End User Computing (EUC)

HCI

NetApp
November 04, 2020

Table of Contents

End User Computing (EUC)

TR-4861: Hybrid Cloud VDI with Virtual Desktop Service

Suresh Thoppay, NetApp

The NetApp Virtual Desktop Service (VDS) orchestrates Remote Desktop Services (RDS) in major public clouds as well as on private clouds. VDS supports Windows Virtual Desktop (WVD) on Microsoft Azure. VDS automates many tasks that must be performed after deployment of WVD or RDS, including setting up SMB file shares (for user profiles, shared data, and the user home drive), enabling Windows features, application and agent installation, firewall, and policies, and so on.

Users consume VDS for dedicated desktops, shared desktops, and remote applications. VDS provides scripted events for automating application management for desktops and reduces the number of images to manage.

VDS provides a single management portal for handling deployments across public and private cloud environments.

Customer Value

The remote workforce explosion of 2020 has changed requirements for business continuity. IT departments are faced with new challenges to rapidly provision virtual desktops and thus require provisioning agility, remote management, and the TCO advantages of a hybrid cloud that makes it easy to provision on-premises and cloud resources. They need a hybrid-cloud solution that:

- Addresses the post-COVID workspace reality to enable flexible work models with global dynamics
- Enables shift work by simplifying and accelerating the deployment of work environments for all employees, from task workers to power users
- Mobilizes your workforce by providing rich, secure VDI resources regardless of the physical location
- Simplifies hybrid-cloud deployment
- Automates and simplifies risk reduction management

Next: Use Cases

Use Cases

Hybrid VDI with NetApp VDS allows service providers and enterprise virtual desktop administrators to easily expand resources to other cloud environment without affecting their users. Having on-premises resources on NetApp HCI provides better control of GPU resources and allows you to expand compute or

storage nodes based on demand.

This solution applies to the following use cases:

- Bursting into the cloud for surges in demand for remote desktops and applications
- Reducing TCO for long running remote desktops and applications by hosting them on-premises with flash storage and GPU resources
- Ease of management of remote desktops and applications across cloud environments
- Experience remote desktops and applications by using a software-as-a- service model with on-premises resources

Target Audience

The target audience for the solution includes the following groups:

- EUC/VDI architects who wants to understand the requirements for a hybrid VDS
- NetApp partners who would like to assist customers with their remote desktop and application needs
- Existing NetApp HCI customers who want to address remote desktop and application demands

Next: NetApp Virtual Desktop Service Overview

NetApp Virtual Desktop Service Overview

NetApp offers many cloud services, including the rapid provisioning of virtual desktop with WVD or Remote Applications, including rapid integration with Azure NetApp Files.

Traditionally, it takes weeks to provision and deliver remote desktop services to customers. Apart from provisioning, it can be difficult to manage applications, user profiles, shared data, group policy objects to enforce policies. Firewall rules can make increase complexity and requires a separate skillset and tools.

With Microsoft Azure Windows Virtual Desktop service, Microsoft takes care of maintenance for Remote Desktop Services components, allowing customers to focus on provisioning workspaces in the cloud. Customers must provision and manage the complete stack which requires special skills to manage VDI environments.

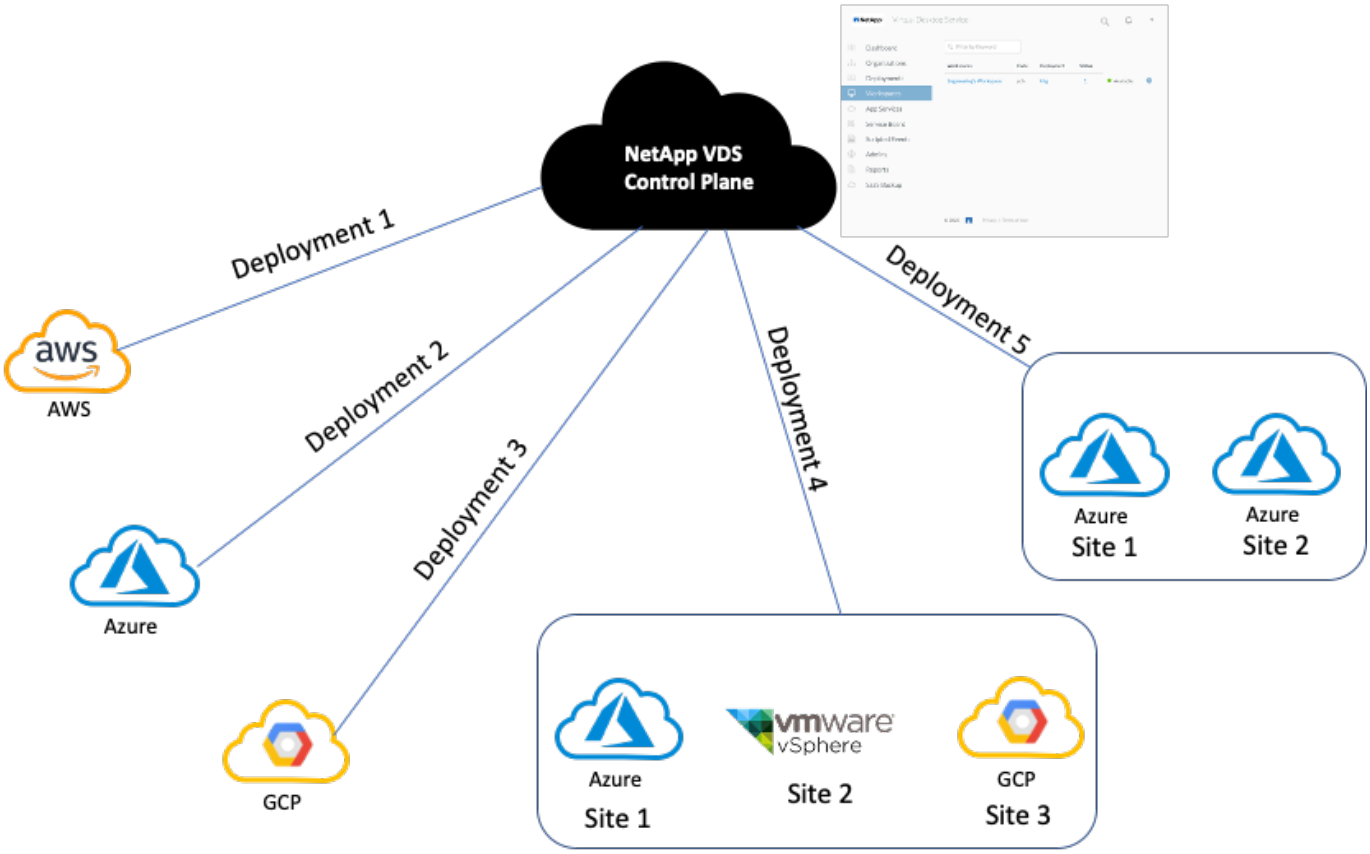
With NetApp VDS, customers can rapidly deploy virtual desktops without worrying about where to install the architecture components like brokers, gateways, agents, and so on. Customers who require complete control of their environment can work with a professional services team to achieve their goals. Customers consume VDS as a service and thus can focus on their key business challenges.

NetApp VDS is a software-as-a-service offering for centrally managing multiple deployments across

AWS, Azure, GCP, or private cloud environments. Microsoft Windows Virtual Desktop is available only on Microsoft Azure. NetApp VDS orchestrates Microsoft Remote Desktop Services in other environments.

Microsoft offers multisession on Windows 10 exclusively for Windows Virtual Desktop environments on Azure. Authentication and identity are handled by the virtual desktop technology; WVD requires Azure Active Directory synced (with AD Connect) to Active Directory and session VMs joined to Active Directory. RDS requires Active Directory for user identity and authentication and VM domain join/management.

A sample deployment topology is shown in the following figure.

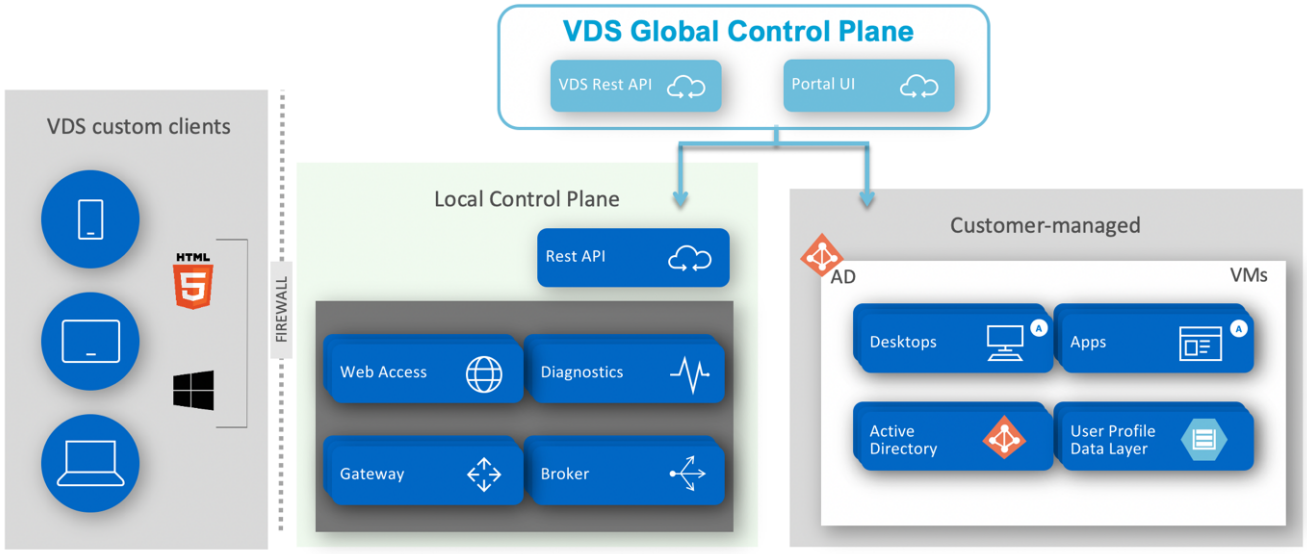


Each deployment is associated with an active directory domain and provides clients with an access entry point for workspaces and applications. A service provider or enterprise that has multiple active directory domains typically has more deployments. A single Active Directory domain that spans multiple regions typically has a single deployment with multiple sites.

For WVD in Azure, Microsoft provides a platform-as-a- service that is consumed by NetApp VDS. For other environments, NetApp VDS orchestrates the deployment and configuration of Microsoft Remote Desktop Services. NetApp VDS supports both WVD Classic and WVD ARM and can also be used to upgrade existing versions.

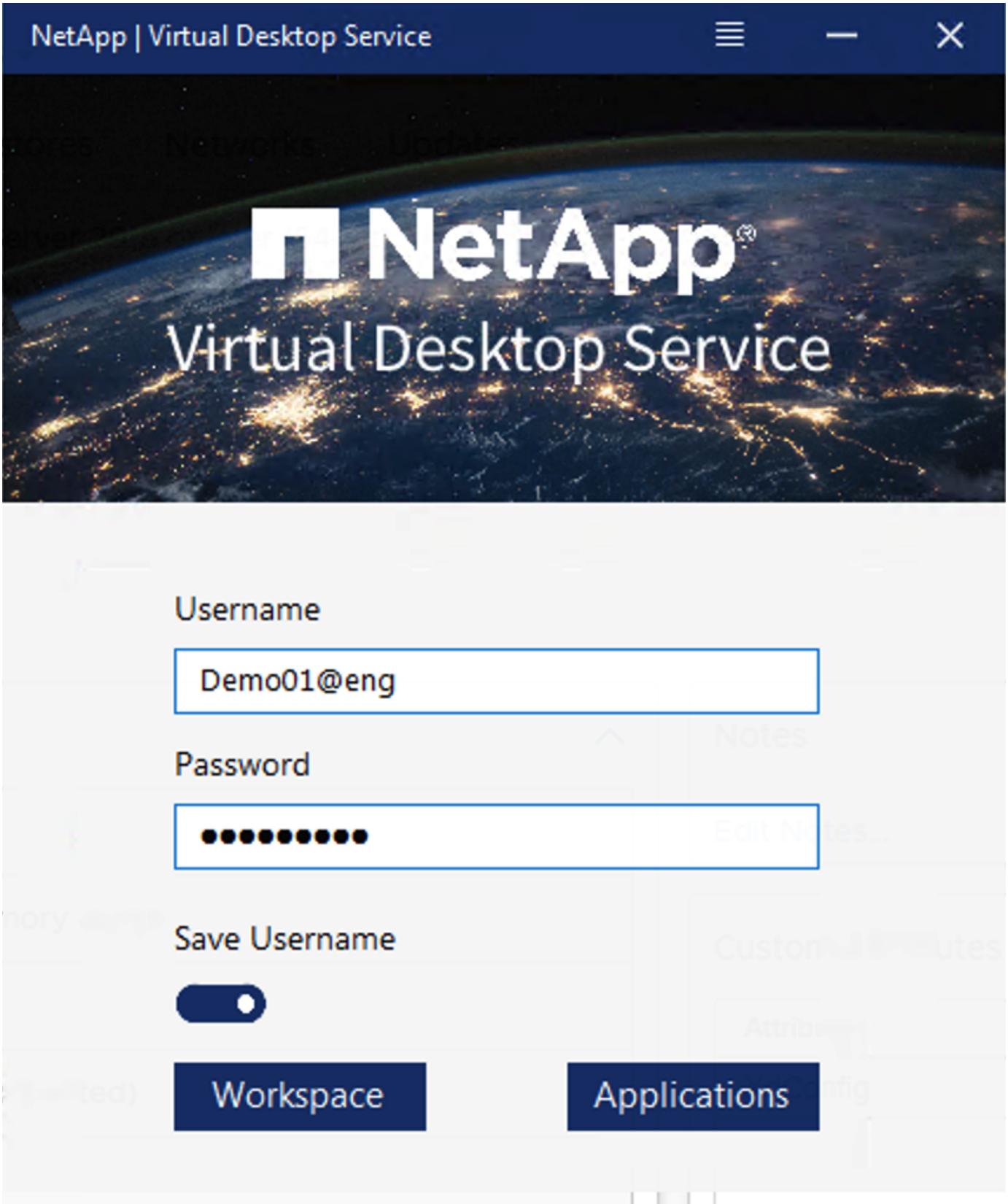
Each deployment has its own platform services, which consists of Cloud Workspace Manager (REST API endpoint), an HTML 5 Gateway (connect to VMs from a VDS management portal), RDS Gateways

(Access point for clients), and a Domain Controller. The following figure depicts the VDS Control Plane architecture for RDS implementation.



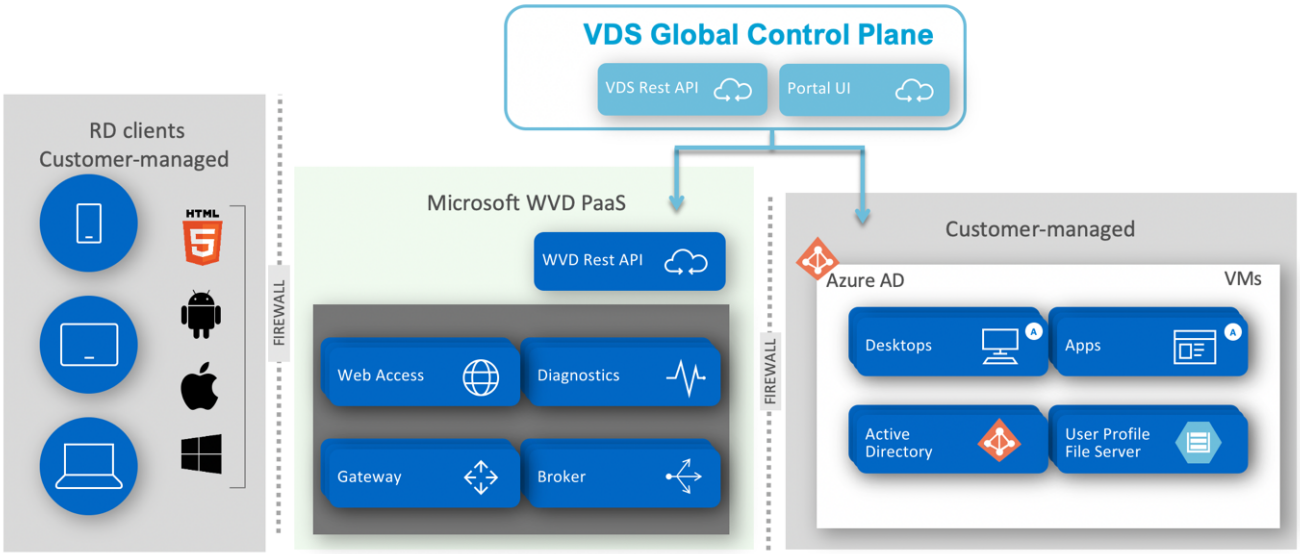
For RDS implementations, NetApp VDS can be readily accessed from Windows and browsers using client software that can be customized to include customer logo and images. Based on user credentials, it provides user access to approved workspaces and applications. There is no need to configure the gateway details.

The following figure shows the NetApp VDS client.



In the Azure WVD implementation, Microsoft handles the access entry point for the clients and can be consumed by Microsoft WVD client available natively for various OS. It can also be accessed from a web-based portal. The configuration of client software must be handled by the Group Policy Object (GPO) or in other ways preferred by customers.

The following figure depicts the VDS Control Plane architecture for Azure WVD implementations.



In addition to the deployment and configuration of required components, NetApp VDS also handles user management, application management, resource scaling, and optimization.

NetApp VDS can create users or grant existing user accounts access to cloud workspace or application services. The portal can also be used for password resets and the delegation of administrating a subset of components. Helpdesk administrators or Level-3 technicians can shadow user sessions for troubleshooting or connect to servers from within the portal.

NetApp VDS can use image templates that you create, or it can use existing ones from the marketplace for cloud-based provisioning. To reduce the number of images to manage, you can use a base image, and any additional applications that you require can be provisioned using the provided framework to include any command-line tools like Chocolatey, MSIX app attach, PowerShell, and so on. Even custom scripts can be used as part of machine lifecycle events.

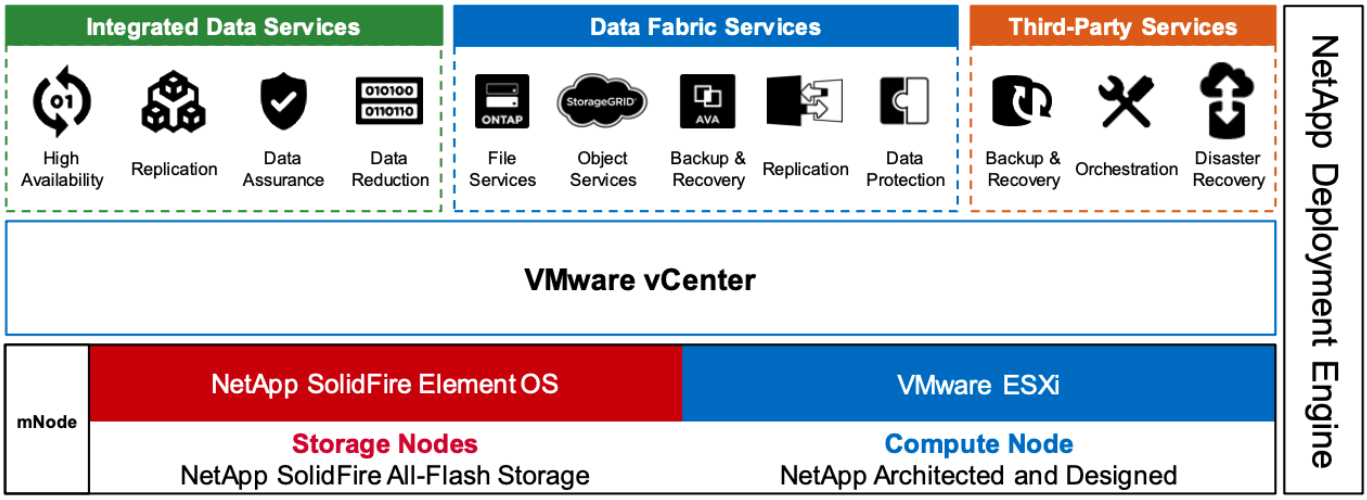
Next: NetApp HCI Overview

NetApp HCI Overview

NetApp HCI is a hybrid cloud infrastructure that consists of a mix of storage nodes and compute nodes. It is available as either a two-rack unit or single-rack unit, depending on the model. The installation and configuration required to deploy VMs are automated with the NetApp Deployment Engine (NDE). Compute clusters are managed with VMware vCenter, and storage clusters are managed with the vCenter Plug-in deployed with NDE. A management VM called the mNode is deployed as part of the NDE.

NetApp HCI handles the following functions:

- Version upgrades
- Pushing events to vCenter
- vCenter Plug-In management
- A VPN tunnel for support
- The NetApp Active IQ collector
- The extension of NetApp Cloud Services to on the premises, enabling a hybrid cloud infrastructure. The following figure depicts HCI components.



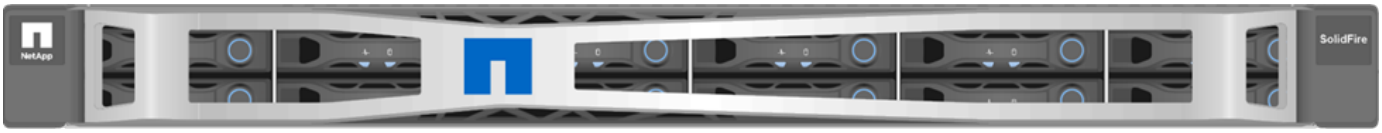
Storage Nodes

Storage nodes are available as either a half-width or full-width rack unit. A minimum of four storage nodes is required at first, and a cluster can expand to up to 40 nodes. A storage cluster can be shared across multiple compute clusters. All the storage nodes contain a cache controller to improve write performance. A single node provides either 50K or 100K IOPS at a 4K block size.

NetApp HCI storage nodes run NetApp Element software, which provides minimum, maximum, and burst QoS limits. The storage cluster supports a mix of storage nodes, although one storage node cannot exceed one-third of total capacity.

Compute Nodes

Compute nodes are available in half-width, full-width, and two rack-unit sizes. The NetApp HCI H410C and H610C are based on scalable Intel Skylake processors. The H615C is based on second-generation scalable Intel Cascade Lake processors. There are two compute models that contain GPUs: the H610C contains two NVIDIA M10 cards and the H615C contains three NVIDIA T4 cards.



The NVIDIA T4 has 40 RT cores that provide the computation power needed to deliver real-time ray

tracing. The same server model used by designers and engineers can now also be used by artists to create photorealistic imagery that features light bouncing off surfaces just as it would in real life. This RTX-capable GPU produces real-time ray tracing performance of up to five Giga Rays per second. The NVIDIA T4, when combined with Quadro Virtual Data Center Workstation (Quadro vDWS) software, enables artists to create photorealistic designs with accurate shadows, reflections, and refractions on any device from any location.

Tensor cores enable you to run deep learning inferencing workloads. When running these workloads, an NVIDIA T4 powered with Quadro vDWS can perform up to 25 times faster than a VM driven by a CPU-only server. A NetApp H615C with three NVIDIA T4 cards in one rack unit is an ideal solution for graphics and compute-intensive workloads.

The following figure lists NVIDIA GPU cards and compares their features.

NVIDIA GPUs Recommended for Virtualization

	V100S	RTX 8000	RTX 6000	T4 (Available on NetApp HCI H615C)	M10 (Available on NetApp HCI H610C)	P6
GPU	1 NVIDIA Volta	1 NVIDIA Turing	1 NVIDIA Turing	1 NVIDIA Turing	4 NVIDIA Maxwell	1 NVIDIA Pascal
CUDA Cores	5,120	4,608	4,608	2,560	2,560 (640 per GPU)	2,048
Tensor Cores	640	576	576	320	—	—
RT Cores	—	72	72	40	—	—
Guaranteed QoS (GPU Scheduler)	✓	✓	✓	✓	—	✓
Live Migration	✓	✓	✓	✓	✓	✓
Multi-vGPU	✓	✓	✓	✓	✓	✓
Memory Size	32/16 GB HBM2	48 GB GDDR6	24 GB GDDR6	16 GB GDDR6	32 GB GDDR5 (8 GB per GPU)	16 GB GDDR5
vGPU Profiles	1 GB, 2 GB, 4 GB, 8 GB, 16 GB, 32 GB	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 16 GB, 24 GB, 48 GB	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 24 GB	1 GB, 2 GB, 4 GB, 8 GB, 16 GB	0.5 GB, 1 GB, 2 GB, 4 GB, 8 GB	1 GB, 2 GB, 4 GB, 8 GB, 16 GB
Form Factor	PCIe 3.0 dual slot and SXM2	PCIe 3.0 dual slot	PCIe 3.0 dual slot	PCIe 3.0 single slot	PCIe 3.0 dual slot	MXM (blade servers)
Power	250 W /300 W (SXM2)	250 W	250 W	70 W	225 W	90 W
Thermal	passive	passive	passive	passive	passive	bare board
vGPU Software Support	Quadro vDWS, GRID vPC, GRID vApps, vComputeServer	Quadro vDWS, GRID vPC, GRID vApps, vComputeServer	Quadro vDWS, GRID vPC, GRID vApps, vComputeServer	Quadro vDWS, GRID vPC, GRID vApps, vComputeServer	Quadro vDWS, GRID vPC, GRID vApps	Quadro vDWS, GRID vPC, GRID vApps, vComputeServer
Use Case	Ultra-high-end rendering, simulation, 3D design with Quadro vDWS; ideal upgrade path for V100	High-end rendering, 3D design and creative workflows with Quadro vDWS	Mid-range to high-end rendering, 3D design and creative workflows with Quadro vDWS	Entry-level to highend 3D design and engineering workflows with Quadro vDWS. High-density, low power GPU acceleration for knowledge workers with NVIDIA GRID software.	Knowledge workers using modern productivity apps and Windows 10 requiring best density and total cost of ownership (TCO), multimonitor support with NVIDIA GRID vPC/vApps	For customers requiring GPUs in a blade server form factor; ideal upgrade path for M6

The M10 GPU remains the best TCO solution for knowledge-worker use cases. However, the T4 makes a great alternative when IT wants to standardize on a GPU that can be used across multiple use cases, such as virtual workstations, graphics performance, real-time interactive rendering, and inferencing. With the T4, IT can take advantage of the same GPU resources to run mixed workloads—for example, running VDI during the day and repurposing the resources to run compute workloads at night.

The H610C compute node is two rack units in size; the H615C is one rack unit in size and consumes less power. The H615C supports H.264 and H.265 (High Efficiency Video Coding [HEVC]) 4:4:4 encoding and decoding. It also supports a VP9 decoder, which is becoming more mainstream; even the WebM container package served by YouTube uses the VP9 codec for video.

The number of nodes in a compute cluster is dictated by VMware; currently, it is 96 with VMware vSphere 7.0 Update 1. Mixing different models of compute nodes in a cluster is supported when

Enhanced vMotion Compatibility (EVC) is enabled.

Next: NVIDIA Licensing

NVIDIA Licensing

When using an H610C or H615C, the license for the GPU must be procured from NVIDIA partners that are authorized to resell the licenses. You can find NVIDIA partners with the partner locator. Search for competencies such as virtual GPU (vGPU) or Tesla.

NVIDIA vGPU software is available in four editions:

- NVIDIA GRID Virtual PC (GRID vPC)
- NVIDIA GRID Virtual Applications (GRID vApps)
- NVIDIA Quadro Virtual Data Center Workstation (Quadro vDWS)
- NVIDIA Virtual ComputeServer (vComputeServer)

GRID Virtual PC

This product is ideal for users who want a virtual desktop that provides a great user experience for Microsoft Windows applications, browsers, high-definition video, and multi-monitor support. The NVIDIA GRID Virtual PC delivers a native experience in a virtual environment, allowing you to run all your PC applications at full performance.

GRID Virtual Applications

GRID vApps are for organizations deploying a Remote Desktop Session Host (RDSH) or other app-streaming or session-based solutions. Designed to deliver Microsoft Windows applications at full performance, Windows Server-hosted RDSH desktops are also supported by GRID vApps.

Quadro Virtual Data Center Workstation

This edition is ideal for mainstream and high-end designers who use powerful 3D content creation applications like Dassault CATIA, SOLIDWORKS, 3Dexcite, Siemens NX, PTC Creo, Schlumberger Petrel, or Autodesk Maya. NVIDIA Quadro vDWS allows users to access their professional graphics applications with full features and performance anywhere on any device.

NVIDIA Virtual ComputeServer

Many organizations run compute-intensive server workloads such as artificial intelligence (AI), deep learning (DL), and data science. For these use cases, NVIDIA vComputeServer software virtualizes the NVIDIA GPU, which accelerates compute-intensive server workloads with features such as error correction code, page retirement, peer-to-peer over NVLink, and multi-vGPU.

A Quadro vDWS license enables you to use GRID vPC and NVIDIA vComputeServer.

Next: Deployment

Deployment

NetApp VDS can be deployed to Microsoft Azure using a setup App available based on the required codebase. The current release is available at https://cwasetup.cloudworkspace.com and the preview release of the upcoming product is available at https://preview.cwasetup.cloudworkspace.com.

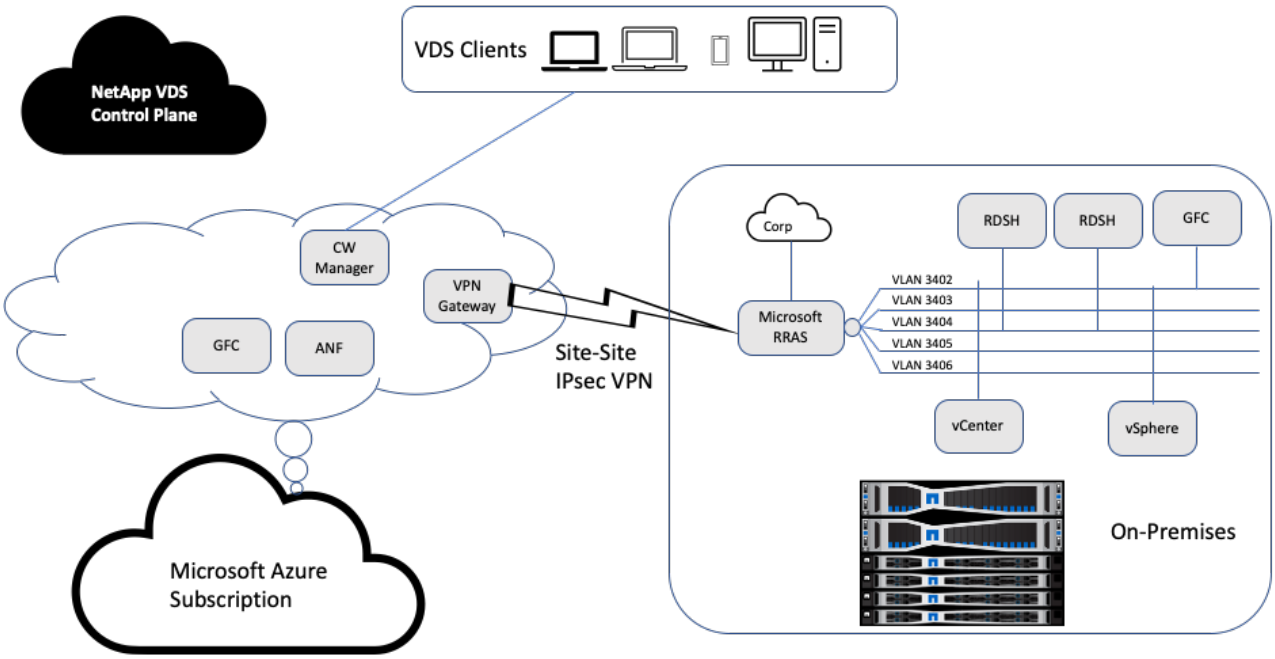
See this video for deployment instructions.

Next: Hybrid Cloud Environment

Hybrid Cloud Environment

NetApp Virtual Desktop Service can be extended to on-premises when connectivity exists between on-premises resources and cloud resources. Enterprises can establish the link to Microsoft Azure using Express Route or a site-to-site IPsec VPN connection. You can also create links to other clouds in a similar way either using a dedicated link or with an IPsec VPN tunnel.

For the solution validation, we used the environment depicted in the following figure.



