Technical Report

NetApp All Flash FAS Solution for Persistent Desktops with Citrix XenDesktop

Chris Gebhardt, Bhavik Desai, Brian Casper, Joe Scott, NetApp

September 2016 | TR-4519

TABLE OF CONTENTS

1 Executive Summary 6

2 Reference Architecture Objectives 7

3 Introduction 7

3.1 Document Overview 7

3.2 NetApp All Flash FAS 8

3.3 ONTAP FlashEssentials 8

3.4 NetApp ONTAP 9 9

3.5 Storage Efficiency 9

3.6 Performance Capacity 10

3.7 Performance Capacity Terms 11

4 Solution Infrastructure 12

4.1 Hardware Infrastructure 13

4.2 Software Components 14

4.3 Login VSI 15

4.4 Full-Clone Virtual Desktops 15

5 Storage Design 16

5.1 Aggregate Layout 17

5.2 Volume Layout 17

5.3 NetApp Virtual Storage Console for VMware vSphere 17

5.4 Importing VSC-Created VMs into XenDesktop 18

6 Testing and Validation: Persistent Full-Clone Desktops 18

6.1 Test Results Overview 18

6.2 Storage Efficiency 19

6.3 Boot Storm Test 20

6.4 Boot Storm Test During Storage Failover 22

6.5 Steady-State Login VSI Test 25

6.6 Test for Patching 1,500 Desktops 39

7 Eight-Hour Test to Determine Performance Capacity 42

7.1 Eight-Hour Test System Utilization Graph 42

7.2 Eight-Hour Test Performance Capacity Used 43

7.3 Eight-Hour Test Performance Capacity Used Estimated Takeover 43

7.4 Eight-Hour Test Latency 44

Conclusion 44
Appendix: Workload Profiles 45
Boot Storm 45
Monday Morning Login and Workload 46
Tuesday Morning Login and Workload 46
Patching Desktops 47
References 48
Acknowledgements 48
Version History 48

LIST OF TABLES

Table 1) Test results for 1,500 persistent desktops. 7
Table 2) Hardware components of server categories. 13
Table 3) Solution software components. 14
Table 4) Virtual desktop configuration. 15
Table 5) Test results overview. 18
Table 6) Efficiency results for each FlexVol volume. 20
Table 7) Results for full-clone boot storm. 20
Table 8) Results for full-clone boot storm during storage failover. 22
Table 9) Results for full-clone Monday morning login and workload. 25
Table 10) Results for full-clone Monday morning login and workload during storage failover. 29
Table 11) Results for full-clone Tuesday morning login and workload. 32
Table 12) Results for full-clone Tuesday morning login and workload during storage failover. 36
Table 13) Results for patching 1,500 persistent full clones on one node. 39

LIST OF FIGURES

Figure 1) Visual representation of inline compression and data compaction. 10
Figure 2) Advanced drive partitioning v2. 10
Figure 3) Performance capacity graphic. 11
Figure 4) Solution infrastructure for persistent desktops. 13
Figure 5) VMware OS optimization tool. 16
Figure 6) SSD layout. 17
Figure 7) Volume layout for nonpersistent desktops. 17
Figure 8) Storage efficiency savings. 19
Figure 9) Throughput and IOPS for full-clone boot storm. 20
Figure 10) Storage controller CPU utilization for full-clone boot storm. 21

Figure 11) Read/write IOPS for full-clone boot storm. ... 21
Figure 12) Read/write ratio for full-clone boot storm. ... 22
Figure 13) Throughput, IOPS, and latency for full-clone boot storm during storage failover. ... 23
Figure 14) Storage controller CPU utilization for full-clone boot storm during storage failover. ... 23
Figure 15) Read/write IOPS for full-clone boot storm during storage failover. ... 24
Figure 16) Read/write ratio for full-clone boot storm during storage failover. ... 24
Figure 17) VSImax results for full-clone Monday morning login and workload. ... 26
Figure 18) Scatterplot of full-clone Monday morning login times. ... 26
Figure 19) Throughput, IOPS, and latency for full-clone Monday morning login and workload. ... 27
Figure 20) Storage controller CPU utilization for full-clone Monday morning login and workload. ... 27
Figure 21) Read/write IOPS for full-clone Monday morning login and workload. ... 28
Figure 22) Read/write ratio for full-clone Monday morning login and workload. ... 28
Figure 23) VSImax results for full-clone Monday morning login and workload during storage failover. ... 29
Figure 24) Scatterplot of full-clone Monday morning login times during storage failover. ... 30
Figure 25) Throughput, IOPS, and latency for full-clone Monday morning login and workload during storage failover. 30
Figure 26) Storage controller CPU utilization for full-clone Monday morning login and workload during storage failover. ... 31
Figure 27) Read/write IOPS for full-clone Monday morning login and workload during storage failover. ... 31
Figure 28) Read/write ratio for full-clone Monday morning login and workload during storage failover. ... 32
Figure 29) VSImax results for full-clone Tuesday morning login and workload. ... 33
Figure 30) Scatterplot of full-clone Tuesday morning login times. ... 33
Figure 31) Throughput, IOPS, and latency for full-clone Tuesday morning login and workload. ... 34
Figure 32) Storage controller CPU utilization for full-clone Tuesday morning login and workload. ... 34
Figure 33) Read/write IOPS for full-clone Tuesday morning login and workload. ... 35
Figure 34) Read/write ratio for full-clone Tuesday morning login and workload. ... 35
Figure 35) VSImax results for full-clone Tuesday morning login and workload during storage failover. ... 36
Figure 36) Scatterplot of full-clone Tuesday morning login times during storage failover. ... 37
Figure 37) Throughput, IOPS, and latency for full-clone Tuesday morning login and workload during storage failover. 37
Figure 38) Storage controller CPU utilization for full-clone Tuesday morning login and workload during storage failover. ... 38
Figure 39) Read/write IOPS for full-clone Tuesday morning login and workload during storage failover. ... 38
Figure 40) Read/write ratio for full-clone Tuesday morning login and workload during storage failover. ... 39
Figure 41) Throughput, IOPS, and latency for patching 1,500 persistent full clones. ... 40
Figure 42) Storage controller CPU utilization for patching 1,500 persistent full clones. ... 40
Figure 43) Read/write IOPS for patching 1,500 persistent full clones. ... 41
Figure 44) Read/write ratio for patching 1,500 persistent full clones. ... 41
Figure 45) Savings between Data ONTAP 8.3.2 and ONTAP 9 (in GB) per FlexVol volume. ... 42
Figure 46) Eight-hour test system utilization graph. ... 43
Figure 47) Eight-hour test performance capacity used graph. ... 43
Figure 48) Eight-hour performance capacity used and estimated takeover graph. ... 44

