Cisco and Hitachi Adaptive Solution with VMware Horizon 8 VDI

Lab Validation Report

MK-SL-268
June 2023

Feedback

Hitachi Vantara welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@HitachiVantara.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

Revision history

Changes	Date
Initial release	June 2, 2023

Lab Validation Report

This lab validation report presents the results and sizing considerations of the Cisco UCS X-Series with Hitachi Virtual Storage Platform (VSP) running ESXi 8.0 and the VMware Virtual Desktop Integration (VDI) solution to support the foundation of various virtualized workloads.

The Cisco and Hitachi Adaptive Solution is the latest converged infrastructure that is managed exclusively by Cisco Intersight. It is designed to meet the requirements of modern applications and improve operational efficiency, agility, and scale with its modular architecture. Additionally, with the introduction of the Cisco UCS X440p PCIe node, end users have additional capabilities of choosing Nvidia Graphical Processing Units (GPUs) to further drive computational requirements of heavy virtualized workloads, which are common for VDI solution deployments.

VDI offers several advantages, such as user mobility, ease of access, flexibility, and greater security. In the past, its high-performance requirements made it costly and challenging to deploy on legacy systems, which posed a barrier for many businesses. However, with the rise in enterprise adoption of converged infrastructure, VDI offers a solution that provides scalability and high performance at a lower cost. Note that for deployments with a few hundred users, HCI is a better fit. The CHAS converged infrastructure solution is more suitable for large companies with thousands of users.

VDI is an ideal match for the Cisco and Hitachi Adaptive solution for converged infrastructure; administrators can instantly consolidate their applications onto an efficient modular, scalable system, which allows cloud-based infrastructure management for the entire stack via Cisco Intersight. Furthermore, VSP storage systems truly enable a simplified approach to managing the datacenter by allowing multiple management options and features, which allow further operational expenditure savings via built-in capacity savings function. Cisco UCS X-Series backed by Hitachi VSP provides customers a future-proof converged infrastructure stack backed by one of the most reliable enterprise storage systems, which guarantees 100% data availability.

Solution design

Cisco and Hitachi Adaptive Solutions for Converged Infrastructure is a validated reference architecture targeting Virtual Server Infrastructure (VSI) implementations. The architecture is built around the Cisco UCS X-Series and Hitachi Virtual Storage Platform connected by Cisco MDS Multilayer SAN switches, and further enabled with Cisco Nexus switches. These components come together to form a powerful and scalable design, built on the best practices of both Enterprise grade companies to create an ideal environment for virtualized systems.

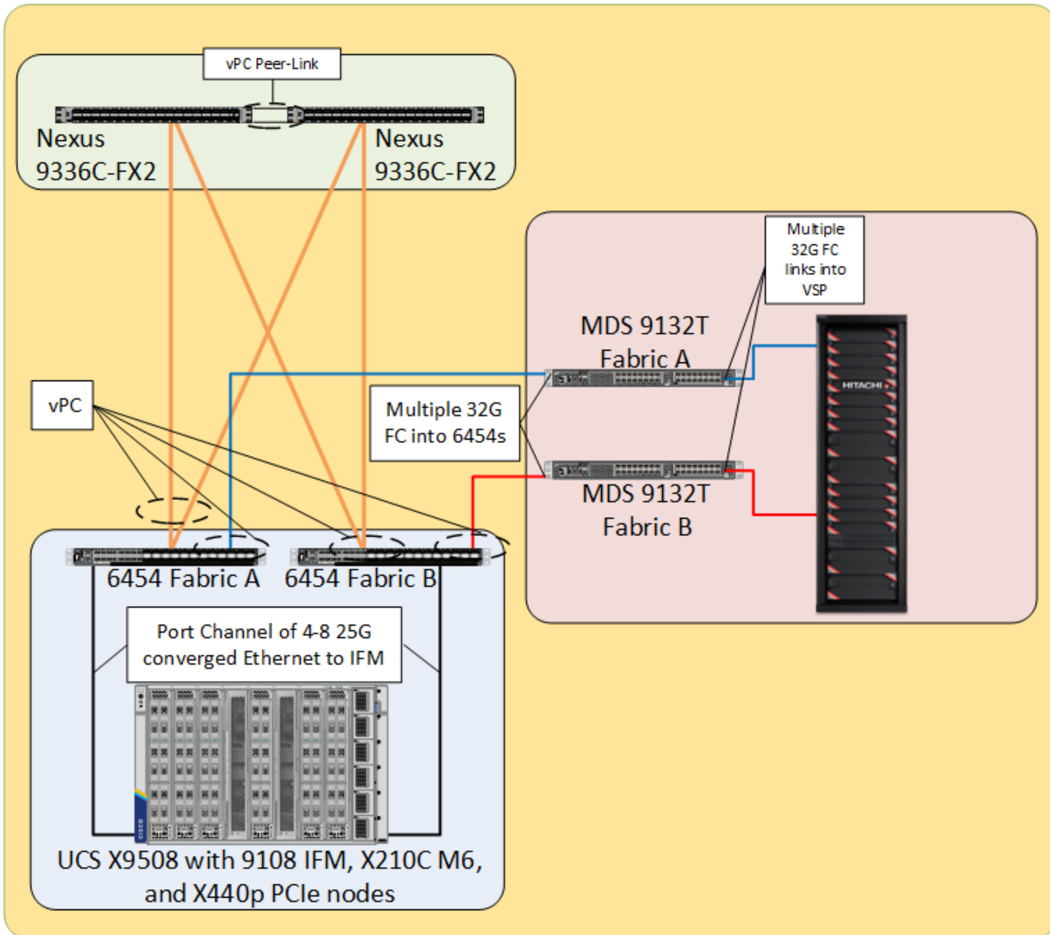
The solution is built and validated for a topology featuring a pair of Cisco UCS Fabric Interconnects as well as a VSP storage system, with both using the same MDS and Nexus switching infrastructure.

The topology shown in the following figure leverages the following products:

- Cisco Nexus 9336C-FX2 — 100 Gb capable, LAN connectivity to the UCS compute resources.
- Cisco UCS 6454 Fabric Interconnect (FI) — Unified management of UCS compute, and the compute networks access to storage and networks.
- Cisco UCS X9508 Chassis - The Cisco UCS X-Series Modular System begins with the Cisco UCS X9508 Chassis engineered to be adaptable and future ready. It is a standard, open system designed to deploy and automate faster in concert with a hybrid cloud environment.
- Cisco UCS 9108 25G Intelligent Fabric Module (IFM) - Brings the unified fabric into the blade server enclosure, providing connectivity between the blade servers and the fabric interconnect, simplifying diagnostics, cabling, and management.
- Cisco UCS X210c M6 Compute Node — High powered, versatile blade server, conceived for virtual computing.
- Cisco Virtual Interface Card (VIC) 14425 - The Cisco UCS VIC 14425 enables a policy-based, stateless, agile server infrastructure that can present to the host PCIe standards-compliant interfaces that can be dynamically configured as either NICs or HBAs.
- Cisco UCS X440p PCIe Node with Nvidia A16 GPU - PCIe resource node to integrate into the Cisco UCS X-Series Modular System for additional compute resources.
- Cisco UCS X9416 X-Fabric Module - UCS X-Fabric Technology module is PCIe Gen functionality and efficient administration.

Note: For PCIe compute node deployments, two X9416 X-Fabric Modules must be installed in the rear backplane of the Cisco UCS X9508 chassis.

Note: Any component mentioned in this reference architecture guide can be supplemented with comparable products as long as they are listed in both Cisco and Hitachi hardware compatibility listings.



The Cisco UCS X210c M6 nodes in this topology are hosted within a Cisco UCS X9508 chassis and connect into the fabric interconnects from the chassis using the Cisco UCS 9108 IFM. The 9108 IFM supports 25 G connections into the 10/25 G ports of the Cisco UCS 6454 FI, delivering a high port availability that may fit well in a branch office setting.

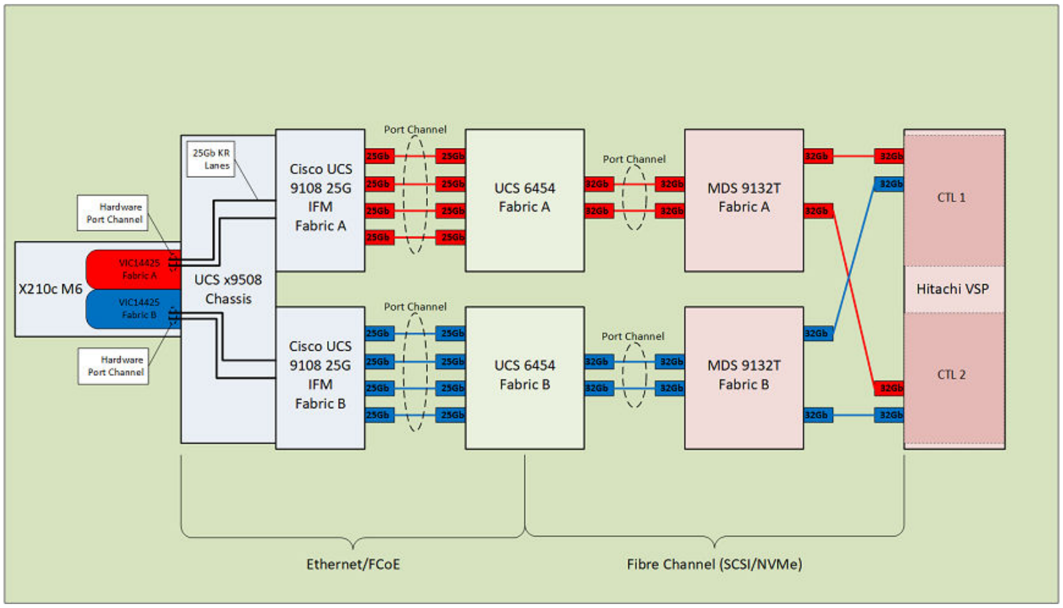
Management components for this architecture also include:

- Cisco Intersight (required) — Comprehensive unified visibility across UCS domains as well as a datacenter ecosystem in conjunction with automation workflows, along with proactive alerts and enablement of expedited Cisco technical assistance center (TAC) communications
- Cisco Data Center Network Manager/Nexus Dashboard Fabric Controller (optional) — Multi-layer network configuration and monitoring

The topology has been validated for vSphere 8.0 to accommodate a larger range of expected customer deployments. Previous and newer versions of vSphere, as well as other vendor hypervisors, might be supported. These additional hypervisors must be within the compatibility and interoperability matrices from Cisco at https://ucshcltool.cloudapps.cisco.com/public/ as well as Hitachi at https://compatibility.hitachivantara.com/.

The following illustration shows the end-to-end data path, where the VSP is connected to the Cisco MDS 9132T at 32 Gb Fibre Channel speeds. Fibre Channel connections are then converged into the 6454 FIs and those connections are then port channeled into the Cisco UCS 9108 IFM to feed into the X210C VIC 14425.

Fabric interconnects are connected to Nexus switches using vPCs. With vPC connections we can achieve both resiliency for hardware (FI or Nexus switch) failure and traffic load balancing.



Hardware and software versions

The following table lists the validated hardware and software versions used for this solution. Component and software version substitution from what is listed is considered acceptable within this reference architecture, but substitutions will need to comply with the hardware and software compatibility matrices from both Cisco and Hitachi.

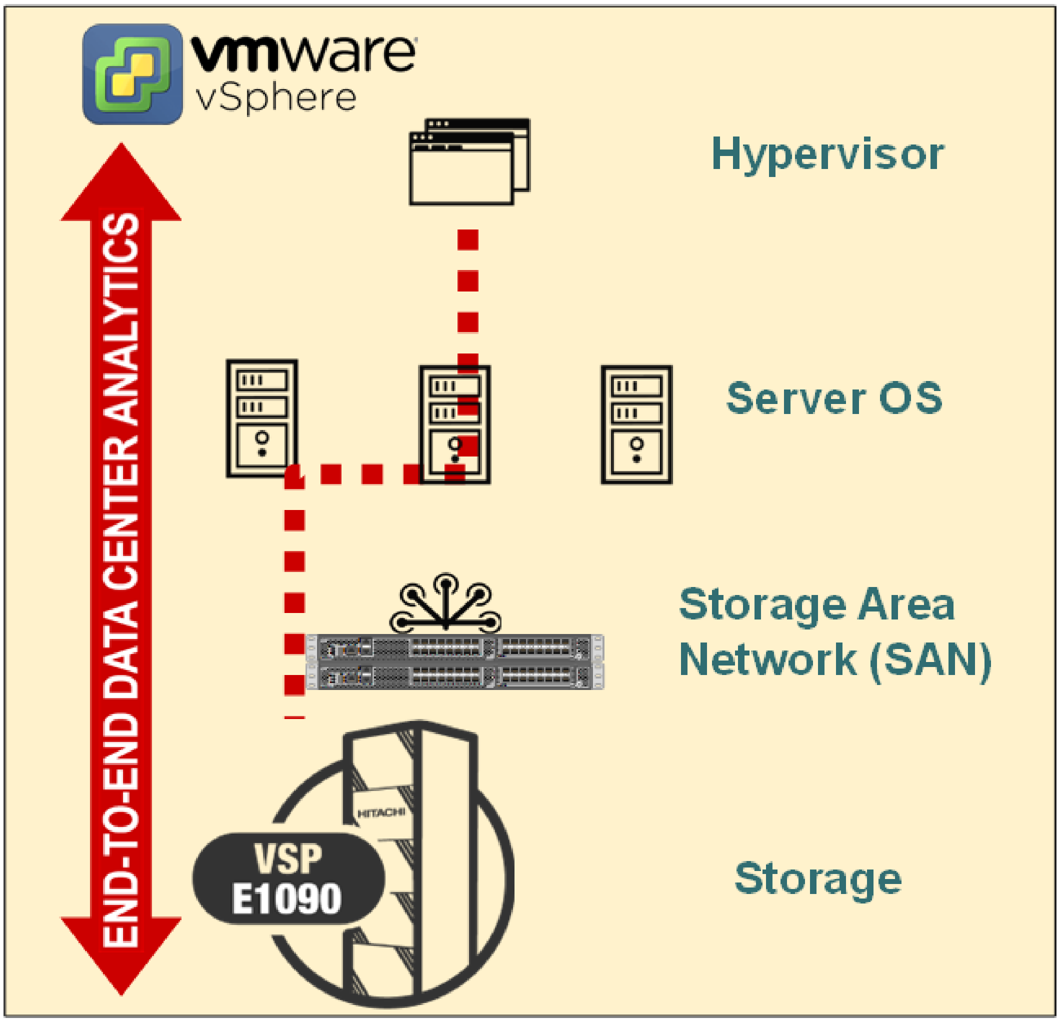
- Cisco UCS Hardware Compatibility Matrix: https://ucshcltool.cloudapps.cisco.com/public/
- Cisco Nexus and MDS Interoperability Matrix: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/mds9000/interoperability/matrix/intmatrx/Matrix1.html
- Cisco Nexus Recommended Releases for Nexus 9K: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/recommended_release/b_Minimum_and_Recommended_Cisco_NX-OS_Releases_for_Cisco_Nexus_9000_Series_Switches.html
- Cisco MDS Recommended Releases: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/mds9000/sw/b_MDS_NX-OS_Recommended_Releases.html
- Hitachi Vantara Interoperability: https://support.hitachivantara.com/en_us/interoperability.html sub-page -> (VSP 5x00, G1X00, F1500, E-series, Gxx0, Fxx0, VSP, HUS VM VMWare Support Matrix)

Component		Software Version/Firmware Version
Network	Cisco Nexus 9336C-FX2	NX-OS 10.2(4)
Compute	Cisco UCS Fabric Interconnect 6454	4.2(3b)
	Cisco UCS 9108 IFM	4.2(3b)
	Cisco UCS X210C M6	5.0(4a)
	Cisco UCS X-Fabric Module UCSX-F-9416	N/A
	Cisco UCS X440p PCIe Node	N/A
	Nvidia A16 GPU and software license	94.07.54.00.45-G171.0200.00.04
	Cisco VIC 14425 UCSX-V4-Q25GML	5.2(3c)
	VMware vSphere ESXi 8	VMware-ESXi-8.0-20513097-Custom-Cisco-4.2.3-b
	ESXi 8.0 NENIC	1.0.45.0-1
	ESXi 8.0 NFNIC	5.0.0.37-1

Component		Software Version/Firmware Version
	Nvidia ESXi driver	NVD-VMware_ESXi_8.0.0_Driver_525.105.14-1OEM.800.1.0.20613240
	Nvidia GPU mgmt. daemon driver	nvdgpumgmtdaemon_525.105.14-1OEM.700.1.0.15843807
	VM Virtual Hardware Version	20
Storage	Hitachi VSP E-1090	93-06-42-x0/01
	Cisco MDS 9132T	9.3(2)
Software	VMware Horizon	8.8.0-21073894 (2212)
	Login VSI	4.1.40
	Windows 10 Enterprise 64 bit	Version 22H2 (OS Build 19045.2846)
	Microsoft SQL Database Server	2019
	Nvidia vGPU Driver (Windows)	528.89
	Nvidia DLS License Server	DLS 2.1

Hitachi Ops Center Analyzer

Hitachi Ops Center Analyzer allows end-to-end metric collection from compute hosts, Fibre Channel switches, and backend storage systems. This software was used during the performance analysis of Horizon VDI being backed by the Hitachi Virtual Storage Platform. See https://knowledge.hitachivantara.com/Documents/Management_Software/Ops_Center/10.9.x/Analyzer for more information on Hitachi Ops Center Analyzer and its capabilities.



Test environment configuration

The following sections describe the configuration that was validated in the Hitachi lab.

Storage

Hitachi Virtual Storage Platform (VSP) configuration is based on UCS with VSP best practices as outlined in the Cisco and Hitachi Adaptive solutions for Converged Infrastructure. Additionally, port mapping can be obtained from the Cisco and Hitachi Adaptive Solution with Cisco UCS X-Series Modular System and Hitachi Virtual Storage platform Reference Architecture Guide in the *Physical cabling for the UCS 6454 with Hitachi Virtual Storage Platform* section.

VSP E1090 configuration

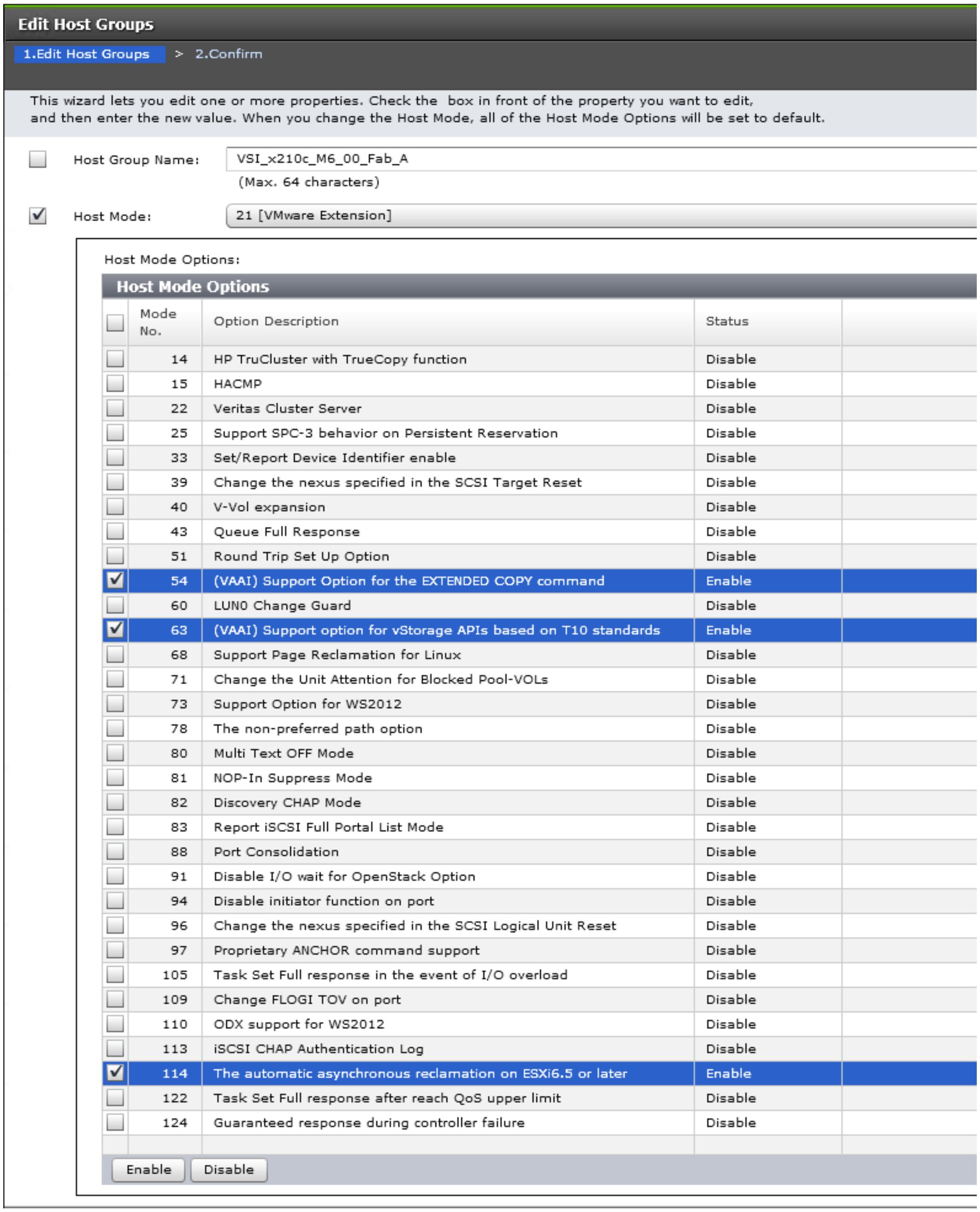
The following table defines the configuration of the VSP E1090 used for validation.

Hitachi Virtual Storage Platform E1090 Configuration
2 × controllers (with controller-based compression)
863 GB cache
24 × 15.3 TB NVMe (SNB5A-R15RNC) drives used
RAID 6 (6D+2P)
Hitachi Dynamic Provision (HDP) Pool
4 × 32 Gbps Fibre Channel ports used (4Port FC32R)

During testing the 4Port FC32R Channel Board (CHB) was used for SCSI protocols. With this CHB, 4G, 8G,16G and 32G SFPs are supported with FC-SCSI. During testing a full 32 GB data path was used with VSP controller connections spanning into fabric A and fabric B MDS, as outlined in the Cisco and Hitachi Adaptive Solutions with Cisco UCS X-Series Reference Architecture Guide at https://www.hitachivantara.com/en-us/pdf/architecture-guide/cisco-adaptive-solutions-with-cisco-ucs-x-series-modular-system-vsp.pdf. In this configuration NVMe mode and SCSI mode can coexist within the same CHB, but during testing all 4 ports were in SCSI mode.

After the VSP has been deployed and configured, Host Groups used for allocation to VMware virtualized environments require the following Host Mode and Host Mode Options (HMO):

- Host Mode: 21 [VMware Extension]
- Host Mode Option: 54 (VAAI) Support Option for EXTENDED COPY command
- Host Mode Option: 63 (VAAI) Support Option for vStorage APIs based on T10 standards
- Host Mode Option: 114 The automatic asynchronous reclamation on ESXi6.5 or later



Edit Host Groups

1.Edit Host Groups > 2.Confirm

This wizard lets you edit one or more properties. Check the box in front of the property you want to edit, and then enter the new value. When you change the Host Mode, all of the Host Mode Options will be set to default.

Host Group Name: VSI_x210c_M6_00_Fab_A (Max. 64 characters)

Host Mode: 21 [VMware Extension]

Host Mode Options:

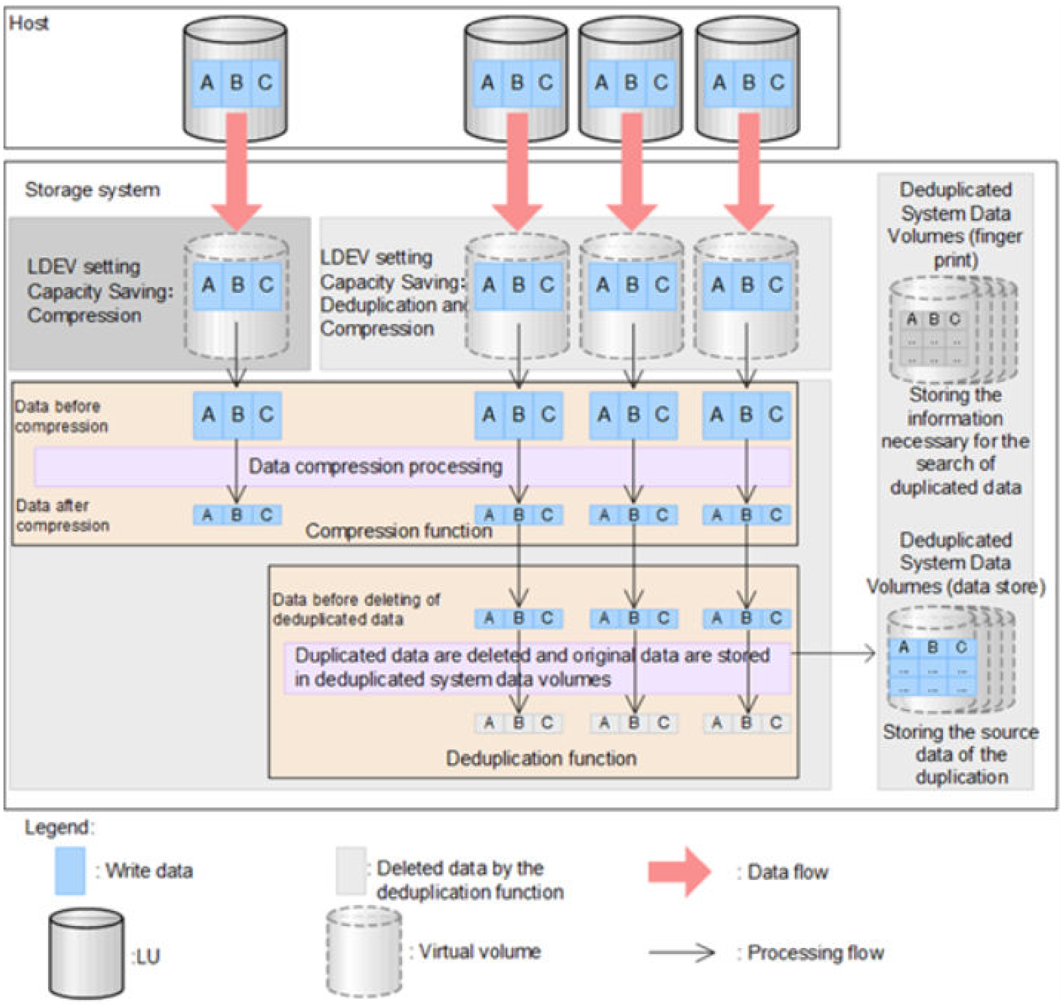
Host Mode Options

Mode No.	Option Description	Status
14	HP TruCluster with TrueCopy function	Disable
15	HACMP	Disable
22	Veritas Cluster Server	Disable
25	Support SPC-3 behavior on Persistent Reservation	Disable
33	Set/Report Device Identifier enable	Disable
39	Change the nexus specified in the SCSI Target Reset	Disable
40	V-Vol expansion	Disable
43	Queue Full Response	Disable
51	Round Trip Set Up Option	Disable
54	(VAAI) Support Option for the EXTENDED COPY command	Enable
60	LUN0 Change Guard	Disable
63	(VAAI) Support option for vStorage APIs based on T10 standards	Enable
68	Support Page Reclamation for Linux	Disable
71	Change the Unit Attention for Blocked Pool-VOLs	Disable
73	Support Option for WS2012	Disable
78	The non-preferred path option	Disable
80	Multi Text OFF Mode	Disable
81	NOP-In Suppress Mode	Disable
82	Discovery CHAP Mode	Disable
83	Report iSCSI Full Portal List Mode	Disable
88	Port Consolidation	Disable
91	Disable I/O wait for OpenStack Option	Disable
94	Disable initiator function on port	Disable
96	Change the nexus specified in the SCSI Logical Unit Reset	Disable
97	Proprietary ANCHOR command support	Disable
105	Task Set Full response in the event of I/O overload	Disable
109	Change FLOGI TOV on port	Disable
110	ODX support for WS2012	Disable
113	iSCSI CHAP Authentication Log	Disable
114	The automatic asynchronous reclamation on ESXi6.5 or later	Enable
122	Task Set Full response after reach QoS upper limit	Disable
124	Guaranteed response during controller failure	Disable

