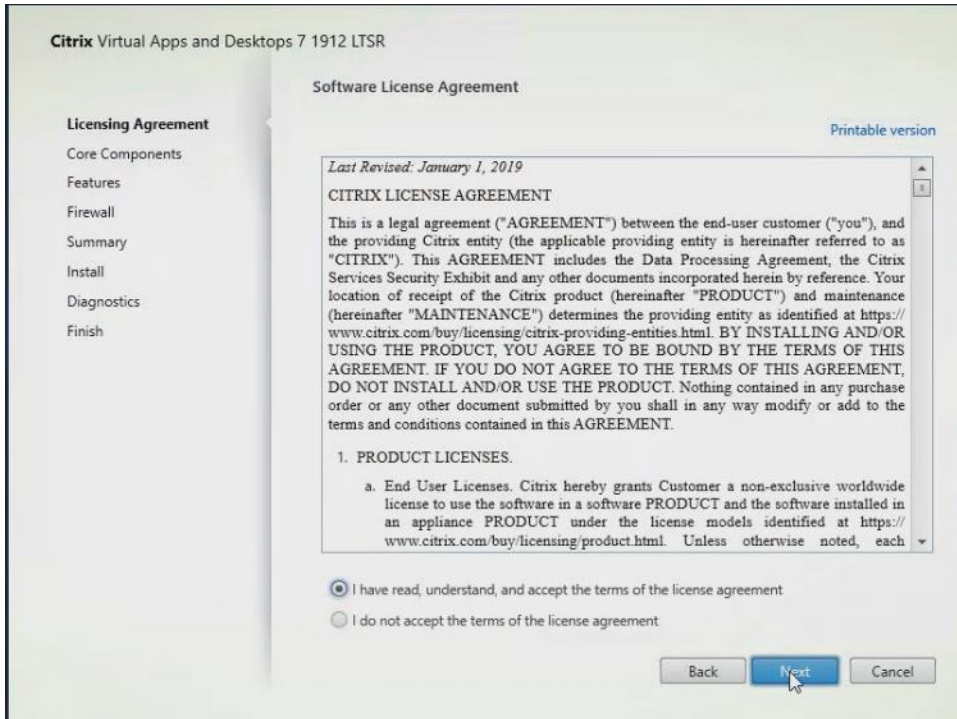


Step 3. The installation wizard presents a menu with three subsections. Click “Get Started - Delivery Controller.”



Step 4. Read the Citrix License Agreement and if acceptable, indicate your acceptance of the license by selecting the “I have read, understand, and accept the terms of the license agreement” radio button.

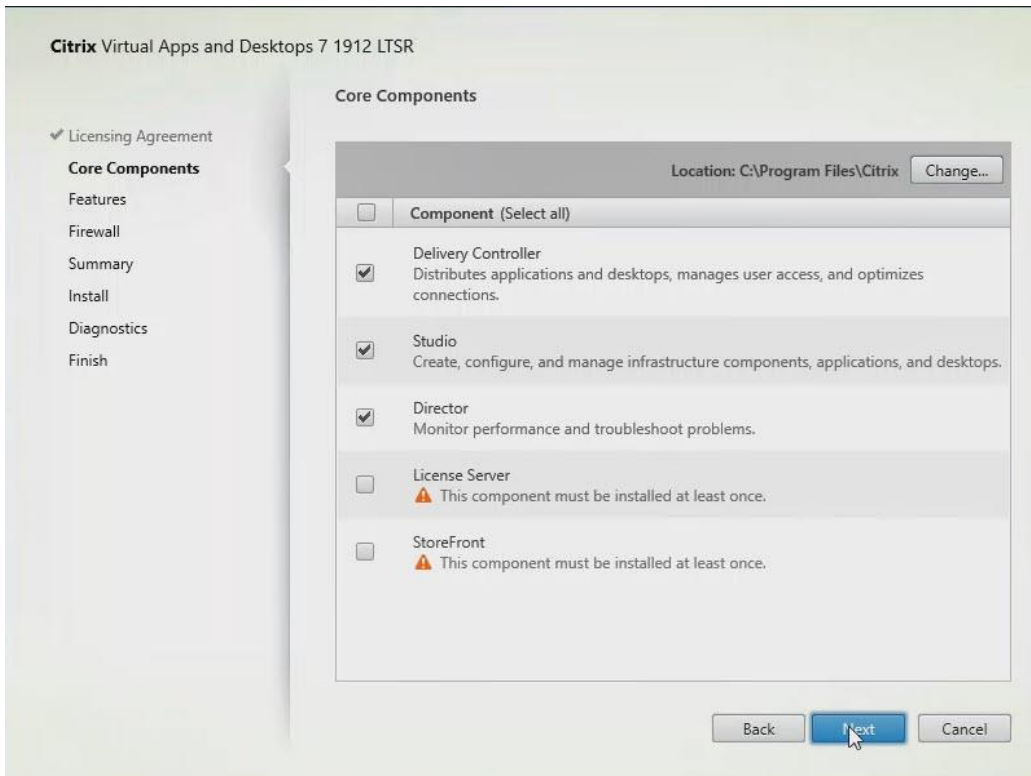
Step 5. Click Next.



Step 6. Select the components to be installed on the first Delivery Controller Server:

- Delivery Controller
- Studio
- Director

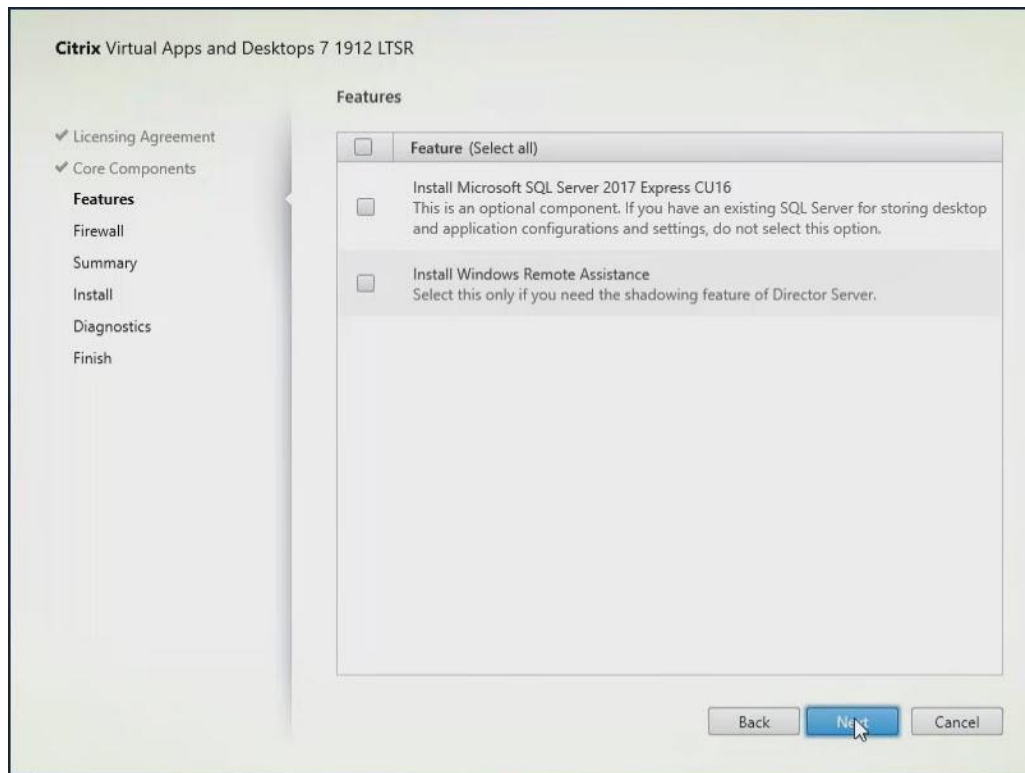
Step 7. Click Next.



Note: Dedicated StoreFront and License servers should be implemented for large-scale deployments.

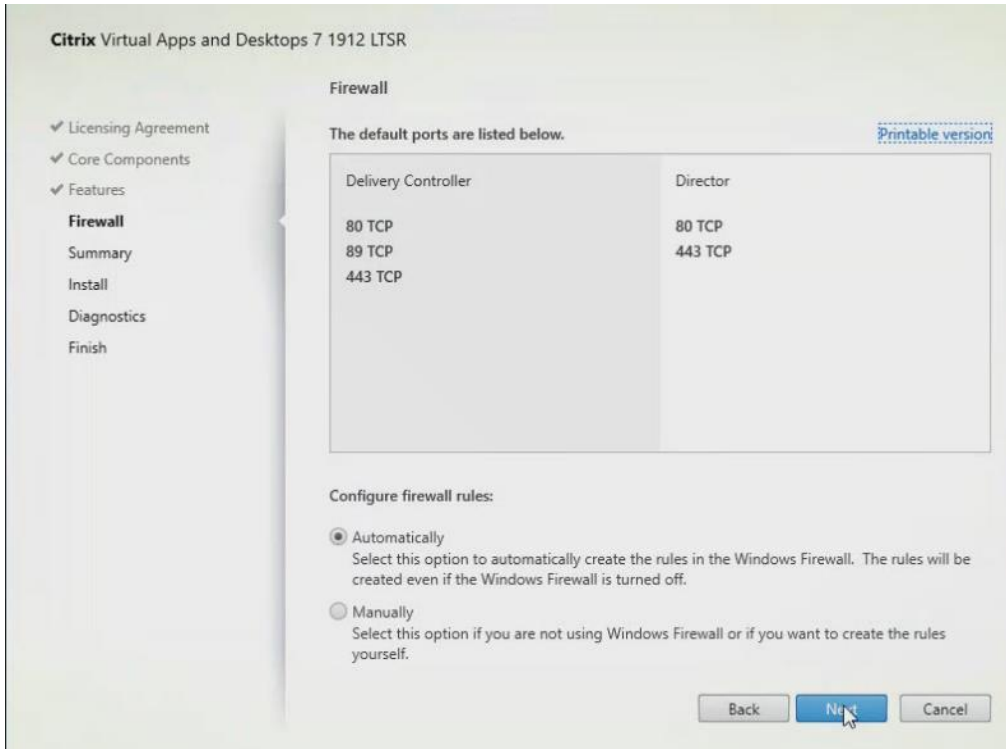
Step 8. Since a SQL Server will be used to Store the Database, leave “Install Microsoft SQL Server 2012 SP1 Express” unchecked.

Step 9. Click Next.

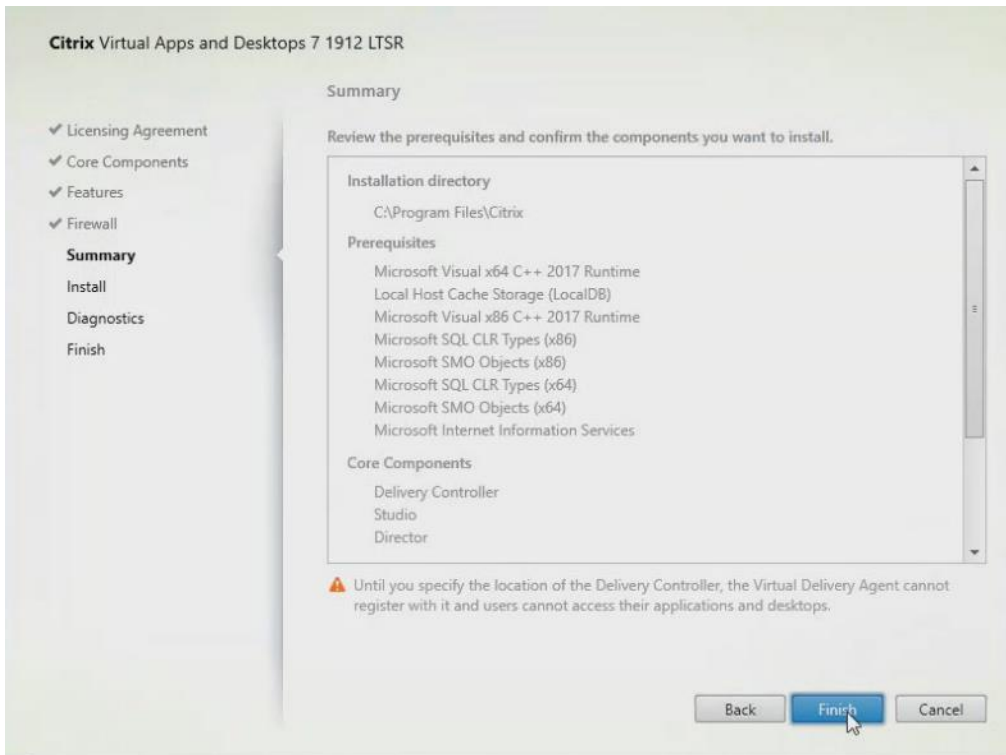


Step 10. Select the default ports and automatically configured firewall rules.

Step 11. Click Next.

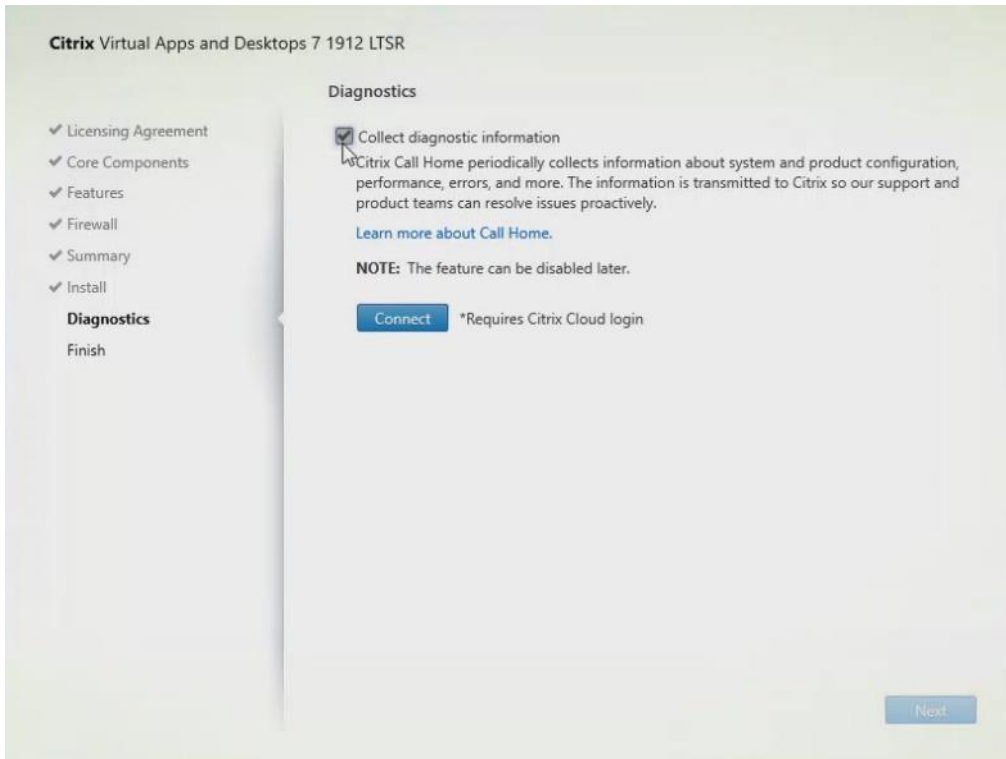


Step 12. Click Install.



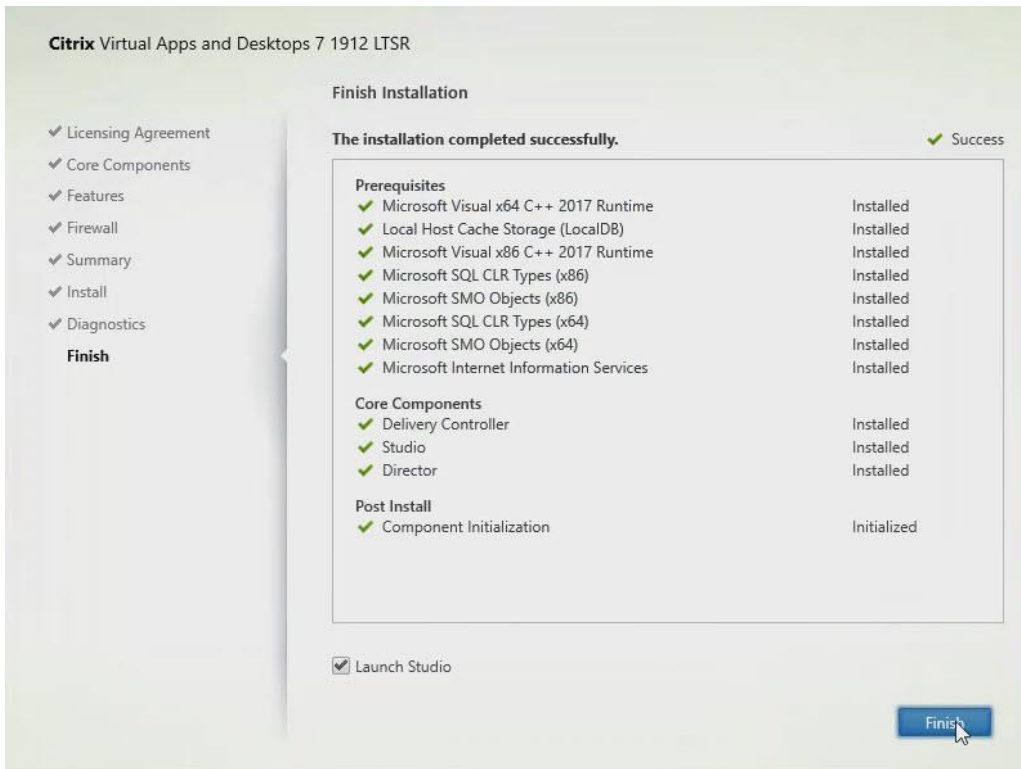
Step 13. (Optional) Click the Call Home participation.

Step 14. Click Next.



Step 15. Click Finish to complete the installation.

Step 16. (Optional) Check Launch Studio to launch Citrix Studio Console.

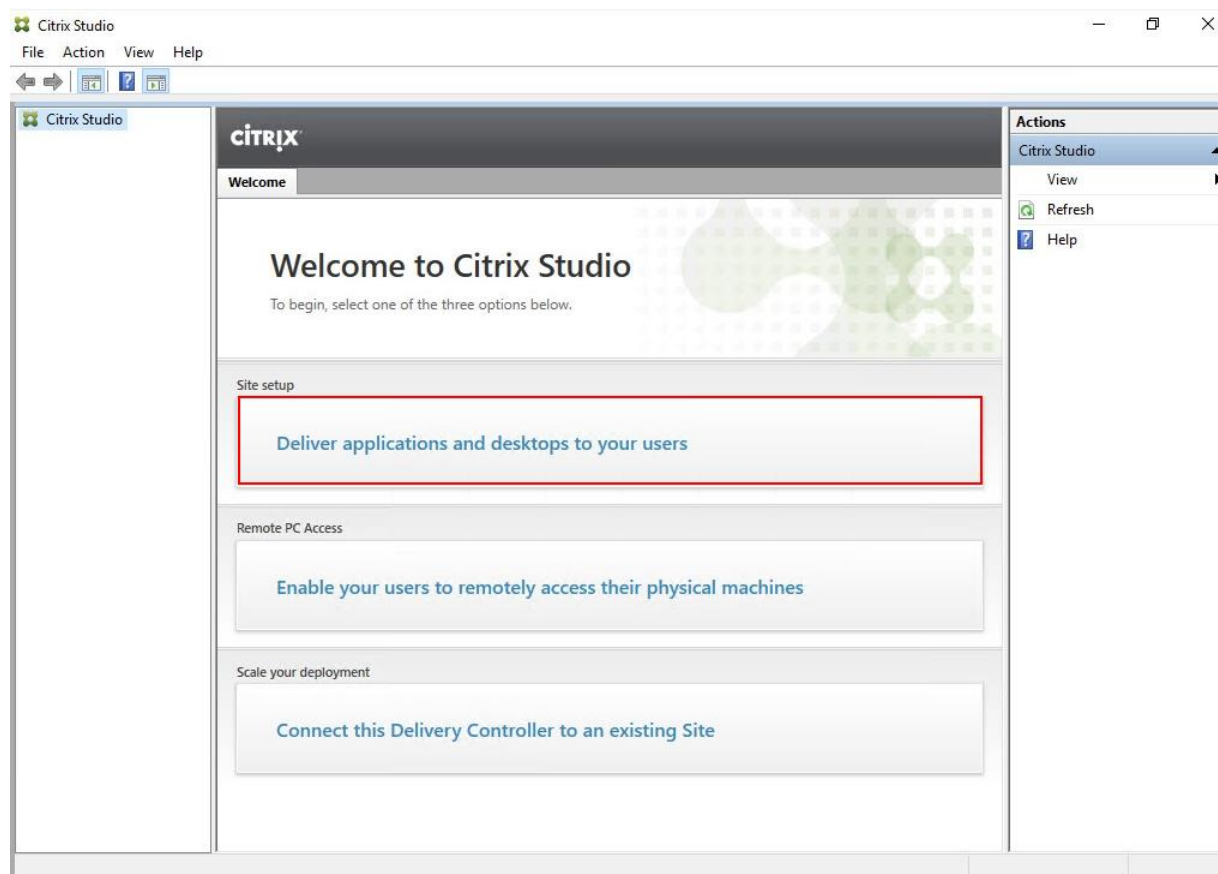


Procedure 4. Configure the Citrix VDI Site

Citrix Studio is a management console that allows you to create and manage infrastructure and resources to deliver desktops and applications. Replacing Desktop Studio from earlier releases, it provides wizards to set up your environment, create workloads to host applications and desktops, and assign applications and desktops to users.

Citrix Studio launches automatically after the Citrix VDI Delivery Controller installation, or if necessary, it can be launched manually. Citrix Studio is used to create a Site, which is the core Citrix VDI environment consisting of the Delivery Controller and the Database.

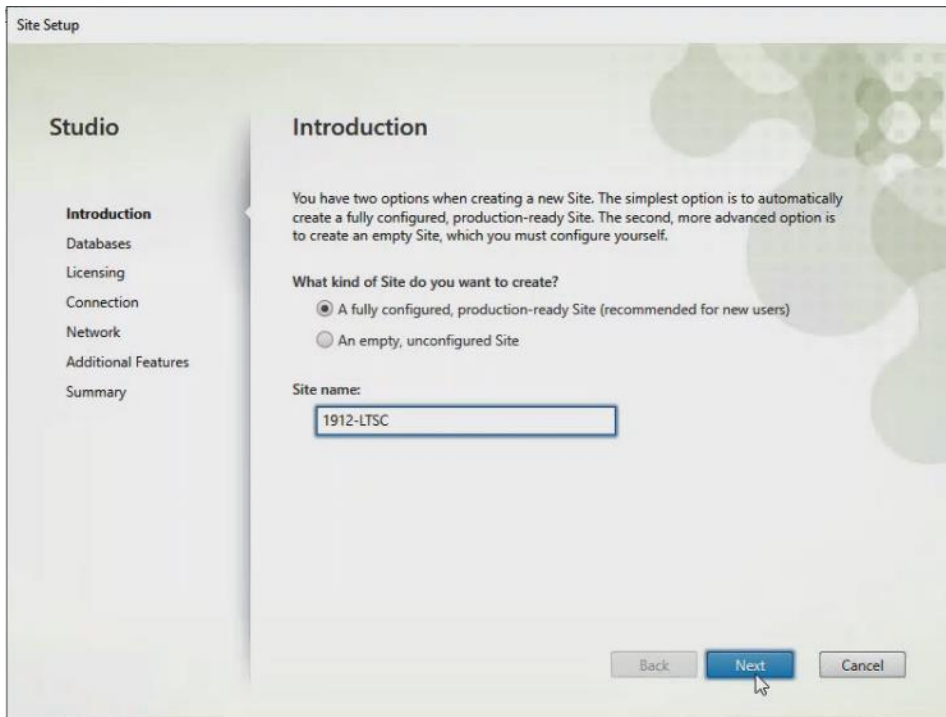
Step 1. From Citrix Studio, click Deliver applications and desktops to your users.



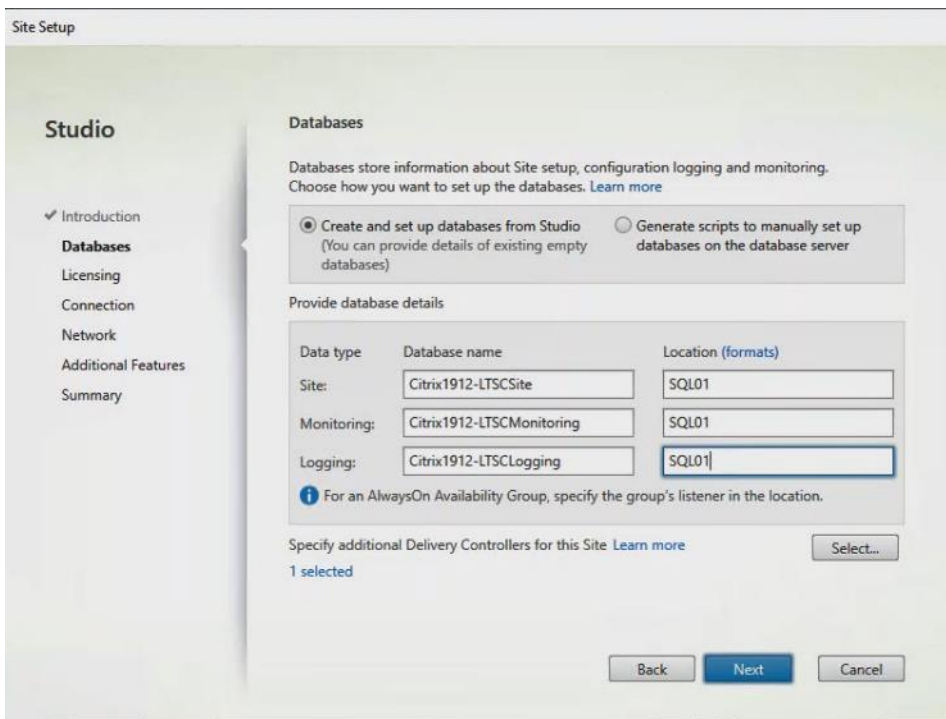
Step 2. Select the “A fully configured, production-ready Site” radio button.

Step 3. Enter a site name.

Step 4. Click Next.



Step 5. Provide the Database Server Locations for each data type and click Next.



Step 6. For an AlwaysOn Availability Group, use the group's listener DNS name.

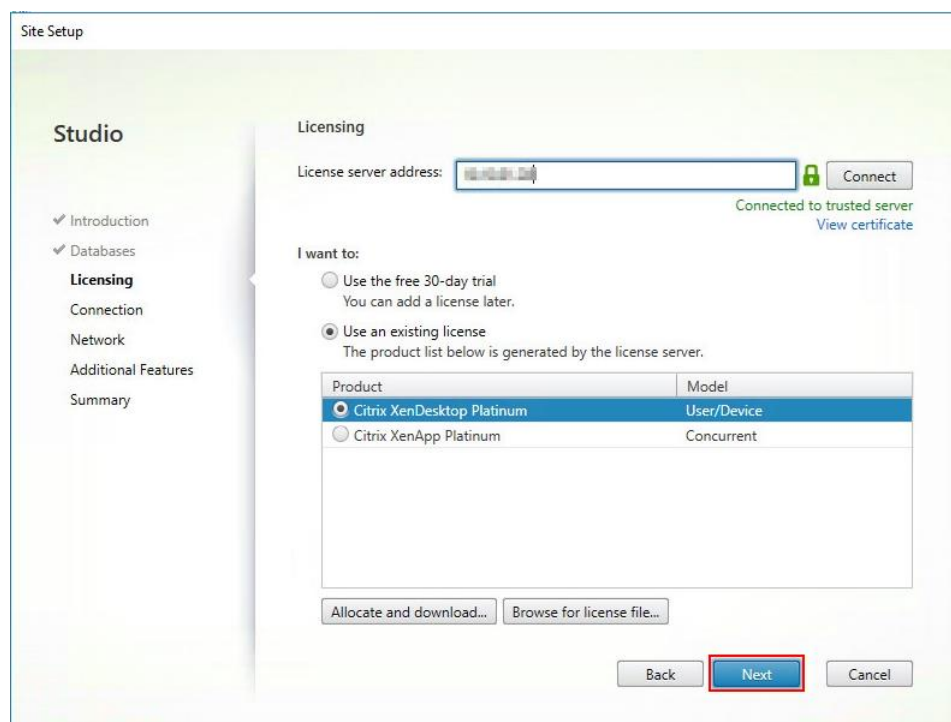
Step 7. Provide the FQDN of the license server.

Step 8. Click Connect to validate and retrieve any licenses from the server.

Note: If no licenses are available, you can use the 30-day free trial or activate a license file.

Step 9. Select the appropriate product edition using the license radio button.

Step 10. Click Next.



Step 11. Select the Connection type of 'Microsoft System Center Virtual Machine Manager.'

Step 12. Enter the Connection Address to the SCVMM Server.

Step 13. Enter the username (in username@domain format) for the vCenter account.

Step 14. Provide the password for the Domain Admin account.

Step 15. Provide a connection name.

Step 16. Select the Studio tools radio button.

Add Connection and Resources

Studio

- Connection
- Storage Management
- Storage Selection
- Network
- Summary

Connection

Use an existing Connection

8x16

Create a new Connection

Connection type: VMware vSphere®

Connection address: Example: https://vmware.example.com/sdk

[Learn about user permissions](#)

User name: Example: domain\username

Password:

Connection name: Example: MyConnection

Create virtual machines using:

Studio tools (Machine Creation Services)
Select this option when using AppDisks, even if you are using Citrix Provisioning.

Other tools

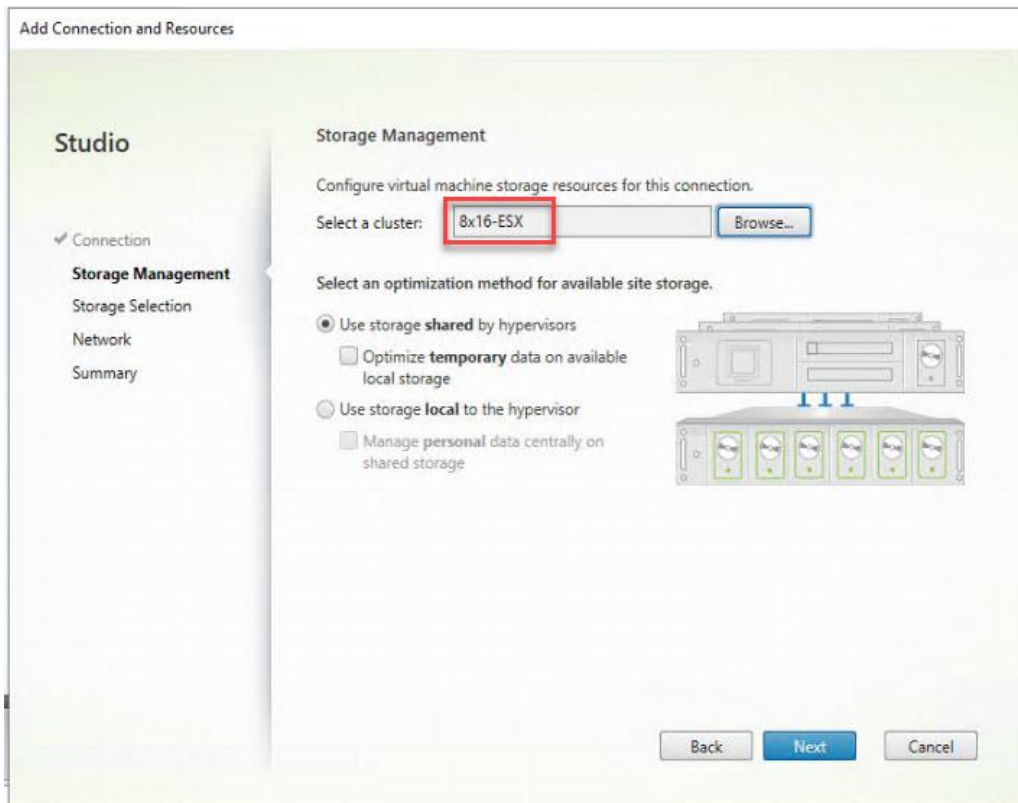
Back Next Cancel

Step 17. Click Next.

Step 18. Select HyperFlex Cluster that will be used by this connection.

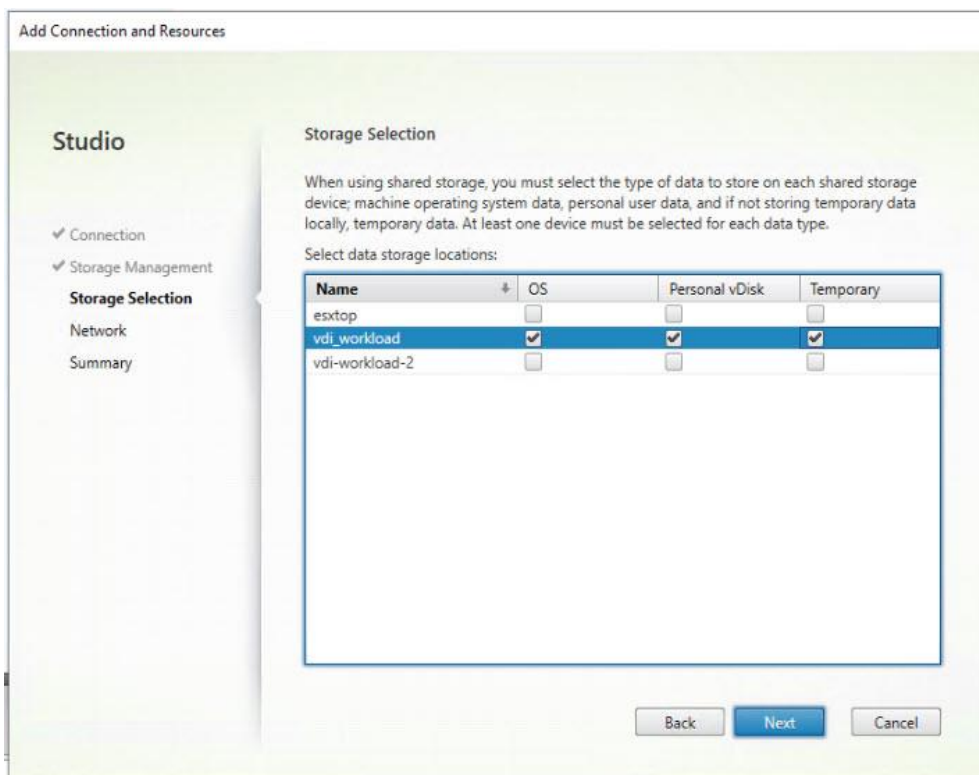
Step 19. Check Studio Tools radio button required to support desktop provisioning task by this connection.

Step 20. Click Next.



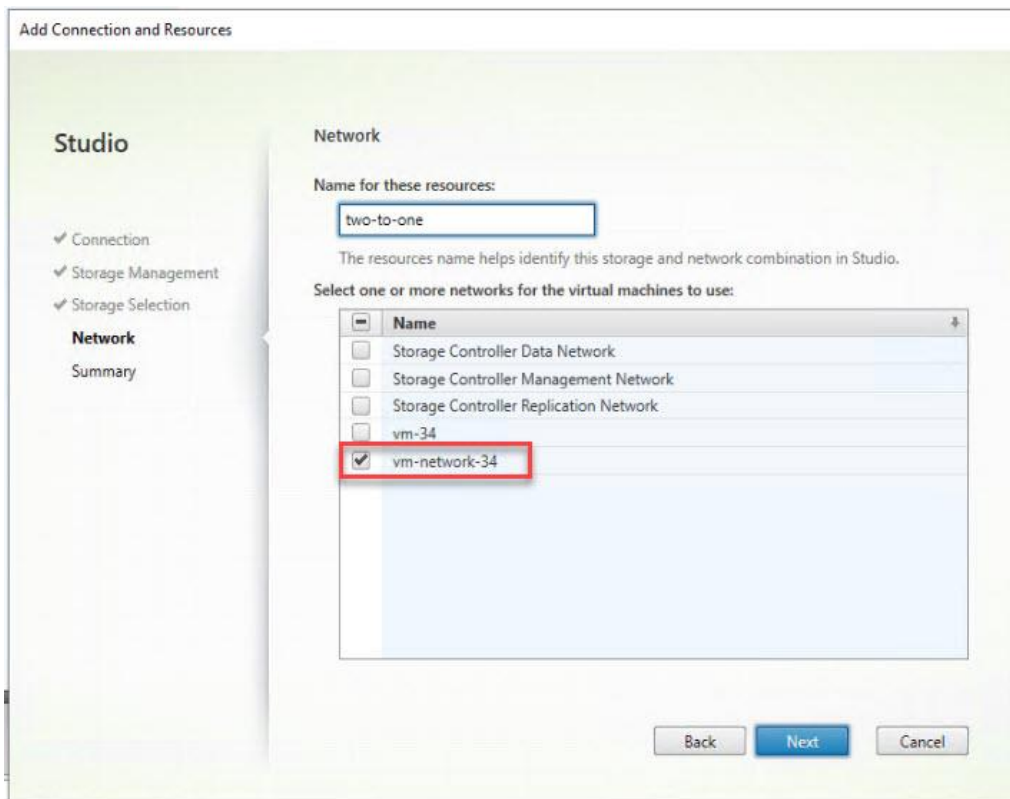
Step 21. Select the Storage to be used by this connection.

Step 22. Click Next.



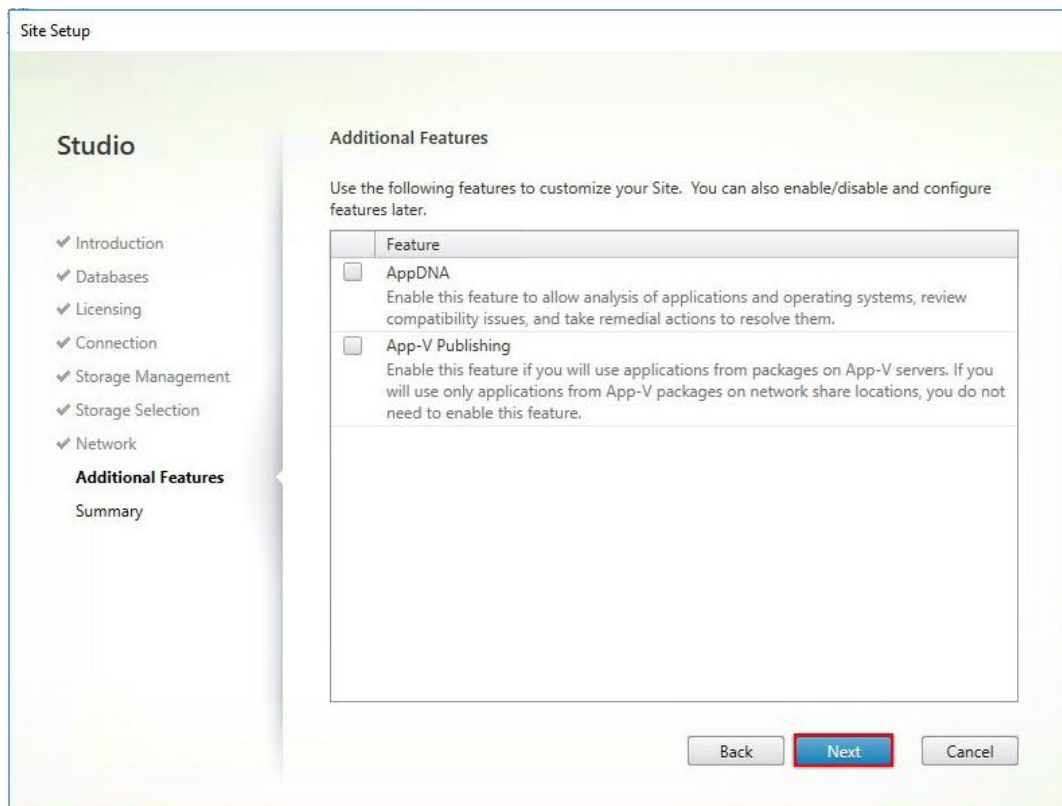
Step 23. Make Network selection to be used by this connection.

Step 24. Click Next.



Step 25. Select Additional features.

Step 26. Click Next.

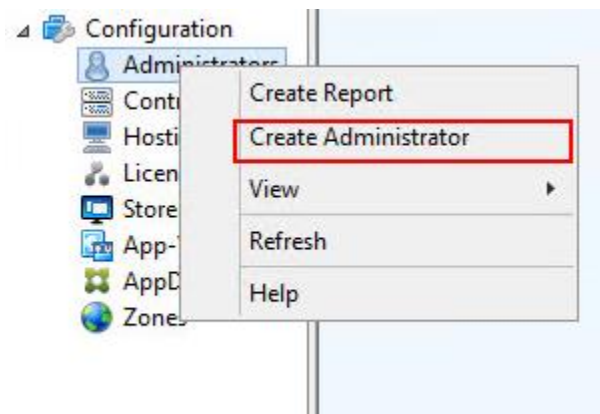


Step 27. Review Site configuration Summary and click Finish.

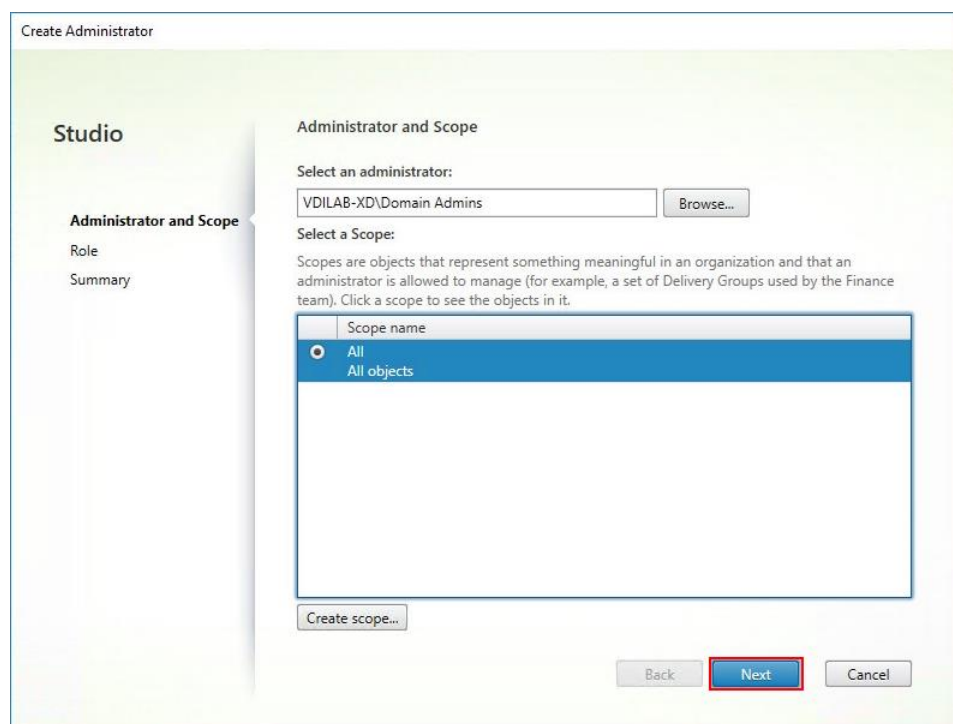
Procedure 5. Configure the Citrix VDI Site Administrators

Step 1. Connect to the Citrix VDI server and open Citrix Studio Management console.