Figure 49) Eight-hour test latency and estimated latency during takeover. 44

1 Executive Summary

The decision to virtualize desktops affects multiple aspects of an IT organization, including infrastructure and storage requirements, application delivery, end-user devices, and technical support. In addition, correctly architecting, deploying, and managing a virtual desktop infrastructure (VDI) can be challenging because of the large number of solution components in the architecture. Therefore, it is critical to build the solution on industry-proven platforms such as NetApp® storage and industry-proven software solutions from Citrix. Citrix and NetApp provide leading desktop virtualization and storage solutions, respectively, for customers to successfully meet these challenges. A desktop virtualization solution provides workspace mobility, centralized management, consolidated and secure delivery of data, and device independence.

The criteria for determining the success of a virtual desktop implementation must include end-user experience. The end-user experience must be as good as or better than any previous experience on a physical PC or virtual desktop.

As a workload, desktop virtualization is highly variable, and it involves cycles of heavy reads, heavy writes, and everything in between, along with varying block sizes for writes. Therefore, it should be no surprise that storage is often the leading culprit of unmet end-user performance expectations. The NetApp All Flash FAS (AFF) solution, with the AFF8000 platform, easily handles highly variable desktop virtualization workloads to deliver an end-user experience that exceeds expectations.

Another criterion of project success is solution cost. Storage has often been the most expensive part of the virtual desktop solution, especially when storage efficiency and flash acceleration technologies were lacking. It was also common practice to forgo an assessment. Skipping this critical step meant that companies often overbought or undersized the storage infrastructure because information is the key to making sound architectural decisions that result in wise IT spending.

NetApp has many technologies that help customers reduce the storage cost of a virtual desktop solution. Technologies such as inline deduplication, inline compression, data compaction, advanced drive partitioning, and thin provisioning help reduce the total amount of storage required for VDI. Storage platforms that scale up and scale out with NetApp ONTAP® software (ONTAP) deliver the right architecture to meet the customer's price and performance requirements. NetApp helps customers achieve their cost and performance goals while providing rich data management features.

With the NetApp All Flash FAS solution, customers might pay as little as US$39 per desktop for storage when deploying at scale. This figure includes the cost of NetApp hardware, software, and three years of 24/7 premium support with four-hour parts replacement. A similarly low ratio of dollars per desktop can be achieved within the portfolio of All Flash FAS platforms.

With Citrix and NetApp, companies can accelerate the virtual desktop end-user experience by using NetApp All Flash FAS storage for Citrix XenDesktop. NetApp All Flash FAS storage, powered by the AFF8000 system, is the optimal platform for using high-performing solid-state drives (SSDs) without adding risk to desktop virtualization initiatives.

When a storage failure prevents users from working, that inactivity translates into lost revenue and productivity. That is why what used to be considered a tier 3 or 4 application is now critical to business operations. Having a storage system with a robust set of data management and availability features is key to keeping the users working and lessens the risk to the business. ONTAP has multiple built-in features, such as active-active high availability (HA) and nondisruptive operations, to help improve availability and to seamlessly move data in the storage cluster without affecting users.

NetApp also provides the ability to easily increase storage system capacity by simply adding disks or shelves. There is no need to purchase additional controllers in order to add users when additional capacity is required. When the platform requires expansion, additional nodes can be added in a scale-out fashion and managed within the same management framework and interface. Workloads can then be nondisruptively migrated or balanced to the new nodes in the cluster without the users ever noticing.

2 Reference Architecture Objectives

The reference architecture described in this document is a 1,500-desktop design using NetApp AFF8040, Citrix XenDesktop 7.6, and VMware vSphere 6.0. It validates a persistent, full-clone, 1,500-desktop architecture created with the NetApp Virtual Storage Console (VSC) and imported into Citrix XenDesktop. The testing covered common administrative tasks, including boot storms, login storms, patching, and steady-state operations. These tests determined time to complete, storage response, and storage utilization, including performance capacity.

In all tests, end-user login time, guest response time, and maintenance activities performance were excellent. The NetApp All Flash FAS system performed well, averaging less than 50% controller utilization during most operations for both use cases. All test categories demonstrated that, based on the 1,500-user workload and maintenance operations, the AFF8040 system is capable of supporting more users while still being able to fail over in the event of a failure and have consistent latencies under 1ms. At a density of 2,666 VMs on an AFF8040 system with the same I/O profile, storage for VDI can be as low as US$39 per desktop. This figure includes the cost of hardware, software, and three years of 24/7 premium support with four-hour parts replacement. A similarly low ratio of dollars per desktop can be achieved within the broader portfolio of All Flash FAS platforms. Table 1 lists the results of testing for 1,500 persistent desktops.

Table 1) Test results for 1,500 persistent desktops.

Test	Time to Complete	Peak IOPS	Peak Throughput	Average Storage Latency
Boot storm test	~10 min	102,437	3.1GBps	1.4ms
Boot storm test during failover	~16 min	67,769	1.9GBps	2.7ms
Login VSI: Monday morning login and workload	7.8 sec/VM	60,533	724MBps	455µs
Login VSI: Monday morning login and workload during failover	6.95 sec/VM	36,573	593MBps	641µs
Login VSI: Tuesday morning login and workload	7.0 sec/VM	26,237	525MBps	470µs
Login VSI: Tuesday morning login and workload during failover	6.65 sec/VM	26,187	477MBps	669µs
Patching desktops with ~900MB of patches (10-sec delay)	~160 min	68,921	1.6GBps	936µs

Note: During the boot storm test, XenDesktop was used to power on the virtual machines.