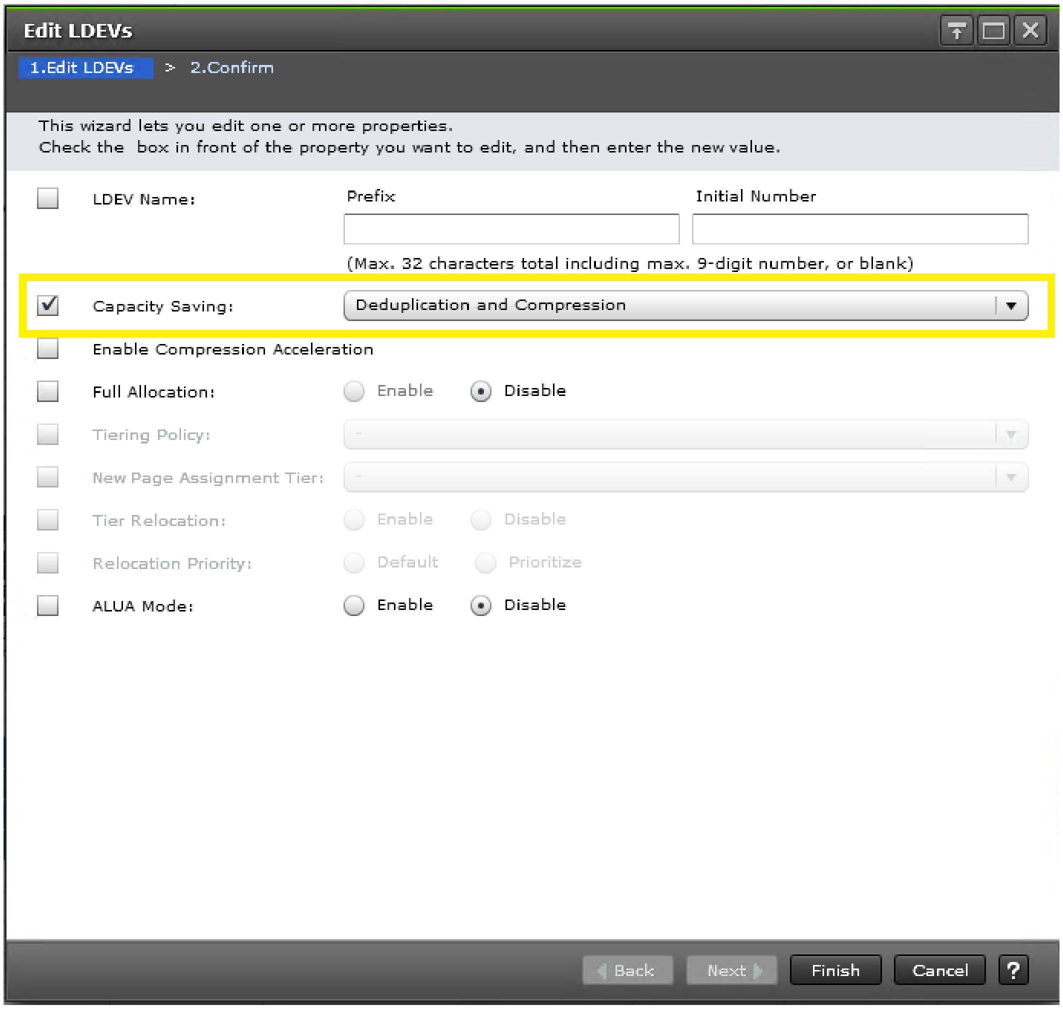
Enable Disable

Capacity saving function

The capacity saving feature is native to the Hitachi Virtual Storage Platform; when enabled, data deduplication and compression is performed to reduce the size of data to be stored. Capacity saving can be enabled on dynamic provisioning volumes in Hitachi Dynamic Provisioning (HDP) pools and Hitachi Dynamic Tiering (HDT) pools. You can use the capacity saving function on internal flash drives only, including data stored on encrypted flash drives and external storage. Within this lab validation report, the capacity saving function was enabled to show the use case benefits in conjunction with VDI deployments. The following image demonstrates how the Hitachi Virtual Storage Platform makes use of the capacity saving function after enabled on the DP pool.



See https://knowledge.hitachivantara.com/Documents/Management_Software/SVOS/9.8.1/Volume_Management_-_VSP_E_Series/Data_Reduction/02_About_capacity_savingfor more information about the Hitachi VSP capacity saving feature. To enable capacity savings, volumes must be selected at the LDEV level, and then under Edit LDEVs select Capacity Saving and Deduplication and Compression.



- Deduplication — The data deduplication function deletes duplicate copies of data written to different addresses in the same pool and maintains only a single copy of the data at one address. The deduplication function is enabled on a Dynamic Provisioning pool and then on the desired DP-VOLs in the pool. When deduplication is enabled, data that has multiple copies on the DP-VOLs assigned to that pool is removed.
- Compression — The data compression function enables you to convert the stored data into data with smaller data size by encoding without reducing the amount of data information. The LZ4 compression algorithm is used for data compression using software with the DKCMAIN microcode. The data compression function can be enabled for each DP-VOL used for Dynamic Provisioning or Dynamic Tiering.

Compute

- Cisco UCS X210c Node

  The following table defines the configuration of the two Cisco UCS X210c nodes used for validation.

Cisco UCS X210c Node Configuration
Intel(R) Xeon(R) Gold 6338 CPU @ 2.00GHz
512 GB Memory
VIC 14425
UCSX-V4-PCIME

- Cisco UCS X440p Node with NVIDIA GPU

  The Cisco UCS X440p PCIe Node is the first PCIe resource node to integrate into the Cisco UCS X-Series Modular System. The Cisco UCS X9508 chassis has eight node slots, up to four of which can be X440p PCIe nodes when paired with a Cisco UCS X210c M6 compute node. The UCS X440p PCIe node supports two x16 full-height, full-length dual slot PCIe cards, or four x8 full-height, full-length single slot PCIe cards and requires both UCS 9416 X-Fabric modules for PCIe connectivity. This provides up to 16 GPUs per chassis to accelerate your applications with the UCS X440p nodes. If your application needs even more GPU acceleration, up to two additional GPUs can be added on each UCS X210c compute node.

  Within this lab validation report, 2 X440p nodes with a single Nvidia A16 GPU in each was used for validation. Each GPU provides the following specifications.

GPU Memory	4 × 16 GB GDDR6 with error-correcting code (ECC)
GPU Memory Bandwidth	4 × 200 GBps
Max Power Consumption	250W
Interconnect	PCI Express Gen 4 x16
Form Factor	Full height, full length (FHFL) dual slot
Thermal	Passive
vGPU Profile Used	nvidia_a16-1b
vGPU Software Support	NVIDIA Virtual PC (vPC), NVIDIA Virtual Applications (vApps), NVIDIA RTX Virtual Workstation (vWS), NVIDIA Virtual Compute Server (vCS), and NVIDIA AI Enterprise
vGPU Profiles Supported	See the https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/solutions/resources/

	documents1/Virtual-GPU-Packaging-and-Licensing-Guide.pdf
NVENC | NVDEC	4x | 8x (includes AV1 decode)
Secure and measured boot with hardware root of trust	Yes (optional)
NEBS Read	Level 3
Power Connector	8-pin CPU

Optionally, the Cisco UCS X440p has the following GPU options:

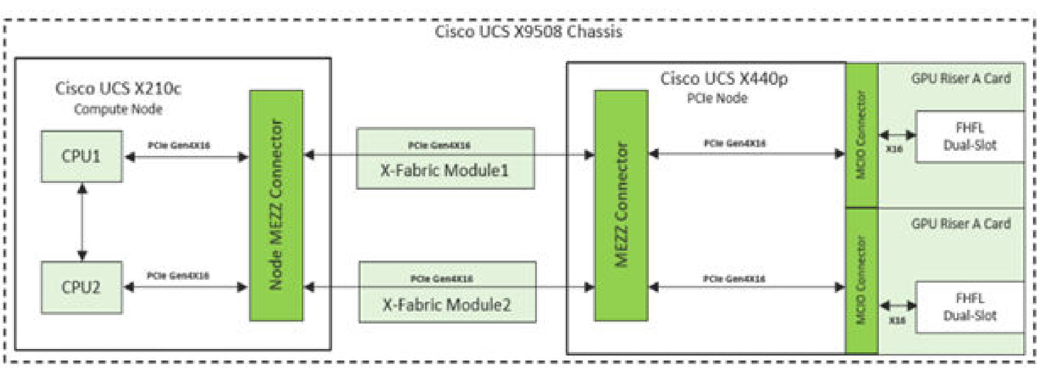
- Nvidia A100 Tensor Core GPU
- Nvidia A16 GPU
- Nvidia A40 GPU
- Nvidia T4 Tensor Core GPU

Note: Riser A configuration for Cisco UCS X440p node will be required if using GPU models A100, A16, and A40. Alternatively, riser B configuration will be required if using T4 tensor core GPU.

Note: For X210c server connectivity into the UCS X440p, a UCSX-V4-PCIME mezzanine connector card needs to be installed into the server so that connectivity into the X9416 X-Fabric module is established.

The following image shows backplane connectivity from the chassis X-Fabric modules into the UCS X440p node. After you install the correct riser configuration for the X440p as well as UCSX-V4-PCIME into the X210c node, an inventory refresh within Intersight will detect the node as well as the GPU installed. Furthermore, during initial installation Intersight will download and install the correct packages for the Nvidia GPU.

Note: In certain cases, after appropriate software and firmware is installed, Intersight may report incompatibilities, which is a known bug that can be referred to on Cisco's bug search tool at https://bst.cloudapps.cisco.com/bugsearch/bug/CSCwe63378.

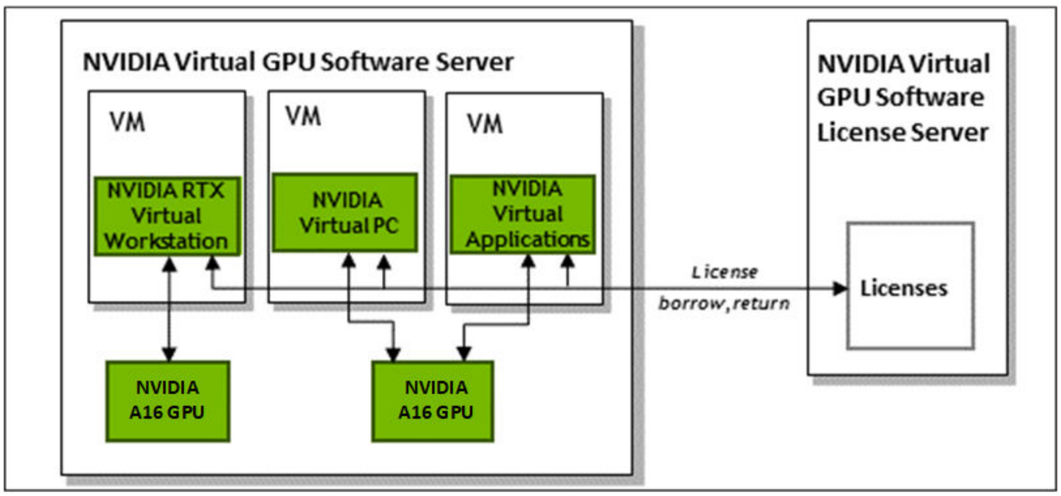


NVIDIA Delegated License Service (DLS) Server

Before using the Nvidia GPU for accelerated workloads, you must configure licensing to be allocated and used by the virtual desktops deployed by VDI. Nvidia enables this in two methods:

- Cloud License Service (CLS) instance. A CLS instance is hosted on the NVIDIA Licensing Portal.
- Delegated License Service (DLS) instance. A DLS instance is hosted on-premises at a location that is accessible from your private network, such as inside your data center.