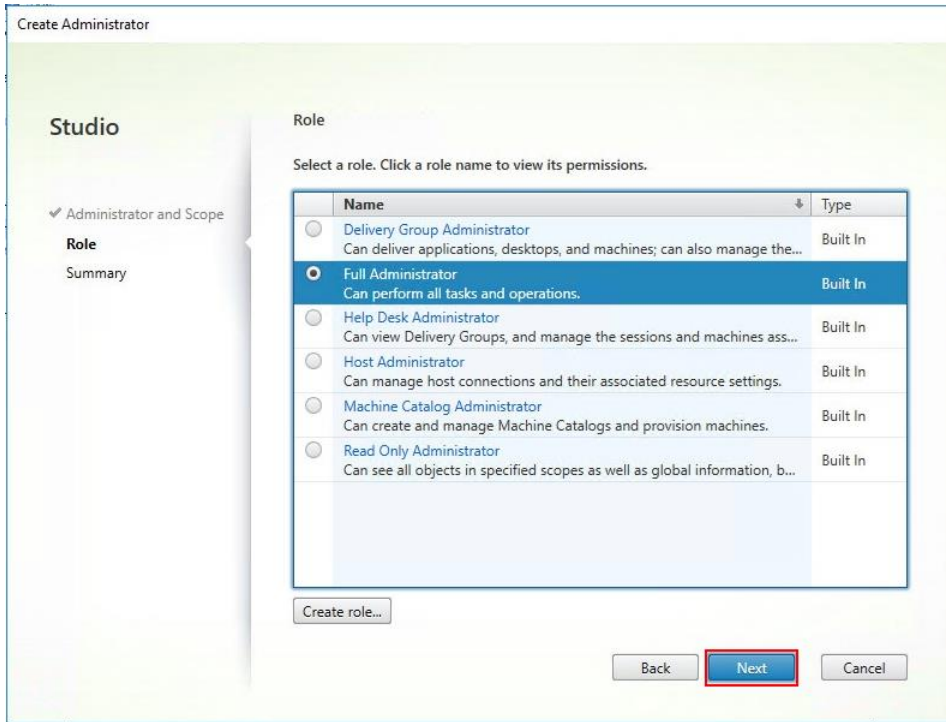
Step 2. From the Configuration menu, right-click Administrator and select Create Administrator from the drop-down list.



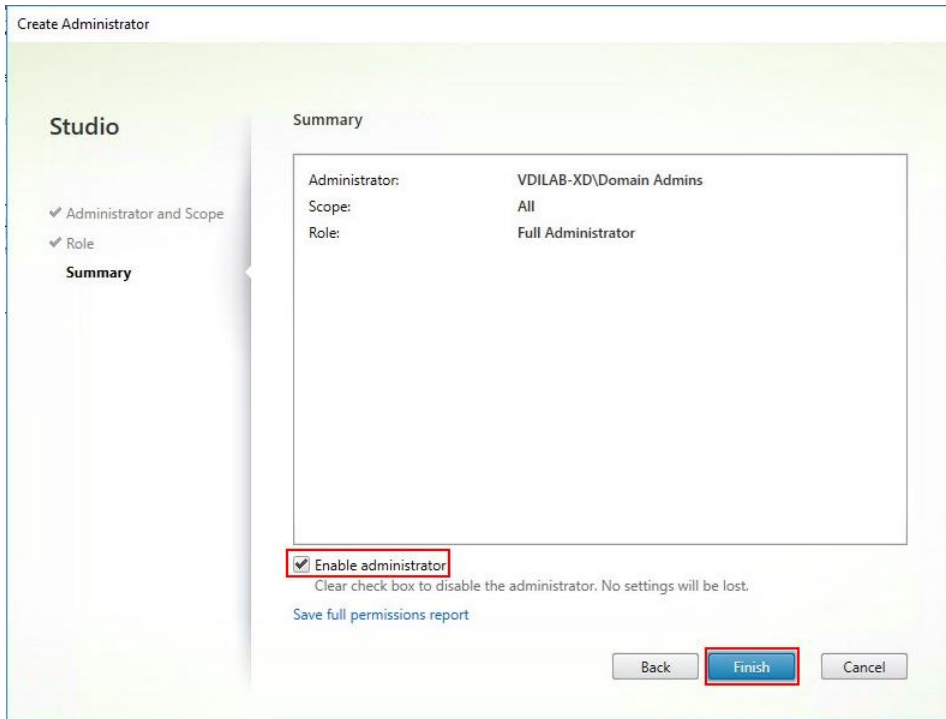
Step 3. Select the appropriate Administrator and Scope and click Next.



Step 4. Select an appropriate Role.



Step 5. Review the Summary, check Enable administrator, and click Finish.



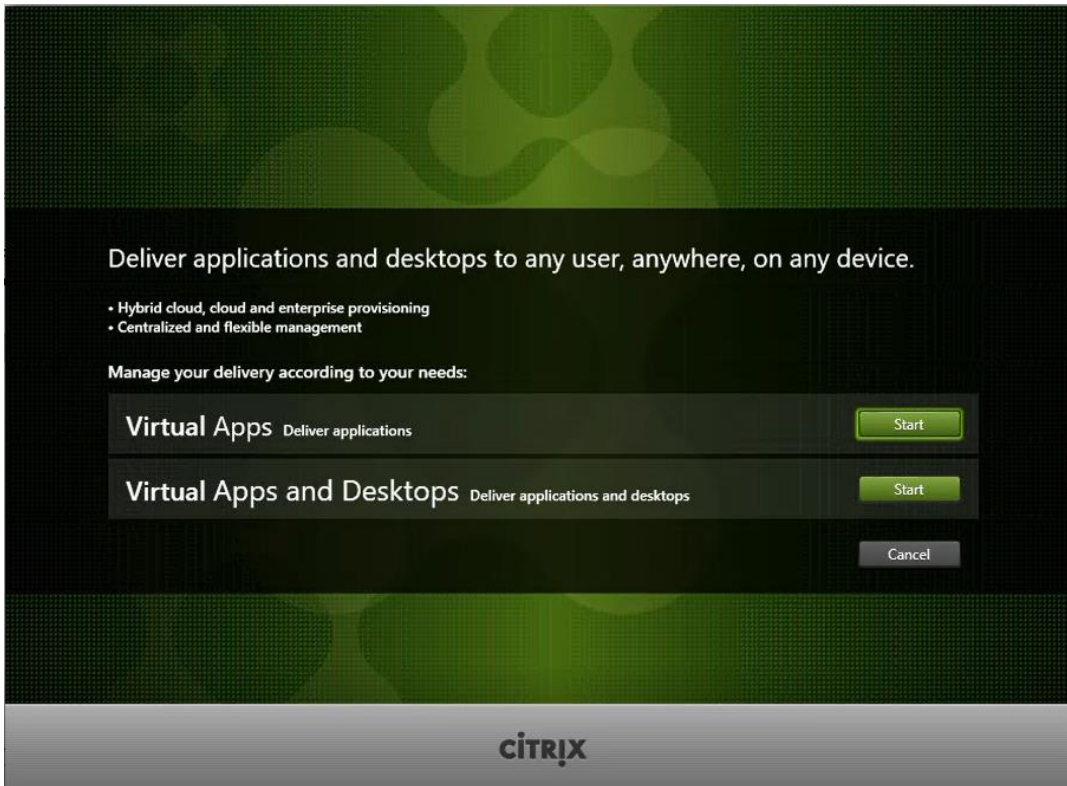
Procedure 6. Configure Additional Desktop Controller

After the first controller is completely configured and the site is operational, you can add additional controllers.

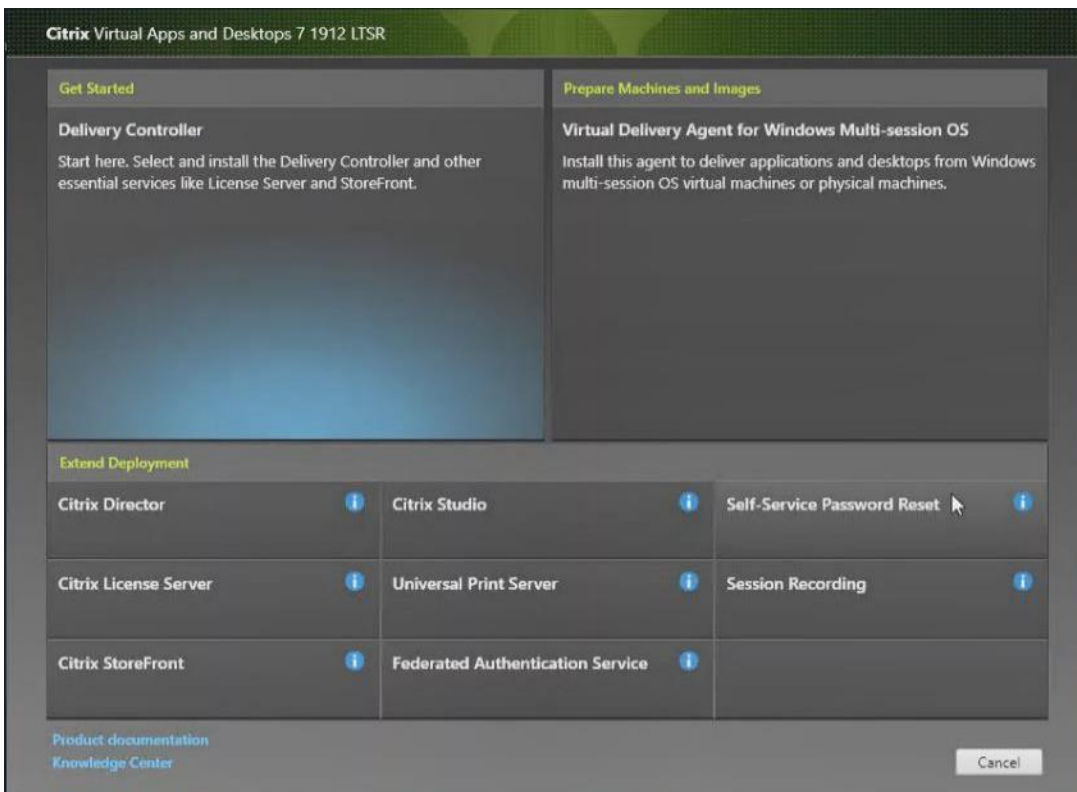
Note: In this CVD, we created two Delivery Controllers.

Step 1. To begin the installation of the second Delivery Controller, connect to the second Citrix VDI server and launch the installer from the Citrix Virtual Apps and Desktops ISO.

Step 2. Click Start.



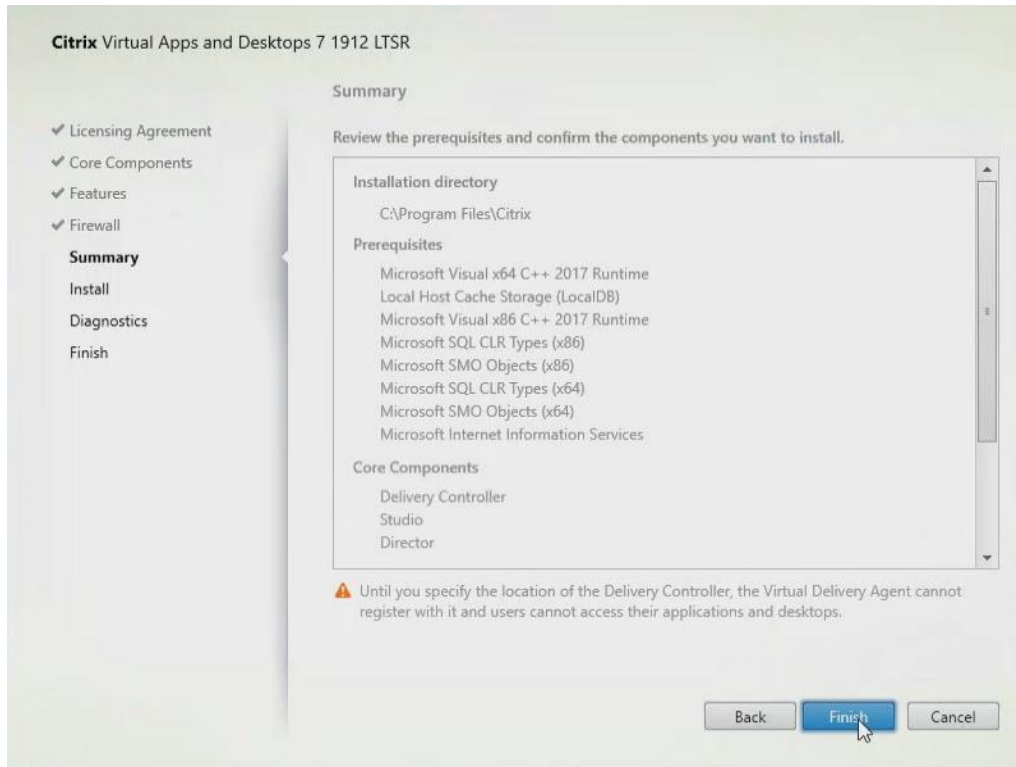
Step 3. Click Delivery Controller.



Step 4. Repeat the steps used to install the first Delivery Controller, including the step of importing an SSL certificate for HTTPS between the controller and HyperV.

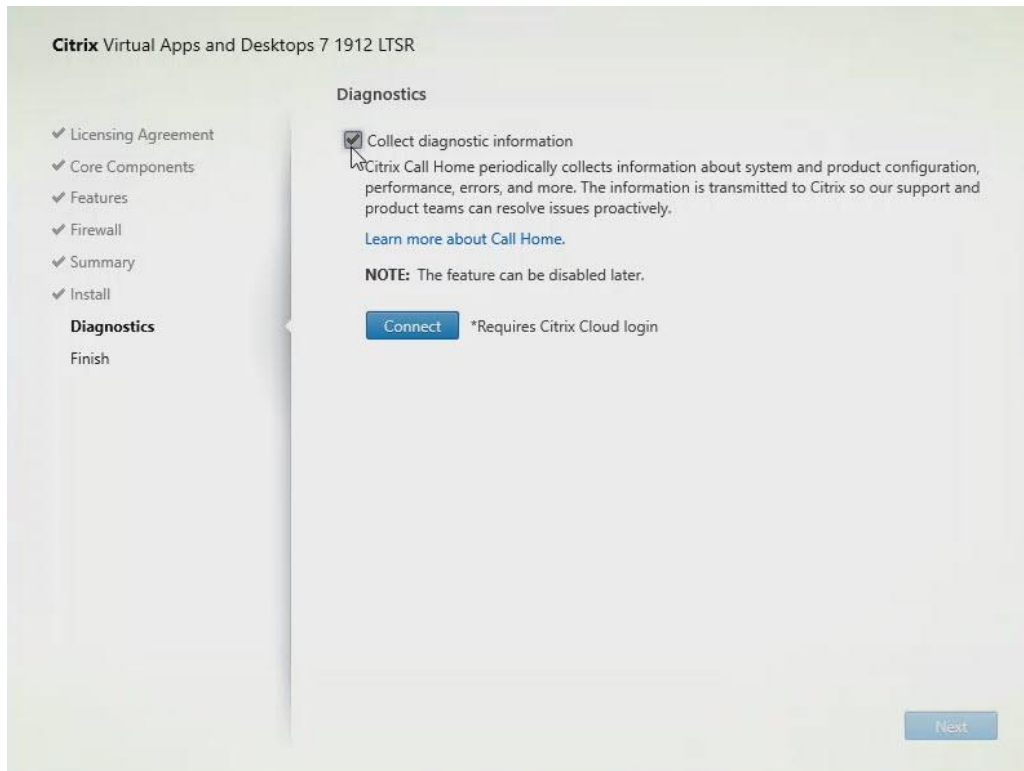
Step 5. Review the Summary configuration.

Step 6. Click Install.



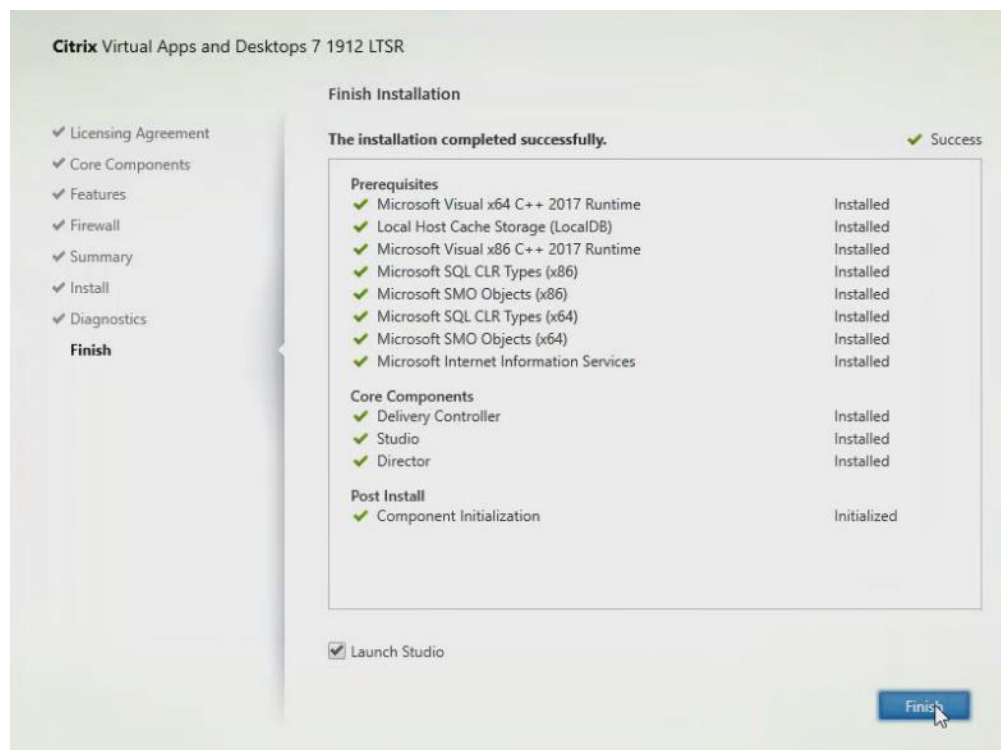
Step 7. (Optional) Click the “Collect diagnostic information.”

Step 8. Click Next.



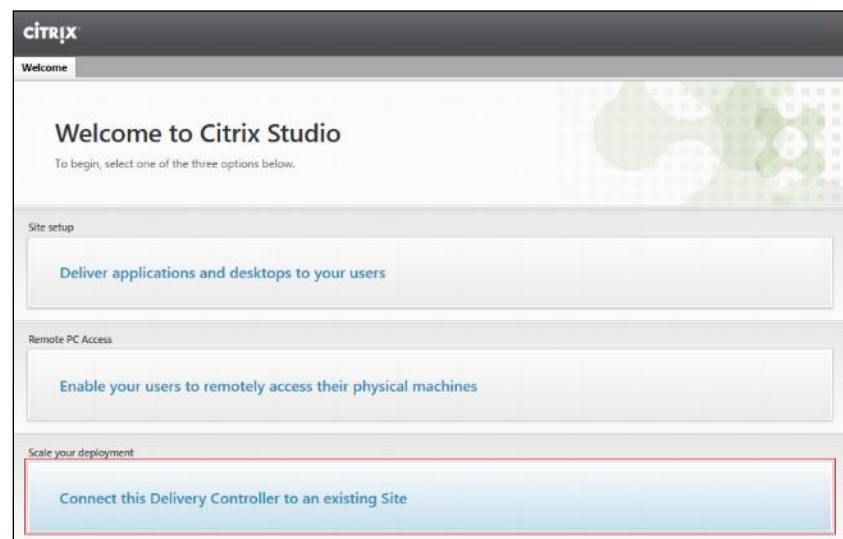
Step 9. Verify the components installed successfully.

Step 10. Click Finish.



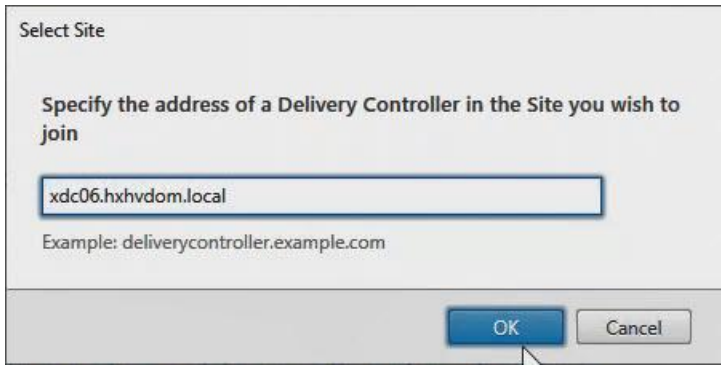
Procedure 7. Add the Second Delivery Controller to the Citrix Desktop Site

Step 1. In Desktop Studio click the “Connect this Delivery Controller to an existing Site” button.



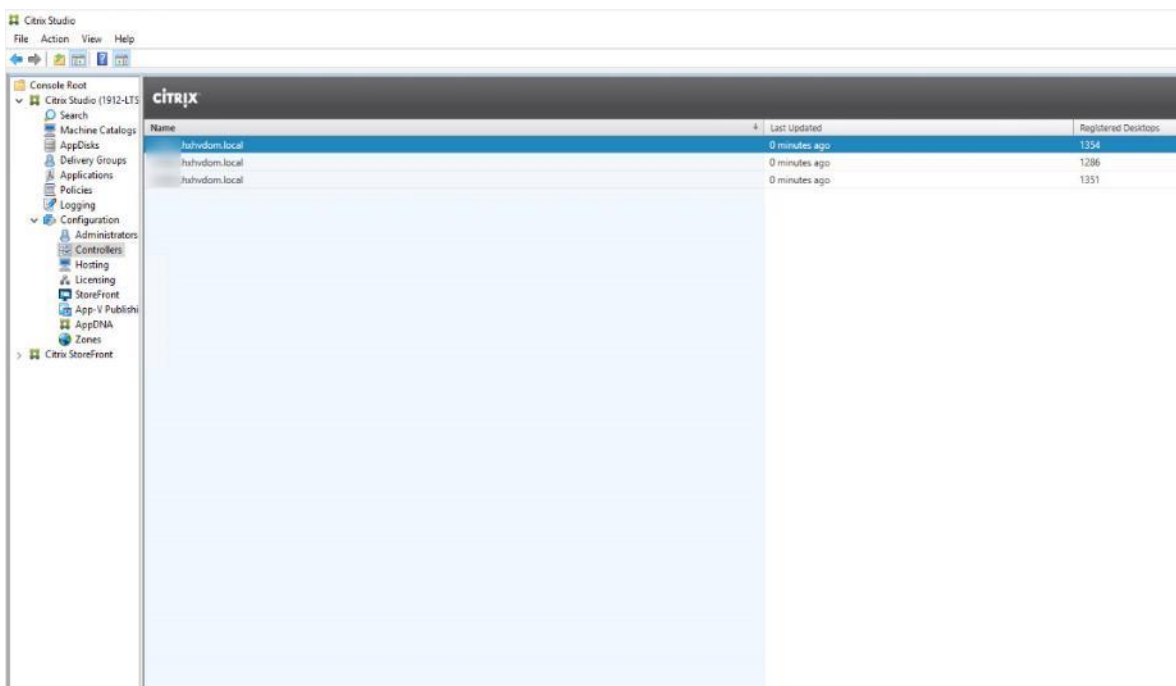
Step 2. Enter the FQDN of the first delivery controller.

Step 3. Click OK.



Step 4. Click Yes to allow the database to be updated with this controller’s information automatically.

Step 5. When complete, test the site configuration and verify the Delivery Controller has been added to the list of Controllers.



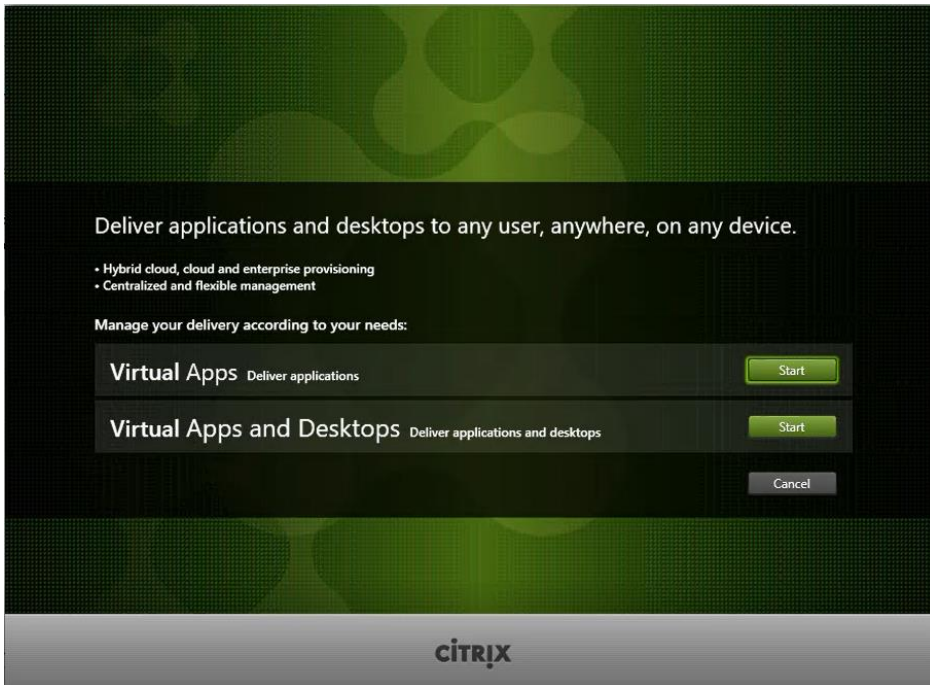
Procedure 8. Install and Configure StoreFront

Citrix StoreFront stores aggregate desktops and applications from Citrix VDI sites, making resources readily available to users.

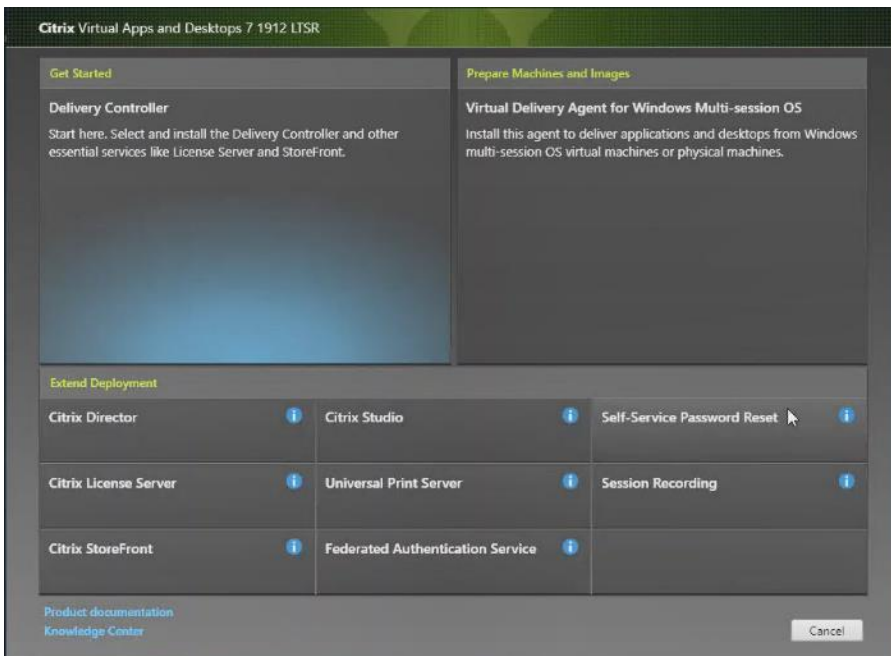
Note: In this CVD, we created two StoreFront servers on dedicated virtual machines.

Step 1. Connect to the first StoreFront server and launch the installer from the Citrix Desktop 1912 LTSR ISO.

Step 2. Click Start.

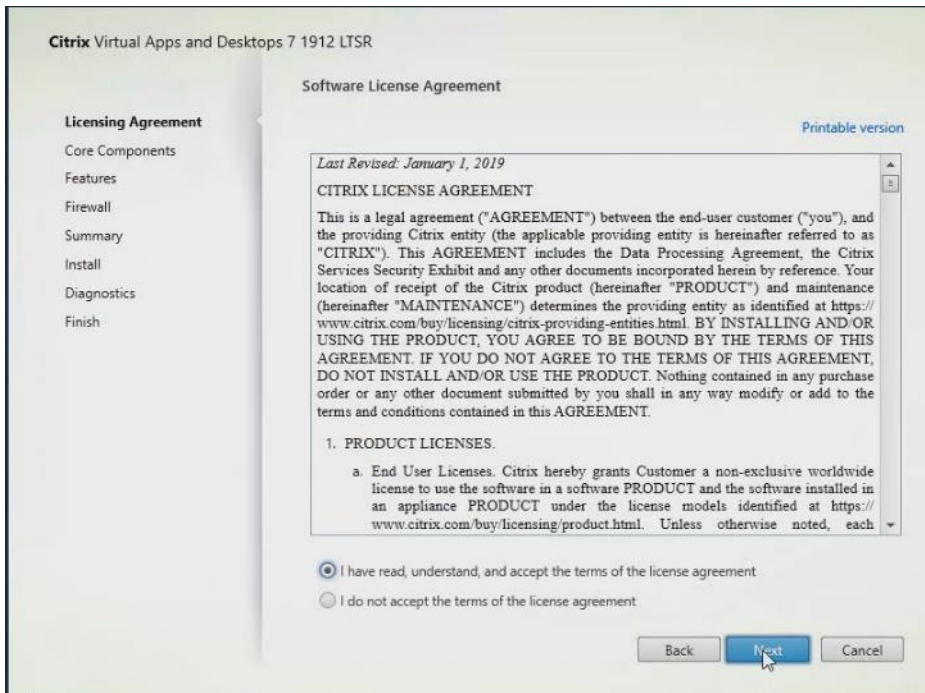


Step 3. Click Extend Deployment Citrix StoreFront.

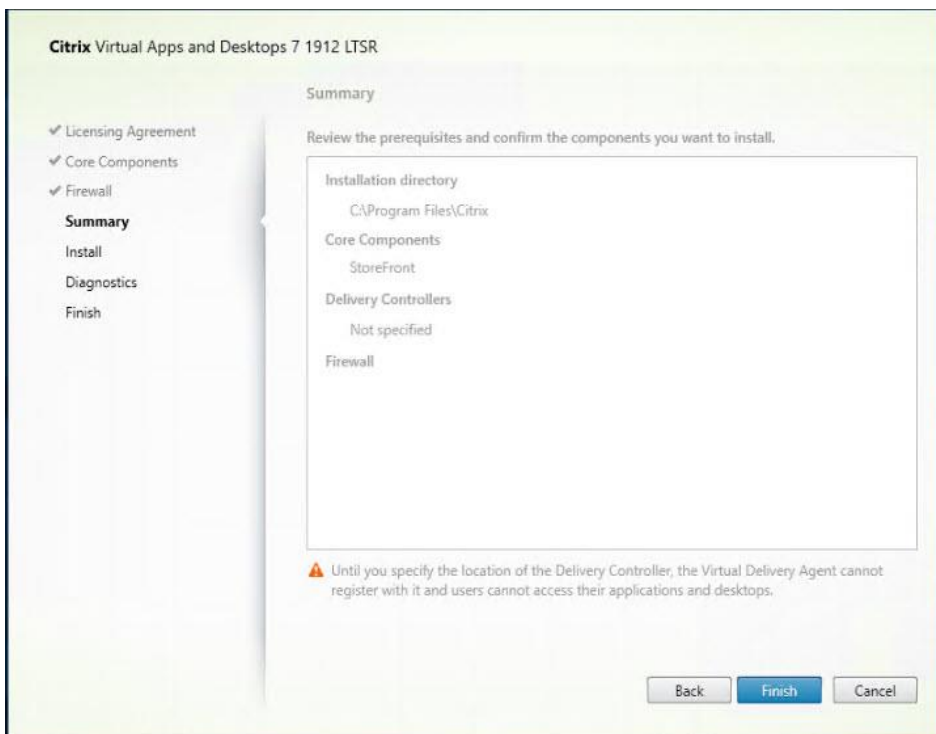


Step 4. If acceptable, indicate your acceptance of the license by selecting the “I have read, understand, and accept the terms of the license agreement” radio button.

Step 5. Click Next.



Step 6. Select Storefront and click Next.



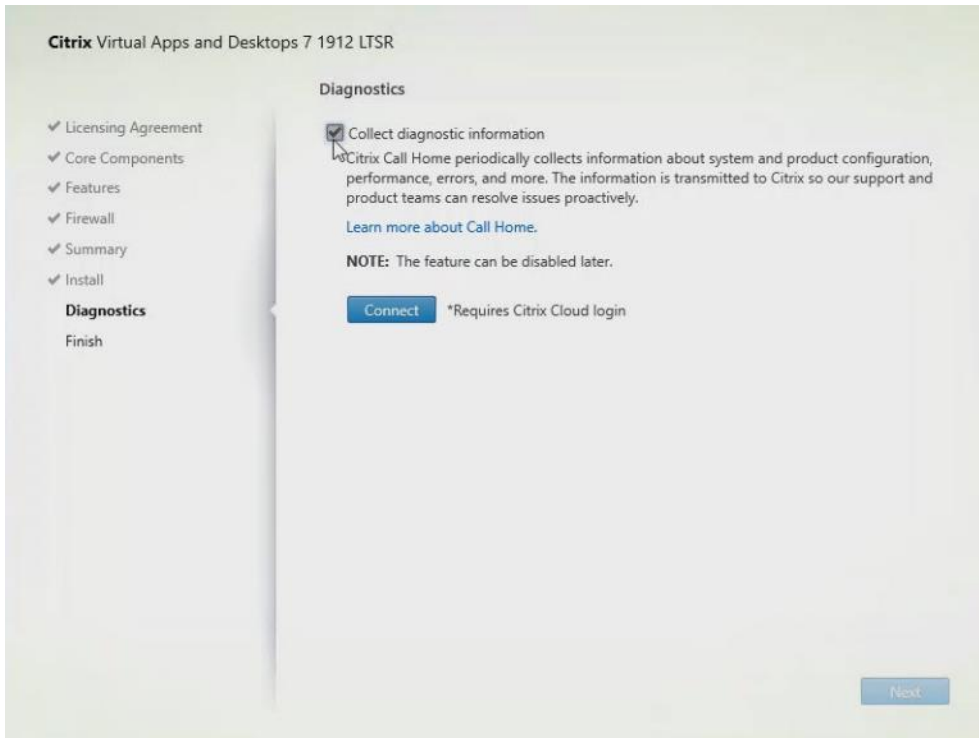
Step 7. Select the default ports and automatically configured firewall rules.

Step 8. Click Next.

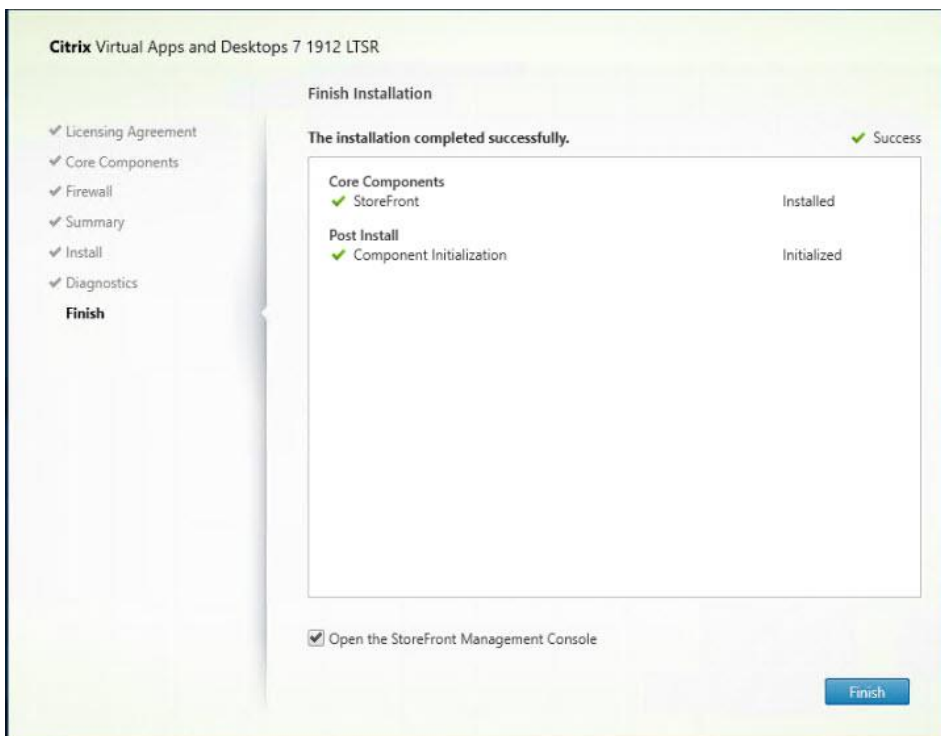
Step 9. Click Install.

Step 10. (Optional) Click "Collect diagnostic information."

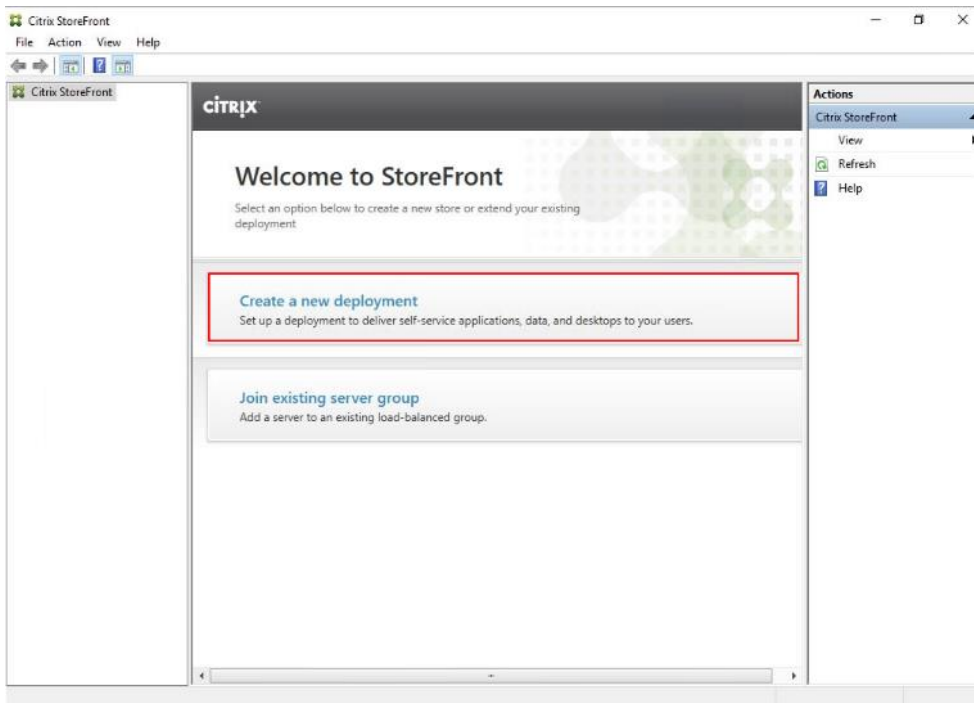
Step 11. Click Next.



Step 12. Click Finish.

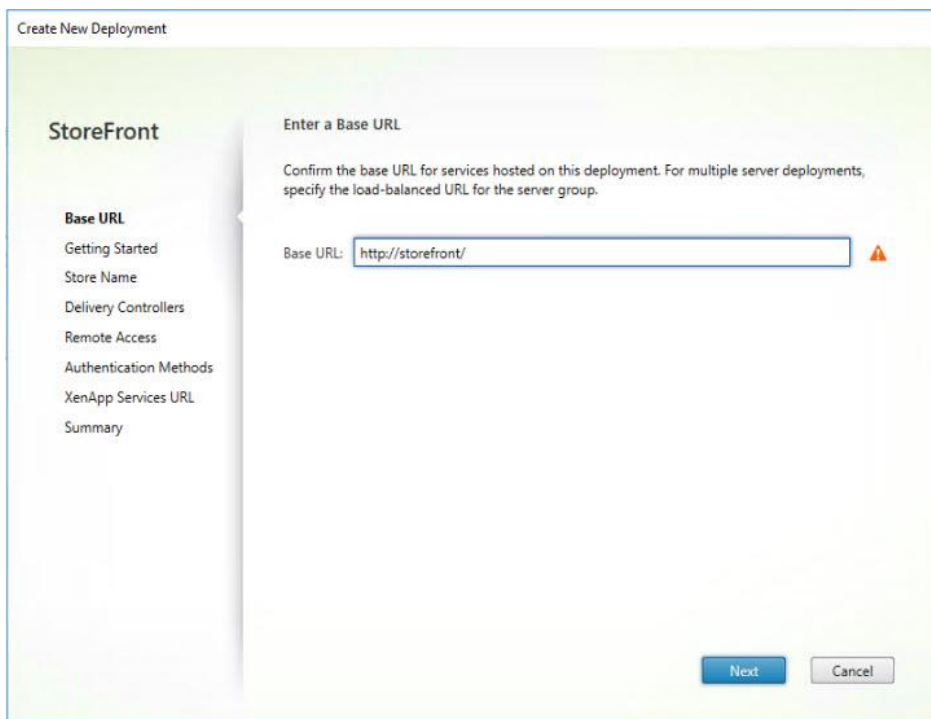


Step 13. Click Create a new deployment.