3 Introduction

This section provides an overview of the NetApp All Flash FAS solution for Citrix XenDesktop, explains the purpose of this document, and introduces Login VSI.

3.1 Document Overview

This document describes the solution components used in the 1,500-user Citrix XenDesktop deployment on NetApp All Flash FAS reference architecture validation. It covers the hardware and software used in the validation, the configuration of the hardware and software, use cases that were tested, and performance results of the tests completed. During these performance tests, many different scenarios were tested to validate the performance of the storage during the lifecycle of a virtual desktop deployment.

The testing included the following criteria:

- Provisioning of 1,500 Citrix XenDesktop full-clone desktops to high-performing, space-efficient NetApp FlexClone® desktops leveraging the NetApp Virtual Storage Console
- Boot storm test of 1,500 desktops (with and without storage node failover)
- Monday morning login and steady-state workload with Login VSI (with and without storage node failover)
- Tuesday morning login and steady-state workload with Login VSI (with and without storage node failover)
- Patching 1,500 desktops
- Eight-hour test using Login VSI to determine performance capacity

Storage performance and end-user acceptance were the main focus of the testing. If a bottleneck occurred in any component of the infrastructure, it was identified and remediated if possible. Best practices were followed to perform staggered patching and maintenance during nonproduction periods of time. NetApp does not recommend patching all desktops at the same time during production hours.

3.2 NetApp All Flash FAS

NetApp ONTAP has evolved to meet the changing needs of customers and help drive their success. ONTAP provides a rich set of data management features and clustering for scale-out, operational efficiency, storage efficiency, and nondisruptive operations to offer customers one of the most compelling value propositions in the industry. The IT landscape is undergoing a fundamental shift to IT as a service. This model requires a pool of computing, network, and storage resources to serve a range of applications and deliver a range of services. Innovations such as ONTAP are fueling this revolution.

The 1,500 desktops were hosted on a 2-node NetApp All Flash FAS active-active storage system running NetApp ONTAP 9 configured with 24 800GB SSDs.

3.3 ONTAP FlashEssentials

NetApp ONTAP FlashEssentials is the power behind the performance and efficiency of All Flash FAS. ONTAP is a well-known operating system, but it is not widely known that ONTAP, with its WAFL® (Write Anywhere File Layout) file system, is natively optimized for flash media.

ONTAP FlashEssentials encapsulates key features that optimize SSD performance and endurance, including the following:

- NetApp data-reduction technologies, including inline compression, inline deduplication, and inline data compaction, can provide significant space savings. Savings can be further increased by using NetApp Snapshot® and NetApp FlexClone technologies. Studies based on customer deployments have shown that total data-reduction technologies have enabled up to 933x space savings.
- Inline data compaction provides continued innovation beyond compression and deduplication, further increasing storage efficiency.
- Coalesced writes to free blocks maximize performance and flash media longevity.
- Flash-specific read path optimizations enable consistently low latency.
- Parallelized processing handles more requests at once.
- Software-defined access to flash maximizes deployment flexibility.
- New advanced drive partitioning (ADPv2) increases storage efficiency and further increases usable capacity by almost 20%.
- Data fabric readiness enables live workload migration between flash and hard disk drive (HDD) tiers on premises or to the cloud.

- Quality of service (QoS) capability safeguards service-level objectives in multiworkload and multitenant environments.

3.4 NetApp ONTAP 9

ONTAP 9 is a major advance in the industry's leading enterprise data management software. It combines new levels of simplicity and flexibility with powerful capabilities and efficiencies. Customers can integrate the best of next-generation and traditional technologies, incorporating flash, cloud, and software-defined architectures while building a foundation for their data fabric. Plus, new customers and existing Data ONTAP® 8.3 environments can quickly and easily use the rich data services delivered by ONTAP 9.

An essential feature for VDI deployed on shared enterprise storage is the ability to deliver consistent and dependable high performance. High performance must be coupled with nondisruptive operations, high availability, scalability, and storage efficiency. Customers can depend on ONTAP 9 and All Flash FAS to provide these essential elements.

Built on ONTAP unified scale-out architecture, All Flash FAS consistently meets or exceeds the high-performance demands of VDI. It also provides rich data management capabilities, such as integrated data protection and nondisruptive upgrades and data migration. These features allow customers to eliminate performance silos and seamlessly integrate All Flash FAS into a shared infrastructure. ONTAP delivers enhanced inline deduplication and a completely new inline data-compaction capability that significantly reduces the amount of flash storage required, with no impact to system performance. It also provides industry-leading ecosystem integration with database applications that makes administration of databases and storage systems far more efficient than other flash storage solutions on the market.

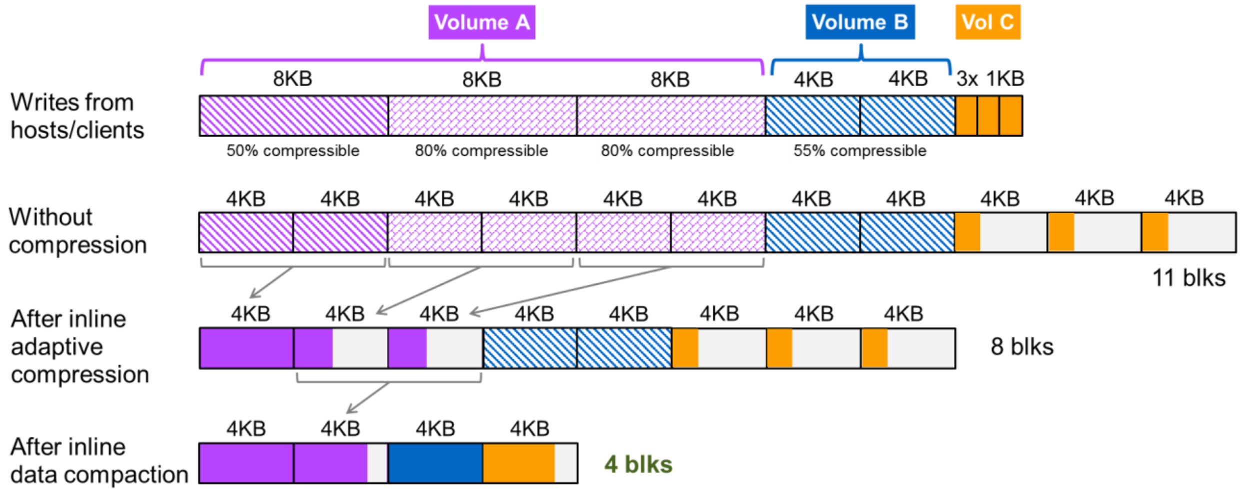
NetApp is a global enterprise scale-out storage and data management fabric provider, and ONTAP has been an industry-leading operating system since 2012. On-site ready but cloud connected, ONTAP is a complete solution that is future-proof in a rapidly changing technology environment.