On-premises, we had multiple VLANs for management, remote-desktop-session hosts, and so on. They were on the 172.21.146-150.0/24 subnet and routed to the corporate network using the Microsoft Remote Routing Access Service. We also performed the following tasks:

1. We noted the public IP of the Microsoft Routing and Remote Access Server (RRAS; identified with IPchicken.com).
2. We created a Virtual Network Gateway resource (route-based VPN) on Azure Subscription.
3. We created the connection providing the local network gateway address for the public IP of the Microsoft RRAS server.
4. We completed VPN configuration on RRAS to create a virtual interface using pre-shared authentication that was provided while creating the VPN gateway. If configured correctly, the VPN should be in the connected state. Instead of Microsoft RRAS, you can also use pfSense or other relevant tools to create the site-to-site IPsec VPN tunnel. Since it is route-based, the tunnel redirects traffic based on the specific subnets configured.

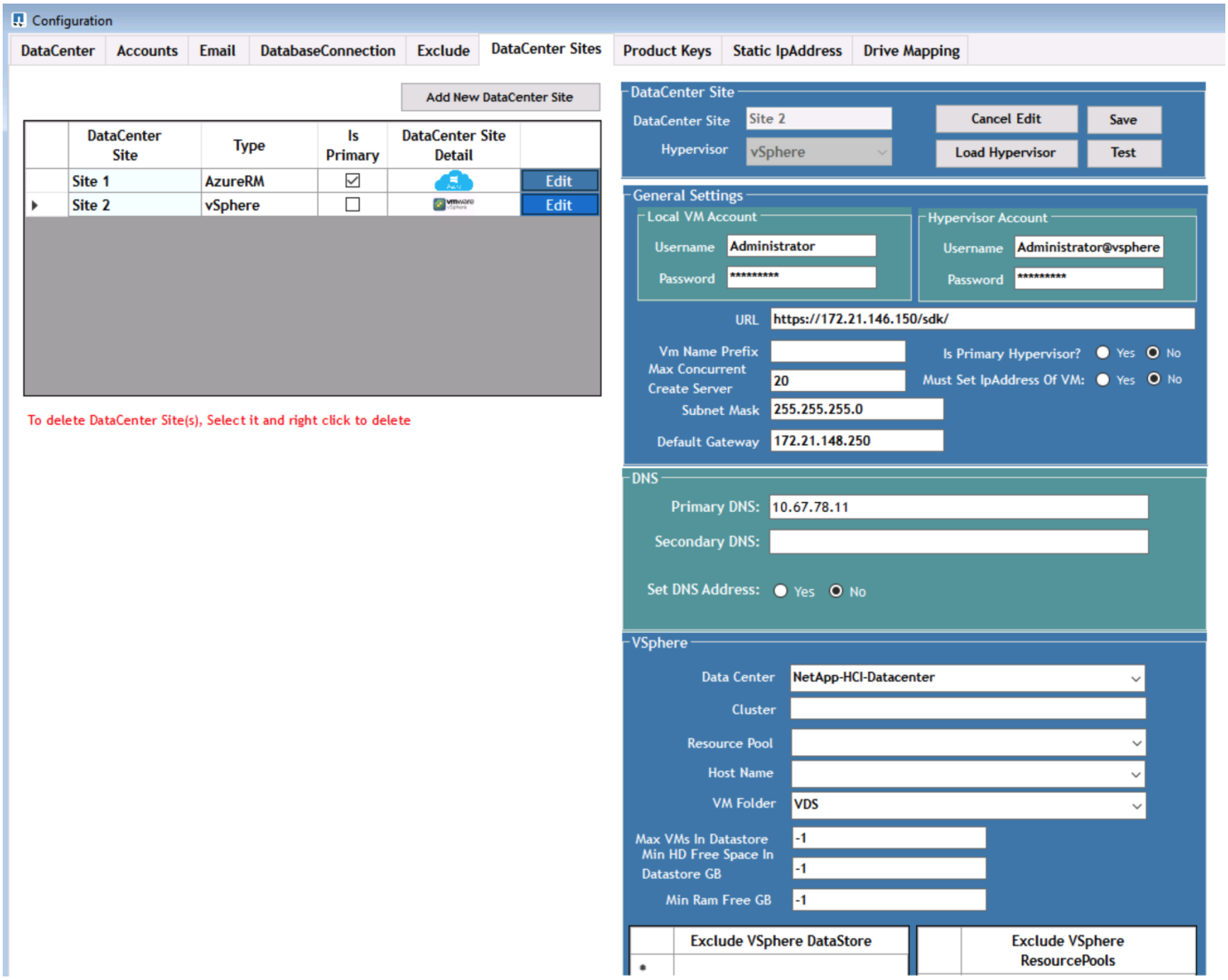
Microsoft Azure Active Directory provides identity authentication based on oAuth. Enterprise client authentications typically require NTLM or Kerberos-based authentication. Microsoft Azure Active Directory Domain Services perform password hash sync between Azure Active Directory and on-prem domain controllers using ADConnect.

For this Hybrid VDS solution validation, we initially deployed to Microsoft Azure and added an additional site with vSphere. The advantage with this approach is that platform services were deployed to Microsoft Azure and were then readily backed up using the portal. Services can then be easily accessed from anywhere, even if the site-site VPN link is down.

To add another site, we used a tool called DCConfig. The shortcut to that application is available on the desktop of the cloud workspace manager (CWMgr) VM. After this application is launched, navigate to the DataCenter Sites tab, add the new datacenter site, and fill in the required info as shown below. The URL points to the vCenter IP. Make sure that the CWMgr VM can communicate with vCenter before adding the configuration.

Make sure that vSphere PowerCLI 5.1 on CloudWorkspace manager is installed to enable communication with VMware vSphere environment.

The following figure depicts on- premises datacenter site configuration.



Note that there are filtering options available for compute resource based on the specific cluster, host name, or free RAM space. Filtering options for storage resource includes the minimum free space on datastores or the maximum VMs per datastore. Datastores can be excluded using regular expressions. Click Save button to save the configuration.

To validate the configuration, click the Test button or click Load Hypervisor and check any dropdown under the vSphere section. It should be populated with appropriate values. It is a best practice to keep the primary hypervisor set to yes for the default provisioning site.

The VM templates created on VMware vSphere are consumed as provisioning collections on VDS. Provisioning collections come in two forms: shared and VDI. The shared provisioning collection type is used for remote desktop services for which a single resource policy is applied to all servers. The VDI type is used for WVD instances for which the resource policy is individually assigned. The servers in a provisioning collection can be assigned one of the following three roles:

- **TSDATA.** Combination of Terminal Services and Data server role.
- **TS.** Terminal Services (Session Host).
- **DATA.** File Server or Database Server. When you define the server role, you must pick the VM

template and storage (datastore). The datastore chosen can be restricted to a specific datastore or you can use the least-used option in which the datastore is chosen based on data usage.

Each deployment has VM resource defaults for the cloud resource allocation based on Active Users, Fixed, Server Load, or User Count.

Next: Single Server Load Test with Login VSI

Single server load test with Login VSI

The NetApp Virtual Desktop Service uses the Microsoft Remote Desktop Protocol to access virtual desktop sessions and applications, and the Login VSI tool determines the maximum number of users that can be hosted on a specific server model. Login VSI simulates user login at specific intervals and performs user operations like opening documents, reading and composing mails, working with Excel and PowerPoint, printing documents, compressing files, and taking random breaks. It then measures response times. User response time is low when server utilization is low and increases when more user sessions are added. Login VSI determines the baseline based on initial user login sessions and it reports the maximum user session when the user response exceeds 2 seconds from the baseline.

The following table contains the hardware used for this validation.

Model	Count	Description
NetApp HCI H610C	4	Three in a cluster for launchers, AD, DHCP, and so on. One server for load testing.
NetApp HCI H615C	1	2x24C Intel Xeon Gold 6282 @2.1GHz. 1.5TB RAM.

The following table contains the software used for this validation.

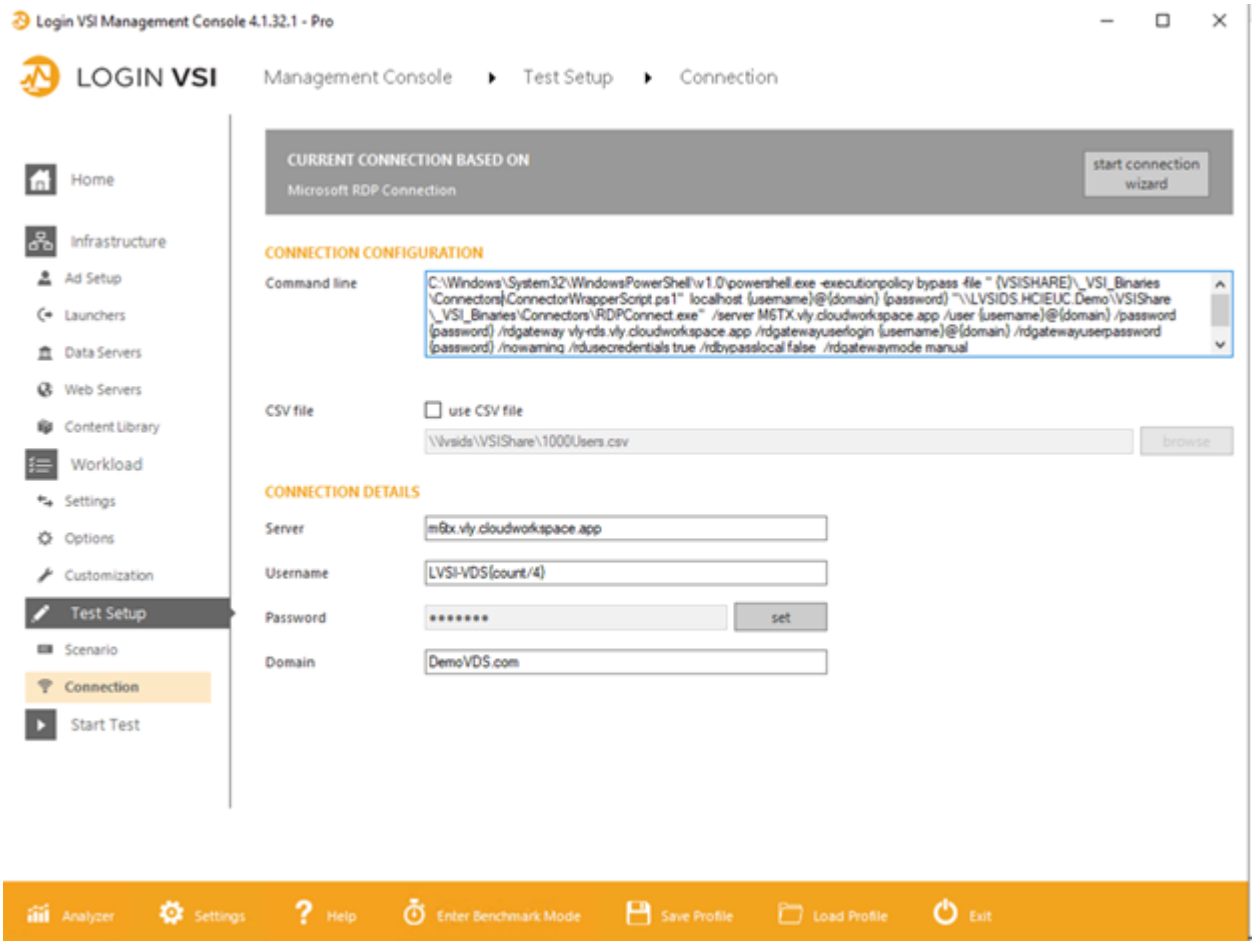
product	Description
NetApp VDS 5.4	Orchestration
VM Template Windows 2019 1809	Server OS for RDSH
Login VSI	4.1.32.1
VMware vSphere 6.7 Update 3	Hypervisor
VMware vCenter 6.7 Update 3f	VMware management tool

The Login VSI test results are as follows:

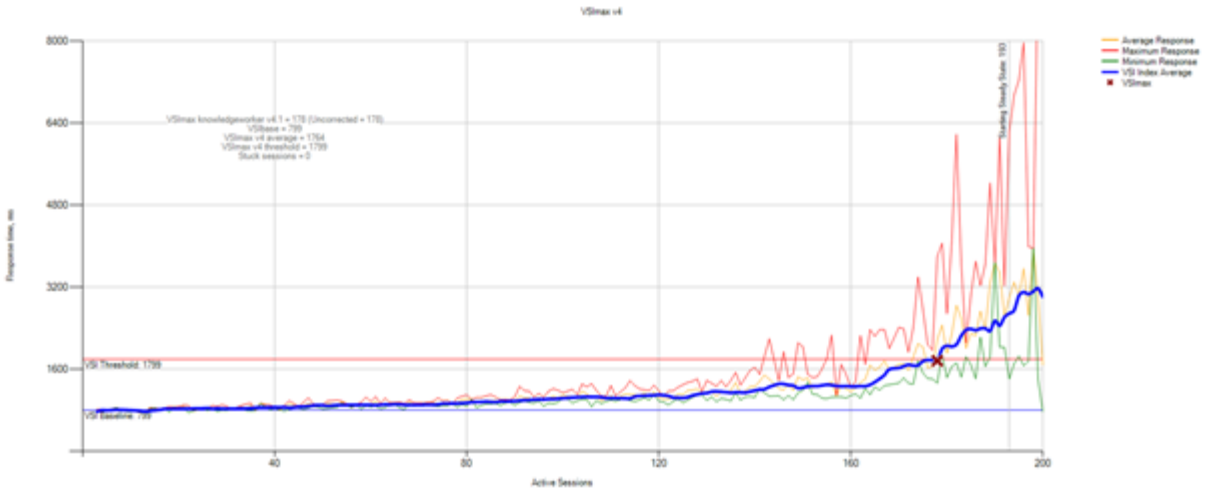
Model	VM configuration	Login VSI baseline	Login VSI Max
H610C	8 vCPU, 48GB RAM, 75GB disk, 8Q vGPU profile	799	178
H615C	12 vCPU, 128GB RAM, 75GB disk	763	272

Considering sub-NUMA boundaries and hyperthreading, the eight VMs chosen for VM testing and configuration depended on the cores available on the host.

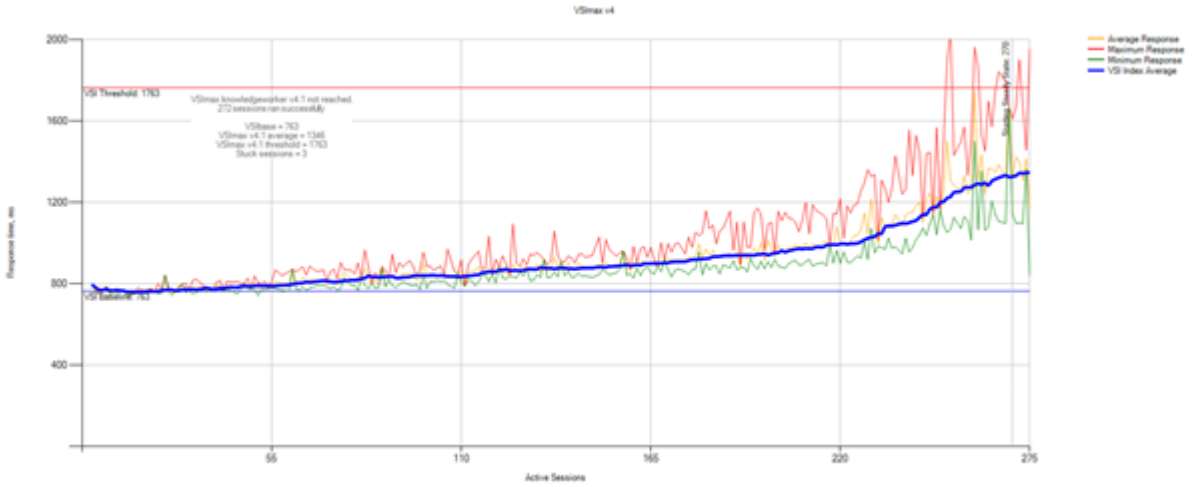
We used 10 launcher VMs on the H610C, which used the RDP protocol to connect to the user session. The following figure depicts the Login VSI connection information.



The following figure displays the Login VSI response time versus the active sessions for the H610C.



The following figure displays the Login VSI response time versus the active sessions for the H615C.



The performance metrics from Cloud Insights during H615C Login VSI testing for vSphere host and VMs is shown in teh following figure.

Next: Management Portal

Management Portal

NetApp VDS Cloud Workspace Management Suite portal is available here and the upcoming version is available here.

The portal allows centralized management for various VDS deployments including one that has sites defined for on-premises, administrative users, the application catalog, and scripted events. The portal is also used by administrative users for the manual provisioning of applications if required and to connect to any machines for troubleshooting.

Service providers can use this portal to add their own channel partners and allow them to manage their own clients.

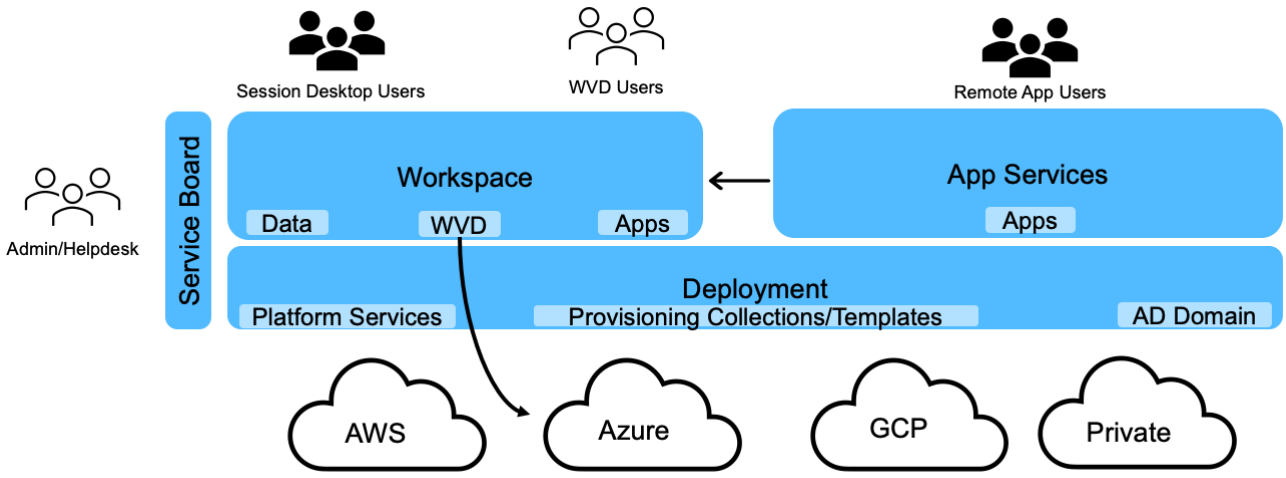
Next: User Management

User Management

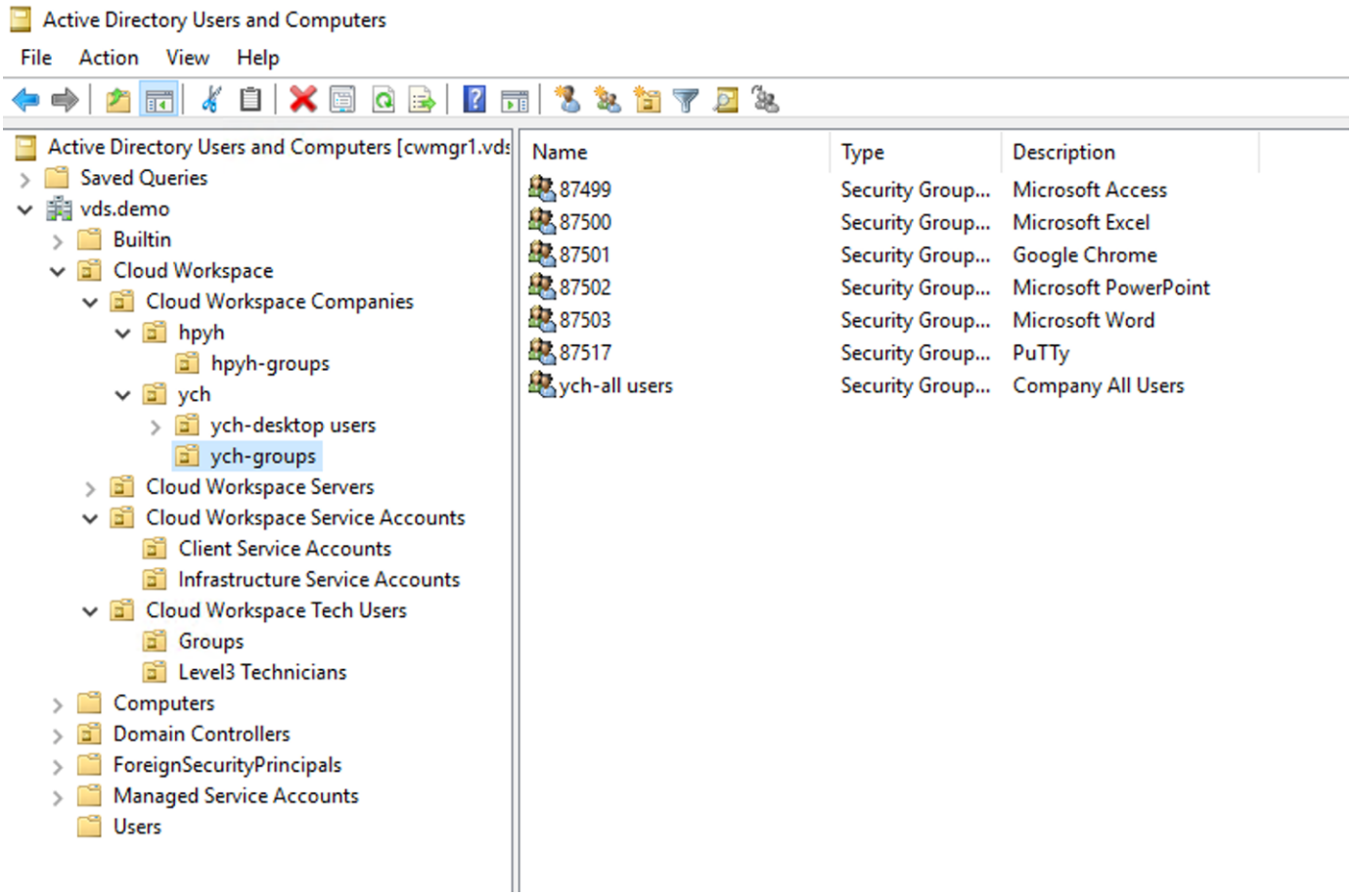
NetApp VDS uses Azure Active Directory for identity authentication and Azure Active Directory Domain Services for NTLM/Kerberos authentication. The ADConnect tool can be used to sync an on-prem Active Directory domain with Azure Active Directory.

New users can be added from the portal, or you can enable cloud workspace for existing users. Permissions for workspaces and application services can be controlled by individual users or by groups. From the management portal, administrative users can be defined to control permissions for the portal, workspaces, and so on.

The following figure depicts user management in NetApp VDS.



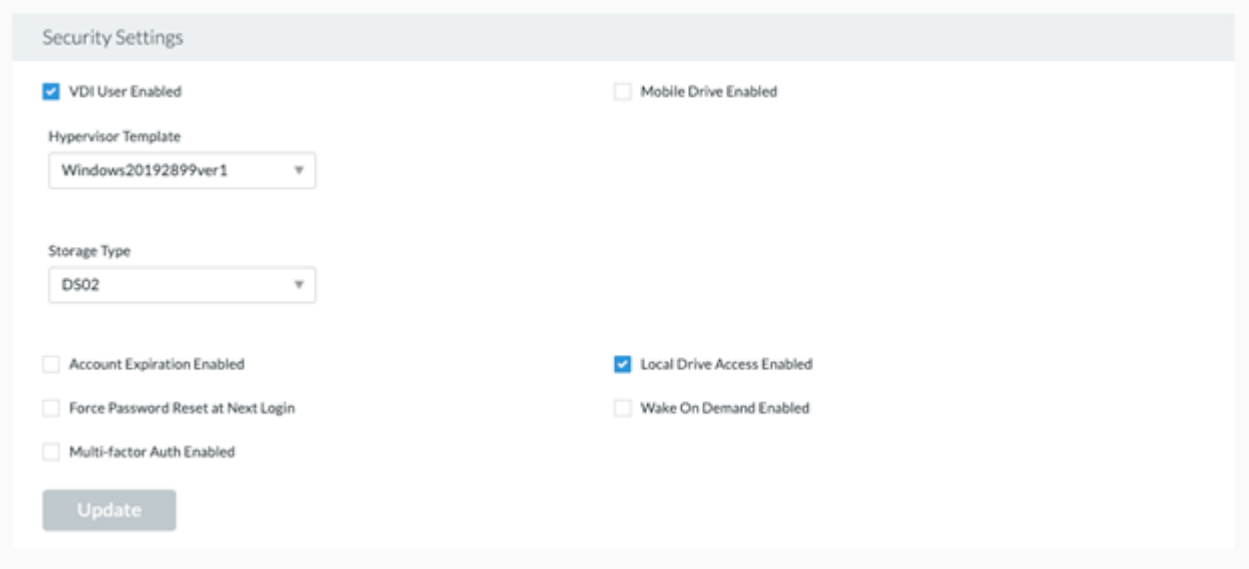
Each workspace resides in its own Active Directory organization unit (OU) under the Cloud Workspace OU as shown in the following figure.



For more info, see this video on user permissions and user management in NetApp VDS.

When an Active Directory group is defined as a CRAUserGroup using an API call for the datacenter, all the users in that group are imported into the CloudWorkspace for management using the UI. As the cloud workspace is enabled for the user, VDS creates user home folders, settings permissions, user properties updates, and so on.

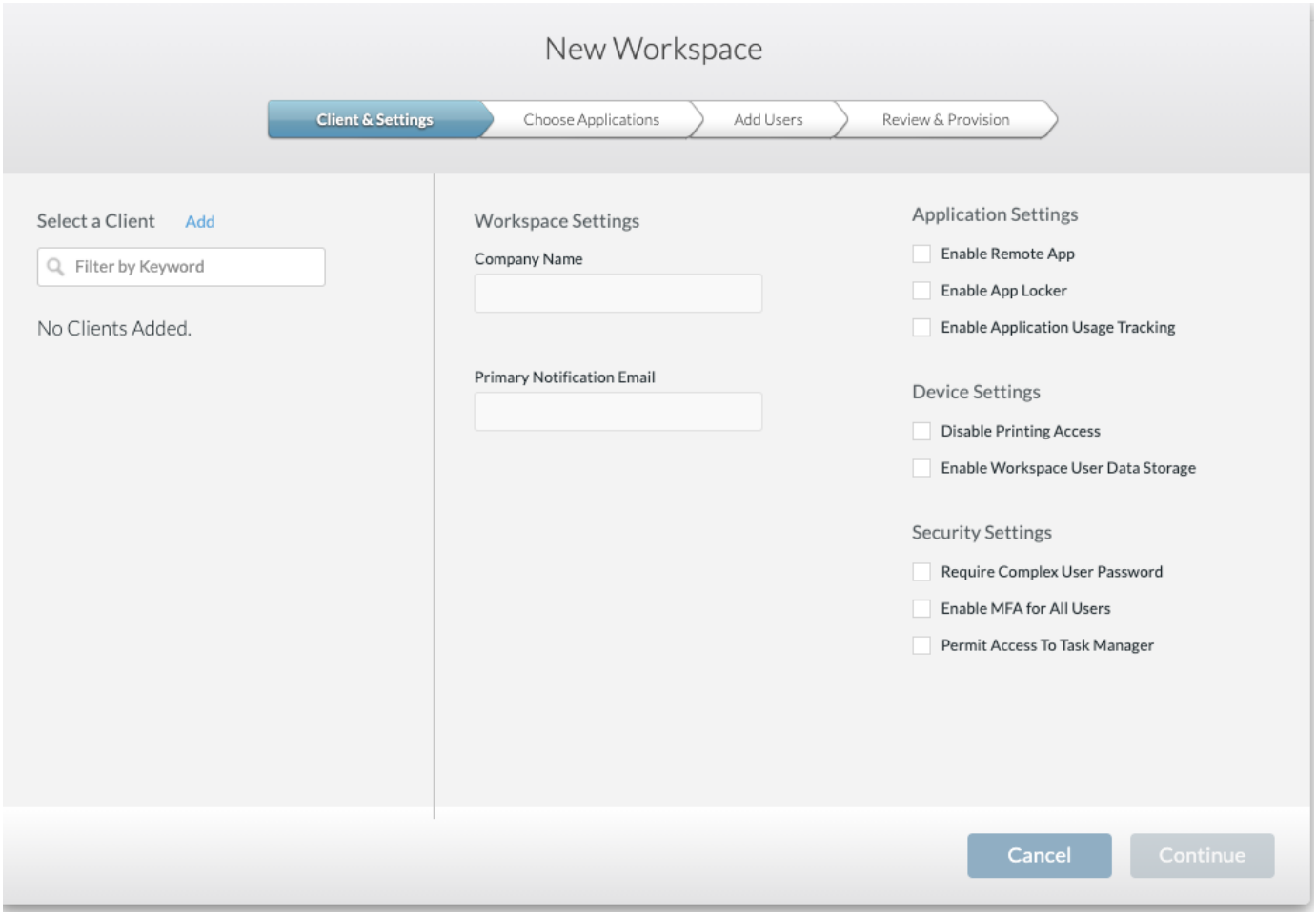
If VDI User Enabled is checked, VDS creates a single-session RDS machine dedicated to that user. It prompts for the template and the datastore to provision.



Next: Workspace Management

Workspace Management

A workspace consists of a desktop environment, which can be shared remote desktop sessions hosted on-premises or on any support cloud environment. With Microsoft Azure, the desktop environment can be persistent with Windows Virtual Desktops. Each workspace is associated with a specific organization or client. Options available when creating a new workspace can be seen in the following figure.



Each workspace is associated with specific deployment.

Workspaces contain associated apps and app services, shared data folders, servers, and a WVD instance. Each workspace can control security options like enforcing password complexity, multifactor authentication, file audits, and so on.

Workspaces can control the workload schedule to power on extra servers, limit the number of users per server, or set the schedule for the resources available for given period (always on/off). Resources can also be configured to wake up on demand.

Workspace can override the deployment VM resource defaults if required. For WVD, WVD host pools (which contains session hosts and app groups) and WVD workspaces can also be managed from the cloud workspace management suite portal. For more info on the WVD Host Pool, see this video.

Next: Application Management

Application Management

Task workers can quickly launch an application from the list of applications made available to them. App services publish applications from the Remote Desktop Services session hosts. With WVD, App Groups provide similar functionality from

multi-session Windows 10 host pools.

For office workers to power users, the applications that they require can be provisioned manually using a service board, or they can be auto-provisioned using the scripted events feature in NetApp VDS.

For more information, see the NetApp Application Entitlement page.

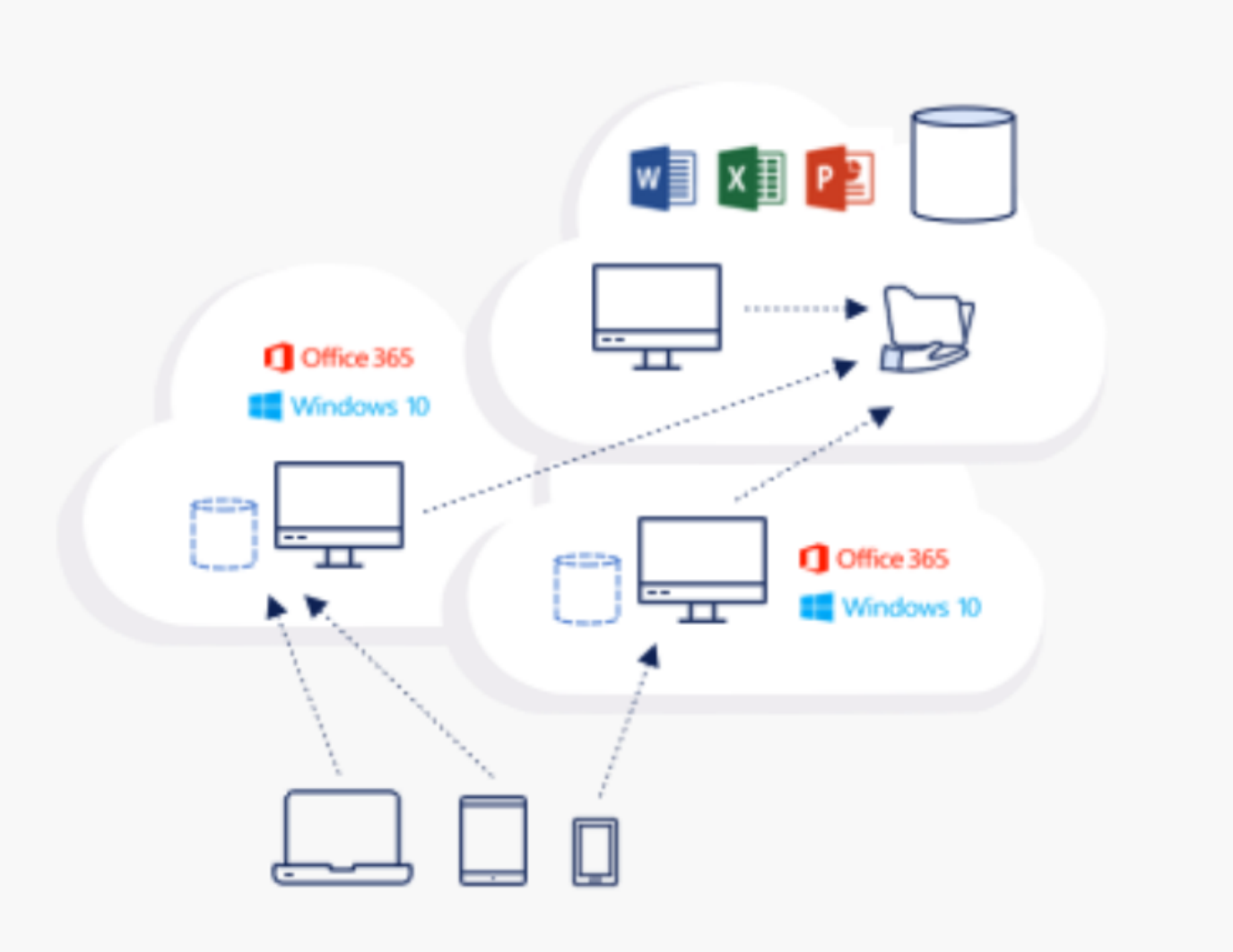
Next: Data Management

Data Management

As a part of deployment, you can choose the file-services method to host the user profile, shared data, and the home drive folder. The available options are File Server, Azure Files, or Azure NetApp Files. However, after deployment, you can modify this choice with the TestvDC tool to point to any SMB share. There are various advantages to hosting with NetApp ONTAP. For more information, see the NetApp Redirecting Storage Platform page.

Global File Cache

When users are spread across multiple sites within a global namespace, Global File Cache can help reduce latency for frequently accessed data. Global File Cache deployment can be automated using a provisioning collection and scripted events. Global File Cache handles the read and write caches locally and maintains file locks across locations. Global File Cache can work with any SMB file servers, including Azure NetApp Files.

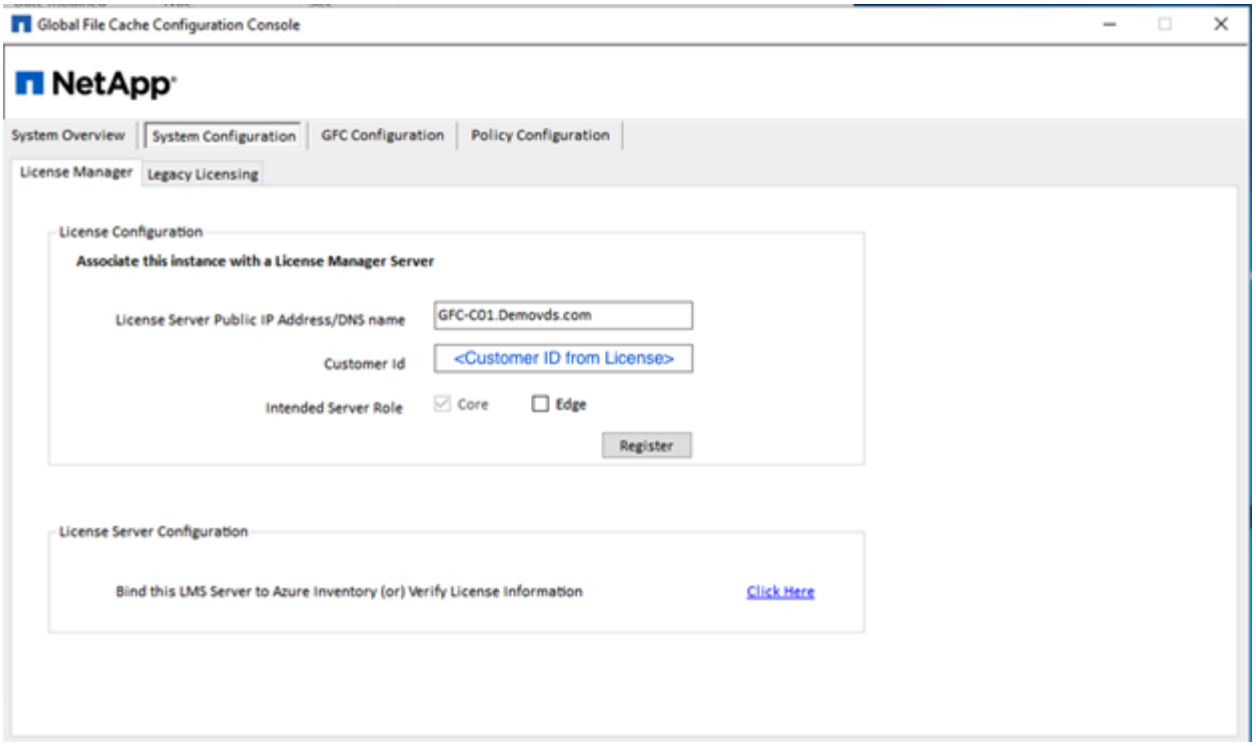


Global File Cache requires the following:

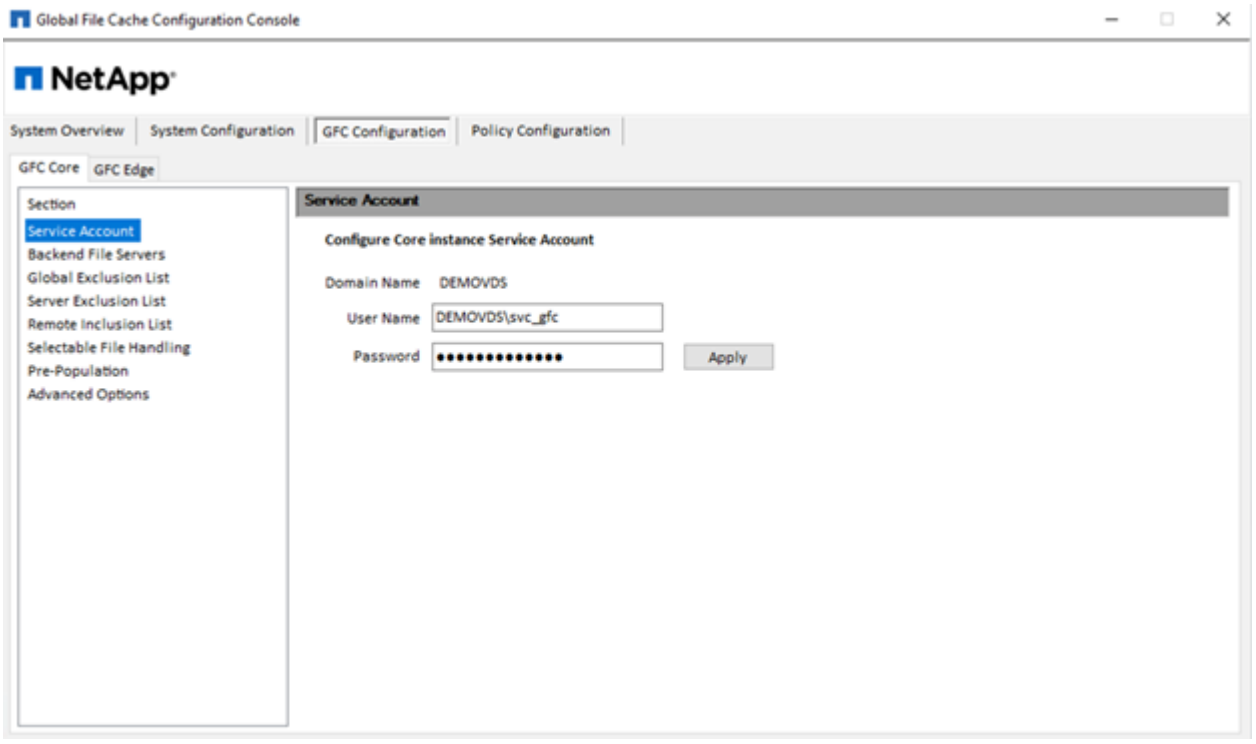
- Management server (License Management Server)
- Core
- Edge with enough disk capacity to cache the data

To download the software and to calculate the disk cache capacity for Edge, see the GFC documentation.

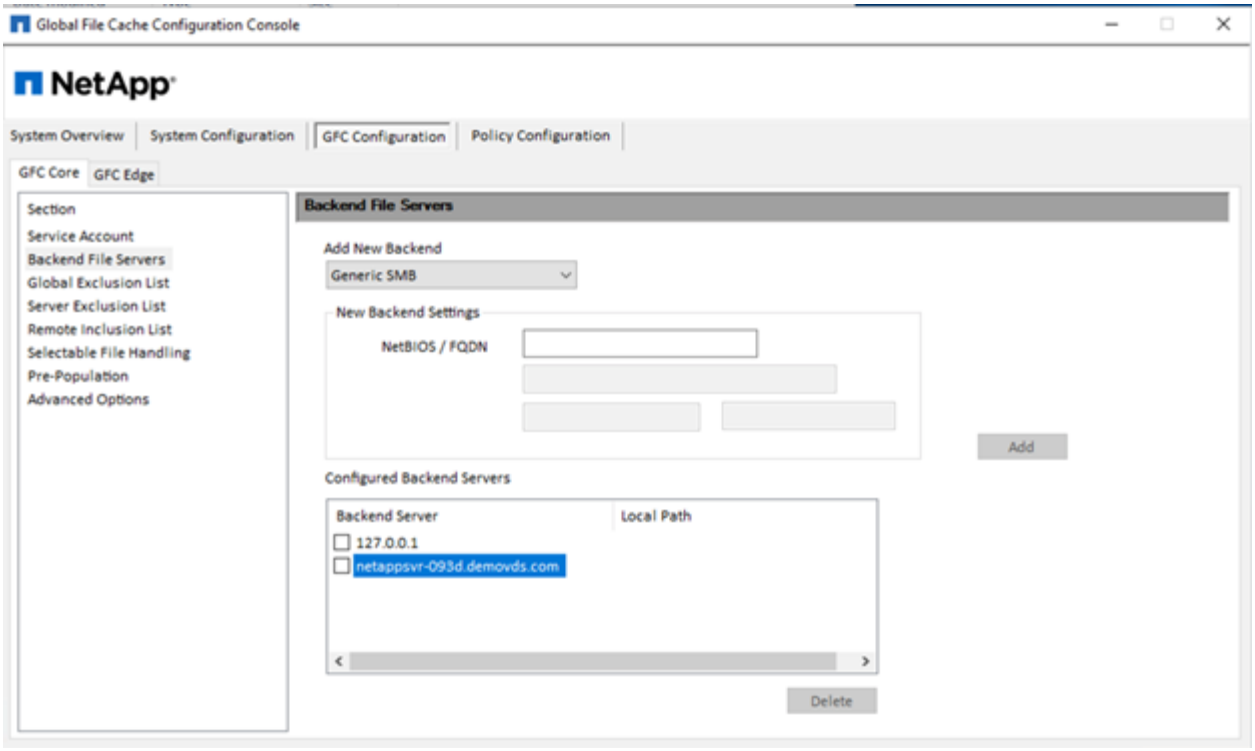
For our validation, we deployed the core and management resources on the same VM at Azure and edge resources on NetApp HCI. Please note that the core is where high-volume data access is required and the edge is a subset of the core. After the software is installed, the license must be activated before use. Under License Configuration section, use the link Click Here to complete the license activation. Then, register the core.



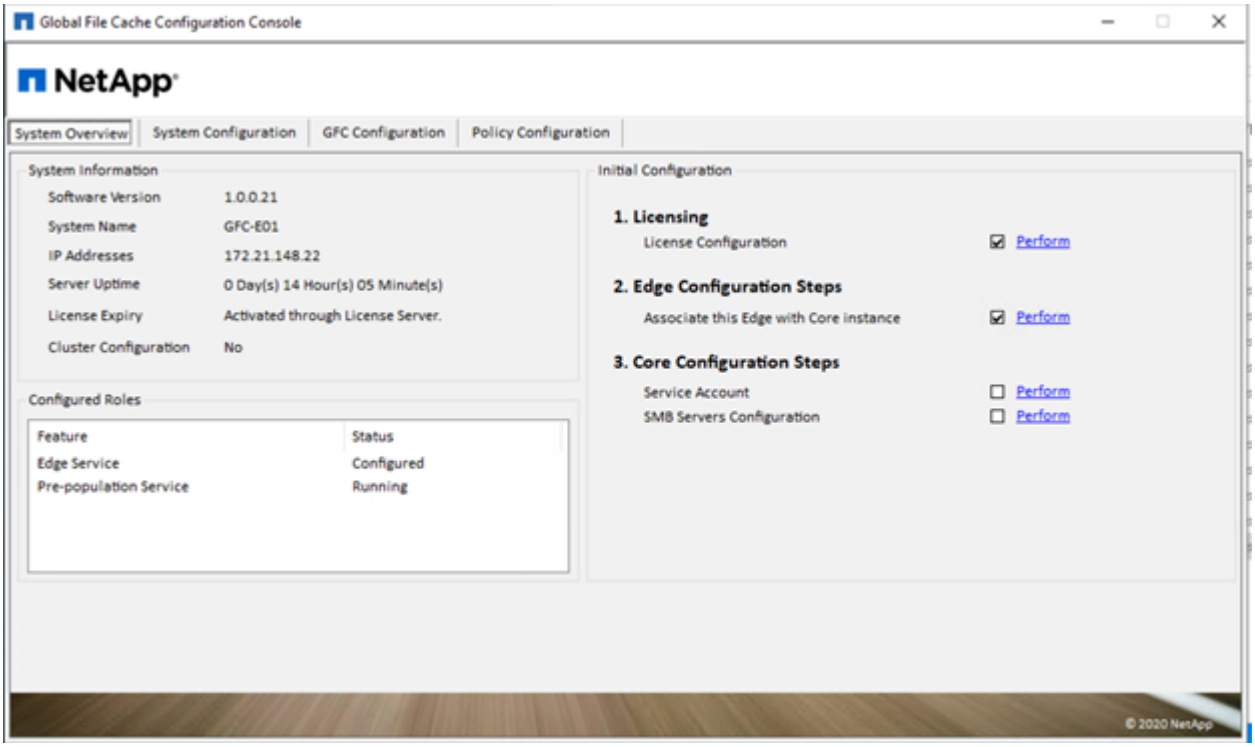
Provide the service account to be used for the Global File Cache. For the required permissions for this account, see the GFC documentation.



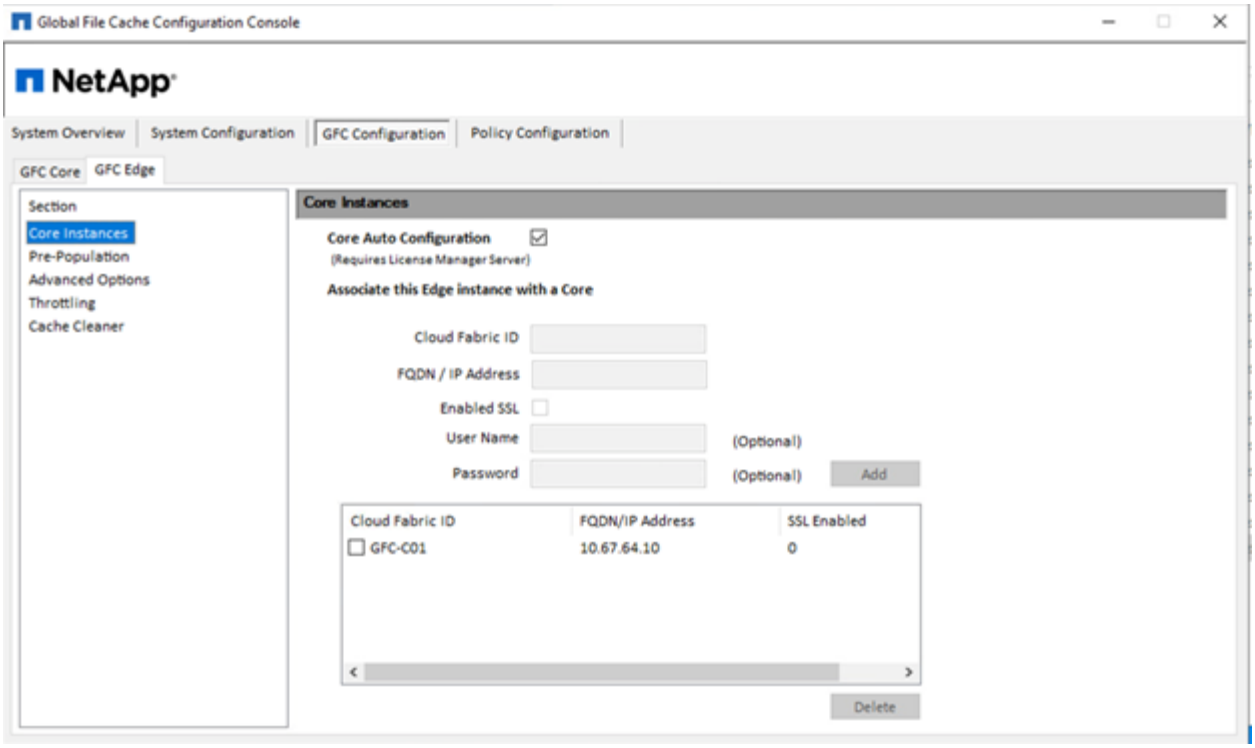
Add a new backend file server and provide the file server name or IP.



On the edge, the cache drive must have the drive letter D. If it does not, use diskpart.exe to select the volume and change drive letter. Register with the license server as edge.

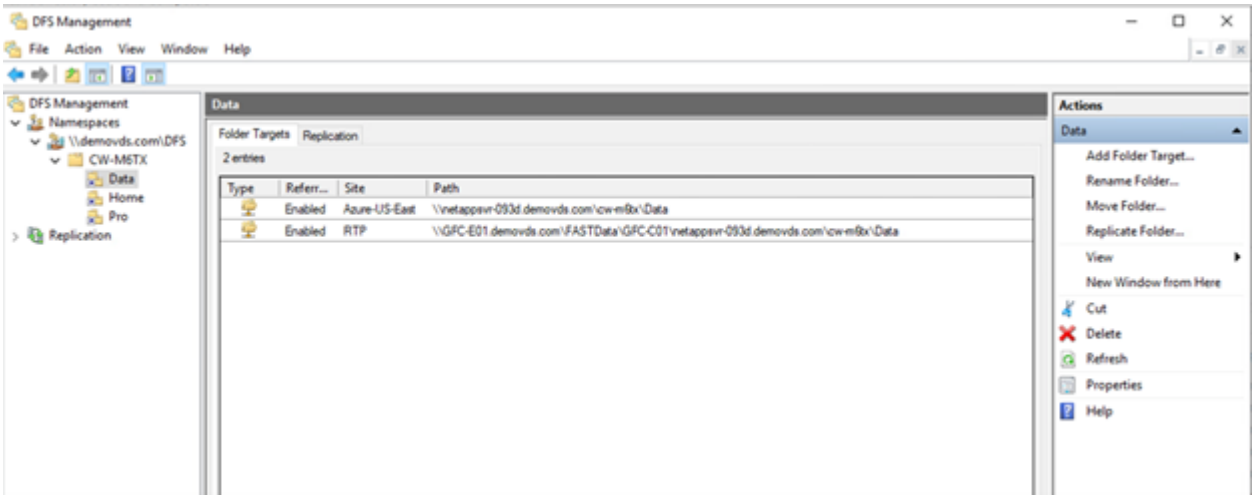


If core auto-configuration is enabled, core information is retrieved from the license management server automatically.

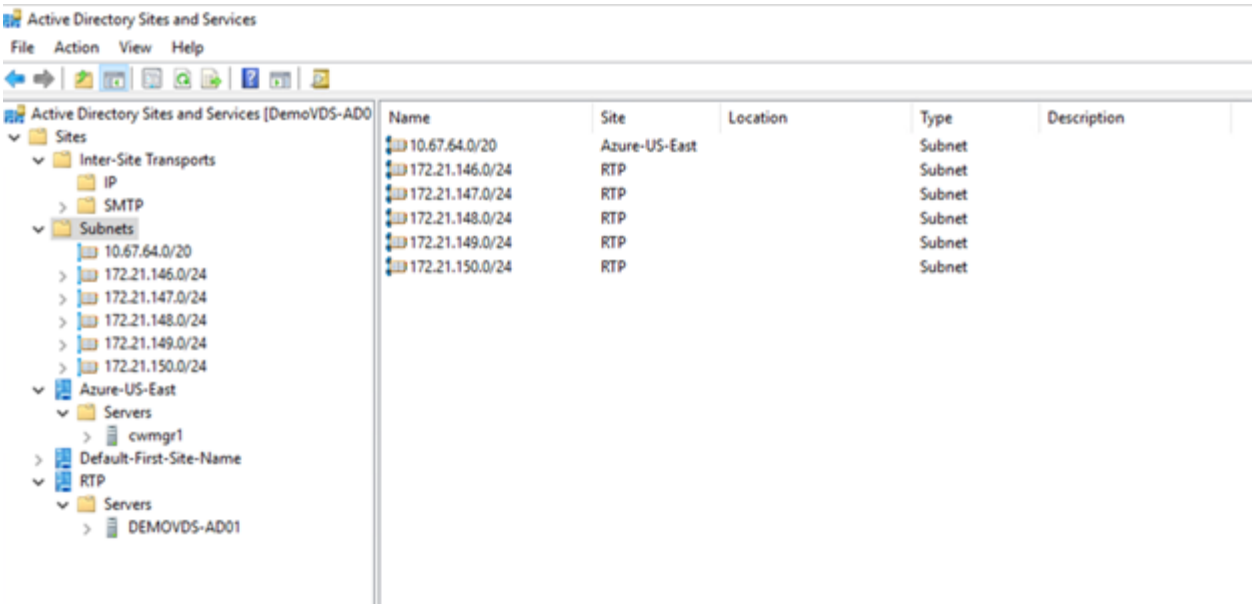


From any client machine, the administrators that use to access the share on file server, can access it via GFC edge using UNC Path `\\<edge server name>\FASTDATA\<core server name>\<backend file server name>\<share name>`. Administrators can include this path in user logonscript or GPO for users drive mapping at the edge location.

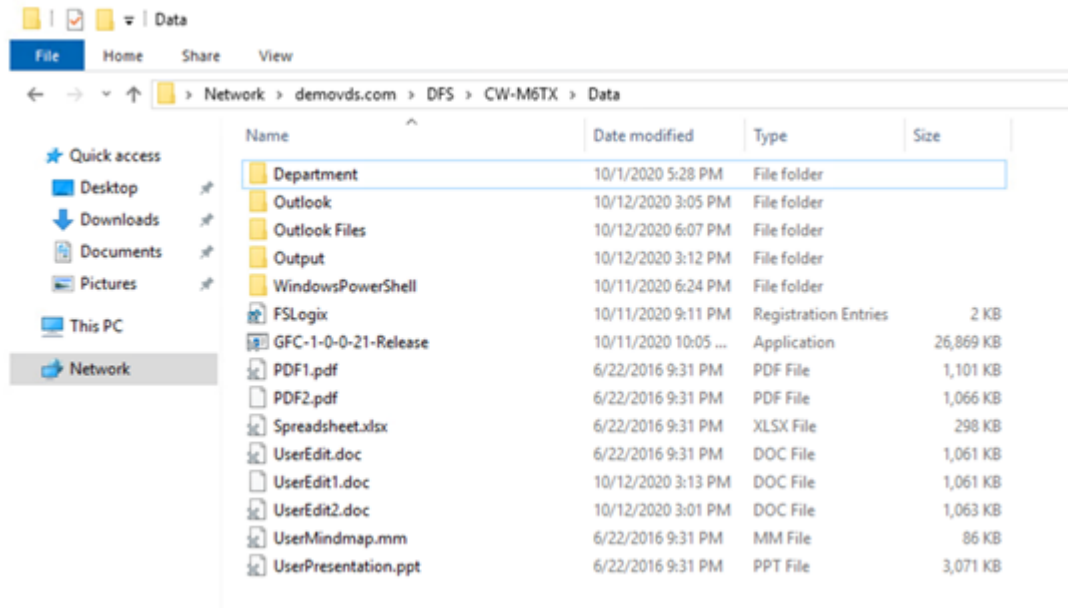
To provide transparent access for users across the globe, an administrator can setup the Microsoft Distributed Filesystem (DFS) with links pointing to file server shares and to edge locations.



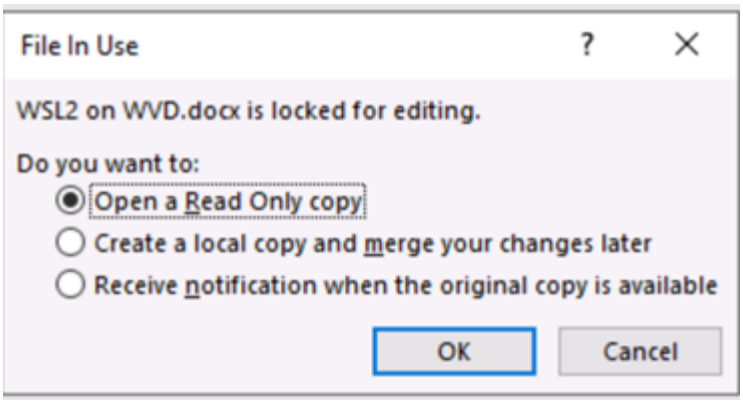
When users log in with Active Directory credentials based on the subnets associated with the site, the appropriate link is utilized by the DFS client to access the data.



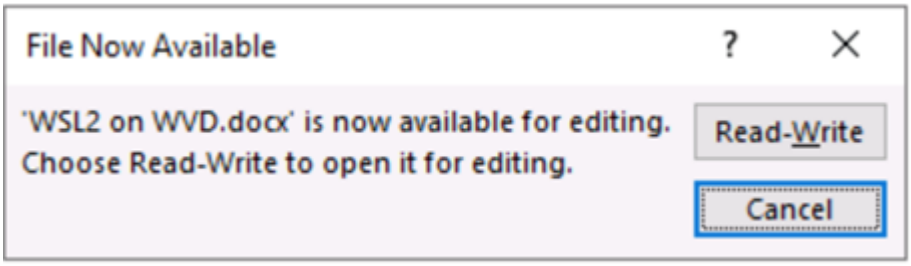
File icons change depending on whether a file is cached; files that are not cached have a grey X on the lower left corner of the icon. After a user in an edge location accesses a file, that file is cached, and the icon changes.



When a file is open and another user is trying to open the same file from an edge location, the user is prompted with the following selection:



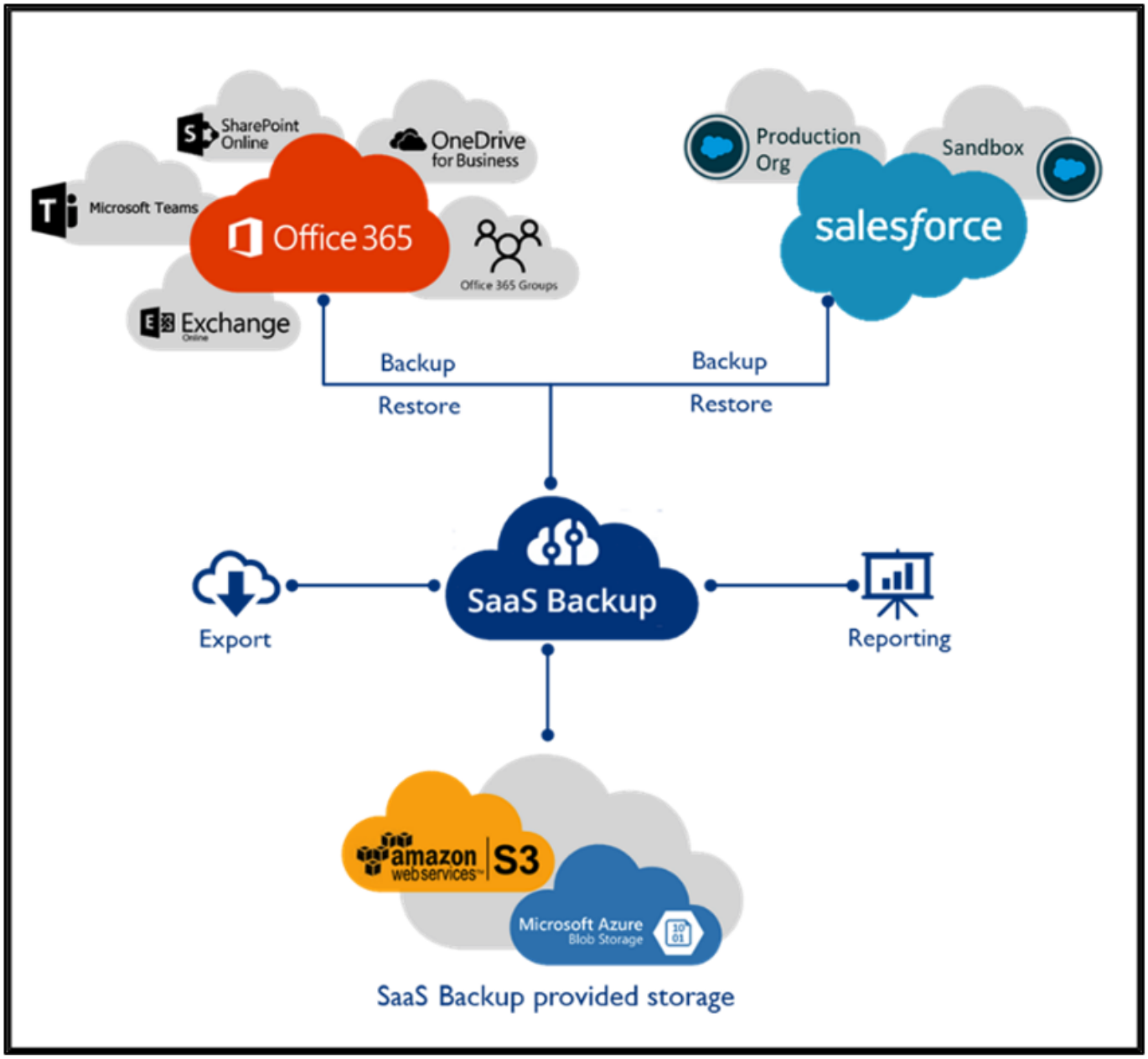
If the user selects the option to receive a notification when the original copy is available, the user is notified as follows:



For more information, see this video on Talon and Azure NetApp Files Deployment.

SaaS Backup

NetApp VDS provides data protection for Salesforce and Microsoft Office 365, including Exchange, SharePoint, and Microsoft OneDrive. The following figure shows how NetApp VDS provides SaaS Backup for these data services.



For a demonstration of Microsoft Office 365 data protection, see this video.

For demonstration of Salesforce data protection, see this video.

Next: Operation Management

Operation Management

With NetApp VDS, administrators can delegate tasks to others. They can connect to deployed servers to troubleshoot, view logs, and run audit reports. While assisting customers, helpdesk or level-3 technicians can shadow user sessions, view process lists, and kill processes if required.

For information on VDS logfiles, see the Troubleshooting Failed VDA Actions page.

For more information on the required minimum permissions, see the VDA Components and

Permissions page.

If you would like to manually clone a server, see the Cloning Virtual Machines page.

To automatically increase the VM disk size, see the Auto-Increase Disk Space Feature page.

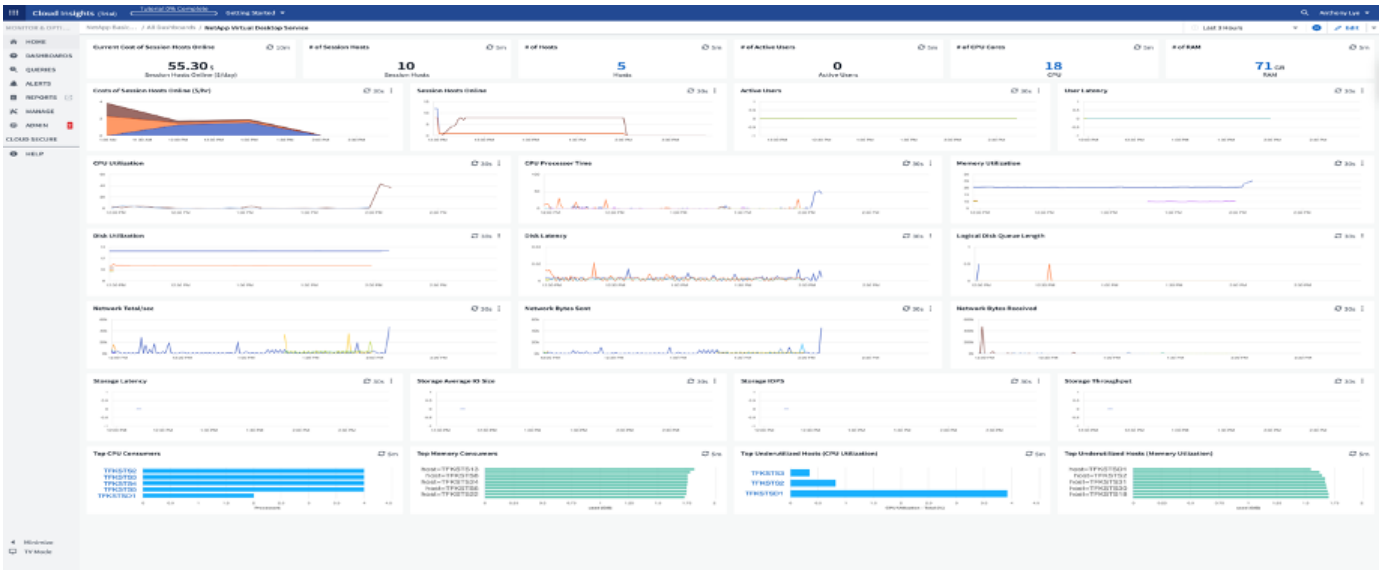
To identify the gateway address to manually configure the client, see the End User Requirements page.

Cloud Insights

NetApp Cloud Insights is a web-based monitoring tool that gives you complete visibility into infrastructure and applications running on NetApp and other third-party infrastructure components. Cloud Insights supports both private cloud and public clouds for monitoring, troubleshooting, and optimizing resources.

Only the acquisition unit VM (can be Windows or Linux) must be installed on a private cloud to collect metrics from data collectors without the need for agents. Agent-based data collectors allow you to pull custom metrics from Windows Performance Monitor or any input agents that Telegraf supports.

The following figure depicts the Cloud Insights VDS dashboard.



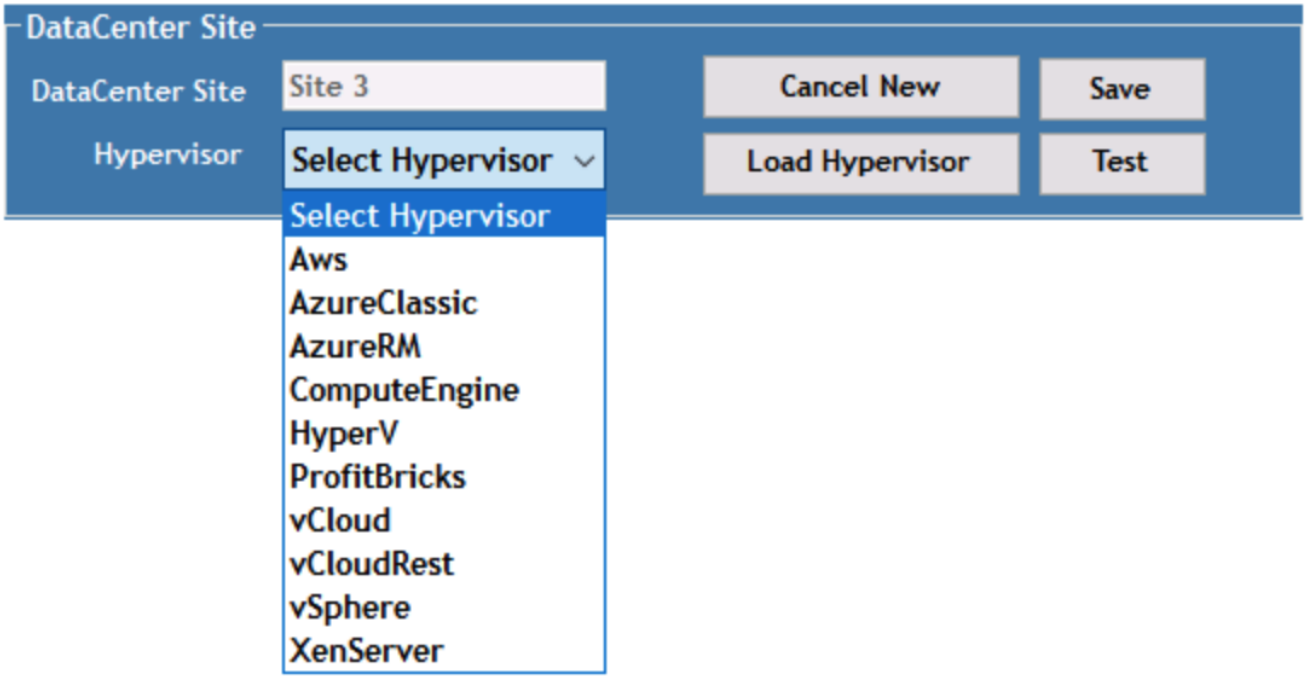
For more info on NetApp Cloud Insights, see this video.

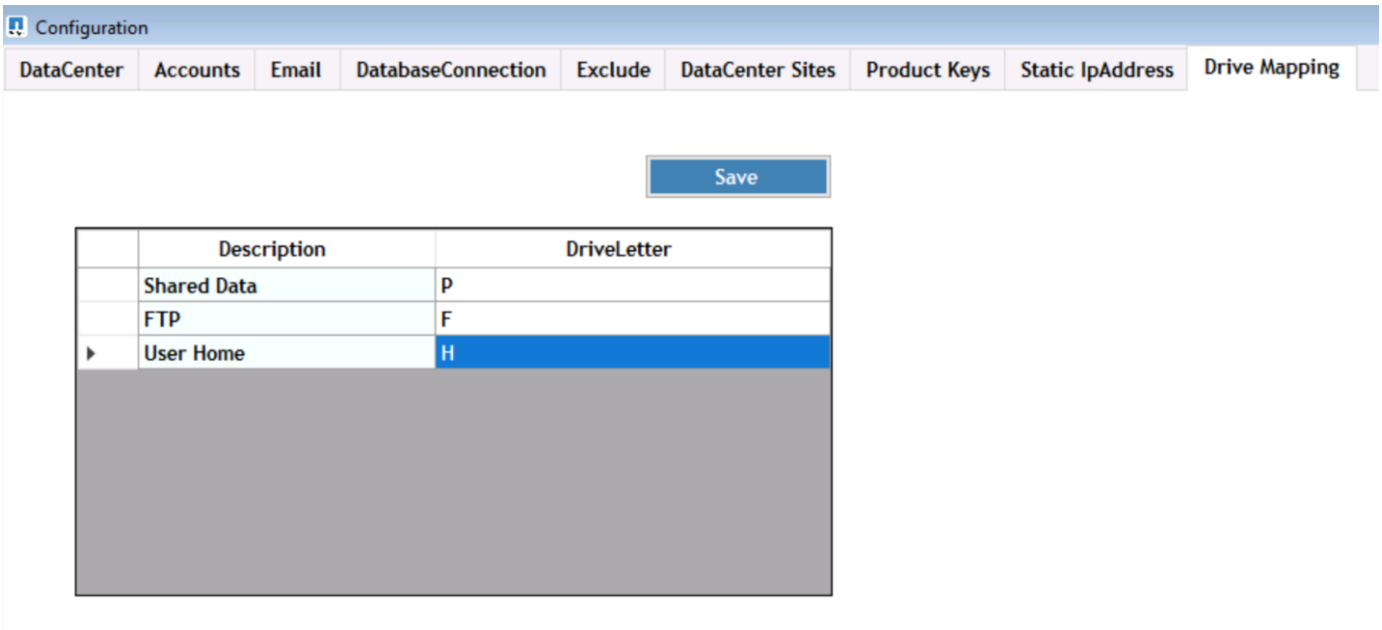
Next: Tools and Logs

Tools and Logs

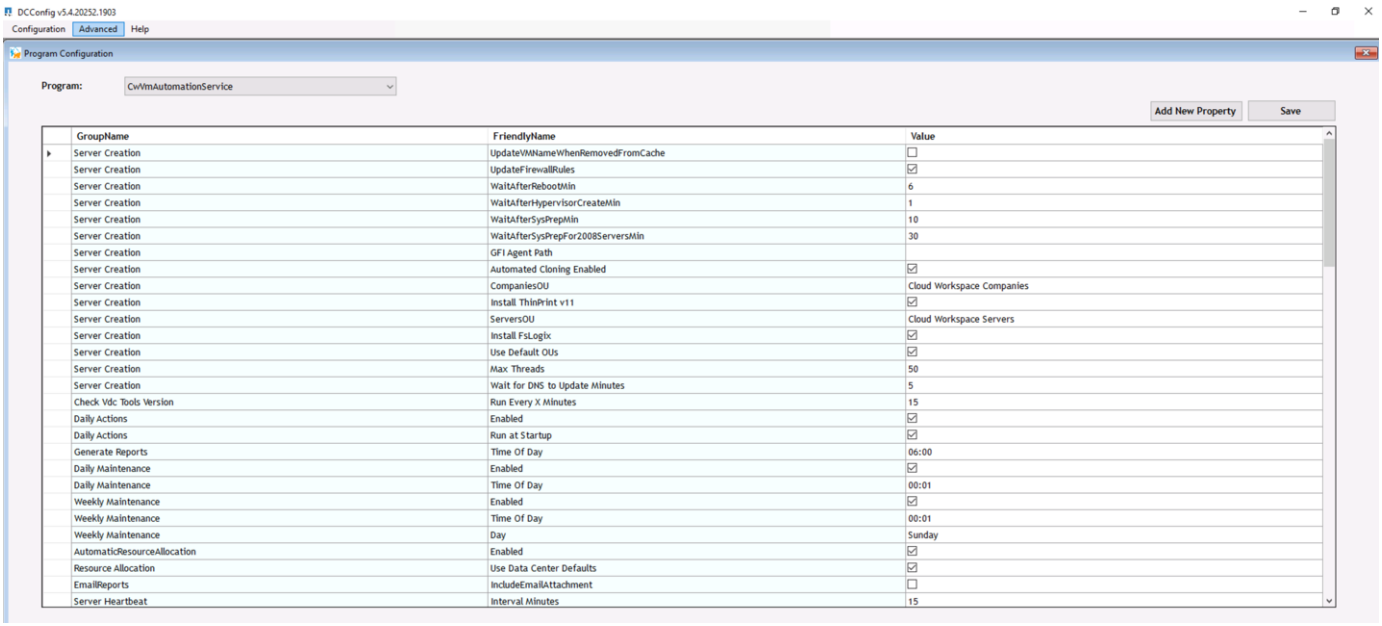
DCConfig Tool

The DCCconfig tool supports the following hypervisor options for adding a site:





Workspace-specific drive letter mapping for shared data can be handled using GPO. Professional Services or the support team can use the advanced tab to customize settings like Active Directory OU names, the option to enable or disable deployment of FSLogix, various timeout values, and so on.



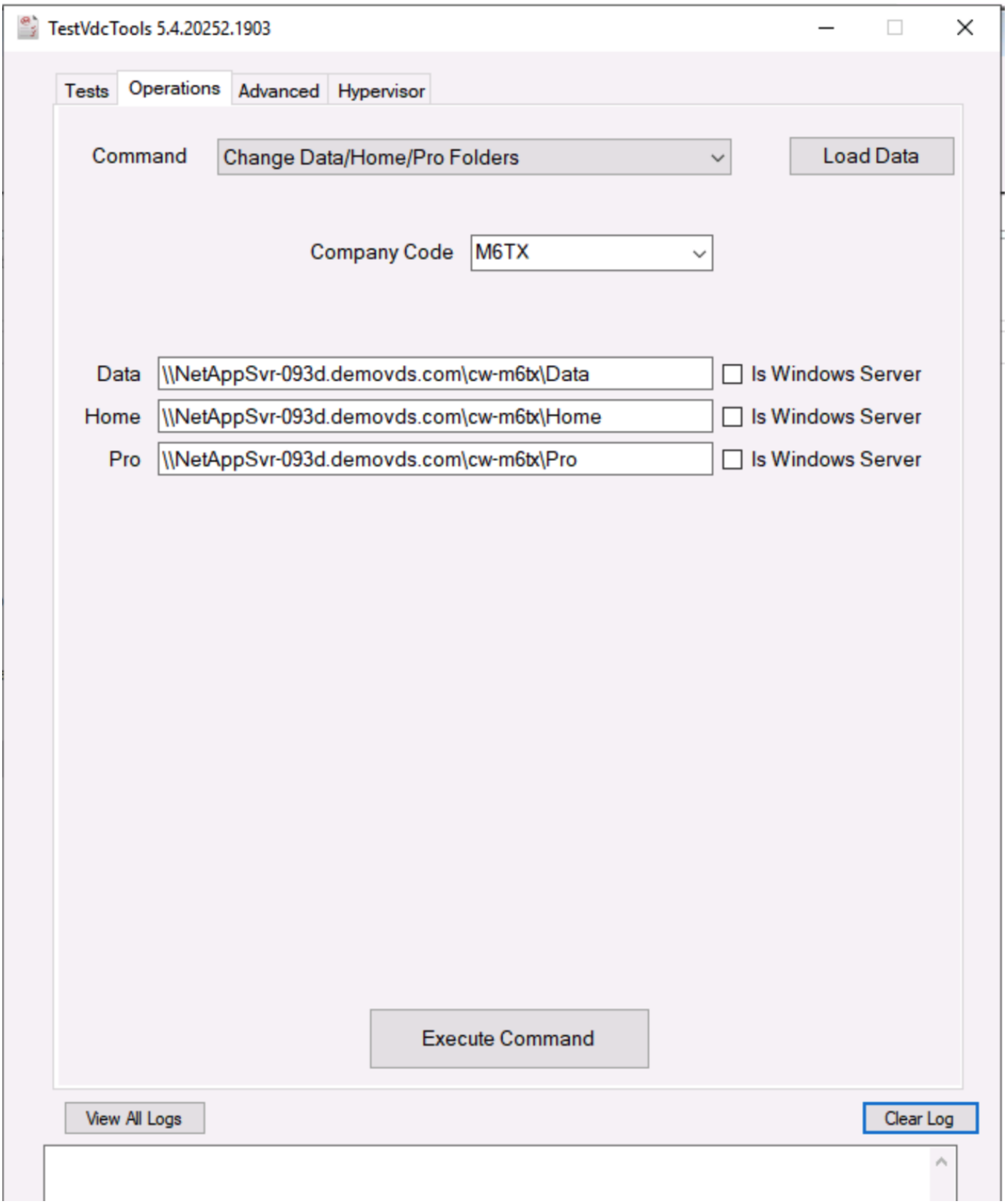
	GroupName	FriendlyName	Value
▸	Server Creation	UpdateVMNameWhenRemovedFromCache	☐
	Server Creation	UpdateFirewallRules	☑
	Server Creation	WaitAfterRebootMin	6
	Server Creation	WaitAfterHypervisorCreateMin	1
	Server Creation	WaitAfterSysPrepMin	10
	Server Creation	WaitAfterSysPrepFor2008ServersMin	30
	Server Creation	GFI Agent Path	
	Server Creation	Automated Cloning Enabled	☑
	Server Creation	CompaniesOU	Cloud Workspace Companies
	Server Creation	Install ThinPrint v11	☑
	Server Creation	ServersOU	Cloud Workspace Servers
	Server Creation	Install FsLogix	☑
	Server Creation	Use Default OUs	☑
	Server Creation	Max Threads	50
	Server Creation	Wait for DNS to Update Minutes	5
	Check Vdc Tools Version	Run Every X Minutes	15
	Daily Actions	Enabled	☑
	Daily Actions	Run at Startup	☑
	Generate Reports	Time Of Day	06:00
	Daily Maintenance	Enabled	☑
	Daily Maintenance	Time Of Day	00:01
	Weekly Maintenance	Enabled	☑
	Weekly Maintenance	Time Of Day	00:01
	Weekly Maintenance	Day	Sunday
	AutomaticResourceAllocation	Enabled	☑
	Resource Allocation	Use Data Center Defaults	☑
	EmailReports	IncludeEmailAttachment	☐
	Server Heartbeat	Interval Minutes	15

TestVdc Tools

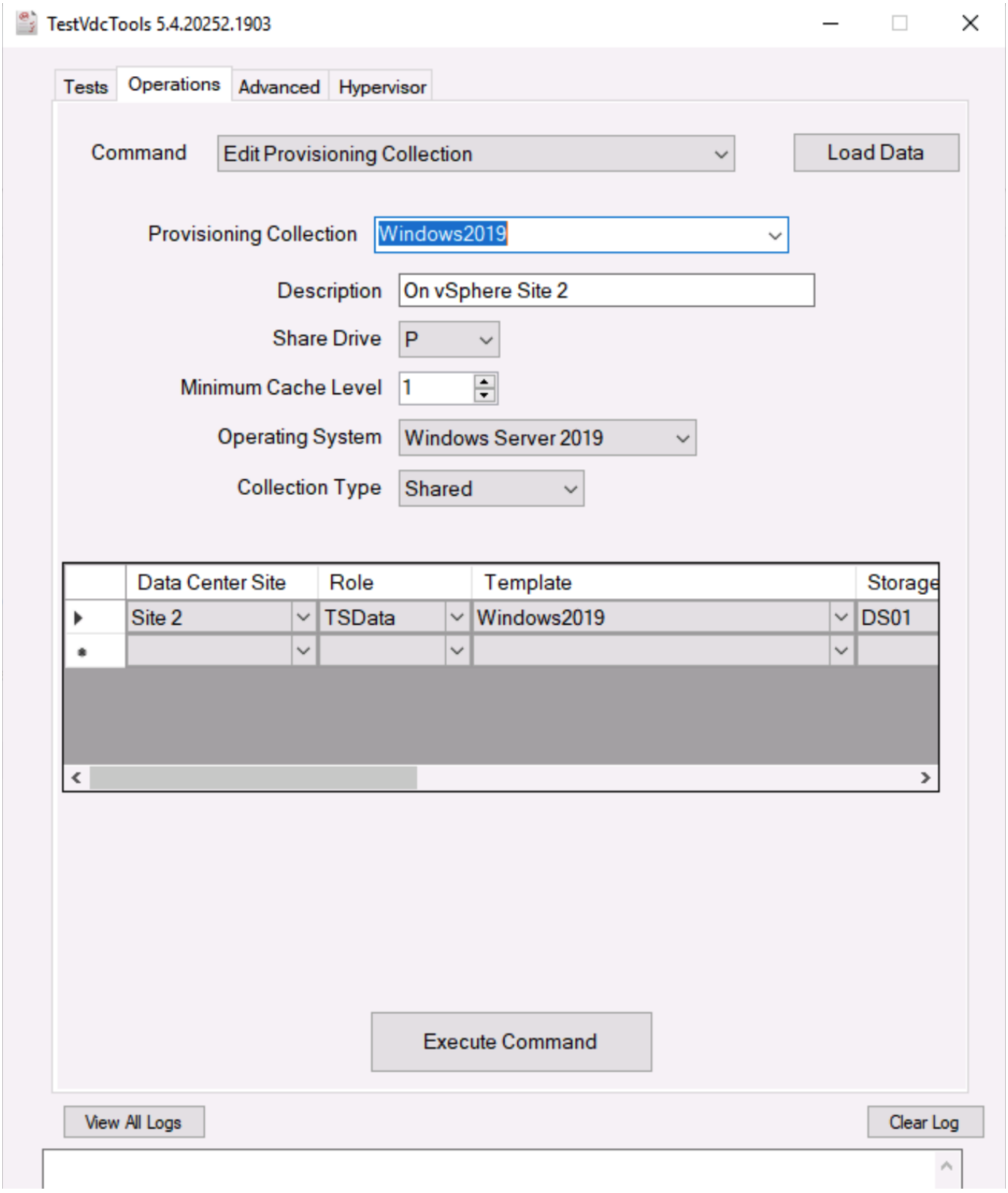
The TestVdc tool is available in the `C:\Program Files\CloudWorkspace\TestVdcTools\` folder.

The following operations can be performed by Professional Services or an administrator:

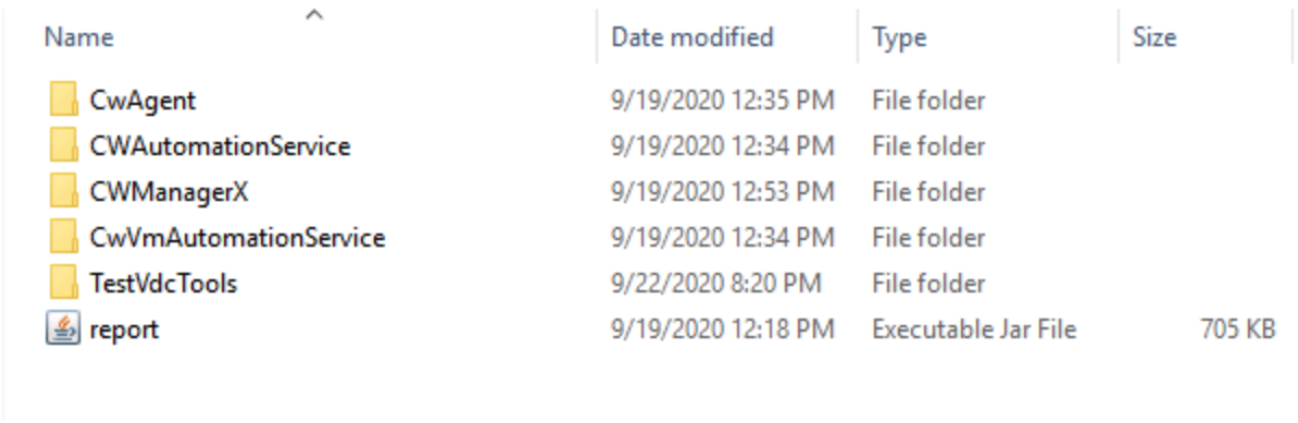
- Change the SMB Path for a workspace.



- Change the site for provisioning collection.


TestVdcTools 5.4.20252.1903
Tests
Operations
Advanced
Hypervisor
Command
Edit Provisioning Collection
Load Data
Provisioning Collection
Windows2019
Description
On vSphere Site 2
Share Drive
P
Minimum Cache Level
1
Operating System
Windows Server 2019
Collection Type
Shared
Data Center Site
Role
Template
Storage
Site 2
TSData
Windows2019
DS01
Execute Command
View All Logs
Clear Log

Log Files



Next: Conclusion

Conclusion

The NetApp Virtual Desktop Service provides an easy-to-consume virtual desktop and application environment with a sharp focus on business challenges. By extending VDS with NetApp HCI, you can use powerful NetApp features in a VDS environment, including in-line deduplication, compaction, thin provisioning, and compression. These features save storage costs and improve performance with all-flash storage. NetApp HCI offers high performance compute, a choice of GPU resources, and with VMware vSphere hypervisor which minimizes the server provisioning time using vSphere API for Array integration. Using the hybrid cloud, customers have the choice to pick the right environment for their demanding workloads and saving expenditure. The desktop session running on-premises can have access to cloud resources based on policy.

Next: Where to Find Additional Information

Where to Find Additional Information

To learn more about the information that is described in this document, review the following documents and/or websites:

- NetApp Cloud

  https://cloud.netapp.com/home

- NetApp VDS Product Documentation

  https://docs.netapp.com/us-en/virtual-desktop-service/index.html

- Connect your on-premises network to Azure with VPN Gateway

https://docs.microsoft.com/en-us/learn/modules/connect-on-premises-network-with-vpn-gateway/

- Azure Portal

  https://portal.azure.com

- Microsoft Windows Virtual Desktop

  https://azure.microsoft.com/en-us/services/virtual-desktop/

- Azure NetApp Files Registration

  https://docs.microsoft.com/en-us/azure/azure-netapp-files/azure-netapp-files-register?WT.mc_id=Portal-Microsoft_Azure_NetApp

End User Computing on NetApp HCI with VMware

Overview of the End User Computing capabilities on NetApp HCI with VMware Horizon.

Learn more about NetApp HCI.

End User Computing (EUC) versus Virtual Desktop Infrastructure (VDI)

Traditionally, the focus on end user computing was centered around the virtualization of desktop infrastructure, or VDI. As VDI evolves, the focus of the conversation has shifted to the accessibility of end user applications and data. To read more about the evolution of VDI and the industry migration to EUC, read the blog discussing the evolution of VDI and infrastructure to EUC and application accessibility.

NetApp Validated Architectures and Technical Reports

End User Computing on NetApp HCI with VMware Horizon is a set of fully validated and supported solutions. Details of the design and deployment considerations are documented in the NetApp Validated Architecture (NVA) documents and Technical Reports (TR).

- NVA: EUC with VMware (Design Guide)
- NVA: EUC with VMware (Deployment Guide)
- NVA: EUC with VMware and NVIDIA GPUs (Design Guide)
- NVA: EUC with VMware and NVIDIA GPUs(Deployment Guide)
- TR: EUC with VMware for 3D Graphics

Additional Material

TR-4854: NetApp HCI for Citrix Virtual Apps and Desktops with Citrix Hypervisor

Suresh Thoppay, NetApp

NetApp HCI infrastructure allows you to start small and build in small increments to meet the demands of virtual desktop users. Compute or storage nodes can be added or removed to address changing business requirements.

Citrix Virtual Apps and Desktops provides a feature-rich platform for end-user computing that addresses various deployment needs, including support for multiple hypervisors. The premium edition of this software includes tools to manage images and user policies.

Citrix Hypervisor (formerly known as Citrix Xen Hypervisor) provides additional features to Citrix Virtual Apps and Desktops compared to running on other hypervisor platforms. The following are key benefits of running on Citrix Hypervisor:

- A Citrix Hypervisor license is included with all versions of Citrix Virtual Apps and Desktops. This licensing helps to reduce the cost of running the Citrix Virtual Apps and Desktops platform.
- Features like PVS Accelerator and Storage Accelerator are only available with Citrix Hypervisor.
- For Citrix solutions, the Citrix Hypervisor is the preferred workload choice.
- Available in Long Term Service Release (LTSR; aligns with Citrix Virtual Apps and Desktops) and Current Release (CR) options.

Abstract

This document reviews the solution architecture for Citrix Virtual Apps and Desktops with Citrix Hypervisor. It provides best practices and design guidelines for Citrix implementation on NetApp HCI. It also highlights multitenancy features, user profiles, and image management.

Solution Overview

Service providers who deliver the Virtual Apps and Desktops service prefer to host it on Citrix Hypervisor to reduce cost and for better integration. The NetApp Deployment Engine (NDE), which performs automated installation of VMware vSphere on NetApp HCI, currently doesn’t support deployment of Citrix Hypervisor. Citrix Hypervisor can be installed on NetApp HCI using PXE boot or installation media or other deployment methods supported by Citrix.

Citrix Virtual Apps and Desktops can automate the provisioning of desktops and session hosts either using Citrix Provisioning (network-based) or by Machine Creation Services (hypervisor storage-based). Both Microsoft Windows-based OSs and popular Linux flavors are supported. Existing physical

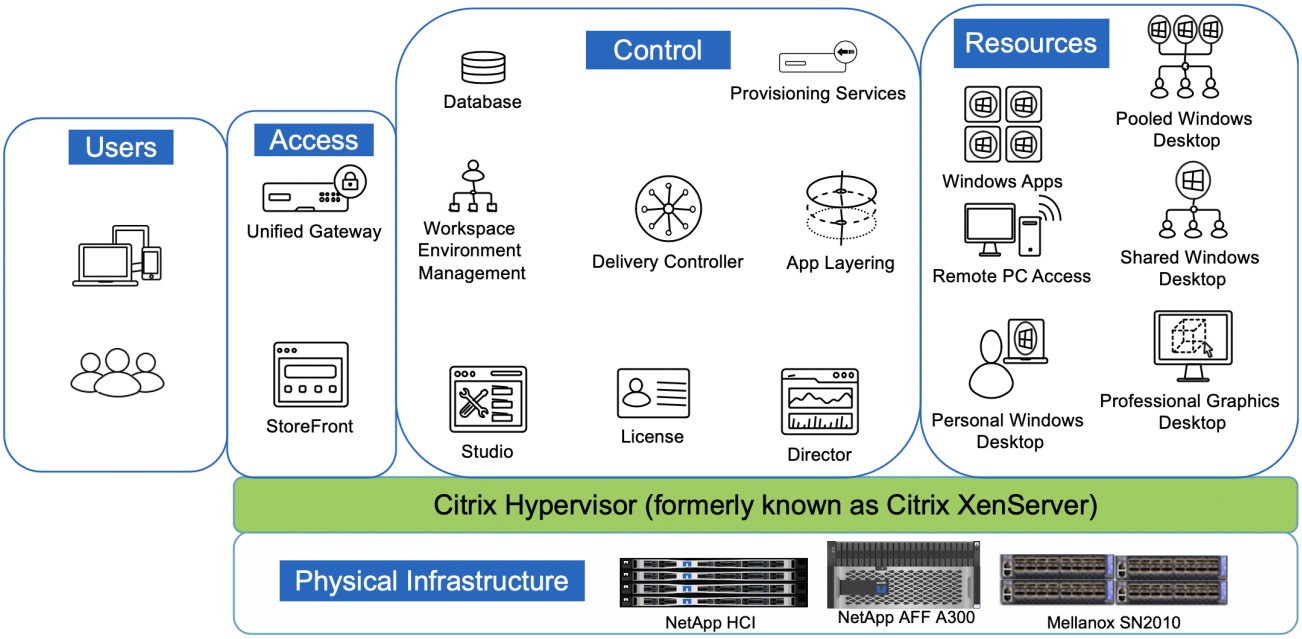
workstations, desktop PCs, and VMs on other hypervisors that are not enabled for auto-provisioning can also be made available for remote access by installing the agents.

The Citrix Workspace Application, a client software used to access Virtual Apps and Desktops, is supported on various devices including tablets and mobile phones. Virtual Apps and Desktops can be accessed using a browser-based HTML5 interface internally or externally to the deployment location.

Based on your business needs, the solution can be extended to multiple sites. However, remember that NetApp HCI storage efficiencies operate on a per-cluster basis.

The following figure shows the high-level architecture of the solution. The access, control, and resource layers are deployed on top of Citrix Hypervisor as virtual machines. Citrix Hypervisor runs on NetApp HCI compute nodes. The virtual disk images are stored in the iSCSI storage repository on NetApp HCI storage nodes.

A NetApp AFF A300 is used in this solution for SMB file shares to store user profiles with FSLogix containers, Citrix profile management (for multisession write-back support), Elastic App Layering images, and so on. We also use SMB file share to mount ISO images on Citrix Hypervisor.



A Mellanox SN2010 switch is used for 10/25/100Gb Ethernet connectivity. Storage nodes use SFP28 transceivers for 25Gb connection, compute nodes use SFP/SFP+ transceivers for 10Gb connection, and interswitch links are QSFP28 transceivers for a 100Gb connection.

Storage ports are configured with multichassis link aggregation (MLAG) to provide total throughput of 50Gb and are configured as trunk ports. Compute node ports are configured as hybrid ports to create a VLAN for iSCSI, XenMotion, and workload VLANs.