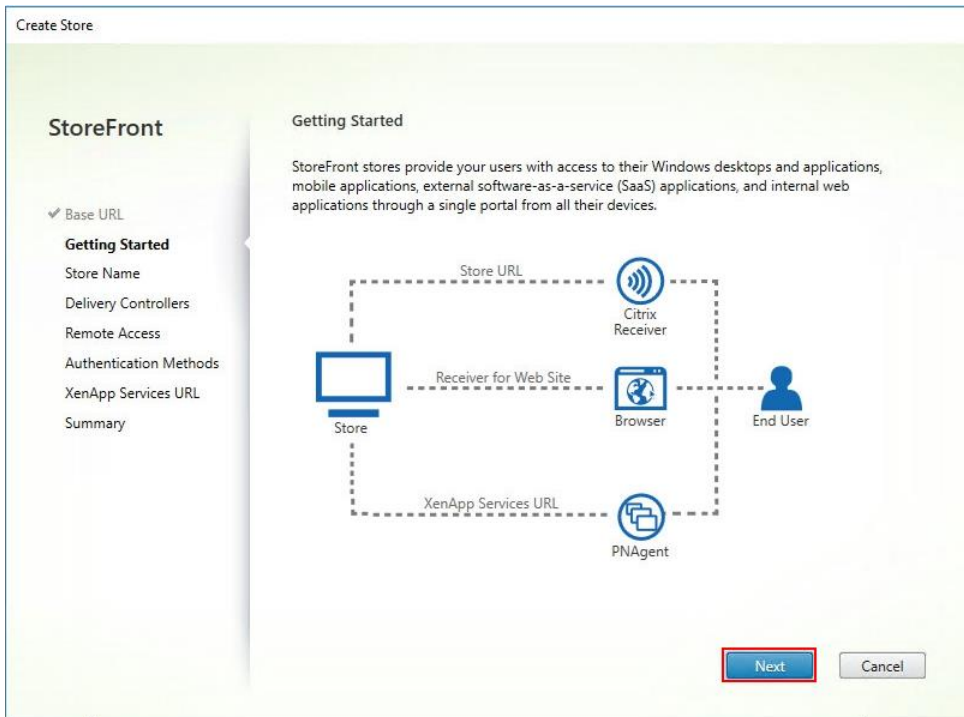


Step 14. Specify the URL of the StoreFront server and click Next.

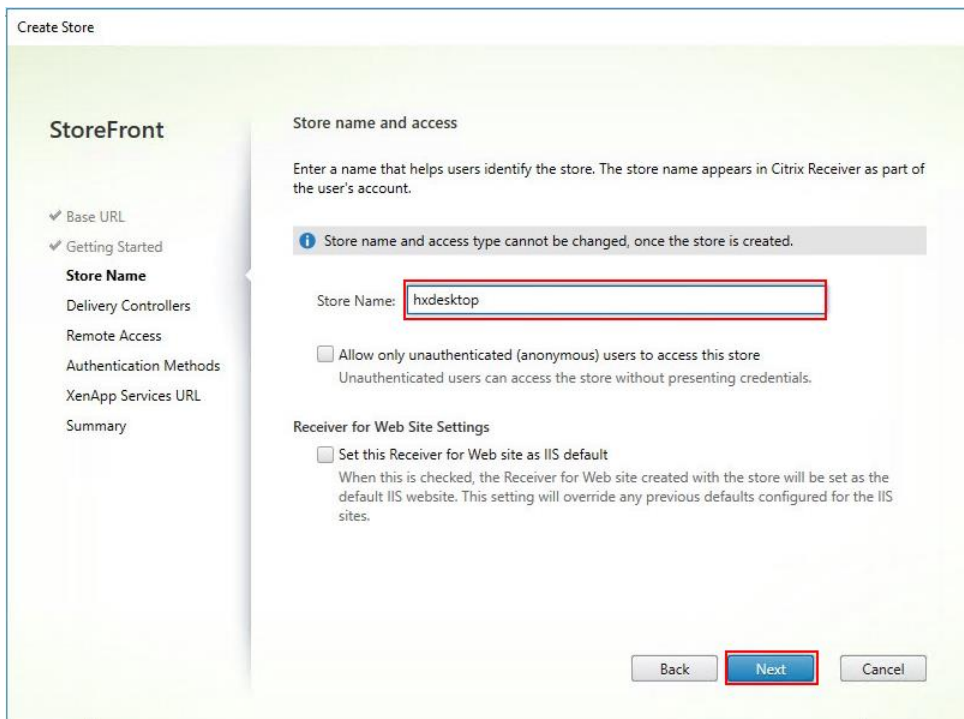
Note: For a multiple server deployment use the load balancing environment in the Base URL box.



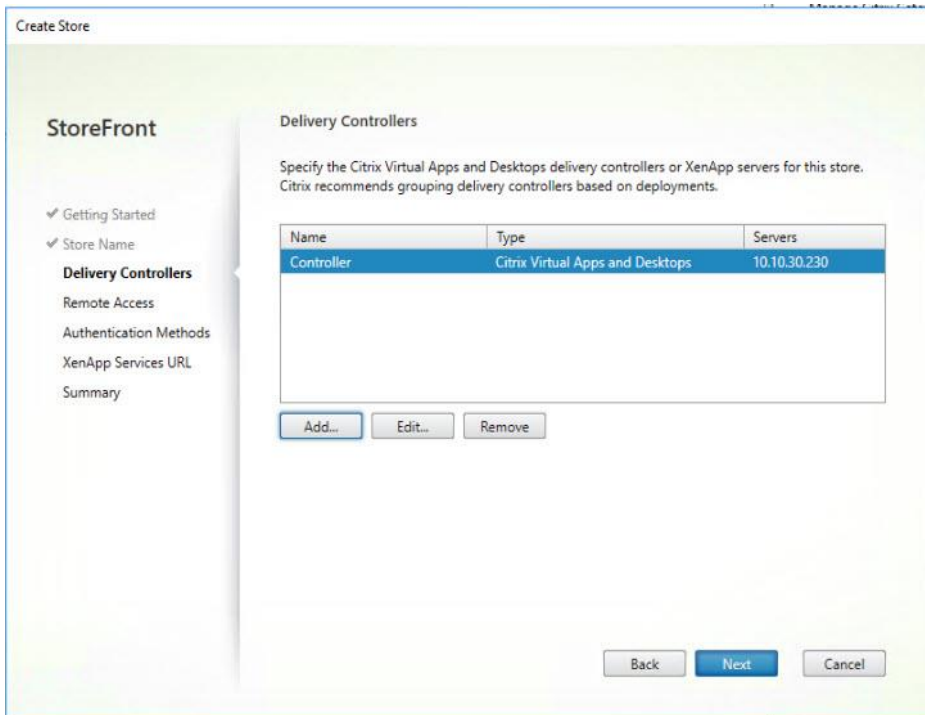
Step 15. Click Next.



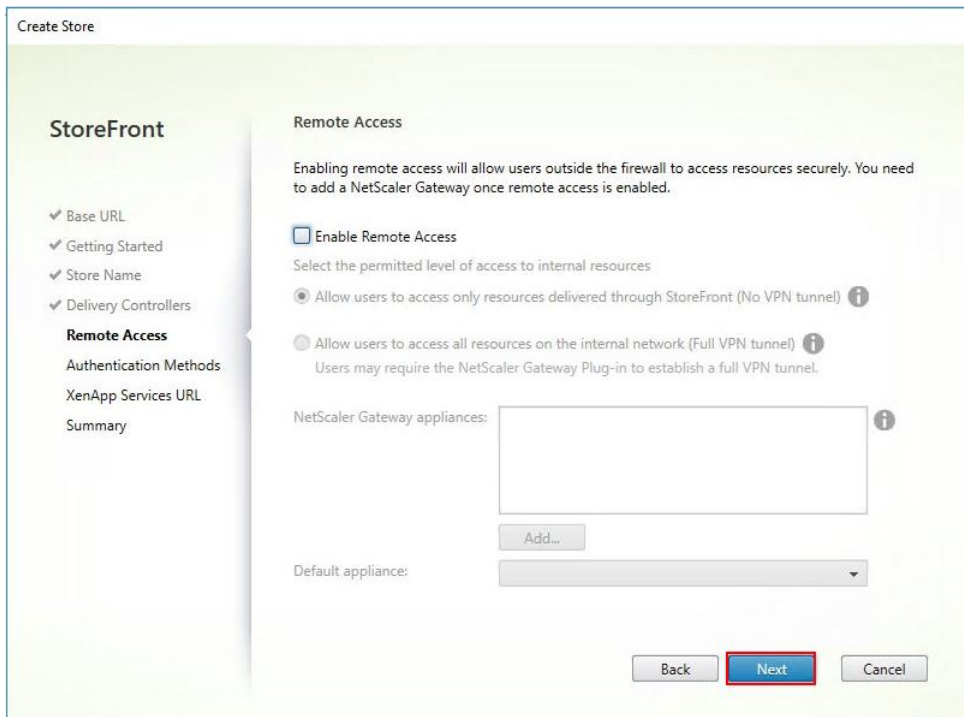
Step 16. Specify a name for your store and click Next.



Step 17. Add the required Delivery Controllers to the store and click Next.



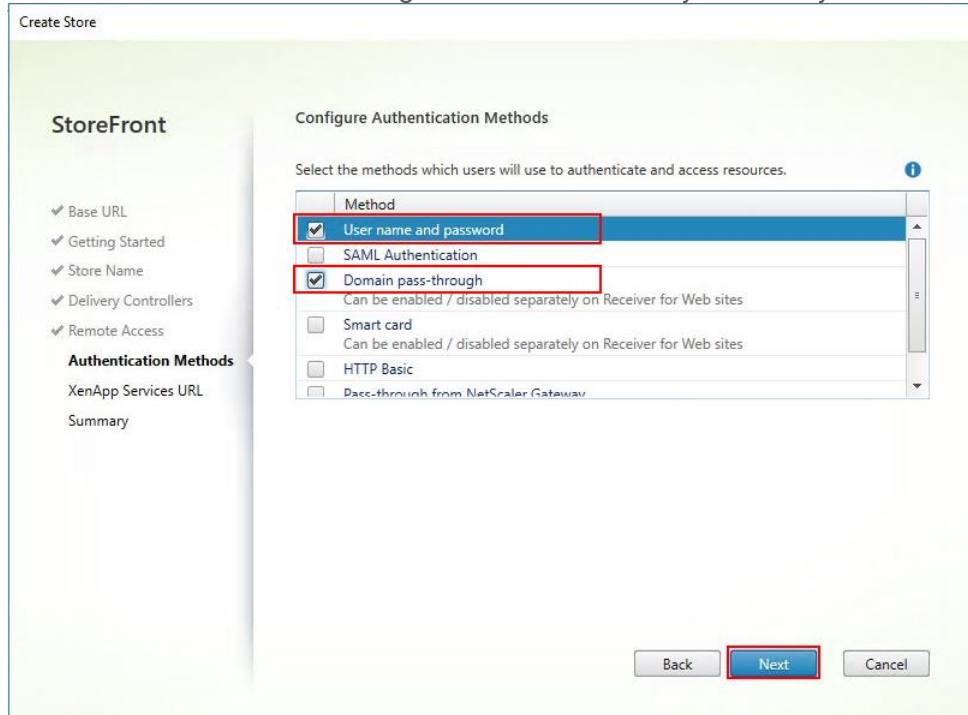
Step 18. Specify how connecting users can access the resources, in this environment only local users on the internal network are able to access the store and click Next.



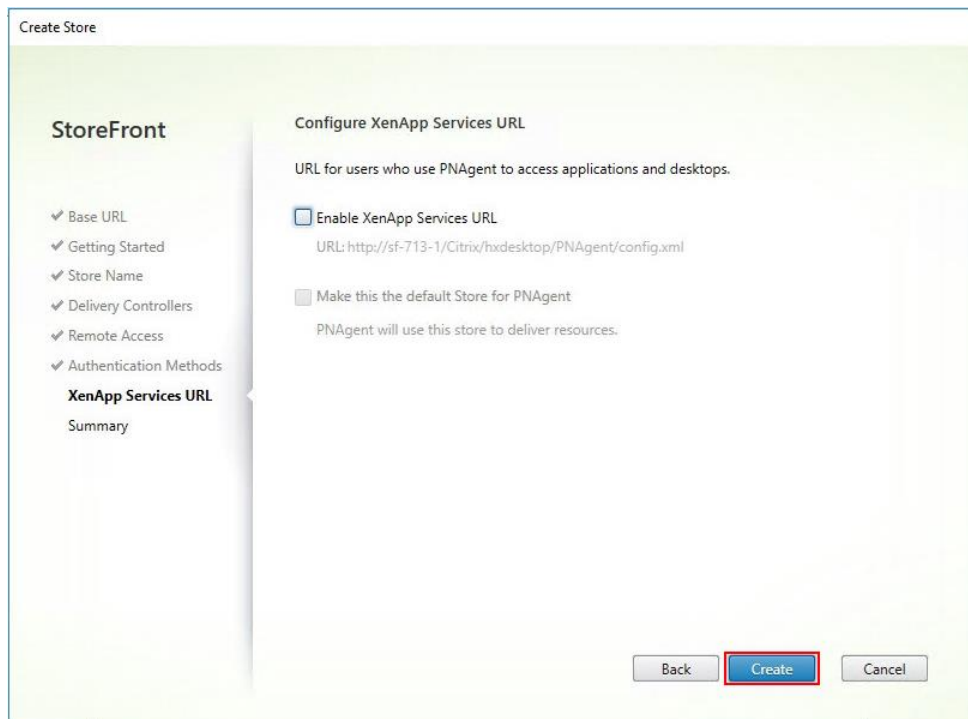
Step 19. On the “Authentication Methods” page, select the methods your users will use to authenticate to the store and click Next. You can select from the following methods as shown below:

- Username and password: Users enter their credentials and are authenticated when they access their stores.

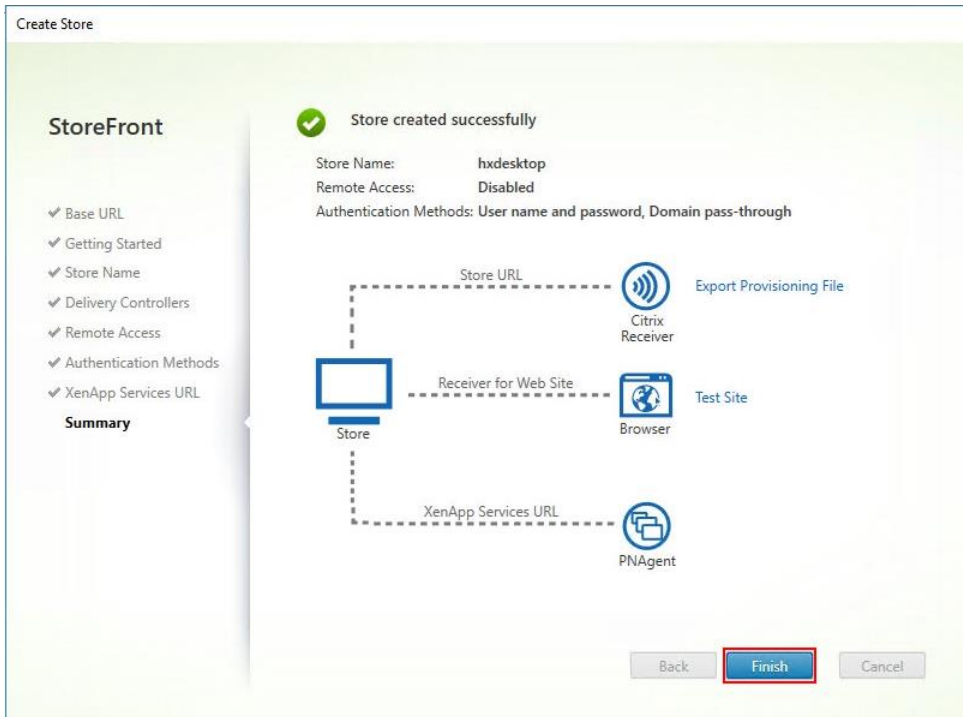
- Domain pass-through: Users authenticate to their domain-joined Windows computers and their credentials are used to log them on automatically when they access their stores.



Step 20. Configure the XenApp Service URL for users who use PNAgent to access the applications and desktops and click Create.



Step 21. After creating the store click Finish.

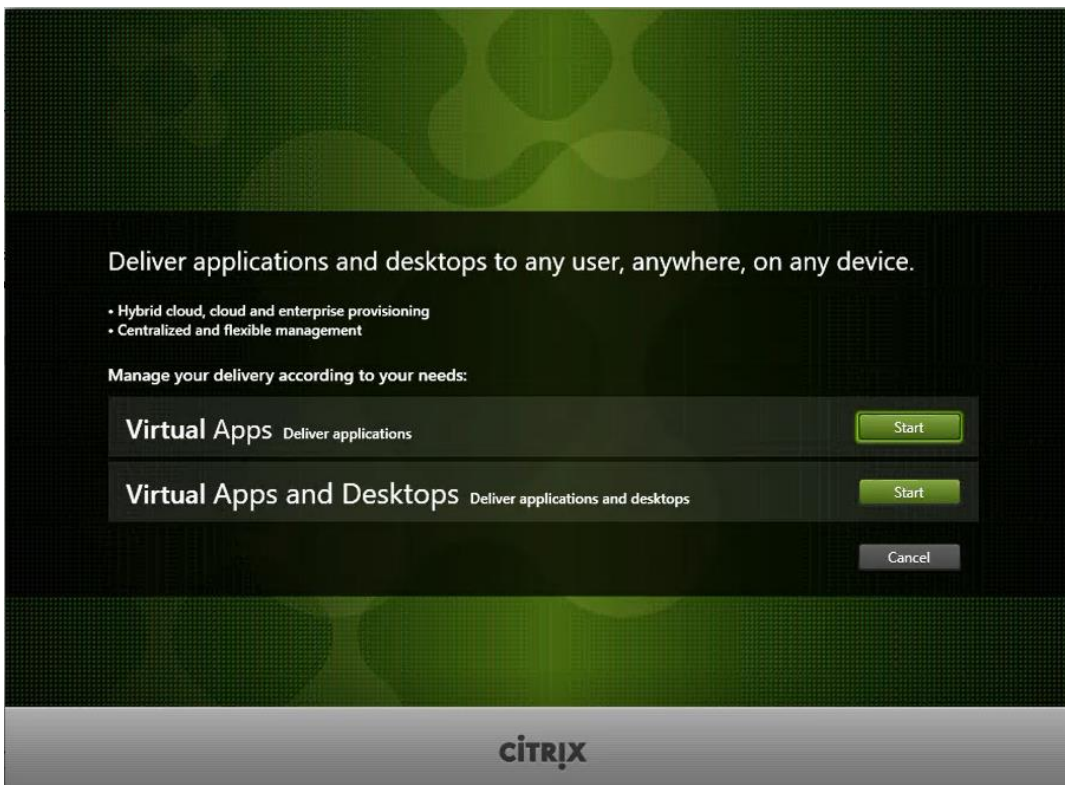


Procedure 9. Additional StoreFront Configuration

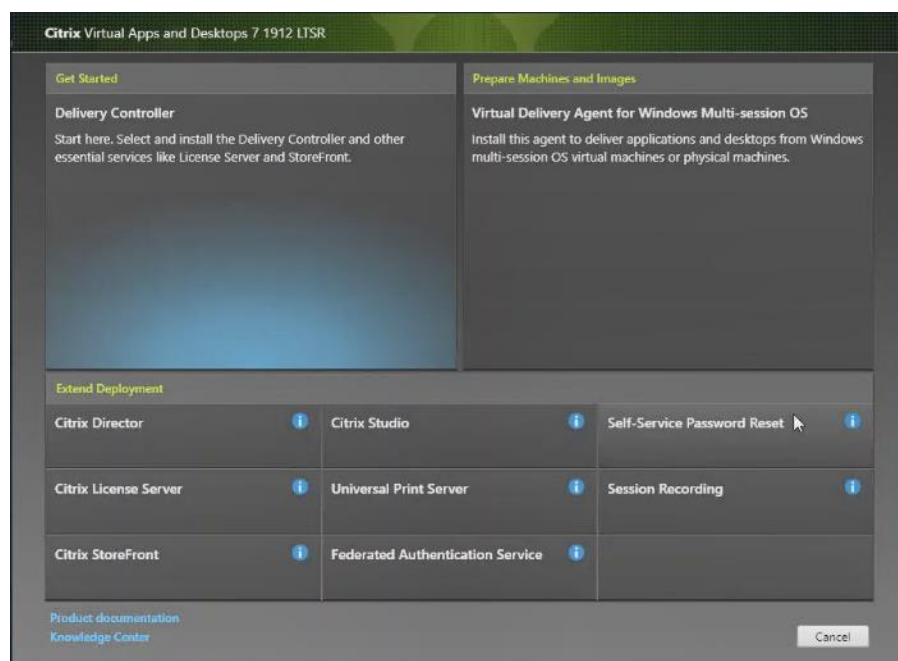
After the first StoreFront server is completely configured and the Store is operational, you can add additional servers.

Step 1. Connect to the second StoreFront server and launch the installer from the Citrix VDI ISO.

Step 2. Click Start.



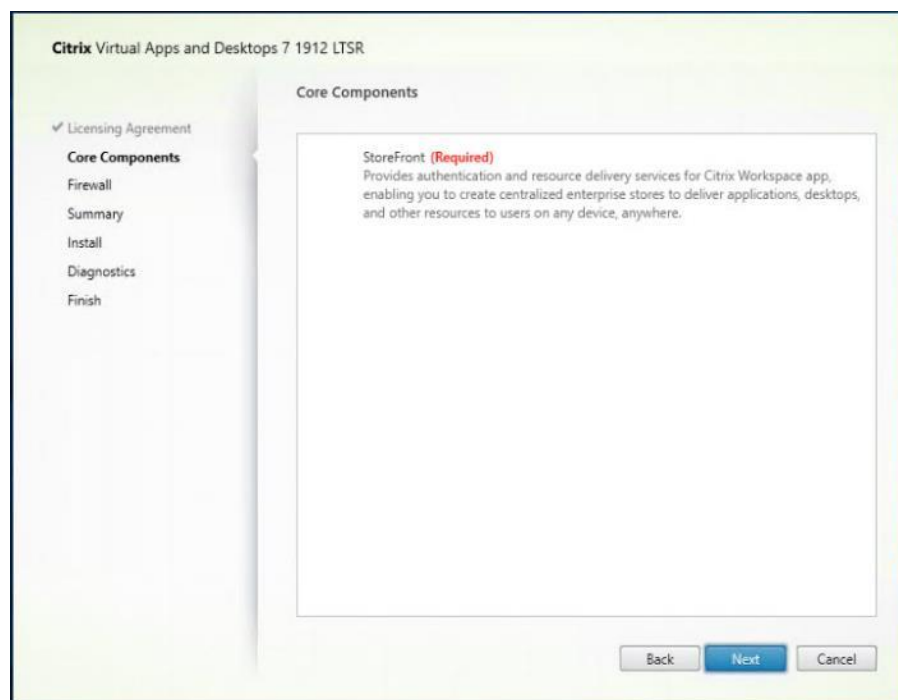
Step 3. Click Extended Deployment Citrix StoreFront.



Step 4. Repeat the steps used to install the first StoreFront.

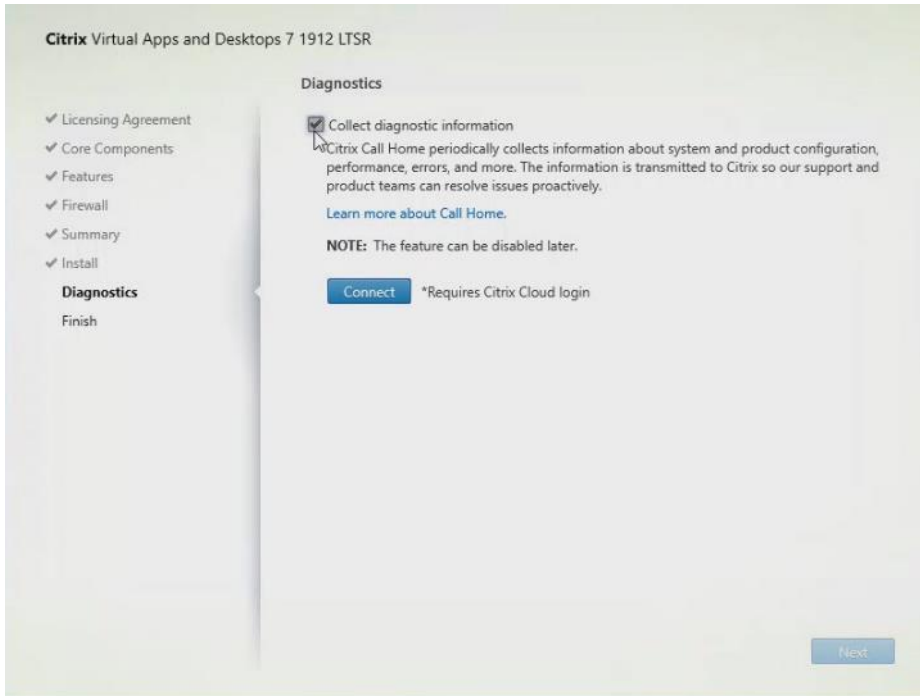
Step 5. Review the Summary configuration.

Step 6. Click Install.



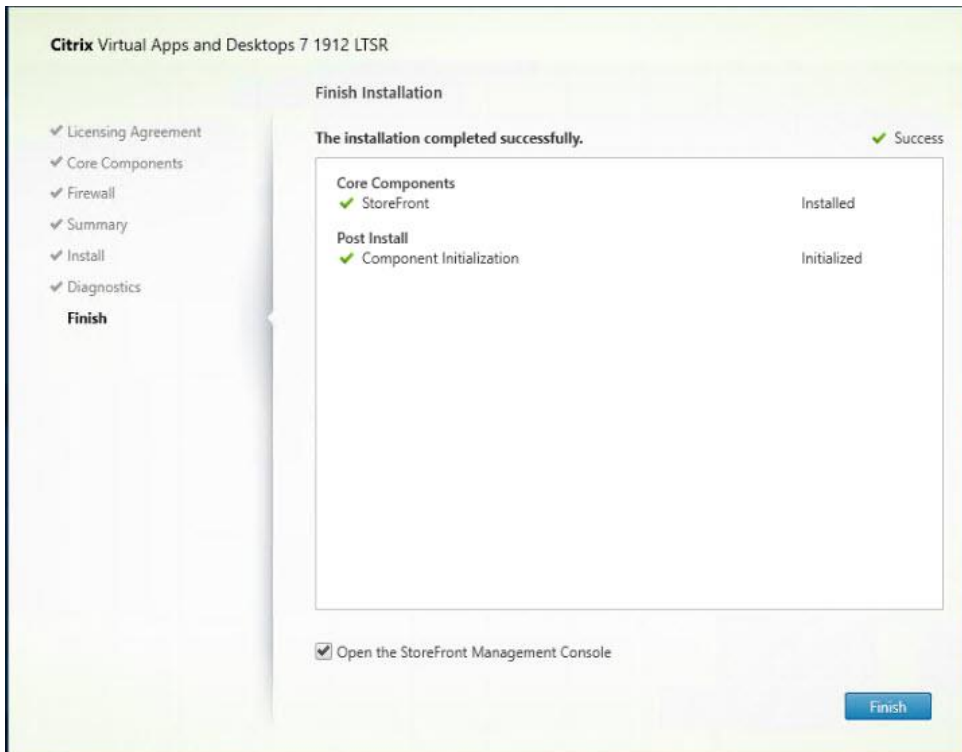
Step 7. (Optional) Click “Collect diagnostic information.”

Step 8. Click Next.



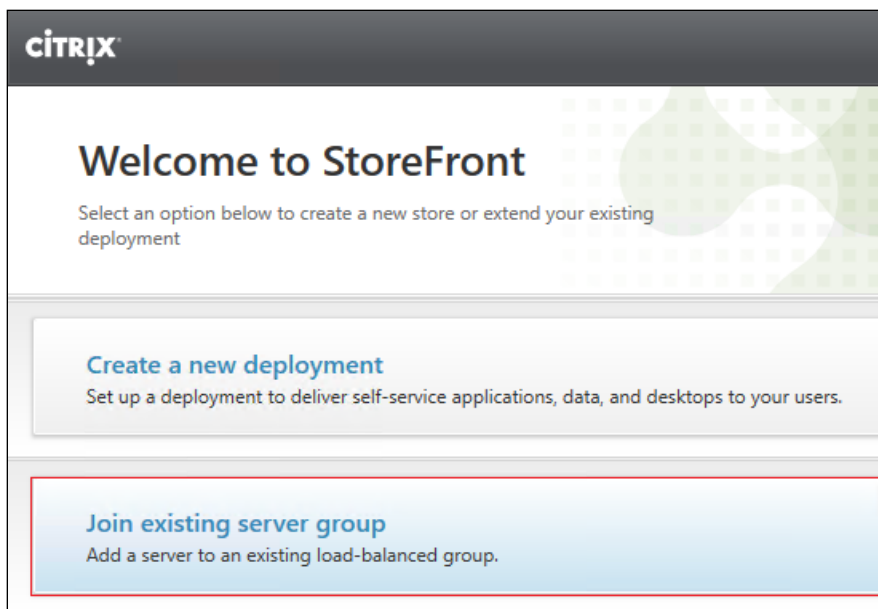
Step 9. Check “Open the StoreFront Management Console.”

Step 10. Click Finish.

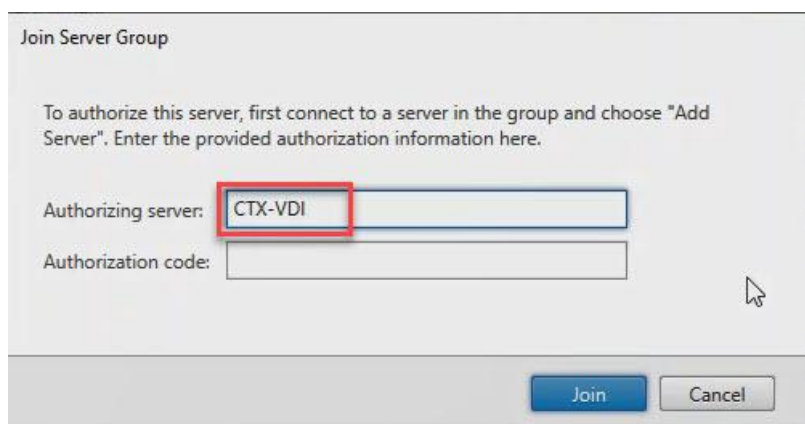


To configure the second StoreFront if used, follow these steps:

Step 1. From the StoreFront Console on the second server select “Join existing server group.”



Step 2. In the Join Server Group dialog, enter the name of the first Storefront server.

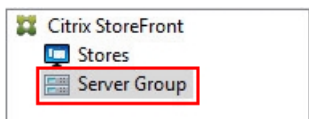


Step 3. Before the additional StoreFront server can join the server group, you must connect to the first Storefront server, add the second server, and obtain the required authorization information.

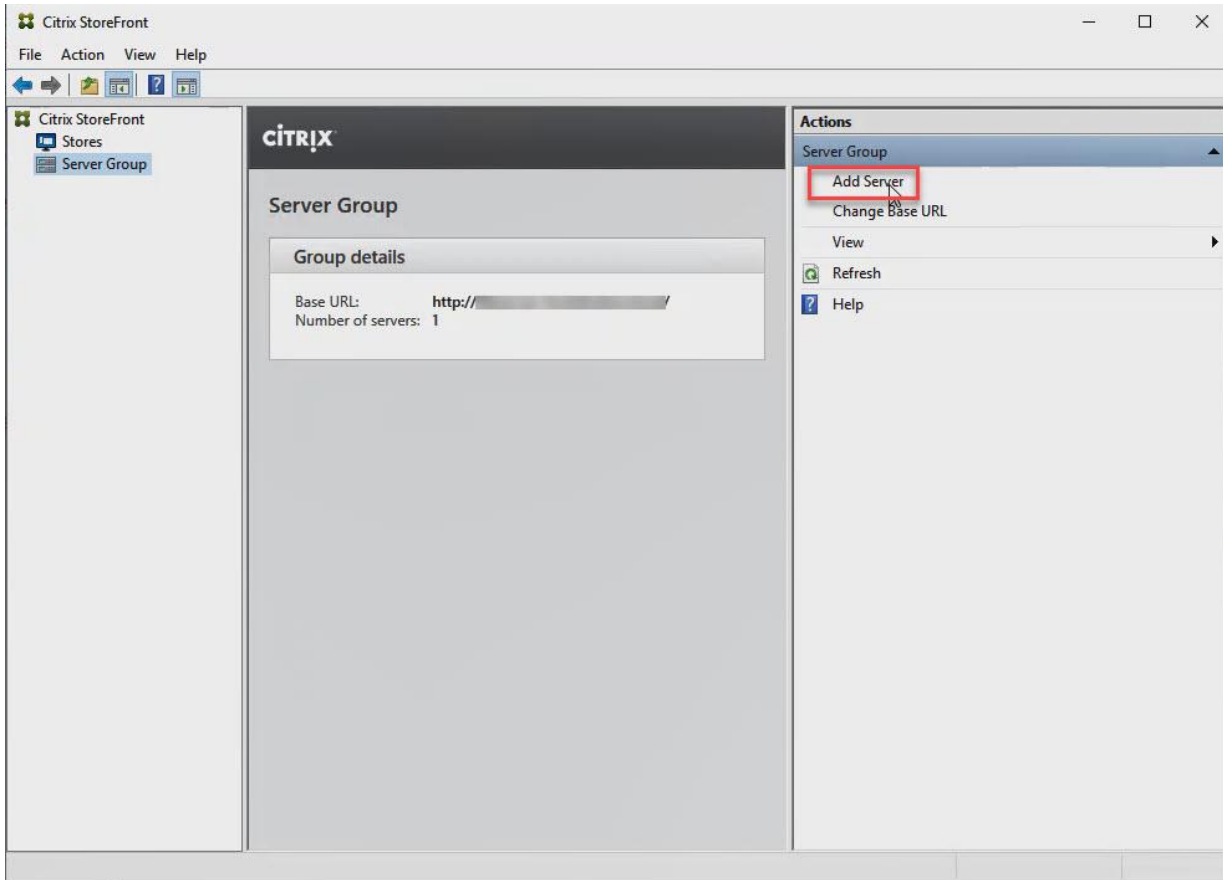
Step 4. Connect to the first StoreFront server.

Step 5. Using the StoreFront menu on the left, you can scroll through the StoreFront management options.

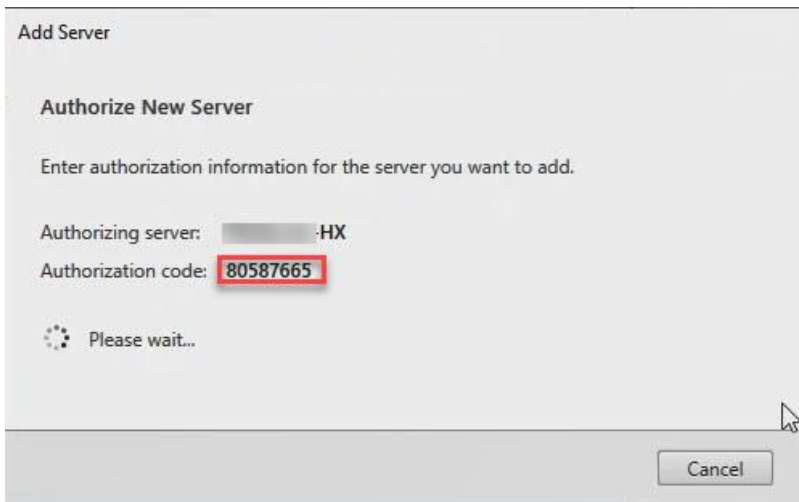
Step 6. Select Server Group from the menu.



Step 7. To add the second server and generate the authorization information that allows the additional StoreFront server to join the server group, select Add Server.

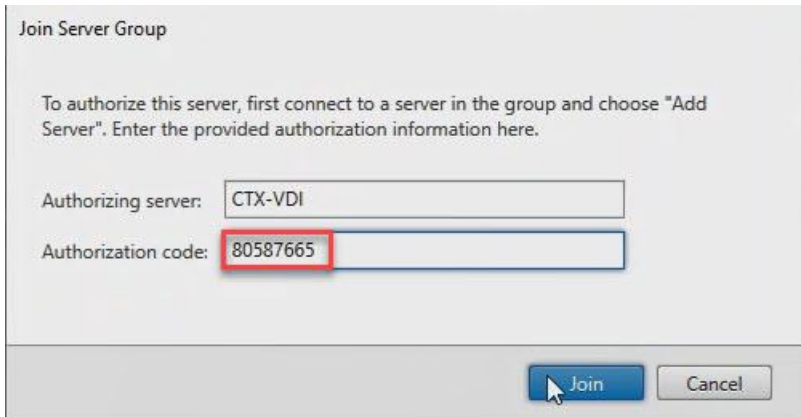


Step 8. Copy the Authorization code from the Add Server dialog.



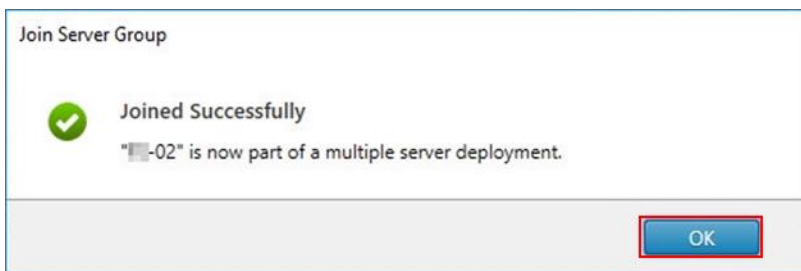
Step 9. Connect to the second Storefront server and paste the Authorization code into the Join Server Group dialog.

Step 10. Click Join.



Step 11. A message appears when the second server has joined successfully.

Step 12. Click OK.



The second StoreFront is now in the Server Group.

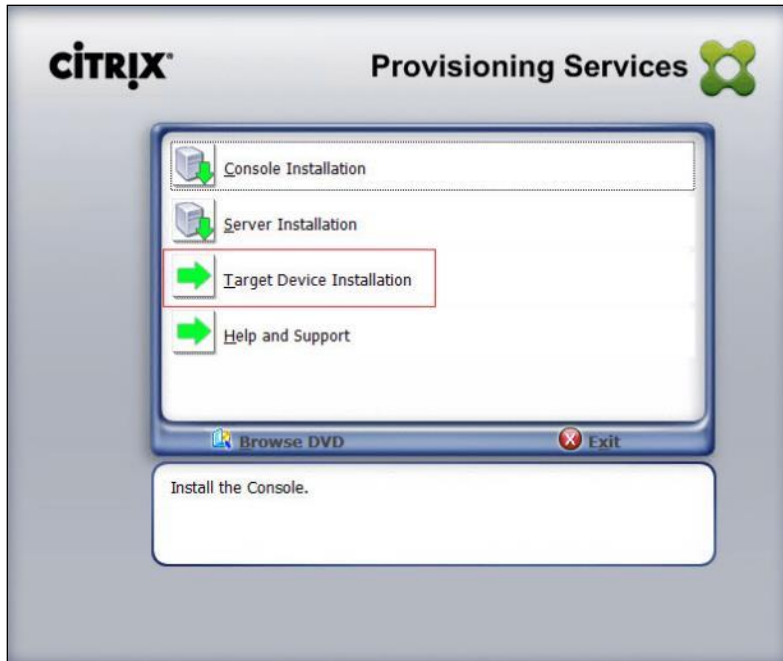
Procedure 10. Install the Citrix Provisioning Services Target Device Software

For non-persistent Windows 10 virtual desktops and Server 2019 RDS virtual machines, Citrix Provisioning Services (PVS) is used for deployment. The Master Target Device refers to the target device from which a hard disk image is built and stored on a vDisk. Provisioning Services then streams the contents of the vDisk created to other target devices. This procedure installs the PVS Target Device software that is used to build the RDS and VDI golden images.

Note: The instructions below outline the installation procedure to configure a vDisk for VDI desktops. When you have completed these installation steps, repeat the procedure to configure a vDisk for RDS.

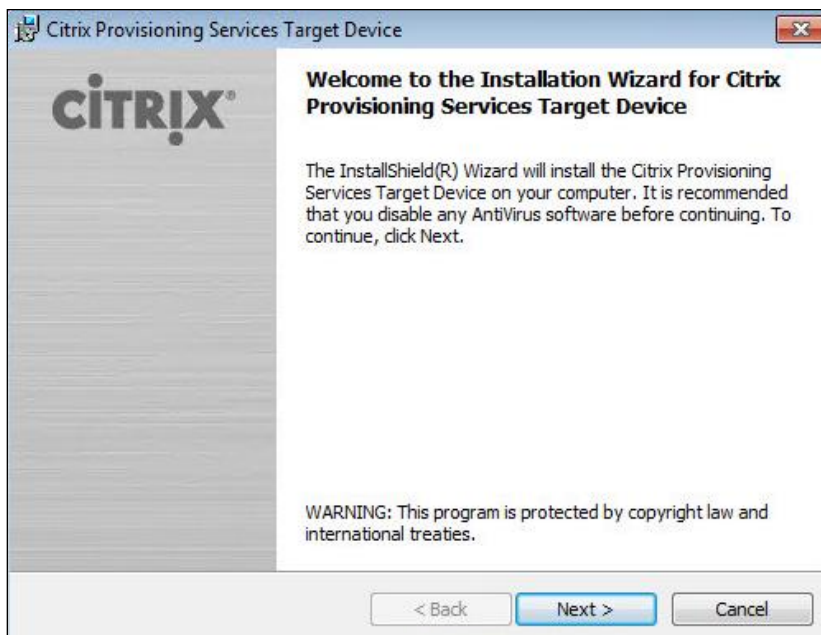
Step 1. On the Window 10 Master Target Device, launch the PVS installer from the Provisioning Services ISO.

Step 2. Click the Target Device Installation button.



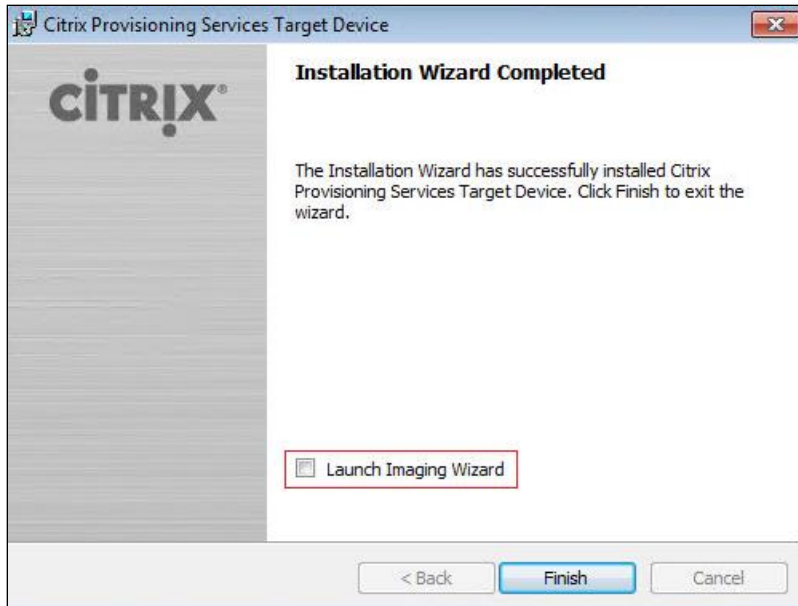
Note: The installation wizard will check to resolve dependencies and then begin the PVS target device installation process.

Step 3. Click Next.



Step 4. Confirm the installation settings and click Install.

Step 5. Deselect the checkbox to launch the Imaging Wizard and click Finish.



Step 6. Reboot the machine.

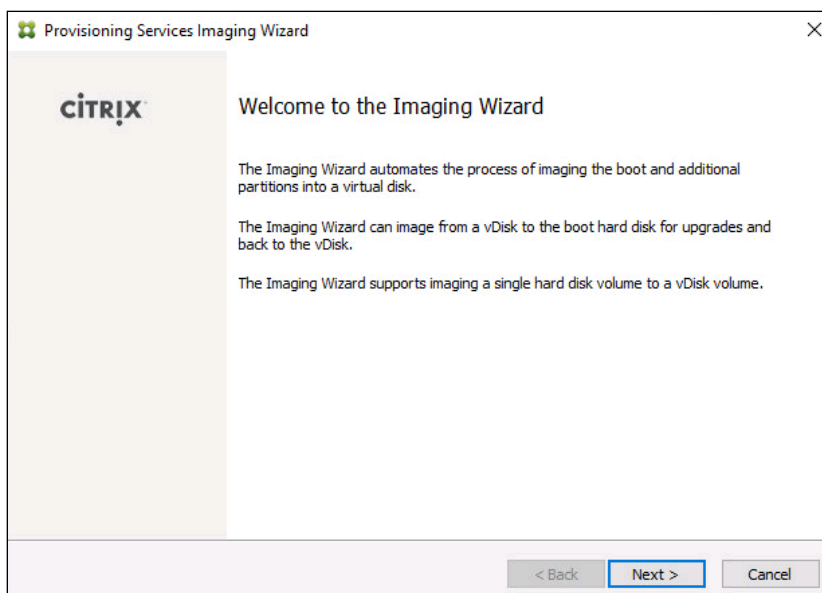
Procedure 11. Create Citrix Provisioning Services vDisks

The PVS Imaging Wizard automatically creates a base vDisk image from the master target device.

Note: The following procedure explains how to create a vDisk for VDI desktops. When you have completed these steps, repeat the procedure to build a vDisk for RDS.

Step 1. The PVS Imaging Wizard's Welcome page appears.

Step 2. Click Next.



Step 3. The Connect to Farm page appears. Enter the name or IP address of a Provisioning Server within the farm to connect to and the port to use to make that connection.

Step 4. Use the Windows credentials (default) or enter different credentials.

Step 5. Click Next.

Provisioning Services Imaging Wizard

Connect to Provisioning Services Site

Enter the Provisioning Services site server name or IP, port, and credentials. Only stores supported by this server will be available for vDisk assignment.

Enter Server Details

Server name or IP:

Port:

Provide Logon Credentials for the Server

Use my Windows credentials

Use these credentials

User name:

Domain:

Password:

< Back Next > Cancel

Step 6. Select Create new vDisk.

Step 7. Click Next.

Provisioning Services Imaging Wizard

Imaging Options

What task do you want to perform?

Create a vDisk
Make a Provisioning Services vDisk from this device's boot hard disk.

Recreate an existing vDisk
Not available because there are no vDisks assigned to the server.

Create an image file
Make an image file from this device's booted disk, for importing into Provisioning Services.

Copy a hard disk volume to a vDisk volume
Not available because there are no vDisks assigned to the server.

< Back Next > Cancel

The Add Target Device page appears.

Step 8. Select the Target Device Name, the MAC address associated with one of the NICs that was selected when the target device software was installed on the master target device, and the Collection to which you are adding the device.

Step 9. Click Next.

Provisioning Services Imaging Wizard

Add Target Device

This device is not a member of the site and needs to be added.

Target device name:
Must be different from the current machine name.

Network connection:
Select the connection that will be used to boot this machine to the server.

Collection name:
Select the site collection that this device will be added to.

< Back Next > Cancel

Step 10. The New vDisk dialog displays. Enter the name of the vDisk.

Step 11. Select the Store where the vDisk will reside. Select the vDisk type, either Fixed or Dynamic, from the drop-down menu. (This CVD used Dynamic rather than Fixed vDisks.)

Step 12. Click Next.

Provisioning Services Imaging Wizard

New vDisk

The new vDisk will be created in the store you select.

vDisk name:

Store name:
Supported by Server: CTX-PVS1

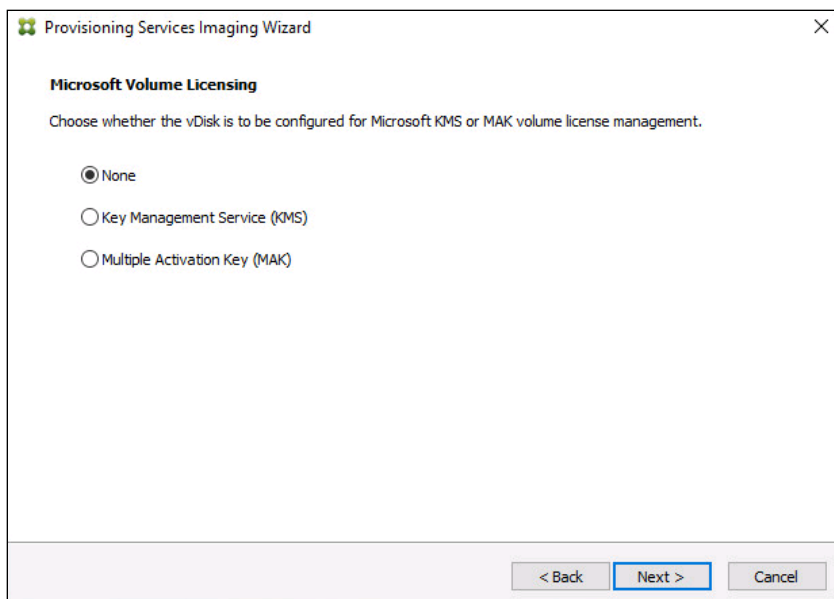
vDisk type:

VHDX
 VHD

< Back Next > Cancel

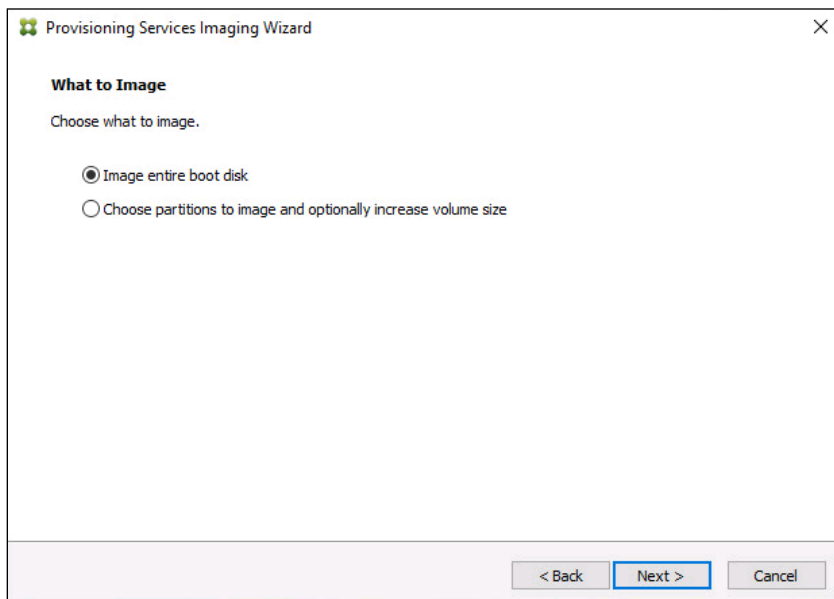
Step 13. On the Microsoft Volume Licensing page, select the volume license option to use for target devices. For this CVD, volume licensing is not used, so the None button is selected.

Step 14. Click Next.



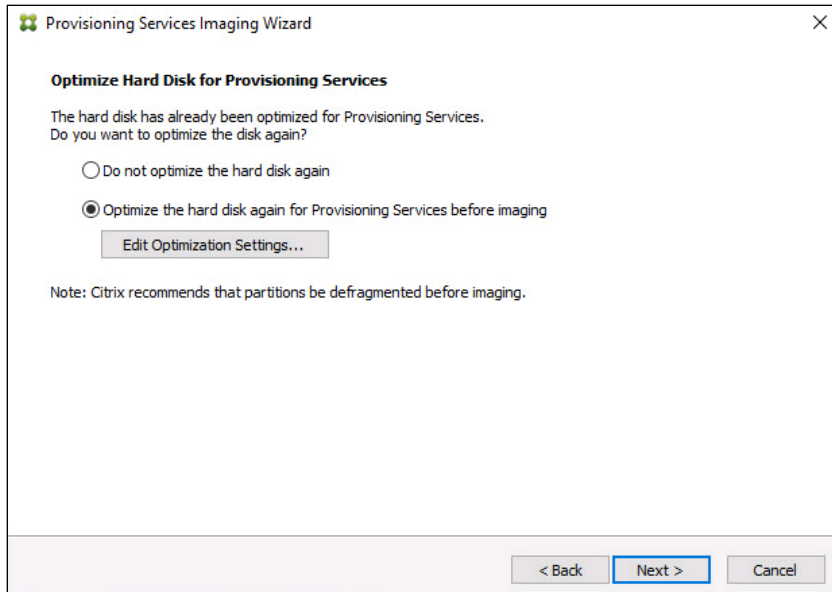
Step 15. Select Image entire boot disk on the Configure Image Volumes page.

Step 16. Click Next.

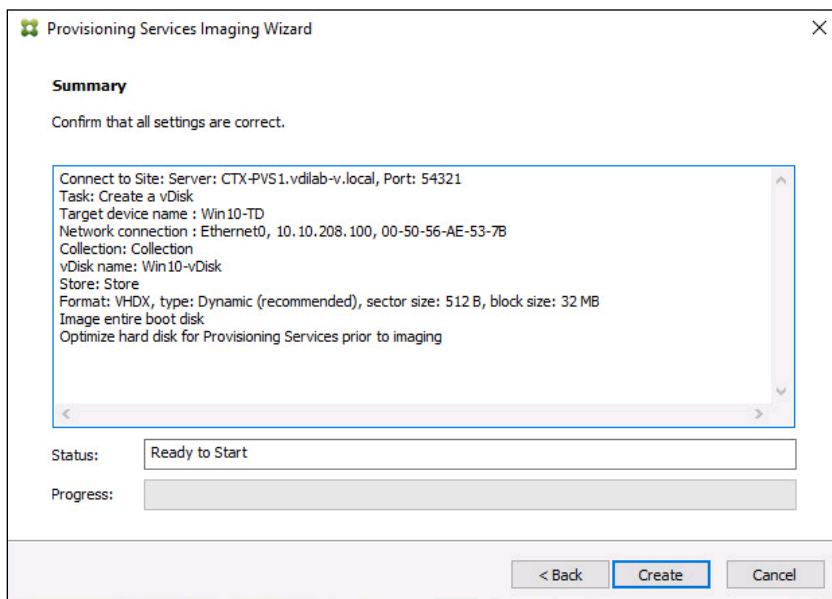


Step 17. Select Optimize for hard disk again for Provisioning Services before imaging on the Optimize Hard Disk for Provisioning Services.

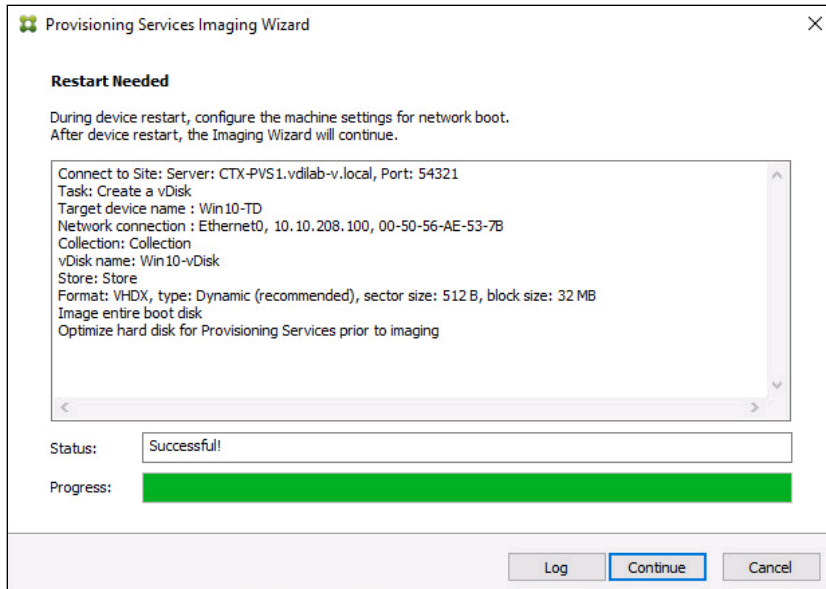
Step 18. Click Next.



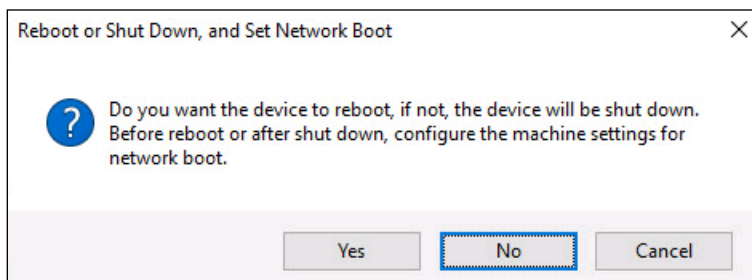
Step 19. Select Create on the Summary page.



Step 20. Review the configuration and click Continue.

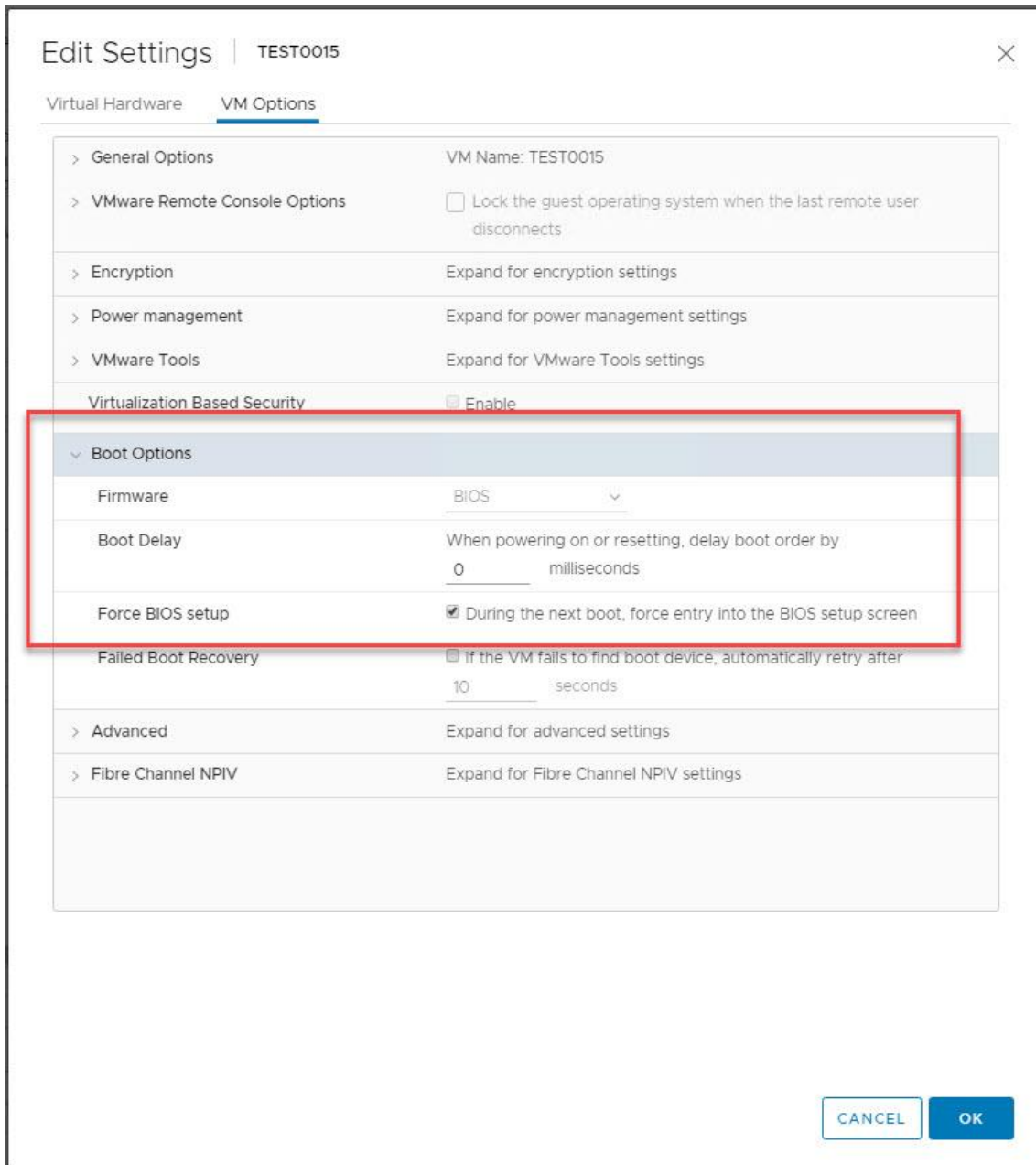


Step 21. When prompted, click No to shut down the machine.



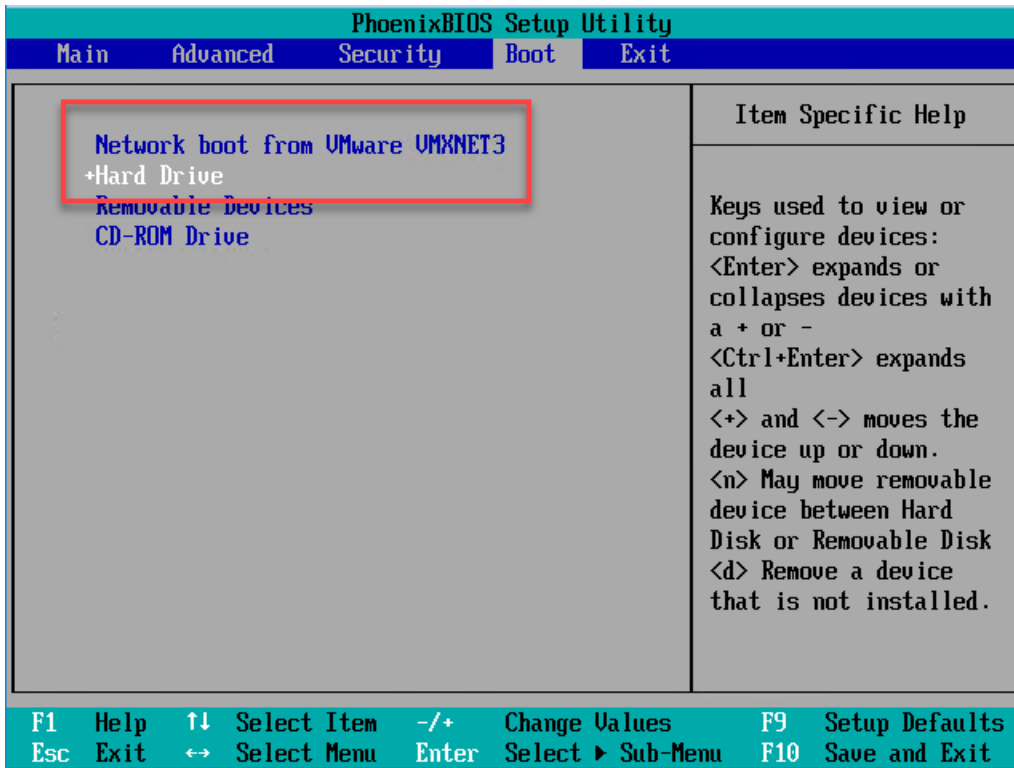
Step 22. Edit the virtual machine settings and select Boot options under VM Options.

Step 23. Select Force BIOS setup.



Step 24. Restart Virtual Machine.

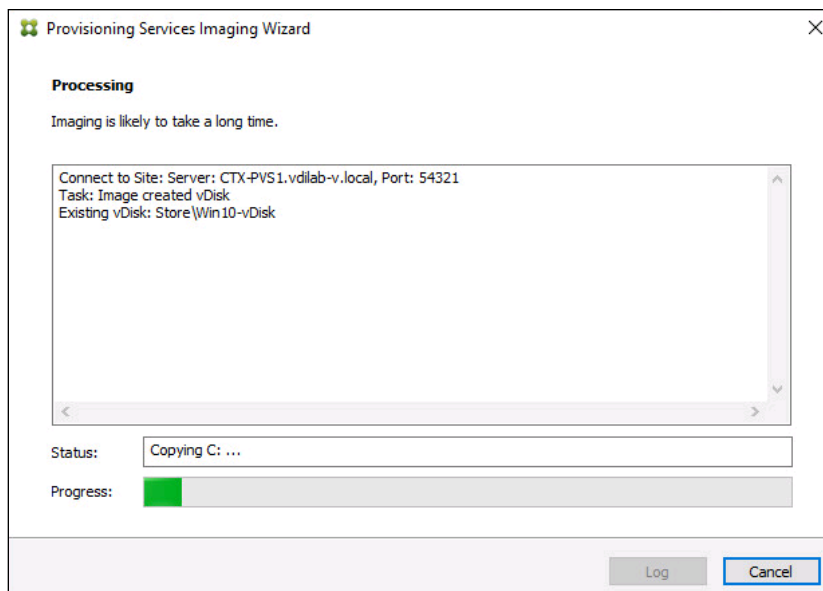
Step 25. When the VM boots into the BIOS, go to 'Boot' menu to move the Network boot from VMware VMXNET3 to the top of the list.



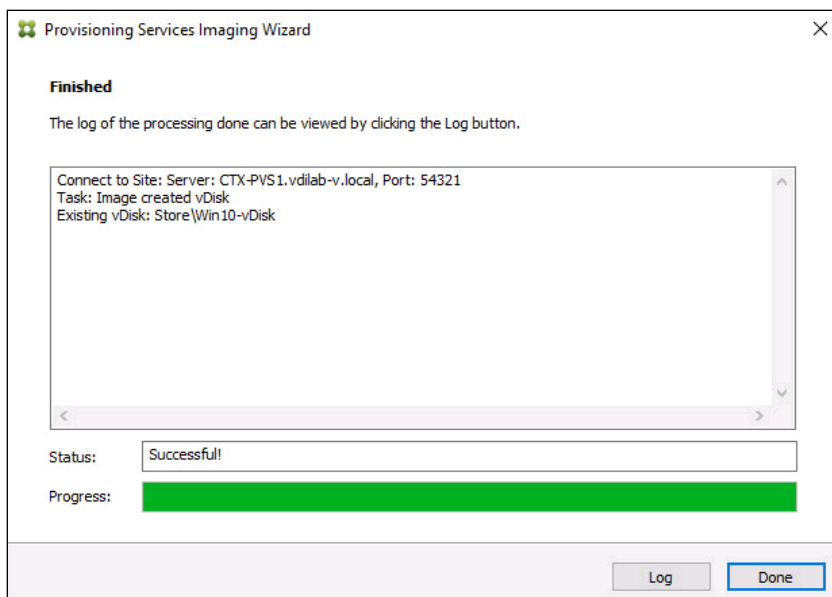
Step 26. Restart Virtual Machine

Note: After restarting the virtual machine, log into the VDI or RDS master target. The PVS imaging process begins, copying the contents of the C: drive to the PVS vDisk located on the server.

Step 27. If prompted to Restart select Restart Later.



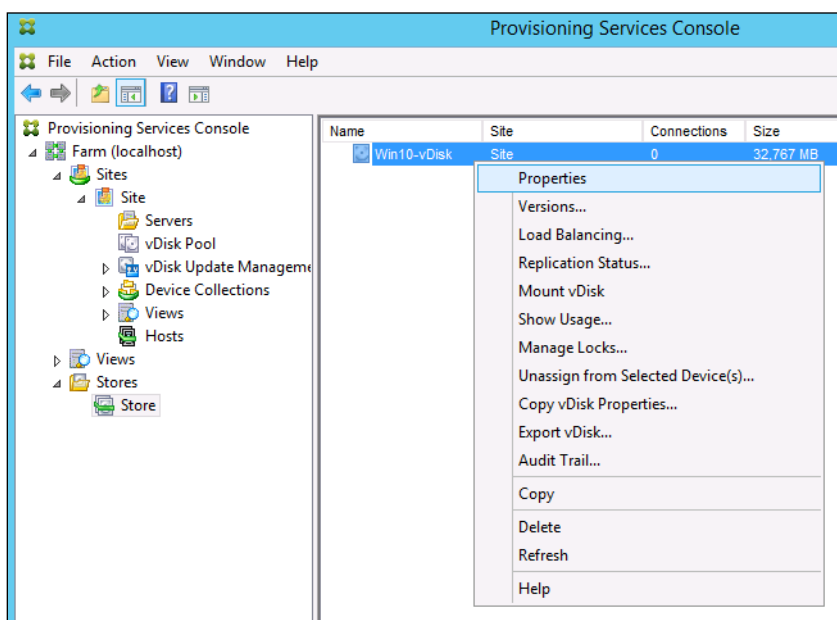
Step 28. A message is displayed when the conversion is complete, click Done.



Step 29. Shutdown the virtual machine used as the VDI or RDS master target.

Step 30. Connect to the PVS server and validate that the vDisk image is available in the Store.

Step 31. Right-click the newly created vDisk and select Properties.



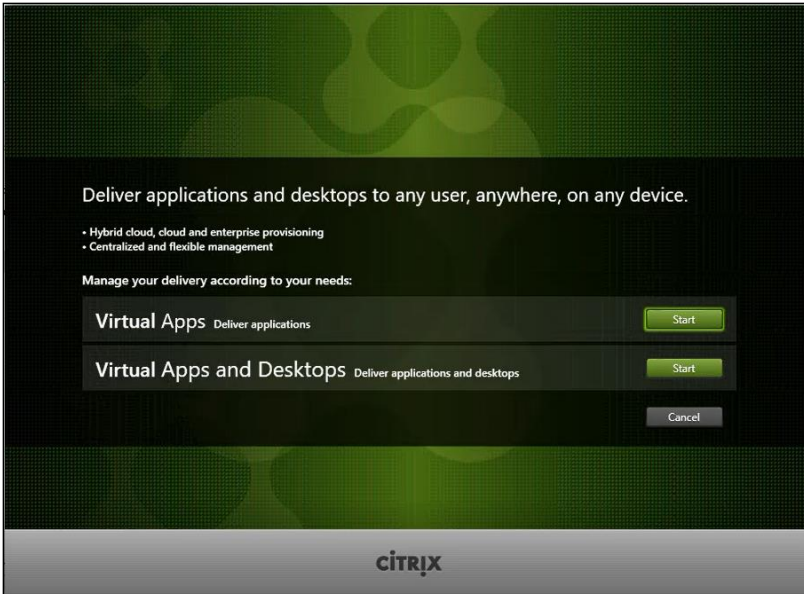
Step 32. On the vDisk Properties dialog, change Access mode to “Private” mode so the Citrix Virtual Desktop Agent can be installed.

Procedure 12. Install Citrix Virtual Apps and Desktop Virtual Desktop Agents

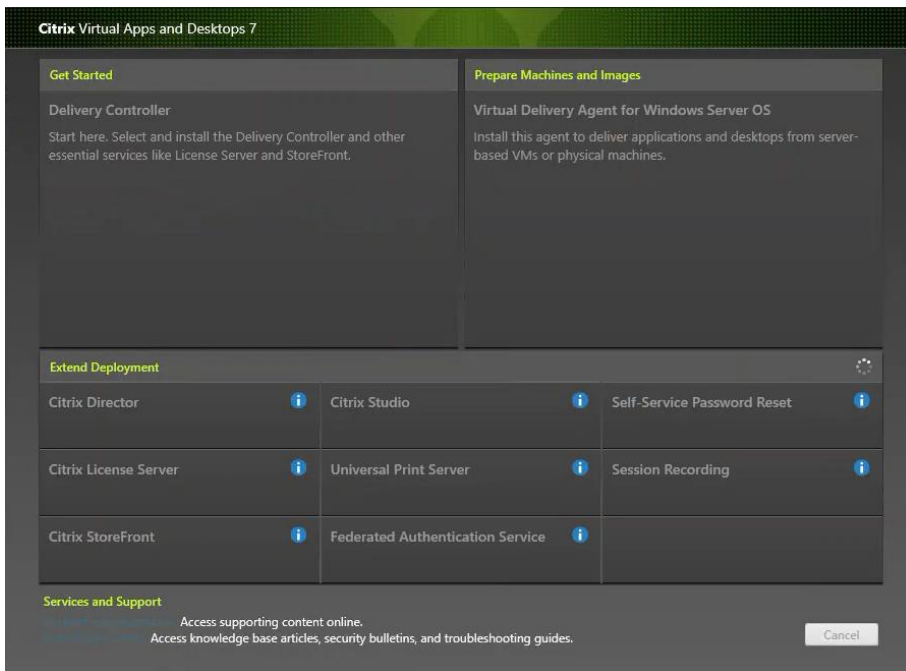
Virtual Delivery Agents (VDAs) are installed on the server and workstation operating systems and enable connections for desktops and apps. The following procedure was used to install VDAs for both HVD and HSD environments.

Step 1. Launch the Citrix Desktop installer from the CVA Desktop 1912 LTSR ISO.

Step 2. Click Start on the Welcome Screen.

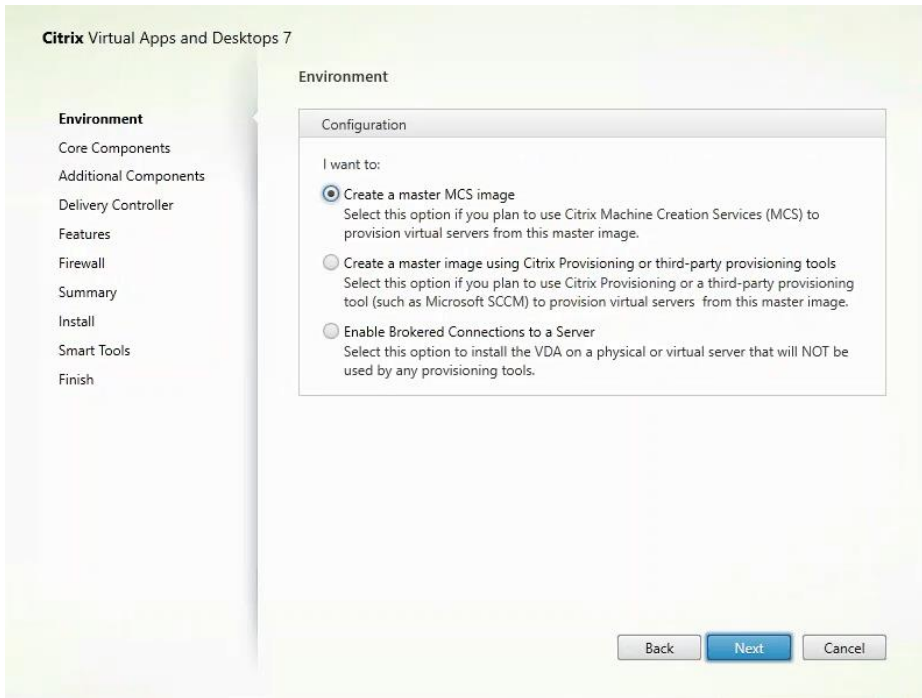


Step 3. To install the VDA for the Hosted Virtual Desktops (VDI), select Virtual Delivery Agent for Windows Desktop OS. After the VDA is installed for Hosted Virtual Desktops, repeat the procedure to install the VDA for Hosted Shared Desktops (RDS). In this case, select Virtual Delivery Agent for Windows Server OS and follow the same basic steps.



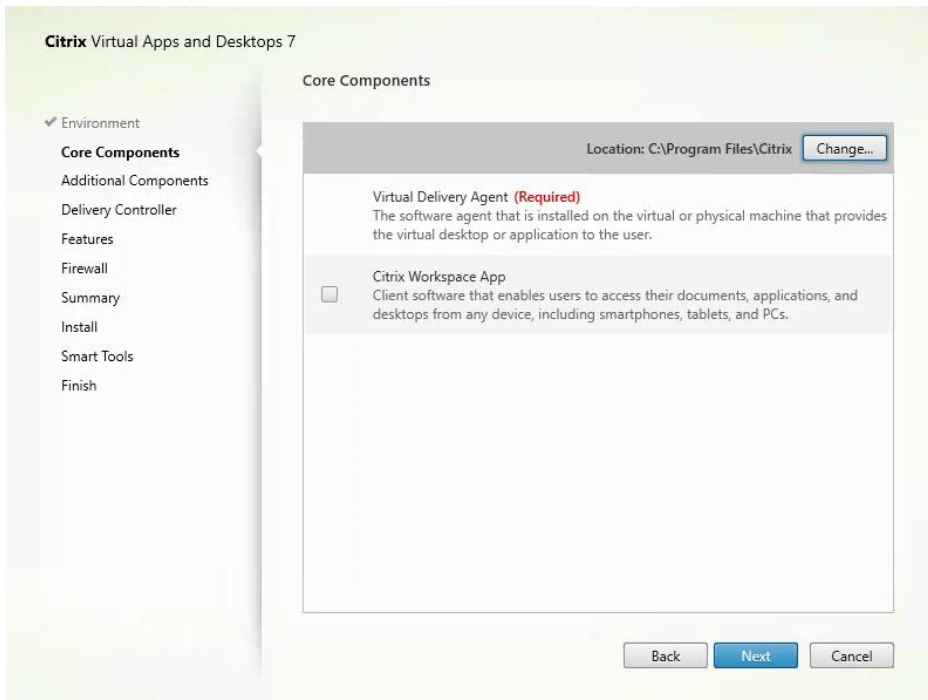
Step 4. Select “Create a Master Image.”(Be sure to select the proper provisioning technology here)

Step 5. Click Next.

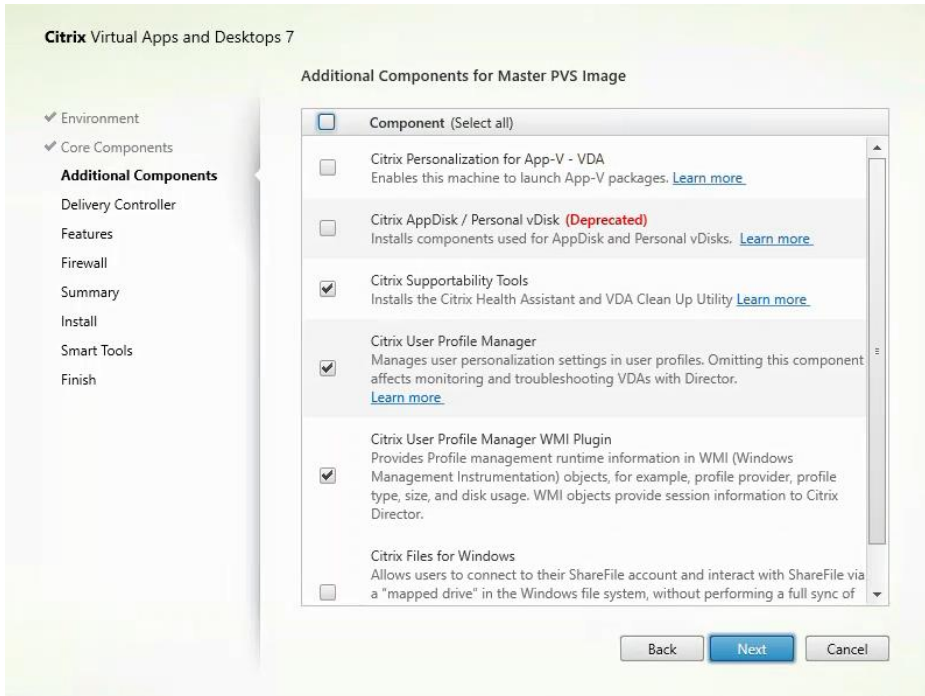


Step 6. Optional: Select Citrix Workspace App.

Step 7. Click Next.

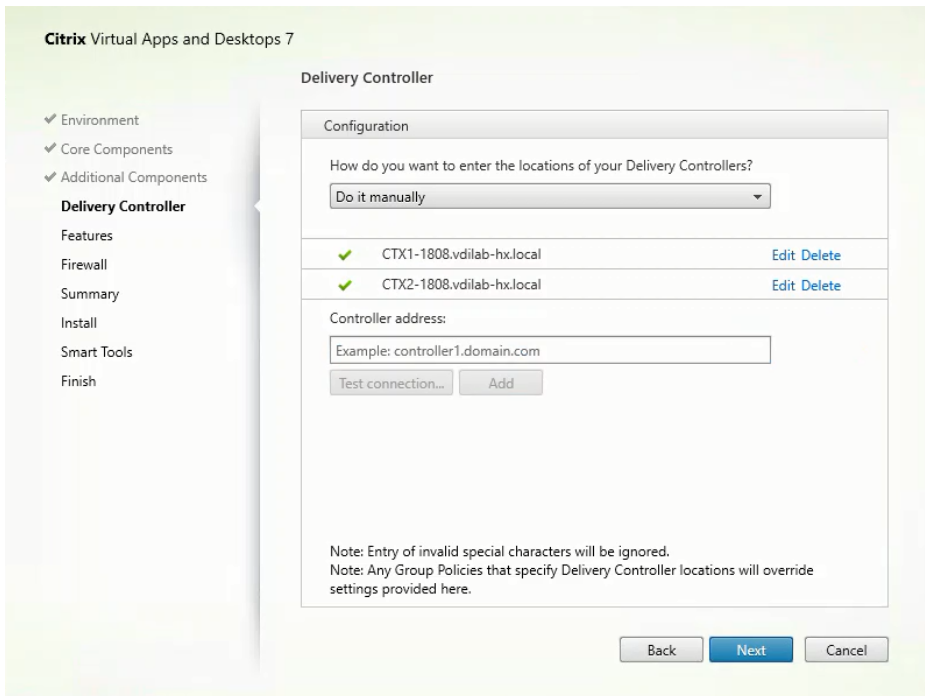


Step 8. Click Next.



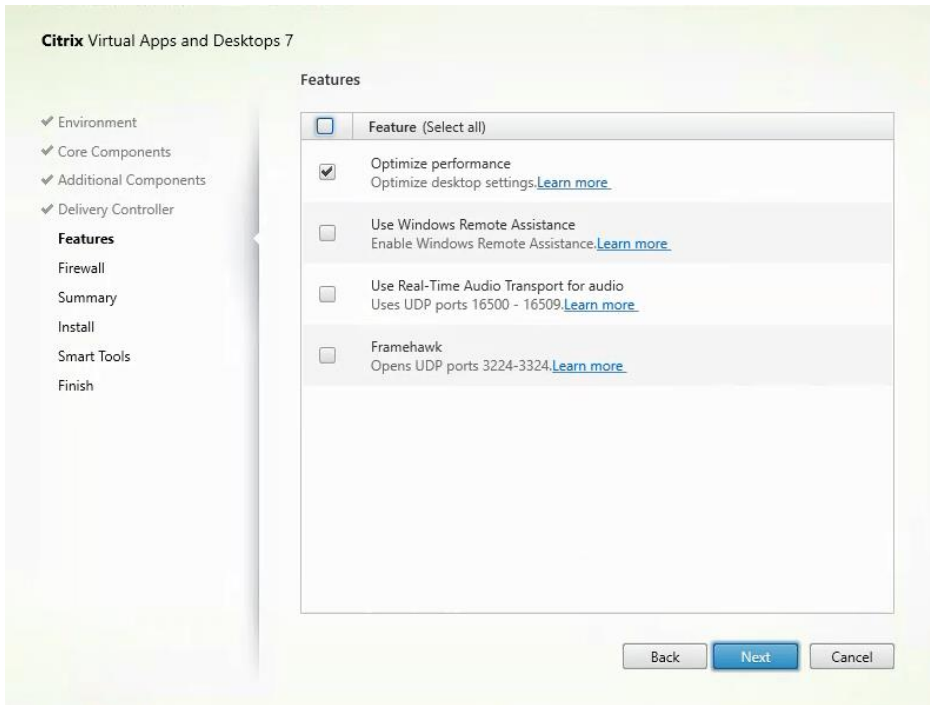
Step 9. Select “Do it manually” and specify the FQDN of the Delivery Controllers.

Step 10. Click Next.



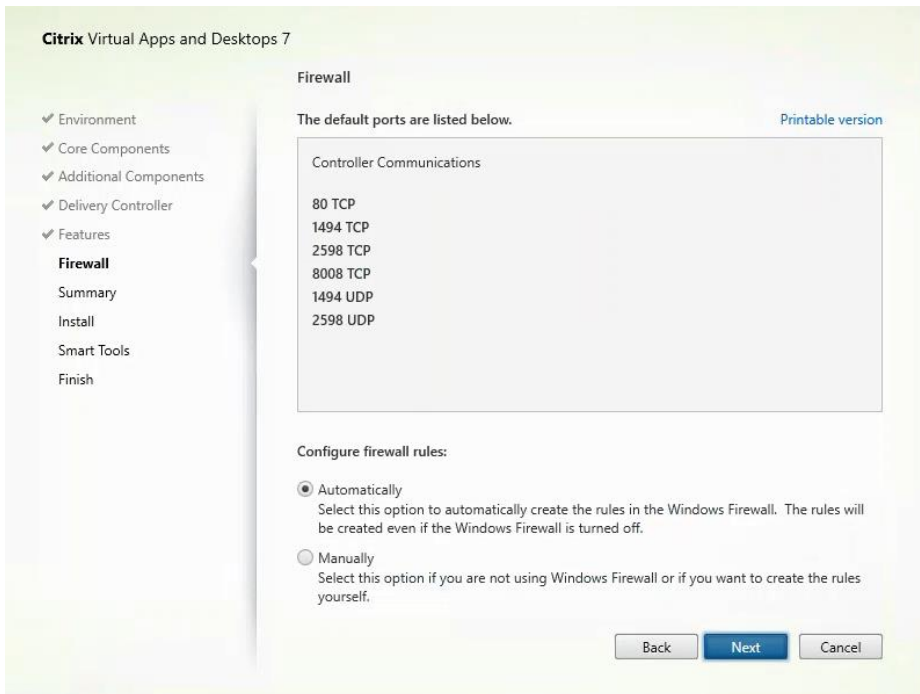
Step 11. Accept the default features.

Step 12. Click Next.

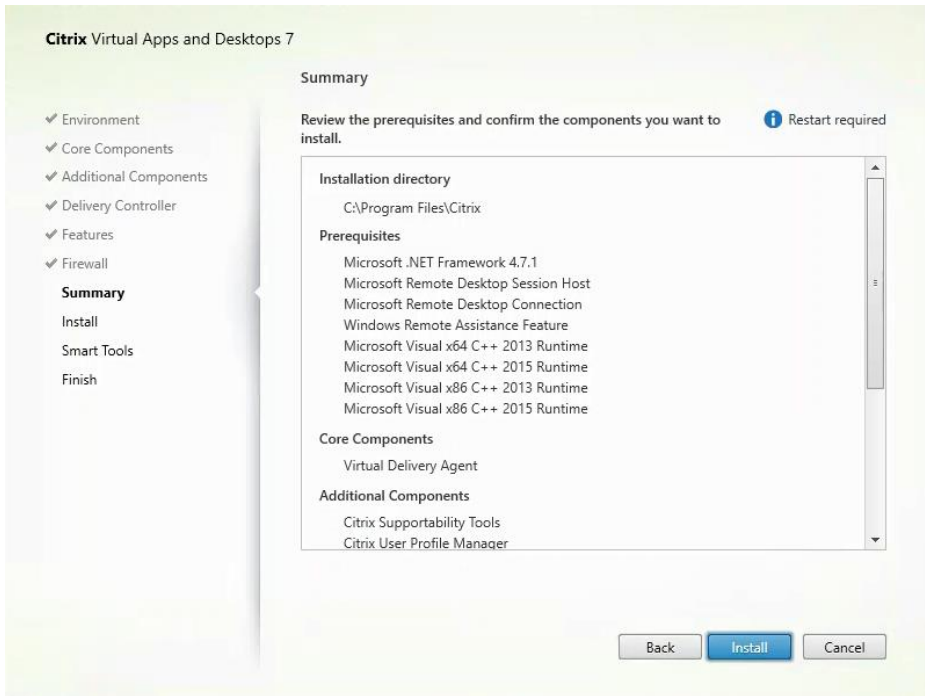


Step 13. Allow the firewall rules to be configured automatically.

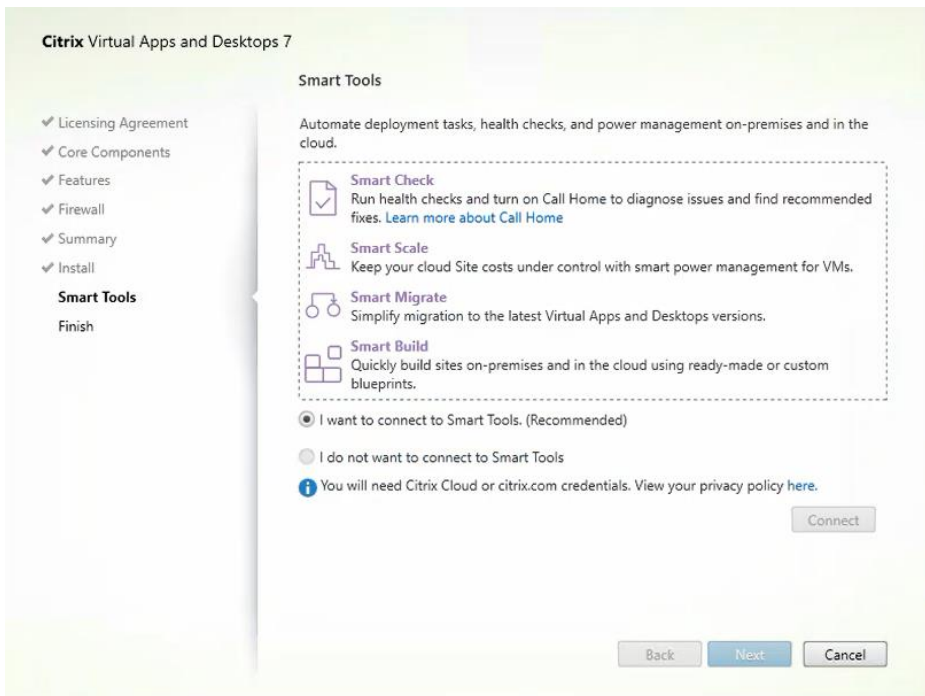
Step 14. Click Next.



Step 15. Verify the Summary and click Install.



Step 16. (Optional) Select Call Home participation.

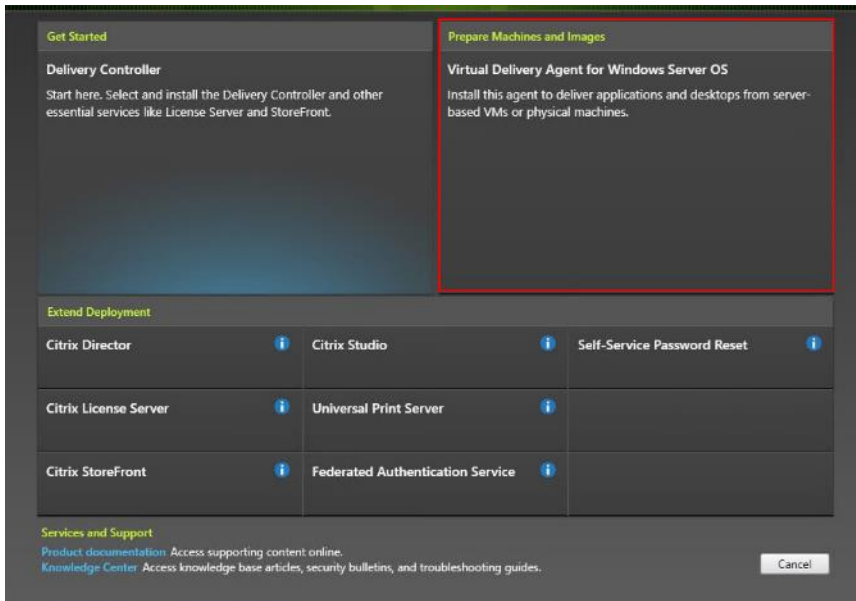


Step 17. (Optional) check “Restart Machine.”

Step 18. Click Finish.

Step 19. Repeat these procedure so that VDAs are installed for both HVD (using the Windows 10 OS image) and the HSD desktops (using the Windows Server 2019 image).

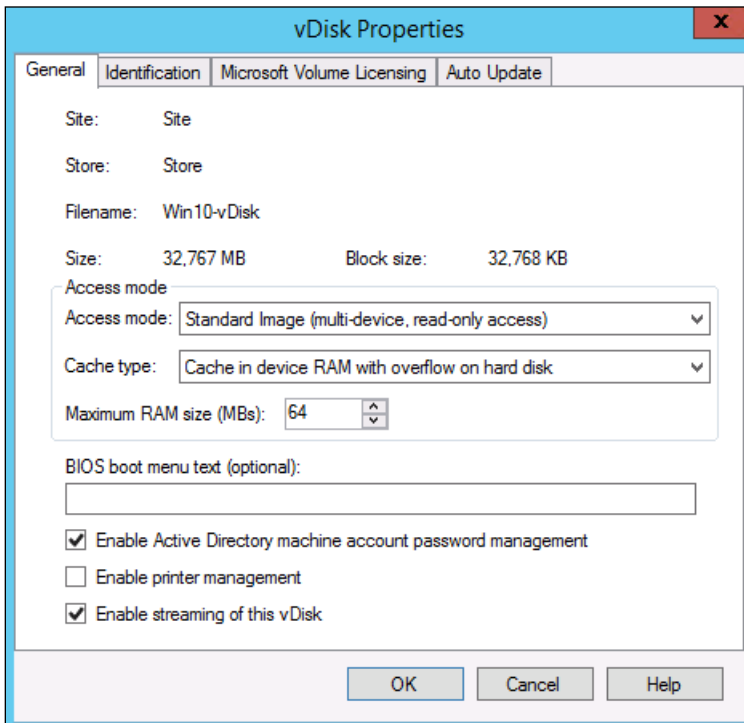
Step 20. Select an appropriate workflow for the HSD desktop.



Step 21. Once the Citrix VDA is installed, on the vDisk Properties dialog, change Access mode to “Standard Image (multi-device, read-only access).”

Step 22. Set the Cache Type to “Cache in device RAM with overflow on hard disk.”

Step 23. Set Maximum RAM size (MBs): 256 for VDI and set 1024 MB for RDS vDisk.



Step 24. Click OK.

Step 25. Repeat this procedure to create vDisks for both the Hosted VDI Desktops (using the Windows 10 OS image) and the Hosted Shared Desktops (using the Windows Server 2019 image).

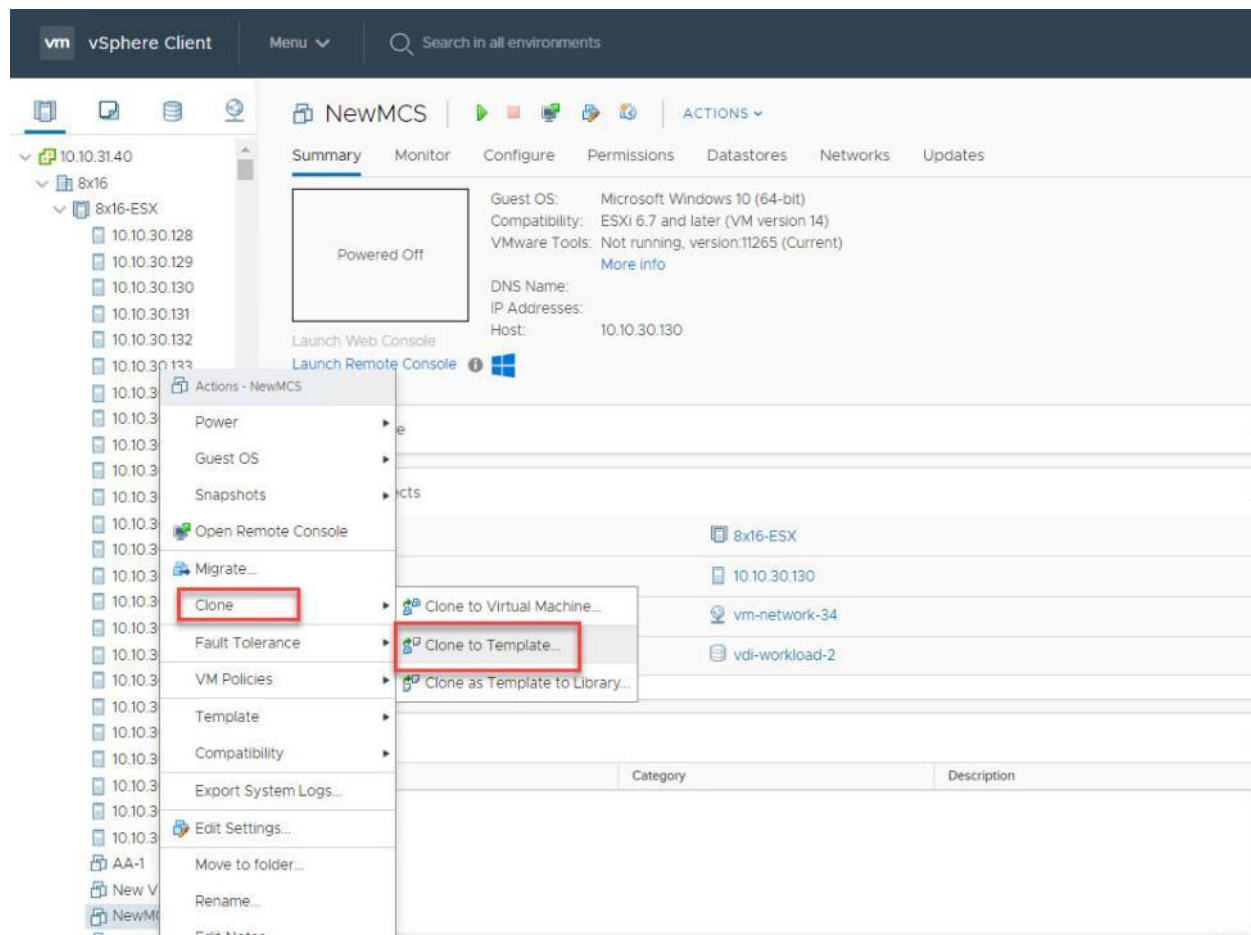
Procedure 13. Provision Virtual Desktop Machines

Step 1. Select the Master Target Device virtual machine from the VCenter Client.

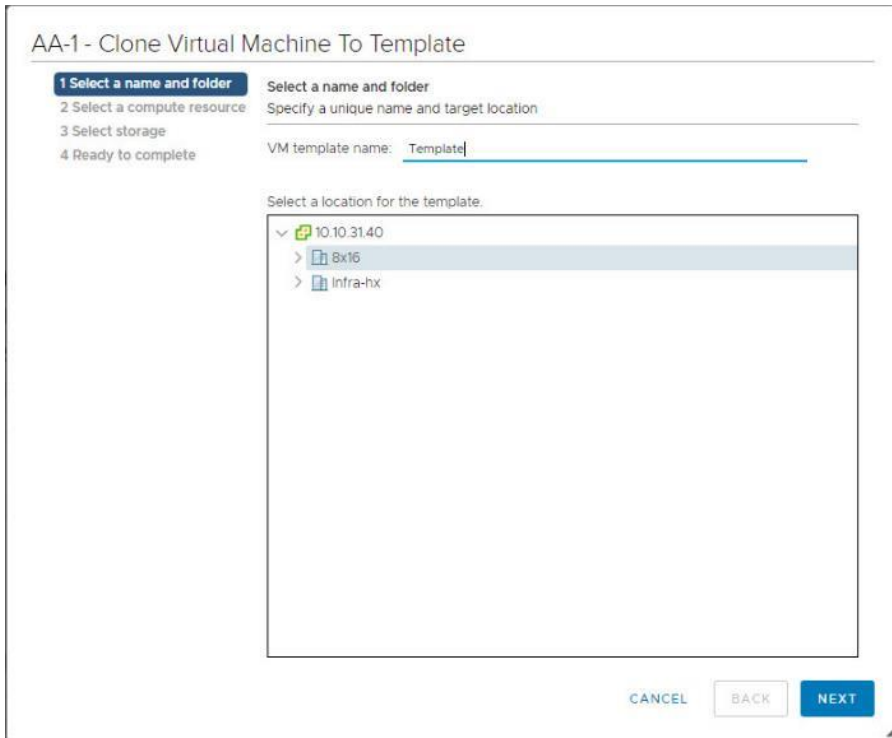
Step 2. Right-click the virtual machine and select ‘Clone > Clone to Template.’

Step 3. Name the cloned ‘Template.’

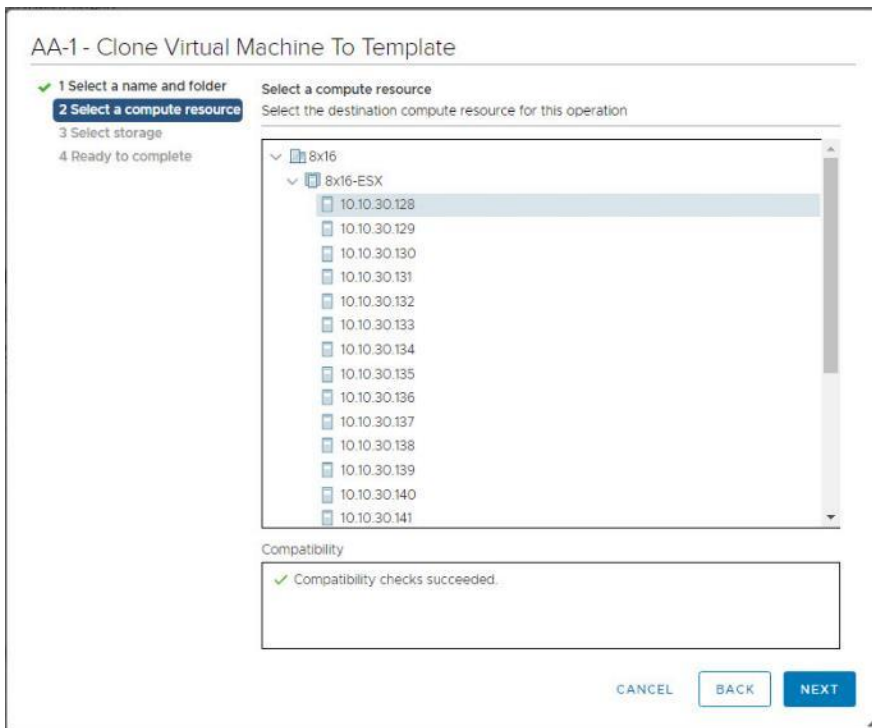
Step 4. Select the cluster and datastore where the first phase of provisioning will occur.



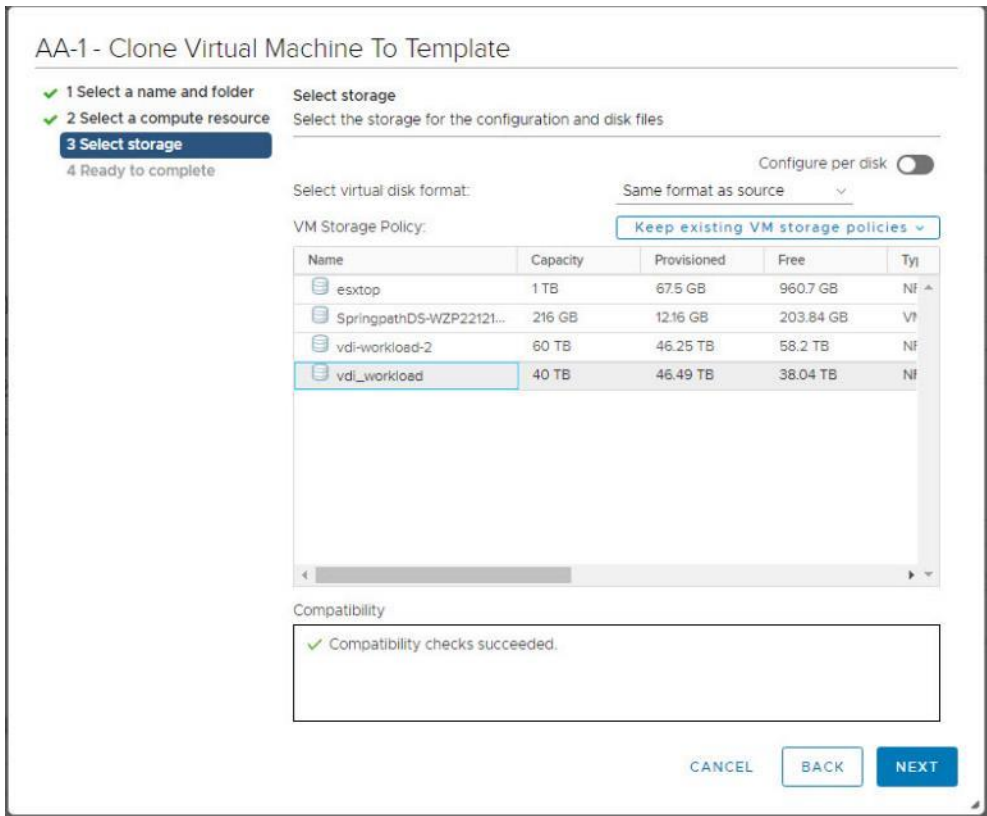
Step 5. Name the template and click Next.



Step 6. Select a host in the cluster to place the template.



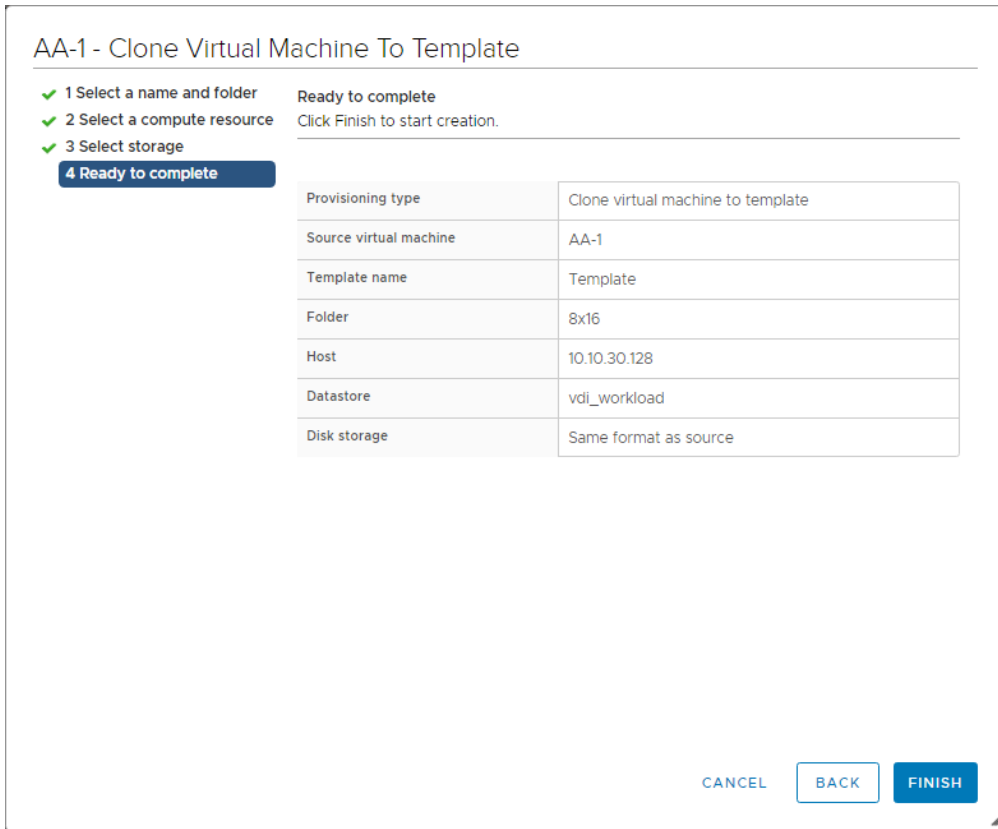
Step 7. Click Next after selecting a datastore.



Step 8. Click Next.

Step 9. Click Next through the remaining screens

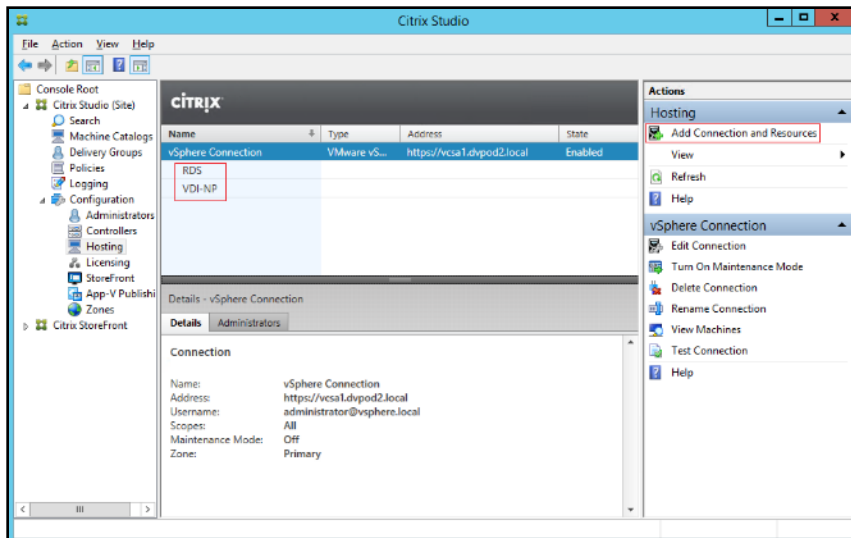
Step 10. Click Finish to create the template.



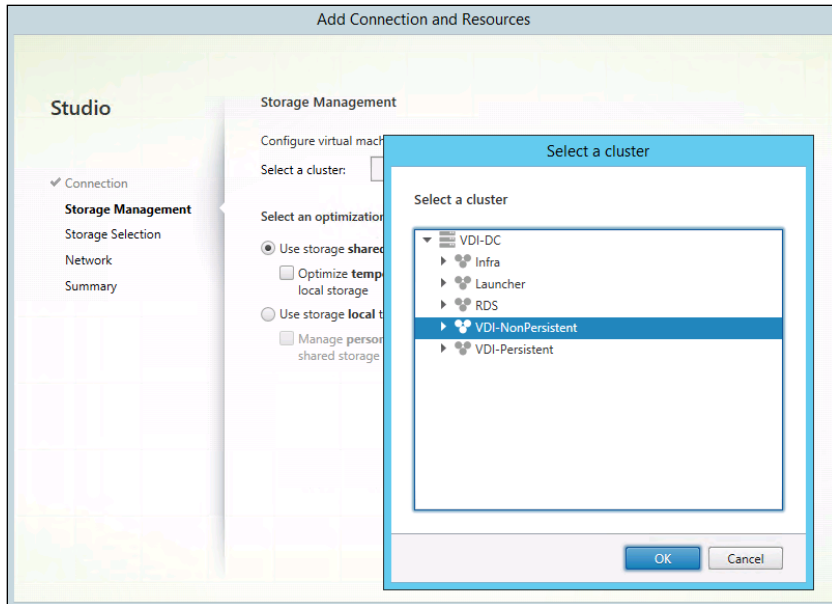
Step 11. From Citrix Studio on the Desktop Controller, select Hosting and Add Connection and Resources.

Step 12. Select Use an existing Connection and click Next.

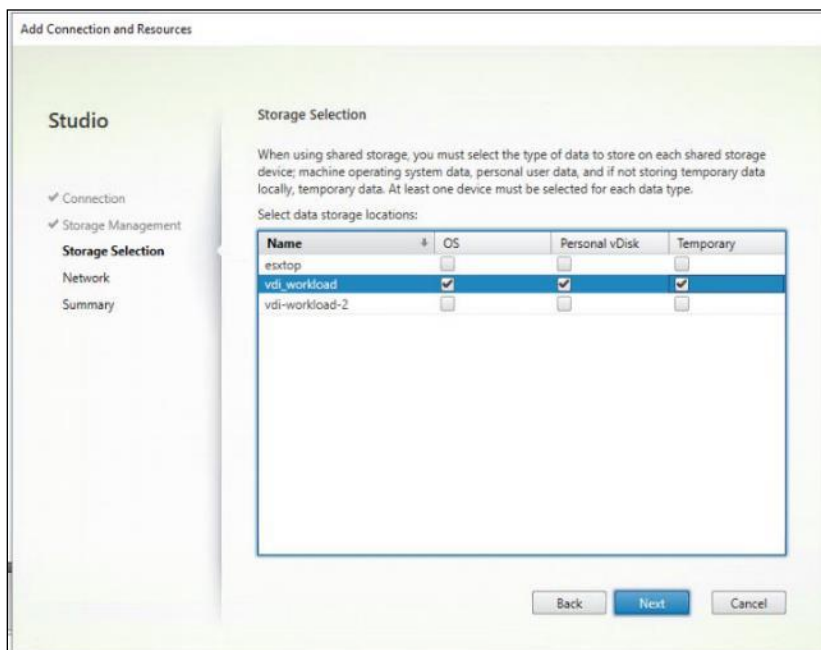
Step 13. Correspond the name of the resource with desktop machine clusters.



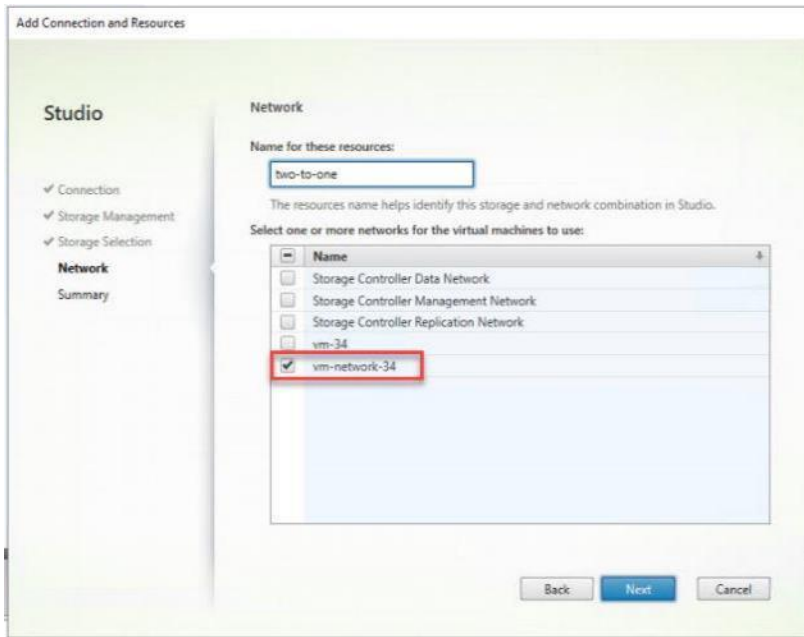
Step 14. Browse and select the VCenter cluster for desktop provisioning and use the default storage method Use storage shared by hypervisors.



Step 15. Select the data storage location for the corresponding resource.



Step 16. Select the VDI networks for the desktop machines and click Next.



Step 17. Click Finish.

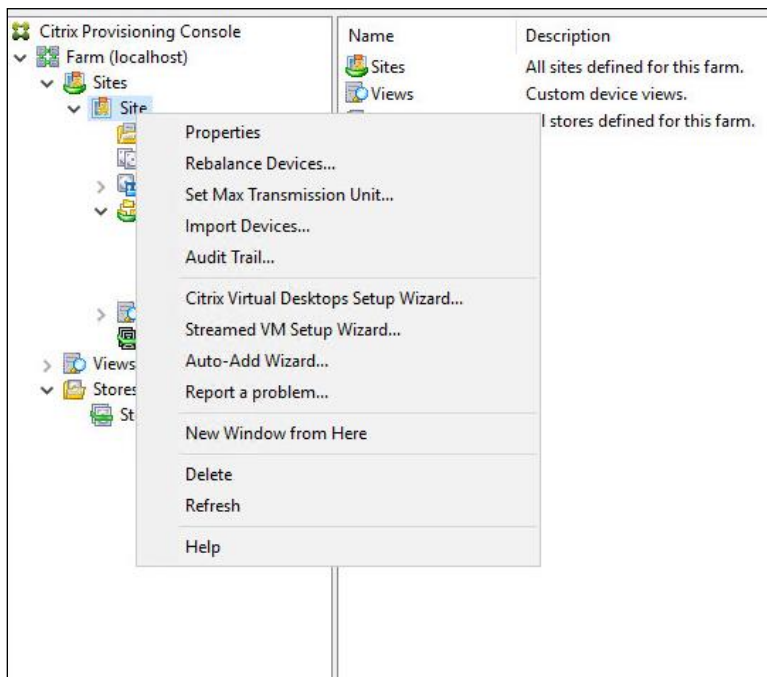
Note: Return to these settings to alter the datastore selection for each set of provisioned desktop machines if you want to create a separate datastore for each image.

Procedure 14. Provision Desktop Machines from Citrix Provisioning Services Console

Step 1. Start the Virtual Desktops Setup Wizard from the Provisioning Services Console.

Step 2. Right-click the Site.

Step 3. Select Virtual Desktops Setup Wizard... from the context menu.



Step 4. Click Next.

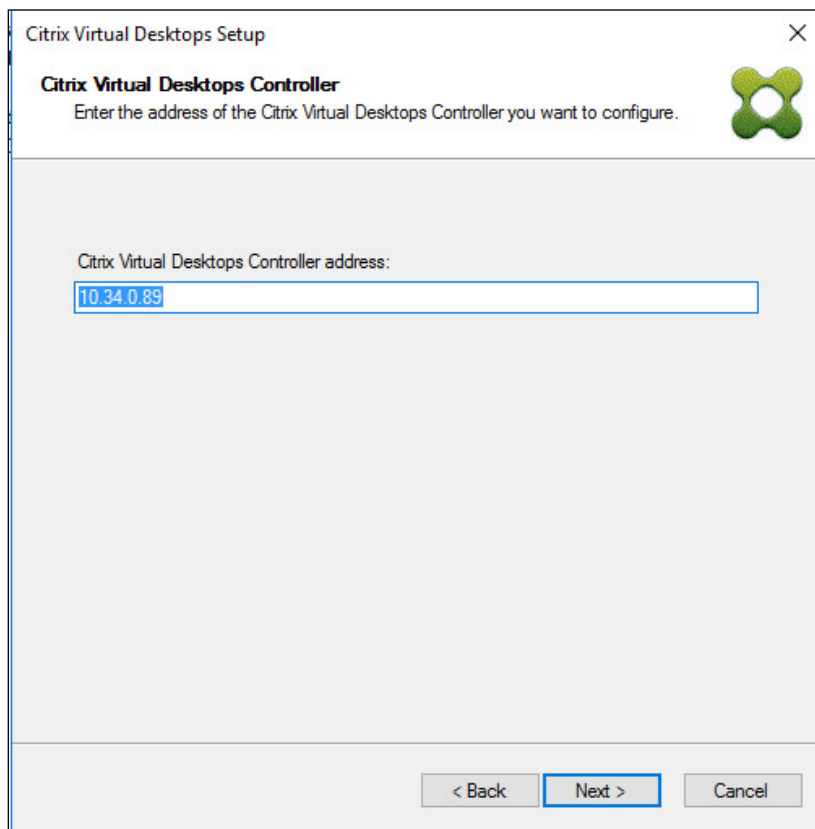
Step 5. Enter the Virtual Desktops Controller address that will be used for the wizard operations.

Step 6. Click Next.



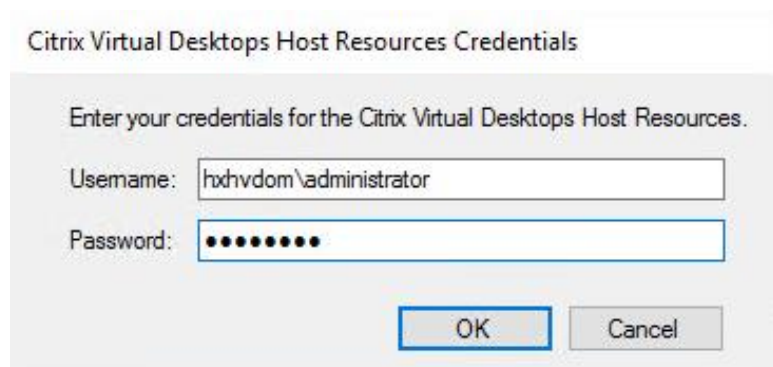
Step 7. Select the Host Resources on which the virtual machines will be created.

Step 8. Click Next.



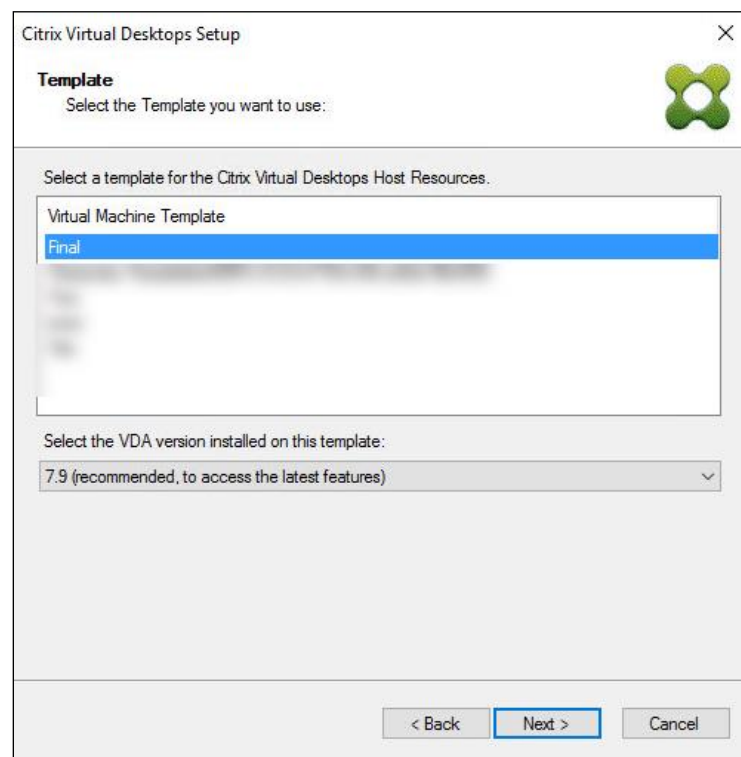
Step 9. Provide the Host Resources Credentials (Username and Password) to the Virtual Desktops controller when prompted.

Step 10. Click OK.



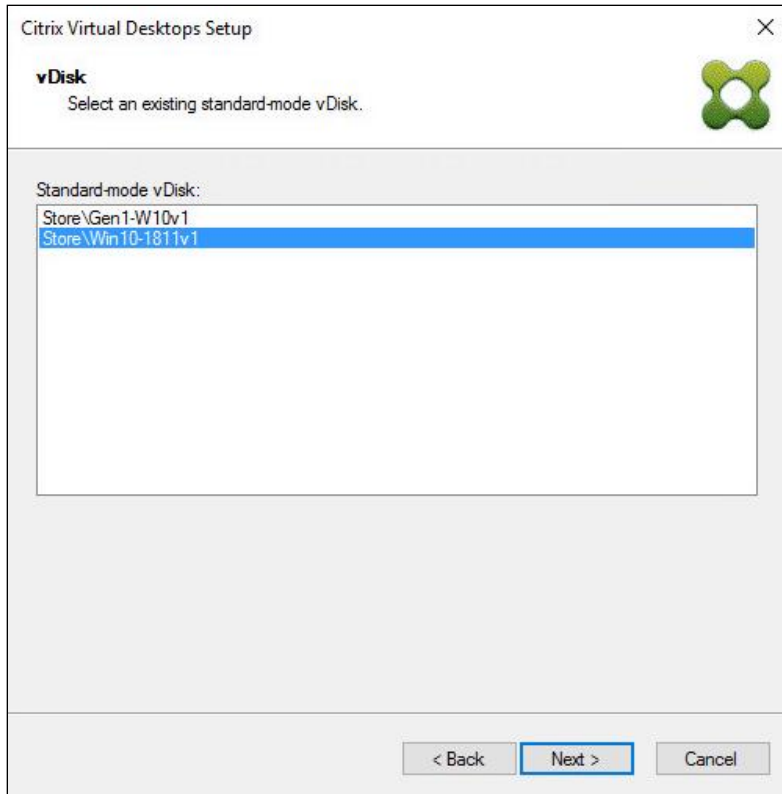
Step 11. Select the Template created earlier.

Step 12. Click Next.



Step 13. Select the vDisk that will be used to stream virtual machines.

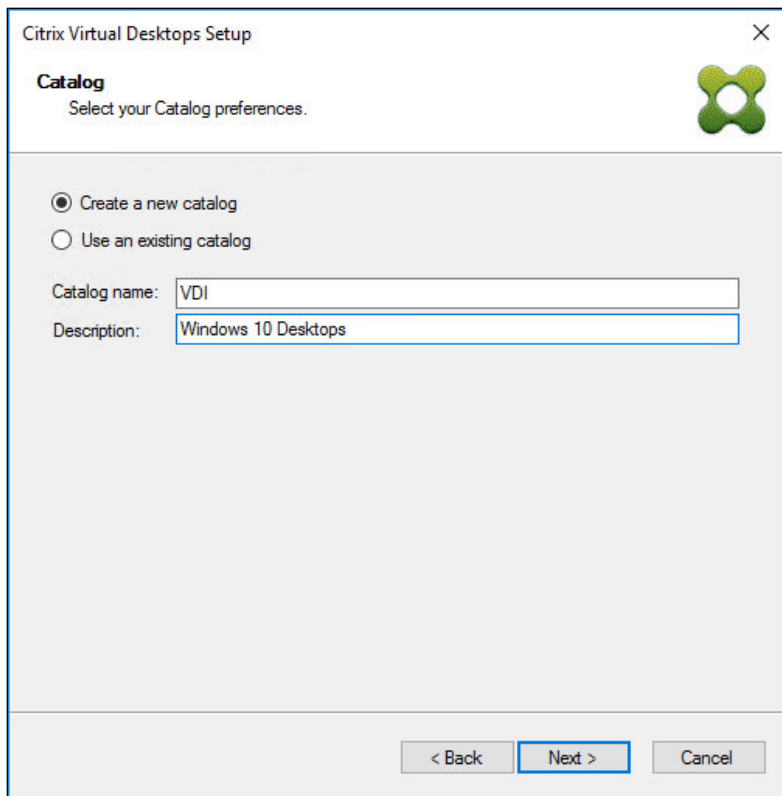
Step 14. Click Next.



Step 15. Select “Create a new catalog.”

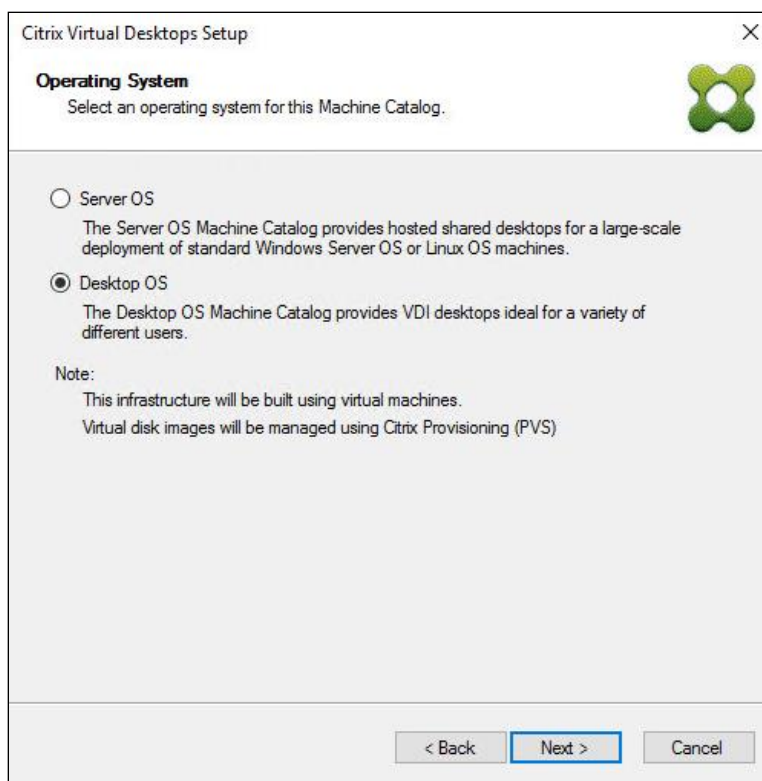
Note: The catalog name is also used as the collection name in the PVS site.

