Dell EMC PowerStore: 3,600 VMware Horizon VDI Users

February 2022

H18239.1

Reference Architecture

Abstract

This document describes an architecture for deploying 3,600 VMware Horizon virtual desktops with Dell EMC PowerStore.

Dell Technologies

D⊄LL Technologies

Contents

Executive summary

Overview

This document details the architecture design, configuration, and implementation considerations when deploying a VDI workload with 3,600 VMware Horizon virtual desktops on the Dell EMC PowerStore 9000T model storage array. This document is primarily storage-focused and details the specific configurations of this system. Also, the network and fabrics that are used for connectivity is explored in detail.

PowerStore layers protocols and services on a flexible, container-based architecture, ensuring that features and services can be added to existing installations with minimal impact. This paper discusses the benefits of a unified storage product that combines block and file services into a virtual desktop infrastructure (VDI) platform. PowerStore allows expansion of both functionality and capacity to protect storage investments.

The information in this document is based on the guidelines in *Dell EMC PowerStore: VMware vSphere Best Practices*. This document includes supplemental information regarding VDI-specific details when running VMware Horizon.

As part of this reference architecture, the solution was tested using the industry-standard Login VSI VDI benchmarking tool. The PowerStore 9000T model appliance delivered sub-millisecond latency to 3,600 virtual desktops running a heavy workload of 30 IOPS per user. This would meet the performance requirements of most organizations.

Audience

This document is intended for IT administrators, storage architects, partners, and Dell Technologies employees. This audience also includes any individuals who may evaluate, acquire, manage, operate, or design a Dell EMC networked storage environment using PowerStore systems.

Revisions

Date	Description
April 2020	Initial release: PowerStoreOS 1.0
February 2022	Template update

We value your feedback

Dell Technologies and the authors of this document welcome your feedback on this document. Contact the Dell Technologies team by email.

Author: Damon Zaylskie

Note: For links to other documentation for this topic, see the PowerStore Info Hub.

Introduction

PowerStore overview

PowerStore storage arrays are designed to enable growth, flexibility, and resiliency. PowerStore services are hardware and protocol independent, allowing file services to be layered without restricting block-storage flexibility.

Dell EMC storage products are integrated into a mature monitoring and reporting suite that is designed to minimize management effort. With robust reporting and notification services, storage management can be as minimal as needed. The integrated dashboards also enable at-a-glance insight into the health and performance of the platform.

PowerStore achieves new levels of operational simplicity and agility. It uses a container-based microservices architecture, advanced storage technologies, and integrated machine learning to unlock the power of your data. PowerStore is a versatile platform with a performance-centric design that delivers multidimensional scale, always-on data reduction, and support for next-generation media.

PowerStore brings the simplicity of public cloud to on-premises infrastructure, streamlining operations with an integrated machine-learning engine and seamless automation. It also offers predictive analytics to easily monitor, analyze, and troubleshoot the environment. PowerStore is highly adaptable, providing the flexibility to host specialized workloads directly on the appliance and modernize infrastructure without disruption. It also offers investment protection through flexible payment solutions and data-in-place upgrades.

With the addition of integrated file services, there is no requirement to purchase an external storage device for file shares. User access through NFS or Windows SMB is integrated, with the ability to share files between Microsoft Windows and UNIX or Linux users. This ability simplifies the control with a single interface for data management.

PowerStore can serve data using iSCSI, Fibre Channel, or both protocols simultaneously to provide seamless integration into most environments.

VMware Horizon overview

VMware Horizon is an industry-standard solution for virtualizing desktops. It provides advanced control and great flexibility for even the largest VDI environments.

Prerequisite reading

The best practices in this document require knowledge from the following resources:

- *Dell EMC PowerStore Host Configuration Guide*
- VMware Horizon overview

Terminology

The following table provides definitions for some of the terms that are used in this document.