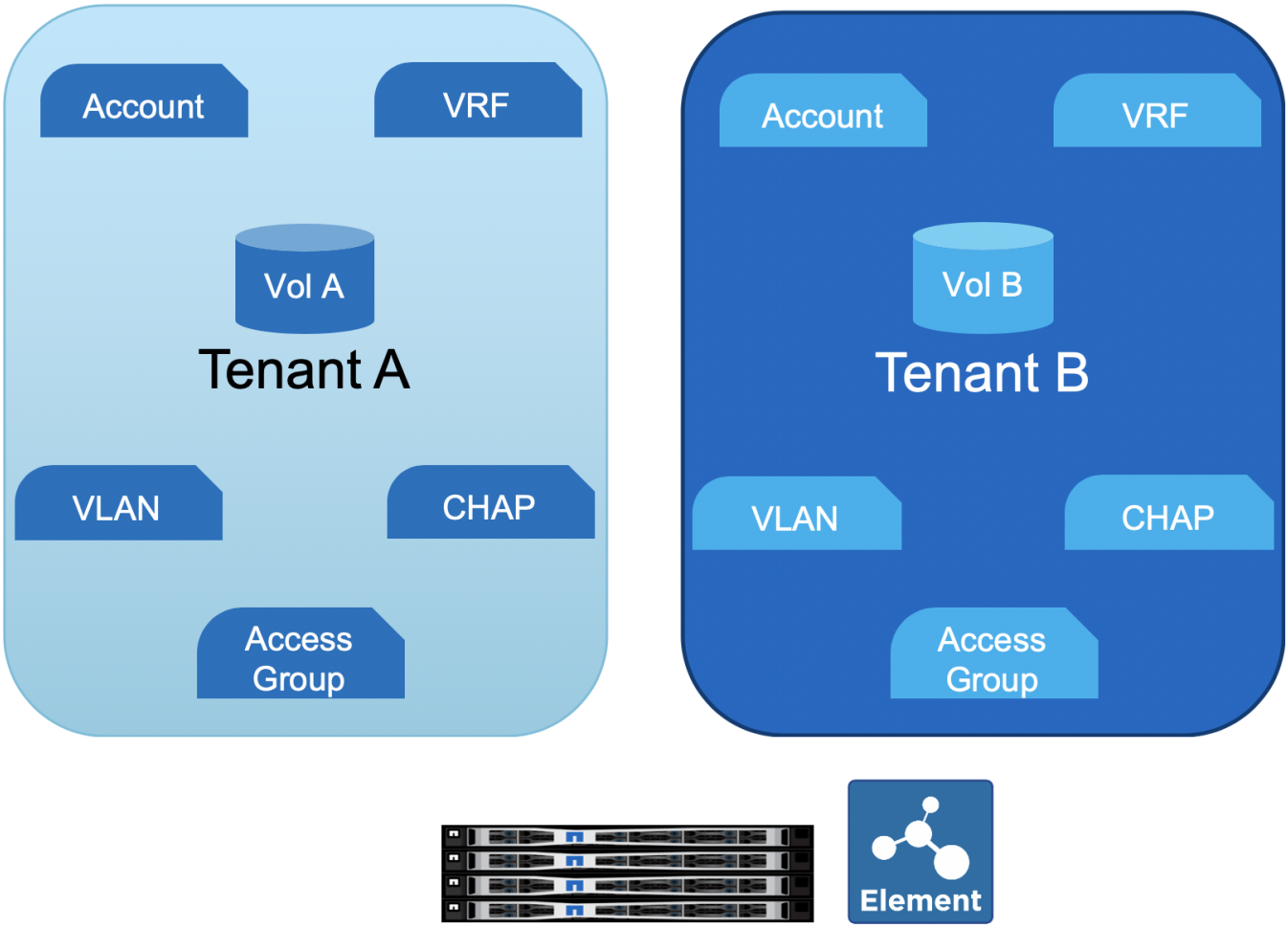
Physical Infrastructure

NetApp HCI

NetApp HCI is available as compute nodes or storage nodes. Depending on the storage node model, a minimum of two to four nodes is required to form a cluster. For the compute nodes, a minimum of two nodes are required to provide high availability. Based on demand, nodes can be added one at a time to increase compute or storage capacity.

A management node (mNode) deployed on a compute node runs as a virtual machine on supported hypervisors. The mNode is used for sending data to ActiveIQ (a SaaS-based management portal), to host a hybrid cloud control portal, as a reverse proxy for remote support of NetApp HCI, and so on.

NetApp HCI enables you to have nondistributive rolling upgrades. Even when one node is down, data is serviced from the other nodes. The following figure depicts NetApp HCI storage multitenancy features.



NetApp HCI Storage provides flash storage through iSCSI connection to compute nodes. iSCSI connections can be secured using CHAP credentials or a volume access group. A volume access group

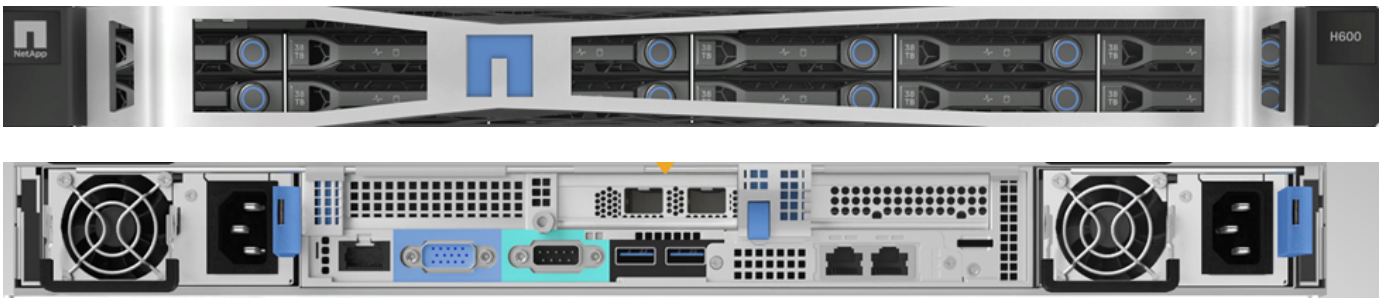
only allows authorized initiators to access the volumes. An account holds a collection of volumes, the CHAP credential, and the volume access group. To provide network-level separation between tenants, different VLANs can be used, and volume access groups also support virtual routing and forwarding (VRF) to ensure the tenants can have same or overlapping IP subnets.

A RESTful web interface is available for custom automation tasks. NetApp HCI has PowerShell and Ansible modules available for automation tasks. For more info, see NetApp.IO.

Storage Nodes

NetApp HCI supports two storage node models: the H410S and H610S. The H410 series comes in a 2U chassis containing four half- width nodes. Each node has six SSDs of sizes 480GB, 960GB, or 1.92TB with the option of drive encryption. The H410S can start with a minimum of two nodes. Each node delivers 50,000 to 100,000 IOPS with a 4K block size. The following figure presents a front and back view of an H410S storage node.

The H610S is a 1U storage node with 12 NVMe drives of sizes 960GB, 1.92TB, or 3.84TB with the option of drive encryption. A minimum of four H610S nodes are required to form a cluster. It delivers around 100,000 IOPS per node with a 4K block size. The following figure depicts a front and back view of an H610S storage node.



In a single cluster, there can be a mix of storage node models. The capacity of a single node can't exceed 1/3 of the total cluster size. The storage nodes come with two network ports for iSCSI (10/25GbE

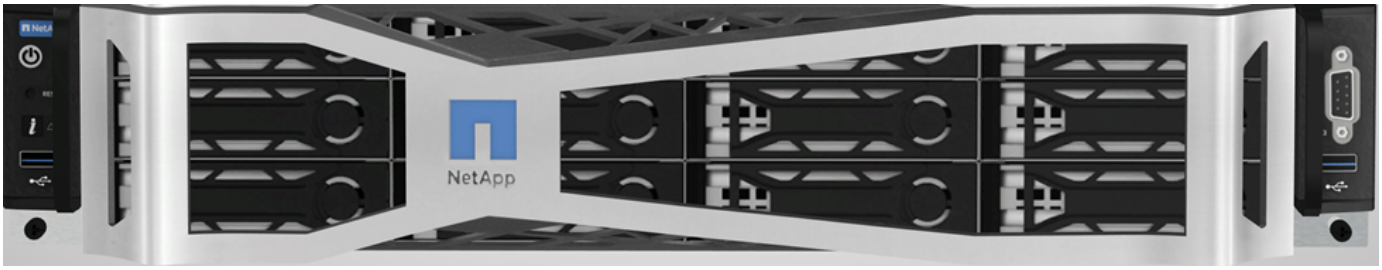
– SFP28) and two ports for management (1/10/GbE – RJ45). A single out-of-band 1GbE RJ45 management port is also available.

Compute Nodes

NetApp HCI compute nodes are available in three models: H410C, H610C, and H615C. Compute nodes are all RedFish API-compatible and provide a BIOS option to enable Trusted Platform Module (TPM) and Intel Trusted eXecution Technology (TXT).

The H410C is a half-width node that can be placed in a 2U chassis. The chassis can have a mix of compute and storage nodes. The H410C comes with first-generation Intel Xeon Silver/Gold scalable processors with 4 to 20 cores in dual-socket configurations. The memory size ranges from 384GB to 1TB. There are four 10/25GbE (SFP28) ports and two 1GbE RJ45 ports, with one 1GbE RJ45 port available for out-of-band management. The following figure depicts a front and back view of an H410C compute node.

The H610C is 2RU and has a dual- socket first generation Intel Xeon Gold 6130 scalable processor with 16 cores of 2.1GHz, 512GB RAM and two NVIDIA Tesla M10 GPU cards. This server comes with two 10/25GbE SFP28 ports and two 1GbE RJ45 ports, with one 1GbE RJ45 port available for out-of-band management. The following figure depicts a front and back view of an H610C compute node.



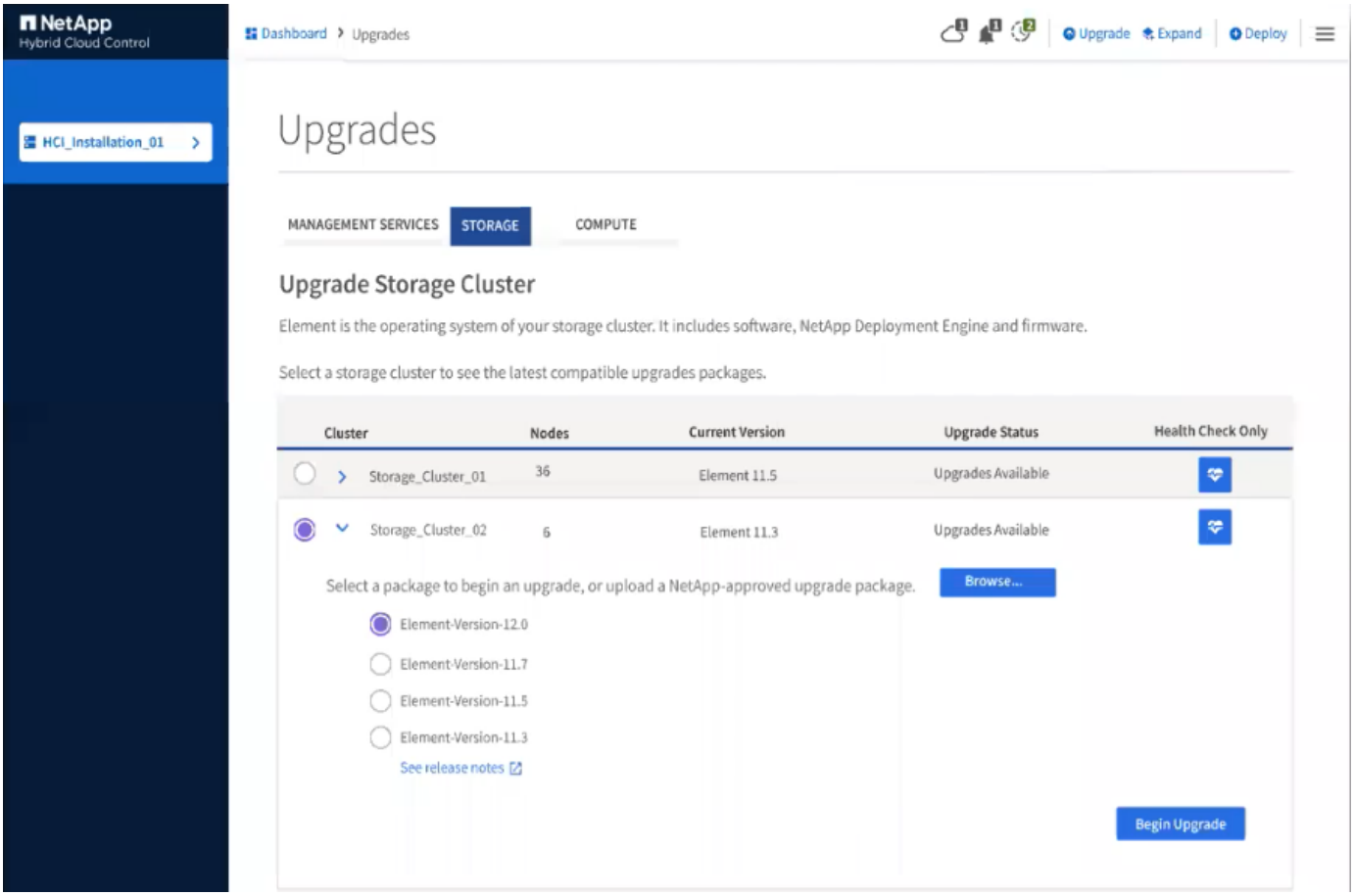
The H610C has two Tesla M10 cards providing a total of 64GB frame buffer memory with a total of 8 GPUs. It can support up to 64 personal virtual desktops with GPU enabled. To host more sessions per server, a shared desktop delivery model is available.

The H615C is a 1RU server with a dual socket for second-generation Intel Xeon Silver/Gold scalable processors with 4 to 24 cores per socket. RAM ranges from 384GB to 1.5TB. One model contains three NVIDIA Tesla T4 cards. The server includes two 10/25GbE (SFP28) and one 1GbE (RJ45) for out-of-band management. The following figure depicts a front and back view of an H615C compute node.

The H615C includes three Tesla T4 cards providing a total of 48GB frame buffer and three GPUs. The T4 card is a general-purpose GPU card that can be used for AI inference workloads as well as for professional graphics. It includes ray tracing cores that can help simulate light reflections.

Hybrid Cloud Control

The Hybrid Cloud Control portal is often used for scaling out NetApp HCI by adding storage or/and compute nodes. The portal provides an inventory of NetApp HCI compute and storage nodes and a link to the ActiveIQ management portal. See the following screenshot of Hybrid Cloud Control.



NetApp AFF

NetApp AFF provides an all-flash, scale-out file storage system, which is used as a part of this solution. ONTAP is the storage software that runs on NetApp AFF. Some key benefits of using ONTAP for SMB file storage are as follows:

- Storage Virtual Machines (SVM) for secure multitenancy
- NetApp FlexGroup technology for a scalable, high-performance file system
- NetApp FabricPool technology for capacity tiering. With FabricPool, you can keeping hot data local and transfer cold data to cloud storage).
- Adaptive QoS for guaranteed SLAs. You can adjusts QoS settings based on allocated or used space.
- Automation features (RESTful APIs, PowerShell, and Ansible modules)
- Data protection and business continuity features including NetApp Snapshot, NetApp SnapMirror, and NetApp MetroCluster technologies

Mellanox Switch

A Mellanox SN2010 switch is used in this solution. However, you can also use other compatible switches. The following Mellanox switches are frequently used with NetApp HCI.

Model	Rack Unit	SFP28 (10/25GbE) ports	QSFP (40/100GbE) ports	Aggregate Throughput (Tbps)
SN2010	Half-width	18	4	1.7
SN2100	Half-width	–	16	3.2
SN2700	Full-width	–	32	6.4

QSFP ports support 4x25GbE breakout cables.

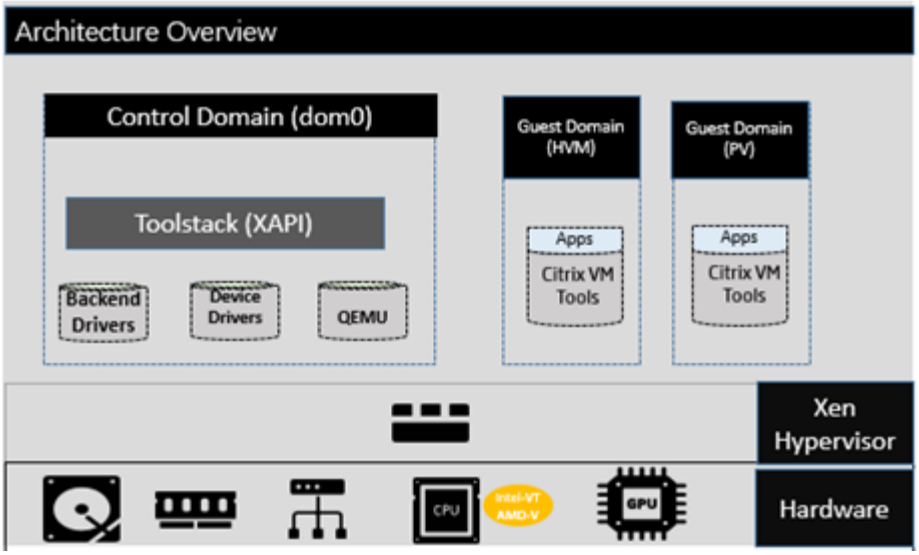
Mellanox switches are open Ethernet switches that allow you to pick the network operating system. Choices include the Mellanox Onyx OS or various Linux OSs such as Cumulus-Linux, Linux Switch, and so on. Mellanox switches also support the switch software development kit, the switch abstraction interface (SAI; part of the Open Compute Project), and Software for Open Networking in the Cloud (SONIC).

Mellanox switches provide low latency and support traditional data center protocols and tunneling protocols like VXLAN. VXLAN Hardware VTEP is available to function as an L2 gateway. These switches support various certified security standards like UC API, FIPS 140-2 (System Secure Mode), NIST 800-181A (SSH Server Strict Mode), and CoPP (IP Filter).

Mellanox switches support automation tools like Ansible, SALT Stack, Puppet, and so on. The Web Management Interface provides the option to execute multi-line CLI commands.

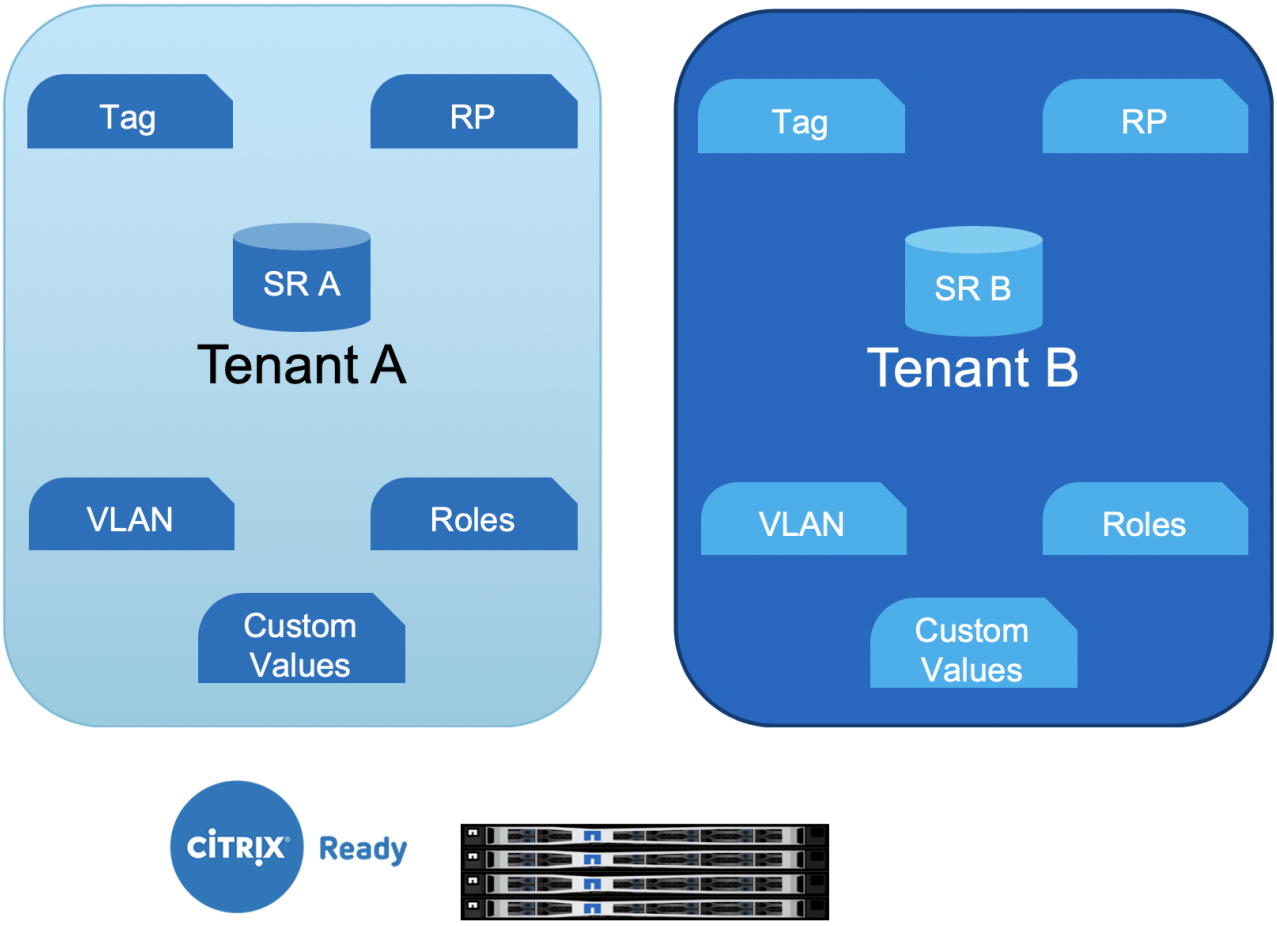
Citrix Hypervisor

Citrix Hypervisor (formerly known as XenServer) is the industry-leading, cost-effective, open-source platform for desktop virtualization infrastructure. XenCenter is a light-weight graphical management interface for Citrix Hypervisor servers. The following figure presents an overview of the Citrix Hypervisor architecture.



Citrix Hypervisor is a type-1 hypervisor. The control domain (also called Domain 0 or dom0) is a secure, privileged Linux VM that runs the Citrix Hypervisor management tool stack known as XAPI. This Linux VM is based on a CentOS 7.5 distribution. Besides providing Citrix Hypervisor management functions, dom0 also runs the physical device drivers for networking, storage, and so on. The control domain can talk to the hypervisor to instruct it to start or stop guest VMs.

Virtual desktops run in the guest domain, sometimes referred as the user domain or domU, and request resources from the control domain. Hardware-assisted virtualization uses CPU virtualization extensions like Intel VT. The OS kernel doesn't need to be aware that it is running on a virtual machine. Quick Emulator (QEMU) is used for virtualizing the BIOS, the IDE, the graphic adapter, USB, the network adapter, and so on. With paravirtualization (PV), the OS kernel and device drivers are optimized to boost performance in the virtual machine. The following figure presents multitenancy features of Citrix Hypervisor.



Resources from NetApp HCI makes up the hardware layer, which includes compute, storage, network, GPUs, and so on.

Compute

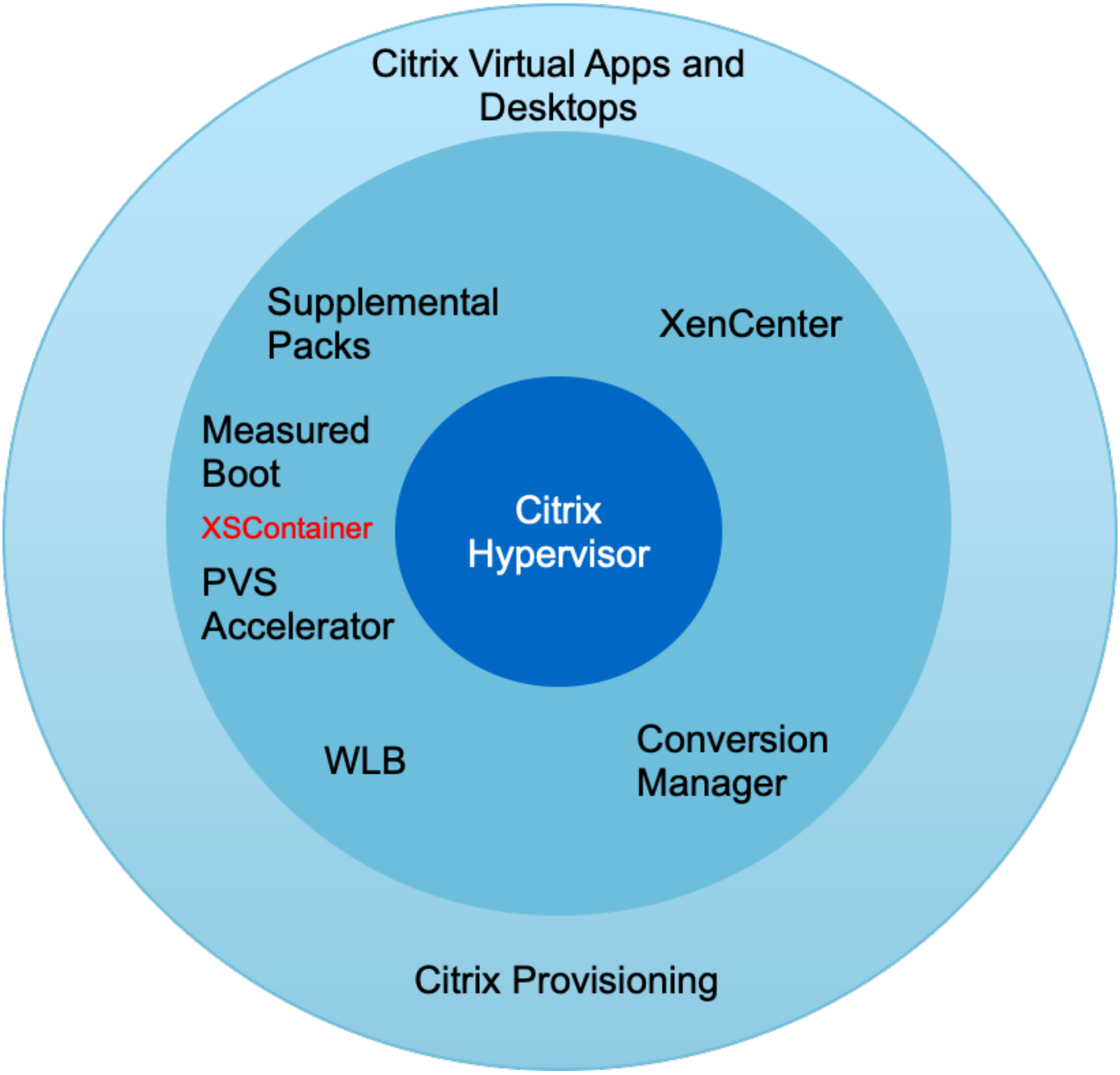
The CPU and memory details of NetApp HCI are covered in the previous section. However, this section focuses on how the compute node is utilized in the Citrix Hypervisor environment.

Each NetApp HCI compute node with Citrix Hypervisor installed is referred as a server. A pool of servers is managed as a resource pool (RP). The resource pools are created with similar model compute nodes to provide similar performance when the workload is moved from one node to another. A resource pool always contains a node designated as master, which exposes the management interface (for XenCenter and the CLI) and which can be routed to other member servers as necessary. When high availability is enabled, master re-election takes place if the master node goes down.

A resource pool can have up to 64 servers (soft limit). However, when clustering is enabled with the GFS2 shared storage resource, the number of servers is restricted to 16.

The resource pool picks a server for hosting the workload and can be migrated to other server using the Live Migration feature. To load balance across the resource pool, the optional WLB management

pack must be installed on Citrix Hypervisor.



Each tenant resource can be hosted on dedicated resource pools or can be differentiated with tags on the same resource pool. Custom values can be defined for operational and reporting purpose.

Storage

NetApp HCI compute nodes have local storage that is not recommended for the storage of any persistent data. Such data should be stored on an iSCSI volume created with NetApp HCI storage or can be on NFS datastore on NetApp AFF.

To use NetApp HCI storage, iSCSI must be enabled on Citrix Hypervisor servers. Using the iQN, register the initiators and create access groups on the Element management portal. Create the volumes

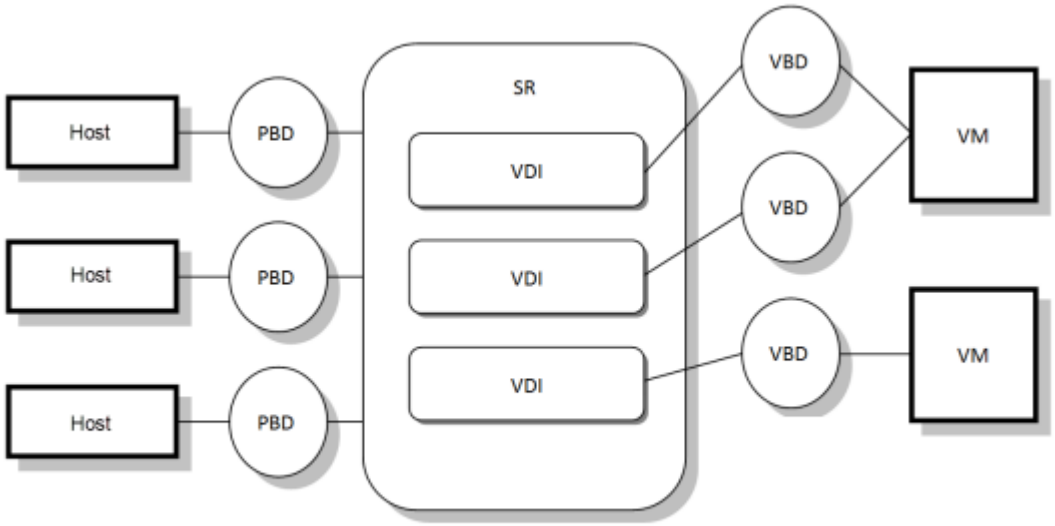
(remember to enable 512e block size support for LVM over iSCSI SR) and assign the account ID and access group.

The iSCSI initiator can be customized using the following command on the CLI:

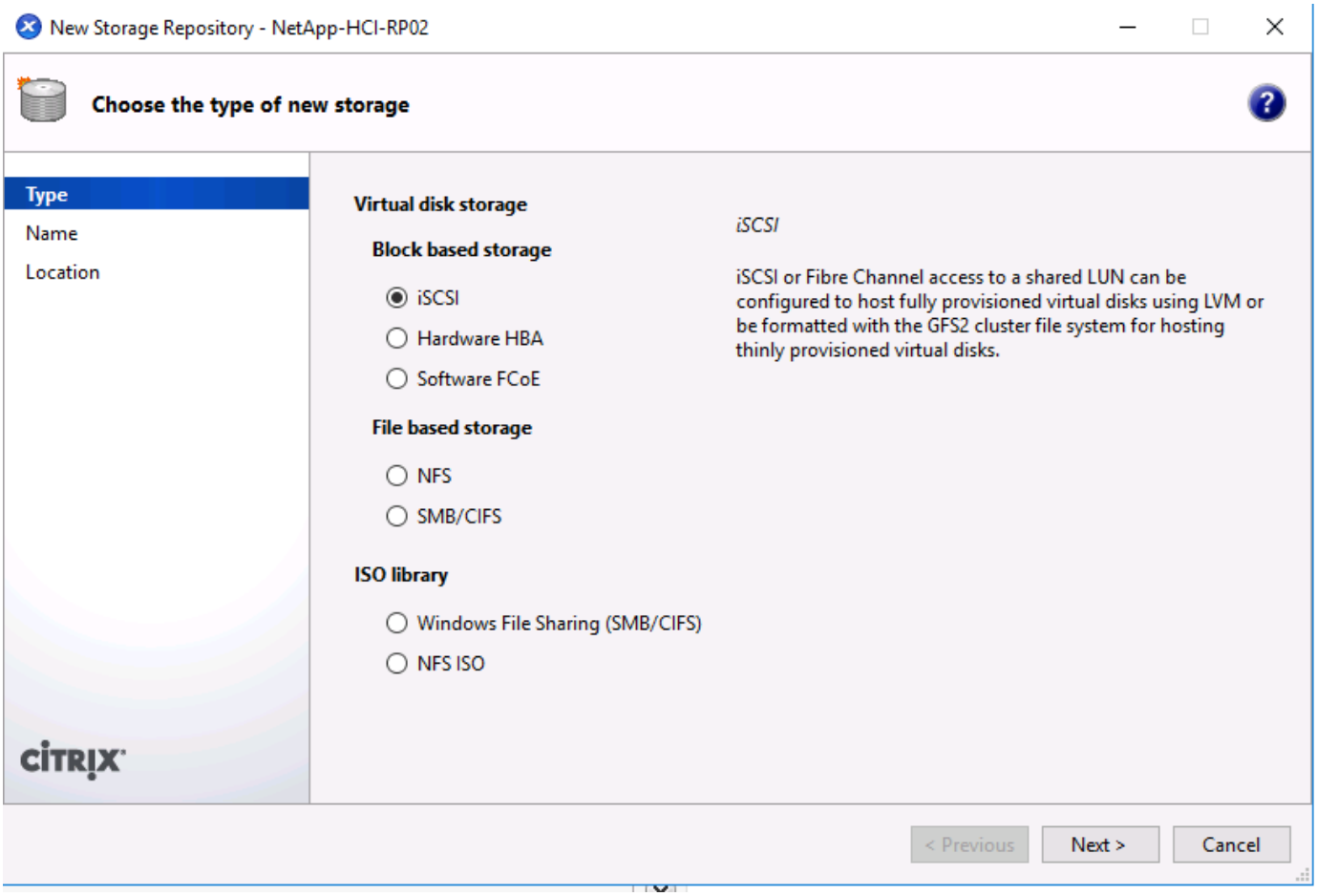
```
xe host-param-set uuid=valid_host_id other-config:iscsi_iqn=new_initiator_iqn
```

Multipathing of iSCSI is supported when multiple iSCSI NICs are configured. iSCSI configuration is performed using XenCenter or by using CLI commands like iscsiadm and multipath. This configuration can also be performed with the various Citrix Hypervisor CLI tools. For iSCSI multipath for single target storage arrays, see CTX138429.

A storage repository (SR) is the storage target in which virtual machine (VM) virtual disk images (VDIs) are stored. A VDI is a storage abstraction that represents a virtual hard disk drive (HDD). The following figure depicts various Citrix Hypervisor storage objects.



The relationship between the SR and host is handled by a physical block device (PBD), which stores the configuration information required to connect and interact with the given storage target. Similarly, a virtual block device (VBD) maintains the mapping between VDIs and a VM. Apart from that, a VBD is also used for fine tuning the quality of service (QoS) and statistics for a given VDI. The following screenshot presents Citrix Hypervisor storage repository types.



With NetApp HCI, the following SR types can be created. The following table provides a comparison of features.

Feature	LVM over iSCSI	GFS2
Maximum virtual disk image size	2TiB	16TiB
Disk provisioning method	Thick Provisioned	Thin Provisioned
Read-caching support	No	Yes
Clustered pool support	No	Yes

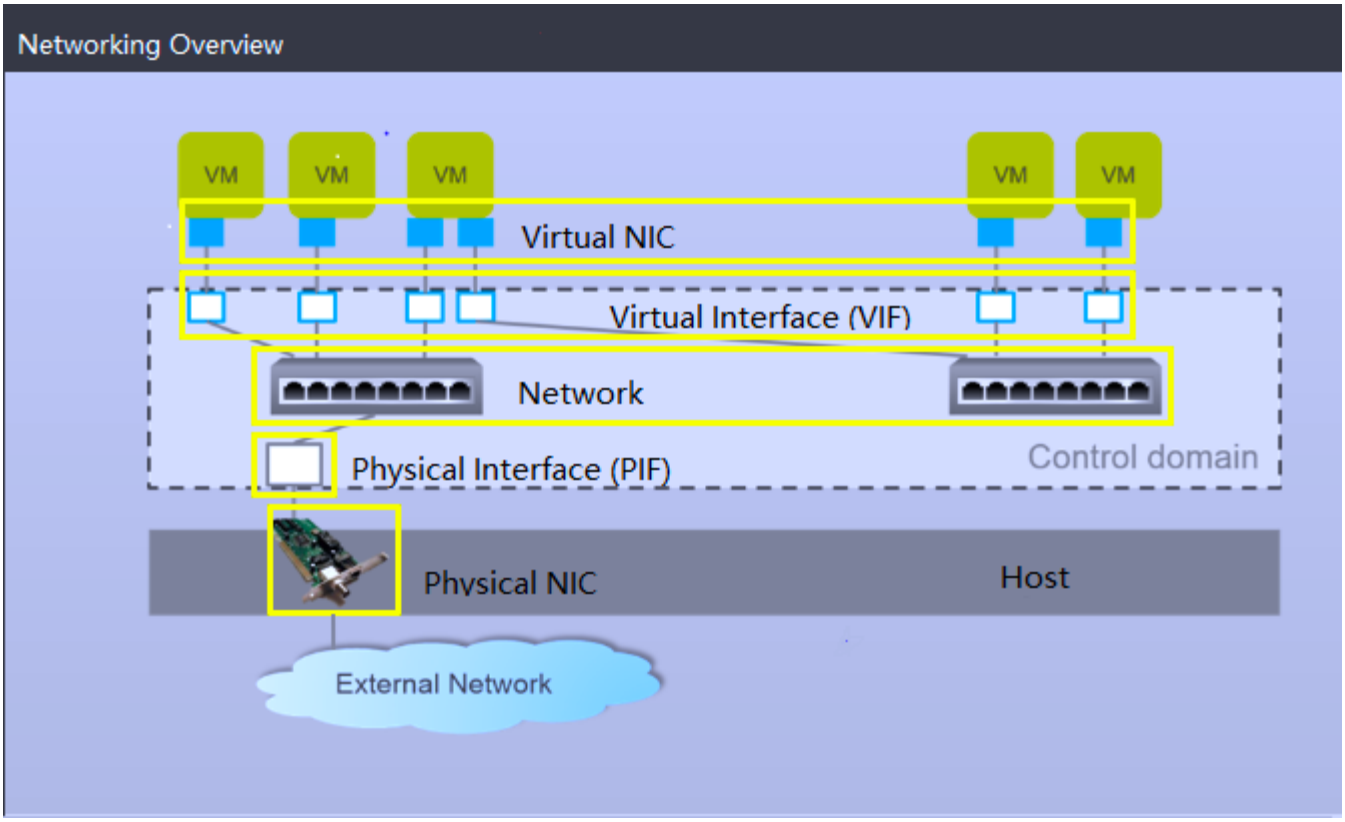
Feature	LVM over iSCSI	GFS2
Known constraints	• Read caching not supported	• VM migration with storage live migration is not supported for VMs whose VDIs are on a GFS2 SR. You also cannot migrate VDIs from another type of SR to a GFS2 SR. • Trim/unmap is not supported on GFS2 SRs. • Performance metrics are not available for GFS2 SRs and disks on these SRs. • Changed block tracking is not supported for VDIs stored on GFS2 SRs. • You cannot export VDIs that are greater than 2TiB as VHD or OVA/OVF. However, you can export VMs with VDIs larger than 2TiB in XVA format. • Clustered pools only support up to 16 hosts per pool.

With the current features available in NetApp HCI, the Intellicache feature of Citrix Hypervisor is not of value to NetApp HCI customers. Intellicache improves performance for file-based storage systems by caching data in a local storage repository.

Read caching allows you to improve performance for certain storage repositories by caching data in server memory. GFS2 is the first iSCSI volume to support read caching.

Network

Citrix Hypervisor networking is based on Open vSwitch with support for OpenFlow. It supports fine grain security policies to control the traffic sent and receive from a VM. It also provides detailed visibility about the behavior and performance of all traffic sent in the virtual network environment. The following figure presents an overview of Citrix Hypervisor networking.



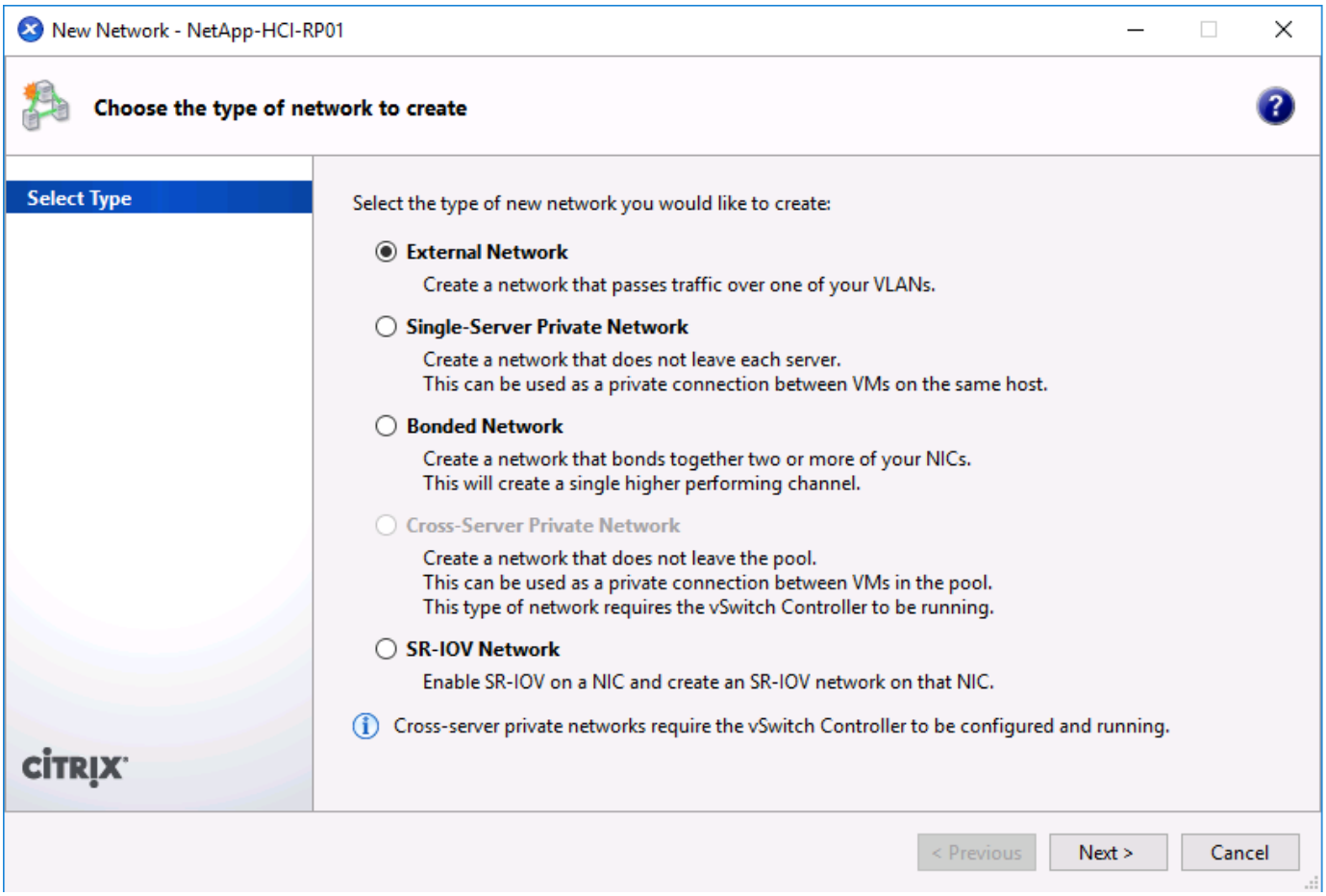
The physical interface (PIF) is associated with a NIC on the server. With Network HCI, up to six NICs are available for use. With the model, which only has two NICs, SR-IOV can be used to add more PIFs. The PIF acts as an uplink port to the virtual switch network. The virtual interface (VIF) connects to a NIC on virtual machines.

Various network options are available:

- An external network with VLANs
- A single server private network with no external connectivity
- Bonded network (active/active – aggregate throughput)
- Bonded network (active/passive – fault tolerant)
- Bonded network (LACP – load balancing based on source and destination IP and port)
- Bonded network (LACP – load balancing based on source and destination mac address)
- Cross-server private network in which the network does not leave the resource pool
- SR-IOV

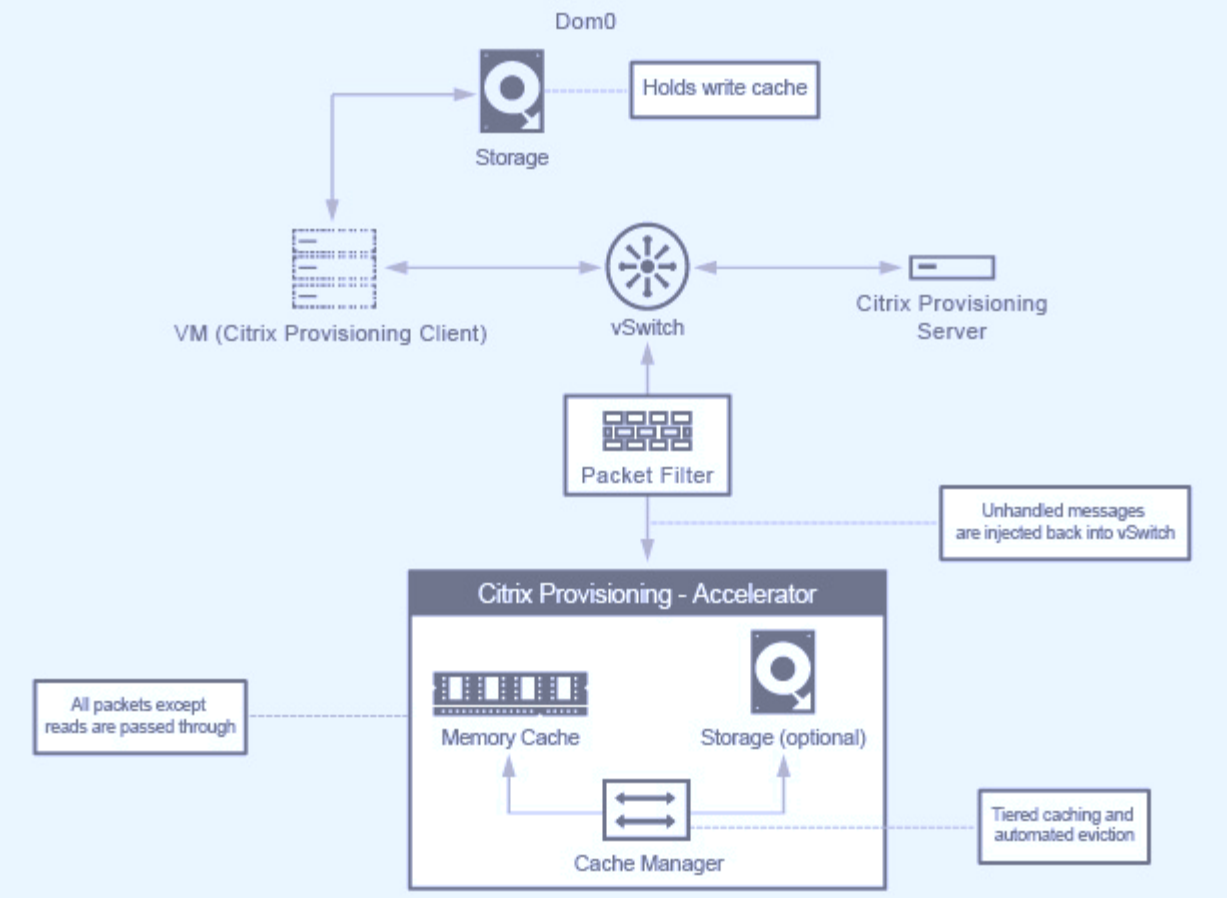
The network configuration created on the master server is replicated to other member servers. Therefore, when a new server is added to the resource pool, its network configuration is replicated from the master.

You can only assign one IP address per VLAN per NIC. For iSCSI multipath, you must have multiple PIFs to assign an IP on the same subnet. For H615C, you can consider SR-IOV for iSCSI.



Because the network on Citrix Hypervisor is based on Open vSwitch, you can manage it with ovs-vsctl and ovs-appctl commands. It also supports NVGRE/VXLAN as an overlay solution for large scale- out environments.

When used with Citrix Provisioning (PVS), PVS Accelerator improves performance by caching Domain 0 memory or by combining memory and a local storage repository.



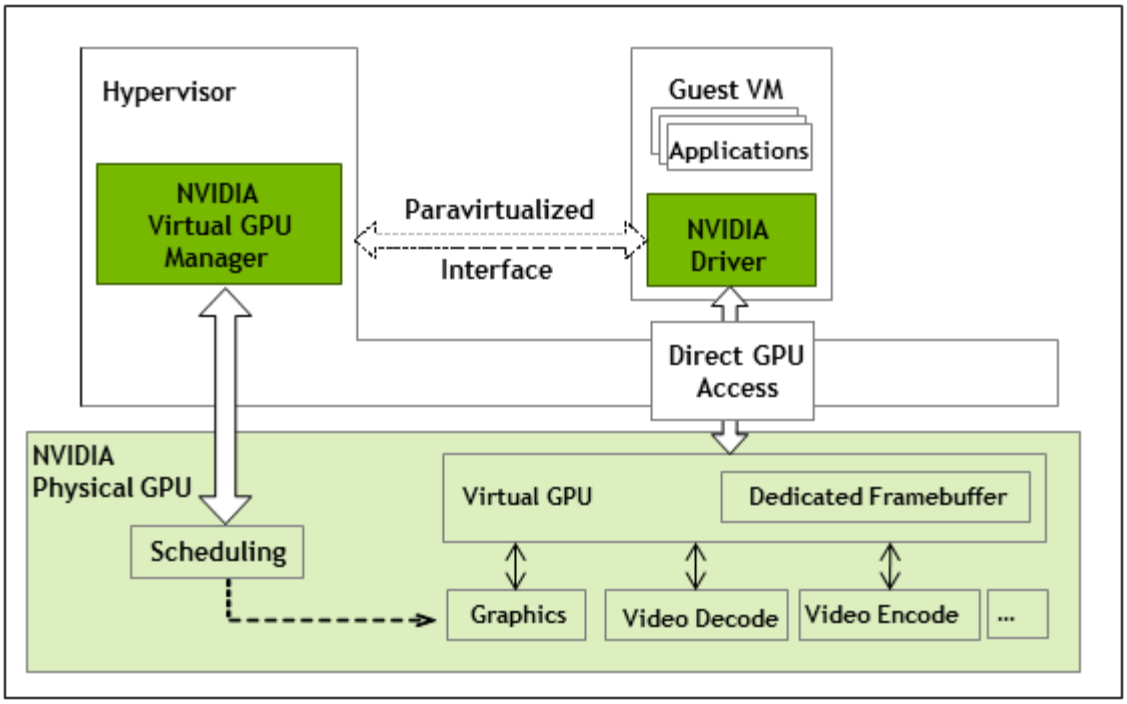
GPU

Citrix Hypervisor was the first to deploy NVIDIA vGPUs, a virtualization platform for GPUs, enabling the sharing of GPU across multiple virtual machines. NetApp HCI H610C (with NVIDIA Tesla M10 cards) and H615C (with NVIDIA Tesla T4 cards) can provide GPU resources to virtual desktops, providing hardware acceleration to enhance the user experience.

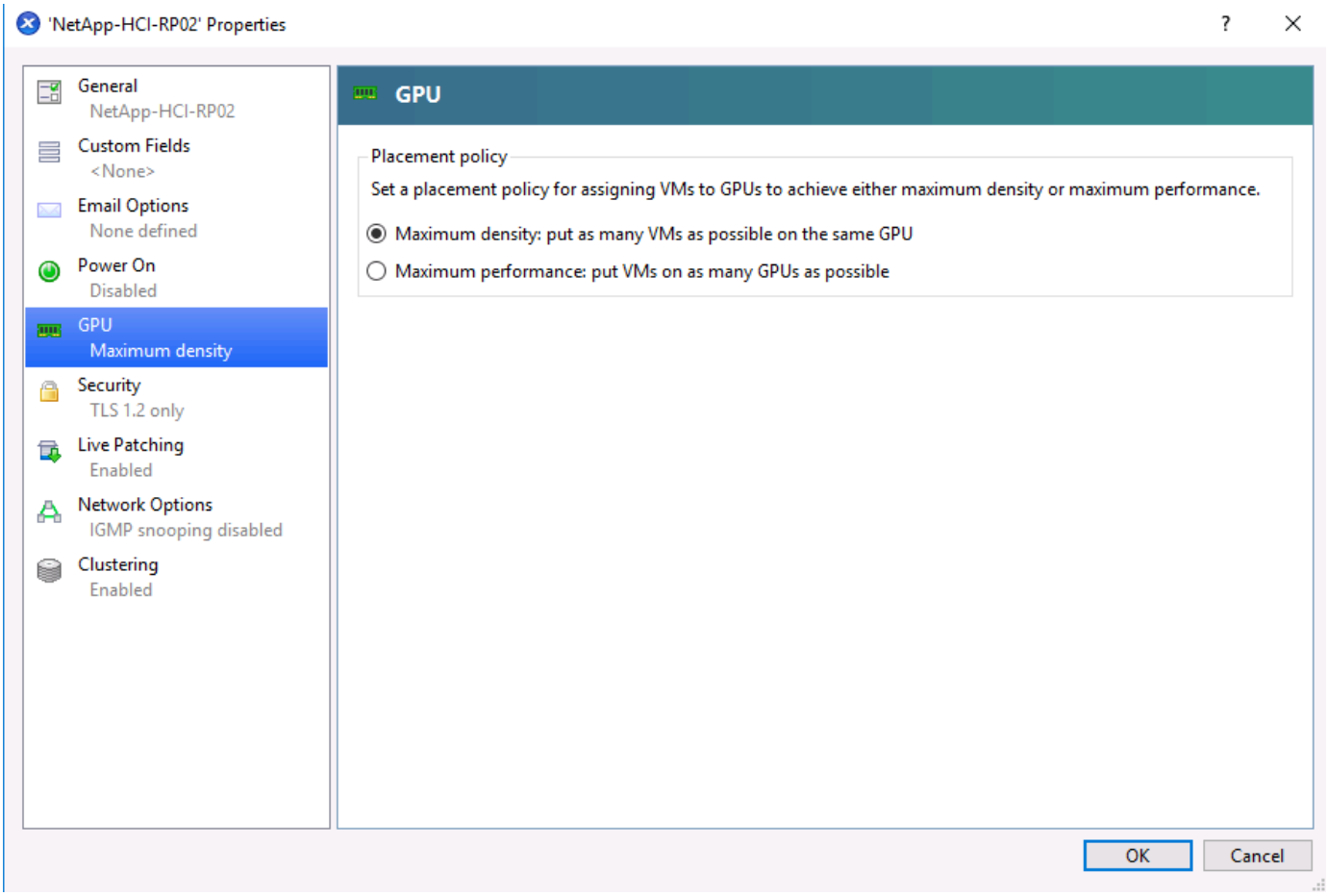
A NetApp HCI GPU can be consumed in a Citrix Hypervisor environment by using pass- through mode, where the whole GPU is presented to a single virtual machine, or it can be consumed using NVIDIA vGPU. Live migration of a VM with GPU pass through is not supported, and therefore NVIDIA vGPU is the preferred choice.

NVIDIA Virtual GPU Manager for Citrix Hypervisor can be deployed along with other management packs by using XenCenter or it can be installed using an SSH session with the server. The virtual GPU gets its own dedicated frame buffers, while sharing the streaming processors, encoder, decoder and so on. It can also be controlled using a scheduler.

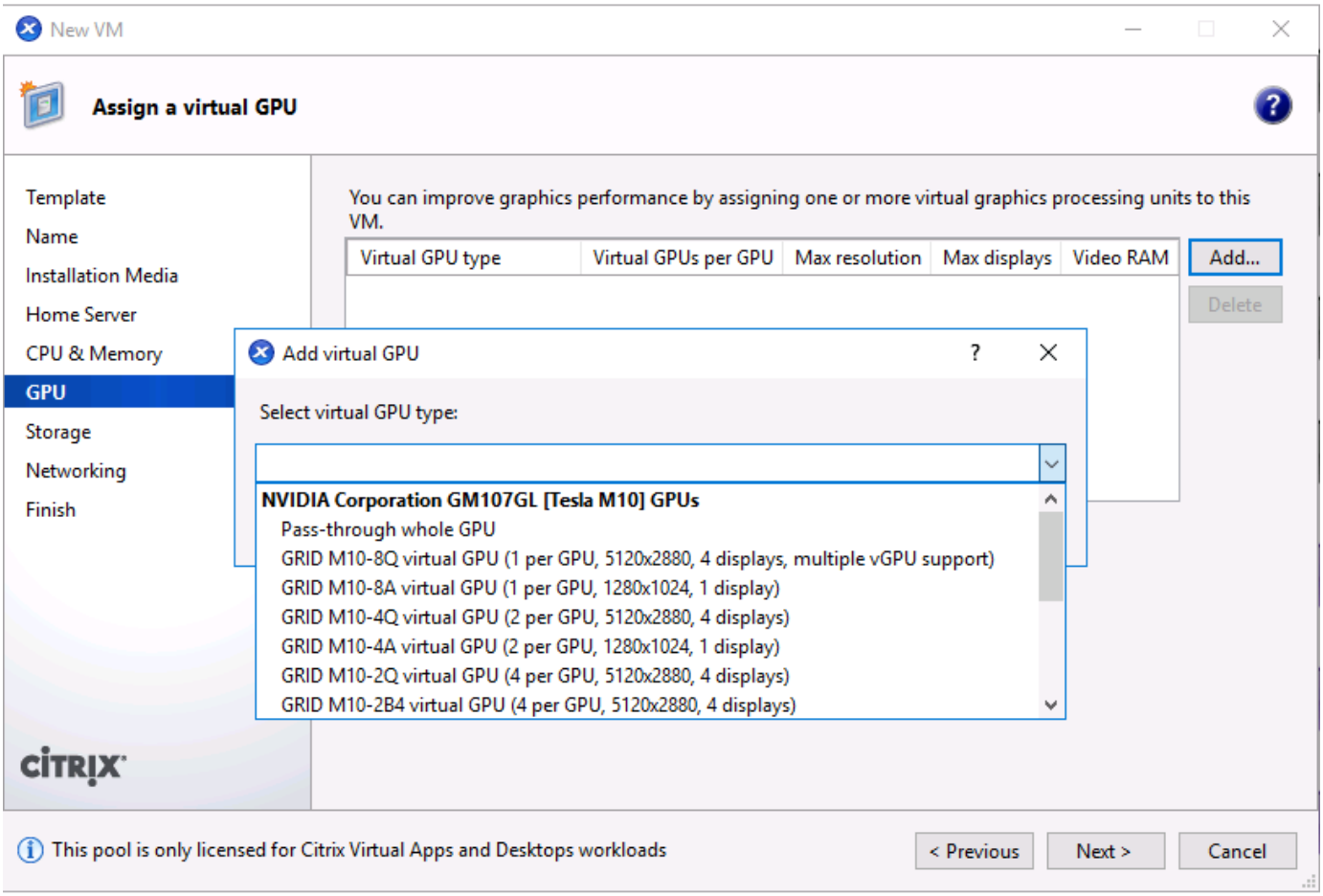
The H610C has two Tesla M10 graphic cards, each with 4 GPUs per card. Each GPU has 8GB of frame buffer memory with a total of 8 GPUs and 64GB of memory per server. H615C has three Tesla T4 cards, each with its own GPU and 16GB frame buffer memory with a total of 3 GPUs and 48GB of graphic memory per server. The following figure presents an overview of the NVIDIA vGPU architecture.



NVIDIA vGPU supports homogenous profiles for each GPU. The placement of virtual machines on a GPU is controlled by a policy that sets either maximum density or maximum performance in response to demand.



When creating a VM, you can set a virtual GPU profile. The vGPU profile you chose is based on the frame buffer memory level needed, the number of displays, and the resolution requirement. You can also set the purpose of a virtual machine, whether it be virtual apps (A), virtual desktops (B), a professional Quadro virtual workstation (Q), or compute workloads (C) for AI inferencing applications.



Independently from XenCenter, the CLI utility on the Citrix Hypervisor nvidia-smi can be used to troubleshoot and for monitoring the performance.

The NVIDIA driver on a virtual machine is required to access the virtual GPU. Typically, the hypervisor driver version and the VM guest driver should have the same vGPU release version. But, starting with vGPU release 10, the hypervisor can have the latest version while the VM driver can be the n-1 version.

Security

Citrix Hypervisor supports authentication, authorization, and audit controls. Authentication is controlled by local accounts as well as by Active Directory. Users and groups can be assigned to roles that control permission to resources. Events and logging can be stored remotely in addition to on the local server.

Citrix Hypervisor supports Transport Layer Security (TLS) 1.2 to encrypt the traffic using SSL certificates.

Because most configuration is stored locally in an XML database, some of the contents, like SMB

passwords, are in clear text, so you must protect access to the hypervisor.

Data Protection

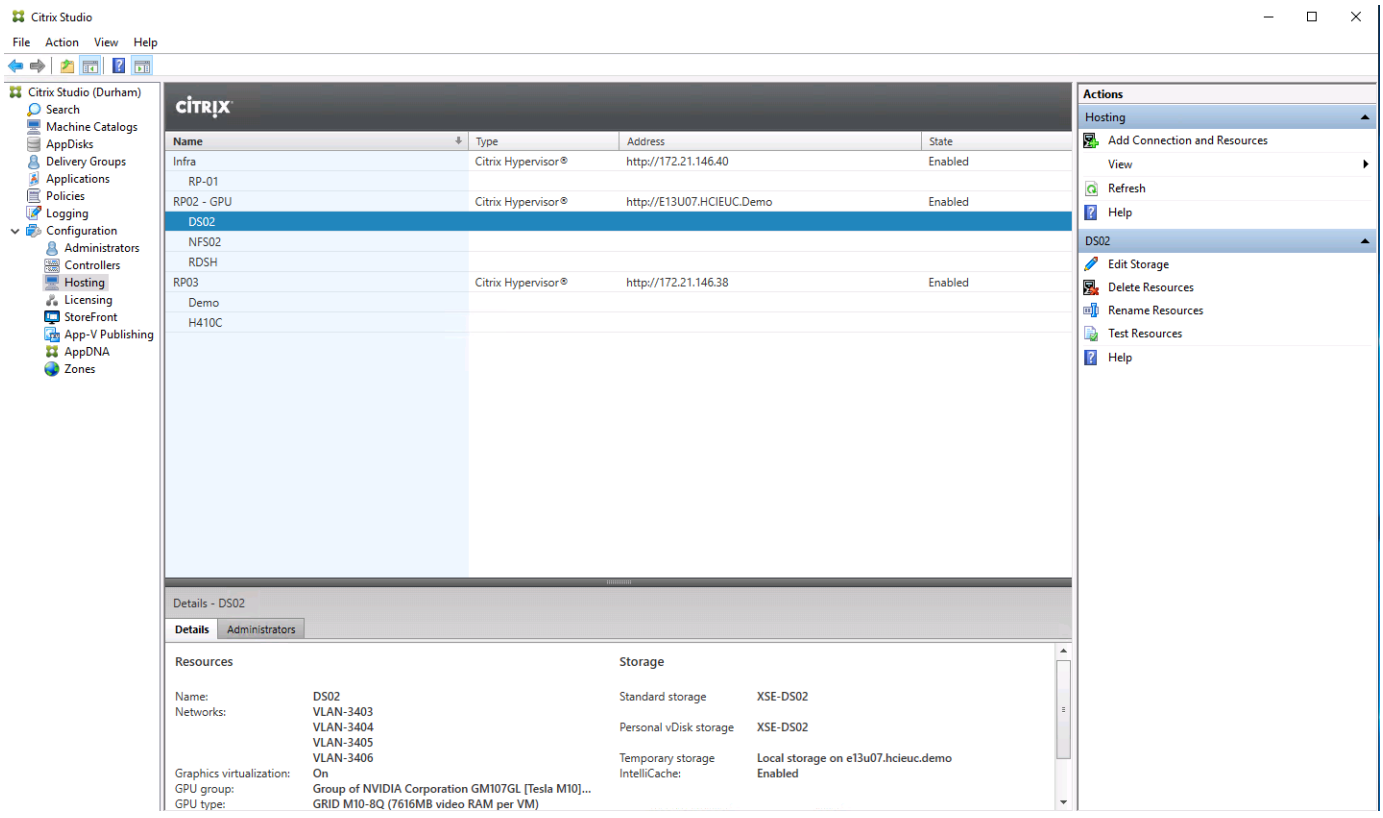
Virtual machines can be exported as OVA files, which can be used to import them to other hypervisors. Virtual machines can also be exported in the native XVA format and imported to any other Citrix Hypervisor. For disaster recovery, this second option is also available along with storage- based replication handled by SnapMirror or native Element OS synchronous or asynchronous replication. With NetApp, HCI storage can also be paired with ONTAP storage for replication.

Storage-based snapshot and cloning features are available to provide crash-consistent image backups. Hypervisor-based snapshots can be used to provide point-in-time snapshots and can also be used as templates to provision new virtual machines.