Step 16. Click Next.



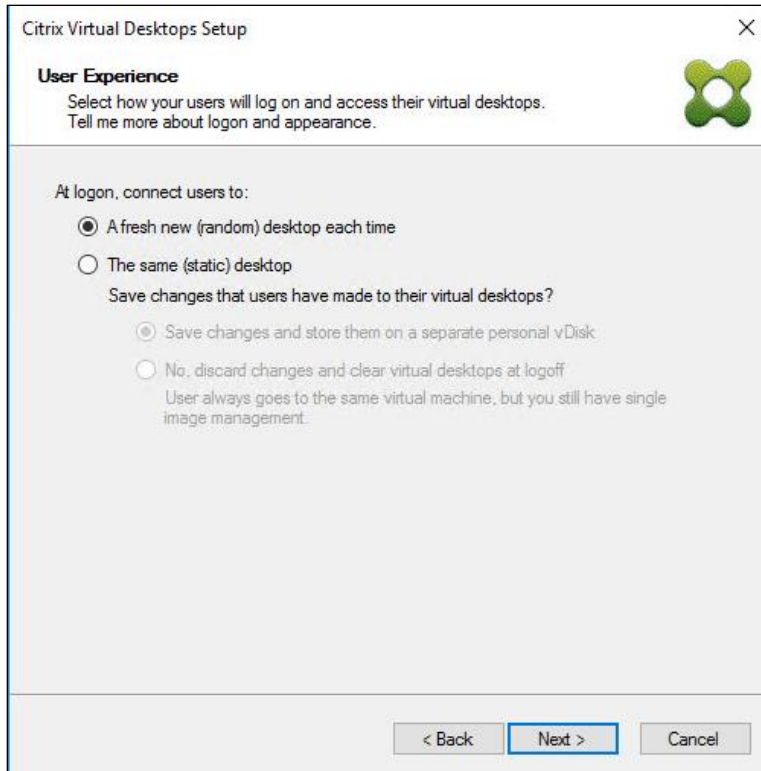
Step 17. On the Operating System dialog, specify the operating system for the catalog. Specify Windows Desktop Operating System for VDI and Windows Server Operating System for RDS.

Step 18. Click Next.



Step 19. If you specified a Windows Desktop OS for VDIs, a User Experience dialog appears. Specify that the user will connect to “A fresh new (random) desktop each time.”

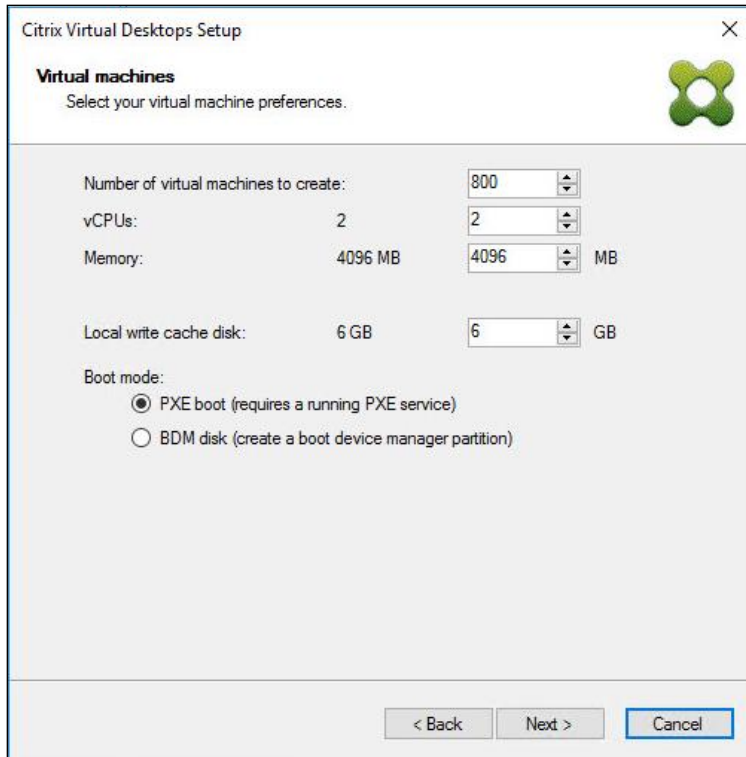
Step 20. Click Next.



Step 21. On the Virtual machines dialog, specify:

- The number of virtual machines to create.
- Number of vCPUs for the virtual machine (2 for VDI, 8 for RDS).
- The amount of memory for the virtual machine (4GB for VDI, 24GB for RDS).
- The write-cache disk size (10GB for VDI, 30GB for RDS).
- PXE boot as the Boot Mode.

Step 22. Click Next.



Citrix Virtual Desktops Setup

Virtual machines
Select your virtual machine preferences.

Number of virtual machines to create: 800

vCPUs: 2

Memory: 4096 MB

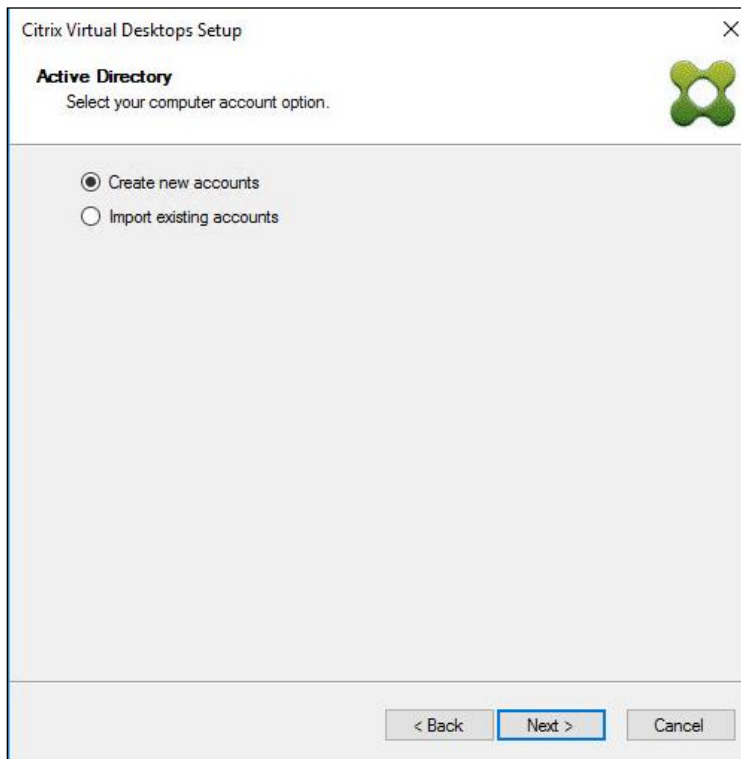
Local write cache disk: 6 GB

Boot mode:
 PXE boot (requires a running PXE service)
 BDM disk (create a boot device manager partition)

< Back Next > Cancel

Step 23. Select the Create new accounts radio button.

Step 24. Click Next.



Citrix Virtual Desktops Setup

Active Directory
Select your computer account option.

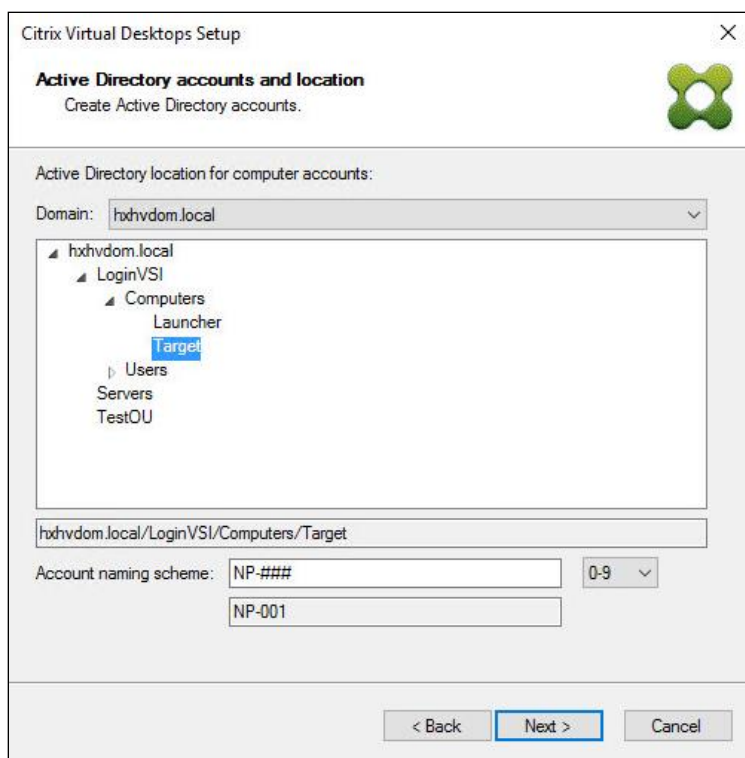
Create new accounts
 Import existing accounts

< Back Next > Cancel

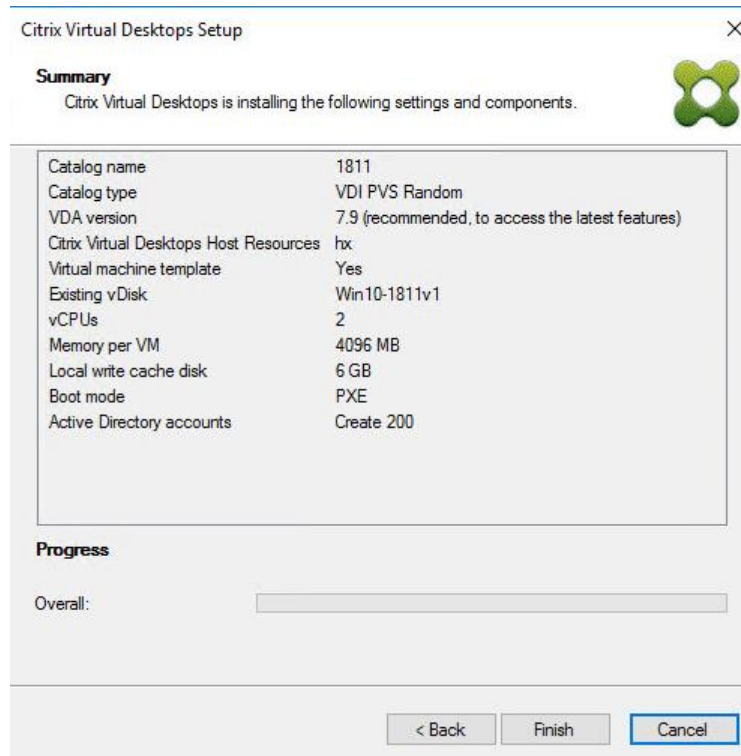
Step 25. Specify the Active Directory Accounts and Location. This is where the wizard should create the computer accounts.

Step 26. Provide the Account naming scheme. An example name is shown in the text box below the name scheme selection location.

Step 27. Click Next.



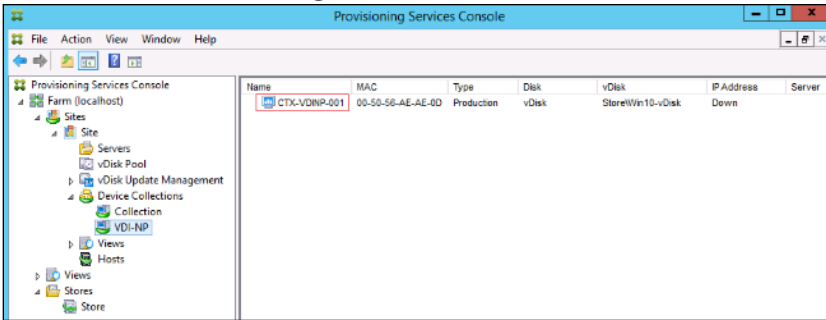
Step 28. Click Finish to begin the virtual machine creation.



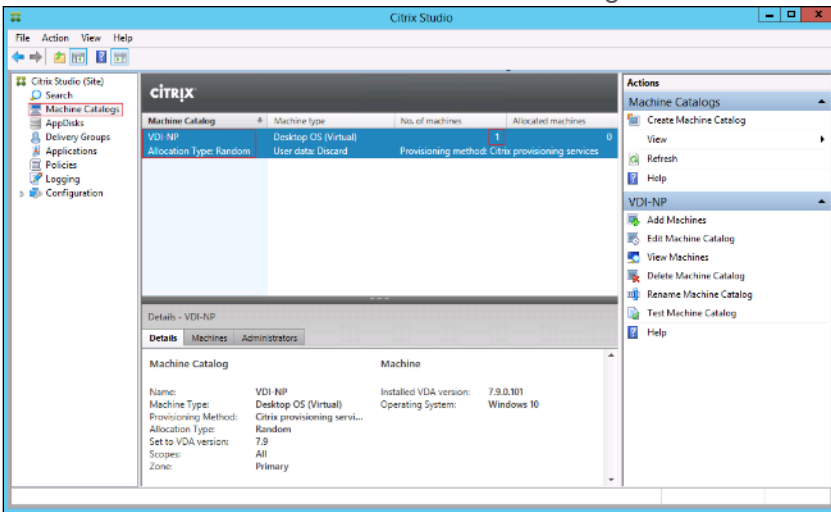
Step 29. When the wizard is done provisioning the virtual machines, click Done.

Step 30. Verify the desktop machines were successfully created in the following locations:

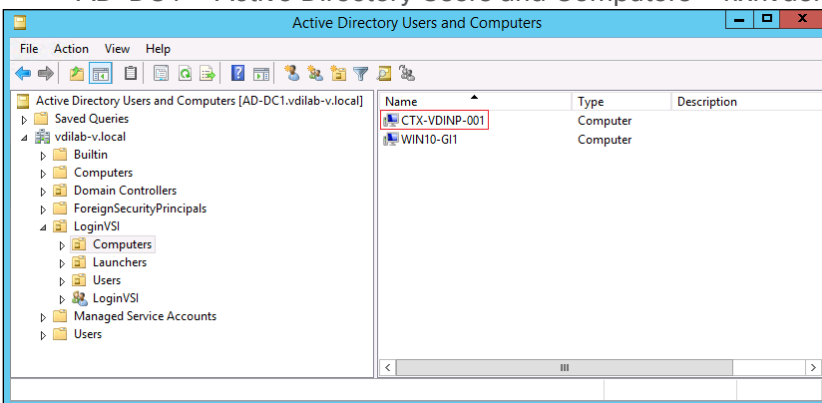
- PVS1 > Provisioning Services Console > Farm > Site > Device Collections > VDI-NP > CTX-VDI-001



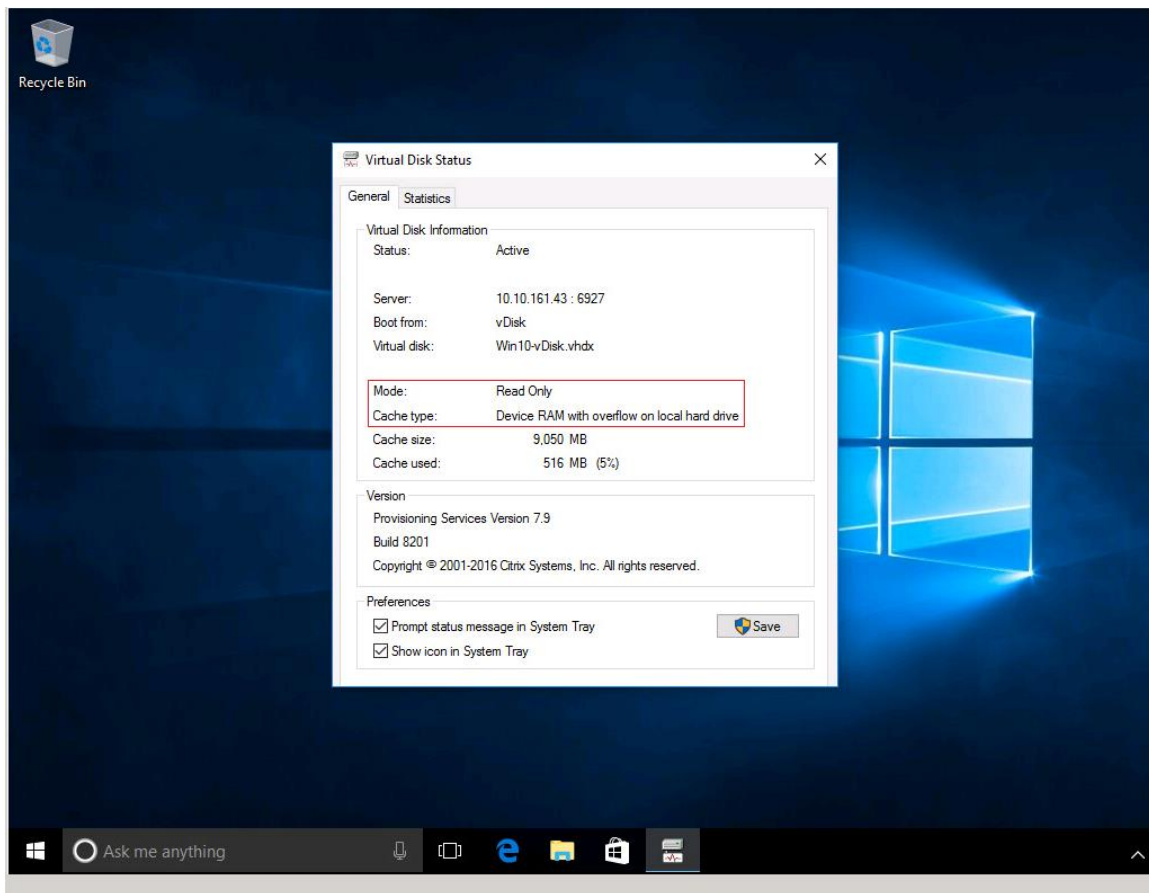
- CTX-XD1 > Citrix Studio > Machine Catalogs > VDI-NP



- AD-DC1 > Active Directory Users and Computers > hxxhvdn.local > Computers > CTX-VDI-001



Step 31. Log into the newly provisioned desktop machine, using the Virtual Disk Status verify the image mode is set to Ready Only and the cache type as Device Ram with overflow on local hard drive.



Procedure 15. Install Citrix Virtual Apps and Desktop Virtual Desktop Agents

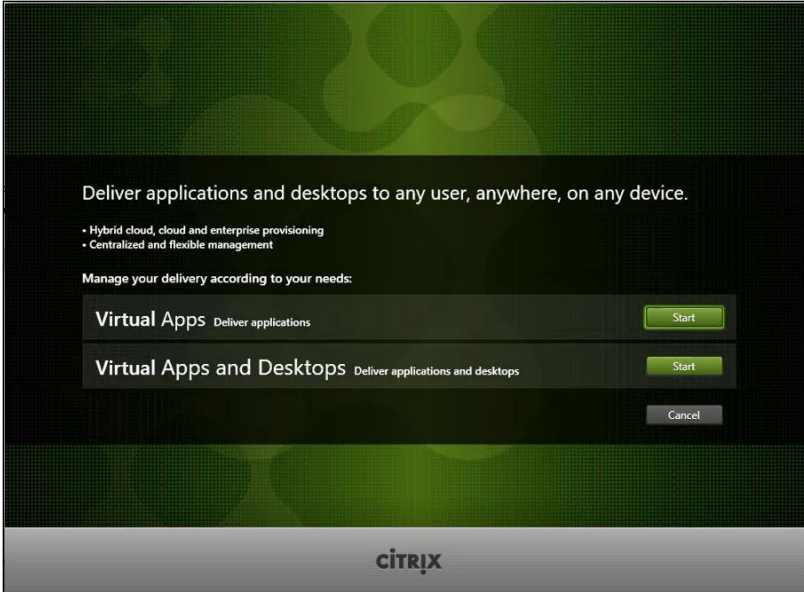
Virtual Delivery Agents (VDAs) are installed on the server and workstation operating systems and enable connections for desktops and apps. The following procedure was used to install VDAs for both HVD and HSD environments.

By default, when you install the Virtual Delivery Agent, Citrix User Profile Management is installed silently on master images.

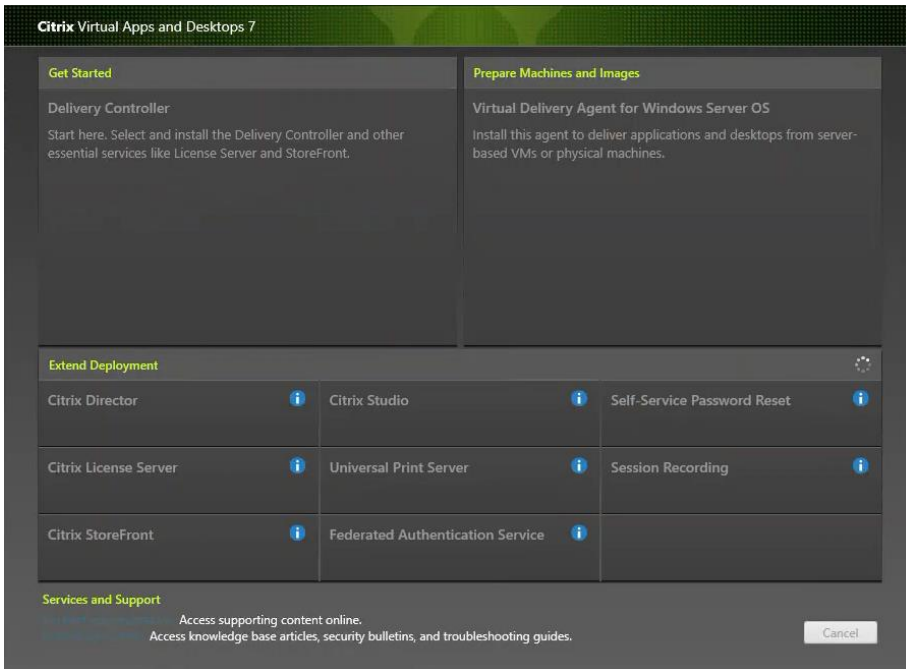
Note: Using profile management as a profile solution is optional but was used for this CVD and is described in a subsequent section.

Step 1. Launch the Citrix Desktop installer from the CVA Desktop 1912 LTSR ISO.

Step 2. Click Start on the Welcome Screen.

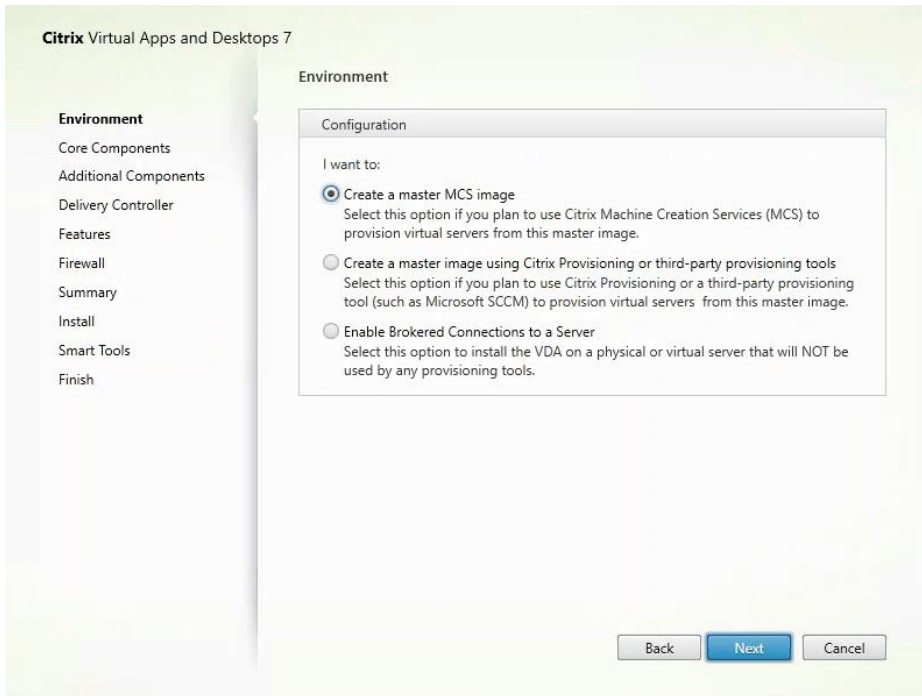


Step 3. To install the VDA for the Hosted Virtual Desktops (VDI), select Virtual Delivery Agent for Windows Desktop OS. After the VDA is installed for Hosted Virtual Desktops, repeat the procedure to install the VDA for Hosted Shared Desktops (RDS). In this case, select Virtual Delivery Agent for Windows Server OS and follow the same basic steps.



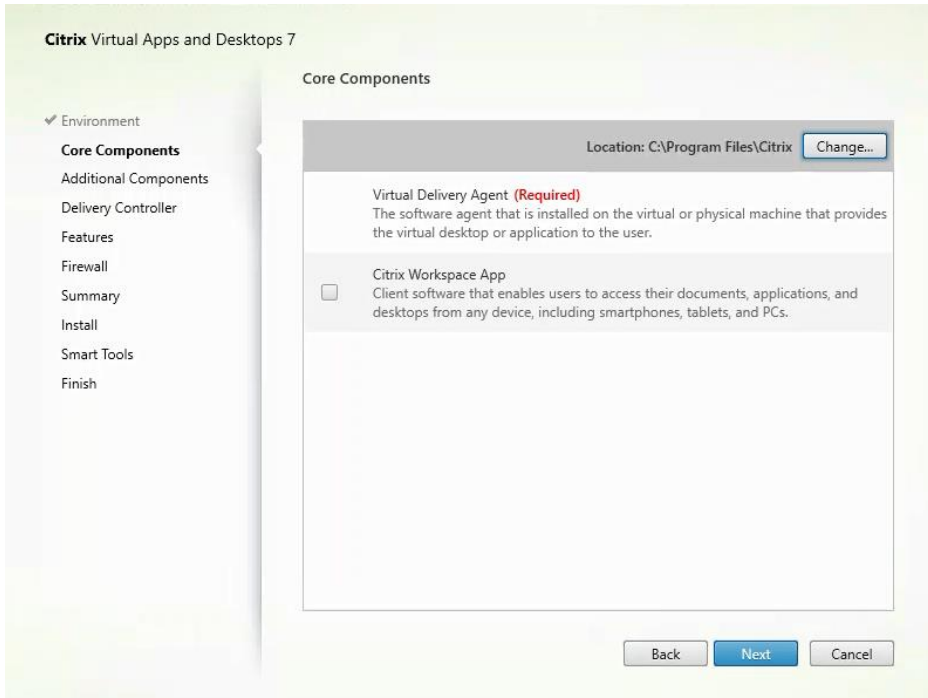
Step 4. Select Create a Master Image (be sure to select the proper provisioning technology).

Step 5. Click Next.

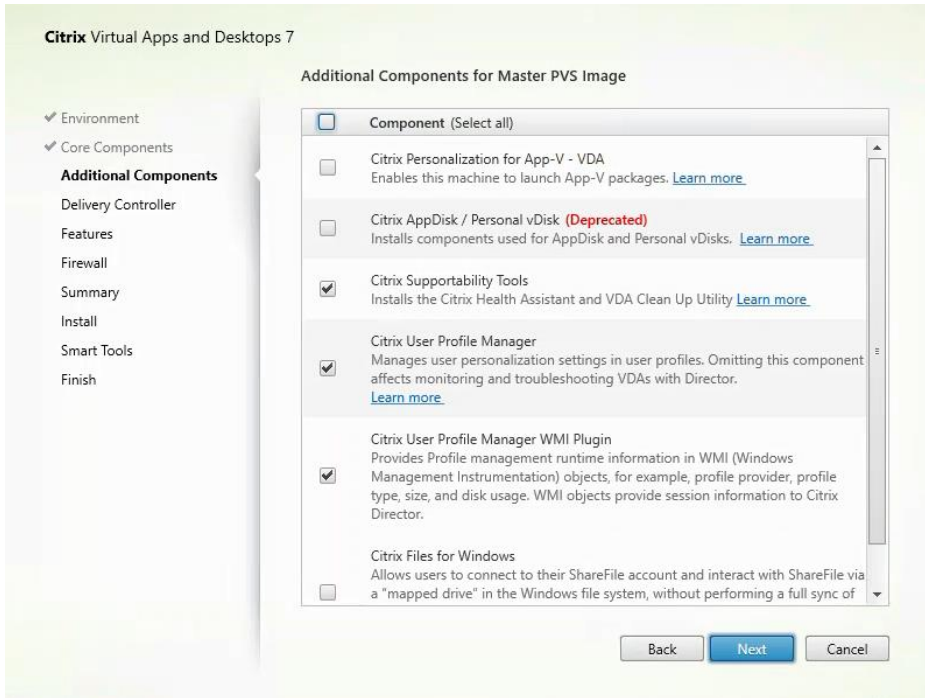


Step 6. Optional: Select Citrix Workspace App.

Step 7. Click Next.

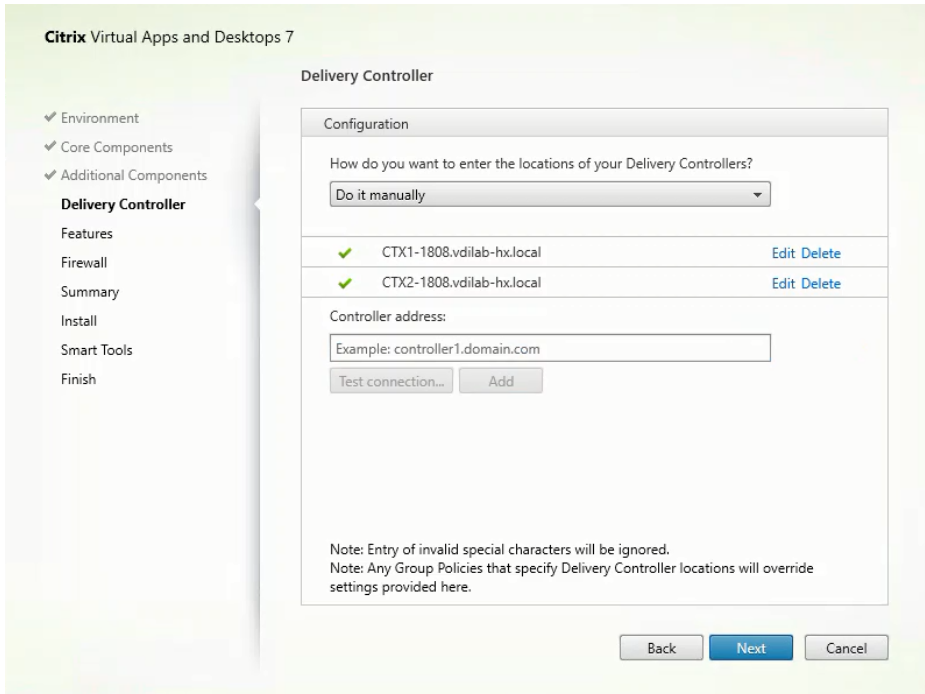


Step 8. Click Next.



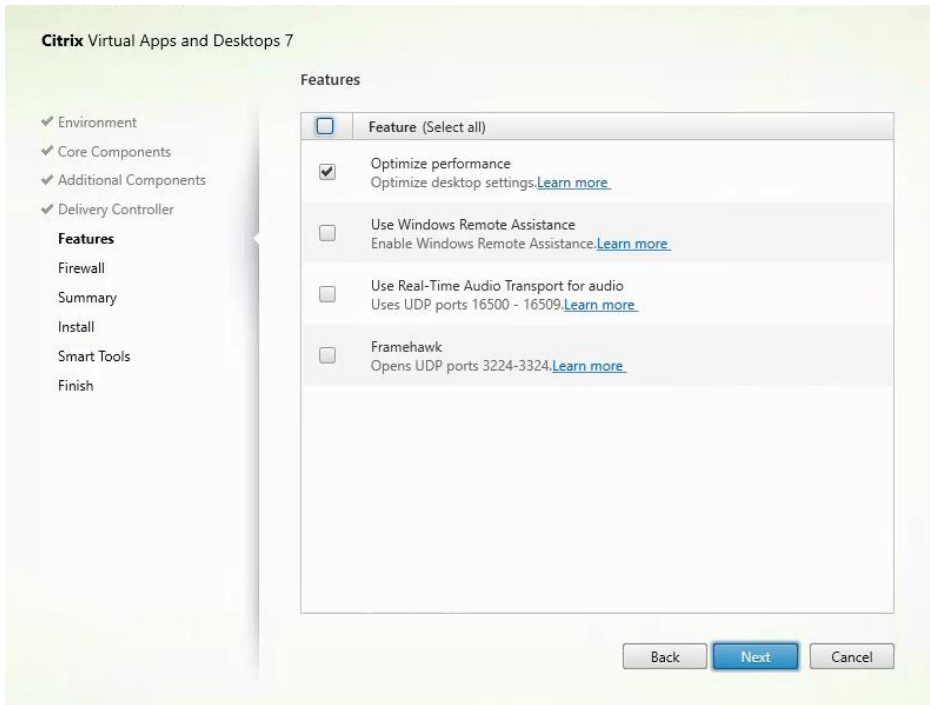
Step 9. Select “Do it manually” and specify the FQDN of the Delivery Controllers.

Step 10. Click Next.



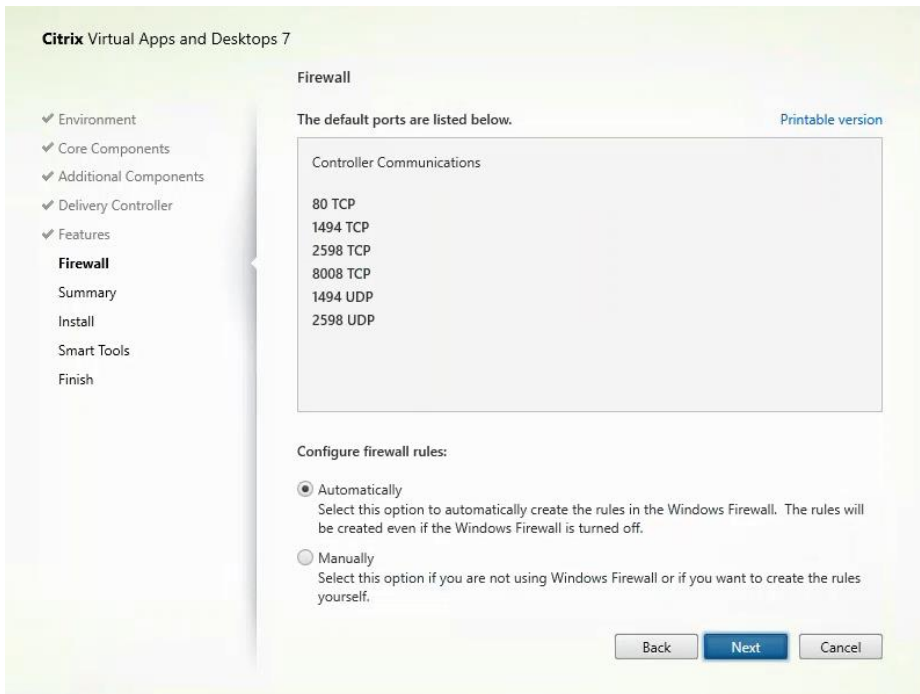
Step 11. Accept the default features.

Step 12. Click Next.

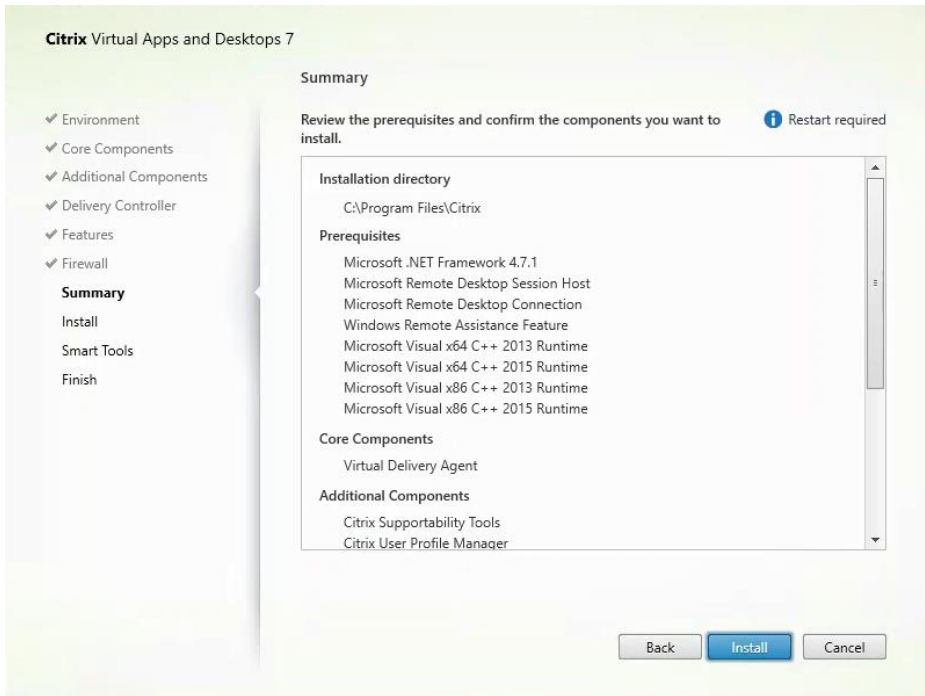


Step 13. Allow the firewall rules to be configured automatically.

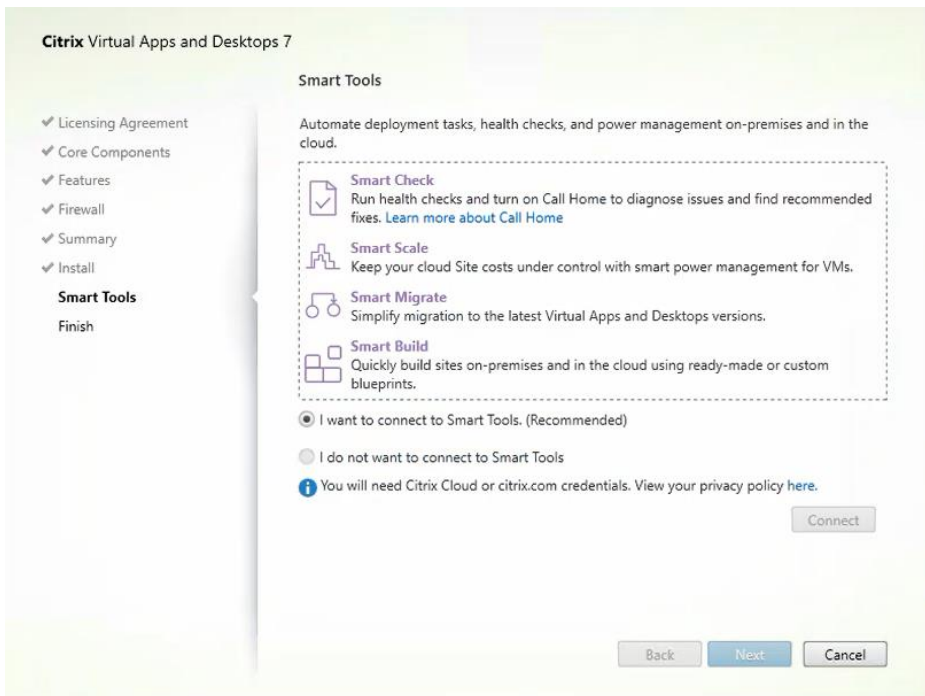
Step 14. Click Next.



Step 15. Verify the Summary and click Install.



Step 16. (Optional) Select Call Home participation.

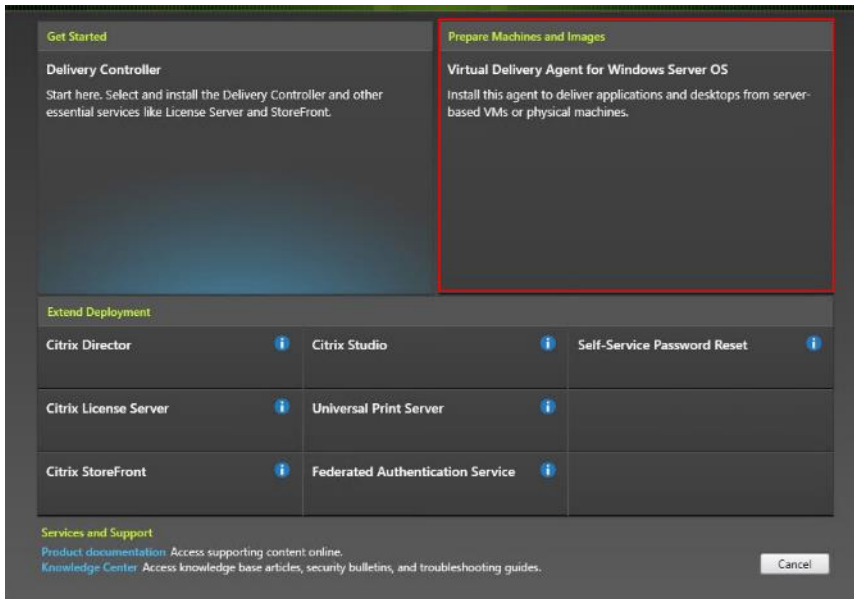


Step 17. (Optional) check “Restart Machine.”

Step 18. Click Finish.

Step 19. Repeat this procedure so that VDAs are installed for both HVD (using the Windows 10 OS image) and the HSD desktops (using the Windows Server 2019 image).

Step 20. Select an appropriate workflow for the HSD desktop.