Table 1.   Terminology

Term	Definition
PowerStore Manager	The web-based user interface (UI) for storage management.
Appliance	A solution containing a base enclosure and attached expansion enclosures. The size of an appliance could include only the base enclosure or the base enclosure plus expansion enclosures.
Node	The component within the base enclosure that contains processors and memory. Each appliance consists of two nodes.
Cluster	One or more PowerStore appliances in a single grouping. Clusters are expandable by adding more appliances (up to four).
Base enclosure	The enclosure containing both nodes (node A and node B) and the NVMe drive slots.
Expansion enclosure	An enclosure that can be attached to a base enclosure to provide 25 additional SAS-based drive slots.
Storage volumes	PowerStore volumes that use block storage. These volumes are displayed in the PowerStore dashboard > Block area.
NAS server	Network-attached storage (NAS) servers allowing block storage to be used for file storage. The NAS service translates file requests to block storage. The NAS can be both Microsoft Windows SMB 3.1.1 or NFS v3/v4.
File system	A system allowing PowerStore to enable file-level access to end users and applications. It also controls user permissions to files, and folders or directories.

Connectivity

Introduction   PowerStore supports both iSCSI and Fibre Channel hosts. While both protocols can be run simultaneously, this configuration can add complexity to routing and failover scenarios. It is recommended to use only one protocol per host.

For this configuration iSCSI was chosen to enable an Ethernet-only environment. The network will be configured as a converged environment to reduce complexity and cost. The connectivity modules used were four port 25 Gb/second Ethernet modules for both the NAS and iSCSI block storage. This provides more than enough bandwidth for the user load. To isolate and allow traffic prioritization, VLANs are configured for each traffic type.

Block storage   PowerStore presents storage to external hosts through either block or file interfaces. Block storage is the most commonly used datastore path for virtual machines due to the various speeds and protocols that are offered, making it ideal for performance. NFS is also used but is less common. PowerStore supports 25 GbE and 32 Gb FC connections for broad compatibility and performance requirements. Instructions and best practices for configuring hosts are in the *Dell EMC PowerStore Host Configuration Guide* on the PowerStore Info Hub.

Multipathing is critical to achieving the best performance from VMware hosts and should be configured properly. For configuration guidance, see the document *Dell EMC PowerStore: VMware vSphere Best Practices* on Dell.com/StorageResources. The large number of small I/Os generated from VDI environments benefit greatly from round robin and rapid path switching. Consider using and perform testing with a low I/O-per-path change, which can help improve VM performance.

The MPIO configuration used was Round Robin as shown:

- Path selection policy: Round Robin (VMware)

- Storage array type policy: VMW_SATP_ALUA

This provides optimal pathing with failover protection and path discovery.

iSCSI

Using iSCSI for block storage allows convergence of storage and networking infrastructure though it does require careful planning for bandwidth requirements and fault isolation. Redundant networks are preferred for data availability.

With iSCSI, Jumbo frames are also highly recommended. They allow greater packet efficiency for higher bandwidth. Using Jumbo frames is not a requirement but should be considered. Combined with the 25 GbE connectivity available in PowerStore, iSCSI can handle demanding workloads and high throughput requirements. All devices in this configuration were configured for Jumbo frames. This reduces network switch loads as a smaller number of large packets are processed.

Another benefit of using the Ethernet infrastructure for iSCSI is the ability to share the network with file services. This use requires planning to ensure that there is enough bandwidth for all services, but it also reduces the hardware requirements. The use of different VLANs also allows traffic prioritization to ensure that block traffic has the highest priority.

The iSCSI traffic was configured for 9000-byte Jumbo frames over 8 x 25Gb Ethernet ports. The ports were configured in pairs across two I/O modules per node for complete redundancy and throughput. This also spreads the load across more I/O modules to improve packet performance.