3.5 Storage Efficiency

Simply stated, storage efficiency enables you to store the maximum amount of data within the smallest possible space at the lowest possible cost. The following NetApp storage efficiency technologies can help you realize maximum space savings:

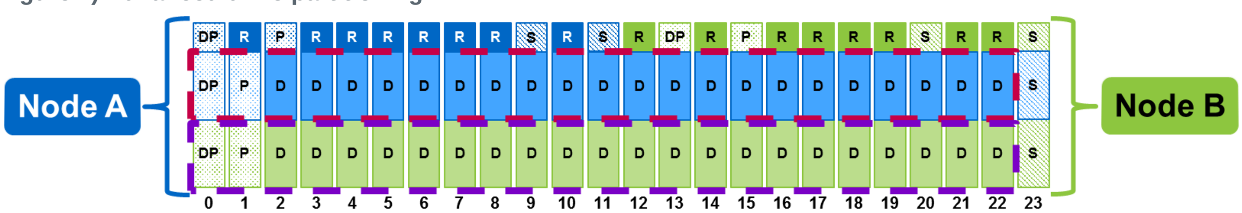
- **Inline compression.** Data compression reduces the disk space required, regardless of storage protocol, application, or storage tier. Inline compression also reduces the data that must be moved to SSDs, thereby reducing the wear on SSDs. Furthermore, the reduced amount of data being written can deliver an overall performance increase.
- **Inline deduplication.** Data deduplication cuts storage requirements by reducing redundancies in primary, backup, and archival data. Inline deduplication of zeros speeds up VM provisioning by 20% to 30%. Further improvements to inline deduplication in ONTAP 9 provide additional efficiency by extending elimination of duplicate data to blocks in memory and SSDs.
- **Inline data compaction.** New in ONTAP 9, NetApp inline data compaction provides significant storage savings by compressing and coalescing small I/O together into single-block writes. This further reduces the disk space required and associated wear on SSDs. Figure 1 demonstrates how compressed data that is smaller than 4K can be stored in a block with other data smaller than 4K.

Figure 1) Visual representation of inline compression and data compaction.



- **Snapshot technology.** NetApp Snapshot technology provides low-cost, instantaneous, point-in-time, space-efficient copies of the file system (volume) or LUN by preserving ONTAP architecture and WAFL consistency points without affecting performance. NetApp SnapManager® software for Oracle automates and simplifies Oracle database management with backup, recovery, restore, and cloning features with no downtime.
- **Thin provisioning.** Thin provisioning, implemented by NetApp at the NetApp FlexVol® volume and LUN level, defers storage purchases by keeping a common pool of free storage available to all applications.
- **Thin replication.** Thin replication is at the center of the NetApp data protection software portfolio, which includes NetApp SnapMirror® and NetApp SnapVault® software. SnapVault thin replication enables more frequent backups that use less storage capacity because no redundant data is moved or stored. SnapMirror thin replication protects business-critical data while minimizing storage capacity requirements.
- **RAID DP.** NetApp RAID DP® technology protects against double disk failure without sacrificing performance or adding disk-mirroring overhead.
- **FlexClone volumes.** FlexClone virtual cloning reduces the need for storage by enabling multiple, instant, space-efficient writable copies.
- **Advanced drive partitioning (v2).** As shown in Figure 2, advanced SSD partitioning with the latest ONTAP 9 release further increases usable capacity by almost 20%.

Figure 2) Advanced drive partitioning v2.



3.6 Performance Capacity

NetApp or partner system engineers should perform initial sizing using NetApp, OS, and application vendor best practices and typically use NetApp internally available tools to determine optimal solution sizing. After initial sizing, NetApp recommends basing all incremental performance sizing, monitoring, capacity planning, and workload placement on available performance capacity. This is a departure from NetApp's previous recommendation, which was to size workloads to use less than 50% CPU utilization. This previous recommendation had the benefit of being easy to make and understand but is far less nuanced and more prone to guesswork than NetApp's current recommendation to use performance capacity planning for sizing.

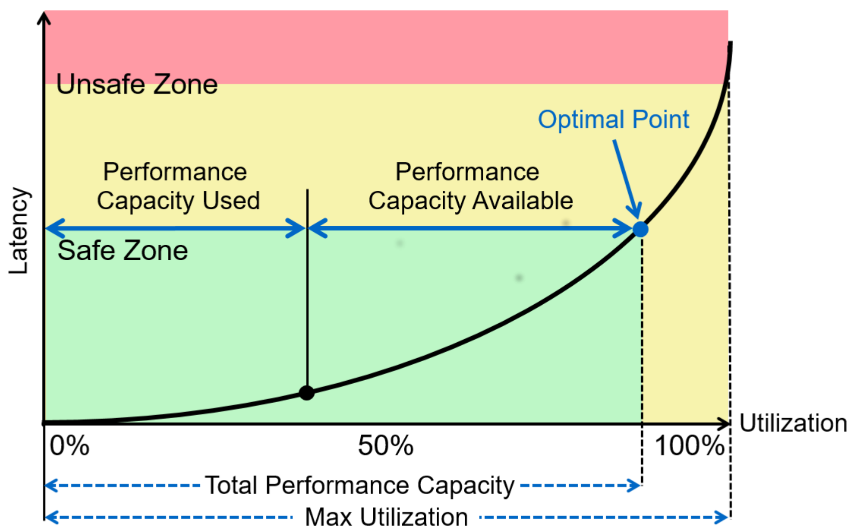
NetApp's best practice for sizing a SAN AFBP environment is to use performance capacity to size each node to less than 50% of performance capacity on each controller in an HA pair. By sizing this way, you can maintain acceptable low latency in the event of a takeover. The cost of this approach is that you sacrifice a little of the steady-state top-level performance.

Before getting started, note that the discussion is based on the assumption of a transactional workload that uses IOPS and latency as principal metrics. With that said, let's define some terms and then finish the discussion with a real-world example of how OnCommand® Performance Monitor (OPM) 7 can be used to determine the total performance capacity and performance capacity available on a given set of controllers.

For more information about performance capacity, see TR-4515: Best Practices for AFF Business-Processing SAN Workloads.

3.7 Performance Capacity Terms

Figure 3) Performance capacity graphic.



The optimal point identifies the total performance capacity a resource has before latency increases more quickly than IOPS do.

The optimal point can be determined by either:

- Finding the "knee" in the performance curve, where an increase in utilization leads to more rapidly increasing latency. Generally, performance curves are fairly flat at lower utilizations, but latency increases as the number of IOPS increases. There is a point in the curve where the rate of increase in latency starts accelerating more rapidly than the increase in the number of IOPS being served.
- An alternative to the knee of the curve method is to target a specific latency value and draw a horizontal line at that latency value. The optimal point is the point where the IOPS curve intersects the latency threshold line you have just drawn.

Total performance capacity (or optimal point) = performance capacity used + performance capacity.

Performance Capacity Used

Performance capacity used can be defined as the amount of the useful capacity of the resource that has been consumed. As noted earlier, the remaining useful capacity is performance capacity available.

Performance capacity used = optimal point – performance capacity available.

Performance Capacity Available

Performance capacity available is derived by subtracting the performance capacity used from the total performance capacity (or simply performance capacity) of a node. The performance capacity is identified by the optimal point. From Figure 3, we see that:

total performance capacity (or optimal point) = performance capacity used + performance capacity available.

Operating Point

The operating point is the location on the curve where your resource is currently operating. This point illustrates the number of IOPS served at a given latency.

Unsafe Zone

The unsafe zone in a performance context can be defined as the portion of the performance graph that is above and to the right of the optimal point. Performance in these areas has higher latencies, and little increases in IOPS yield larger increases in latency.

Safe Zone

The safe zone is the area of the performance graph that is below and to the left of the optimal point. This is the area of the graph where you see the highest throughput relative to latencies and is the area within which you want to operate. In order to maintain consistent low-latency high performance, the operating point needs to stay inside the safe zone.

Differences Between Total Performance Capacity and Utilization

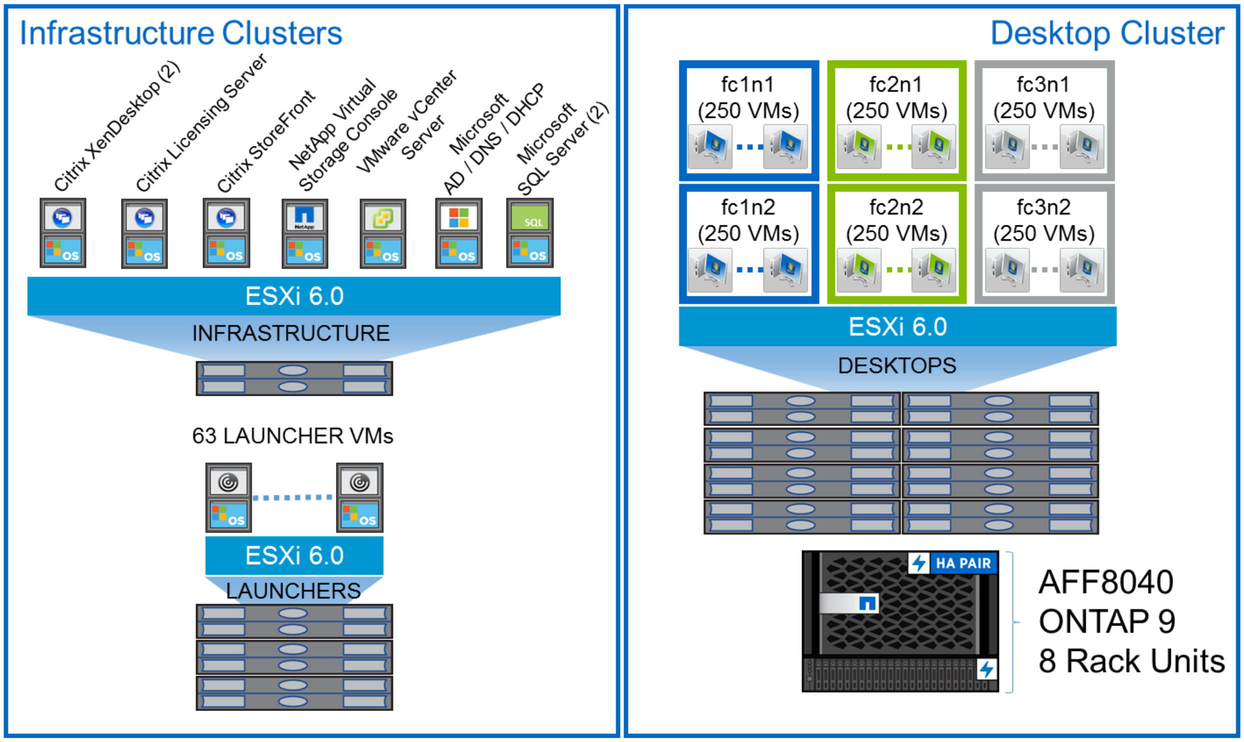
Performance capacity used incorporates both utilization and latency. Performance capacity of up to 100% is ok. However, performance capacity can rise above 100%, at which point a little more output increases latency much more quickly. Utilization as a measure is fine for workloads that are not sensitive to latency such as batch workloads.