Resource Layer

Compute

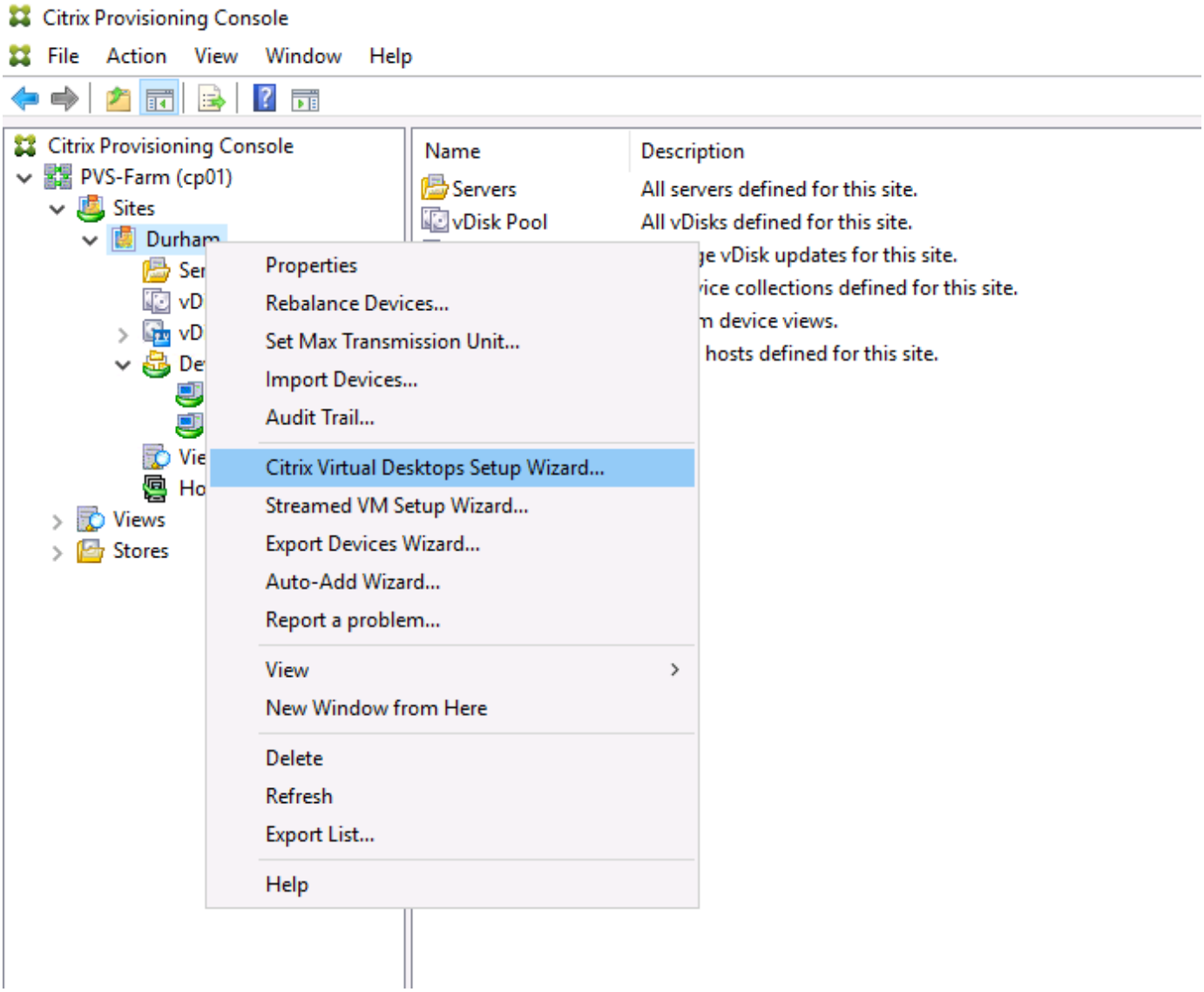
To host virtual apps and desktop resources, a connection to a hypervisor and resource details should be configured in Citrix Studio or with PowerShell. In the case of Citrix Hypervisor, a resource pool master node DNS or IP address is required. For a secure connection, use HTTPS with SSL certificates installed on the server. Resources are defined with selection the of storage resources and networks.



When additional compute capacity is required, a hypervisor server can be added to existing resource pool. Whenever you add a new resource pool and you need to make it available for hosting virtual apps and desktops, you must define a new connection.

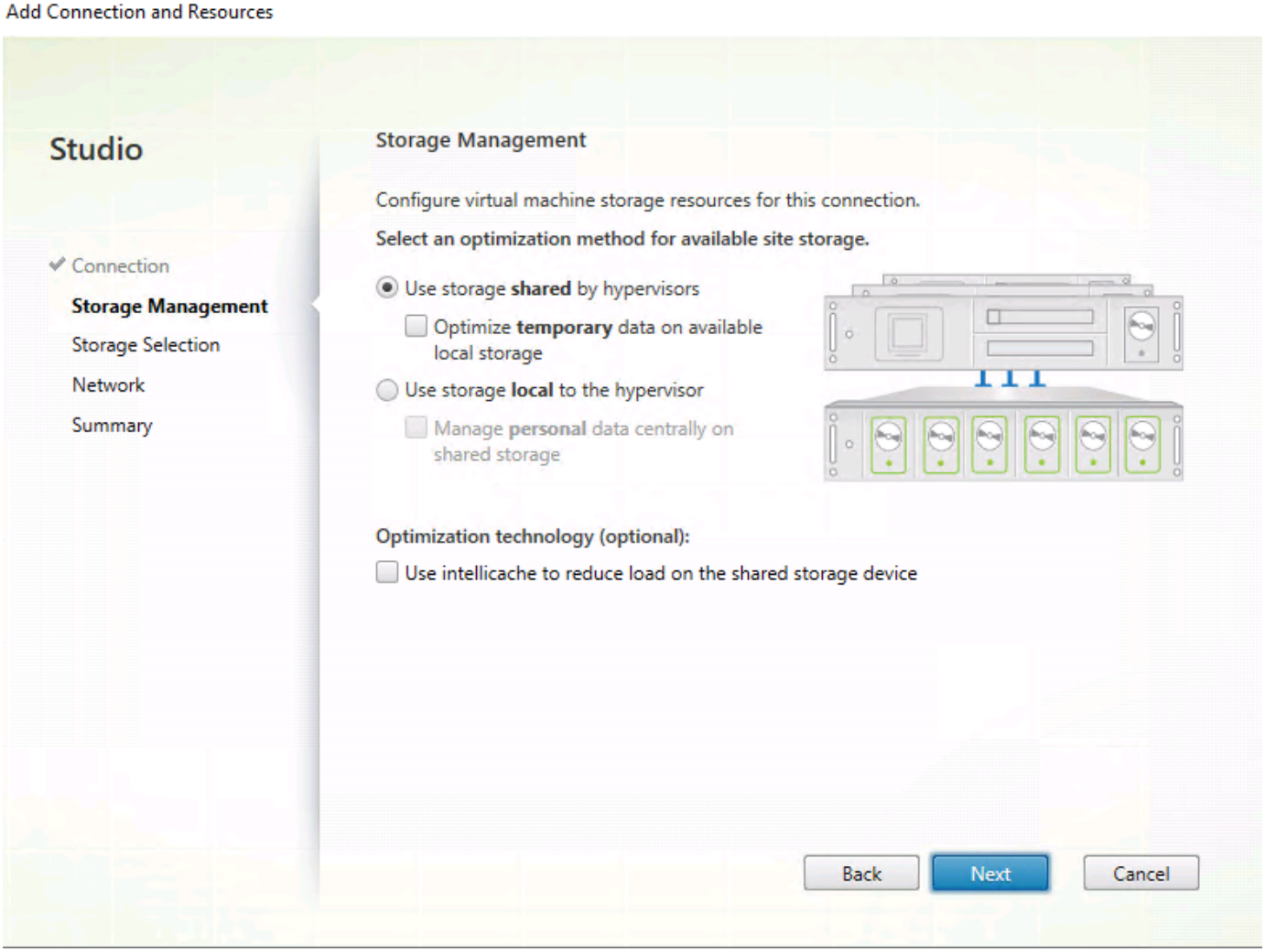
A site is where the SQL database resides and is known as the primary zone. Additional zones are added to address users in different geographic locations to provide better response time by hosting on local resources. A satellite zone is a remote zone that only has hypervisor components to host virtual apps or desktops with optional delivery controllers.

Citrix Provisioning also uses the connection and resources information when using the Citrix Virtual Desktops Setup Wizard.

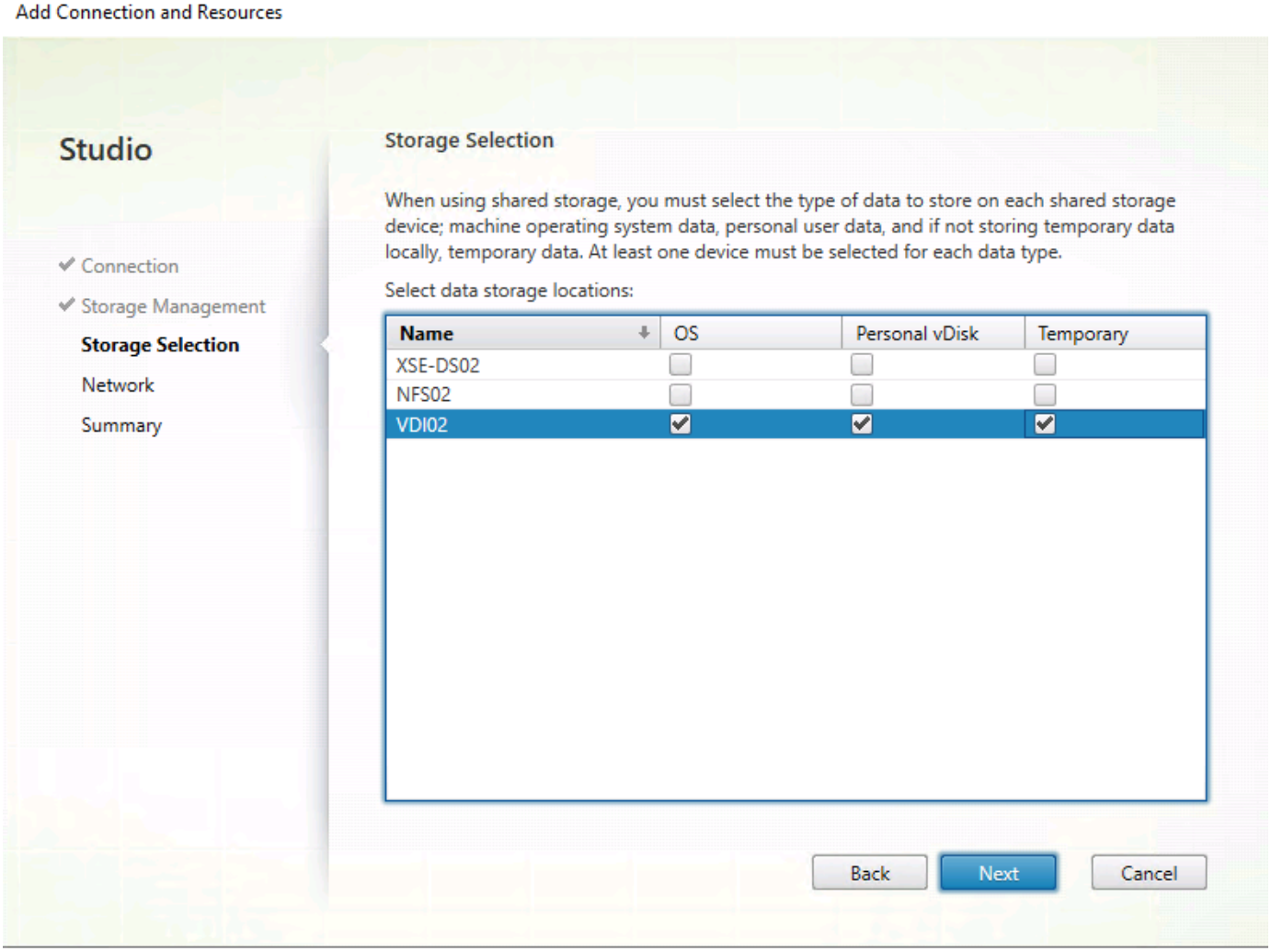


Storage

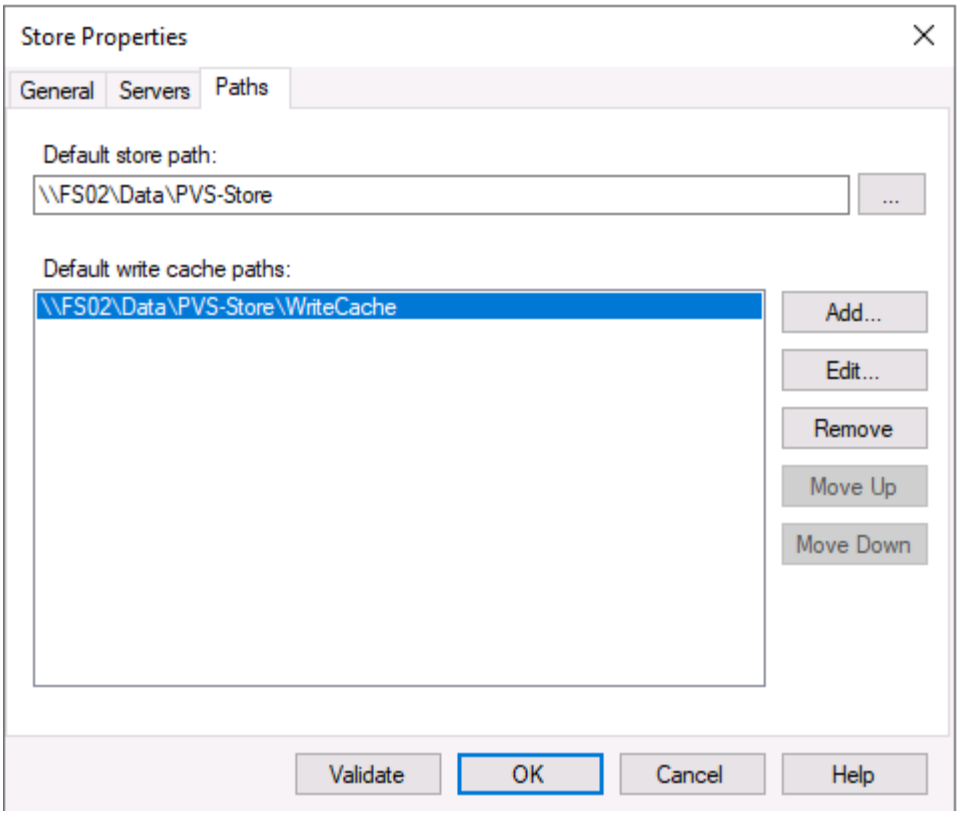
The storage repository for Virtual Apps and Desktops is controlled using the connection and resources covered in the section Compute. When you define the resource, you have the option to pick the shared storage and enable Intellicache with Citrix Hypervisor.



There is also an option to pick resources for the OS, the personal vDisk, and temporary data. When multiple resources are selected, CItrix Virtual Apps and Desktops automatically spreads the load. In a multitenant environment, a dedicated resource selection can be made for each tenant resource.

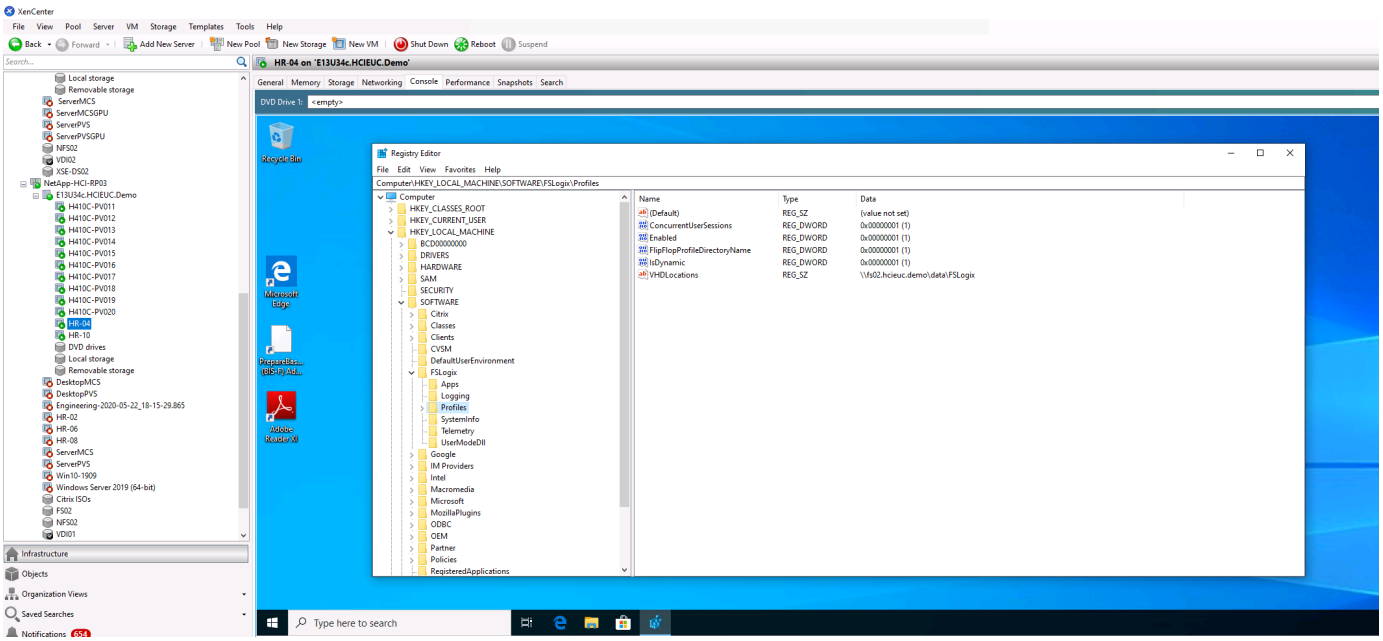


Citrix Provisioning requires an SMB file share to host the vDisks for the devices. We recommend hosting this SMB share on a FlexGroup volume to improve availability, performance, and capacity scaling.



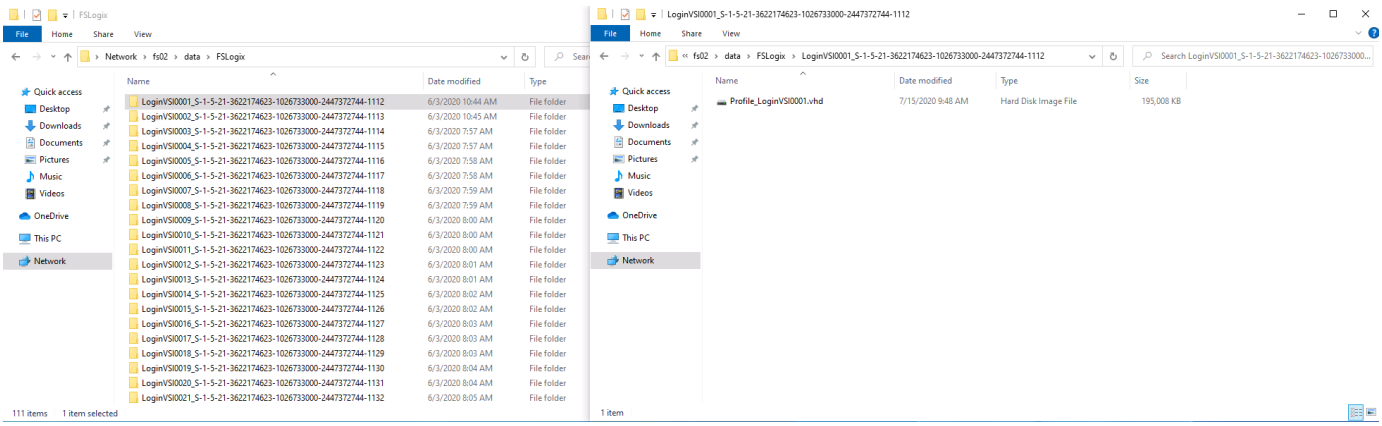
FSLogix

FSLogix allows users to have a persistent experience even in non-persistent environments like pooled desktop deployment scenarios. It optimizes file I/O between the virtual desktops and the SMB file store and reduces login time. A native (local) profile experience minimizes the tasks required on the master image to set up user profiles.



FSLogix keeps user settings and personal data in its own container (VHD file). The SMB file share to store the FSLogix user profile container is configured on a registry that is controlled by group policy

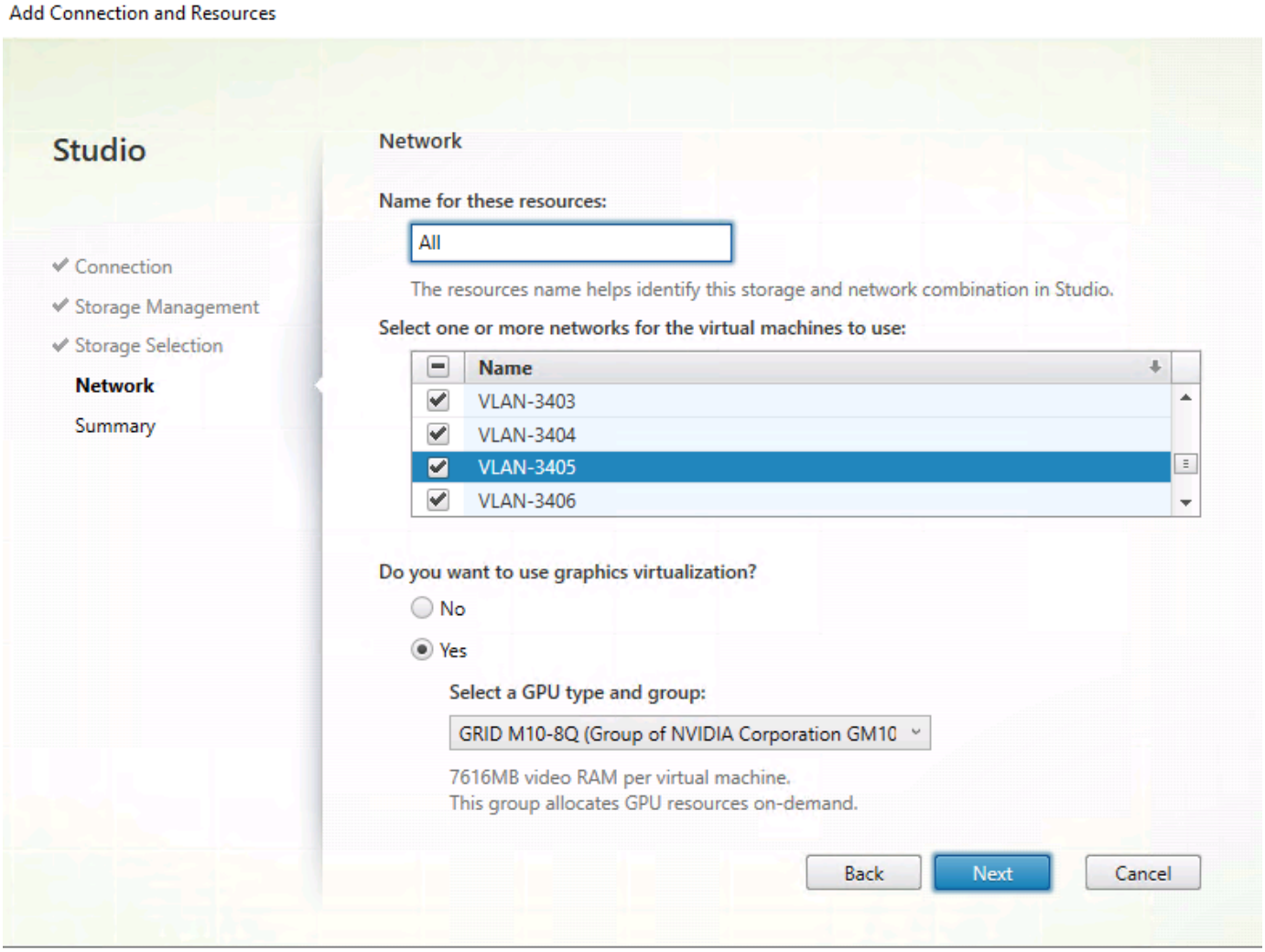
object. Citrix User Profile Management can be used along with FSLogix to support concurrent sessions with virtual desktops at the same time on virtual apps.



This figure shows the content of the FSLogix SMB location. Note that we switched the directory name to show the username before the security identifier (sid).

Network

Virtual Apps and Desktops require a connection and resources to host, as covered in the section Compute. When defining the resource, pick the VLANs that must be associated with the resource. During machine catalog deployment, you are prompted to associate the VM NIC to the corresponding network.



GPU

As indicated in the previous section, when you determine whether the hypervisor server has a GPU resource, you are prompted to enable graphics virtualization and pick the vGPU profile.

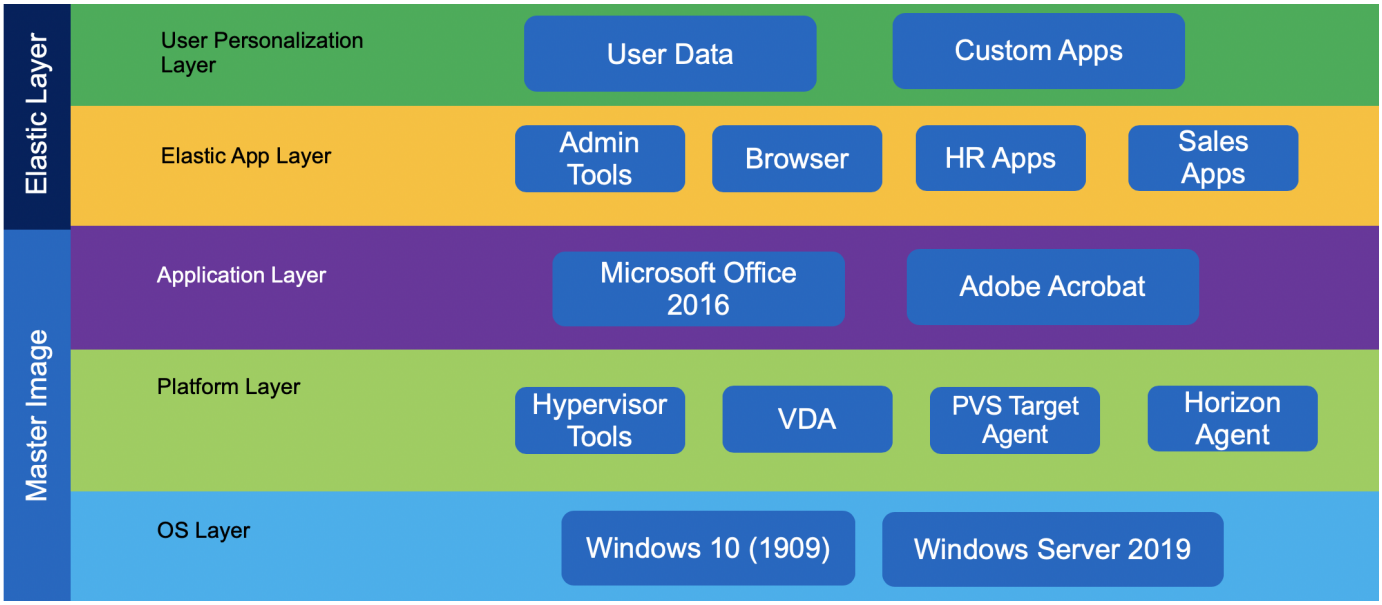
Control Layer

App Layering

Layering is a technology to separate the OS, applications, and user settings and data, each hosted on its own virtual disks or group of virtual disks. These components are then merged with the OS as if they were all on same machine image. Users can continue with their work without any additional training. Layers make it easy to assign, patch, and update. A layer is simply a container for file system and registry entries unique to that layer.

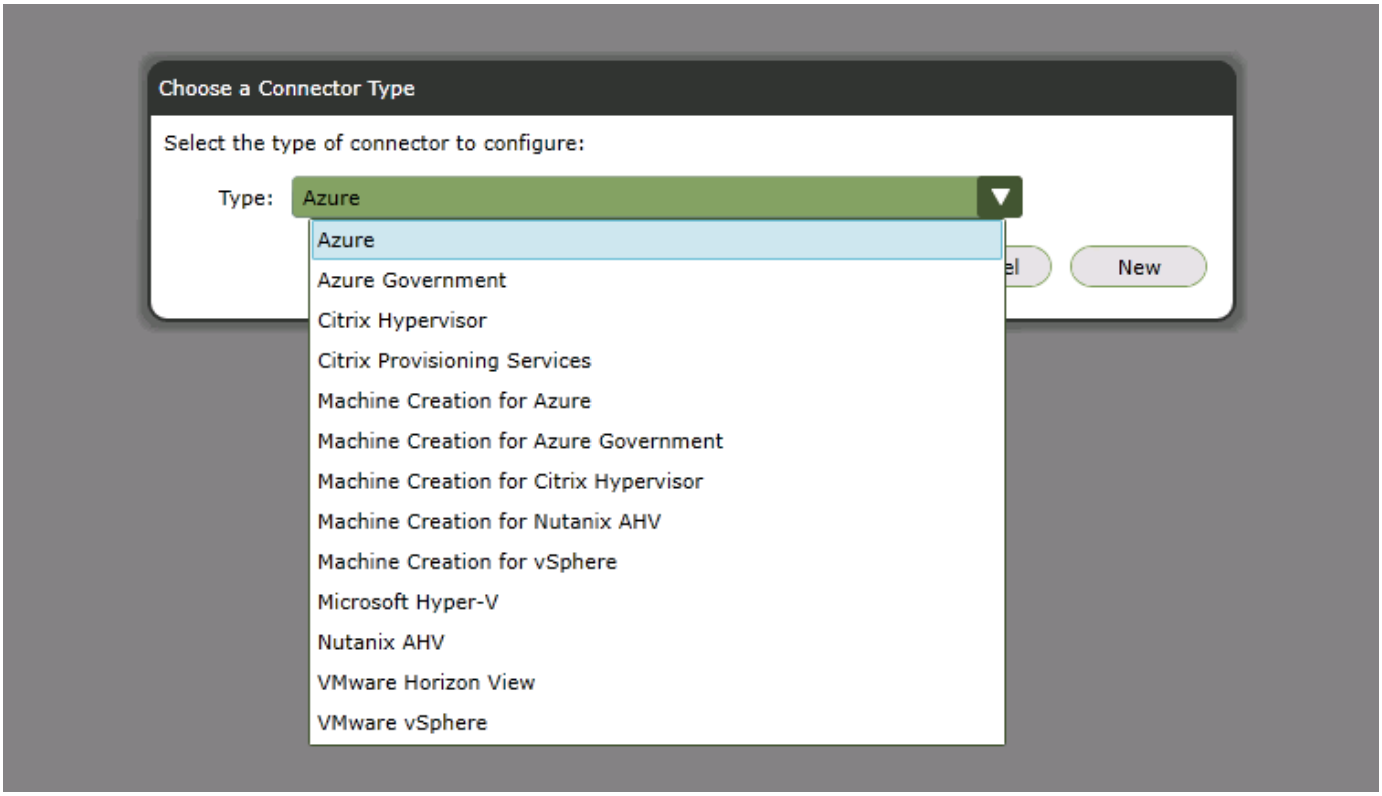
Citrix App Layering allows you to manage master images for Citrix Virtual Apps and Desktops as well as for the VMware Horizon environment. App layering also allows you to provision applications to users on demand; these apps are attached while logging in. The user personalization layer allows users to install custom apps and store the data on their dedicated layer. Therefore, you can have a personal desktop experience even when you are using a shared desktop model.

Citrix App Layering creates merged layers to create the master image and does not have any additional performance penalty. With Elastic Layers, the user login time increases.