Create Delivery Groups

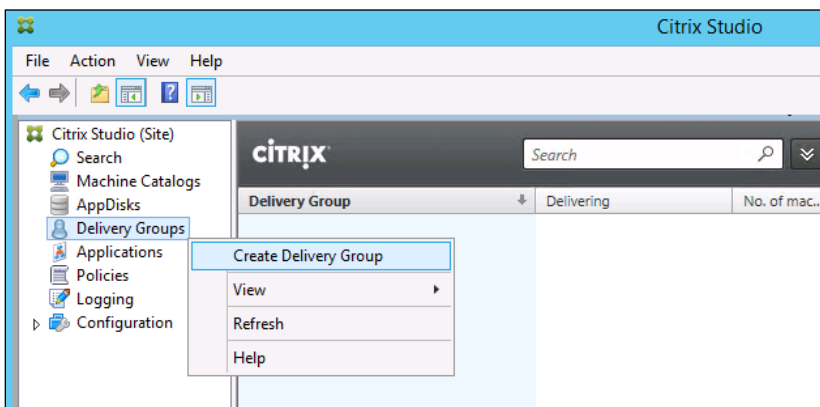
Delivery Groups are collections of machines that control access to desktops and applications. With Delivery Groups, you can specify which users and groups can access which desktops and applications.

Procedure 1. Create Delivery Groups

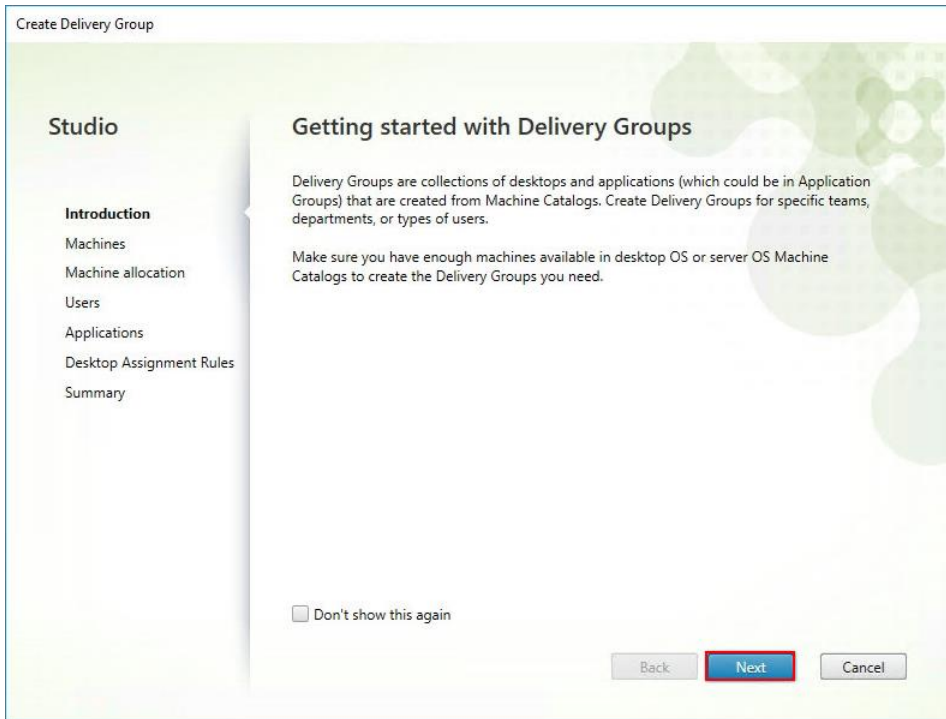
Note: The instructions below outline the steps to create a Delivery Group for VDI desktops. When you have completed these steps, repeat the procedure to a Delivery Group for HVD desktops.

Step 1. Connect to a Virtual Desktops server and launch Citrix Studio.

Step 2. Select Create Delivery Group from the drop-down list.



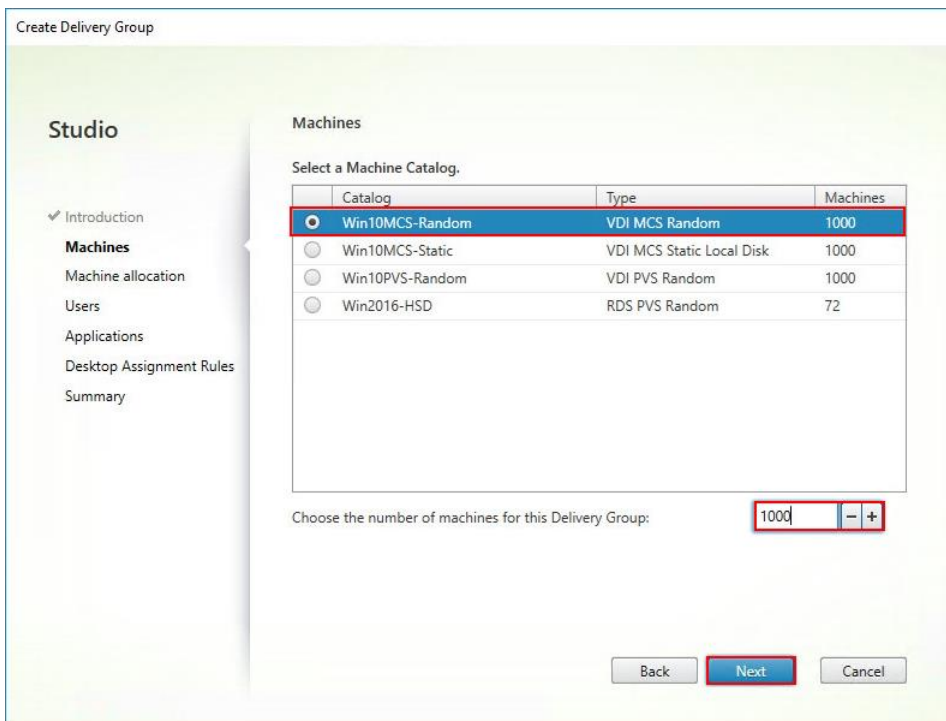
Step 3. Click Next.



Step 4. Select Machine catalog.

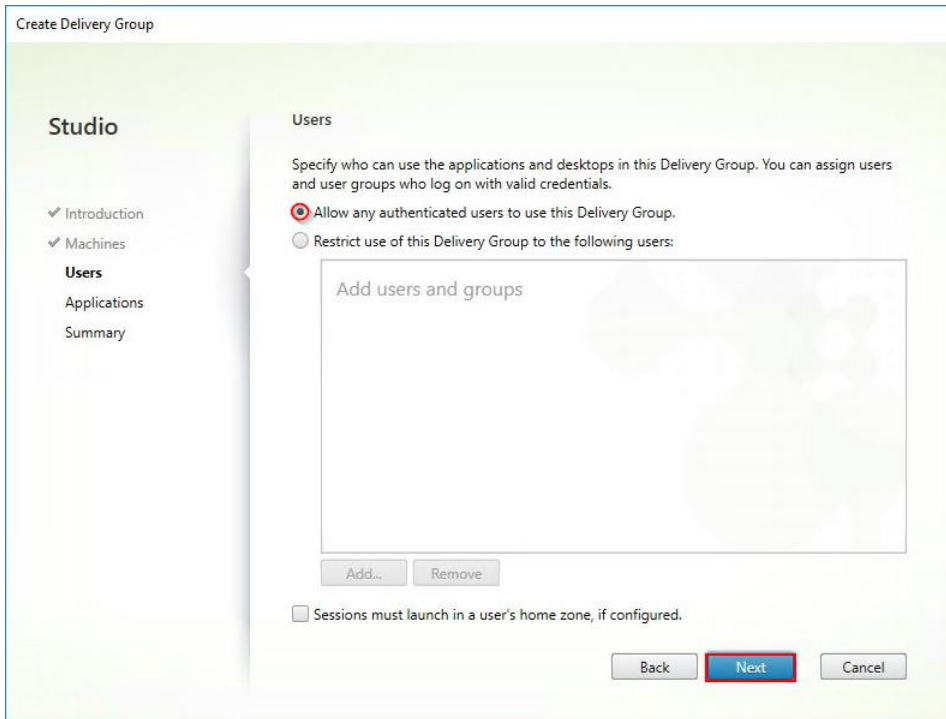
Step 5. Provide the number of machines to be added to the delivery Group.

Step 6. Click Next.



Step 7. To make the Delivery Group accessible, you must add users, select Allow any authenticated users to use this Delivery Group.

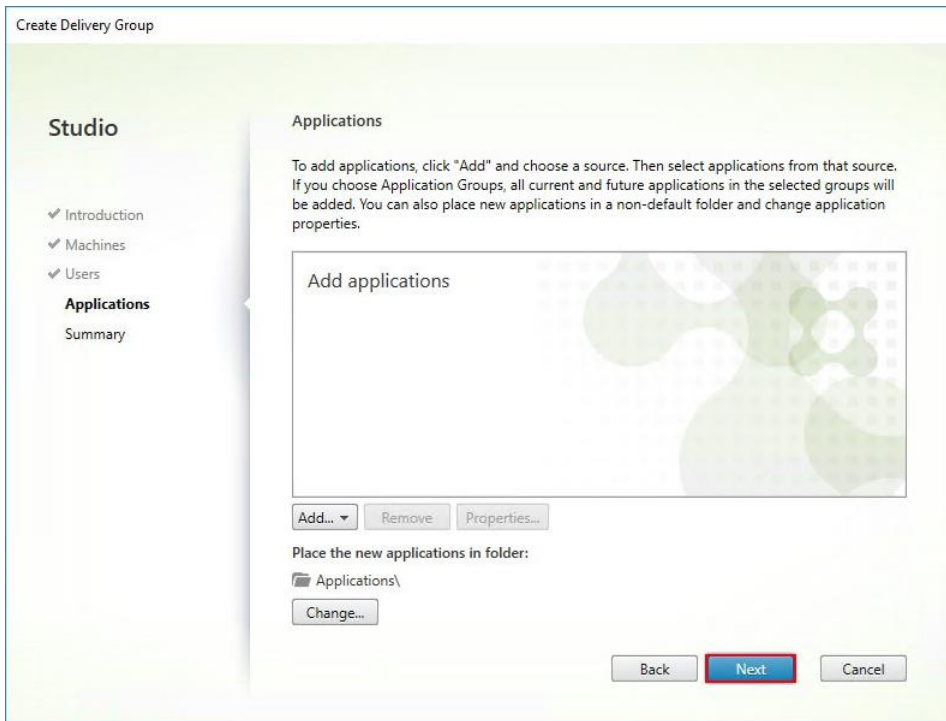
Step 8. Click Next.



Note: User assignment can be updated any time after Delivery group creation by accessing Delivery group properties in Desktop Studio.

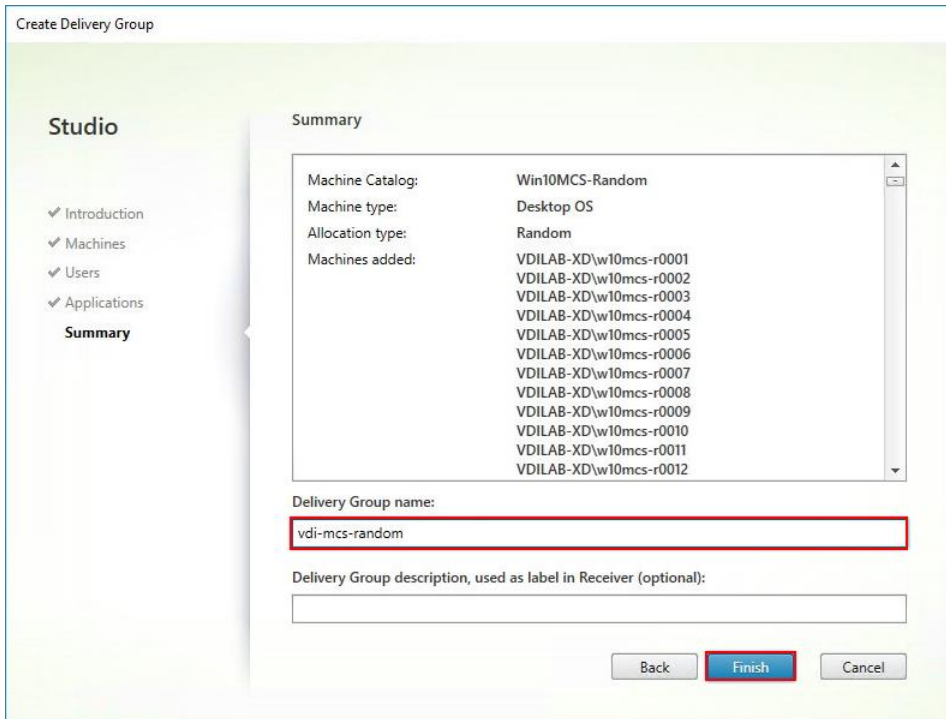
Step 9. (Optional) specify Applications catalog will deliver.

Step 10. Click Next.

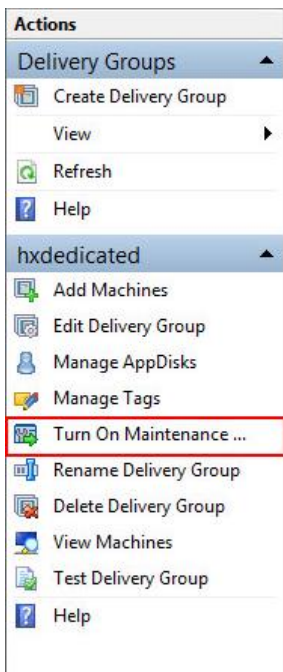


Step 11. On the Summary dialog, review the configuration. Enter a Delivery Group name and a Display name (for example, HVD or HSD).

Step 12. Click Finish.



Step 13. Citrix Studio lists the created Delivery Groups and the type, number of machines created, sessions, and applications for each group in the Delivery Groups tab. Select Delivery Group and in Action List, select “Turn on Maintenance Mode.”



Citrix Virtual Desktops Policies and Profile Management

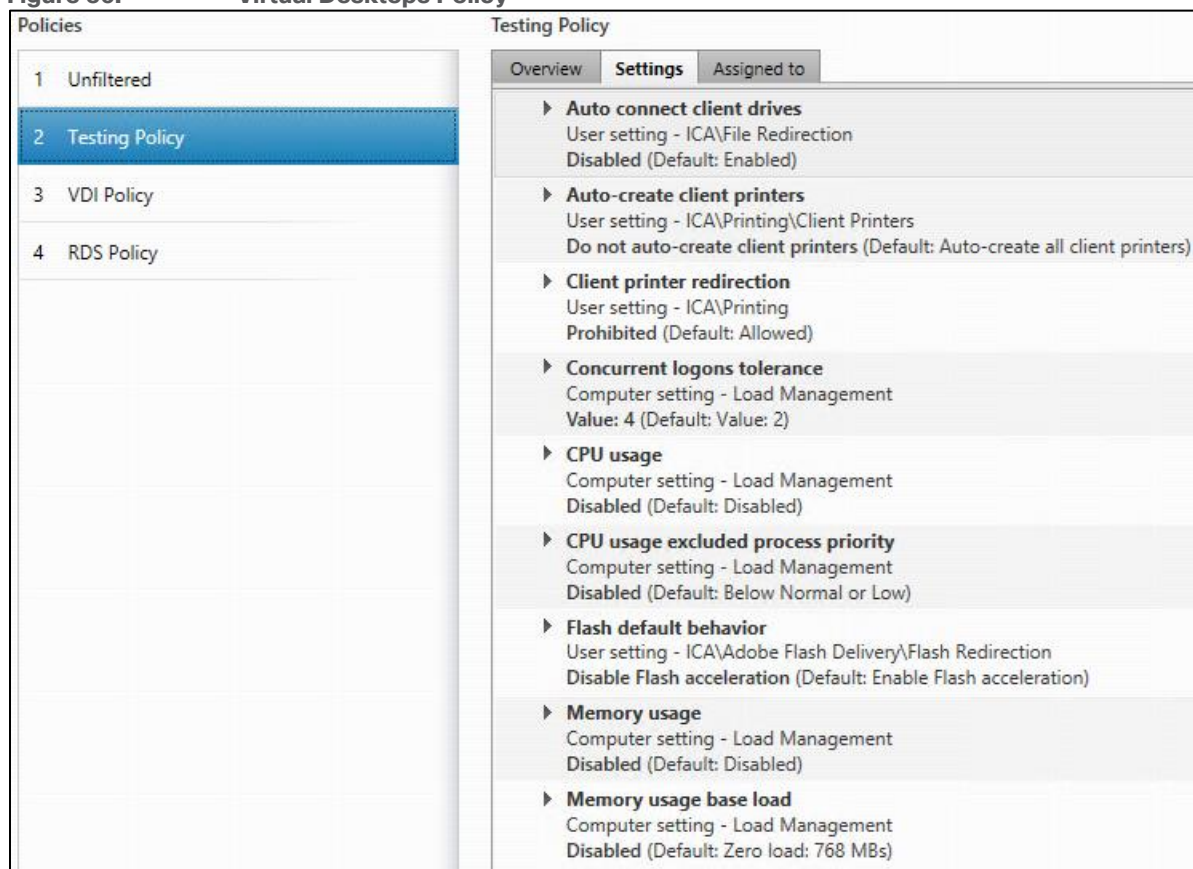
Policies and profiles allow the Citrix Virtual Desktops environment to be easily and efficiently customized.

Configure Citrix Virtual Desktops Policies

Citrix Virtual Desktops policies control user access and session environments and are the most efficient method of controlling connection, security, and bandwidth settings. You can create policies for specific groups of users,

devices, or connection types with each policy. Policies can contain multiple settings and are typically defined through Citrix Studio. (The Windows Group Policy Management Console can also be used if the network environment includes Microsoft Active Directory and permissions are set for managing Group Policy Objects). [Figure 36](#) shows policies for Login VSI testing in this CVD.

Figure 36. Virtual Desktops Policy



Configure User Profile Management

Profile management provides an easy, reliable, and high-performance way to manage user personalization settings in virtualized or physical Windows environments. It requires minimal infrastructure and administration and provides users with fast logons and logoffs. A Windows user profile is a collection of folders, files, registry settings, and configuration settings that define the environment for a user who logs on with a particular user account. These settings may be customizable by the user, depending on the administrative configuration. Examples of settings that can be customized are:

- Desktop settings such as wallpaper and screen saver
- Shortcuts and Start menu setting
- Internet Explorer Favorites and Home Page
- Microsoft Outlook signature
- Printers

Some user settings and data can be redirected by means of folder redirection. However, if folder redirection is not used these settings are stored within the user profile.

The first stage in planning a profile management deployment is to decide on a set of policy settings that together form a suitable configuration for your environment and users. The automatic configuration feature simplifies some of this decision-making for Virtual Desktops deployments. Screenshots of the User Profile Management interfaces that establish policies for this CVD's RDS and VDI users (for testing purposes) are shown below.

Basic profile management policy settings are documented here: <https://docs.citrix.com/en-us/citrix-virtual-apps-desktops>

Figure 37. VDI User Profile Manager Policy

Policies		VDI Policy	
		Overview	Settings
1	Unfiltered		
2	Testing Policy		
3	VDI Policy		<ul style="list-style-type: none"> ▶ Active write back Computer setting - Profile Management\Basic settings Enabled (Default: Disabled) ▶ Delete locally cached profiles on logoff Computer setting - Profile Management\Profile handling Enabled (Default: Disabled) ▶ Enable Profile management Computer setting - Profile Management\Basic settings Enabled (Default: Disabled) ▶ Exclusion list - directories Computer setting - Profile Management\File system\Exclusions AppData\Local;AppData\LocalLow;AppData\Roaming;\$Recycle.Bin (Default:) ▶ Path to user store Computer setting - Profile Management\Basic settings \\10.10.62.92\Profile-VDI01\$\#SAMAccountName# (Default: Windows) ▶ Process logons of local administrators Computer setting - Profile Management\Basic settings Enabled (Default: Disabled)
4	RDS Policy		

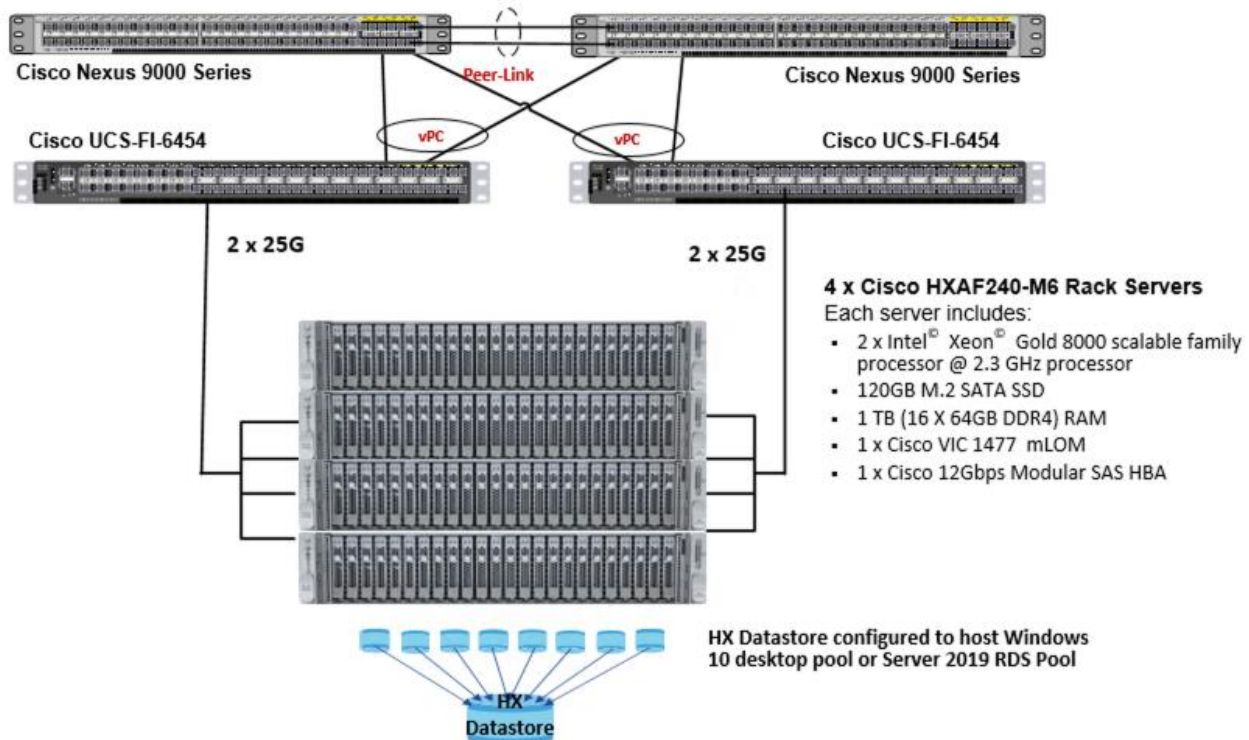
Test Setup and Configurations

This chapter is organized into the following subjects:

- [Test Methodology and Success Criteria](#)
- [Test Procedure](#)

In this project, we tested a 4 node Cisco HyperFlex cluster running four Cisco UCS HXAF240-M6 Rack Servers. This solution is tested to illustrate linear scalability for each workload studied. We tested performance and failover for VDI on the cluster.

Cisco HyperFlex VDI, Reference Architecture



Hardware Components:

- 2 x Cisco UCS 6454 Fabric Interconnects
- 2 x Cisco Nexus 93108YCPX Access Switches
- 4 x Cisco UCS HXAF240-M6 Rack Servers (2 Intel Xeon Gold 8000 scalable family processor at 2.3 GHz, with 1 TB of memory per server [64 GB x 16 DIMMs at 3200 MHz])
- Cisco VIC 1477 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)
- 8 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 5.0
- Citrix Virtual Desktops 1912 LTSR
- Citrix User Profile Management
- Microsoft SQL Server 2016
- Microsoft Windows 10
- Microsoft Windows 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.

Along with regular performance testing for VDI workloads, we also tested disaster recovery and failover functionality of a stretched cluster environment.

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

Windows 10 virtual machines for VDI and Windows 2019 Server for RDS were deployed on both sites of the stretch cluster. Fifty percent of the total number on each site. All machines were shut down utilizing the Citrix Virtual Desktops 1912 LTSR 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.

Procedure 1. Test Run Protocol

To simulate severe, real-world environments, Cisco requires the logon 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 500 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.

Step 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, Citrix Connection brokers, image mgmt., and so on)

Step 2. Time 0:00:10 Start Storage Partner Performance Logging on Storage System.

Step 3. Time 0:05: Boot VDI Machines using Citrix Virtual Desktops 1912 LTSR Administrator Console.

Step 4. Time 0:06 First machines boot.

Step 5. Time 0:35 Single Server or Scale target number of VDI Servers registered on Desktop Studio.

Note: No more than 60 Minutes of rest time is allowed after the last desktop is registered and available on Citrix Virtual Desktops 1912 LTSR Administrator Console dashboard. Typically, a 20-30 minute rest period for Windows 10 desktops and 10 minutes for RDS virtual machines is sufficient.

Step 6. Time 1:35 Start Login VSI 4.1.40 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).

Step 7. Time 2:23 Single Server or Scale target number of desktop virtual machines desktops launched (48-minute benchmark launch rate).

Step 8. Time 2:25 All launched sessions must become active.

Note: All sessions launched must become active for a valid test run within this window.

Step 9. Time 2:40 Login VSI Test Ends (based on Auto Logoff 900 Second period designated above).

Step 10. Time 2:55 All active sessions logged off.

Note: All sessions launched and active must be logged off for a valid test run. The Citrix Virtual Desktops 1912 LTSR Administrator Dashboard must show that all desktops have been returned to the registered/available state as evidence of this condition being met.

Step 11. Time 2:57 All logging terminated; Test complete.

Step 12. Time 3:15 Copy all log files off to archive; Set virtual desktops to maintenance mode through broker; Shutdown all Windows 7 machines.

Step 13. Time 3:30 Reboot all hypervisors.

Step 14. 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 Citrix 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.)

Note: We will publish Cisco Validated Designs with our recommended workload following the process above 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 Citrix Virtual Desktops 1912 LTSR Hosted Shared Desktop with Citrix Virtual Desktops 1912 LTSR Composer provisioning using Microsoft Windows Server 2016 sessions on Cisco UCS HXAF220c-M4S, Cisco UCS 220 M4 and Cisco UCS B200 M4 servers.

The information contained 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 Citrix 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.

All of these standard Login VSI CVD Testing results for VDI will be evaluated against each failure scenario in the Stretch Cluster. Each test should pass in each failure scenario for this to be considered a successful test.

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:

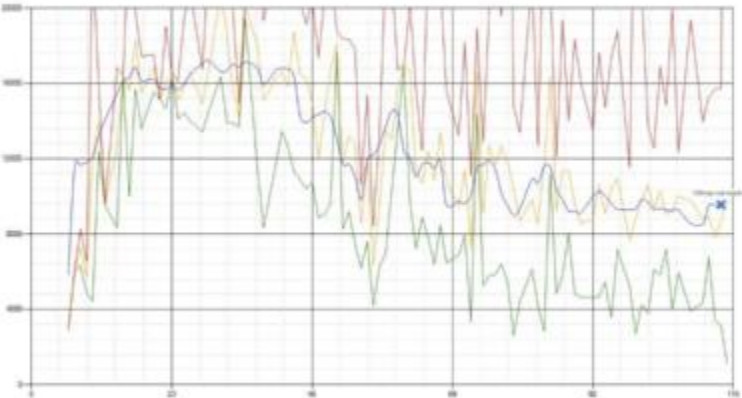
- Notepad File Open (NFO)
Loading and initiating VSINotepad.exe and opening the openfile dialog. This operation is handled by the OS and by the VSINotepad.exe itself through execution. This operation seems almost instant from an end-user’s point of view.
- Notepad Start Load (NSLD)
Loading and initiating 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 38. Sample of a VSI Max Response Time Graph, Representing a Normal Test



Figure 39. 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 VSImax + 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 an 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 VSI_{max} v4.1.x, and the higher VSI_{max} is, the larger overall user capacity can be expected.

With Login VSI 4.1.x a new VSI_{max} method is introduced: VSI_{max} v4.1. This methodology gives much better insight in system performance and scales to extremely large systems.

Test Results

This chapter is organized into the following subjects:

- [Boot Storms](#)
- [Recommended Maximum Workload and Configuration Guidelines](#)
- [ESXi Host Performance Counters](#)

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 750-1000 desktops and measure the time it takes for the 1000th virtual machine to register as available in the Virtual Desktops Administrator console.

The Cisco HyperFlex HXAF240c-M6 based All-Flash cluster running Data Platform version 5.0 software can accomplish this task in 35 minutes.

Recommended Maximum Workload and Configuration Guidelines

This subject details the recommended workloads and configuration guidelines.

Four Node Cisco HXAF240c-M6 Rack Server and HyperFlex All-Flash Cluster

For Citrix Virtual Apps RDS Hosted Shared Desktop and Hosted Virtual Desktop use case, the recommended maximum workload was determined based on both Login VSI Knowledge Worker workload end user experience measures and HXAF240c-M6 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.

Note: 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 workload over a 48 minutes duration.
Steady state	The steady state phase is where all users are logged in and performing various workload 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.

Four Node Cisco HXAF240c-M6 Rack-Mount Server, and HyperFlex All-Flash Cluster

For Citrix Virtual Apps RDS Hosted Shared Desktop and Hosted Virtual Desktop use case, the recommended maximum workload was determined based on both Login VSI Knowledge Worker workload end user experience measures and HXAF240c-M6 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.

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 workload 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.

Note: The recommended maximum workload for a Cisco HyperFlex cluster configured on Cisco HXAF240c-M6 with Intel Xeon Gold 8000 scalable family processors and 1 TB of RAM for Windows 10 desktops with Office 2016 is 750 virtual desktops.

Note: The recommended maximum workload for a Cisco HyperFlex cluster configured on Cisco HXAF240c-M6 with Intel Xeon Gold 8000 scalable family processors and 1 TB of RAM for Windows Server 2019 RDS desktop sessions with Office 2016 is 1000 virtual desktops.

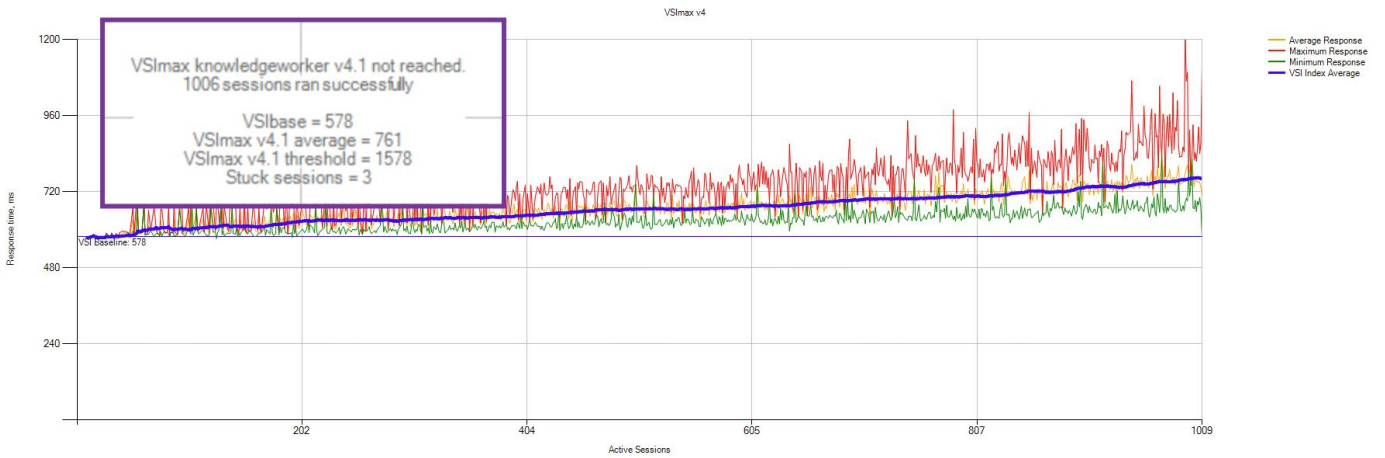
1000 RDS Sessions and 750 Windows 10 Citrix PVS Non-Persistent Testing on 24 Node Cisco HyperFlex Cluster

Hosted Shared desktops with 1000 user sessions on 45 Windows Server 2019 virtual machines on a 4-node HyperFlex cluster.

Test results for 1000 user sessions on Citrix RDS highlights include:

- 0.578 second baseline response time
- 0.761 second average response time with 1000 desktops running
- Average CPU utilization of 35 percent during steady state
- Average of 400 GB of RAM used out of 1 TB available
- 22,000 peak I/O operations per second (IOPS) per cluster at steady state
- 650MBps peak throughput per cluster at steady state

Figure 40. Login VSI Analyzer Chart for 1000 Windows 2019 Citrix Shared Desktops



Summary Settings VSImax v4 VSImax v4 Detailed VSImax v4 Detailed Weighted VSImax v4 Scatter UMEM IO CPU ZLC ZHC NFP NFO NSL

RDS04a

Successfully completed Login VSI test with **1006 knowledgeworker** sessions. VSImax (system saturation) was not reached.

Test result review

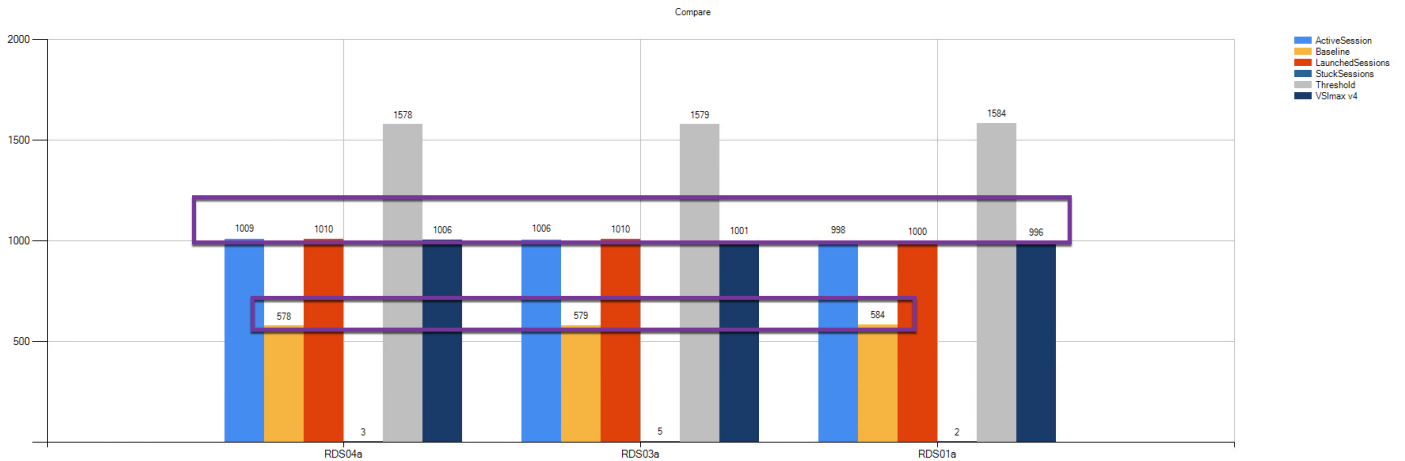
- 1010 sessions were configured to be launched in 2880 seconds.
- In total 4 sessions failed during the test:
 - 0 sessions was/were not successfully launched
 - 1 launched sessions failed to become active
 - 1009 sessions were active during the test
 - 3 sessions got stuck during the test (before VSImax threshold) > [Click Here](#)

With 1006 sessions the maximum capacity VSImax (v4.1) **knowledgeworker** was not reached with a Login VSI baseline performance score of 578

Login VSI index average score is 871 lower than threshold. It might be possible to launch more sessions in this configuration.

Baseline performance of **578** is: **Very good**

Figure 41. Three Consecutive Login VSI Analyzer Chart for 1000 Windows 2019 Citrix Shared Desktops

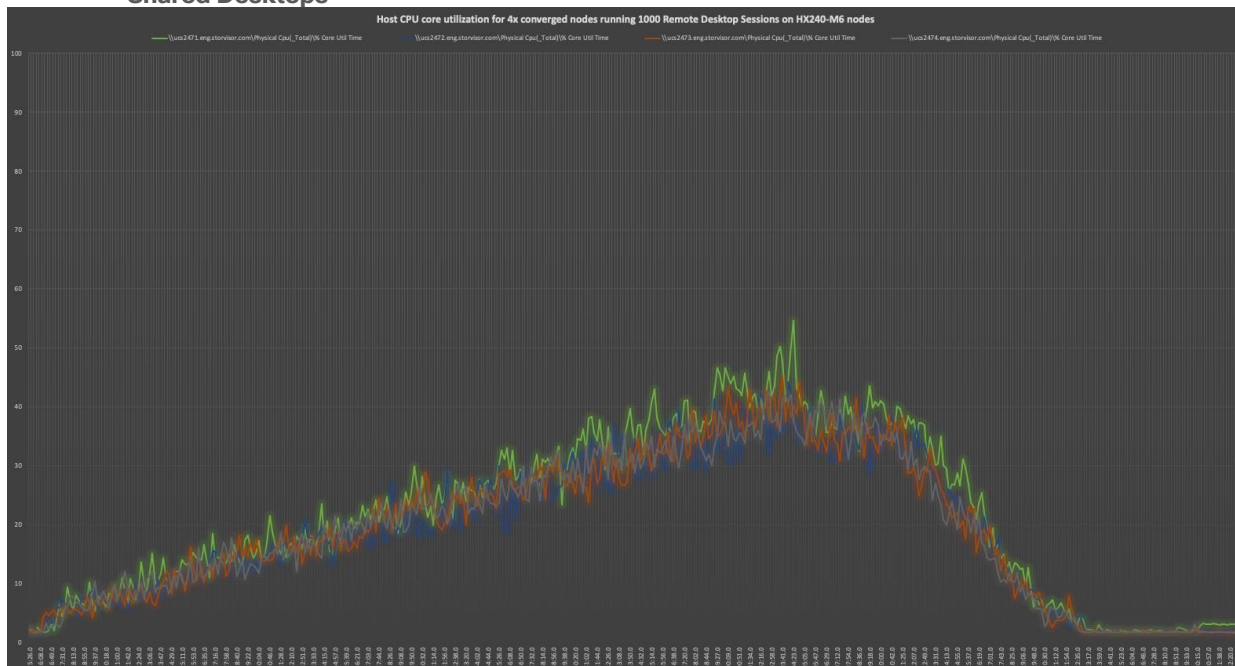


ESXi Host Performance Counters

When running a VMware ESXi environment for our Citrix 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 vswitch 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 42. Four HX240-M6 Hosts CPU Core Utilization Running 1000 Windows Server 2019 Citrix Hosted Shared Desktops



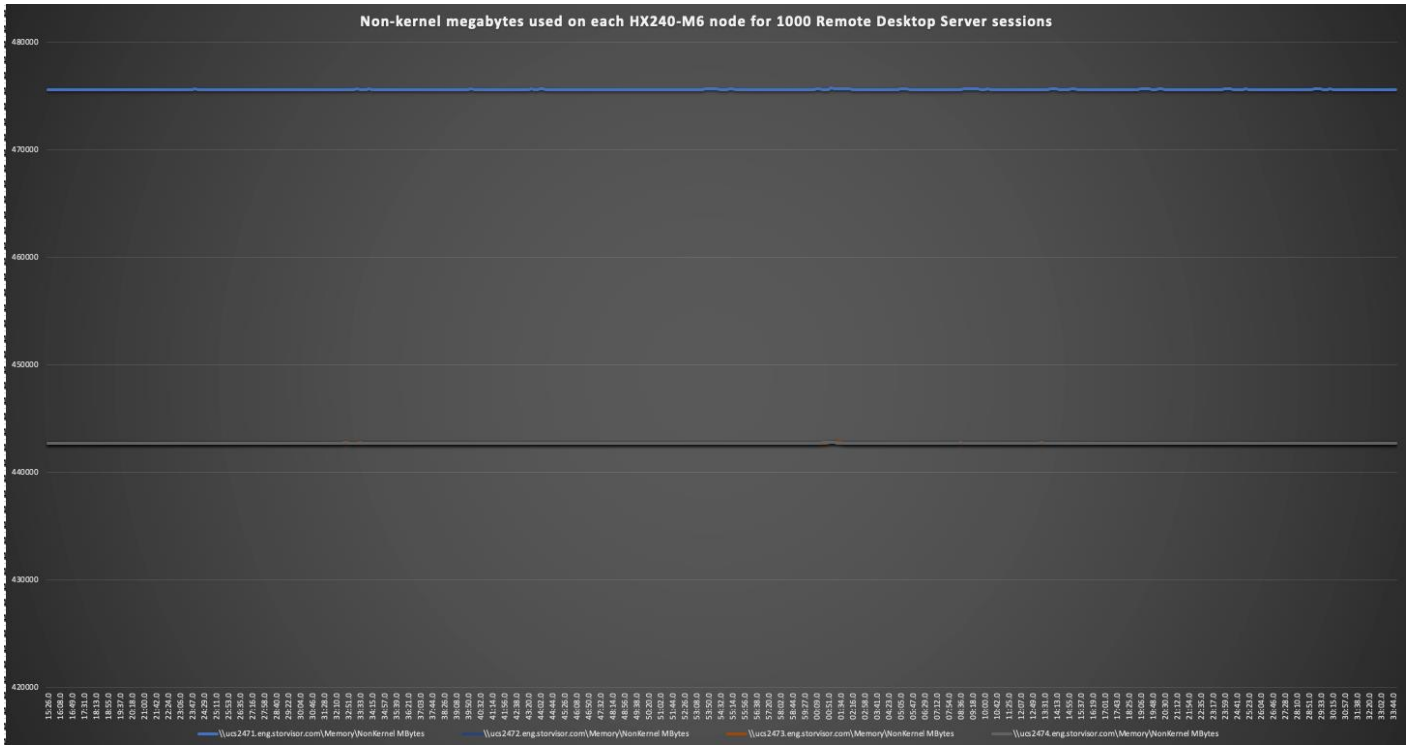
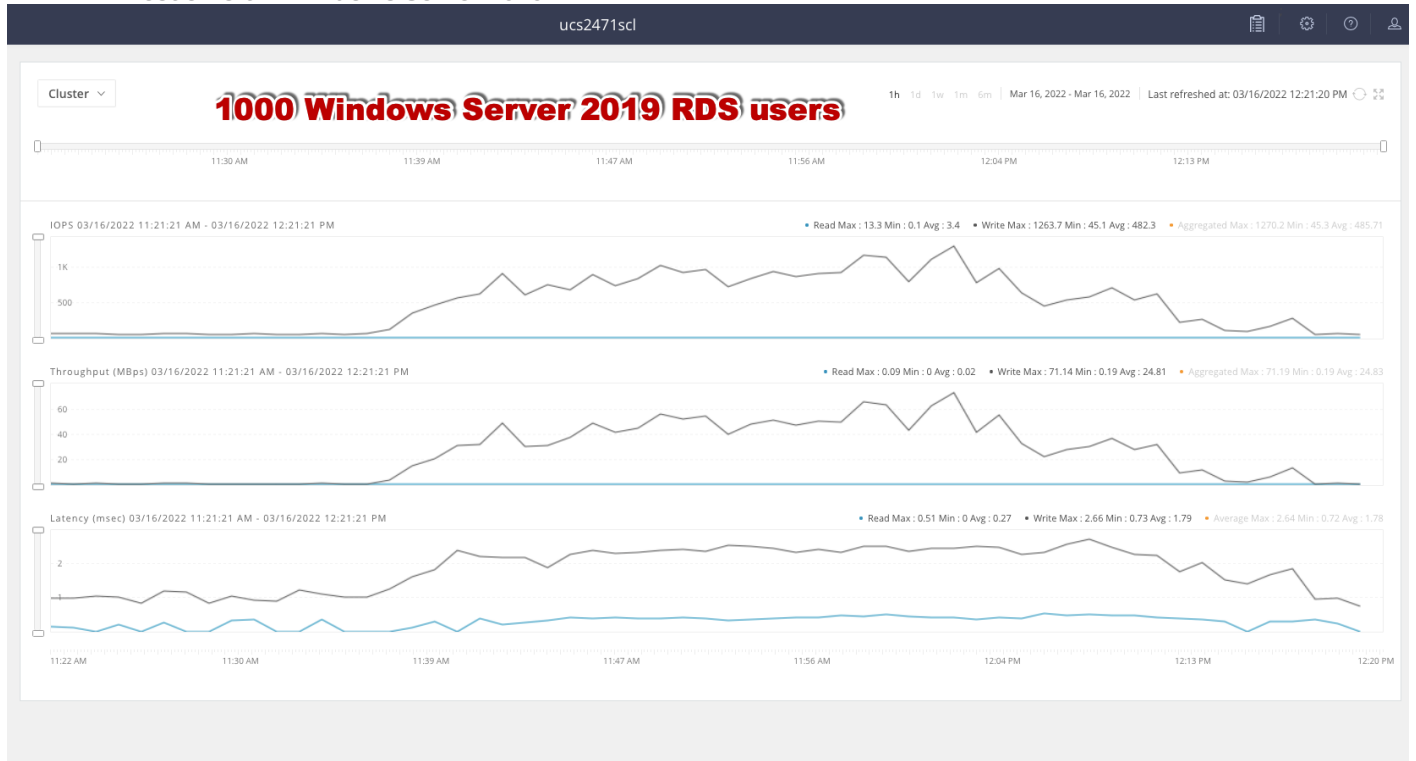


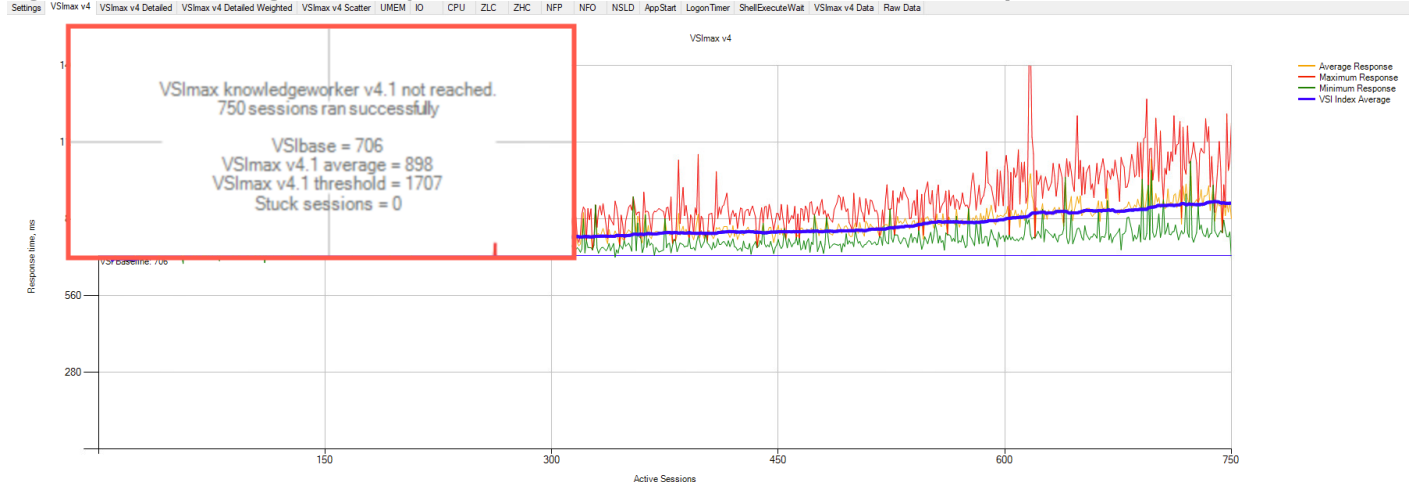
Figure 43. HyperFlex Cluster UI Performance Chart for Knowledge Worker Workload Running 1000 User Test on Citrix Windows Server 2019



Test results for 750 Citrix VDI Desktops highlights include:

- 0.706 second baseline response time
- 0.898 second average response time with 750 desktops running
- Average CPU utilization of 45 percent during steady state
- Average of 350 GB of RAM used out of 768 GB available
- 4500 peak I/O operations per second (IOPS) per cluster at steady state
- 150MBps peak throughput per cluster at steady state

Figure 44. Login VSI Analyzer Chart for 750 Windows 10 Citrix Virtual Desktops



NP-750-005a-nodes

Successfully completed Login VSI test with **750 knowledgeworker** sessions. VSImax (system saturation) was not reached. All Login VSI users completed the test.

Test result review

750 sessions were configured to be launched in **2880** seconds.

In total **0** sessions failed during the test:

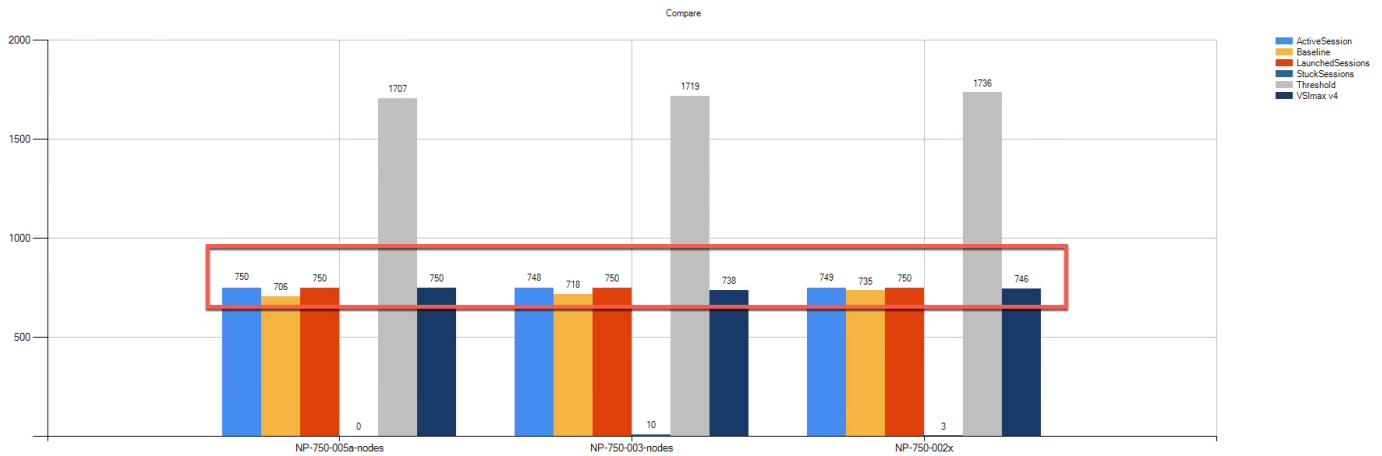
- **0** sessions was/were not successfully launched
- **0** launched sessions failed to become active
- **750** sessions were active during the test
- **0** sessions got stuck during the test (before VSImax threshold)

With **750** sessions the maximum capacity VSImax (v4.1) **knowledgeworker** was not reached with a Login VSI baseline performance score of **706**

Login VSI index average score is **866** lower than threshold. It might be possible to launch more sessions in this configuration.

Baseline performance of **706** is: **Very good**

Figure 45. Three Consecutive Login VSI Analyzer Chart for 750 Windows 10 Citrix PVS Non-persistent Virtual Desktops



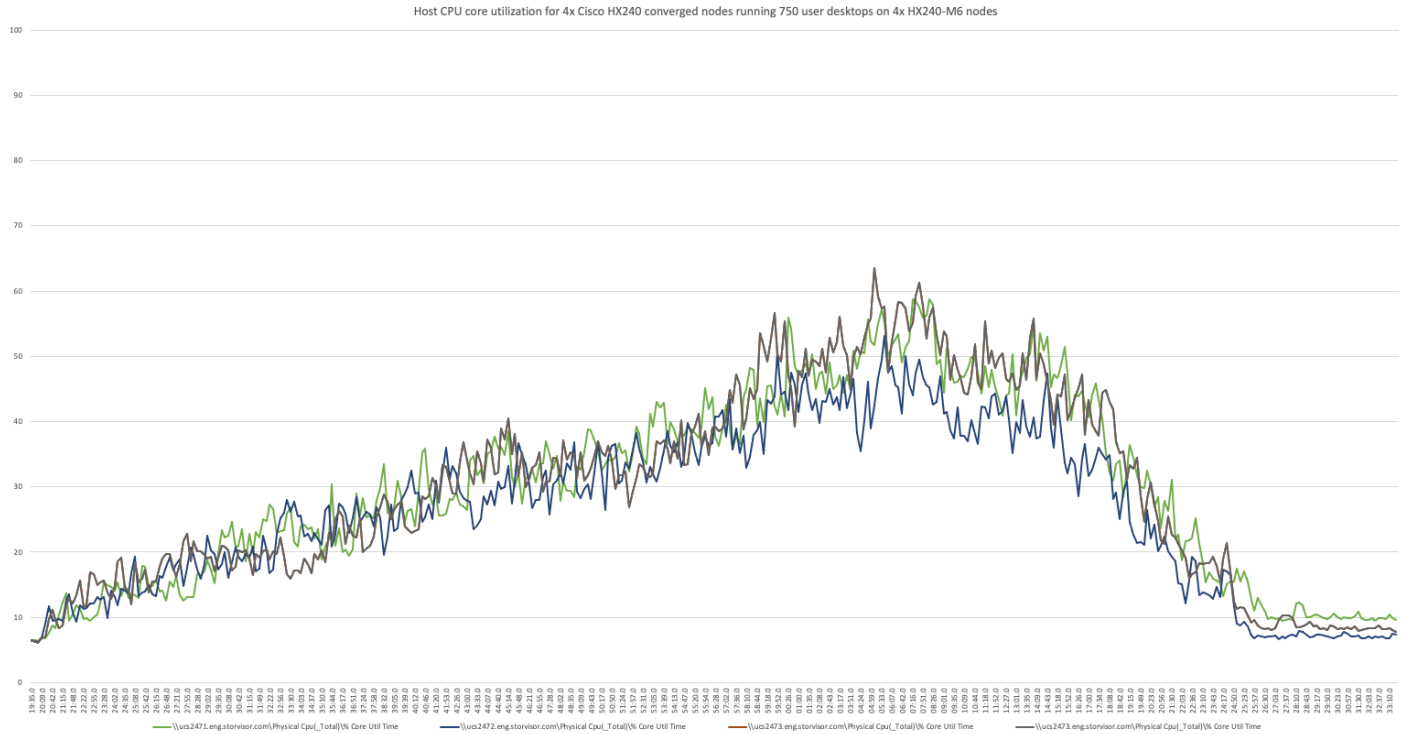
ESXi Host Performance Counters

When running a VMware ESXi environment for our Citrix 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 vswitch 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 46. Four HX240-M6 Hosts CPU Core Utilization Running 750 Windows 10 Citrix PVS Non-persistent Virtual Desktops



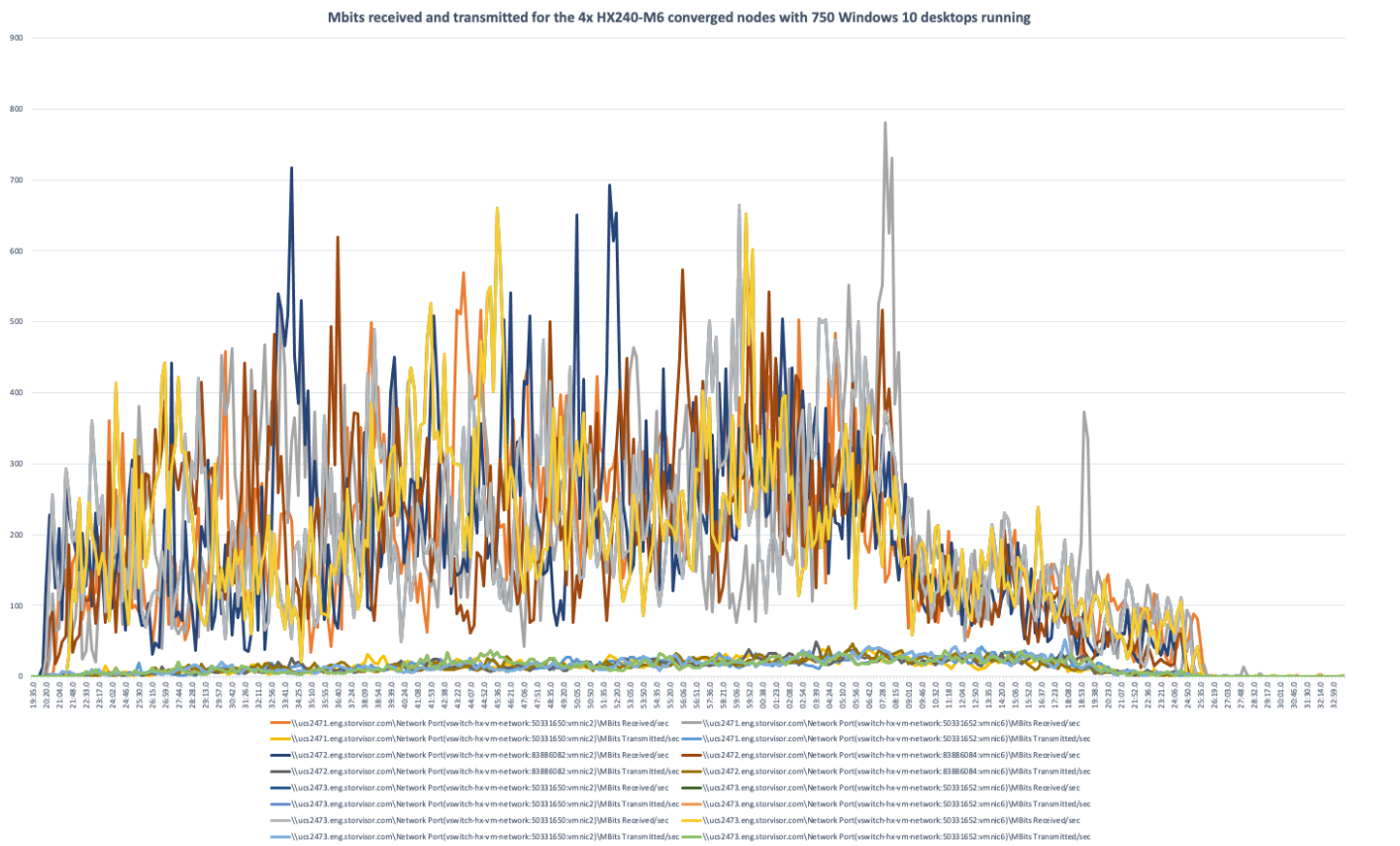
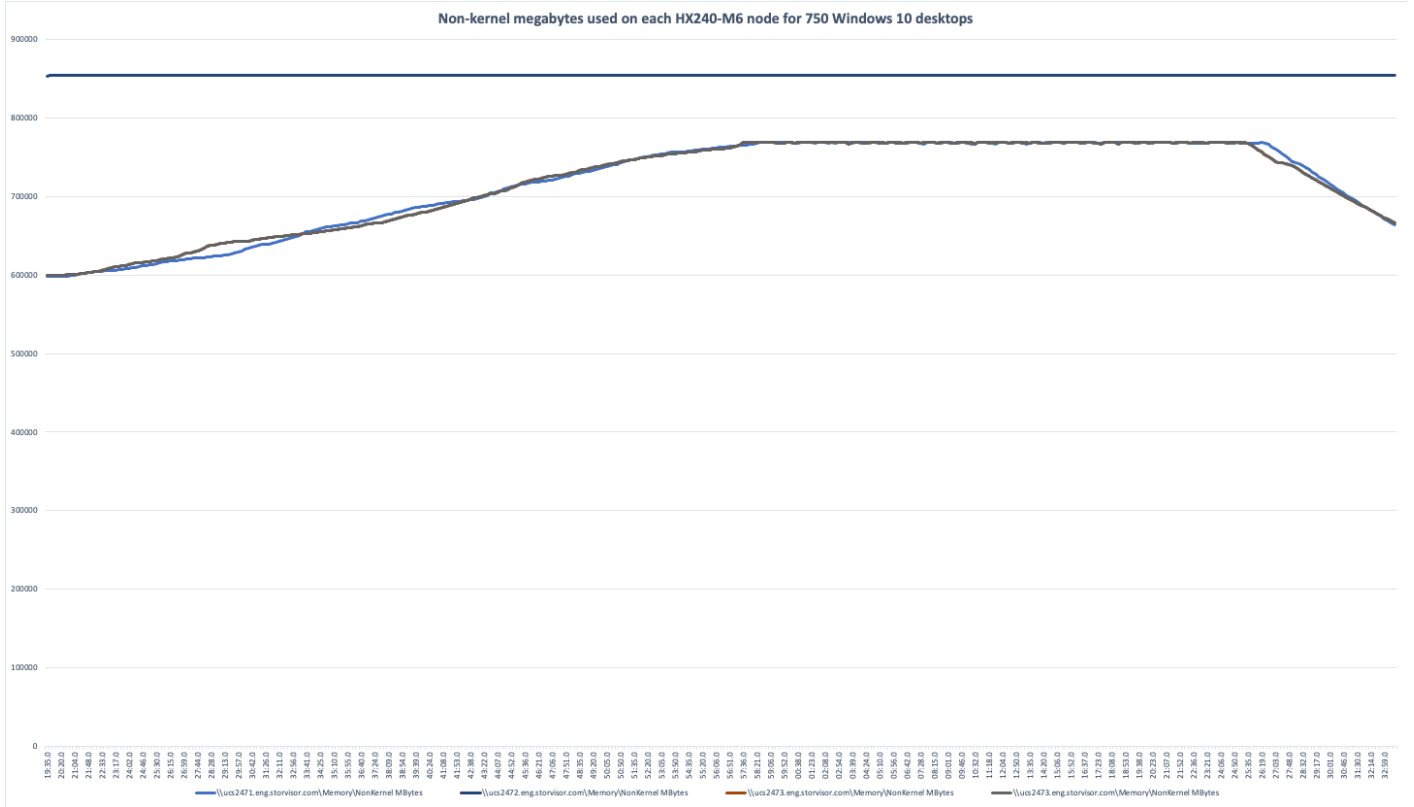
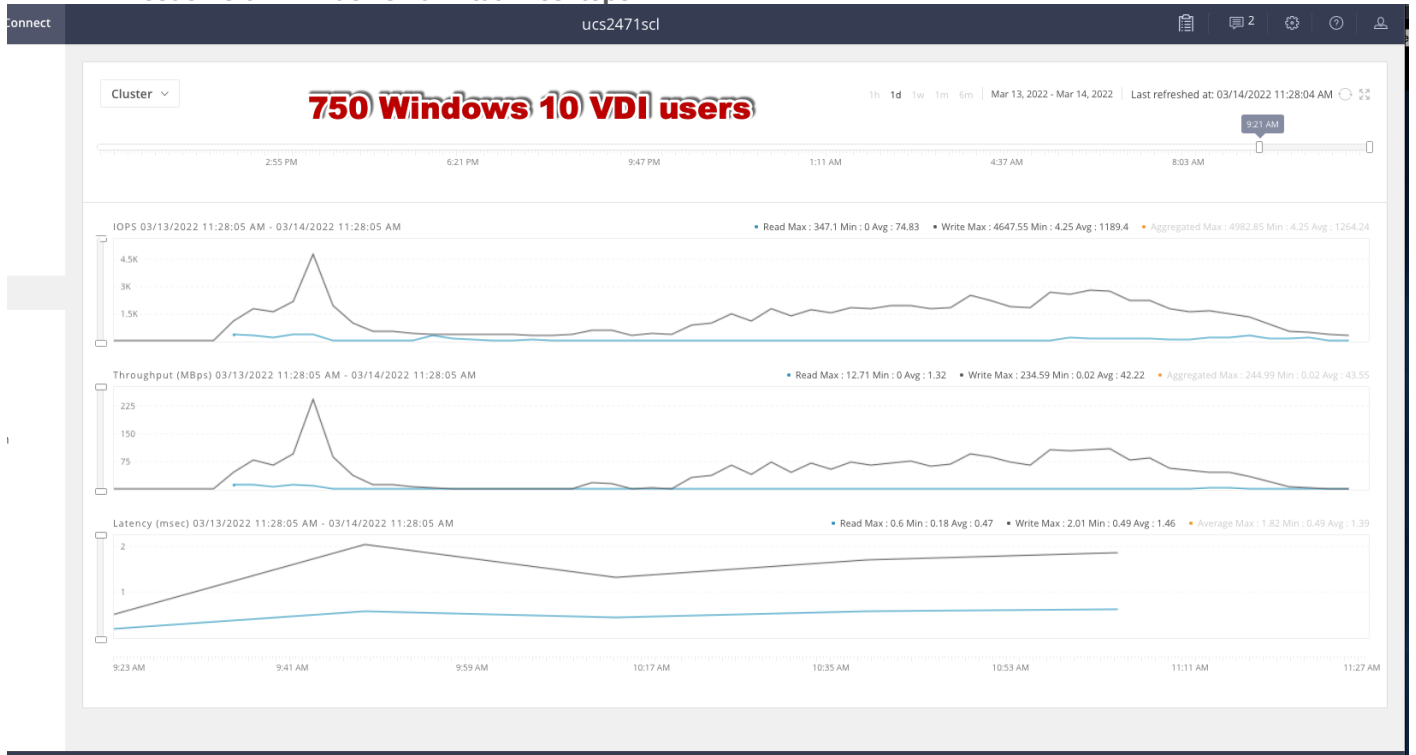


Figure 47. HyperFlex Cluster UI Performance Chart for Knowledge Worker Workload Running 750 User Test on Citrix Windows 10 Virtual Desktops



Summary

This Cisco HyperFlex solution addresses urgent needs 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.

Delivering responsive, resilient, high-performance Citrix Virtual Desktops provisioned Microsoft Windows 10 Virtual Machines and Microsoft Windows Server for hosted Apps or desktops has many advantages for desktop virtualization administrators.

The solution is fully capable of supporting graphics accelerated workloads. 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 Citrix Virtual Desktops.

Virtual desktop end-user experience, as measured by the Login VSI tool in benchmark mode, is outstanding with Intel Xeon scalable family processors and Cisco 3200 Mhz memory. In fact, we have set a new industry standard in performance for Desktop Virtualization on a hyper-converged platform.

About the Authors

Jeff Nichols, Technical Marketing Engineer, Desktop Virtualization and Graphics Solutions, Cisco Systems, Inc.

Jeff Nichols is a Cisco Unified Computing System architect, focusing on Virtual Desktop and Application solutions with extensive experience with Microsoft ESX/Hyper-V, Virtual Desktops, Virtual Apps and Microsoft Remote Desktop Services. He has expert product knowledge in application, desktop, and server virtualization across all three major hypervisor platforms and supporting infrastructures including but not limited to Windows Active Directory and Group Policies, User Profiles, DNS, DHCP and major storage platforms.

Acknowledgements

For their support and contribution to the design, validation, and creation of this Cisco Validated Design, the authors would like to thank:

- Manish Agarwal, Director, Product Management, Cisco Systems, Inc.
- John McAbel, Senior Product Manager, Cisco Systems, Inc.

Appendices

This appendix is organized into the following:

- [Appendix A - Glossary of Acronyms](#)
- [Appendix B - Glossary of Terms](#)

Appendix A - Glossary of Acronyms

- **AAA**—Authentication, Authorization, and Accounting
- **ACP**—Access-Control Policy
- **ACI**—Cisco Application Centric Infrastructure
- **ACK**—Acknowledge or Acknowledgement
- **ACL**—Access-Control List
- **AD**—Microsoft Active Directory
- **AFI**—Address Family Identifier
- **AMP**—Cisco Advanced Malware Protection
- **AP**—Access Point
- **API**—Application Programming Interface
- **APIC**— Cisco Application Policy Infrastructure Controller (ACI)
- **ASA**—Cisco Adaptative Security Appliance
- **ASM**—Any-Source Multicast (PIM)
- **ASR**—Aggregation Services Router
- **Auto-RP**—Cisco Automatic Rendezvous Point protocol (multicast)
- **AVC**—Application Visibility and Control
- **BFD**—Bidirectional Forwarding Detection
- **BGP**—Border Gateway Protocol
- **BMS**—Building Management System
- **BSR**—Bootstrap Router (multicast)
- **BYOD**—Bring Your Own Device
- **CAPWAP**—Control and Provisioning of Wireless Access Points Protocol
- **CDP**—Cisco Discovery Protocol
- **CEF**—Cisco Express Forwarding
- **CMD**—Cisco Meta Data
- **CPU**—Central Processing Unit
- **CSR**—Cloud Services Routers

-
- **CTA**—Cognitive Threat Analytics
 - **CUWN**—Cisco Unified Wireless Network
 - **CVD**—Cisco Validated Design
 - **CYOD**—Choose Your Own Device
 - **DC**—Data Center
 - **DHCP**—Dynamic Host Configuration Protocol
 - **DM**—Dense-Mode (multicast)
 - **DMVPN**—Dynamic Multipoint Virtual Private Network
 - **DMZ**—Demilitarized Zone (firewall/networking construct)
 - **DNA**—Cisco Digital Network Architecture
 - **DNS**—Domain Name System
 - **DORA**—Discover, Offer, Request, ACK (DHCP Process)
 - **DWDM**—Dense Wavelength Division Multiplexing
 - **ECMP**—Equal Cost Multi Path
 - **EID**—Endpoint Identifier
 - **EIGRP**—Enhanced Interior Gateway Routing Protocol
 - **EMI**—Electromagnetic Interference
 - **ETR**—Egress Tunnel Router (LISP)
 - **EVPN**—Ethernet Virtual Private Network (BGP EVPN with VXLAN data plane)
 - **FHR**—First-Hop Router (multicast)
 - **FHRP**—First-Hop Redundancy Protocol
 - **FMC**—Cisco Firepower Management Center
 - **FTD**—Cisco Firepower Threat Defense
 - **GBAC**—Group-Based Access Control
 - **GbE**—Gigabit Ethernet
 - **Gbit/s**—Gigabits Per Second (interface/port speed reference)
 - **GRE**—Generic Routing Encapsulation
 - **GRT**—Global Routing Table
 - **HA**—High-Availability
 - **HQ**—Headquarters
 - **HSRP**—Cisco Hot-Standby Routing Protocol
 - **HTDB**—Host-tracking Database (SD-Access control plane node construct)
 - **IBNS**—Identity-Based Networking Services (IBNS 2.0 is the current version)

-
- **ICMP**– Internet Control Message Protocol
 - **IDF**–Intermediate Distribution Frame; essentially a wiring closet.
 - **IEEE**–Institute of Electrical and Electronics Engineers
 - **IETF**–Internet Engineering Task Force
 - **IGP**–Interior Gateway Protocol
 - **IID**–Instance-ID (LISP)
 - **IOE**–Internet of Everything
 - **IoT**–Internet of Things
 - **IP**–Internet Protocol
 - **IPAM**–IP Address Management
 - **IPS**–Intrusion Prevention System
 - **IPSec**–Internet Protocol Security
 - **ISE**–Cisco Identity Services Engine
 - **ISR**–Integrated Services Router
 - **IS-IS**–Intermediate System to Intermediate System routing protocol
 - **ITR**–Ingress Tunnel Router (LISP)
 - **LACP**–Link Aggregation Control Protocol
 - **LAG**–Link Aggregation Group
 - **LAN**–Local Area Network
 - **L2 VNI**–Layer 2 Virtual Network Identifier; as used in SD-Access Fabric, a VLAN.
 - **L3 VNI**– Layer 3 Virtual Network Identifier; as used in SD-Access Fabric, a VRF.
 - **LHR**–Last-Hop Router (multicast)
 - **LISP**–Location Identifier Separation Protocol
 - **MAC**–Media Access Control Address (OSI Layer 2 Address)
 - **MAN**–Metro Area Network
 - **MEC**–Multichassis EtherChannel, sometimes referenced as **MCEC**
 - **MDF**–Main Distribution Frame; essentially the central wiring point of the network.
 - **MnT**–Monitoring and Troubleshooting Node (Cisco ISE persona)
 - **MOH**–Music on Hold
 - **MPLS**–Multiprotocol Label Switching
 - **MR**–Map-resolver (LISP)
 - **MS**–Map-server (LISP)
 - **MSDP**–Multicast Source Discovery Protocol (multicast)

-
- **MTU**—Maximum Transmission Unit
 - **NAC**—Network Access Control
 - **NAD**—Network Access Device
 - **NAT**—Network Address Translation
 - **NBAR**—Cisco Network-Based Application Recognition (NBAR2 is the current version).
 - **NFV**—Network Functions Virtualization
 - **NSF**—Non-Stop Forwarding
 - **OSI**—Open Systems Interconnection model
 - **OSPF**—Open Shortest Path First routing protocol
 - **OT**—Operational Technology
 - **PAgP**—Port Aggregation Protocol
 - **PAN**—Primary Administration Node (Cisco ISE persona)
 - **PCI DSS**—Payment Card Industry Data Security Standard
 - **PD**—Powered Devices (PoE)
 - **PETR**—Proxy-Egress Tunnel Router (LISP)
 - **PIM**—Protocol-Independent Multicast
 - **PITR**—Proxy-Ingress Tunnel Router (LISP)
 - **PnP**—Plug-n-Play
 - **PoE**—Power over Ethernet (Generic term, may also refer to IEEE 802.3af, 15.4W at PSE)
 - **PoE+**—Power over Ethernet Plus (IEEE 802.3at, 30W at PSE)
 - **PSE**—Power Sourcing Equipment (PoE)
 - **PSN**—Policy Service Node (Cisco ISE persona)
 - **pxGrid**—Platform Exchange Grid (Cisco ISE persona and publisher/subscriber service)
 - **PxTR**—Proxy-Tunnel Router (LISP - device operating as both a PETR and PITR)
 - **QoS**—Quality of Service
 - **RADIUS**—Remote Authentication Dial-In User Service
 - **REST**—Representational State Transfer
 - **RFC**—Request for Comments Document (IETF)
 - **RIB**—Routing Information Base
 - **RLOC**—Routing Locator (LISP)
 - **RP**—Rendezvous Point (multicast)
 - **RP**—Redundancy Port (WLC)
 - **RP**—Route Processor

-
- **RPF**—Reverse Path Forwarding
 - **RR**—Route Reflector (BGP)
 - **RTT**—Round-Trip Time
 - **SA**—Source Active (multicast)
 - **SAFI**—Subsequent Address Family Identifiers (BGP)
 - **SD**—Software-Defined
 - **SDA**—Cisco Software Defined-Access
 - **SDN**—Software-Defined Networking
 - **SFP**—Small Form-Factor Pluggable (1 GbE transceiver)
 - **SFP+**— Small Form-Factor Pluggable (10 GbE transceiver)
 - **SGACL**—Security-Group ACL
 - **SGT**—Scalable Group Tag, sometimes reference as Security Group Tag
 - **SM**—Spare-mode (multicast)
 - **SNMP**—Simple Network Management Protocol
 - **SSID**—Service Set Identifier (wireless)
 - **SSM**—Source-Specific Multicast (PIM)
 - **SSO**—Stateful Switchover
 - **STP**—Spanning-tree protocol
 - **SVI**—Switched Virtual Interface
 - **SVL**—Cisco StackWise Virtual
 - **SWIM**—Software Image Management
 - **SWP**—Scalable Group Tag Exchange Protocol
 - **Syslog**—System Logging Protocol
 - **TACACS+**—Terminal Access Controller Access-Control System Plus
 - **TCP**—Transmission Control Protocol (OSI Layer 4)
 - **UCS**— Cisco Unified Computing System
 - **UDP**—User Datagram Protocol (OSI Layer 4)
 - **UPoE**—Cisco Universal Power Over Ethernet (60W at PSE)
 - **UPoE+**— Cisco Universal Power Over Ethernet Plus (90W at PSE)
 - **URL**—Uniform Resource Locator
 - **VLAN**—Virtual Local Area Network
 - **VN**—Virtual Network, analogous to a VRF in SD-Access
 - **VNI**—Virtual Network Identifier (VXLAN)

- **vPC**—virtual Port Channel (Cisco Nexus)
- **VPLS**—Virtual Private LAN Service
- **VPN**—Virtual Private Network
- **VPNv4**—BGP address family that consists of a Route-Distinguisher (RD) prepended to an IPv4 prefix
- **VPWS**—Virtual Private Wire Service
- **VRF**—Virtual Routing and Forwarding
- **VSL**—Virtual Switch Link (Cisco VSS component)
- **VSS**—Cisco Virtual Switching System
- **VXLAN**—Virtual Extensible LAN
- **WAN**—Wide-Area Network
- **WLAN**—Wireless Local Area Network (generally synonymous with IEEE 802.11-based networks)
- **WoL**—Wake-on-LAN
- **xTR**—Tunnel Router (LISP - device operating as both an ETR and ITR)

Appendix B - Glossary of Terms

This glossary addresses some terms used in this document, for the purposes of aiding understanding. This is not a complete list of all multicloud terminology. Some Cisco product links are supplied here also, where considered useful for the purposes of clarity, but this is by no means intended to be a complete list of all applicable Cisco products.

<p>aaS/XaaS (IT capability provided as a Service)</p>	<p>Some IT capability, X, provided as a service (XaaS). Some benefits are:</p> <ul style="list-style-type: none"> • The provider manages the design, implementation, deployment, upgrades, resiliency, scalability, and overall delivery of the service and the infrastructure that supports it. • There are very low barriers to entry, so that services can be quickly adopted and dropped in response to business demand, without the penalty of inefficiently utilized CapEx. • The service charge is an IT OpEx cost (pay-as-you-go), whereas the CapEx and the service infrastructure is the responsibility of the provider. • Costs are commensurate to usage and hence more easily controlled with respect to business demand and outcomes. <p>Such services are typically implemented as “microservices,” which are accessed via REST APIs. This architectural style supports composition of service components into systems. Access to and management of aaS assets is via a web GUI and/or APIs, such that Infrastructure-as-code (IaC) techniques can be used for automation, for example, Ansible and Terraform.</p> <p>The provider can be any entity capable of implementing an aaS “cloud-native” architecture. The cloud-native architecture concept is well-documented and supported by open-source software and a rich ecosystem of services such as training and consultancy. The provider can be an internal IT department or any of many third-party companies using and supporting the same open-source platforms.</p> <p>Service access control, integrated with corporate IAM, can be mapped to specific users and business activities, enabling consistent policy controls across services, wherever they are delivered from.</p>
<p>Ansible</p>	<p>An infrastructure automation tool, used to implement processes for instantiating and configuring IT service components, such as VMs on an IaaS platform. Supports the consistent execution of processes defined in YAML “playbooks” at scale, across multiple</p>

	<p>targets. Because the Ansible artefacts (playbooks) are text-based, they can be stored in a Source Code Management (SCM) system, such as GitHub. This allows for software development like processes to be applied to infrastructure automation, such as, Infrastructure-as-code (see IaC below).</p> <p>https://www.ansible.com</p>
<p>AWS (Amazon Web Services)</p>	<p>Provider of IaaS and PaaS.</p> <p>https://aws.amazon.com</p>
<p>Azure</p>	<p>Microsoft IaaS and PaaS.</p> <p>https://azure.microsoft.com/en-gb/</p>
<p>Co-located data center</p>	<p>“A colocation center (CoLo)...is a type of data center where equipment, space, and bandwidth are available for rental to retail customers. Colocation facilities provide space, power, cooling, and physical security for the server, storage, and networking equipment of other firms and also connect them to a variety of telecommunications and network service providers with a minimum of cost and complexity.”</p> <p>https://en.wikipedia.org/wiki/Colocation_centre</p>

Containers (Docker)	<p>A (Docker) container is a means to create a package of code for an application and its dependencies, such that the application can run on different platforms which support the Docker environment. In the context of aaS, microservices are typically packaged within Linux containers orchestrated by Kubernetes (K8s).</p> <p>https://www.docker.com</p> <p>https://www.cisco.com/c/en/us/products/cloud-systems-management/containerplatform/index.html</p>
DevOps	<p>The underlying principle of DevOps is that the application development and operations teams should work closely together, ideally within the context of a toolchain that automates the stages of development, test, deployment, monitoring, and issue handling. DevOps is closely aligned with IaC, continuous integration and deployment (CI/CD), and Agile software development practices.</p> <p>https://en.wikipedia.org/wiki/DevOps</p> <p>https://en.wikipedia.org/wiki/CI/CD</p>
Edge compute	<p>Edge compute is the idea that it can be more efficient to process data at the edge of a network, close to the endpoints that originate that data, or to provide virtualized access services, such as at the network edge. This could be for reasons related to low latency response, reduction of the amount of unprocessed data being transported, efficiency of resource utilization, and so on. The generic label for this is Multi-access Edge Computing (MEC), or Mobile Edge Computing for mobile networks specifically.</p> <p>From an application experience perspective, it is important to be able to utilize, at the edge, the same operations model, processes, and tools used for any other compute node in the system.</p> <p>https://en.wikipedia.org/wiki/Mobile_edge_computing</p>
IaaS (Infrastructure as-a-Service)	<p>Infrastructure components provided aaS, located in data centers operated by a provider, typically accessed over the public Internet. IaaS provides a base platform for the deployment of workloads, typically with containers and Kubernetes (K8s).</p>
IaC (Infrastructure as-Code)	<p>Given the ability to automate aaS via APIs, the implementation of the automation is typically via Python code, Ansible playbooks, and similar. These automation artefacts are programming code that define how the services are consumed. As such, they can be subject to the same code management and software development regimes as any other body of code. This means that infrastructure automation can be subject to all of the quality and consistency benefits, CI/CD, traceability, automated testing, compliance checking, and so on, that could be applied to any coding project.</p> <p>https://en.wikipedia.org/wiki/Infrastructure_as_code</p>
IAM (Identity and Access Management)	<p>IAM is the means to control access to IT resources so that only those explicitly authorized to access given resources can do so. IAM is an essential foundation to a secure multicloud environment.</p> <p>https://en.wikipedia.org/wiki/Identity_management</p>
IBM (Cloud)	<p>IBM IaaS and PaaS.</p> <p>https://www.ibm.com/cloud</p>
Intersight	<p>Cisco Intersight™ is a Software-as-a-Service (SaaS) infrastructure lifecycle management platform that delivers simplified configuration, deployment, maintenance, and support.</p> <p>https://www.cisco.com/c/en/us/products/servers-unified-computing/intersight/index.html</p>

GCP (Google Cloud Platform)	Google IaaS and PaaS. https://cloud.google.com/gcp
Kubernetes (K8s)	Kubernetes is an open-source system for automating deployment, scaling, and management of containerized applications. https://kubernetes.io
Microservices	A microservices architecture is characterized by processes implementing fine-grained services, typically exposed via REST APIs and which can be composed into systems. The processes are often container-based, and the instantiation of the services often managed with Kubernetes. Microservices managed in this way are intrinsically well suited for deployment into IaaS environments, and as such, are the basis of a cloud native architecture. https://en.wikipedia.org/wiki/Microservices
PaaS (Platform-as-a-Service)	PaaS is a layer of value-add services, typically for application development, deployment, monitoring, and general lifecycle management. The use of IaC with IaaS and PaaS is very closely associated with DevOps practices.
Private on-premises data center	A data center infrastructure housed within an environment owned by a given enterprise is distinguished from other forms of data center, with the implication that the private data center is more secure, given that access is restricted to those authorized by the enterprise. Thus, circumstances can arise where very sensitive IT assets are only deployed in a private data center, in contrast to using public IaaS. For many intents and purposes, the underlying technology can be identical, allowing for hybrid deployments where some IT assets are privately deployed but also accessible to other assets in public IaaS. IAM, VPNs, firewalls, and similar are key technologies needed to underpin the security of such an arrangement.
REST API	Representational State Transfer (REST) APIs is a generic term for APIs accessed over HTTP(S), typically transporting data encoded in JSON or XML. REST APIs have the advantage that they support distributed systems, communicating over HTTP, which is a well-understood protocol from a security management perspective. REST APIs are another element of a cloud-native applications architecture, alongside microservices. https://en.wikipedia.org/wiki/Representational_state_transfer
SaaS (Software-as-a-Service)	End-user applications provided “aaS” over the public Internet, with the underlying software systems and infrastructure owned and managed by the provider.
SAML (Security Assertion Markup Language)	Used in the context of Single-Sign-On (SSO) for exchanging authentication and authorization data between an identity provider, typically an IAM system, and a service provider (some form of SaaS). The SAML protocol exchanges XML documents that contain security assertions used by the aaS for access control decisions. https://en.wikipedia.org/wiki/Security_Assertion_Markup_Language
Terraform	An open-source IaC software tool for cloud services, based on declarative configuration files. https://www.terraform.io

Feedback

For comments and suggestions about this guide and related guides, join the discussion on [Cisco Community](https://cs.co/en-cvds) at <https://cs.co/en-cvds>.

CVD Program

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

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

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, Cisco TelePresence, Cisco WebEx, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn and Cisco Store are service marks; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unified Computing System (Cisco UCS), Cisco UCS B-Series Blade Servers, Cisco UCS C-Series Rack Servers, Cisco UCS S-Series Storage Servers, Cisco UCS Manager, Cisco UCS Management Software, Cisco Unified Fabric, Cisco Application Centric Infrastructure, Cisco Nexus 9000 Series, Cisco Nexus 7000 Series, Cisco Prime Data Center Network Manager, Cisco NX-OS Software, Cisco MDS Series, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study, LightStream, Linksys, MediaTone, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trade-marks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries. (LDW_P1)

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

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

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

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

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

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/go/trademarks>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)