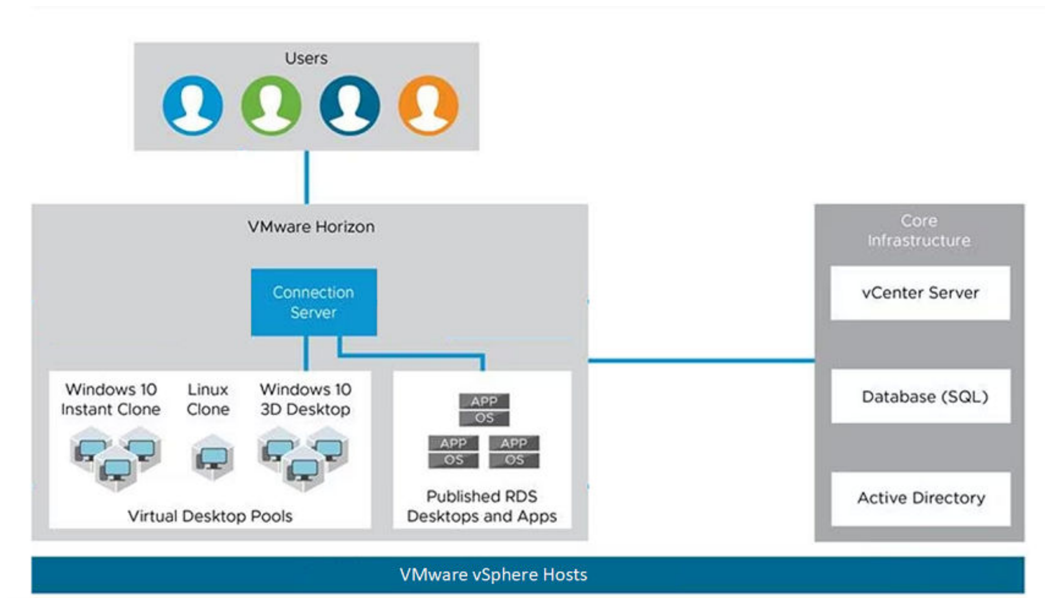
Within this validation report, a DLS instance was used, The following figure shows the software enforcement of the Nvidia vGPU licensing. The DLS instance can be deployed in the form of an OVA package. Also, you must install the correct Windows driver for the Nvidia graphics card on their master image. This allows VMs deployed by VDI to automatically lease and release licenses when spun up or deleted. Both DLS OVA and any drivers for the OS or ESXi host can be obtained from Nvidia's licensing portal at Nvidia's licensing portal https://ui.licensing.nvidia.com/.



Application

- VMware Horizon VDI

  Virtual Desktop Infrastructure (VDI) is an innovative technology that uses virtual machines to manage and provide virtual desktops. Rather than being confined to a specific physical device, a virtual desktop is a preconfigured image of an operating system, and its applications can be accessed from any compatible device. With VDI, desktop environments can be hosted on a centralized server and deployed to end-users on demand. Within this lab validation report, we have leveraged VMware Horizon 8 connection server backed by core infrastructure, which includes VMware vCenter 8 as well as Microsoft SQL database server and Microsoft active directory services to support VDI. Optionally, you have the ability to use published applications; but in the Hitachi lab, validation was done using instant clones. In summary, VDI allows organizations to deploy VMware VDI desktops that can be accessed from anywhere, at any time, making it an ideal solution for remote and on-prem users.



VMware optimization tool

The Windows OS Optimization Tool for VMware Horizon optimizes golden images by automating the removal of unnecessary settings and disabling unnecessary features typically present in a physical Windows machines to improve VDI performance and user experience. The follow image shows the process of using the VMware optimization tool.

Within the context of this report, the VMware optimization tool was used to achieve better login VSI scores as outlined in the following sections.

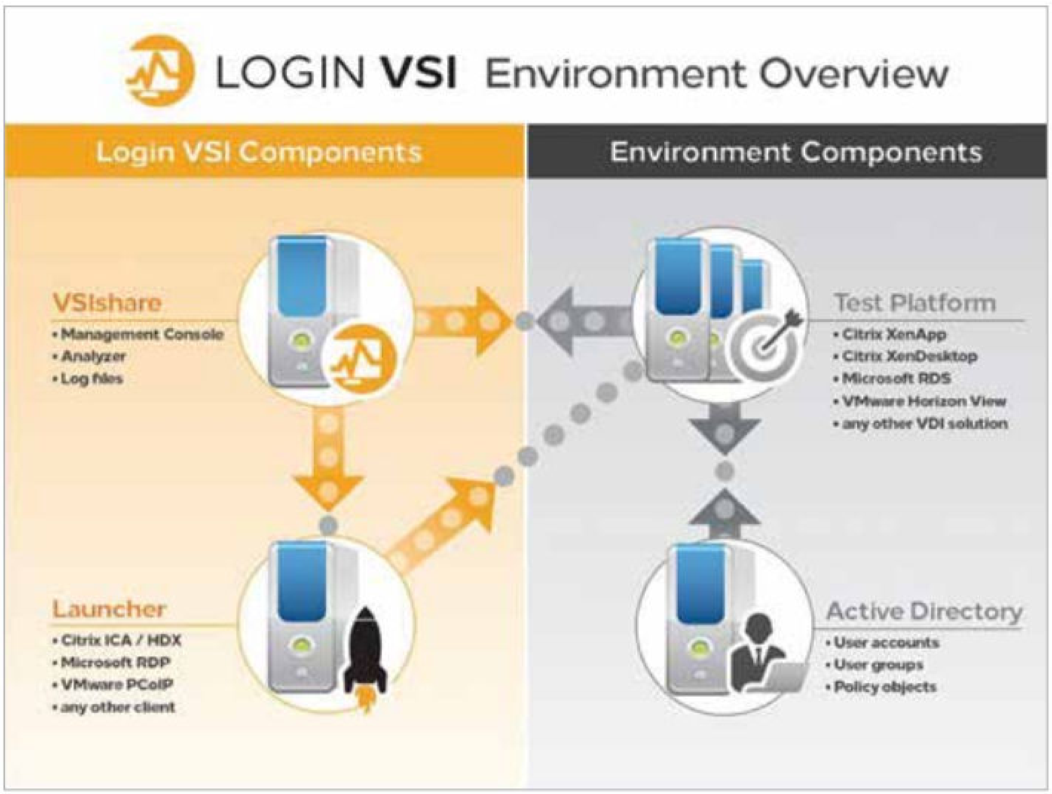
Login VSI

Login VSI successfully predicts, validates, and manages the performance of virtualized desktop environments making it easy to load, test, benchmark, and plan capacity to improve end user experience and productivity for even the most complex virtualized desktop environments. Login VSI tests performance using virtual users, so your real users benefit from consistently great performance.

With Login VSI, you can gain performance insights that enable you to do the following:

- Predict the performance impact of necessary updates and upgrades
- Know the maximum user capacity of your current infrastructure
- Understand the end users' perspective on performance

With agentless installation and minimal infrastructure requirements, Login VSI works in any Windows-based virtualized desktop environment including VMware Horizon, Citrix XenDesktop, and Microsoft Remote Desktop Services (Terminal Services). The following is a architecture diagram that represents both Login VSI and environment components.



When used for benchmarking, the product measures the total response times of several specific user operations being performed within a desktop workload, in a scripted loop. The baseline is the measurement of the specific operational response times performed in the desktop workload, measured in milliseconds (ms). Two values are very important: VSIbase and VSImax.

- VSIbase — A score reflecting the response time of specific operations performed in the desktop workload when there is little or no stress on the system. A low baseline indicates a better user experience and a well-tuned desktop image, resulting in applications responding faster within the environment.
- VSImax — The maximum number of desktop sessions attainable on the host before experiencing degradation in both host and desktop performance.

Both values, VSIbase and VSImax, offer undeniable proof (vendor independent, industry standard, and easy to understand) to innovative technology vendors of the power, the scalability, and the benefits of their software and hardware solutions, in a virtual desktop environment. The following table shows Login VSI base score ratings.

Login VSI Base Score Range	Rating
0-799	Very Good
800-1199	Good
1200-1599	Fair
1600-1999	Poor
2000-9999	Very Poor

Analysis of results

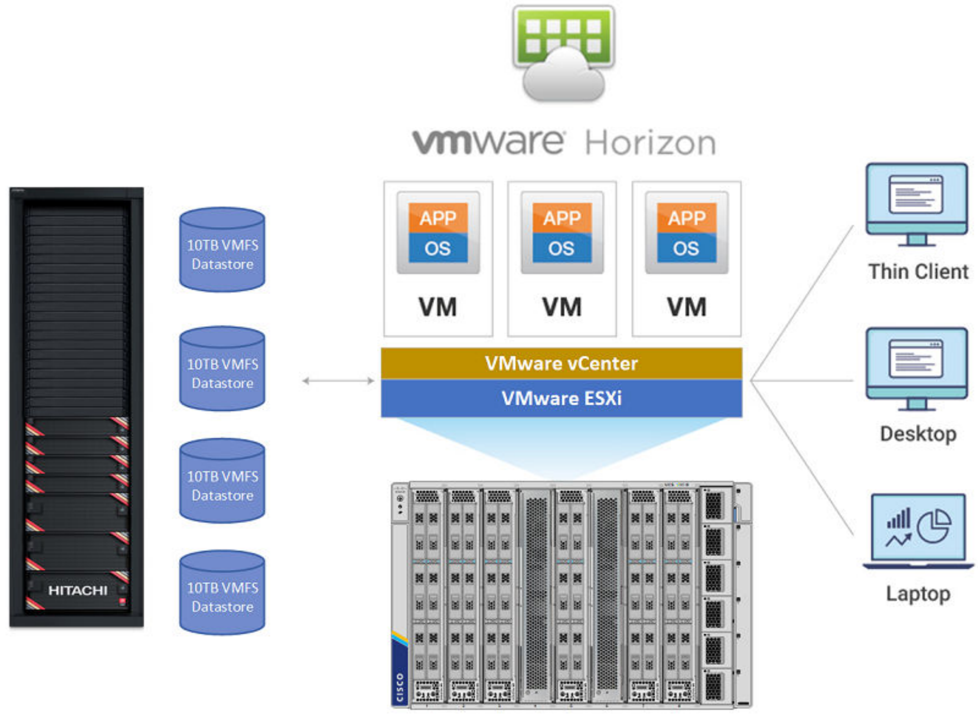
This section provides test details.

Test methodology

Testing within this lab validation report was done in two phases, which include without and with GPUs. For GPU-less test cases various user workloads were benchmarked such as task, knowledge, and power users to validate the behavior of both compute and storage resources. For GPU test cases, only the power user was utilized as in most cases low resource workloads such as task and knowledge are not beneficial, for hardware resources of the NVIDIA A16 GPU.

Additionally, both phases were conducted with 2 ESXi hypervisors being managed by VMware vCenter. This deployment configuration also used a Storage Area Network (SAN) fabric backed by the Hitachi VSP E1090 system. The Hitachi VSP provided 4 VMFS datastores from a HDP pool to enable VMware Horizon VDI to deploy VMs that host user files and applications.

Note: It is recommended that management nodes are kept on a separate cluster.



Within each gold master template used by VMware Horizon to create user sessions, applications such as Microsoft Office 2019, Adobe Reader, Adobe Flash, and Java were utilized to mimic real user scenarios during testing. The following table indicates the resources used by each user type during testing.

User Profiles	CPU	Memory	Disk Size
Task User	2 vCPU	2 GB	40 GB
Knowledge User	2 vCPU	4 GB	60 GB
Power User	4 vCPU	8 GB	80 GB

Sizing considerations

For production environments, deployed configurations will vary based on the amount of resources on the ESXi hypervisor as well as the NVIDIA GPU available for allocation. In the following test cases this table shows how many VMs were deployed within the test bed for benchmarking.

Test Case	Number of VMs Deployed
Task User (Non-GPU)	400
Knowledge User (Non-GPU)	210
Power User (Non-GPU)	110
Power User (GPU)	110

Non-GPU testing

The following results were run within the Hitachi lab, which validated performance of various user workloads using no GPUs.

Task User

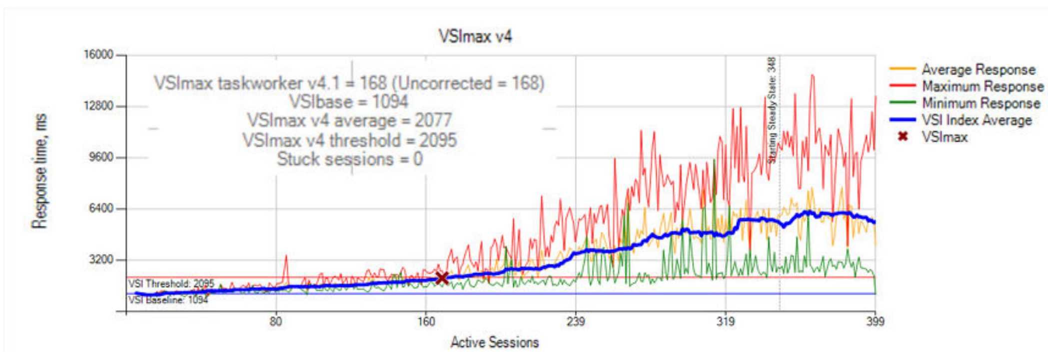
The task user workload runs fewer applications and starts/stops them less frequently than the other workloads, resulting in lower CPU, RAM, and IO usage based on density of VMs deployed. The Task Worker workload uses the following applications:

- Adobe Reader
- Internet Explorer
- Microsoft Excel
- Microsoft Outlook
- Microsoft Word

Results

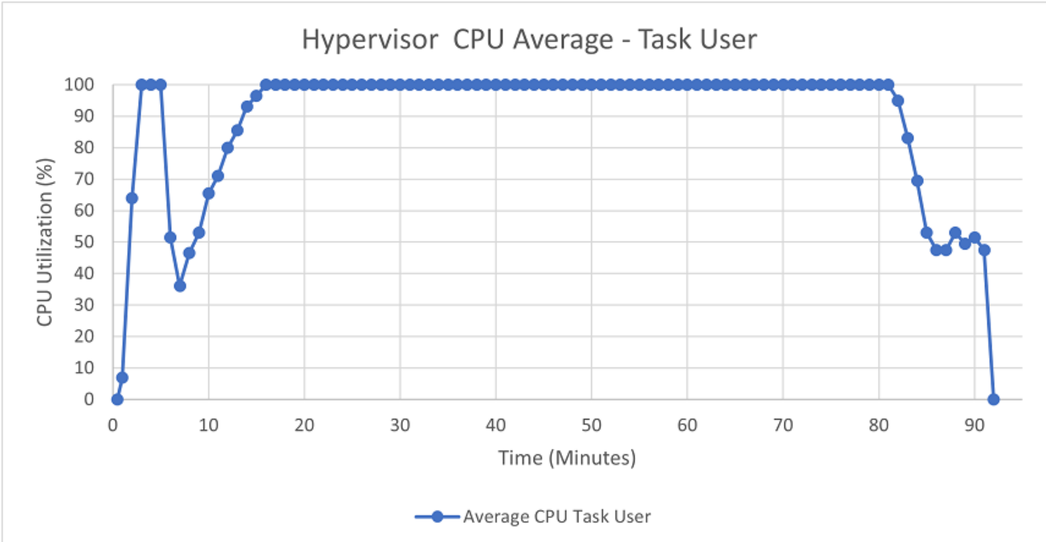
Login VSI

The Login VSI Max user experience score shown in the following figure for the test with 400 task users was at 170 desktops. This means that the system was saturated because of resource constraints when 170 users were utilizing the VDI environment, even though it was able to complete the test with a login VSI base score of 1094.



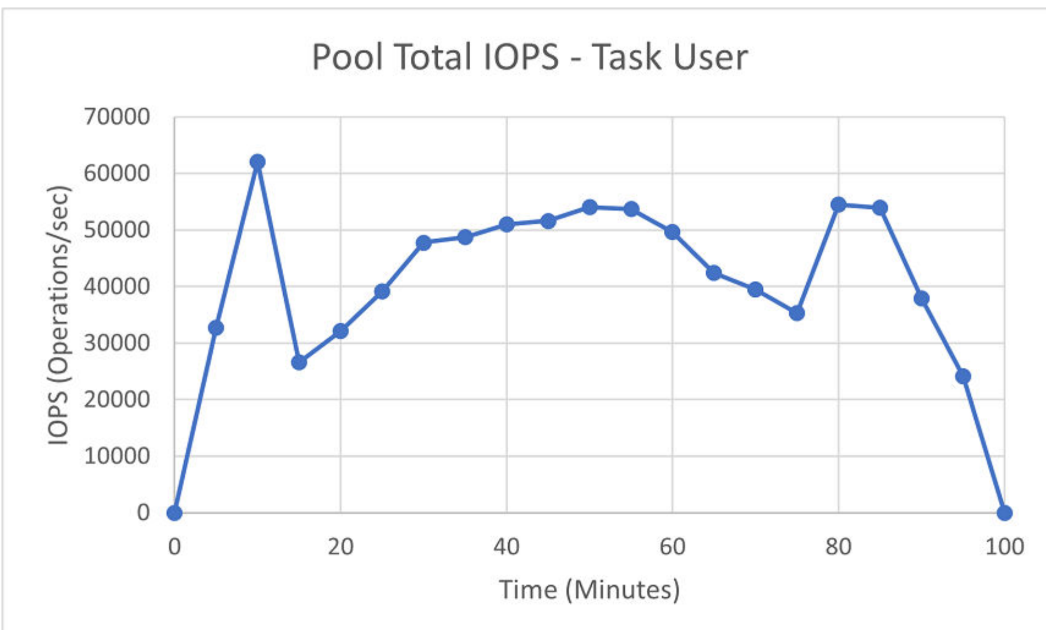
Hypervisor performance

Because the maximum number of VMs possible were deployed with the number of resources available on the ESXi hypervisor, 400 VMs with the task user profile produced a high load on the hypervisor CPU. In scenarios where administrators must deploy a high density of desktops for VDI, they should confirm that the resources on the hypervisor are sufficient enough to provide all users with enough back-end resources. The following graph shows that during steady state, the hypervisor CPU was maxed out.

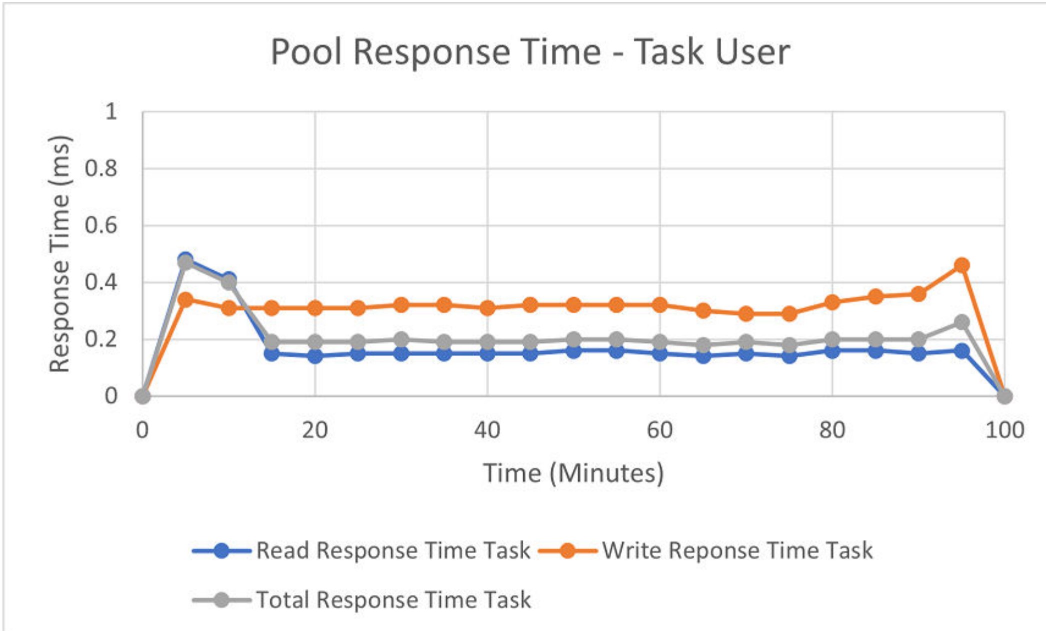


Storage performance

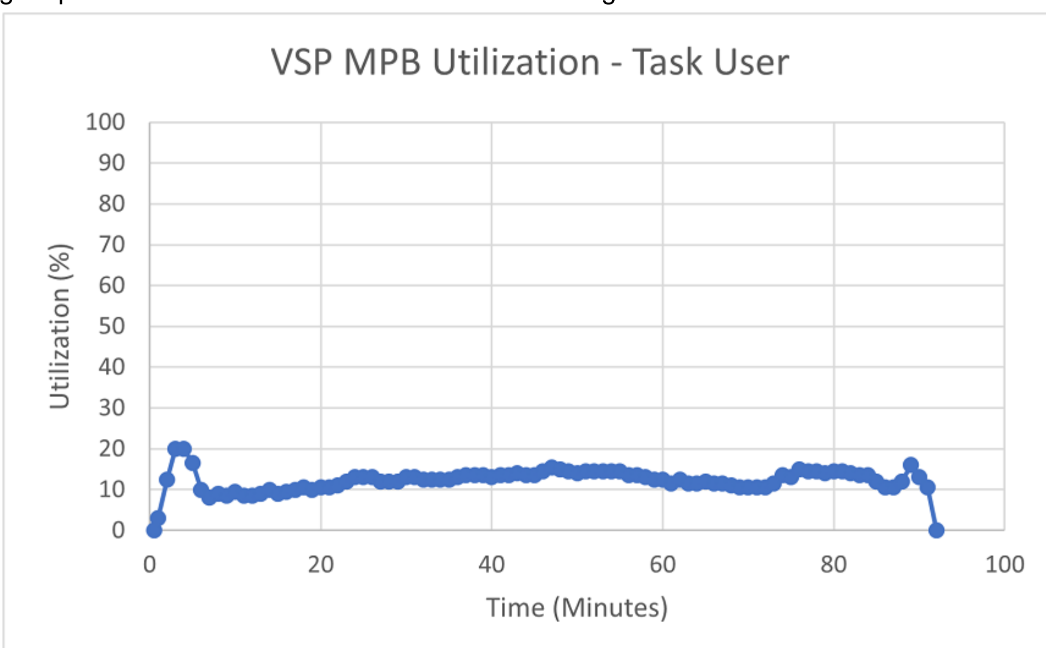
The following figures show IOPS, latency, and MPU usage at the VSP storage pool level during the log on, steady state, and log off phases during testing. During the log on phase of task user, 62000 IOPS was reached for the 400 VMs. During steady state and log off period, IOPS dipped to about 53000 IOPS.



Hitachi Virtual Storage Platform showed sub-millisecond response times during task user testing. During peak login period the VSP provided a max total response time of .47 milliseconds, while during steady state and log off period performance improved to .2 milliseconds.



Max MPB utilization was during peak log on period, which was 20%. During steady state and log off periods MPB utilization decreased to an average of 12%.



Knowledge user

The Knowledge Worker is designed for 2(v)CPU environments. This is a well-balanced intensive workload that stresses the system smoothly, resulting in higher CPU, RAM, and IO usage.

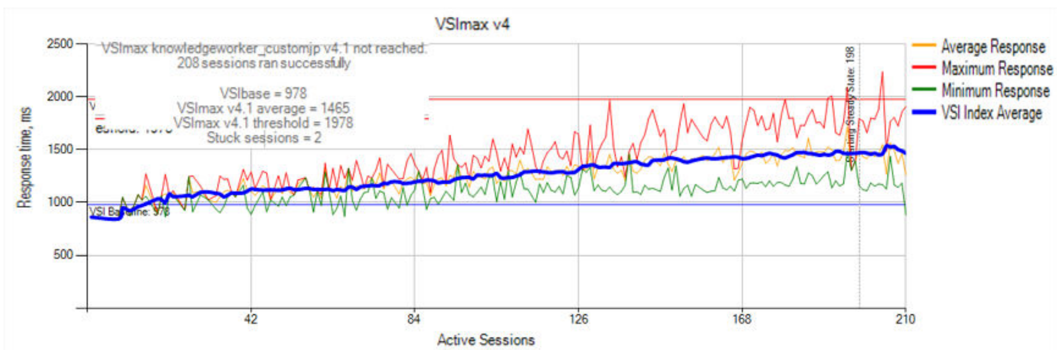
The Knowledge Worker workload uses the following applications:

- Adobe Reader
- Freemind/Java
- Internet Explorer
- Microsoft Excel
- Microsoft Outlook
- Microsoft PowerPoint
- Microsoft Word
- Photo Viewer

Results

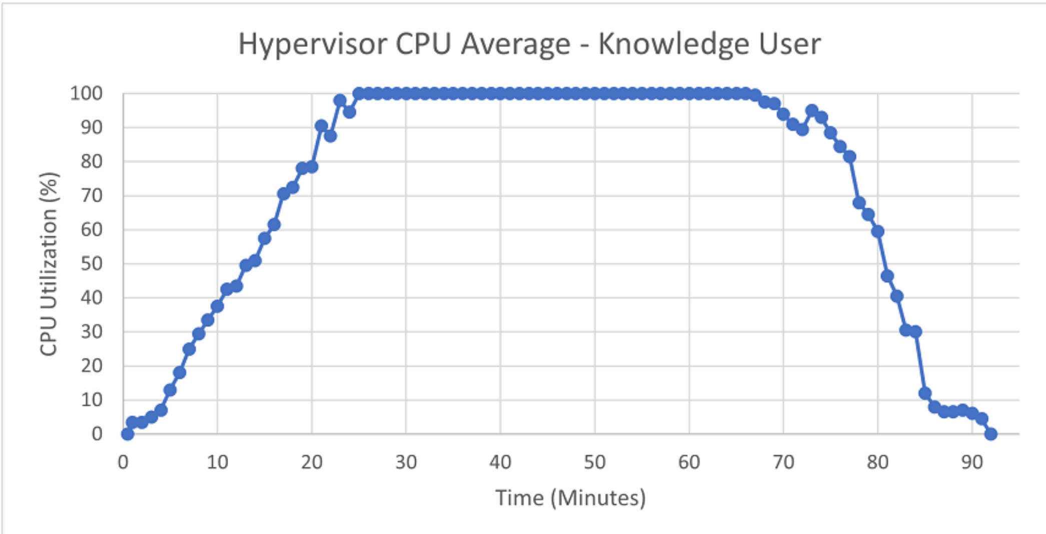
Login VSI

The Login VSI Max user experience score was not reached as shown in the following figure for the test with 210 knowledge users. This means that the system was not saturated because of resource constraints and received a login VSI base score of 978.



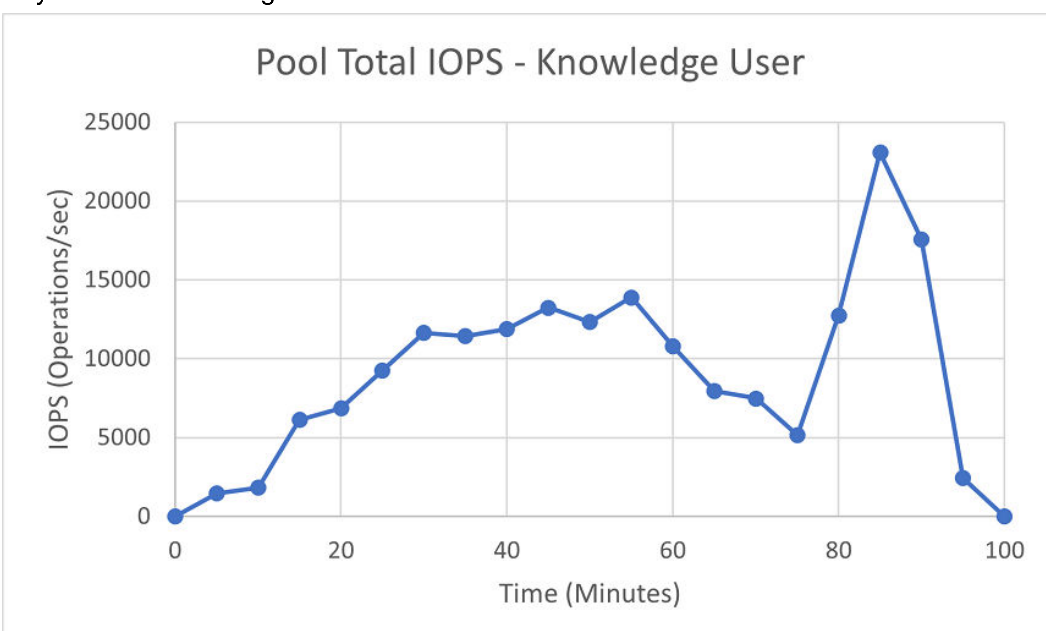
Hypervisor performance

During knowledge profile testing, a total of 210 VMs were deployed. CPU utilization of the hypervisors during this time showed 100% CPU utilization during steady state, but during log on and log off durations the hypervisor was not saturated in terms of compute resources. In the scenarios where admins are required to deploy a high density of desktops for VDI, it is recommended to confirm that the resources on the hypervisor are sufficient enough to provide all users with enough back-end resources. The following graph shows CPU utilization percentage of the hypervisor during log on, steady state, and log off.

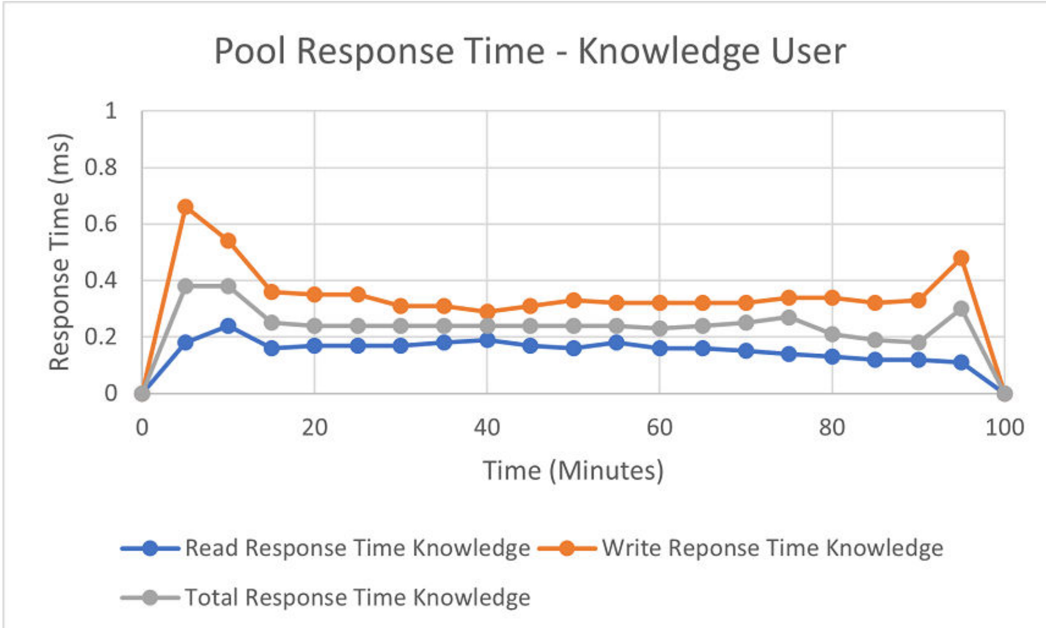


Storage performance

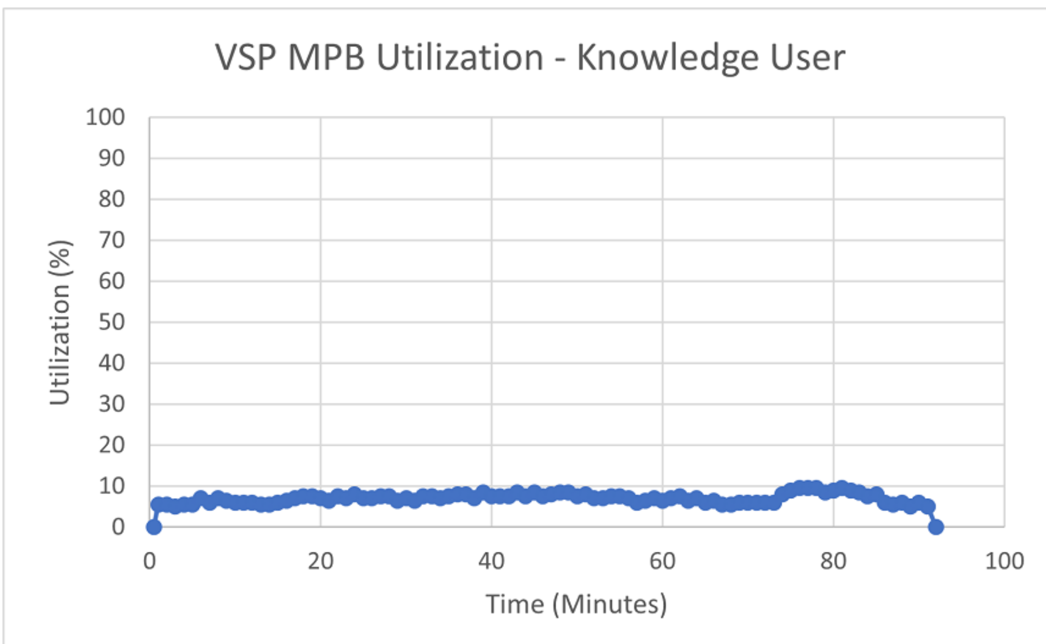
The following figures show IOPS, latency, and MPU usage at the VSP storage pool level during the log on, steady state, and log off phases during testing. During the log on phase of knowledge user testing, 1800 IOPS was reached for the 210 VMs. During steady state, the knowledge user was pushing 13000 IOPS and log off IOPS spiked to about 23000 IOPS briefly before normalizing.



Hitachi Virtual Storage Platform showed sub-millisecond response times during knowledge user testing. During peak log on period the VSP provided a max total response time of .38 milliseconds, while during steady state and log off period performance improved to .24 milliseconds.



MPB utilization was consistent during all phases of testing, which was an average of 7.5%.



Power user

The Power Worker is designed for 4(v)CPU+ environments. This is a very intensive workload that puts maximum stress on the system, resulting in very high CPU, RAM, and IO usage.

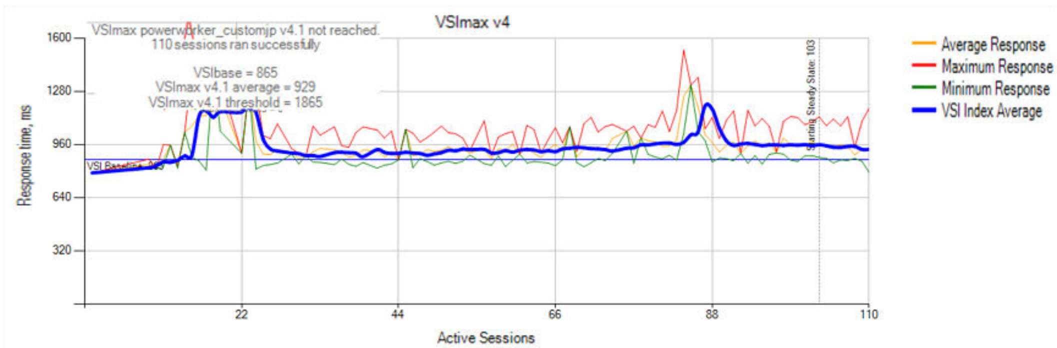
The Power Worker workload uses the following applications:

- Adobe Reader
- Freemind/Java
- Internet Explorer
- Microsoft Excel
- Microsoft Outlook
- Microsoft PowerPoint
- Microsoft Word
- Photo Viewer
- Simulated application install

Results

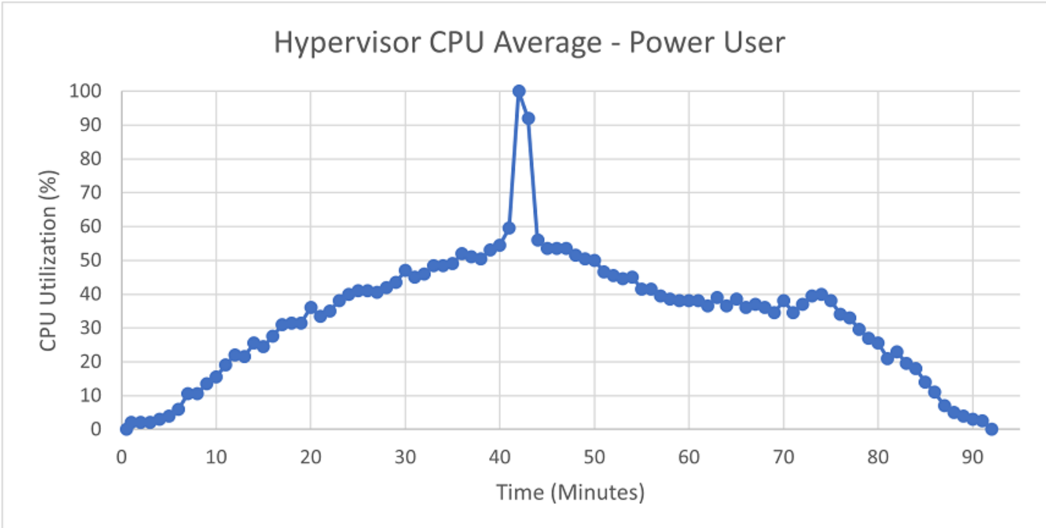
Login VSI

The Login VSI Max user experience score was not reached as shown in the following figure for the test with 110 power users. This means that the system was not saturated because of resource constraints and received a login VSI base score of 865.



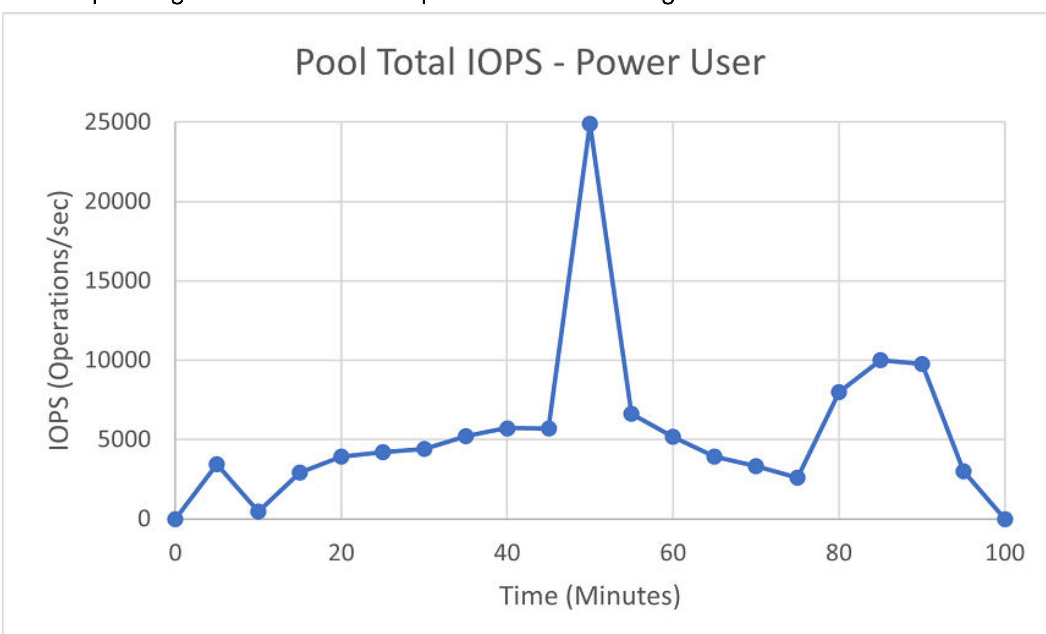
Hypervisor performance

During power profile testing, a total of 110 VMs were deployed. CPU utilization of the hypervisors during this time showed 50% CPU utilization during steady state with a brief spike to 100%, but during log on and log off durations the hypervisor was not saturated in terms of compute resources and was below 50%. The following graph shows CPU utilization percentage of the hypervisor during log on, steady state, and log off.

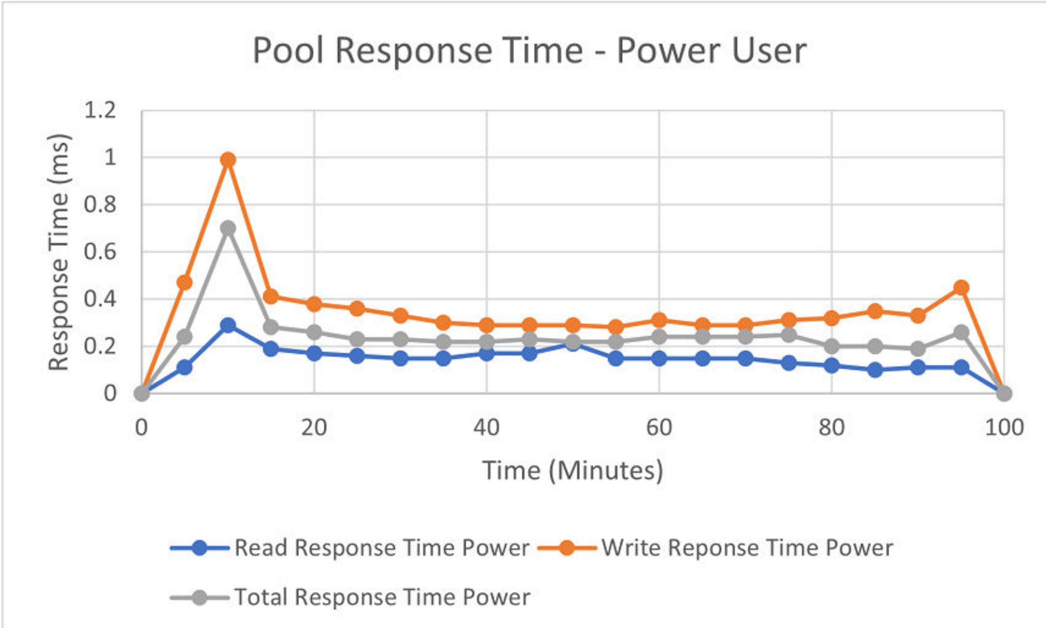


Storage performance

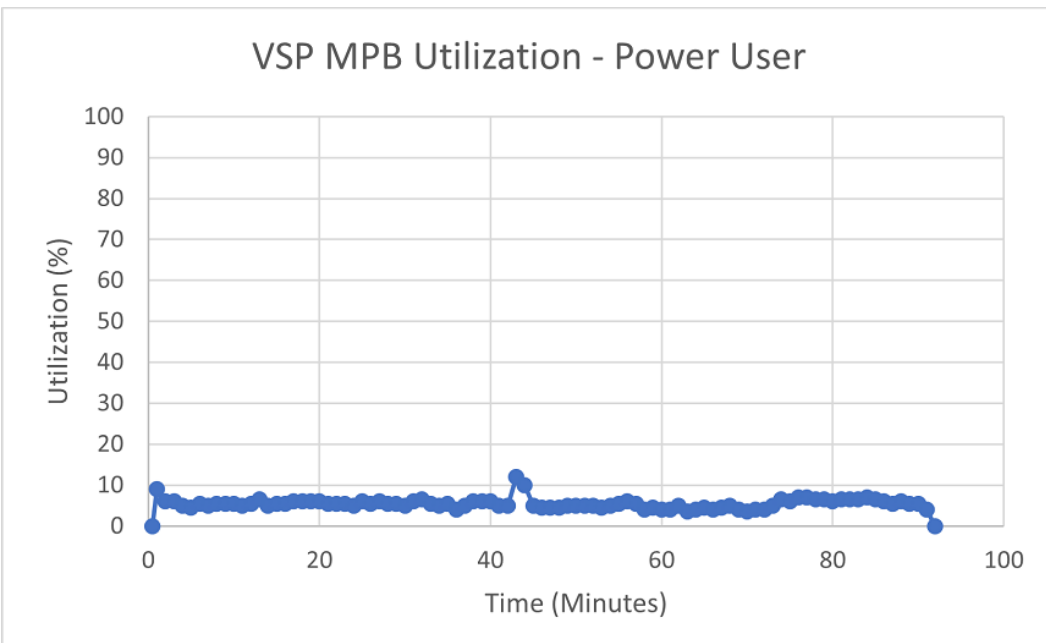
The following figures show IOPS, latency, and MPU usage at the VSP storage pool level during the log on, steady state, and log off phases of testing. During the log on phase of power user testing, 3000 IOPS was reached for the 110 VMs. During steady state, the power user was pushing 5000 IOPS with a spike to 25000 and log off was 10000 IOPS.



Hitachi Virtual Storage Platform showed sub-millisecond response times during power user testing. During peak log on period the VSP provided a max total response time of .7 milliseconds, while during steady state and log off period performance improved to .23 milliseconds.



MPB utilization was consistent during all phases of testing, which was an average of 6%.



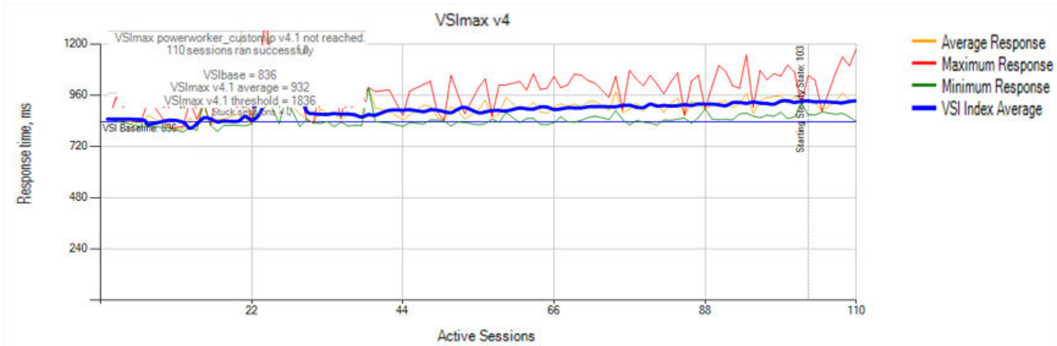
With GPU testing

The following results were run within the Hitachi lab, which validates the performance of power user profile with Nvidia A16 GPUs. Profile configuration for the power user can be viewed in the previous section.

Results

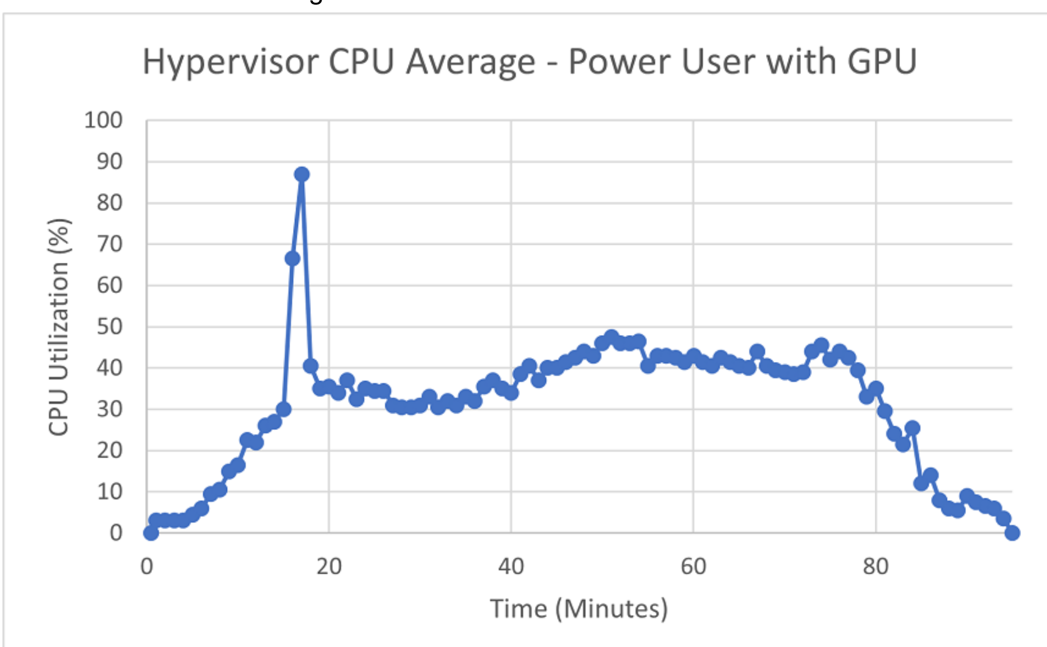
Login VSI

The Login VSI Max user experience score was not reached as shown in the following figure for the test with 110 power users with GPU. This means that the system was not saturated because of resource constraints and received a login VSI base score of 836. When comparing this score with the same profile used with no GPU, there was a slight improvement in user session experience by a factor of 29.

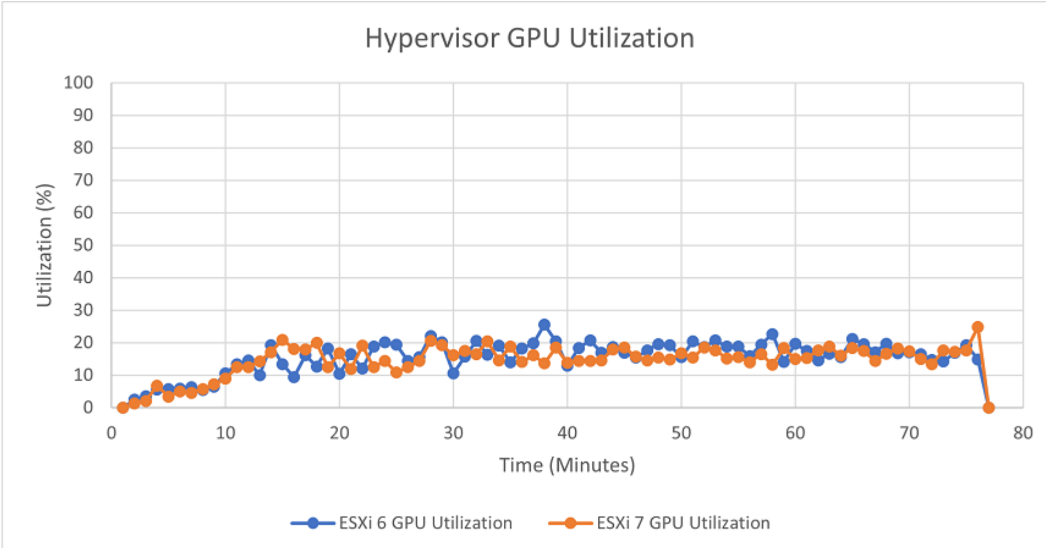


Hypervisor performance

During power profile testing with GPUs, a total of 110 VMs were deployed. CPU utilization of the hypervisors during this time showed 40% CPU utilization during steady state with a brief spike to 87% during log on. Aside from the minor spike during log on, the hypervisor was not saturated in terms of compute resources and was below 50% during the entire test run. The following graph shows CPU utilization percentage of the hypervisor during log on, steady state, and log off. When compared to power profile testing with no GPU, a 10% CPU cycle gain can be seen when using a GPU.

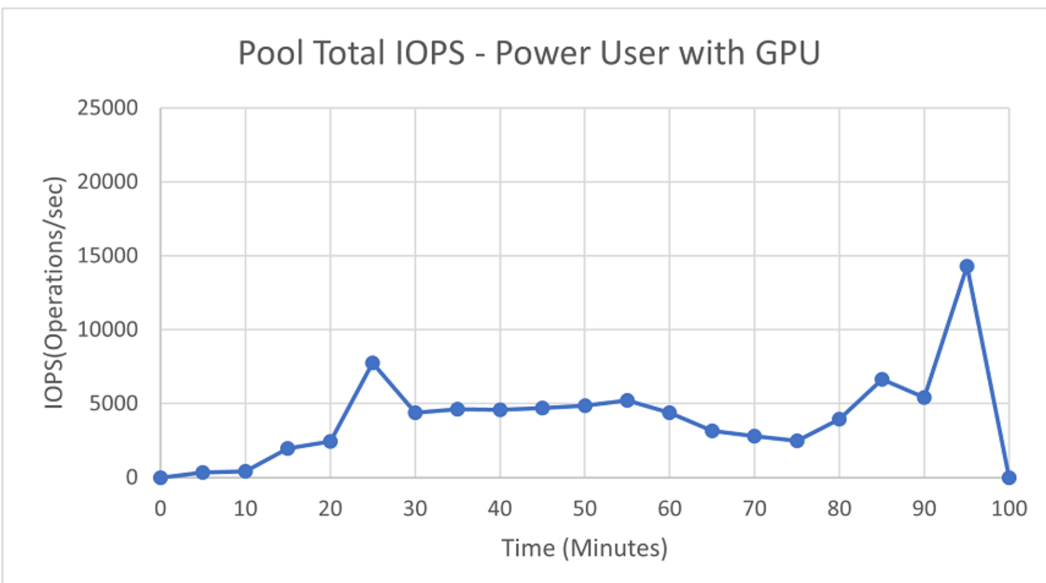


During power profile testing with GPUs, both hypervisors showed an average utilization of 18% as indicated in the following graph.

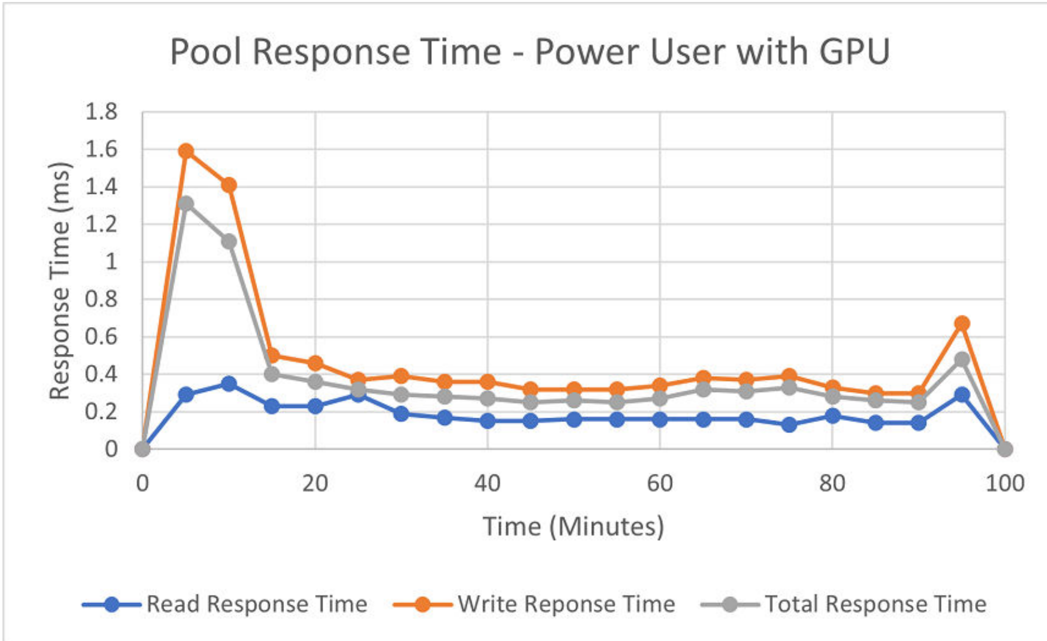


Storage performance

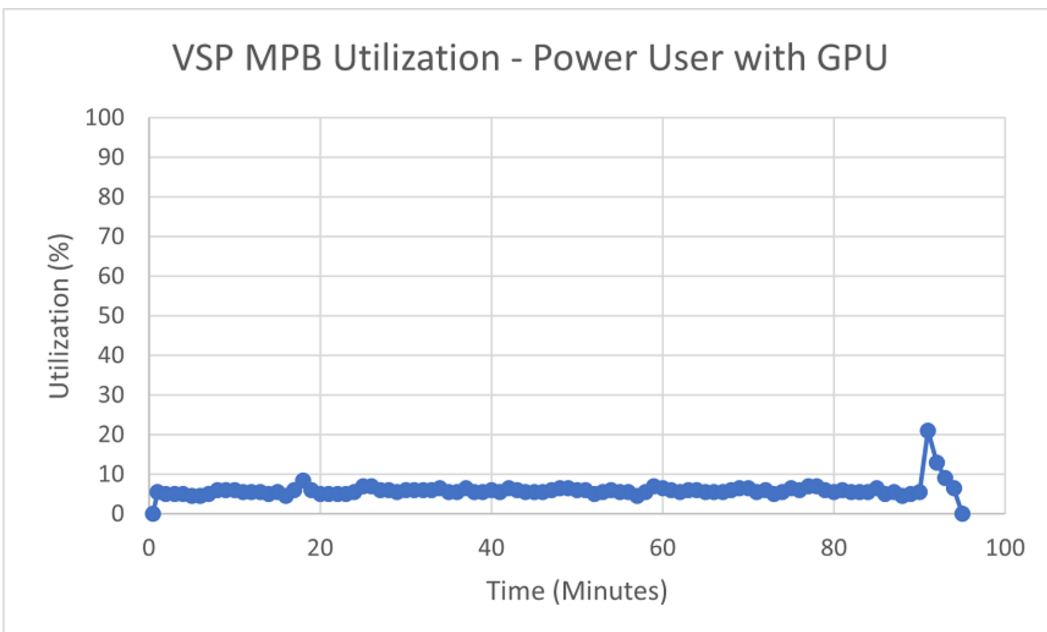
The following figures show IOPS, latency, and MPU usage at the VSP storage pool level during the log on, steady state, and log off phases of testing. During the log on phase of power user testing with GPU, 400 IOPS was reached for the 110 VMs. During steady state, the power user pushed about 5000 IOPS with a spike to 14000 IOPS during log off.



Hitachi Virtual Storage Platform showed sub-millisecond response times during power user testing. During peak login period the VSP provided a max total response time of 1.3 milliseconds, while during steady state and log off period performance improved to .30 milliseconds.

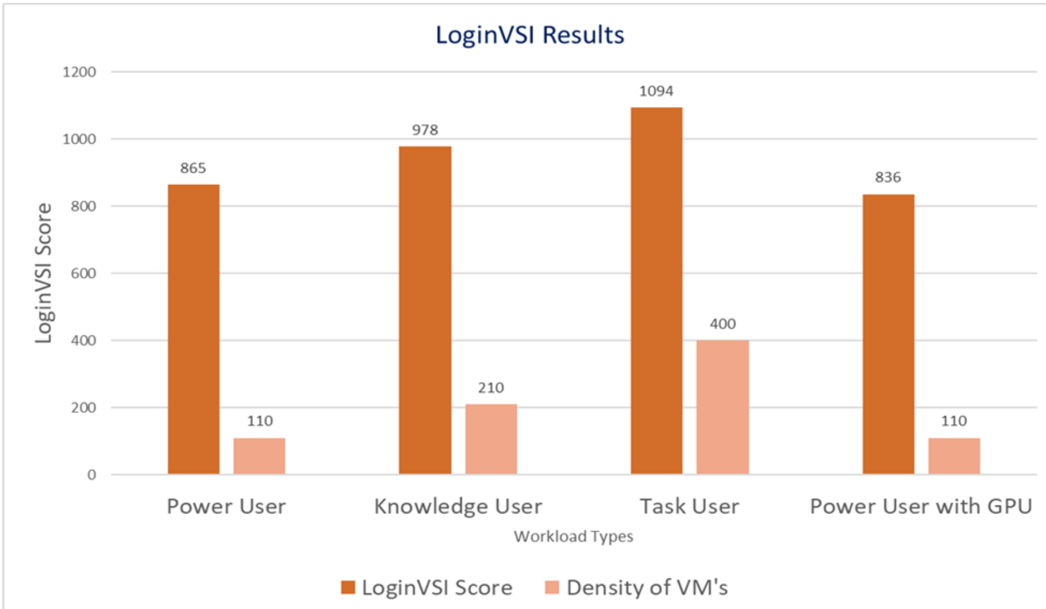


MPB utilization was consistent during all phases of testing, which was an average of 6%. The results show behavior similar to the power user profile tested with no GPU.



Login VSI scores

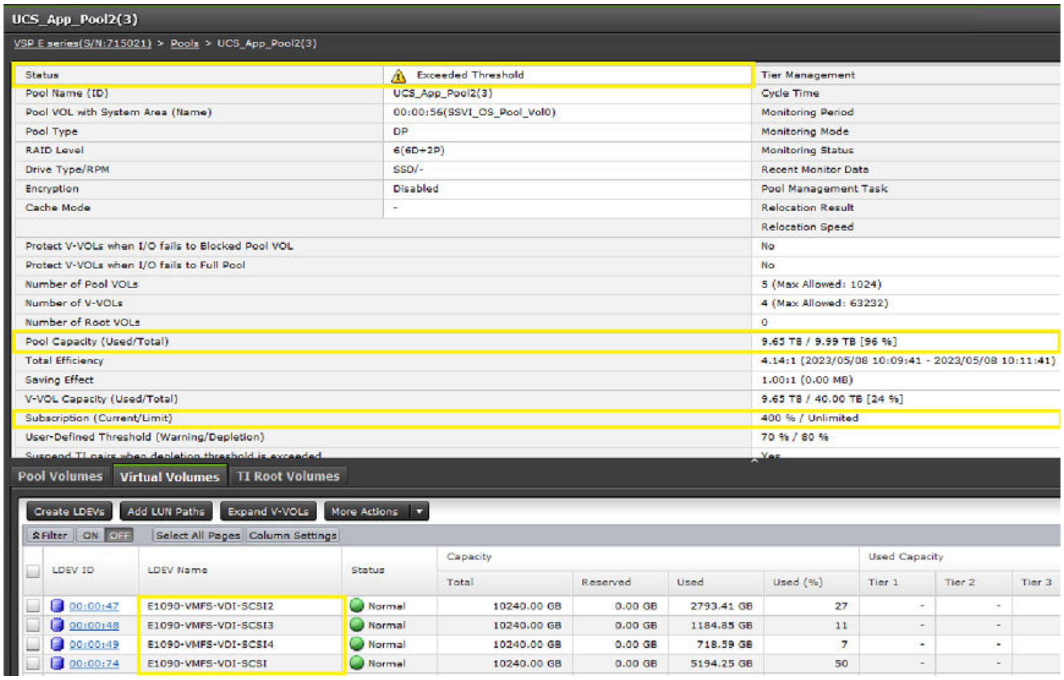
The following figure shows various user workloads that ran in the Hitachi Lab with the density of VMs and Login VSI score. All tests backed by Hitachi Virtual Storage Platform are in the good range. See the following KB article https://support.loginvsi.com/hc/en-us/articles/360001069519-Login-VSI-Login-VSI-VSImax-VSImax-VSItimer-threshold-NFO-NSLD-ZHC-ZLC-CPU- for additional information about Login VSI test results.



Capacity saving function

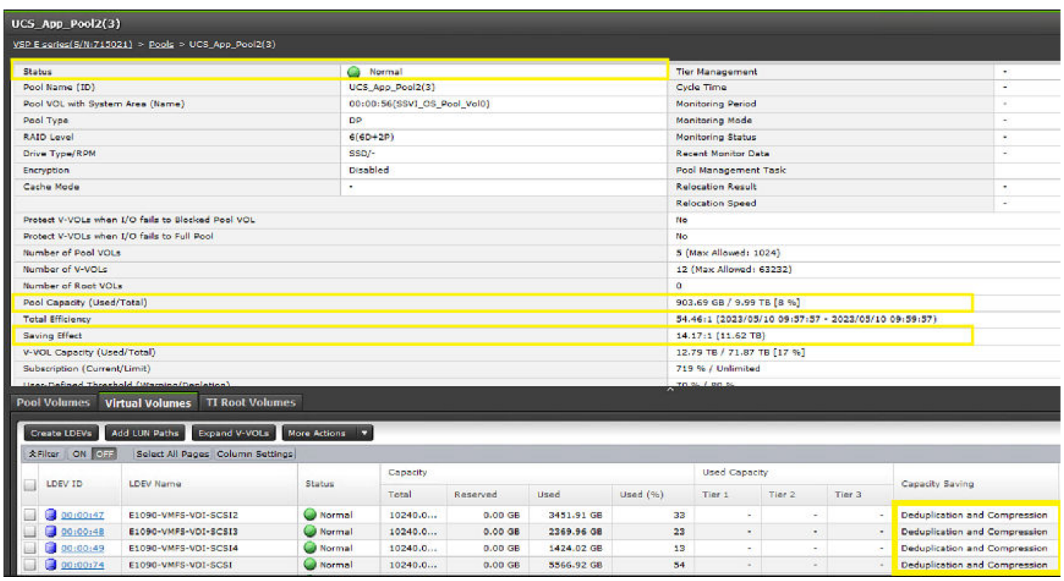
The capacity saving function was enabled on all LDEVs within the pool supporting the VDI deployment for both deduplication and compression. The pool supporting the VDI deployment had a physical capacity of 10 Tb from which four 10 Tb LDEVs in the form of VMFS datastores was provisioned to VMware vCenter. This means that from the pool perspective, storage was overallocated by 400%. A total of 400 VDI desktops were deployed using the task user with thin provisioned disk of 100 Gb during testing.

- Before



Prior to enabling the capacity saving function for both deduplication and compression on the Hitachi Virtual Storage Platform, the VSP pool exceeded capacity thresholds and had a Used/Total capacity of 9.65 Tb of 9.99 Tb.

- After



After capacity savings was enabled, it took 24-72 hours to obtain results because time was required for the algorithm to scan blocks and deduplicate and compress data as needed. With these settings enabled, pool capacity was reduced from 96% to 8% with the effective used capacity respectively dropping from 9.65 Tb to 903.69 Gb. This provides a savings effect of 14.17:1 ratio, indicating that for VDI deployments that share common data sets at the OS and file system level Hitachi Virtual Storage Platform can drastically save back-end capacity.

Conclusion

The Cisco and Hitachi adaptive solution for converged infrastructure shows that it is a great candidate for VDI workloads. Throughout all testing, Hitachi Virtual Storage Platform performance for total response time was less than 1 ms with profiles that included task, knowledge, and power users. Additionally, VSP MPB utilization was low during all workloads, which indicates that the VSP can manage much more load than the compute resources allowed within the testing done in the Hitachi lab.

During VDI sizing and deployment it is critical that administrators take compute resources into consideration, as testing has shown. Even though a lower profile such as task may be used, performance is directly impacted by the quantity of user sessions deployed. For users looking to drive additional high graphic workloads, a GPU will assist in driving better performance as well as reducing the load on the hypervisor CPU. Within the Hitachi lab and the resources deployed, the average CPU savings for the hypervisor was 10%. Additionally, for VDI deployments, the VSP capacity saving function drastically saves capacity and provides end users the ability to use the Hitachi HDP pool overprovisioning feature much more effectively. This can reduce the need to purchase additional storage, which translates into reduced operational expenditures.

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA

HitachiVantara.com/contact