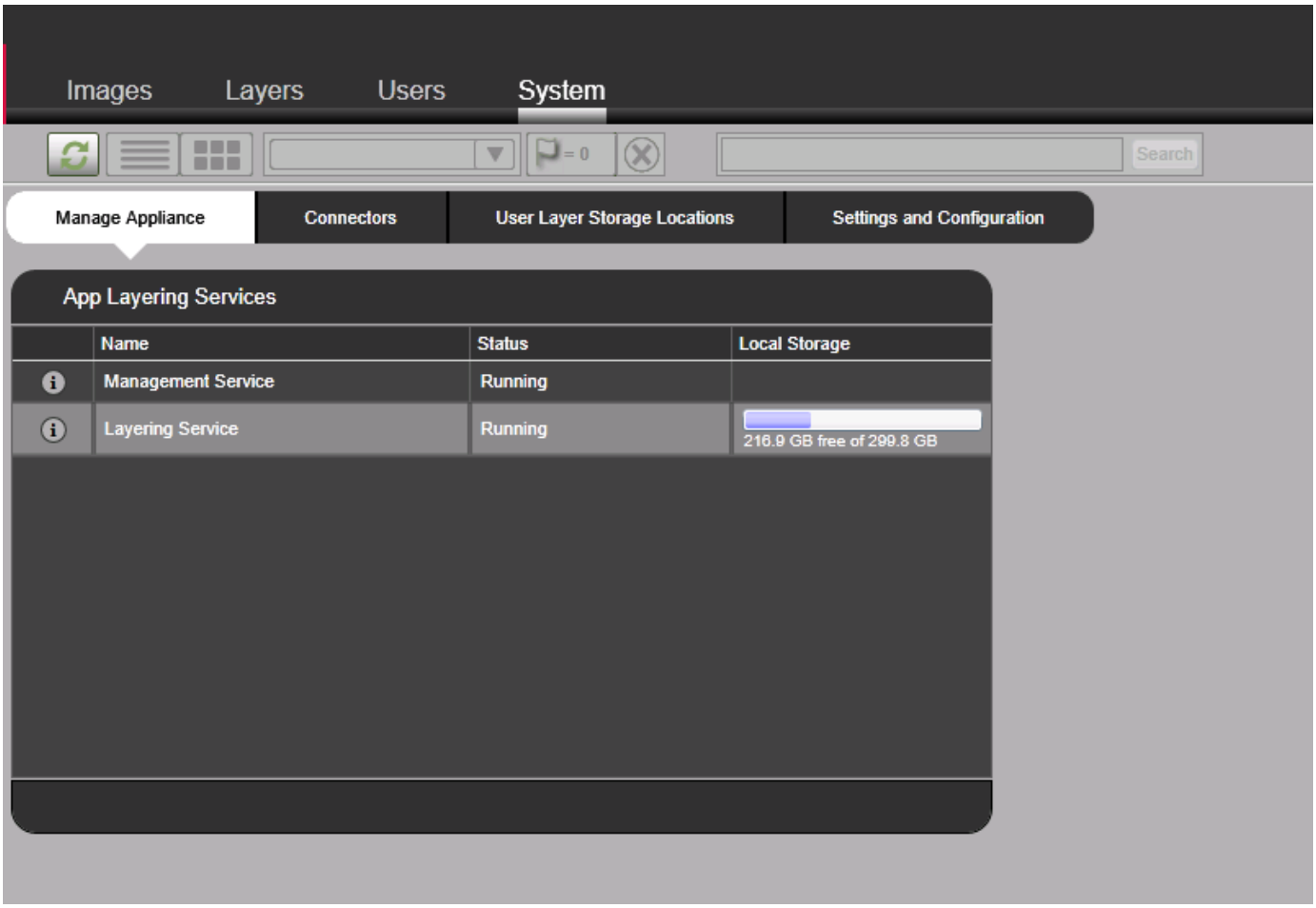


Citrix App Layering uses a single virtual appliance to manage the layers and hands off using the image and application delivery to another platform. The Citrix Enterprise Layer Manager (ELM) portal must be accessed from web browsers that supports Microsoft Silverlight 4.0. A cloud- based management portal is also available if local management interface requirements cannot be met.

Initial configuration includes the creation of platform connectors of two types; the first is a platform connector for layer creation, and the other is a platform connector for image publishing.



A layer repository is an SMB file share configured with ELM where Elastic Layers are stored. A layer work disk is where all the layers created by ELM are stored. The disk is attached to the appliance and is consumed as a block device on which a local Linux file system is used. The layer work disk is used as scratch area where the layer images are put together. After the master image is created, it is pushed to the provisioning platform.



When there are common or shared files on multiple layers, by default the high priority layer ID wins. Layer ID is incremented whenever a new layer is created. If would like to control layer priority, use the support utility on the Citrix LayerPriority Utility page.

ELM also supports authentication and role- based access control with integration with Active Directory and LDAP.

Delivery Controller

The delivery controller is responsible for user access, brokering, and optimizing connections. It also provides Machine Creation Services (MCS) for provisioning virtual machines in an effective manner. At least one delivery controller is required per site, and typically additional controllers are added for redundancy and scalability.

Virtual desktop agents (VDA) must register with the delivery controller to make it available to users. During VDA deployment, the initial registration options can be provided manually through GPO based

on the Active Directory OU. This process can also be handled with MCS.

Delivery controllers keep a local host cache in case a controller loses its connectivity to database server.

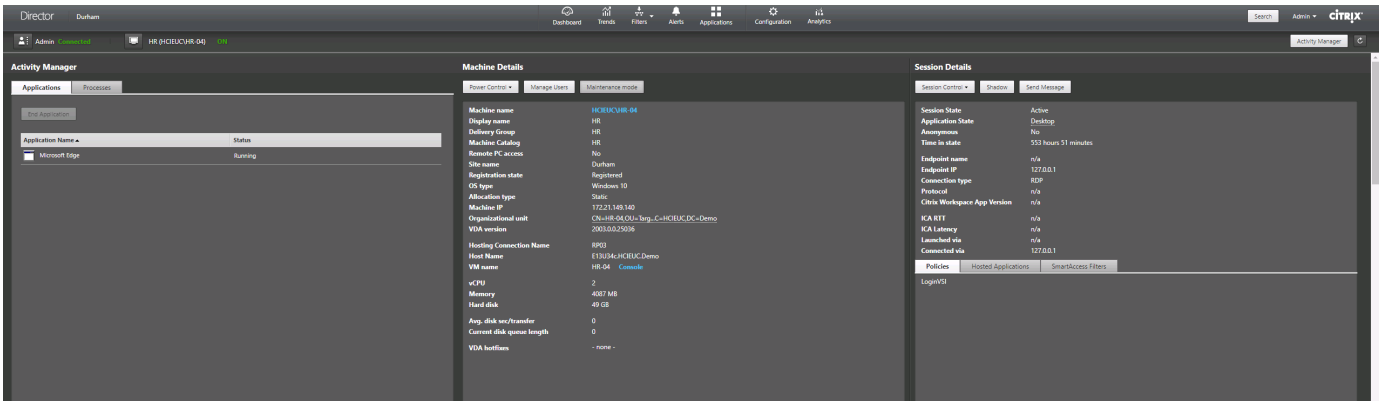
Database

A SQL Server database is used for site configuration data, logging, and monitoring. There should be at least one database per site. To provide high availability, use Microsoft SQL Server features like AlwaysOn availability groups, database mirroring, or SQL clustering. At a minimum, consider using the hypervisor high- availability feature for a SQL VM.

Even though the controller has a local host cache, it doesn't affect any existing connections. However, for new connections, NetApp recommends database connectivity.

Director

Citrix Director provides a monitoring solution for Citrix Virtual Apps and Desktops. Help Desk users can search for a specific user session and get a complete picture for troubleshooting. When Citrix Virtual Apps and Desktop Resources are hosted on Citrix Hypervisor, Help Desk users have the option to launch a console session from the Director portal.

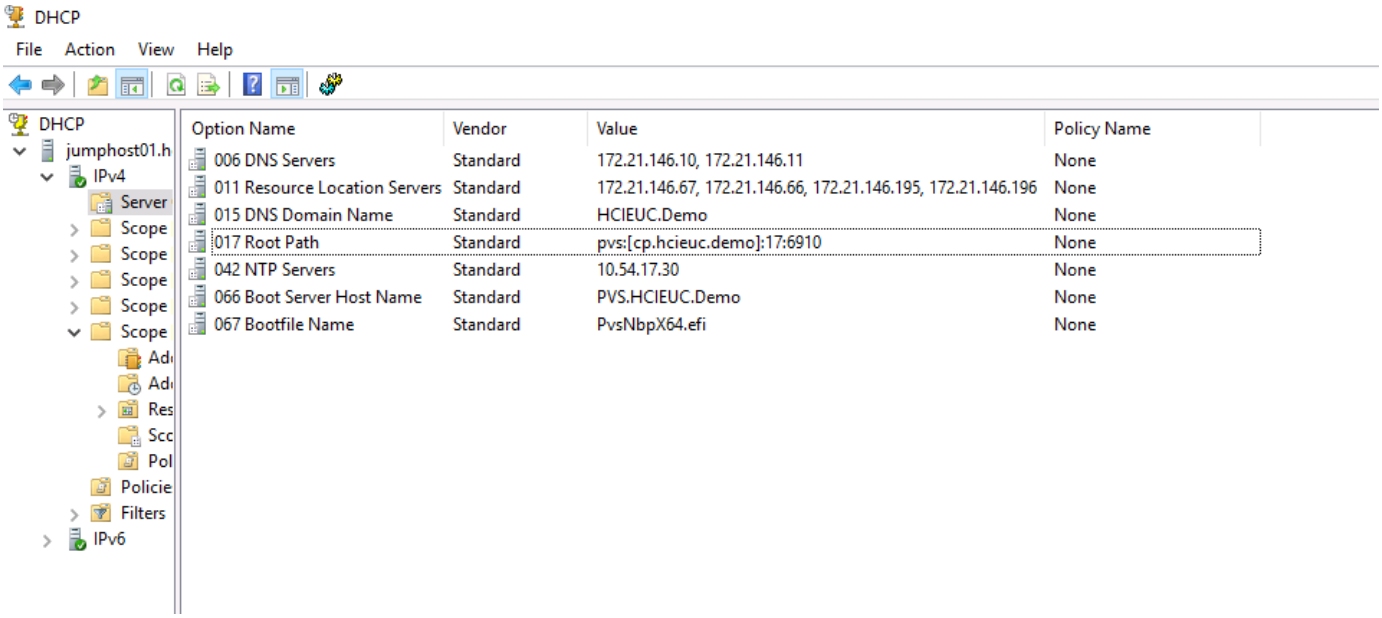


License

The Citrix license server manages the repository of all Citrix licenses so that licenses can be easily consumed by applications. The license server provides a management portal for advanced troubleshooting. For regular operations, Citrix Studio can also be used.

Provisioning Services

Provisioning services enable the provisioning of desktop images even to bare metal workstations by using PXE boot. An ISO or CDROM-based boot option is also available to support environments in which network changes aren't allowed for PXE boot. The DHCP server options that we used in our lab is provided in the following figure. CP.HCIEUC.Demo and PVS.HCIEUC.Demo are the load balancer virtual IPs that point to two provisioning servers. When option 011 and 017 are available, options 066 and 067 are ignored.



The high-level operation to create a machine catalog based on Citrix provisioning is as follows:

1. On the template VM, install the target agent before installing VDA.
2. Assign an additional disk for caching and format it with MBR. This step is optional. At least verify that the PVS store has a write cache path.
3. Start the Target Image Wizard and respond to its questions. Remember to provide a single Citrix Provisioning server when prompted.
4. The device boots with PXE or with ISO. The Imaging wizard continues to capture the image.
5. Select the vDisk that is created and right click to select Load Balancing and enable it.
6. For vDisk Properties, change the access mode to Standard and the Cache Type to Cache in Device RAM with Overflow on Hard Disk.
7. Right click on the site to pick the Create Virtual Desktops Setup Wizard and respond to the questions.

Studio

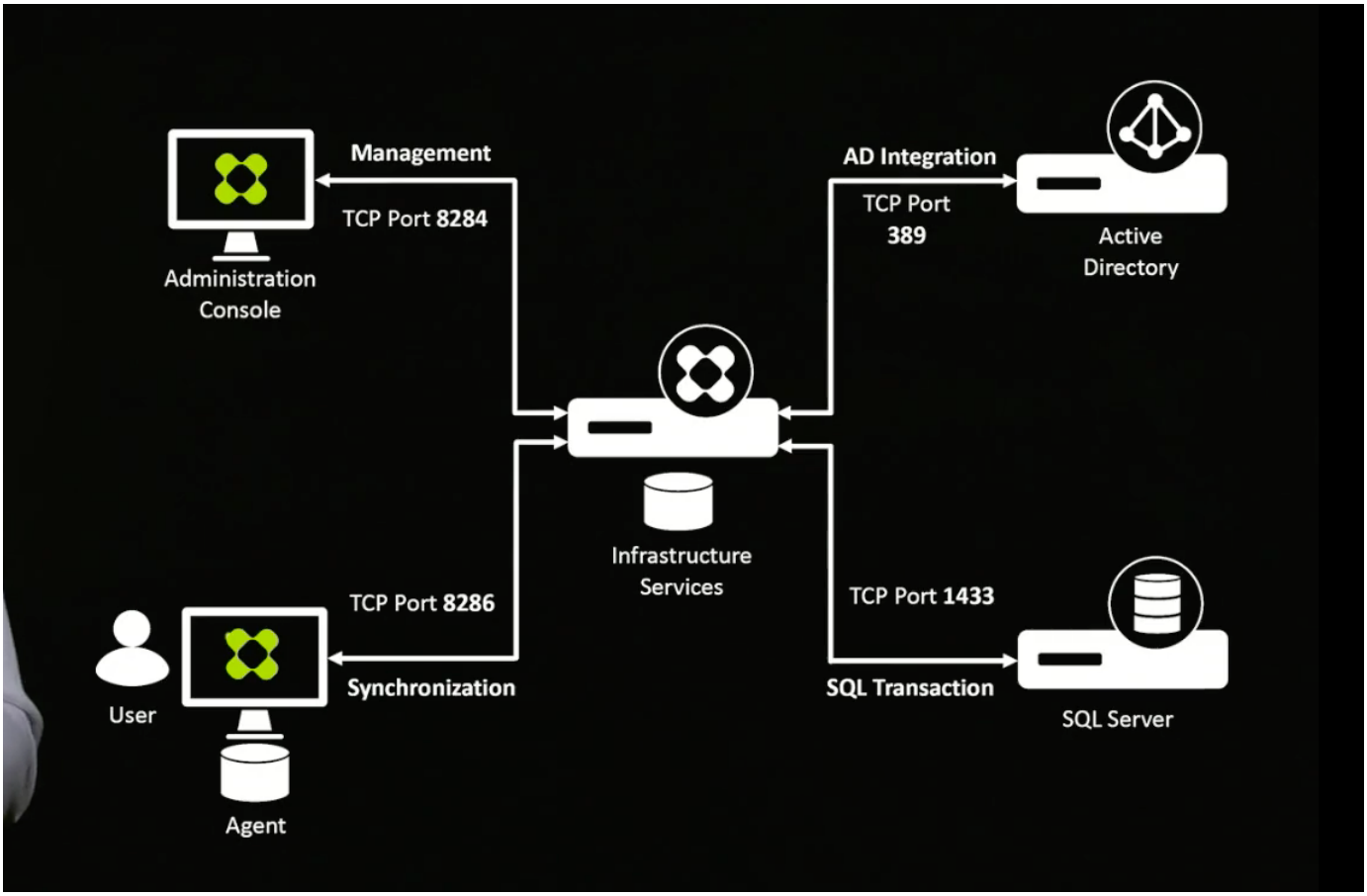
Citrix Studio is the central management console used by the Citrix Virtual Apps and Desktops. The management of machine catalogs, delivery groups, applications, policies, and the configuration of resource hosting, licenses, zones, roles, and scopes are handled by the Citrix Studio. Citrix Studio also provides PowerShell snap-ins to manage Citrix Virtual Apps and Desktops.

Workspace Environment Management

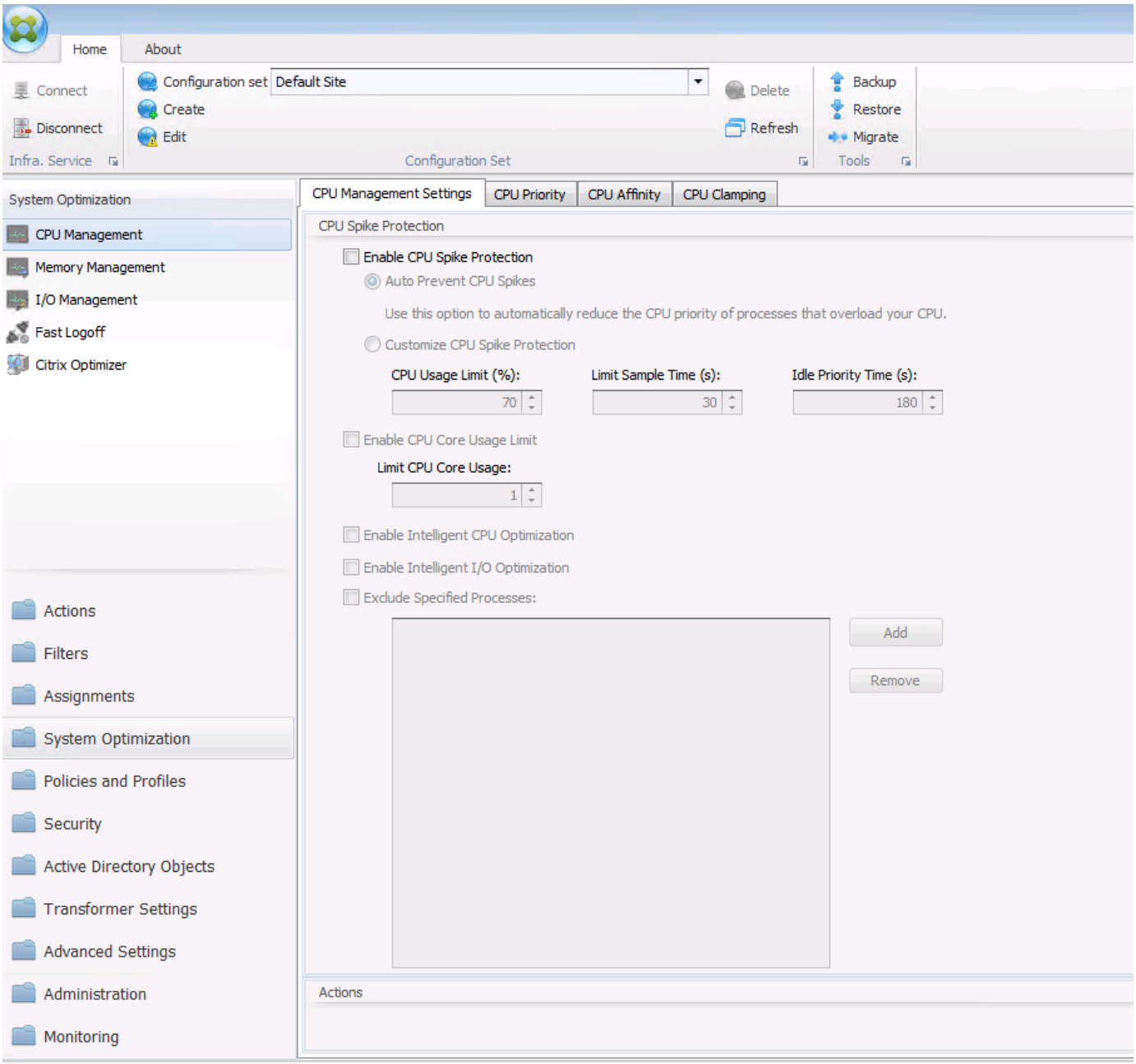
Workspace Environment Management (WEM) provides intelligent resource management and profile management technologies to deliver the best possible performance, desktop login, and application response times for Citrix Virtual Apps and Desktops in a software-only, driver-free solution.

WEM requires a SQL database to store configuration information. To provide high availability to

infrastructure services, multiple instances are used with a load balancer virtual server connection. The following figure depicts the WEM architecture.



The following figure depicts the WEM console.



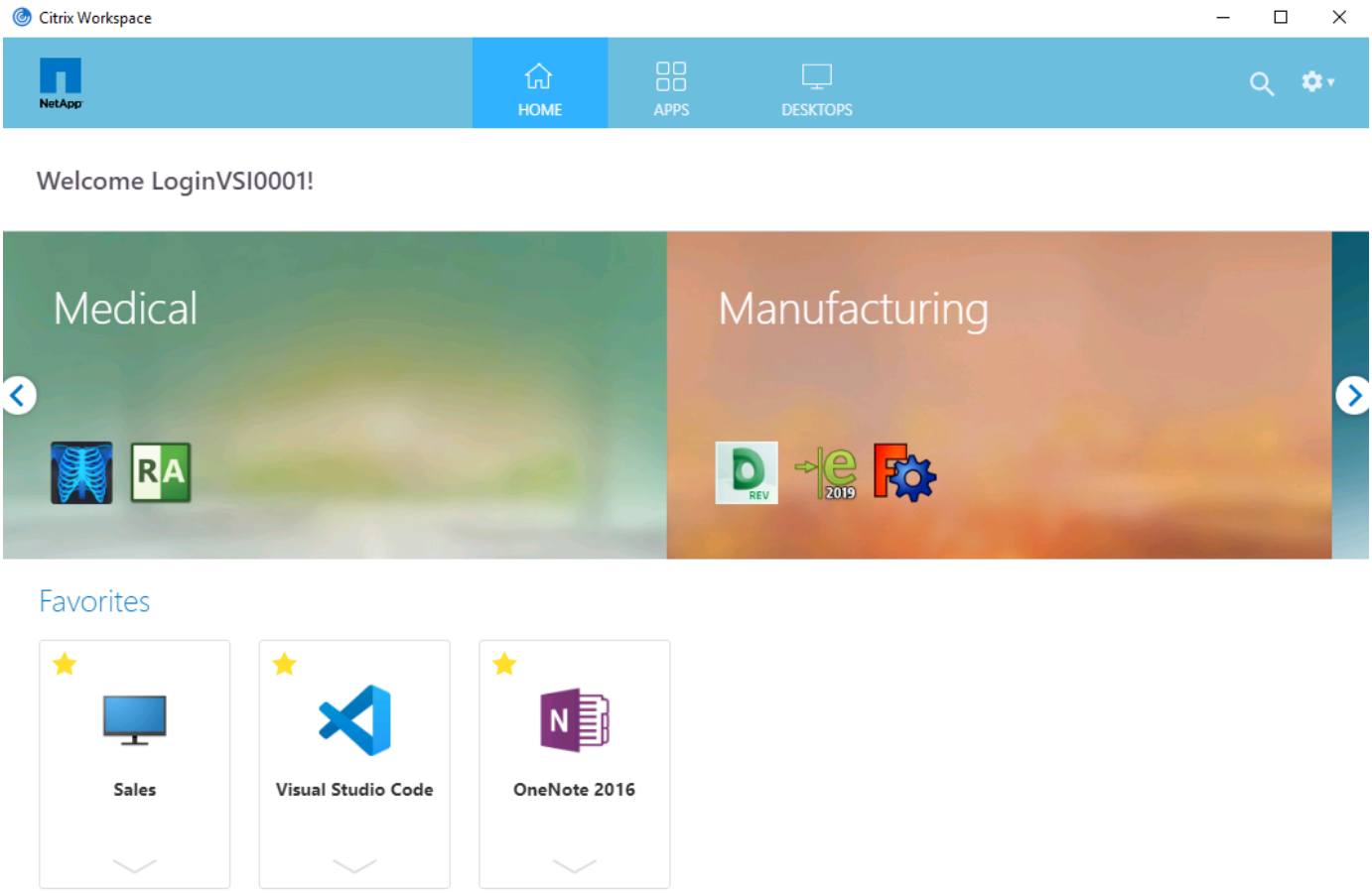
The key features of WEM are as follows:

- The ability to control resources for certain tasks or applications
- An easy interface to manage windows icons, network drives, start menu items, and so on
- The ability to reuse an old machine and manage it as a thin client
- Role-based access control
- Control policies based on various filters

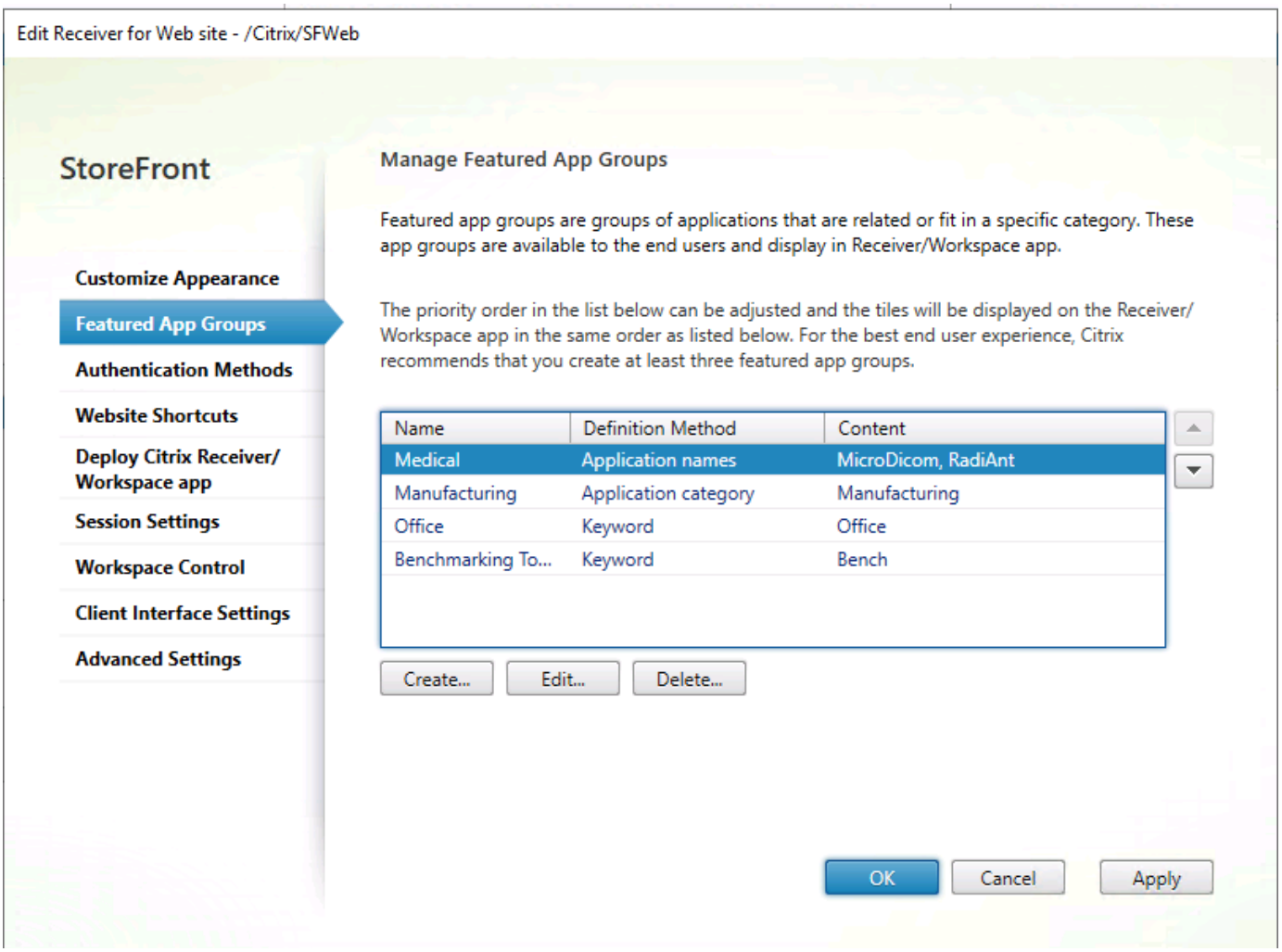
Access Layer

StoreFront

StoreFront consolidates resources published from multiple delivery controllers and presents unique items to users. Users connect to StoreFront and hides the infrastructure changes on the backend.



Users connect to StoreFront with the Citrix Workspace application or with a web browser. The user experience remains the same. An administrator can manage StoreFront using Microsoft Management Console. The StoreFront portal can be customized to meet customer branding demands. Applications can be grouped into categories to promote new applications. Desktops and applications can be marked as favorites for easy access. Administrators can also use tags for ease of troubleshooting and to keep track of resources in multitenant environments. The following screenshot depicts featured app groups.



Unified Gateway

To provide secure access to Citrix Virtual Apps and Desktops from the public internet to resources hosted behind a corporate firewall, Unified Gateway is deployed in a DMZ network. Unified Gateway provides access to multiple services like an SSL VPN, a reverse proxy to intranet resources, load balancer and so on by using a single IP address or URL.

Users have the same experience whether they are accessing the resources internally or externally to an organization. Application Delivery Controller (ADC) provides enhanced networking features for Virtual Apps and Desktops, and HDX Network Insights enhances HDX monitoring information with Citrix Director.

User Layer

Citrix Virtual Apps and Desktops enables users to access their workspace environment from anywhere with internet access and from any device with a web browser that has HTML5 support or with the Citrix Workspace application.

Users can be categorized as task workers, office workers, knowledge workers, and power users. Task workers primarily use predefined applications throughout the day for their work. Hosted Windows

Apps can serve their needs. Office workers requires desktop interfaces that run office applications, a web browser, and so on. Typically, they are not allowed to install applications on their workspace. They are best served by either a shared desktop with multi-session on server OS or with pooled desktops.

Knowledge workers typically require a desktop experience working with multiple applications simultaneously and must be able to persist the applications that they installed on their workspace. Static desktops (also referred to as personal desktops) allow this. Power users typically work on graphic-intensive applications or other applications that requires more hardware resources. Static desktops created with an appropriate master image address the needs of power users.

NetApp Value

Data Fabric

Infrastructure built with the data fabric powered by NetApp allows you to migrate data or perform disaster recovery from one site to another (including the cloud). The data in Citrix Virtual Apps and Desktops can be categorized as follows:

- Infrastructure components
- Machine images
- Applications
- User profiles
- User data

Based on your needs, sites can be configured as active/active or active/passive. Infrastructure components can be on-premises or in the cloud and accessed as a service. VM templates must be distributed to each site to provision desktop and application pools. Application layers, user profiles, and data are stored in SMB file shares that must be available on each site.

You can create a global namespace using Azure NetApp Files, NetApp Cloud Volumes ONTAP, and FlexGroup volumes at the location where most of your users reside. Other locations can use Global FileCache to cache the content locally on a file server. If Citrix ShareFile is preferred, NetApp StorageGRID provides high-performance, S3-compatible storage to host data on-premises with NAS gateway access.

Cloud Insights

Cloud Insights allows you to monitor, optimize, and troubleshoot resources deployed in the public cloud as well as on private datacenters.

Cloud Insights helps you in the following ways:

- **Reduce the mean time to resolution by as much as 90%.** Stop lengthy log hunting and failing to manually correlate infrastructure; use our dynamic topology and correlation analysis to pinpoint

the problem area immediately.

- **Reduce cloud infrastructure costs by an average of 33%.** Remove inefficiencies by identifying abandoned and unused resources and right-size workloads to optimized performance and cost tiers.
- **Prevent as much as 80% of cloud issues from affecting end users.** Stop searching through vast amounts of data to find the relevant item by using advanced analytics and machine learning to identify issues before they become critical outages.

Appendix iSCSI Device Configuration

Edit the multipath configuration file at `/etc/multipath.conf` as follows:

```
# This is a basic configuration file with some examples, for device mapper
# multipath.
## Use user friendly names, instead of using WWIDs as names.
defaults {
user_friendly_names yes
}
##
devices {
device {
vendor "SolidFir"
product "SSD SAN"
path_grouping_policy multibus path_selector "round-robin 0"
path_checker tur hardware_handler "0"
failback immediate rr_weight uniform rr_min_io 10 rr_min_io_rq 10
features "0"
no_path_retry 24
prio const
}
}
## Device black list
## Enter devices you do NOT want to be controlled by multipathd
## Example: internal drives
#blacklist {
#}
```

Where to Find Additional Information

To learn more about the information that is described in this document, review the following documents and/or websites:

- NetApp Cloud Central
  https://cloud.netapp.com/home

- NetApp Element Software Configuration for Linux
  https://www.netapp.com/us/media/tr-4639.pdf
- NetApp Product Documentation
  https://docs.netapp.com
- Citrix Security Recommendations
  https://www.citrix.com/content/dam/citrix/en_us/documents/white-paper/security-recommendations-when-deploying-citrix-xenserver.pdf
- Citrix Monitoring in Healthcare Environment with Goliath
  https://goliathtechnologies.com/webinar/on-demand/
- Citrix User Profile and FSLogix Integration
  https://youtu.be/dFpWdXIytJI
- Citrix App Layering Login VSI Test Results
  https://youtu.be/rWF5e84To4E
- Citrix App Layering FAQ
  https://www.citrix.com/blogs/2020/03/02/citrix-tips-citrix-app-layering-webinar-qa/
- Citrix App Layering Reference Architecture
  https://docs.citrix.com/en-us/tech-zone/design/reference-architectures/app-layering.html
- Citrix App Layering
  https://docs.citrix.com/en-us/citrix-app-layering/4/app-layering.pdf
- Multi-session write back to FSLogix Profile Container
  https://www.deyda.net/index.php/en/2020/03/27/citrix-virtual-apps-and-desktops-wem-2003-is-released/ - MultiSession_writeback_for_FSLogix_Profile_Container
- Citrix XAPI Backup
  https://support.citrix.com/article/CTX217618