Networking   The configuration that is tested is an Ethernet-only solution. The connections are 25 Gb Ethernet, with trunking for management redundancy and performance. The NAS servers used dual 25 Gb ports per node for an aggregate bandwidth of 100 Gb/sec.
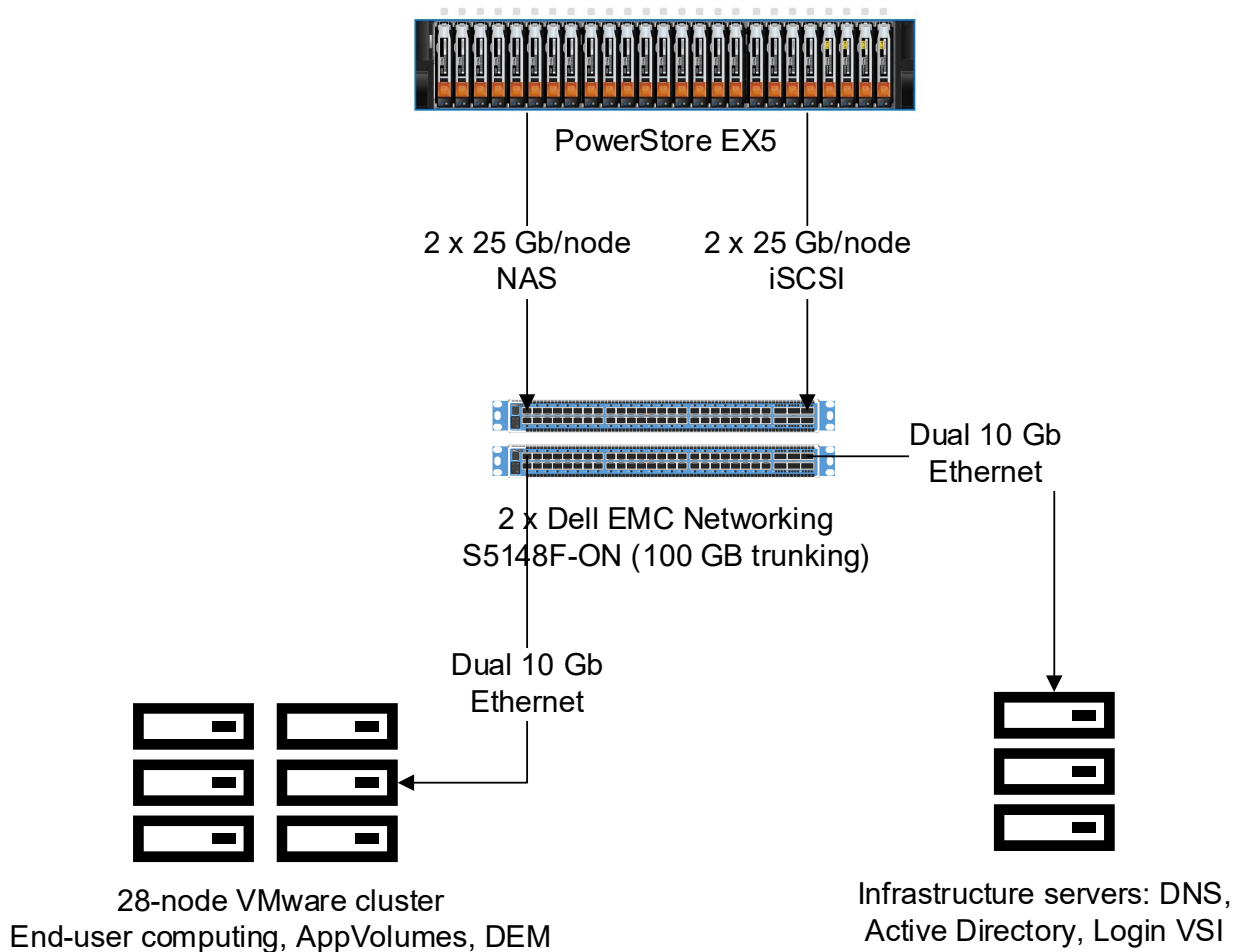
Figure 1.   Logical network

NAS servers

A key feature of PowerStore is the ability to support multiple NAS servers on the same array. This support has multiple benefits, including the following:

- NAS servers can enable multi-tenancy.

- NAS servers are logically separated from each other.

- Clients of one NAS server do not have access to data on other NAS servers.

    Each NAS server has its own independent configuration (for example, DNS, LDAP, NIS, interfaces, protocols, and so forth).

With each NAS instance being isolated, multiple environments can share a single system. Enforced isolation helps increase security and stability. Because a NAS instance runs on a single node to get maximum performance, a minimum of two NAS servers is recommended. The load can then be split between the two instances. The type of data being stored determines the best method of balancing the load.

For user profiles, a tool like VMware DEM (Dynamic Environment Manager) can be used to balance the data across multiple file systems. This tool allows setting dynamic rules to place data based on user-defined criteria. This use automates data redirection for greater control and centralized management.

There are two NAS servers that are configured, one per node in the appliance. This was for performance reasons. Each NAS server is an isolated process. Each NAS server had a single file system created. Multiple file systems can be created per NAS instance if desired.

The use of DEM allows condition-based file redirection. With condition-based rules, the data placement of the user profiles is configured for odd-numbered (such as user1, user3) user profiles to one share with even based to the second file share. This is not a requirement but a guideline. In the context of this testing, it was done to ensure load balancing across the storage nodes.
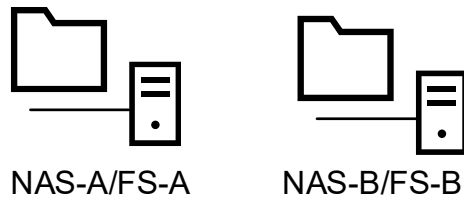

NAS-A/FS-A   NAS-B/FS-B

Figure 2.   NAS/file systems

File systems

PowerStore arrays support multiple file systems for flexibility, security, and availability. This support allows for creating file shares that are based on user type or operating environment. A single NAS and file-system instance can support both SMB and NFS simultaneously.

For this testing, two NAS servers were created. Each NAS server had a single file share that is created and used in a Round-Robin file sharing configuration. This load balanced the file activity and reduced the number of files per share for backup purposes.

In this environment, SMB was the only protocol tested. The scenario is designed to simulate Windows desktop users accessing files from a Windows server.

Storage design

Introduction

The size and type of traffic that is generated from a VDI environment are noticeably different than most business applications. The difference in the size and pattern of VDI traffic is fairly consistent but requires a few changes to the storage design for proper management.

Storage array

The PowerStore 9000T model is the most powerful array in the PowerStore platform. It can provide excellent performance and functionality with advanced capabilities. With a breadth of features and monitoring it can serve the needs of a broad spectrum of customers.

The system used was configured as shown here. There are many options for connectivity to almost any environment.

PowerStore 9000T model:

- 25 flash-based drives (4 NVMe NVRAM plus 21 NVMe SSD)

- 2 x 25 Gb Ethernet ports per node for iSCSI

- 2 x 25 Gb Ethernet ports per node for NAS

- Dual controller nodes for fault tolerance and availability

- Automated RAID for ease of use

Management and iSCSI ports   NAS ports

Figure 3.   PowerStore 9000T model hardware

VMware best practices

Most VMware best practices apply to VDI environments, but there are a few changes required due to the nature of high-density VM configurations. For VMware best practices, see the document *Dell EMC PowerStore: VMware vSphere Best Practices.*

With most applications, storage traffic is typically 70% reads. There are applications that generate all reads or all writes, but the average tends to be 70/30 R/W. VDI is typically 70-80% writes. The array-sizing changes are based on the higher write ratio.

For best practices for Horizon on PowerStore, see the document *Dell EMC PowerStore: VMware Horizon Best Practices* guide on Dell.com/StorageResources.

Volume count

There are several factors that determine the optimal volume count with Horizon on PowerStore. There is not a single volume count that fits all scenarios. Performance and management overhead are typically the two most important factors.

For performance reasons, a minimum of 16 volumes is recommended. This configuration ensures good queue balancing and reduces I/O bottlenecks.

This volume recommendation is offset by the need to keep the volume count low for minimized management. The VMware guidance for volume count should also be considered. Internal testing shows that 16 volumes can support 3,600 users from a performance perspective but may have an impact on VM recovery from a backup. Using array-based snapshots reduces recovery time and minimizes the impact of higher user-per-volume counts.

The testing for this configuration was done with 16 volumes for VM storage and 6 volumes for application delivery. The Adobe Reader application was made available using VMware AppVolumes to demonstrate layered application delivery. This also provides application I/O load balancing using a round-robin storage group in AppVolumes.

I/O size

The average I/O size of storage traffic in a VDI environment typically ranges from 24k to 32k. This size can vary if host-based caching is enabled or allowed. Some VM configurations do not allow host-based caching.

The average I/O size from Horizon VMs during the workload test was 18k at the array level.

Heavy writes

VDI environments are typically heavily biased towards writes. This bias occurs because once the VMs are booted, the I/O traffic includes mostly changes from each VM. The traffic consists of file changes, swap file writes, memory paging, and updates to user preferences. This traffic pattern makes VDI one of the more demanding applications.

During the steady-state portion of the testing, the write ratio was approximately 70%. This is a higher write ratio than most other workloads that a storage array would support.

Instant clone pools

Instant clones generate a large volume of traffic during the provisioning process due to the method used to create the VMs. Since the in-memory VMFork technology creates the machines quickly, the volume of traffic is significant. The I/O required to complete the creation is brief but large. Plan for bursts of high traffic when creating or refreshing instant clone machines.

Since instant clone machines reset on logout, as users log off, their virtual desktops refresh. This occurs throughout the workday and should be considered in the design.

For this design, three Instant Clone pools were used, with a dedicated Horizon Connection server per pool. This is to ensure load balancing across the connection servers. A randomization was applied to the login order to help spread the login load evenly across the connection servers to ensure best performance and behave more like a production scenario.

Capacity

The density of data on VDI tends to be very high because many Horizon VDI environments use data-reduced clones. With profile-redirection technology like VMware Dynamic Environment Manager (DEM), clone pools appear as persistent to users. DEM enables users to personalize their desktop, but pools can still employ the benefits of

nonpersistent VMs. Keeping the VMs small and clean improves performance and recovery time.

The system was configured with 16 2TB datastores for virtual desktops and 6x200GB volumes for application storage. Since any volume can be expanded at any time, these are only starting points. All volumes are also thin provisioned and have space reclaim enabled. This maximizes storage efficiency.

Virtual machine types

In a virtual desktop environment for end users, Horizon supports many different types of virtual workspaces. They tend to be grouped together under the term Virtual Desktop Infrastructure (VDI). However, the various types of virtual desktops have very different characteristics.

Full clones

Full clones were the first type of clone used. The use of full clones requires no special tools. It is as simple as cloning a VM in vCenter. If the clone is a Windows VM it requires a Sysprep first, or changing the name and IP address before joining a domain. If Horizon is used, it creates unique clones.

This is also the largest clone type and requires individual updating. Each VM is a unique machine that must be managed.

PowerStore can perform data reduction to reduce the disk footprint, but the VM still needs regular updating. During Windows Update cycles, each VM must perform a full update.

The footprint of full clones is the largest before data reduction. There are no shared pages between machines.

Linked clones

Linked clones are the most common type of clone that is used by VMware Horizon. They can be the least I/O intensive. Persistent linked clones overall generate the least I/O. The small number of base images relative to the pool size reduces the number of disk pages accessed.

The update process is less impactful as only the base image runs Windows update. Then the VMs can be recomposed and the entire pool is updated to a new base image snapshot. This allows VMs to be updated very quickly and with minimal administrative input.

Going forward, linked clones do not have support in cloud environments. This prevents users from doing hybrid cloud deployments using this technology.

Instant clones

Instant clones are the fastest-growing clone type that is used with Horizon. The improvements in instant clone technology have enabled them to replace other clone type in most configurations. They are the most I/O intensive of the clone type. Each VM refreshes on every user logout. This has trade-offs.

The big advantage of instant clones is each machine is completely refreshed on every user logout to the base image. The update process can leverage this and update each

VM to a new base image on refresh. This allows frequent updating of VMs without user interruption. The virtual machines deploy with the fresh base image on user logout.

The caveat is the load that is created while this process is happening. A correctly sized environment does not have an issue, but the CPU and disk requirements must be accounted for. The architecture in this document accounts for the extra load.

With the advantages of instant-clone deployment and updating, the time spent managing the environment is the lowest of the technologies. The end-user impact can also be one of the lowest.

Windows I/O patterns

The evolution of Windows has brought new features to users but has created additional load for systems hosting it. From Windows 7 to Windows 10, the load increased dramatically. A larger portion of the increase was due to added services. There are other encryption and memory-protection schemes that are built into Windows. These serve to reduce the attack surface but also generate additional load.

As Windows 10 has evolved more security and features have been added that have increased the load even further. From the introduction to build 1909 the I/O generated by Windows 10 has doubled. This constant increase in I/O load requires matching increases storage performance.

The I/O patterns of VDI are a shift from the average read/write ratio of 70/30. The typical I/O pattern of VDI is 70-80% writes while users are logged in. Where the patterns differ between machine types is during the logout phase.

Instant clones

Instant clones perform a shutdown and refresh on every logout. This results in a new clone for each VM. This requires user data redirection to provide a persistent user experience. This VM type has the highest I/O load with the refresh cycles. It is very predictable as the load is directly proportional to user-log-out rates.

Linked clones

Linked clones use a common base image with a disk digest architecture to redirect changes. All changes to the VM are written to the digest. This can be a persistent architecture. If persistence is required, all changes are saved. Otherwise, all changes are flushed on logout. The disk load is higher in a non-persistent configuration as a new user profile gets created on each login.

The disk footprint of linked clones is very small. The base image is copied to each datastore in the pool and all machines share the base images. This greatly reduces duplication of the base image.

Full clones

The I/O profile of full clones is that of full machines, with reads and writes spread across the entire dataset. Full clones typically save user data to the local profile. This results in a persistent virtual experience. This VM type typically has the lowest overall storage impact since data is saved between sessions. This is offset somewhat by the update requirements of full clones.

Since each VM is unique, it requires updates to be applied to each instance. The impact of thousands of machines updating can be significant and must be planned very carefully. Maintenance windows should be defined and groups of machines scheduled. The impact of each round of updates varies based on the updates that are released. This makes it difficult to predict the impact.

Guests

Operating system

Windows 10 build 1809 was used for this test. The processor requirements for Windows 10 change based on the build number and patch level. The image was optimized using the VMware OS Optimization Tool. There were two processor cores and 3 GB of RAM assigned to each VM.

One of the best ways to reduce CPU and disk load from a Windows virtual machine is to use the VMware OS Optimization Tool. This tool is a consolidated interface that can be used to change the behavior of virtual machines to reduce their hardware requirements. Be careful if choosing the most-restrictive configurations, which can cause applications or even Windows features to stop working. This tool is designed to reduce the effort of optimizing Windows, but it requires testing of the configuration for all user applications.

User profiles

Another tool that was mentioned previously in this paper is VMware DEM. This tool redirects user data to one or more file shares that are based on certain conditions. This capability allows directing different types of data to different locations. User-profile data and documents can be pointed to home directories that are replicated with the temporary data of applications that are placed on volumes with no data protection. This approach reduces replication traffic and the overall storage load.

All user profiles were redirected to the unified file shares to capture all user load. The ability of the PowerStore array to host file shares consolidates management and reduces backup complexity. A single storage appliance can host all end-user data.

Applications

The applications can be layered using VMware AppVolumes. This greatly improves application management and updating by removing the applications from the base image and managing them independently. For this architecture Adobe Reader was presented using AppVolumes. When an update is required the layered application is updated externally and the updated version is applied to the environment.

Layering applications enables rapid application provision and updating. The individual applications are managed independently with the ability to present groups of applications to user groups. This greatly speeds application deployments to new users or additional applications to existing users.

RAID configurations and storage containers

PowerStore automatically manages the underlying storage for maximum performance and capacity, eliminating the need for administrators to configure RAID or the storage pool. Manually setting or configuring these options is unnecessary in PowerStore. The underlying RAID configuration and drive management are automatic and require no management.

Horizon configuration

Horizon does not require any special configuration to work with PowerStore. The storage is presented as block storage volumes discoverable through vCenter. When the pool is created, the volumes are visible as usable disk.

The application volumes are used in a round robin fashion to provide the most consistent performance as well. With AppVolumes, distributing access across all six volumes makes the access time consistent for all users.

vCenter provisioning settings

Based on performance testing under load, the following settings are used:

Advanced Settings

Specify the concurrent operation limits.

Max concurrent vCenter provisioning operations:	30
Max concurrent power operations:	30
Max concurrent View Composer maintenance operations:	20
Max concurrent View Composer provisioning operations:	16
Max concurrent Instant Clone Engine provisioning operations:	30

Figure 4.   Horizon server settings for provisioning

The numbers were derived from latency under load and the impact of instant clone provisioning. Under full load, the system performed well even with user logouts and VM cloning.

Secondary pools

One simple way to help users quickly return to work is to keep additional VMs ready in a recovery pool. This pool can be assigned to end users if the primary systems are down. This pool may not have the same performance characteristics but can aid in business continuity. Because of the rapid deployment nature of virtual clone VMs, it is also possible to have the VMs created on demand in a failover scenario. This reduces the storage requirement. Since only one environment would be generating load, the IOP requirement would be the same once all the users are logged in. The load would be higher during virtual machine creation.

Data encryption

Many applications have data-encryption requirements, specifically on data at rest. Data at Rest Encryption (D@RE) can be used as an encryption solution for VDI without requiring VM changes. This ability also avoids potential performance impact to the applications and has no performance impact on the array.

D@RE is enabled by default on the PowerStore array, so no configuration steps are necessary to protect the drives.

Test results

Login VSI

Login phase

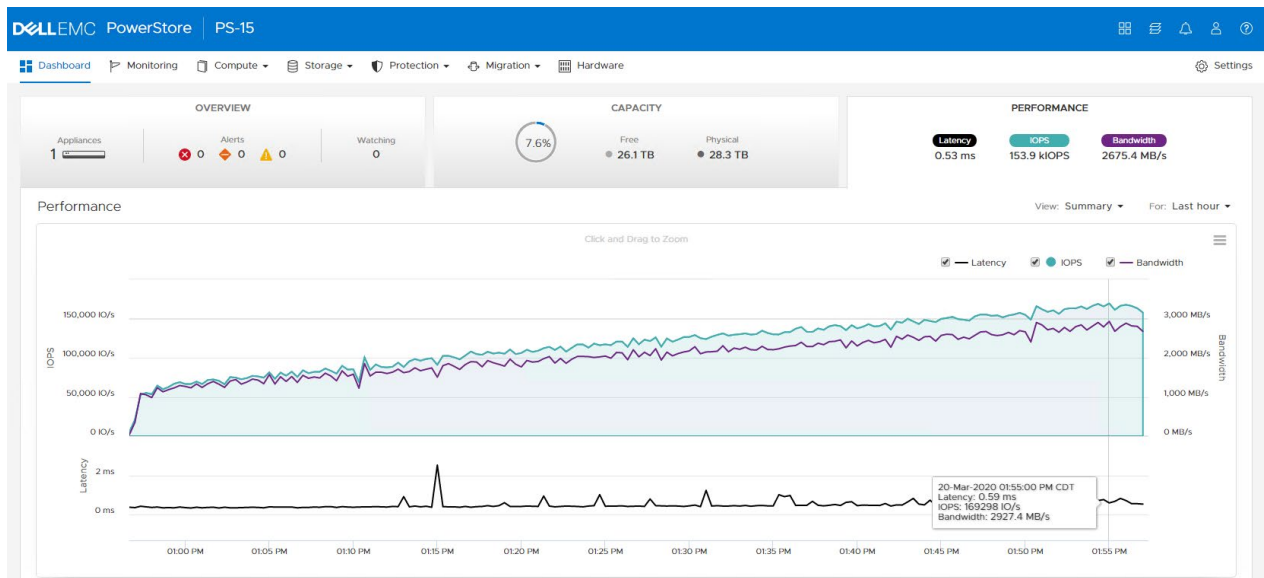The following figure shows the login phase.

Figure 5.   Workload ramp-up during login phase

Steady-state workload

Once all the users have logged in and are performing their daily tasks, the load is still over 110,000 IOPS at less than 0.5 ms of system latency. The following figure shows the steady-state workload.
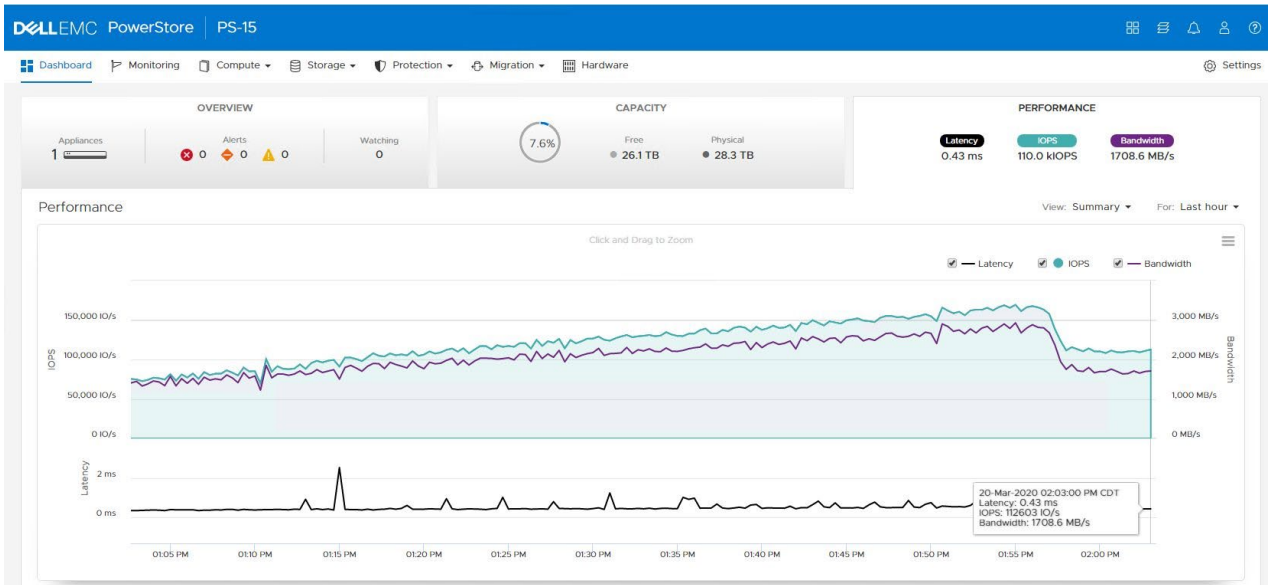
Figure 6.   Steady-state performance

The Login VSI test results, displayed in the following figure, show how well the PowerStore system performs under load.
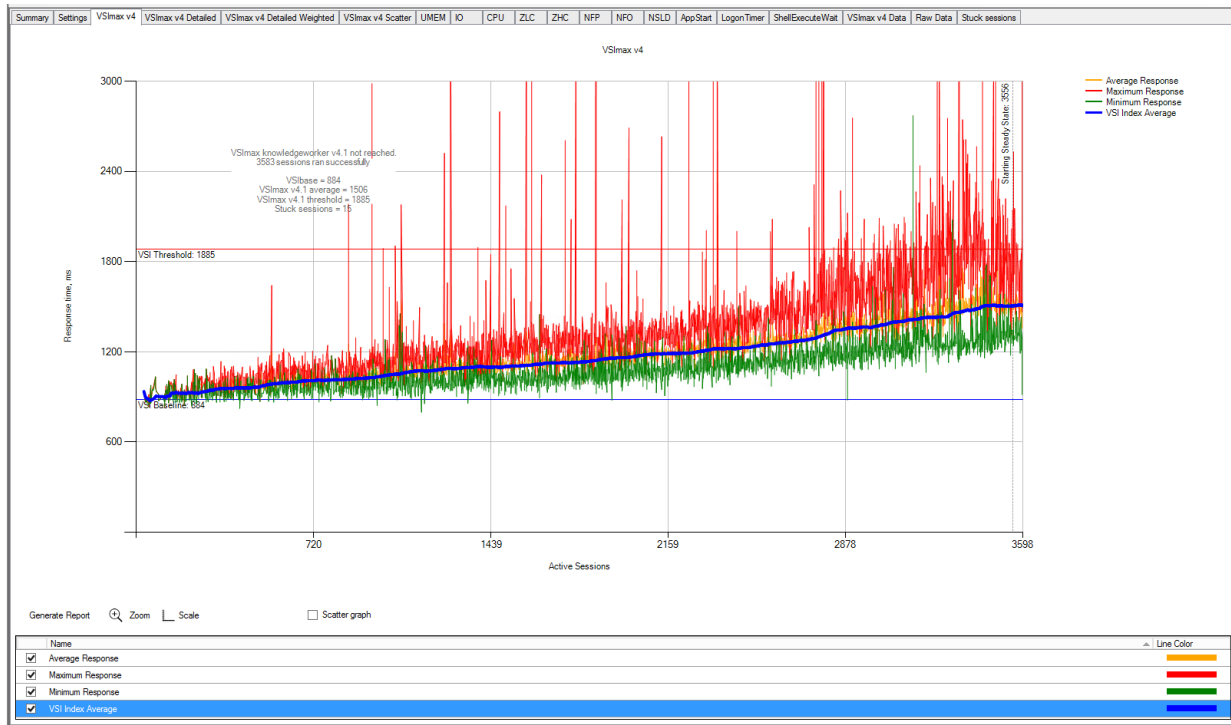

Figure 7.   Login VSI test results

Summary

The test results show that the PowerStore 9000T model array is extremely capable and performs well under heavy loads. The low latency and predictable throughput make it an excellent platform for VMware Horizon workloads.

Along with the excellent performance and predictability, PowerStore has strong monitoring and alerting capabilities. These are discussed in more detail in the document *Dell EMC PowerStore: VMware Horizon VDI Best Practices* on Dell.com/StorageResources. In summary, the PowerStore platform brings peace of mind with its solid performance and alerting.

Appendix: 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 EMC storage platforms.

For more information, see the following related resources:

- VMware Documentation
- Dell Technologies VDI info hub