4 Solution Infrastructure

This section describes the software and hardware components of the solution. Figure 4 shows the solution infrastructure, which includes two Citrix XenDesktop VMs, two Citrix StoreFront VMs, and one Citrix licensing server.

Scale performance tests, such as boot, login, and steady-state tests, were performed to help validate that this All Flash FAS array is capable of hosting 1,500 users with the standard configuration of two storage controllers.

Figure 4) Solution infrastructure for persistent desktops.



4.1 Hardware Infrastructure

During solution testing, 22 servers were used to host the infrastructure and the desktop VMs. The desktops and infrastructure servers were hosted on discrete resources so that the workload to the NetApp All Flash FAS system could be measured precisely. It is both a NetApp and an industry best practice to separate the desktop VMs from the infrastructure VMs because noisy neighbors or bully virtual desktops can affect the infrastructure and have a negative impact on all users, applications, and performance results.

Options to neutralize this problem include leveraging intelligent quality of service policies in ONTAP to eliminate noisy neighbor behavior, using intelligent sizing to account for infrastructure VMs, or putting infrastructure VMs on an existing or separate NetApp FAS storage system. For this lab validation, we used a separate NetApp FAS storage system (not shown) to host the infrastructure and the launcher. In the real world, infrastructure VMs can be on the same All Flash FAS system that is hosting the virtual desktops. Table 2 lists the hardware specifications of each server category.

Table 2) Hardware components of server categories.

Hardware Components	Configuration
Infrastructure Servers and Launcher Servers	
Server quantity	8 Fujitsu Primergy RX300 S8
CPU model	Intel Xeon CPU E5-2630 v2 at 2.60GHz (6-core)
Total number of cores per server	12 cores (24 with hyperthreading)
Memory per server	64GB–256GB
Desktop Servers	

Hardware Components	Configuration
Server quantity	14 Fujitsu Primergy RX2540 M1
CPU model	Intel Xeon CPU E5-2670 v2 at 2.30GHz (12-core)
Total number of cores per server	24 cores (48 with hyperthreading)
Memory per server	512GB
Storage	
NetApp controller	AFF8040A
Disk shelf	1 DS2246
Disk drives	24 800GB SSDs

4.2 Software Components

This section describes the purpose of each software product used to test the NetApp All Flash FAS system and provides configuration details. Table 3 lists the software components and identifies the version of each component.

Table 3) Solution software components.

Software	Version
NetApp FAS	
NetApp ONTAP	9.0 RC
NetApp VSC for VMware	6.2
VMware Software	
VMware ESXi	6.0.0, 2494585
VMware vCenter Server (VCSA)	6.0.0
Citrix Software	
Citrix XenDesktop	7.6
Citrix Licensing Server	11.12.1 build 14008
Citrix StoreFront Server	2.6.0.5031
Workload Generation Utility	
Login VSI Professional	Login VSI 4.1.4.2
Database Server	
Microsoft SQL Server	11.12.1 build 14008
Microsoft SQL Server Management Studio	2.6.0.5031

4.3 Login VSI

Login Virtual Session Indexer (Login VSI) is the industry-standard load-testing tool for testing 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 major hardware and software vendors and is recommended by both leading IT analysts and the technical community. Login VSI is vendor independent and works with standardized user workloads; therefore, conclusions based on Login VSI test data are objective, verifiable, and replicable.

SBC-oriented and VDI-oriented vendor organizations that are committed to enhancing end-user experience in the most efficient way use Login VSI as an objective method of testing, benchmarking, and improving the performance and scalability of their solutions. VSImax provides the proof (vendor independent, industry standard, and easy to understand) to innovative technology vendors to demonstrate the power, the scalability, and the gains of their solutions.

Login VSI–based test results are published in technical white papers and presented at conferences. Login VSI is used by end-user organizations, system integrators, hosting providers, and testing companies. It is also the standard tool used in all tests executed in the internationally acclaimed Project Virtual Reality Check.

For more information about Login VSI or for a free test license, refer to the Login VSI website.

4.4 Full-Clone Virtual Desktops

The desktop VM template was created with the virtual hardware and software listed in Table 4. The VM hardware and software were installed and configured according to Login VSI documentation.

Table 4) Virtual desktop configuration.

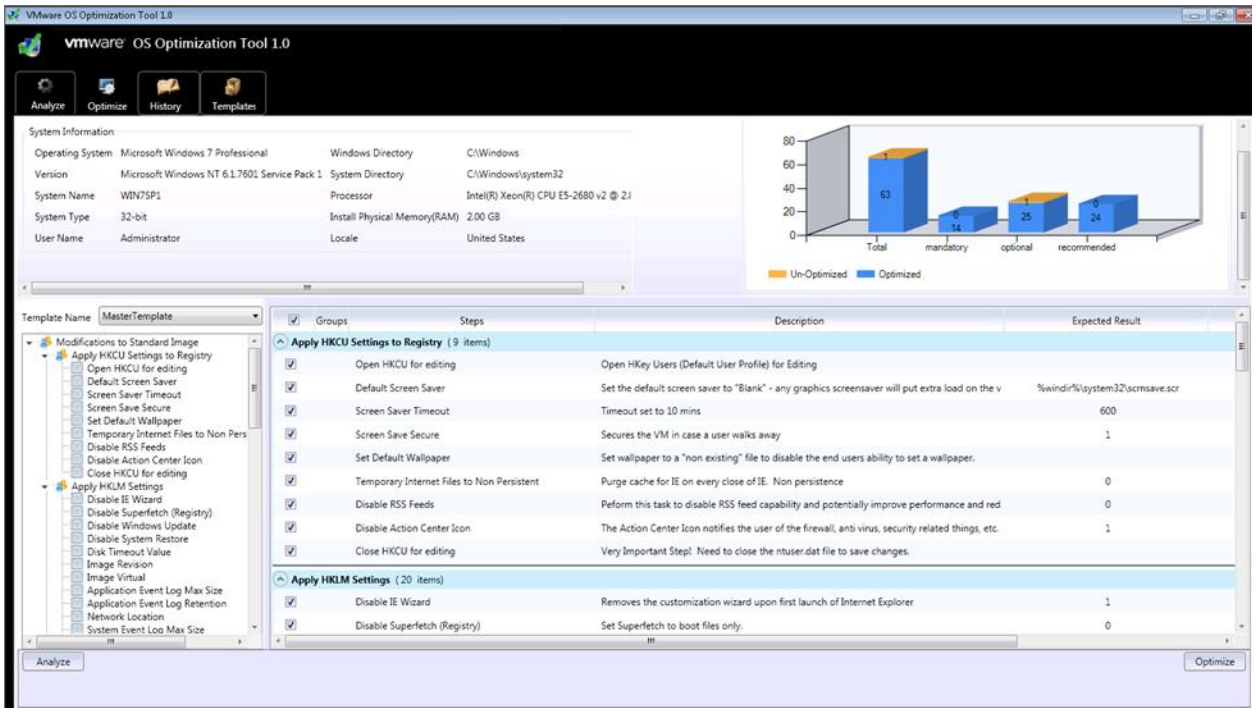
Desktop	Configuration
Desktop VM	
VM quantity	1,500
VM hardware version	11
vCPU	2 vCPUs
Memory	2GB
Hard disk size	32GB
Hard disk type	Thin provisioned
Desktop Software	
Guest OS	Windows 7 (64-bit)
VM hardware version	ESXi 6.0 and later (VM version 11)
VMware tools version	10.0.0.3000743
Microsoft Office	2013 version 15.0.4569.1506
Microsoft .NET Framework	4.5.1
Adobe Acrobat Reader	11.0.00
Adobe Flash Player	11.5.502.146

Desktop	Configuration
Java	7.0.130
Doro PDF	1.82
Citrix XenDesktop Agent	7.6
Citrix Receiver	14.1.200.13
Login VSI target software	4.1.4.2

Guest Optimization

Guest OS optimizations were applied to the template VMs used in this reference architecture. Figure 5 shows the VMware OS optimization tool that was used to perform the guest optimizations.

Figure 5) VMware OS optimization tool.



Although it might be possible to run desktops without guest optimizations, the impact of not optimizing must first be understood. Many recommended optimizations address services and features (such as hibernation, Windows update, or system restore) that do not provide value in a virtual desktop environment. To run services and features that do not add value would decrease the overall density of the solution and increase cost because these services and features would consume CPU, memory, and storage resources in relation to both capacity and I/O.

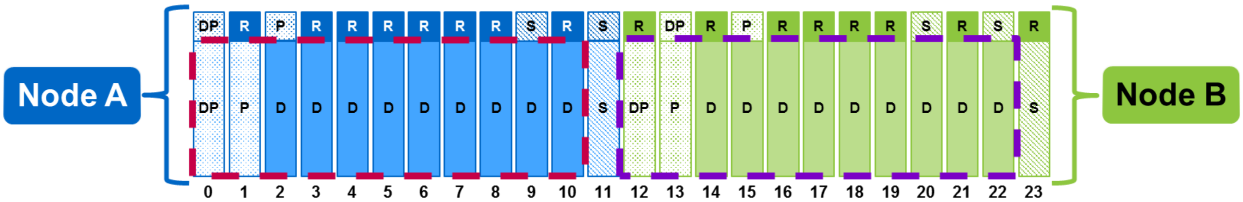
5 Storage Design

This section provides an overview of the storage design, the aggregate and volume layout, and VSC.

5.1 Aggregate Layout

In this reference architecture, we used 24 800GB SSDs divided across the 2 nodes of an AFF8040 controller using advanced drive partitioning version 1. This architecture is shown in Figure 6.

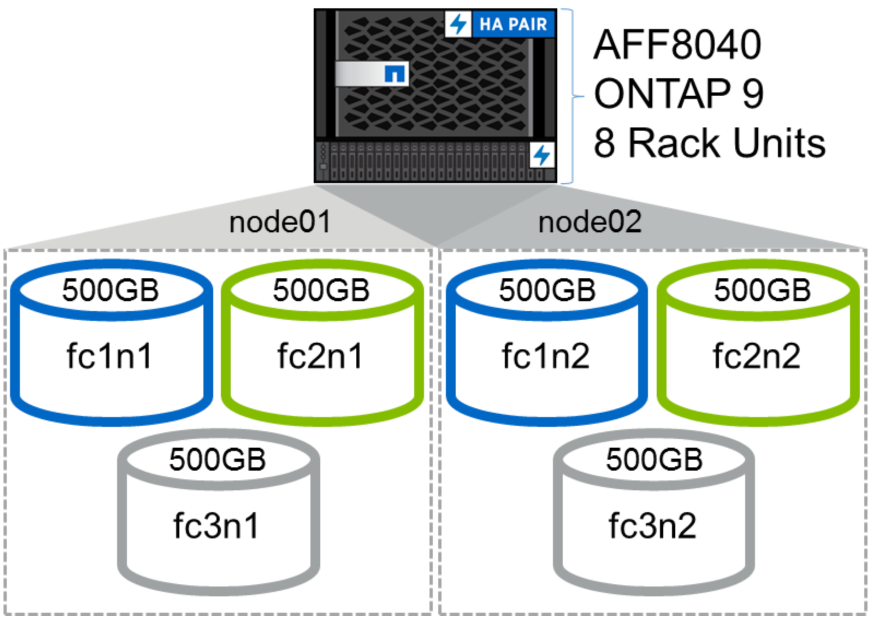
Figure 6) SSD layout.



5.2 Volume Layout

To adhere to NetApp best practices, all volumes were provisioned with NetApp VSC. Figure 7 shows how the volumes were arranged.

Figure 7) Volume layout for nonpersistent desktops.



Note: A root volume for the virtual desktop SVM is present but is not depicted.

5.3 NetApp Virtual Storage Console for VMware vSphere

NetApp VSC was used to provision the datastores and the desktop machines in this reference architecture. NetApp VSC for VMware vSphere provides integrated, comprehensive, end-to-end virtual storage management for your VMware infrastructure, including discovery, health monitoring, capacity management, provisioning, cloning, backup, restore, and disaster recovery.

Note: When using the NetApp VSC to set the NFS host optimizations, it sets the NFS.MaxQueueDepth to 64 for all versions of ESX. The default setting for vSphere 6 is 128, and the default in 5.x is 4,294,967,295. NetApp recommends using 64 when using a single-processor FAS with spinning media. For All Flash FAS, NetApp recommends that the NFS.MaxQueueDepth should be set to at least 128 if not higher to eliminate any queuing bottlenecks. The exception to this recommendation occurs when a customer has both a FAS2500 system as well as an All Flash FAS system connected to the same ESX cluster. If the setting were to remain at 64, the customer would not get the maximum performance from the All Flash FAS system. If the customer sets the NFS.MaxQueueDepth to a higher value, there is a risk that the datastore might disconnect on the FAS2500 system. Careful consideration should be given before making any changes to NFS.MaxQueueDepth.

```
Get-VMHost | Sort Name | Get-AdvancedSetting -Name NFS.MaxQueueDepth | Set-AdvancedSetting -Value 256
-Confirm:$false
```

5.4 Importing VSC-Created VMs into XenDesktop

Here is a truncated example of the csv file that was used to import the persistent VSC-created desktops into XenDesktop.

```
[VirtualMachinePath],[ADComputerAccount],[AssignedUser]
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM001.vm,vdi.rtp.netapp.com\FCV1N1VM0
01,VDI\Login VSI0001
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM002.vm,vdi.rtp.netapp.com\FCV1N1VM0
02,VDI\Login VSI0002
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM003.vm,vdi.rtp.netapp.com\FCV1N1VM0
03,VDI\Login VSI0003
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM004.vm,vdi.rtp.netapp.com\FCV1N1VM0
04,VDI\Login VSI0004
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM005.vm,vdi.rtp.netapp.com\FCV1N1VM0
05,VDI\Login VSI0005
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM006.vm,vdi.rtp.netapp.com\FCV1N1VM0
06,VDI\Login VSI0006
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM007.vm,vdi.rtp.netapp.com\FCV1N1VM0
07,VDI\Login VSI0007
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM008.vm,vdi.rtp.netapp.com\FCV1N1VM0
08,VDI\Login VSI0008
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM009.vm,vdi.rtp.netapp.com\FCV1N1VM0
09,VDI\Login VSI0009
XDHyp:\connections\vCenter\VDI.datacenter\Desktops.cluster\fcv1n1VM010.vm,vdi.rtp.netapp.com\FCV1N1VM0
10,VDI\Login VSI0010
```

6 Testing and Validation: Persistent Full-Clone Desktops

This section describes the testing and validation of persistent Citrix XenDesktop full-clone desktops.

6.1 Test Results Overview

Table 5 lists the high-level results that were achieved during the reference architecture testing. One item to note is that throughout the paper, graphs displaying system utilization are shown. With ONTAP, system utilization is not absolute. This means that during a given workload, system utilization on both nodes could exceed 50% and could perform excellently during failover without increasing latency. For example, node 1 is running at 70%, and node 2 is running at 60%; during failover, the combined controllers may be running at 85%, still serving data in <1ms.

Table 5) Test results overview.

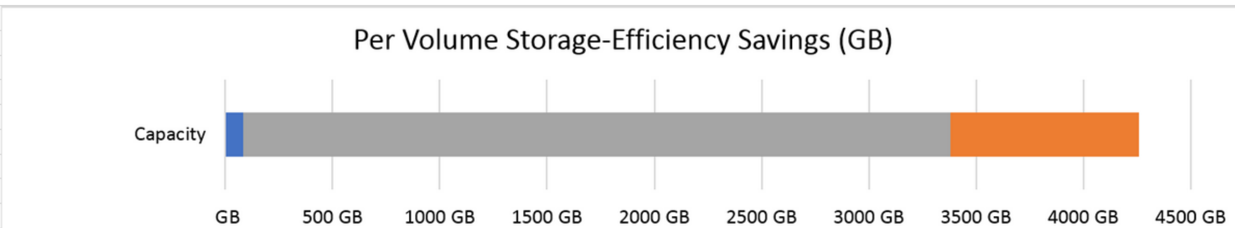
Test	Time to Complete	Peak IOPS	Peak Throughput	Average Storage Latency
Boot storm test	~10 min	102,437	3.1GBps	1.4ms
Boot storm test during failover	~16 min	67,769	1.9GBps	2.7ms
Login VSI Monday morning login and workload	7.8 sec/VM	60,533	724MBps	455µs

Test	Time to Complete	Peak IOPS	Peak Throughput	Average Storage Latency
Login VSI Monday morning login and workload during failover	6.95 sec/VM	36,573	593MBps	641µs
Login VSI Tuesday morning login and workload	7.0 sec/VM	26,237	525MBps	470µs
Login VSI Tuesday morning login and workload during failover	6.65 sec/VM	26,187	477MBps	669µs
Patching desktops with ~900MB of patches (10-sec delay)	~160 min	68,921	1.6GBps	936µs

6.2 Storage Efficiency

During the tests, FlexClone technology was used to provision the VMs, and deduplication was enabled. On average, a 50:1 deduplication efficiency ratio, or 98% storage efficiency, was observed. This means that 50 virtual desktops consumed the storage of one desktop on disk. These high rates are due not only to inline deduplication but also to the ability of FlexClone technology to instantaneously create storage-efficient virtual desktops. Without these technologies, traditional storage environments would have consumed 25TB of storage. With deduplication and FlexClone technology, 1,500 desktops consumed only 490GB of storage, a savings of over 98%. At the aggregate level, which accounts for physical blocks stored, we saw an additional 9% savings due to compaction in addition to the 98% seen at FlexVol level. Figure Figure 8 shows the significant difference in storage efficiency savings.

Figure 8) