Reference Architecture

VMware Horizon VDI Reference Architecture for Dell EMC PowerFlex family

Abstract

This Reference Architecture demonstrates the integration of VMware Horizon with Dell EMC PowerFlex single-layer (HCI) deployment for VDI solutions.

April 2021

Revisions

Date	Description
August 2020	Initial release
November 2020	Added new content related to Multimedia Worker ─ Test results and analysis
April 2021	• Updated Section 4.1 VDI test tools – Login VSI • Added Section 4.5 About Login VSI • Added Login VSI logo

Acknowledgements

Author: Mohanraj Gnanamani

Contributors: Colin Byrne, Kevin M Jones, and Nicholas Busick

Table of contents

Executive summary

The Dell EMC PowerFlex family provides flexible consumption and deployment models for different types of workloads, the HCI deployment model scales compute, storage, and networking as one, enabling VDI environments to start with a lower investment without compromising on performance.

The single-layer (HCI) deployment model enables customers to size their environments without overprovisioning hardware for their VDI environment, providing the flexibility to scale horizontally and vertically.

This document provides the reference architecture for VMware Horizon software with Dell EMC PowerFlex family to create the virtual desktop infrastructure environments. It includes the Login VSI test environment configuration and best practices for systems that have undergone testing.

1 Introduction

1.1 Objectives

This reference architecture provides an overview of the benefits when integrating the PowerFlex family with VMware Horizon based VDI solutions. It provides technical and networking design with the hardware specifications. Validated Login VSI test results, run with a selection of workloads, are provided to assist customers in sizing their VDI environment.

1.2 Audience

The audience of this document includes decision makers, managers, architects, and technical administrators of IT environments and anyone else who are interested in gaining in-depth understanding of VMware Horizon VDI solutions with PowerFlex family.

Readers are expected to have an understanding and working knowledge of virtualization, virtual desktop infrastructure, VMware Horizon, and VMware vSphere technologies.

1.3 Terminology

The following table defines acronyms and terms that are used throughout this document:

Table 1 Terms and definitions

Term	Definition
VDI	Virtual Desktop Infrastructure
HCI	Hyperconverged Infrastructure
MDM	Meta Data Manager
SDS	Storage Data Server
SDC	Storage Data Client
SVM	Storage Virtual Machine
OS	Operating System
SSD	Solid-state Drive
SAN	Storage Area Network
GPU	Graphics Processing Unit
VM	Virtual Machine
ITOM	IT Operations Management
PAAC	Performance Analysis and Characterization
EUE	End-user Experience
VIB	vSphere Installation Bundle

2 Product overview

2.1 PowerFlex family

PowerFlex is a software-defined storage platform designed to significantly reduce operational and infrastructure complexity, empowering organizations to move faster by delivering flexibility, elasticity, and simplicity with predictable performance and resiliency at scale. The PowerFlex family provides a foundation that combines compute as well as high performance storage resources in a managed unified fabric. PowerFlex comes in flexible deployment options - rack, appliance, or ready nodes-that enables disaggregated (two-layer), HCI (single-layer), or mixed architectures. PowerFlex is ideal for high performance applications and databases, building an agile private cloud, or consolidating resources in heterogeneous environments.

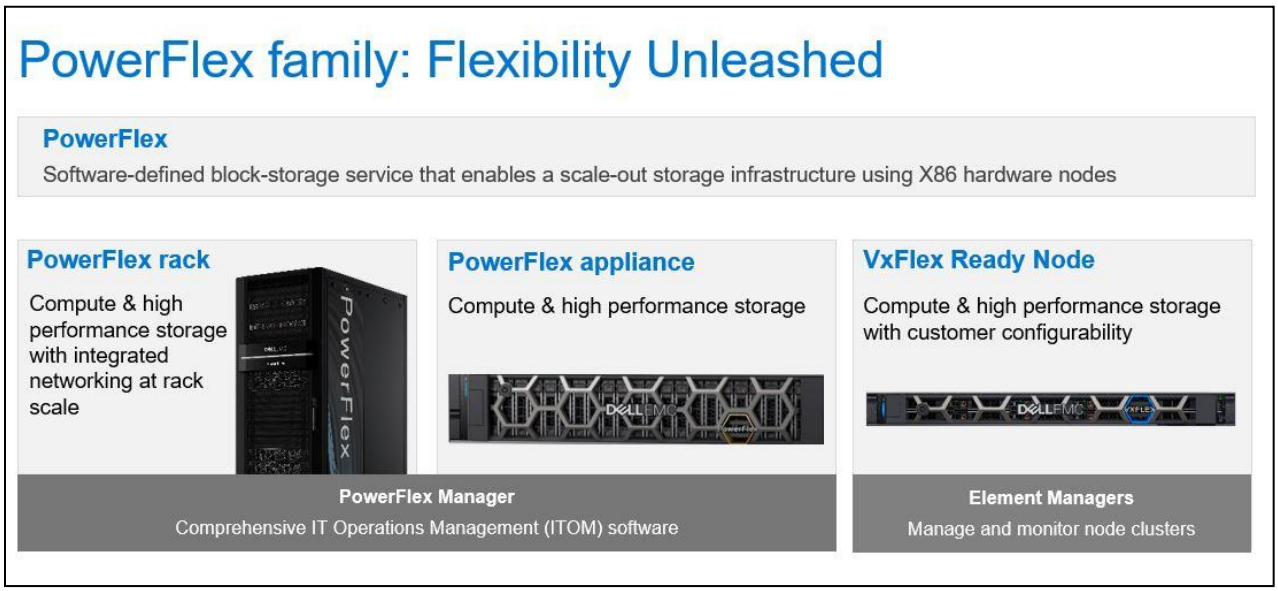


Figure 1 PowerFlex family

2.2 PowerFlex software components

Software is the key differentiation and the “secret sauce” in the PowerFlex offering. PowerFlex software components not only provide software-defined storage services, they also help simplify infrastructure management and orchestration with comprehensive ITOM and LCM capabilities that span compute as well as storage infrastructure, from BIOS and Firmware to nodes, software and networking.

The core foundational component in the PowerFlex family that enables Software Defined Storage (SDS) services is called simply PowerFlex, to represent the core value it enables for the platform. Additionally, PowerFlex Manager is a comprehensive IT Operational Management (ITOM) and Life Cycle Management (LCM) tool that drastically simplifies management and ongoing operation.

2.2.1 PowerFlex

PowerFlex **(previously VxFlex OS)** is the software foundation of PowerFlex software-defined storage. It is a scale-out block storage service designed to deliver flexibility, elasticity, and simplicity with predictable high performance and resiliency at scale.

PowerFlex consists of the following components:

Storage Data Client (SDC)

- Provides front-end volume access to applications and file system
- Installed on servers consuming storage
- Maintains peer-to-peer connections to every SDS managing a pool of storage

Storage Data Server (SDS)

- Abstracts local storage, maintains storage pools, and presents volumes to the SDCs
- Installed on servers contributing local storage to the PowerFlex cluster

Meta Data Manager (MDM)

- Oversees storage cluster configurations, monitoring, rebalances, and rebuilds
- Highly available, independent cluster installed on three or five different nodes
- May reside alongside SDCs or SDSs, or on separate nodes
- Sits outside the data path

Gateway

- Performs installation and configuration checks
- Acts as an endpoint for API calls and passes them to MDM

2.2.2 PowerFlex Manager

PowerFlex Manager is the software component in PowerFlex family that enables ITOM automation and life cycle management capabilities for PowerFlex systems.

2.3 PowerFlex consumption options

PowerFlex SDS platform is available in multiple consumption options to help customers meet their project and data center requirements. PowerFlex appliance and PowerFlex rack provide customers comprehensive IT Operations Management (ITOM) and life cycle management (LCM) of the entire infrastructure stack in addition to sophisticated high-performance, scalable, resilient storage services. PowerFlex appliance and PowerFlex rack are the two preferred and proactively marketed consumption options. PowerFlex is also available on VxFlex Ready Nodes without the ITOM and LCM capabilities.

Note: The brand for Ready Nodes continues to be VxFlex.

2.3.1 PowerFlex rack

PowerFlex rack is a software-defined storage platform designed to deliver flexibility, elasticity, and simplicity with predictable performance and resiliency at scale by combining compute as well as high performance storage resources in a managed unified network. This rack-scale engineered system, with integrated networking, enables customers to achieve the scalability and management requirements of a modern data center.

2.3.2 PowerFlex appliance

PowerFlex appliance is a software-defined storage platform designed to deliver flexibility, elasticity, and simplicity with predictable performance and resiliency at scale by combining compute as well as high performance storage resources in a managed unified network. This turnkey offer allows customers the flexibility and savings to bring their own compatible networking. With PowerFlex, customers deploy to match their initial needs and easily expand with massive scale potential, without having to compromise on performance and resiliency.

2.3.3 VxFlex Ready Nodes

VxFlex Ready Nodes are validated server building blocks configured for use with PowerFlex. They are available with thousands of configuration options and are available for customers who prefer to build their own environments.

2.4 PowerFlex deployment architectures

PowerFlex software-defined storage offers flexibility of deployment architecture to help best meet the specific deployment and architectural requirements. PowerFlex can be deployed in a two-layer (Server SAN), single-layer (HCI), or in storage-only architectures.

2.4.1 Two-layer architecture

In a two-layer architecture, nodes that provide storage capacity and host datasets are separated from nodes that host applications and workloads. PowerFlex manager provides LCM and IOTM for the entire infrastructure, including nodes that provide storage and nodes that host the applications. Compute and storage resources can be scaled by adding respective node to the cluster. This segregation of compute and storage resources can be helpful to minimize software licensing costs in certain situations. This architecture could be most suitable for hosting high-performance high-value databases and application workloads.

2.4.2 Single-layer (HCI) architecture

In this architecture, each node in the cluster contributes storage resources as well as hosts applications and workloads. This architecture allows you to scale your infrastructure uniformly and with a pre-defined building block that adds both storage and compute resources. PowerFlex Manager provides ITOM and LCM capabilities for the entire infrastructure. This architecture is most suitable for data center and workload consolidation.

2.4.3 Mixed architecture

Using PowerFlex storage-only nodes, a software-defined block storage environment is created that can be accessed and consumed by a number of applications and workloads that are hosted outside of the PowerFlex cluster. PowerFlex Manager provides LCM and ITOM for the storage infrastructure. This is a suitable architecture where the customer has existing compute infrastructure but needs high-performance SDS. This can be a starting point with the customer and may expand to a two-layer Server SAN deployment in the future as the external compute is migrated to PowerFlex.

2.5 VMware vSphere 6.7

VMware vSphere provides a powerful, flexible, and secure foundation for business agility that accelerates the digital transformation to cloud computing and promotes success in the digital economy.

vSphere provides the following benefits for VDI applications:

- **Improved appliance management** – The vCenter Server Appliance management interface provides CPU and memory statistics, network and database statistics, disk space usage, and health data. These features reduce reliance on a command-line interface for simple monitoring and operational tasks.
- **VMware vCenter Server native high availability** – This solution for vCenter Server Appliance consists of active, passive, and witness nodes that are cloned from the existing vCenter Server instance. The vCenter HA cluster can be enabled, disabled, or destroyed at any time. Maintenance mode prevents planned maintenance from causing an unwanted failover. The vCenter Server database uses Native PostgreSQL synchronous replication, while key data outside the database uses a separate asynchronous file system replication.
- **Backup and restore** – Native backup and restore for the vCenter Server Appliance enables users to back up vCenter Server and Platform Services Controller appliances directly from the VAMI or API. The backup consists of a set of files that is streamed to a selected storage device using SCP, HTTP(S), or FTP(S) protocols. This backup fully supports VCSA instances with both embedded and external Platform Services Controller instances.
- **VMware vSphere HA Support for NVIDIA vGPU-configured VMs** – vSphere HA protects VMs with the NVIDIA vGPU shared pass-through graphics acceleration. In the event of a failure, vSphere HA tries to restart the VMs on another host that has an identical NVIDIA vGPU profile. If no available healthy host meets this criterion, the VM fails to power on.
- **VMware Log Insight** – Provides log management, actionable dashboards, and refined analytics, which enable deep operational visibility and faster troubleshooting.

For more information, see VMware vSphere resources page.

2.6 VMware Horizon 7

The architecture that this guide describes is based on VMware Horizon 7, which provides a complete end-to-end solution delivering Microsoft Windows and Linux virtual desktops to users on a wide variety of endpoint devices. Virtual desktops can be dynamically assembled on demand, providing users with pristine, yet personalized, desktops each time they log in.

VMware Horizon 7 provides a complete virtual desktop delivery system by integrating several distributed components with advanced configuration tools that simplify the creation and real-time management of the virtual desktop infrastructure.

The core Horizon 7 components include:

- **Horizon Connection Server (HCS)** – Installed on servers in the data center. The Horizon Connection Server brokers client connections, authenticates users, entitles users by mapping them to desktops or pools, establishes secure connections from clients to desktops, supports single sign-on, and sets and applies policies.
- **Horizon Administrator** – Provides administrator functions such as deployment and management of Horizon desktops and pools, setting and controlling user authentication, and more.

- **Horizon Agent** – Installed on all VMs, physical machines, and Terminal Service servers that are used as a source for Horizon desktops. On VMs, the agent is used to communicate with the Horizon client to provide services such as USB redirection, printer support, and more.
- **Horizon Client** – Installed on endpoints for creating connections to Horizon desktops that can be run from tablets, Windows, Linux, or Mac PCs or laptops, thin clients, and other devices.
- **Unified Access Gateway** – Provides a way to securely deliver connections that require a higher level of security to access, such as remote connections from the Internet.
- **Horizon Portal** – Provides access to links for downloading full Horizon clients. Enable the HTML access feature to run a Horizon desktop inside a supported browser.
- **vCenter Server** – Provides centralized management and configuration to the entire virtual desktop and host infrastructure. It facilitates configuration, provisioning, and management services.

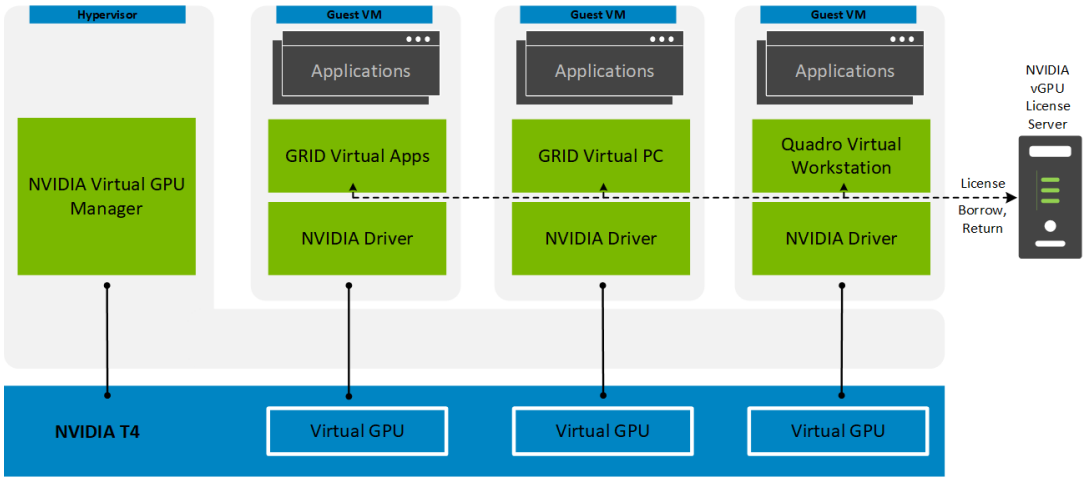
2.6.1 Horizon 7 clone technology

- **Full clones** – Full clones are typically used for testing purposes or to create management VMs. Full clones are not ideal for VDI because full copies have no connection to the original VM. Updates must be performed on each VM with this approach.
- **Instant clones** – This technology provisions a VM the instant a user requests one. The result is a far easier approach to operating system updates and patch management because the VM is created when it is needed. You can use the combination of Just-in-Time Management Platform (JMP) features such as App Volumes and User Environment Manager to emulate persistence.

For more information, see VMware Horizon resource page.

2.7 NVIDIA vGPU

NVIDIA vGPU brings the full benefit of NVIDIA hardware-accelerated graphics to virtualized solutions. This technology provides exceptional graphics performance for virtual desktops equivalent to local PCs when sharing a GPU among multiple users. It also enables aggregation of multiple GPUs assigned to a single VM to power the most demanding workloads.



NVIDIA vGPU is the industry's most advanced technology for virtualizing true GPU hardware acceleration to share them between multiple virtual desktops, or aggregate them to assign to a single virtual desktop, without compromising the graphics experience. NVIDIA vGPU offers three software variants to enable graphics for different virtualization techniques:

- **GRID Virtual PC (GRID vPC)** – Designed to provide full virtual desktops with up to dual 4K monitor support or single 5K monitor support.

- **GRID Virtual Applications (GRID vApps)** – Designed to deliver graphics accelerated applications using RDSH.

- **Quadro Virtual DataCenter Workstation (Quadro vDWS)** – Designed to provide workstation-grade performance in a virtual environment with support for up to four quad 4K or 5K monitors or up to two 8K monitors.

- **NVIDIA Virtual Compute Server (vCS)** – Designed to accelerate server virtualization so that the most compute-intensive workloads, such as artificial intelligence, deep learning, and data science, can be run in a VM.

Dell EMC Ready Solutions for VDI can be configured with the following NVIDIA GPUs:

- **NVIDIA M10 (Maxwell)** – Recommended for GRID vApps or GRID vPC environments, each card is equipped with 32 GB of frame buffer with the maximum available frame buffer per user at 8 GB. Dell Technologies recommends hosting a maximum of 32 Windows 10 users per card. While some PowerFlex hyperconverged node configurations support three cards, consider sizing with a maximum of two cards per node. Configure systems with less than 1 TB of memory when using the M10.

- **NVIDIA T4 Tensor Core** – NVIDIA's Turing architecture is available in the T4 GPU, which is considered the universal GPU for data center workflows. The T4 GPU is flexible enough to run knowledge worker VDI or professional graphics workloads. Add up to six GPU cards into your R740xd node to enable up 96 GB of graphics frame buffer. For modernized data centers, use this card in off-peak hours to perform your inferencing workloads with NVIDIA vCS software.

- **NVIDIA Quadro RTX 6000 and 8000** – Select the Turing-based RTX 6000 or RTX 8000 for the best virtualized graphics performance for professional graphics and rendering workloads. These two GPUs have similar specifications, but the RTX 6000 has 24 GB of graphics frame buffer while the RTX 8000 has 48 GB per card. Add up to three RTX 6000 or 8000 GPU cards into your R740xd node to enable up to 72 GB or 144 GB respectively.

3 Solution architecture

This section provides an architecture overview and guidance on managing and scaling a VMware Horizon 7 environment on PowerFlex single-layer (HCI) architecture.

3.1 Logical architecture

The following figure illustrates the logical view of the VDI management infrastructure and virtual desktop VMs deployed in the cluster using three PowerFlex hyperconverged nodes, all the nodes are configured in the cluster shares both the compute and storage.

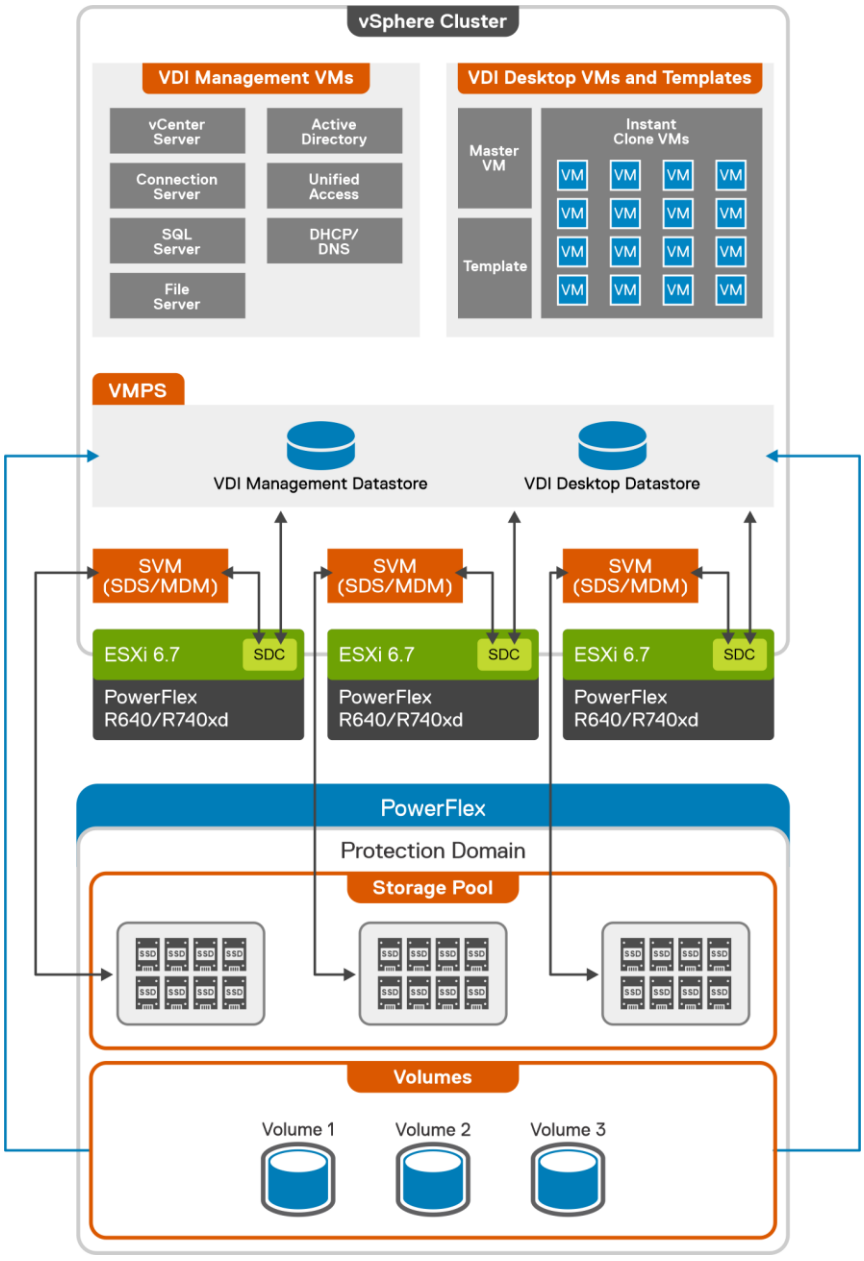


Figure 2 PowerFlex HCI cluster with Horizon

The vSphere ESXi hypervisor is installed on all nodes in the cluster, the SVM is deployed on the hypervisor to provide SDS and MDM services that make up the PowerFlex storage cluster. A single protection domain is created, and a storage pool is carved out of the SSDs available in the nodes. The SDC component is the PowerFlex client software VIB installed in the ESXi to access the volumes that are created from the storage pool.

VDI management and desktop datastores are PowerFlex volumes that are formatted with VMFS to store all the virtual machines deployed in the cluster.

Note: Dell EMC recommends deploying all the VDI management infrastructure on dedicated nodes from the virtual desktop compute nodes within the same or separate PowerFlex cluster.

3.2 Network architecture

The following figure shows the physical and virtual networking configured to the PowerFlex hyperconverged nodes with 4 x 25 GbE network in an active/active teaming mode.

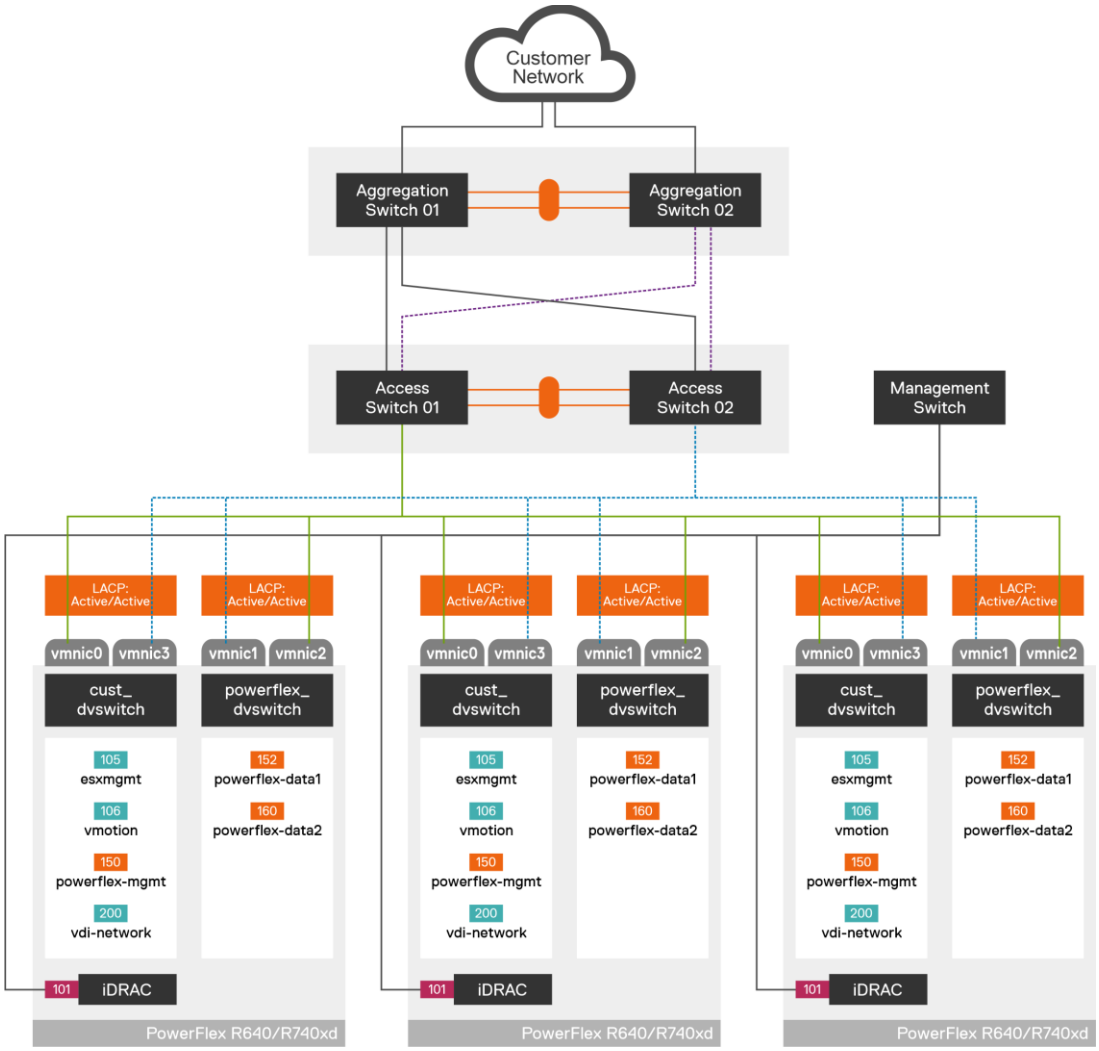


Figure 3 PowerFlex HCI physical and virtual networking

The following table lists the network VLAN used for this deployment:

Table 2 Port group and VLAN details

Port group	VLAN	Purpose
esxmgmt	105	ESXi management traffic
vmotion	106	Virtual machine migration traffic
powerflex-mgmt	150	PowerFlex management traffic
powerflex-data1	152	PowerFlex SDS and SDC traffic
powerflex-data2	160	PowerFlex SDS and SDC traffic
vdi-network	200	VDI traffic

3.3 Hardware recommendation

PowerFlex uses industry-leading Dell EMC PowerEdge servers that are rigorously tested and integrated into the PowerFlex system with a broad range of configuration options based on the requirements.

The following hyperconverged nodes are recommended for this solution:

- **PowerFlex R640** – a 1U compact and powerful node, ideal for regular (non-graphics intensive) virtual desktops, occupying less space in the data center resulting in more density per rack.
- **PowerFlex R740xd** – a 2U node that supports NVIDIA GPU hardware, the NVIDIA virtual GPU software enables a physical GPU to be shared across multiple virtual desktops. Recommended for graphics-intensive VDI use cases.

For more information about supported servers and switches, see PowerFlex specification sheet.

3.4 PowerFlex storage provisioning

The PowerFlex cluster allows you to create two types of Storage Pools:

- **Medium Granularity (MG)** –This storage pool is recommended for VDI. MG storage pool volumes provide the supreme I/O performance with the least latency to the virtual machines and applications.
- **Fine Granularity (FG)** –This storage pool is recommended where space efficiency is more valuable than performance, it enables the compression feature to reduce the size of the volume data depends on the data compressible and allows it to use by the other volumes.

The following tables compare both the features:

Table 3 MG vs FG comparison

Feature	MG storage pool	FG storage pool
Volume provision type	Thick and thin	Thin
Space allocation block	1 MB	4 KB
Compression	False	True
Supported drives	All-flash	All-flash
NVDIMM required	False	True

4 Test and performance analysis

This section provides a detailed summary and description of the tests performed to validate the optimal number of VMs density per host for the following test profiles.

- Testprofile1 - Task worker
- Testprofile2 - Knowledge worker
- Testprofile3 - Power worker
- Testprofile4 - Multimedia worker

For desktop VM configuration used for different test profiles, see Configuration details.

4.1 VDI test tools – Login VSI

Login VSI is the industry standard load-testing tool for measuring the performance and scalability of centralized Windows desktop environments, such as server-based computing (SBC) and VDI. Login VSI is used for testing and benchmarking by all the major hardware and software vendors and is recommended by both leading IT analysts and the technical community. Login VSI is 100% vendor independent and works with standardized user workloads and statistics—making all conclusions based on Login VSI test data objective, verifiable, and repeatable.

Login VSI-based test results are used and published in multiple technical white papers and presented at various IT-related conferences by our vendor customers. The product Login VSI is also widely used by end-user organizations, system integrators, hosting providers, and testing companies. It is also the standard testing tool used in all tests executed in the internationally acclaimed research project VDI Like a Pro (formerly known as Project Virtual Reality Check).

4.2 VDI test methodology

To ensure the optimal combination of EUE and cost-per-user, PAAC on Dell EMC VDI solutions is carried out using a carefully designed, holistic methodology that monitors both hardware resource utilization parameters and EUE during load-testing.

Login VSI – the launchers and Login VSI environment are configured and managed by a centralized management console. Additionally, the following login and boot paradigm is used:

- Data collection interval for non vSAN datastore is 1 minute while for vSAN metrics data collection interval is 5 minutes.
- User logon and workload are two separate phases which are staggered to start every 5 seconds.
- All desktops are pre-booted in advance of logins commencing.
- Data collection is a combination of the automated framework and manual scripts.

4.3 Profiles and workloads

It is important to understand user workloads and profiles when designing a desktop virtualization solution to understand the density numbers that the solution can support. At Dell EMC, we use five Login VSI workload/profile levels workload/profile, each of which is bound by specific metrics and capabilities with two targeted at graphics-intensive use cases. We will present more detailed information in relation to these workloads and profiles below, but first it is useful to define the terms “profile” and “workload” as they are used in this document.

- **Profile:** Is the configuration of the virtual desktop - number of vCPUs and amount of RAM configured on the desktop and available to the user.
- **Workload:** Is the set of applications used by performance analysis and characterization (PAAC) of Dell EMC VDI solutions (for example Microsoft Office applications, PDF Reader, Internet Explorer, and so on.)

Load-testing on each profile is carried out using an appropriate workload that is representative of the relevant use case and summarized in the table below:

Table 4 Login VSI

Workload Name	Application Set
Task Worker	2-7 Apps
Knowledge Worker	5-9 Apps and 360p Video
Power Worker	8-12 Apps and 720p Video
Multimedia Worker	Google Earth, Bouncing 3D Objects and HD Video

4.3.1 Resource monitoring

The following sections explain respective component monitoring used across all Dell EMC solutions where applicable.

- VMware vCenter is used for VMware vSphere-based solutions to gather key data (CPU, Memory, Disk, and Network usage) from each of the compute hosts during each test run. This data is exported to .csv files for single hosts and then consolidated to show data from all hosts (when multiple is tested). While the report does not include specific performance metrics for the Management host servers, these servers are monitored during testing to ensure they are performing at an expected performance level with no bottlenecks.
- PowerFlex REST API captures the storage performance data (Read & Write), IOPS, and latency through PowerFlex Gateway during each test run.

4.3.2 Resource utilization thresholds

The purpose of this test is to determine the user density at a reasonable system load. Testing to system failure is out of scope. To achieve a reasonable system load, target thresholds for system resources are set in the table below. These thresholds reflect a system that is well utilized but not near failure.

Table 5 Resource utilization thresholds

Metrics	Target Threshold
Avg CPU Usage	85%
Avg Memory Utilization (Active)	85%
Consumed Memory	<100%
Network Throughput	85%
Storage Latency	20 milliseconds

4.4 Standard VDI test results and analysis

Management cluster – Deployed management infrastructure VMs (vCenter server, Horizon servers, Active Directory, DHCP, DNS, Login VSI, Launcher VMs) on a separate cluster.

Compute cluster – Three PowerFlex R740xd hyperconverged nodes are used in the cluster to create desktop pools with different pool sizes for testing multiple workloads.

Table 6 Desktop pool setting applied

Attribute	Specification
Pool Type	Automated
VM Type	Instant clone
User Assignment	Floating
Datastore	4 TB
Display Protocol	VMware BLAST

For hardware and software specifications used for this solution, see Configuration details.

All test results graphs include the performance of the platform during the deletion and recreation of the instant clone virtual machines after all users log off when the test run has completed. The different phases of the test cycle are displayed in the test results graphs later in this document as ‘Logon’, ‘Steady State’, ‘Logoff’, and ‘Recreate Clones’. All graphs have time on the X-axis and the metric displayed on the Y-axis.

4.5 About Login VSI

Login VSI (www.loginvsi.com) is the industry standard in synthetic load testing and active monitoring for SBC, VDI and DaaS environments (infrastructure and applications). Typical customers are enterprises with centralized desktop environments and/or business critical applications running in VDI, and all the major IT vendors that offer well-performing solutions for VDI.

The Login VSI Enterprise Edition offers a unique combination of synthetic load-testing and pro-active monitoring capabilities, allowing enterprises to design, build and maintain VDI environments (both infrastructure and applications) that can provide, and safeguard, the optimal end-user experience.

The Login VSI load-testing solution generates a large number of synthetic users to test and protect the performance and scalability of your new and existing VDI, SBC and DaaS deployments. The Login PI active monitoring solution uses a single synthetic user running 24/7, to safeguard performance and availability of your virtual desktop infrastructure and applications. The Login AT bulk application testing solution checks the availability and health of large numbers of applications, fast and efficient.

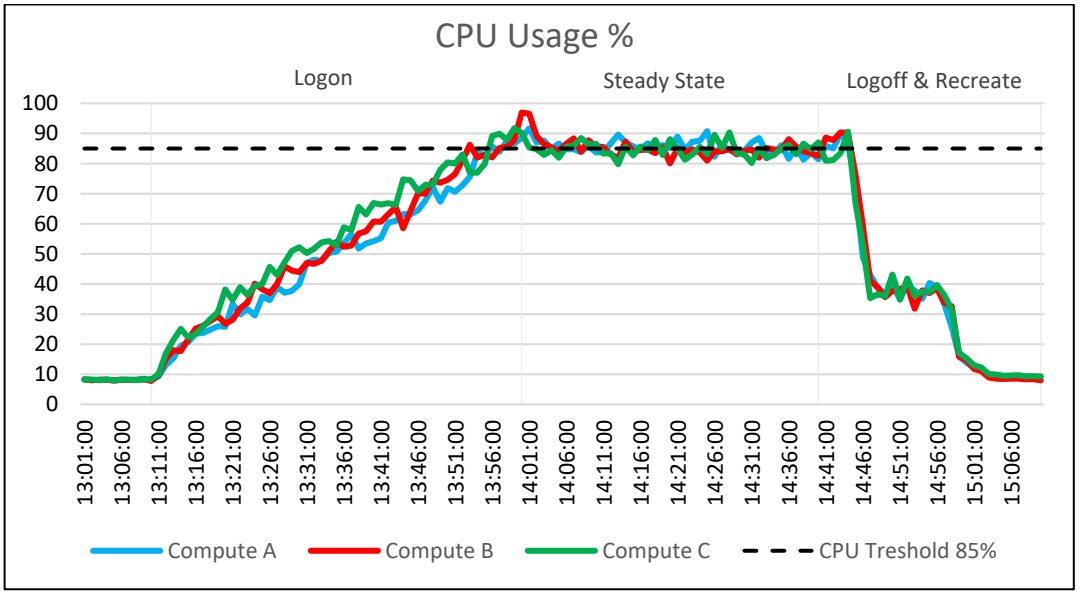
For more information, or a free trial, please visit www.loginvsi.com.

Login VSI accepts no responsibility regarding this publication in any way and cannot be held accountable for any damages following from, or related to, any information contained within this publication, or any conclusions that may be drawn from it.

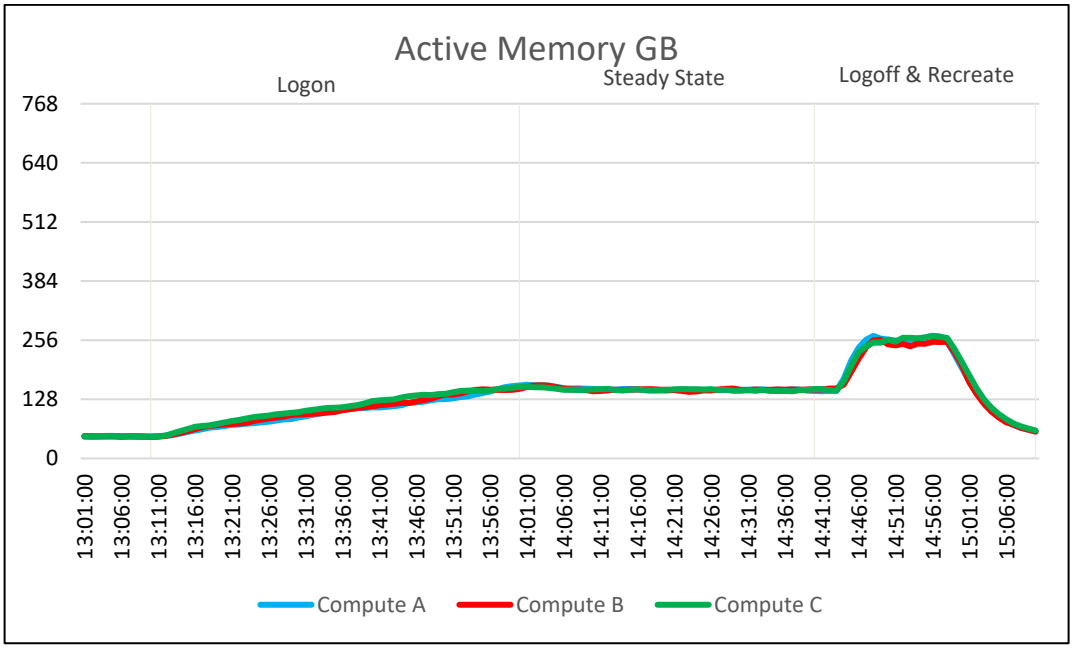
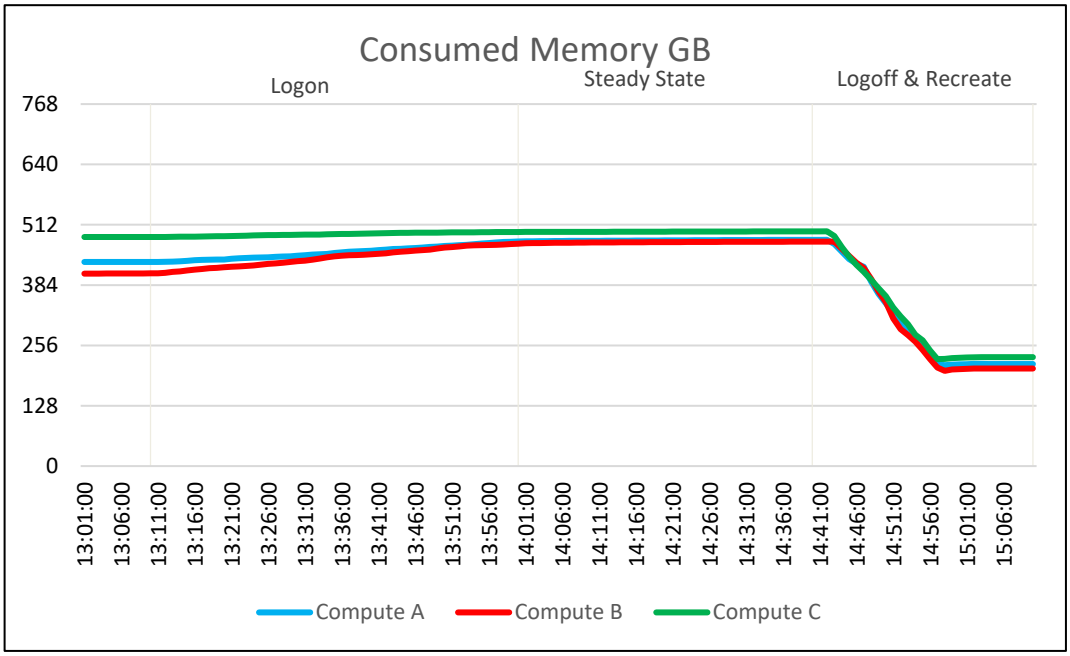
4.5.1 Testprofile1 - Task Worker

The total user density achieved for this test is 154 per host and 462 in the cluster.

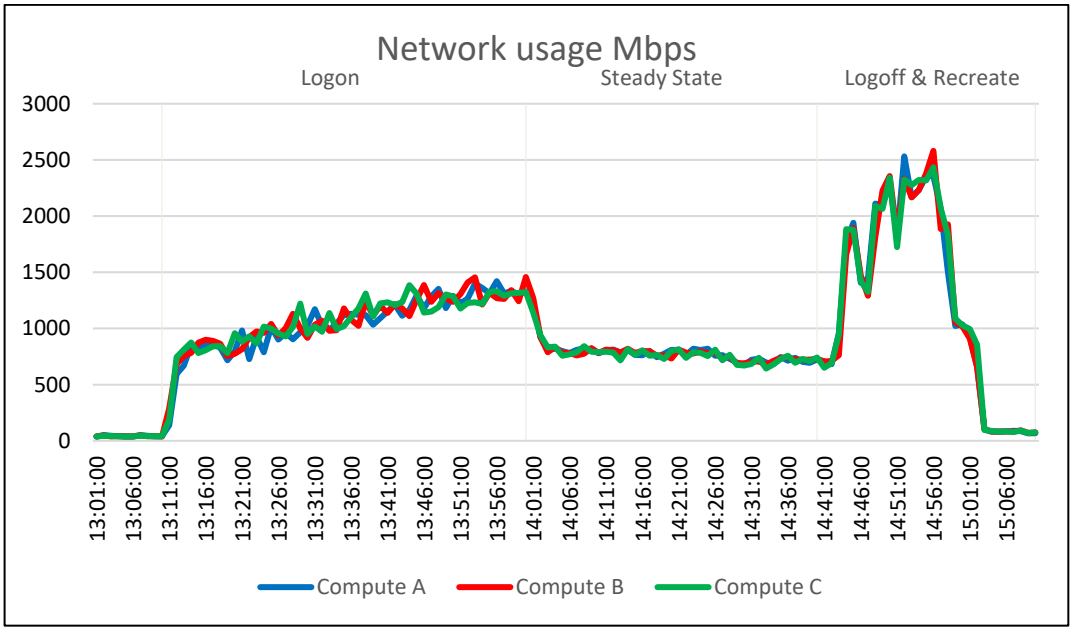
CPU Usage – Login VSI Task Worker

Metric		CPU usage graph
CPU usage at start	8%	
CPU usage at steady state	85%	
Analysis		CPU usage steadily increased during logon phase and converged at the defined threshold during steady state phase of the test. The compute nodes only showed slight discrepancies in usage during the logon phase which appears to be a normal distribution of usage. There were no anomalies seen during testing.

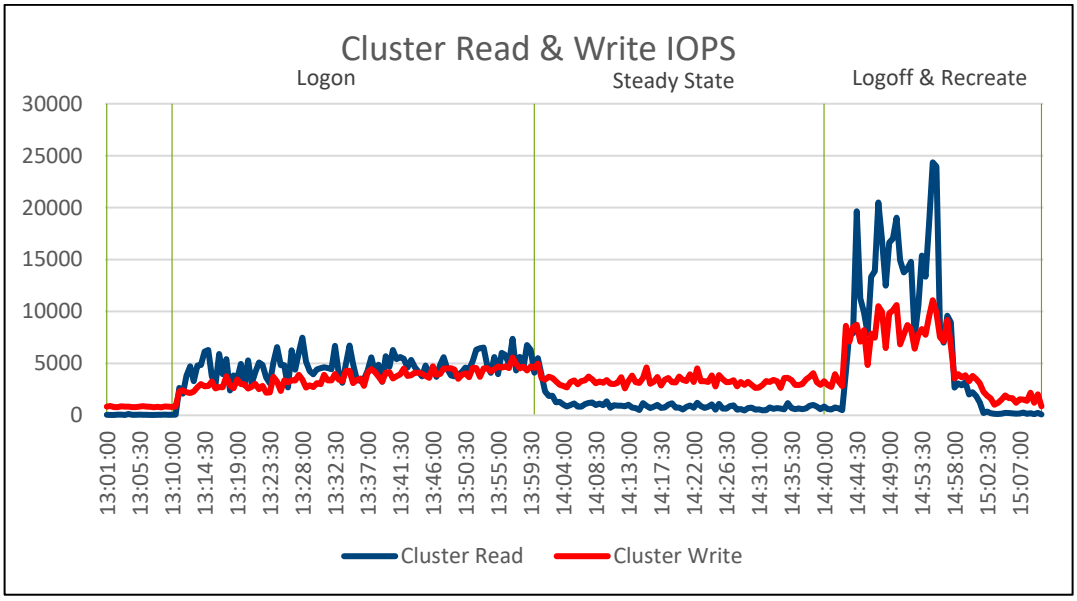
Memory – Login VSI Task Worker

Metric		Active and consumed memory graph
Total memory	768 GB	
Max memory consumed	498 GB	
Max Active memory usage	265 GB	
Analysis		Active and consumed memory remained below the physical capacity installed on each host. Memory swapping and paging was not observed in this test which indicates the system was performing as expected. There were no anomalies seen during testing.

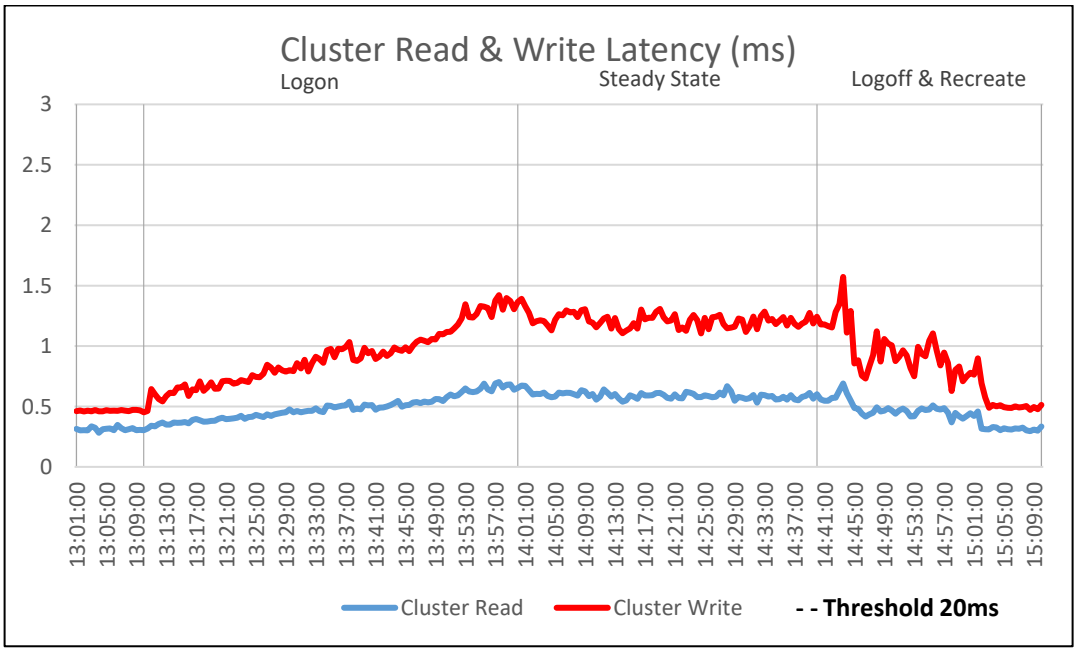
Network – Login VSI Task Worker

Metric		Network Usage
Steady state peak	1,457 Mbps	
Max host peak	2,580 Mbps	
Steady state average	791 Mbps	
Analysis		Network bandwidth steadily increased during logon phase and converged during steady state phase of the test. The compute nodes only showed slight discrepancies in bandwidth during the logon phase which appears to be a normal distribution of network traffic. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones. There were no anomalies seen during testing.

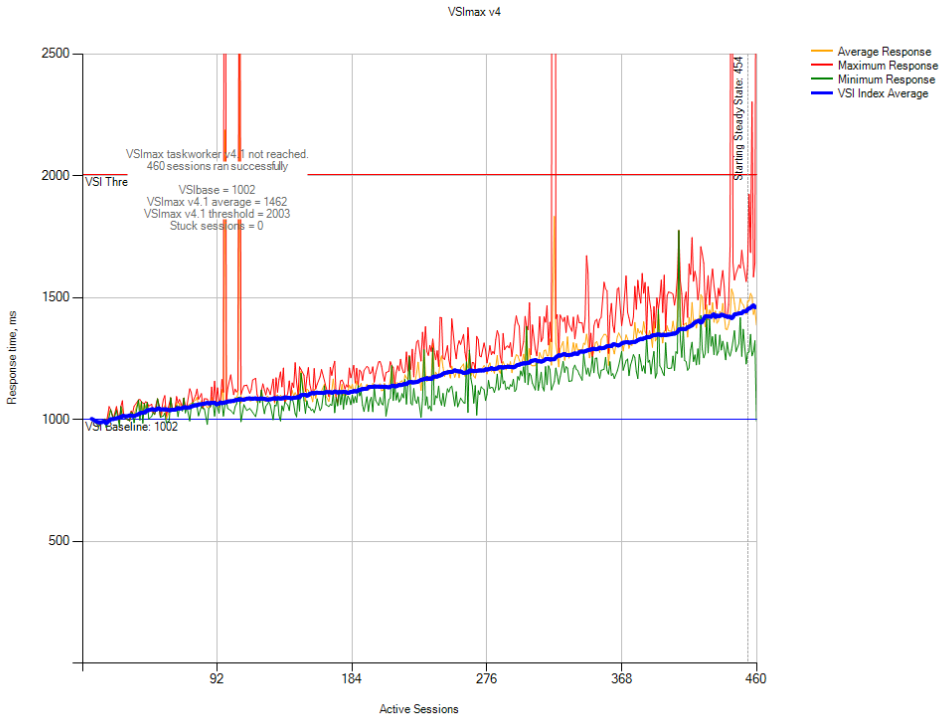
IOPS – Login VSI Task Worker

Metric		Cluster IOPS graph
Max IOPS	35,444	
Steady state peak	5,738	
Steady state average	4,180	
Analysis		Total IOPS generated is as expected. There are more read IOPS than write IOPS as users logged on to their desktops and read their profile data. During steady state this pattern changed where there are more write IOPS than read IOPS. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones.

Latency – Login VSI Task Worker

Metric		Cluster Latency
Max read + write latency	2.26	
Avg steady state latency	1.81	
Analysis		Cluster latency remained at an acceptable level and below the defined threshold. There were no anomalies seen during testing.

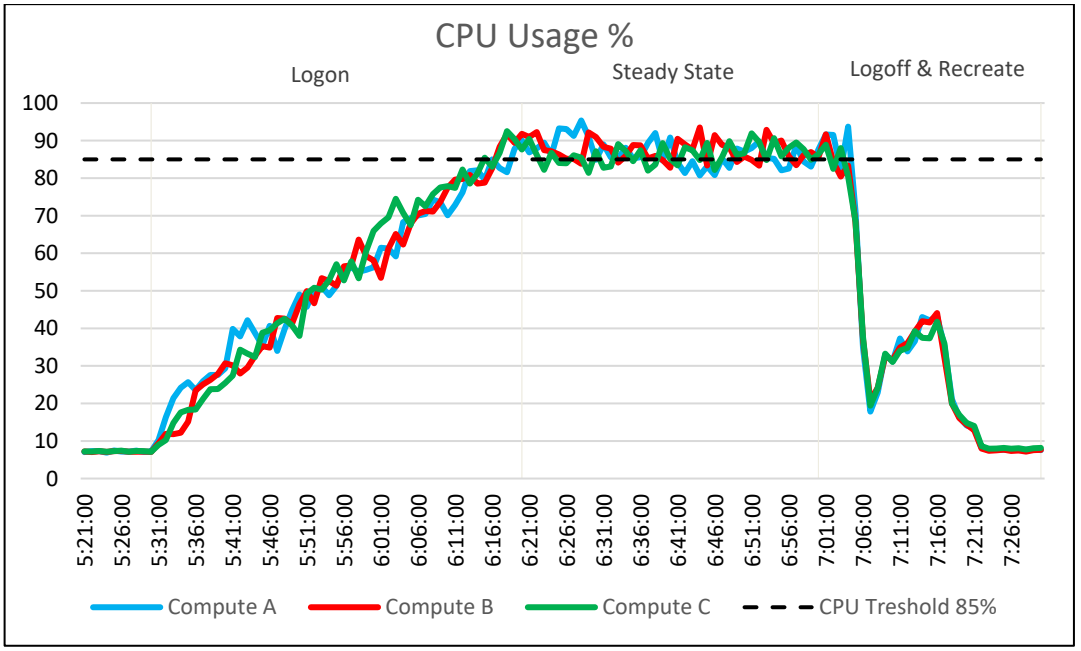
User Experience – Login VSI Task Worker

Metric		Login VSI Max user experience graph
VSI base	1,002	
Avg VSI max	1,462	
VSI threshold	2,003	
Stuck sessions	0	
Analysis		The VSI base is at an acceptable score. The VSI index average shows that the system responds appropriately as additional load is added and the system behaves in a predictable manner. The response times are tightly grouped; meaning that there is a consistent user experience across VDI sessions.

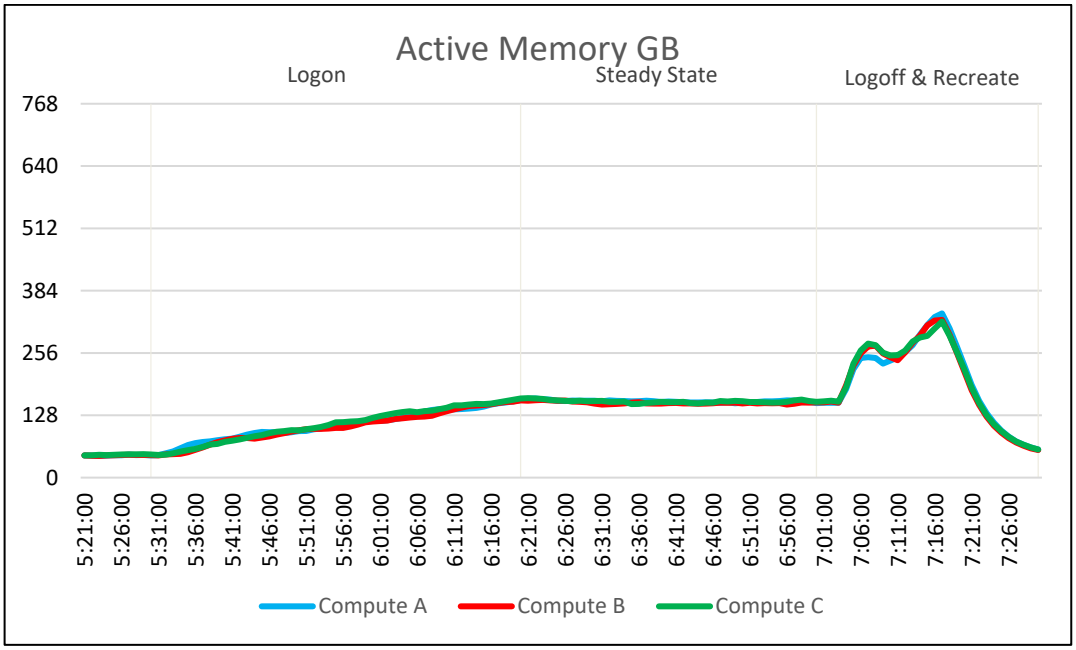
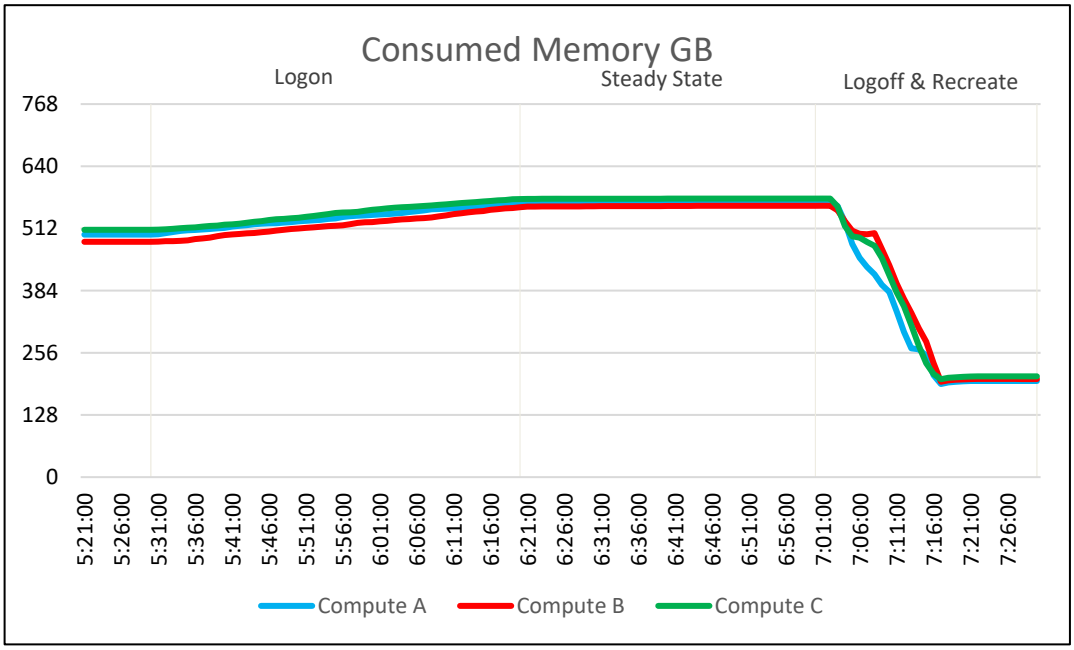
4.5.2 Testprofile2 - Knowledge Worker

The total user density achieved for this test is 135 per host and 405 in the cluster.

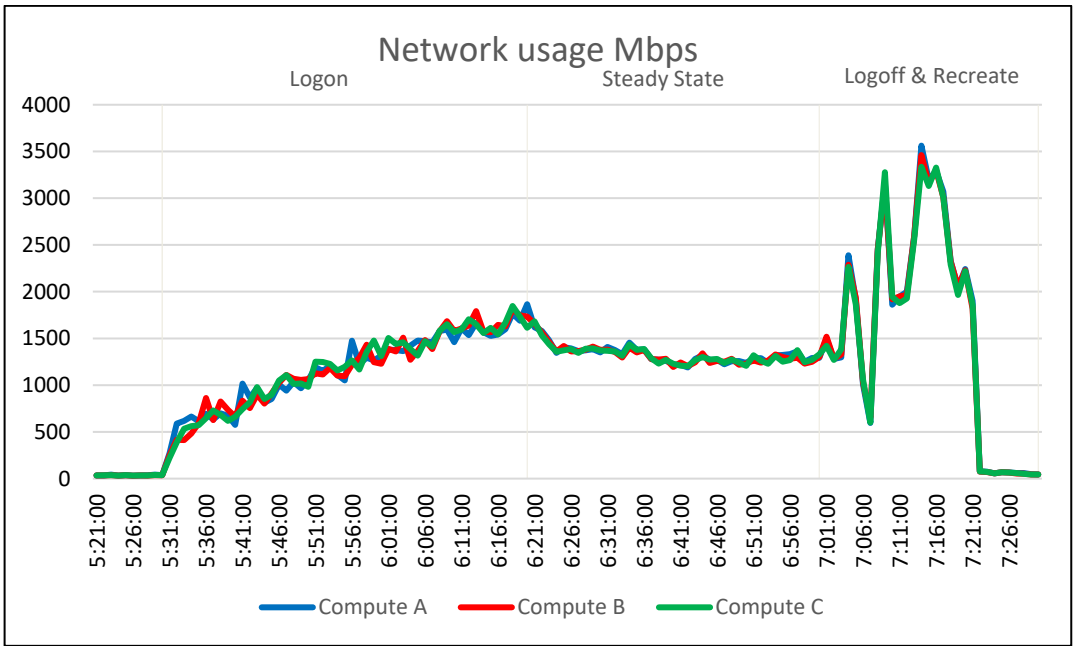
CPU Usage – Login VSI Knowledge Worker

Metric		CPU usage graph
CPU usage at start	8%	
CPU usage at steady state	86%	
Analysis		CPU usage steadily increased during logon phase and converged at the defined threshold during steady state phase of the test. The compute nodes only showed slight discrepancies in usage during the logon phase which appears to be a normal distribution of usage. There were no anomalies seen during testing.

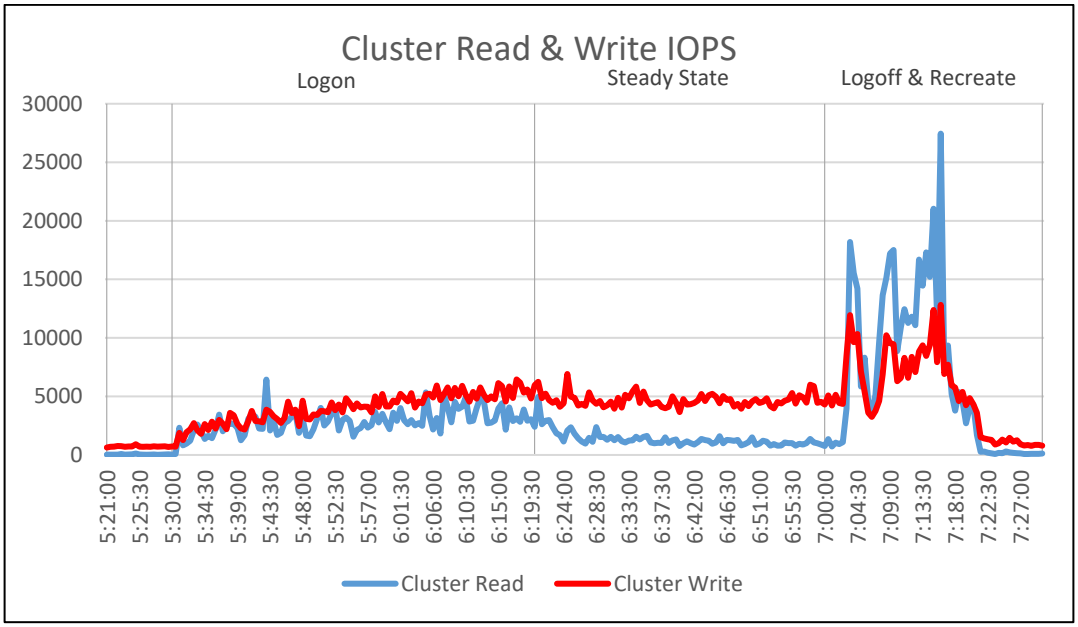
Memory – Login VSI Knowledge Worker

Metric		Active and consumed memory graph
Total memory	768 GB	
Max memory consumed	573 GB	
Max Active memory usage	337 GB	
Analysis		Active and consumed memory remained below the physical capacity installed on each host. Memory swapping and paging was not observed in this test which indicates the system was performing as expected. There were no anomalies seen during testing.

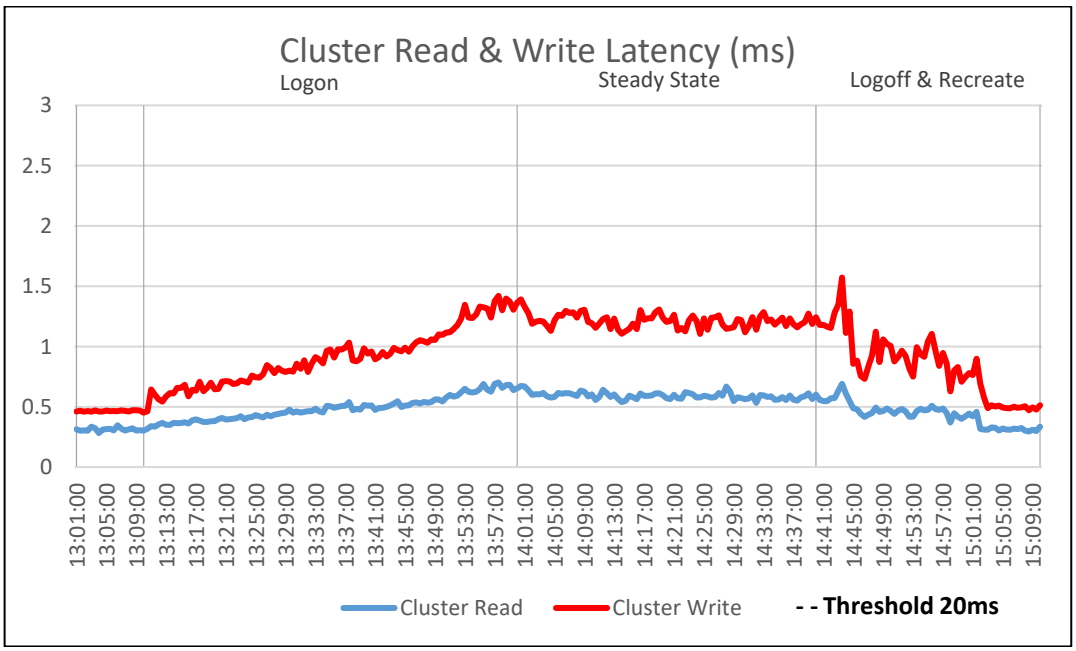
Network – Login VSI Knowledge Worker

Metric		Network Usage
Steady state peak	1,866 Mbps	
Max host peak	3,562 Mbps	
Steady state average	1,333 Mbps	
Analysis		Network bandwidth steadily increased during logon phase and converged during steady state phase of the test. The compute nodes only showed slight discrepancies in bandwidth during the logon phase which appears to be a normal distribution of network traffic. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones. There were no anomalies seen during testing.

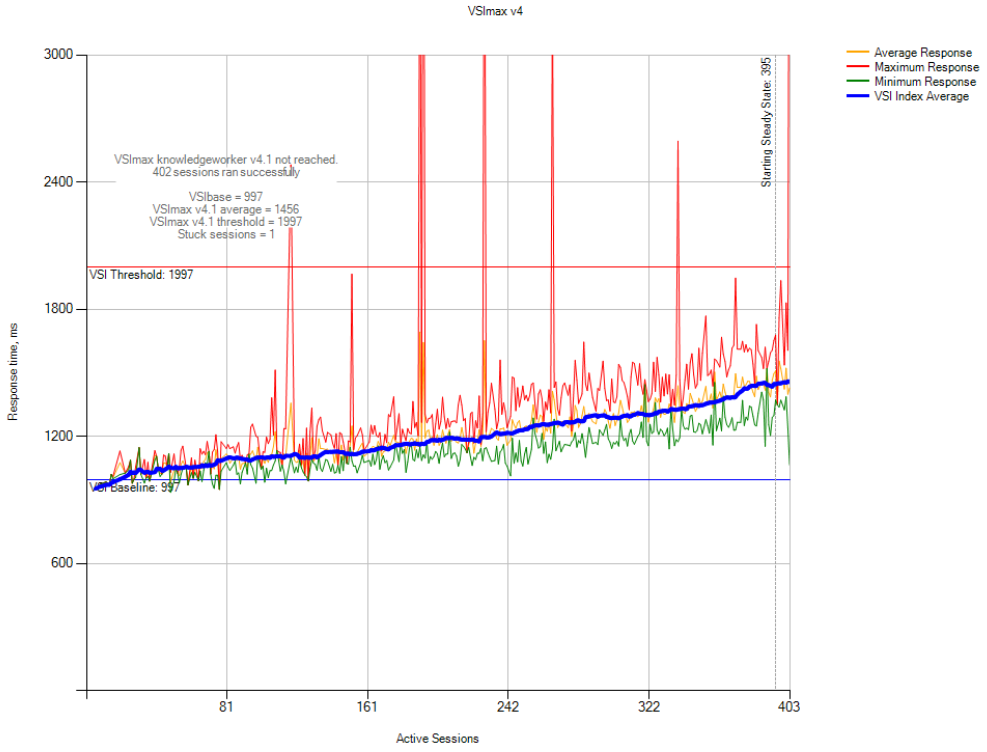
IOPS – Login VSI Knowledge Worker

Metric		Cluster IOPS graph
Max IOPS	40,263	
Steady state peak	11,196	
Steady state average	6,040	
Analysis		Total IOPS generated is as expected. There are more read IOPS than write IOPS as users logged on to their desktops and read their profile data. During steady state this pattern changed where there are more write IOPS than read IOPS. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones.

Latency – Login VSI Knowledge Worker

Metric		Cluster Latency
Max read + write latency	2.91	
Avg steady state latency	1.92	
Analysis		Cluster latency remained at an acceptable level and below the defined threshold. There were no anomalies seen during testing.

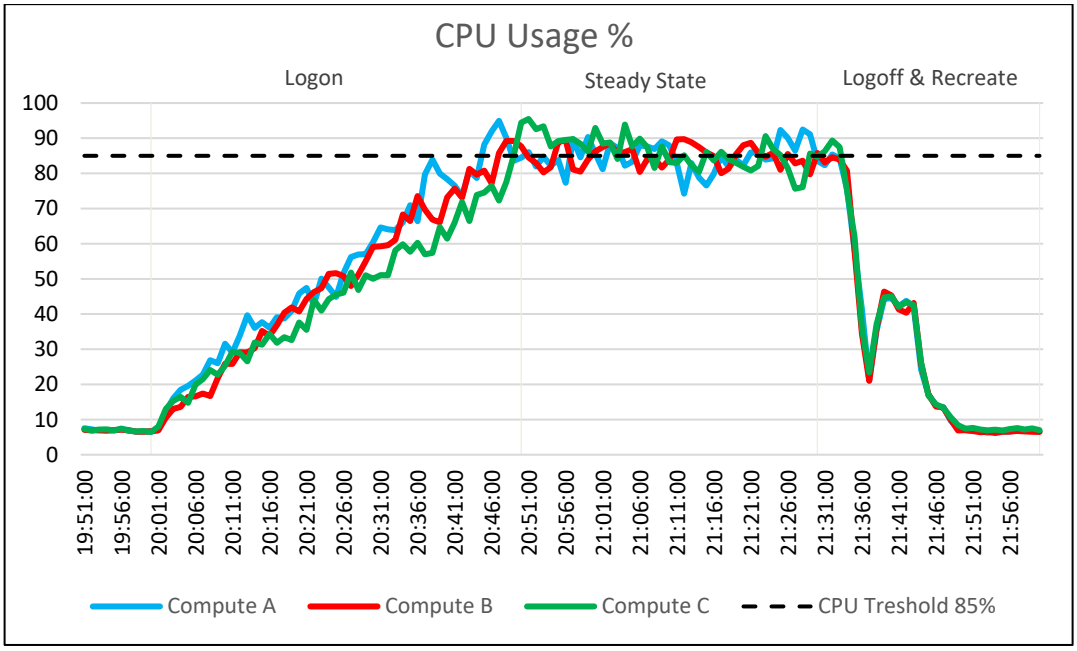
User Experience – Login VSI Knowledge Worker

Metric		Login VSI Max user experience graph
VSI base	997	
Avg VSI max	1,456	
VSI threshold	1,997	
Stuck sessions	1	
Analysis		The VSI base is at an acceptable score. The VSI index average shows that the system responds appropriately as additional load is added and the system behaves in a predictable manner. The response times are tightly grouped; meaning that there is a consistent user experience across VDI sessions.

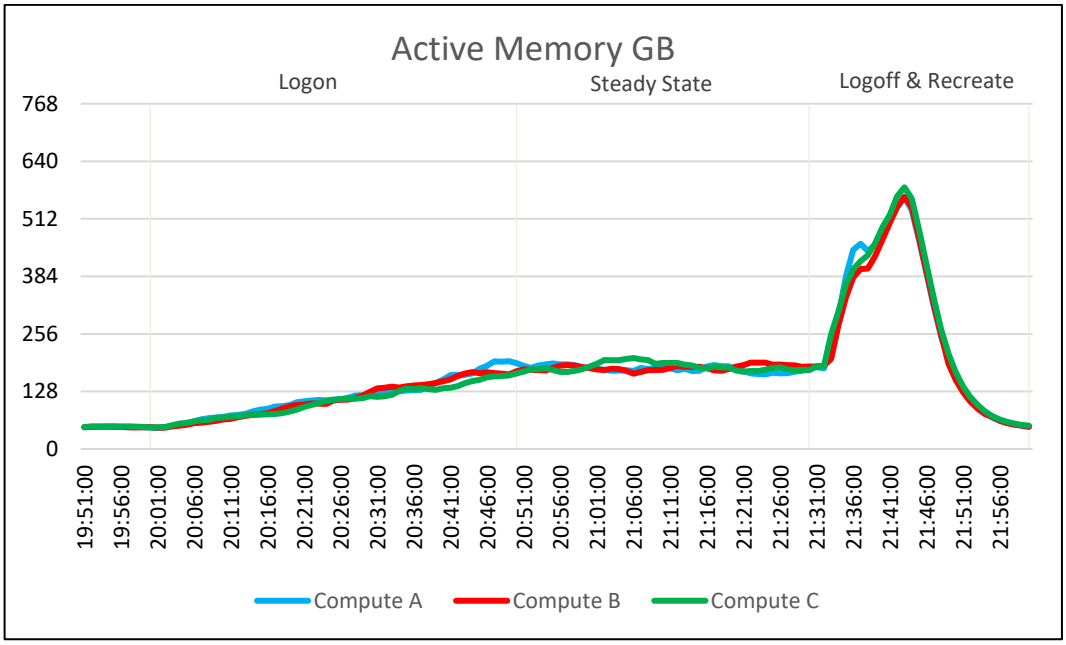
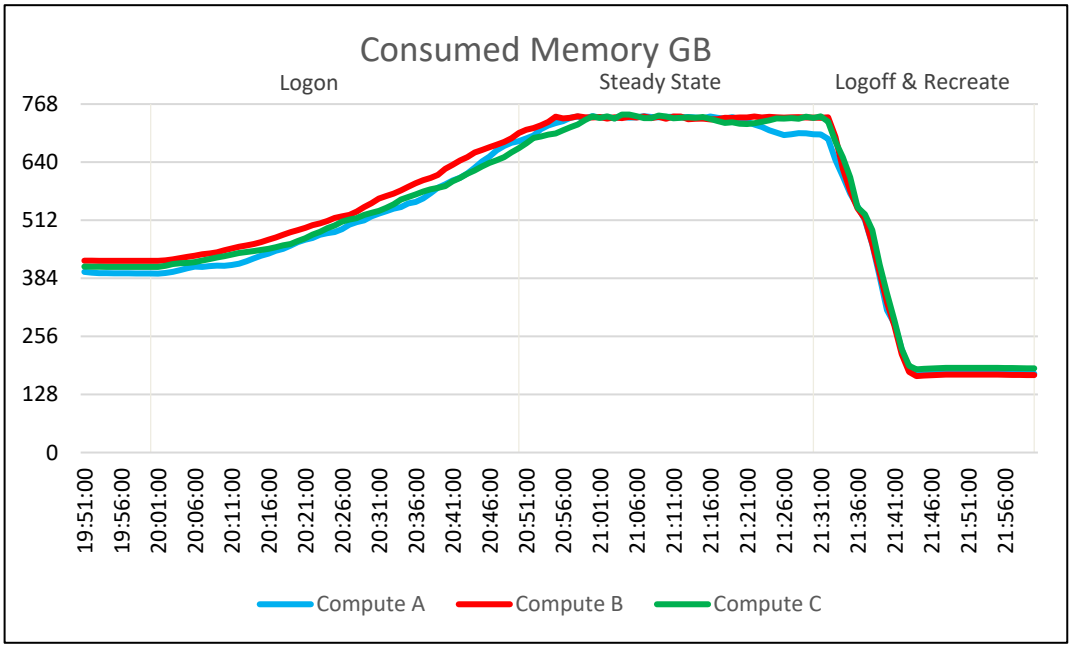
4.5.3 Testprofile3 - Power Worker

The total user density achieved for this test is 112 per host and 336 in the cluster.

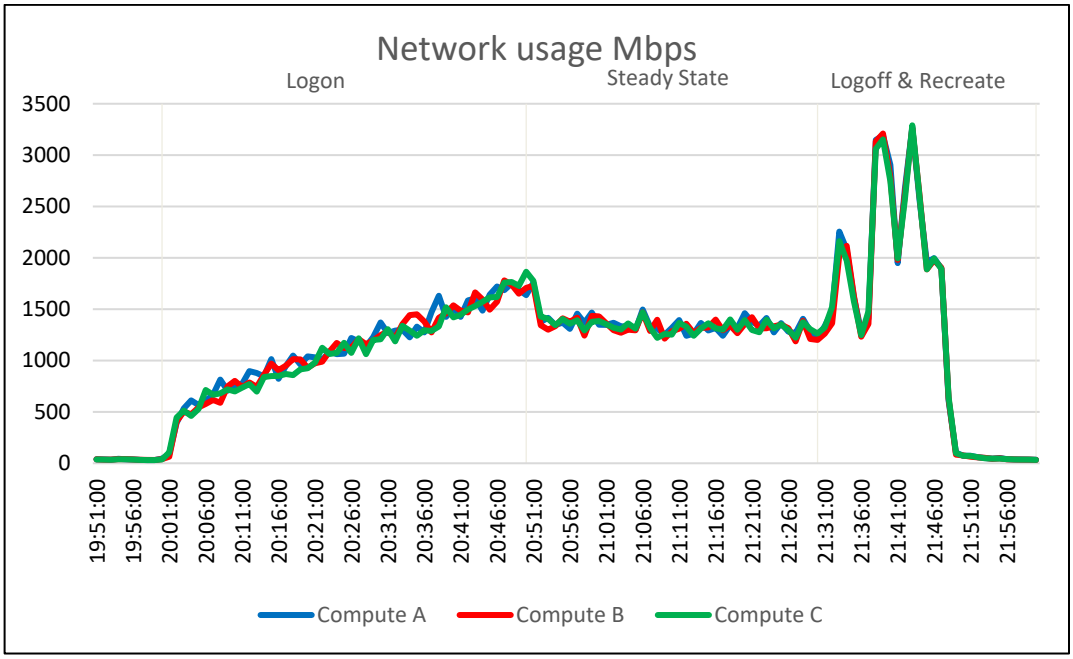
CPU Usage – Login VSI Power Worker

Metric		CPU usage graph
CPU usage at start	7%	
CPU usage at steady state	85%	
Analysis		CPU usage steadily increased during logon phase and converged at the defined threshold during steady state phase of the test. The compute nodes only showed slight discrepancies in usage during the logon phase which appears to be a normal distribution of usage. There were no anomalies seen during testing.

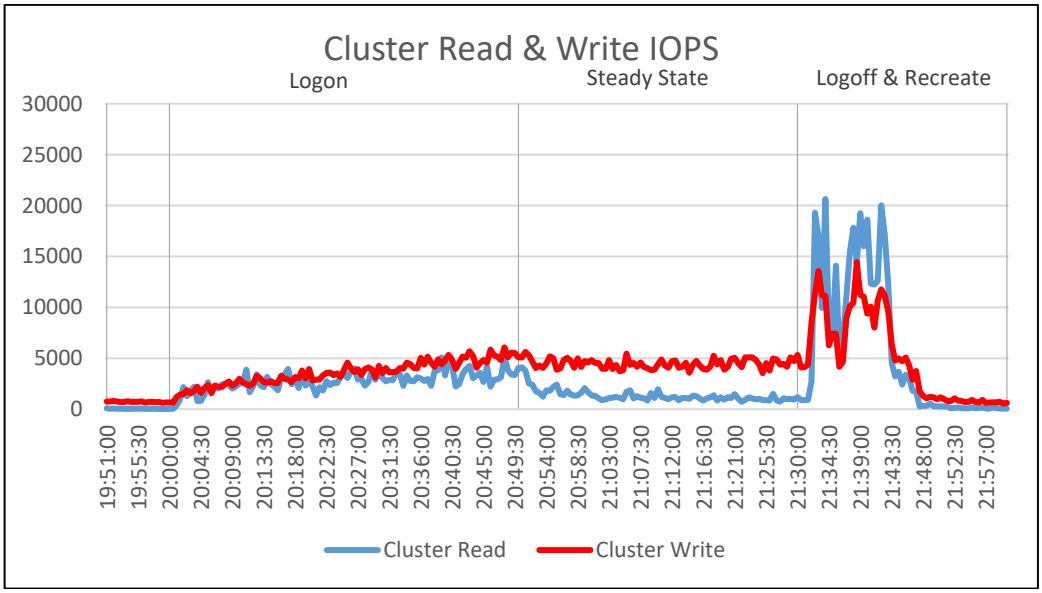
Memory – Login VSI Power Worker

Metric		Active and consumed memory graph
Total memory	768 GB	
Max memory consumed	744 GB	
Max Active memory usage	582 GB	
Analysis		Active and consumed memory remained below the physical capacity installed on each host. Memory swapping and paging was not observed in this test which indicates the system was performing as expected. There were no anomalies seen during testing.

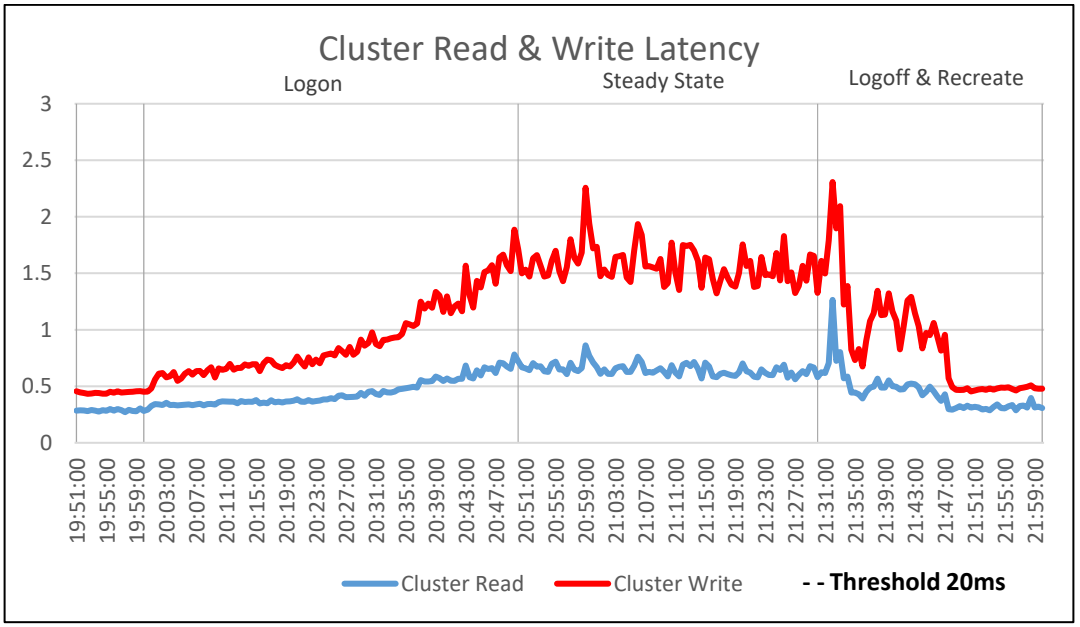
Network – Login VSI Power Worker

Metric		Network Usage
Steady state peak	1,864 Mbps	
Max host peak	3,289 Mbps	
Steady state average	1,353 Mbps	
Analysis		Network bandwidth steadily increased during logon phase and converged during steady state phase of the test. The compute nodes only showed slight discrepancies in bandwidth during the logon phase which appears to be a normal distribution of network traffic. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones. There were no anomalies seen during testing.

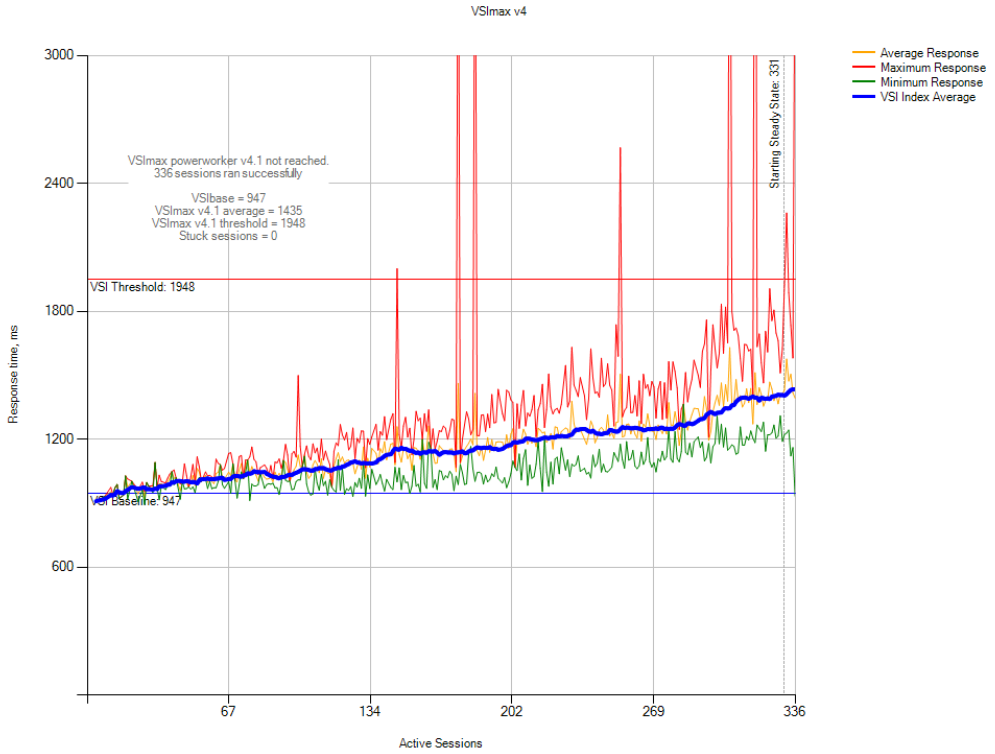
IOPS – Login VSI Power Worker

Metric		Cluster IOPS graph
Max IOPS	31,803	
Steady state peak	9,440	
Steady state average	5,861	
Analysis		Total IOPS generated is as expected. There are more read IOPS than write IOPS as users logged on to their desktops and read their profile data. During steady state this pattern changed where there are more write IOPS than read IOPS. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones.

Latency – Login VSI Power Worker

Metric		Cluster Latency
Max read + write latency	3.57	
Avg steady state latency	2.22	
Analysis		Cluster latency remained at an acceptable level and below the defined threshold. There were no anomalies seen during testing.

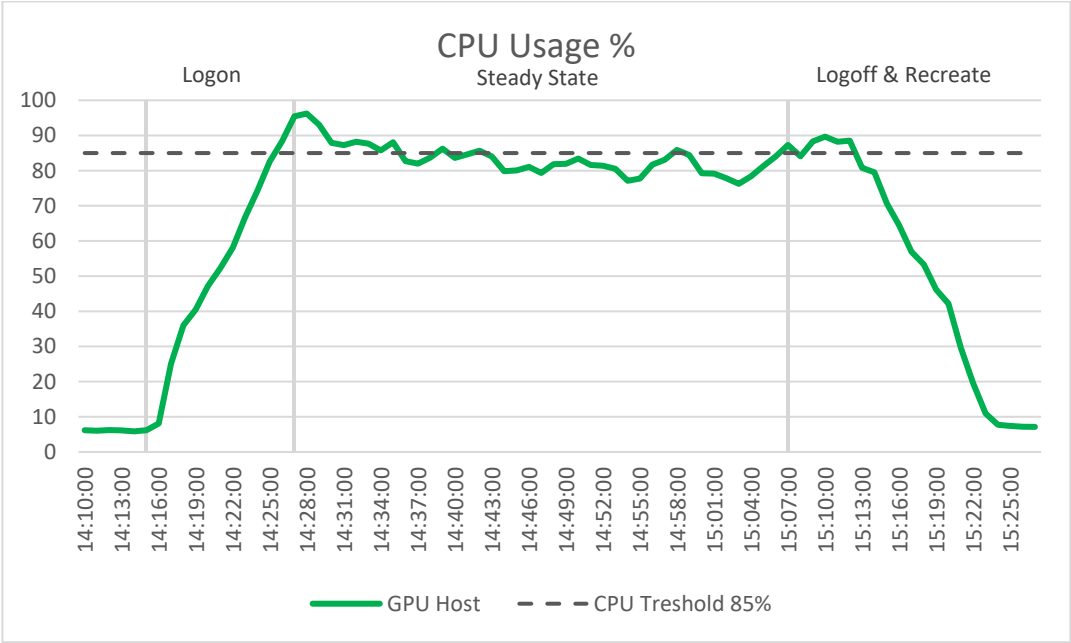
User Experience – Login VSI Power Worker

Metric		Login VSI Max user experience graph
VSI base	947	
Avg VSI max	1,435	
VSI threshold	1,948	
Stuck sessions	0	
Analysis		The VSI base is at an acceptable score. The VSI index average shows that the system responds appropriately as additional load is added and the system behaves in a predictable manner. The response times are tightly grouped; meaning that there is a consistent user experience across VDI sessions.

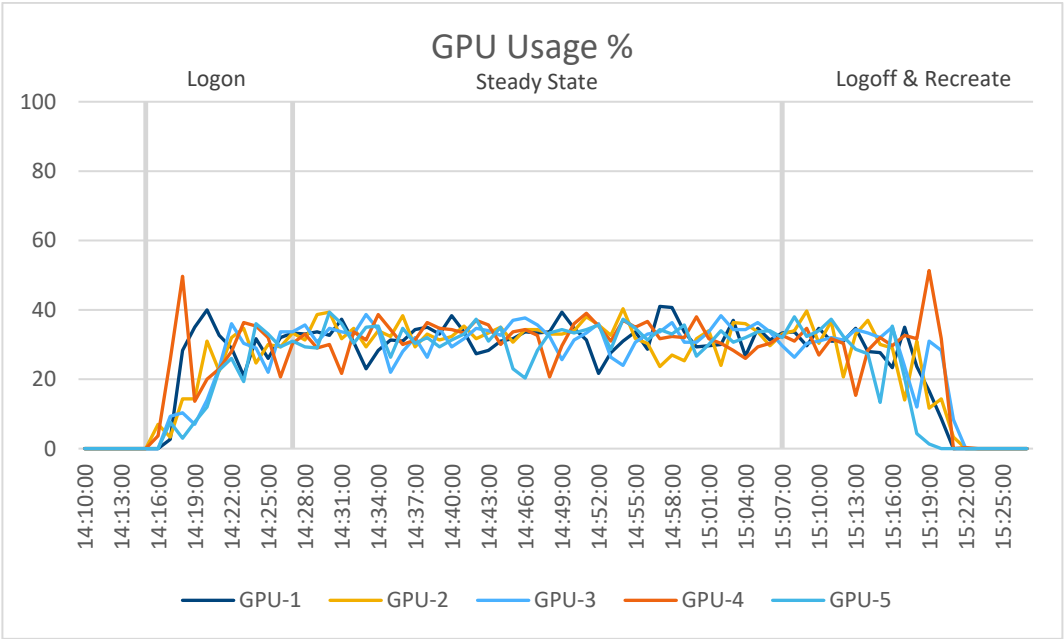
4.5.4 Testprofile4 - Multimedia Worker

The total user density achieved for this test is 70 (1B vGPU profile) desktops per host, 5 x NVIDIA T4 cards installed and Fixed share scheduler (vGPU scheduling policy) is used for testing.

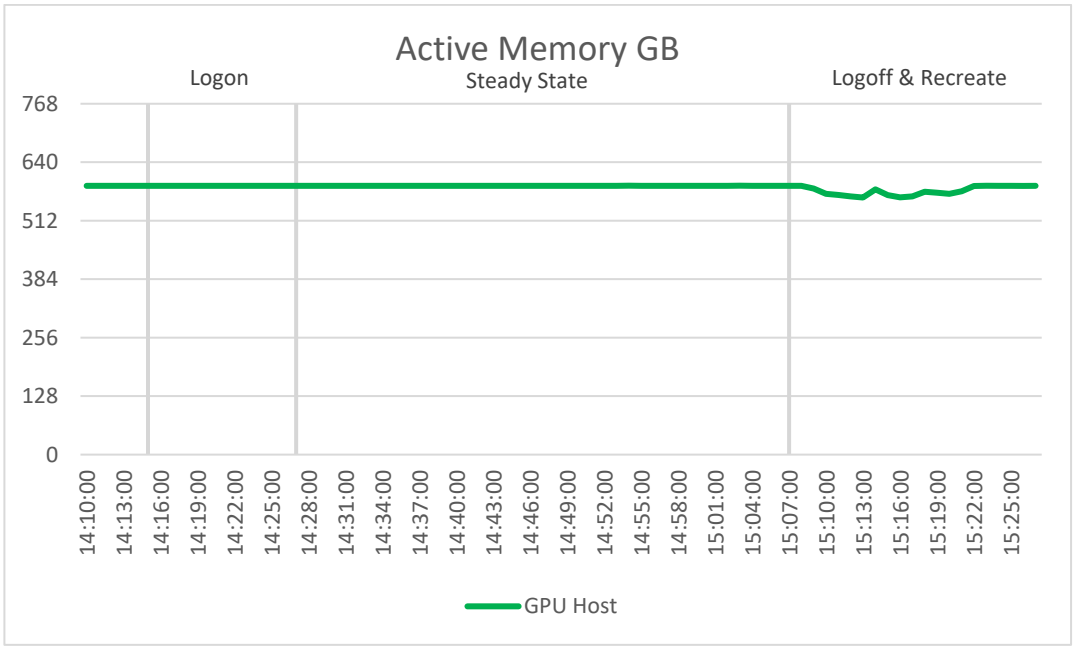
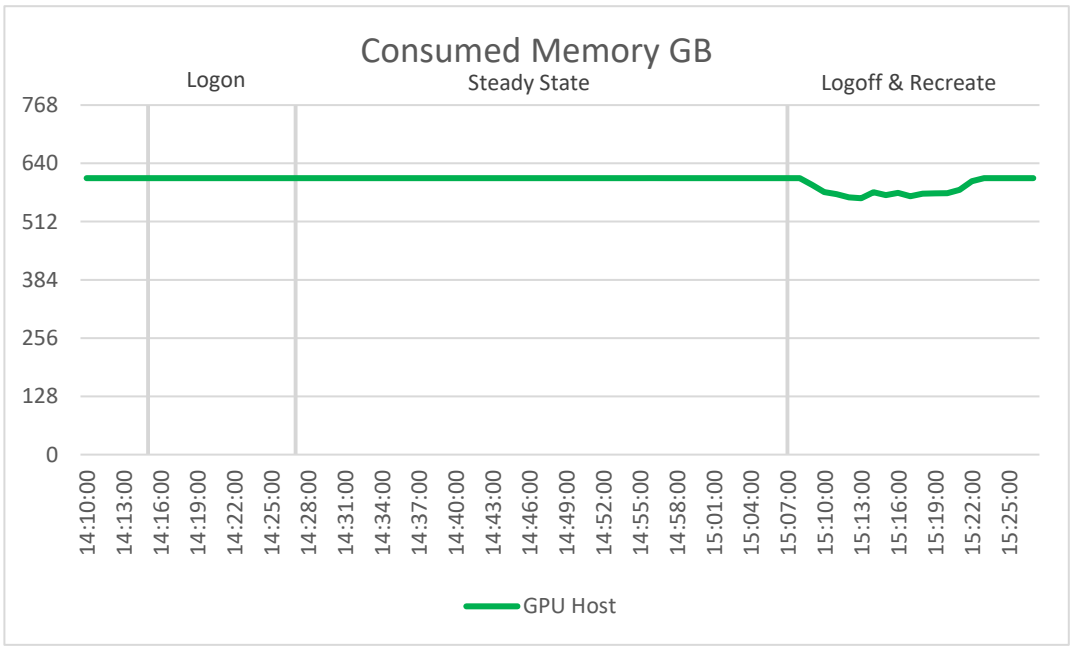
CPU Usage – Login VSI Multimedia Worker

Metric		CPU usage graph
CPU usage at start	6%	
CPU usage at steady state	85%	
Analysis		CPU usage steadily increased during logon phase and converged at the defined threshold during steady state phase of the test. The compute nodes only showed slight discrepancies in usage during the logon phase which appears to be a normal distribution of usage. There were no anomalies seen during testing.

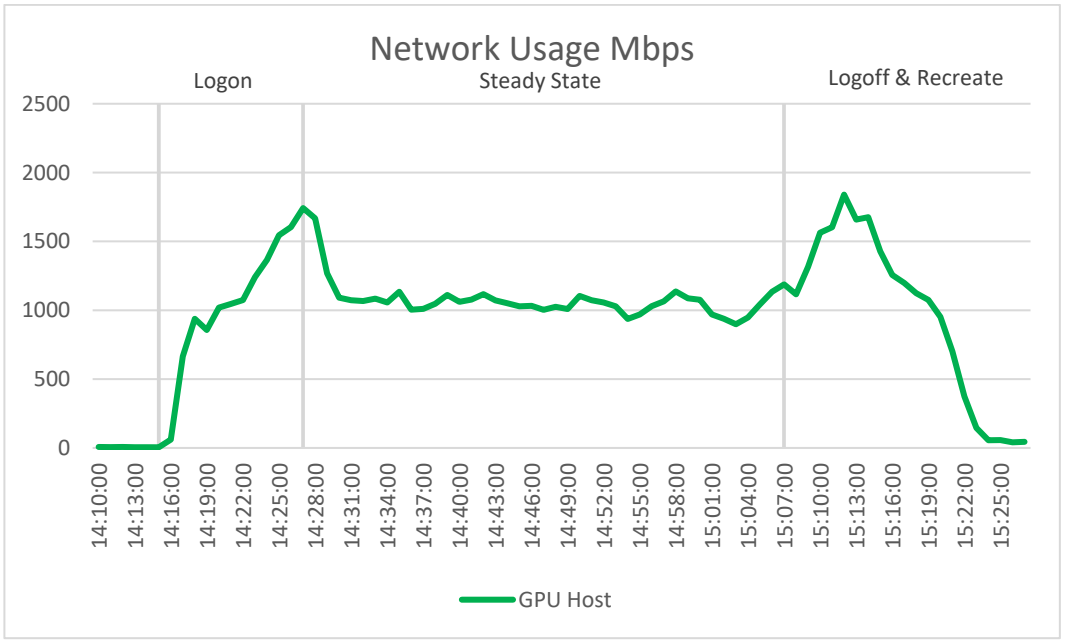
GPU Usage – Login VSI Multimedia Worker

Metric		GPU usage graph
GPU usage at start	0%	
GPU usage at steady state	33%	
Analysis		GPU remained at a reasonable level throughout the test indicating that there are enough GPU resources for all users. No throttling of the GPUs was experienced during the test.

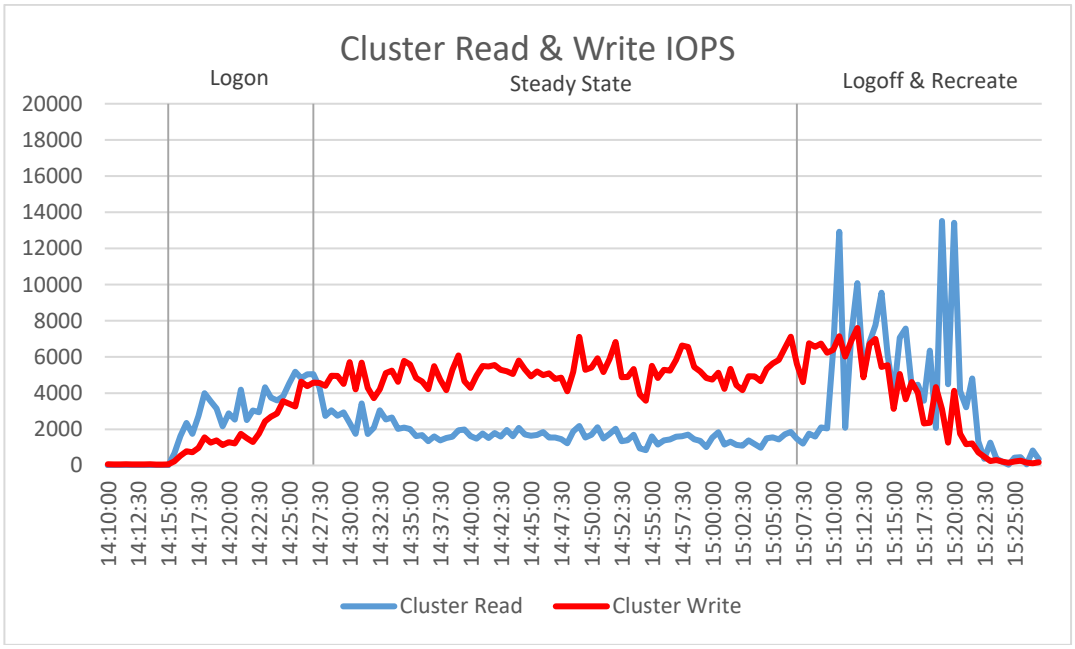
Memory – Login VSI Multimedia Worker

Metric		Active and consumed memory graph
Total memory	768 GB	
Max memory consumed	588 GB	
Max Active memory usage	607 GB	
Analysis		Active and consumed memory remained below the physical capacity installed on each host. Memory swapping and paging was not observed in this test which indicates the system was performing as expected. There were no anomalies seen during testing.

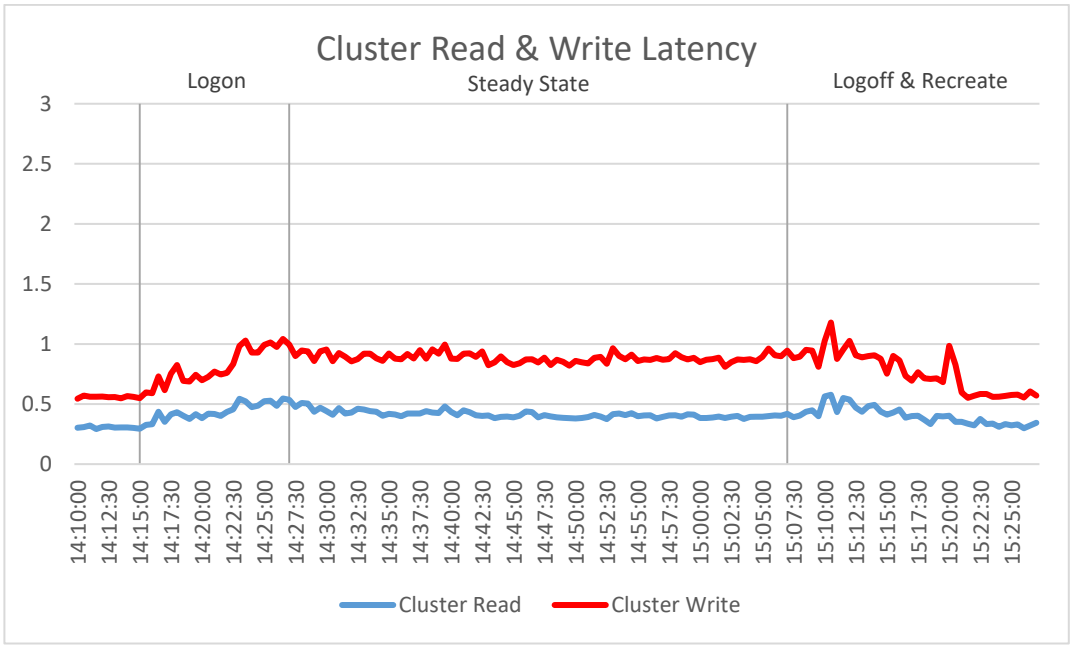
Network – Login VSI Multimedia Worker

Metric		Network Usage
Steady state peak	1,742 Mbps	
Max host peak	1,840 Mbps	
Steady state average	1,085 Mbps	
Analysis		Network bandwidth steadily increased during logon phase and converged during steady state phase of the test. The compute nodes only showed slight discrepancies in bandwidth during the logon phase which appears to be a normal distribution of network traffic. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones. There were no anomalies seen during testing.

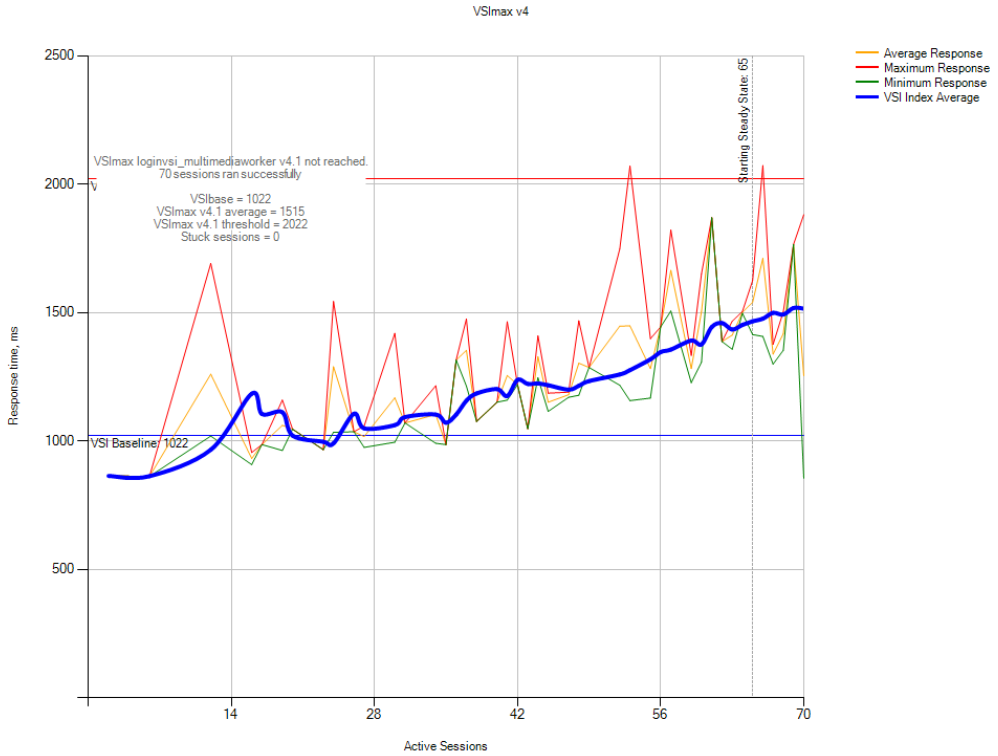
IOPS – Login VSI Multimedia Worker

Metric		Cluster IOPS graph
Max IOPS	20,064	
Steady state peak	9,627	
Steady state average	6,935	
Analysis		Total IOPS generated is as expected. There are more read IOPS than write IOPS as users logged on to their desktops and read their profile data. During steady state this pattern changed where there are more write IOPS than read IOPS. The large spike during the logoff and re-create clones phases is due to the recreation of instant clones.

Latency – Login VSI Multimedia Worker

Metric		Cluster Latency
Max read + write latency	1.75	
Avg steady state latency	1.30	
Analysis		Cluster latency remained at an acceptable level and below the defined threshold. There were no anomalies seen during testing.

User Experience – Login VSI Multimedia Worker

Metric		Login VSI Max user experience graph
VSI base	1,022	
Avg VSI max	1,515	
VSI threshold	2,022	
Stuck sessions	0	
Analysis		The VSI base is at an acceptable score. The VSI index average shows that the system responds appropriately as additional load is added and the system behaves in a predictable manner. The response times are tightly grouped; meaning that there is a consistent user experience across VDI sessions.

4.6 Summary

The following table summarizes the test results for the compute hosts using the various workloads and configurations:

Table 7 Test result summary for Horizon instant clone desktops

Workload	Density Per Host	Avg CPU %	Avg Memory Consumed GB	Avg Memory Active GB	Avg IOPS/User	Avg Net Mbps/User
Task Worker	154	85	483	149	10	6
Knowledge Worker	135	86	565	155	15	10
Power Worker	112	85	729	180	17	12
Multimedia Worker (1B vGPU profile)	70	85	607	588	98	15

- **Density Per Host** – The number of users per compute host that successfully completed the workload test within the acceptable resource limits for the host.
- **CPU Usage** – The figure shown in the table, 'Avg' CPU %', is the combined average CPU usage of all compute hosts over the steady state period.
- **Consumed Memory** – Consumed memory is the average consumed amount of memory host physical memory consumed by a virtual machine, host, or cluster. The figure 'Avg' Memory Consumed GB' in the table is the across all compute hosts over the steady state period.
- **Active Memory** – Active Memory is the amount of memory that is actively used, as estimated by VMkernel based on recently touched memory pages. The figure 'Avg' Memory Active GB' in the table is the average amount of guest "physical" memory actively used across all compute hosts over the steady state period.
- **Disk IOPS** – Disk IOPS are calculated from the total Disk IOPS figure at the start of the steady state period divided by the number of users to produce the 'IOPS/User' figure.
- **Network Usage** – Network Usage per User. The figure that is shown in the table 'Avg' Net Mbps/User' is the combined average of all compute hosts over the steady state period divided by the number of users on a host in Megabits per second.

5 Best practices

This section provides the recommendations and best practices for VMware Horizon VDI environment.

5.1 Management infrastructure

Combining management infrastructure and virtual desktops in the same cluster is acceptable for small-scale and pilot environments to validate the VDI benefits. Dell EMC recommends having dedicated nodes for the management infrastructure and VDI compute within the same or different clusters in the large-scale environment to avoid performance issues.

The following table lists the sizing requirement for management server components:

Table 8 Management VMs

VM	vCPU	Memory (GB)	Hard disk (GB)
vCenter Appliance	4	16	340
Connection Server	4	12	60
View Composer	4	8	60
Active Directory	4	8	60
DHCP/DNS	4	8	60
File Server	4	8	2048
SQL Server	4	8	200
SVM	8	20	16
PowerFlex Gateway	2	4	60
PowerFlex Manager	8	32	200

Note: Number of VMs and configuration may differ depends on the size of the overall deployment.

5.2 Windows 10 optimization

Windows 10 optimization is a necessary step to improve the overall VDI performance and increase the VM density per host. Remove the unwanted applications from the base image and disable the services which are not required. Use optimization tools provided by VDI vendors to achieve better performance.

For more information, see VMware OS Optimization Tool.

5.3 Scaling guidelines

The PowerFlex family provides great flexibility in scaling the VDI environment to reduce the initial and future cost of ownership. Adding additional nodes to the cluster to increase resources (CPU, memory, disk capacity, network and GPU) allows for both horizontal and vertical scaling to match requirements. The boundaries are based on the number of nodes supported for a vSphere 6.7 cluster which is 64 nodes, although the PowerFlex cluster can scale much larger, well beyond the boundaries of the hypervisor in use.

GPU and non-GPU compute hosts can co-exist in the same cluster, adding or removing a graphics workload within a running cluster can be achieved with no issues. Only PowerFlex R740xd supports graphic cards.

PowerFlex Manager greatly simplifies the steps required to add nodes to an existing cluster by providing templates to complete the process.

5.4 NVIDIA vGPU considerations

Best practices for sizing and configuring solutions requiring graphics accelerators include:

- GRID vPC licenses support up to 2 GB of frame buffer and up to two 4K monitors, or a single 5K monitor to cover most traditional VDI users. Maximum node density for graphics accelerated use can typically be calculated as the available frame buffer per node divided by the frame buffer size.
- Addition of GPU cards does not necessarily reduce CPU utilization. Instead, it enhances the user experience and offloads specific operations best performed by the GPU.
- Dell Technologies recommends using the BLAST protocol for vGPU enabled desktops. NVIDIA GPUs are equipped with encoders that support BLAST.
- When configuring NVIDIA M10 GPU cards in a solution, Dell Technologies recommends a maximum memory capacity of 768 GB, due to limitations in the Maxwell architecture. Pascal and Turing architectures do not have the same limitation.

NVIDIA vGPU scheduling

NVIDIA vGPU supports various scheduling policies that allow VMware vSphere Shared Direct (passthrough graphics) feature to share the vGPU resources optimally with multiple workload vms.

- **Best effort scheduler –** default policy in vSphere, allows a vGPU virtual machine to use the GPU processing cycles are not used by other vGPUs virtual machines.
- **Equal share scheduler –** shares the GPU processing equally to all the vGPU virtual machines are running.
- **Fixed share scheduler –** each vGPU has the fixed share of the GPU allocated to the virtual machine.

Equal share scheduler or Fixed share scheduler is recommended for VDI solutions, which reduces the overall CPU overhead to handle the vGPU scheduling for the virtual machines running in the ESXi host and improves the end-user graphics performance equivalently.

6 Conclusion

This reference architecture provides an overview of the PowerFlex family with the integration of Horizon for VDI solutions. The HCI deployment offering benefits customers by providing an initial lower hardware investment cost but provides flexibility to scale the VDI environment as the requirement grows without compromising the performance.

Validated Login VSI test results demonstrate the performance of multiple workloads hosted on the PowerFlex family to deliver the demanding IOPS to the virtual desktops with minimum latency. The test summary for multiple workload profiles assists the customers to size the VDI environment efficiently.

A Appendix

A.1 Configuration details

The table below describes the hardware and software components of the infrastructure used for performance analysis and characterization test.

Table 9 Component details

Component	Details
Compute Host Hardware	3x PowerFlex R740xd hyperconverged node • 2x 6248 Intel Xeon Gold @ 2.5 GHz, 20-Core Processors • 768 GB Memory @ 2933 MT/s. (12 x 64 GB DDR4) • 2x Mellanox Connect X-4 LX 25 GbE
Storage	BOSS S1 • 2x 240 GB SSDs per host Dell HBA330 • 6x 1.6TB SSDs per host
Network	S5248-ON Switch
Protocol	BLAST Extreme H.264 + Switch Codec
Broker	VMWare Horizon 7.12
Hypervisor	vSphere ESXi 6.7 U3 (15160138)
SQL	Microsoft SQL Server 2016
Desktop OS	Microsoft Windows 10 Enterprise 64-Bit, 1909 version
OS Patches	KB4552931, KB4513661, KB4516115, KB4517245, KB4528759, KB4537759, KB4552152, KB4556799
Office	Office 2019
Anti-virus	Windows Defender
Management OS	Microsoft Window Server 2016
Login VSI	Login VSI 4.1.40.1
PowerFlex	R3_5.0.436
NVIDIA GPU	5 x NVIDIA T4 GPU cards installed in one host
NVIDIA GRID Virtual PC Software version	11.1

The following table summarizes the standard desktop VM configurations currently used for the various profiles/workloads tested:

Table 10 Standard desktop VM configurations

Workload	VM Profiles				
	vCPUs	**RAM**	**RAM Reserved**	**Desktop Video Resolution**	**Operating System**
Task Worker	2	3 GB	1.5 GB	1280 x 720	Windows 10 Enterprise 64-bit
Knowledge Worker	2	4 GB	2 GB	1920 x 1080	Windows 10 Enterprise 64-bit
Power Worker	4	8 GB	4 GB	1920 x 1080	Windows 10 Enterprise 64-bit
Multimedia Worker (1B vGPU profile)	4	8 GB	8 GB	1920 x 1080	Windows 10 Enterprise 64-bit

B Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

Storage technical documents and videos provide expertise that helps to ensure customer success on Dell Technologies storage platforms.

B.1 Related resources

The following list of documents and other assets that are referenced in the paper:

- Dell EMC PowerFlex Product Overview
- PowerFlex specification sheet
- VMware vSphere
- VMware Horizon
- NVIDIA Virtual GPU Software Quick Start Guide
- Login VSI Overview
- VMware OS Optimization Tool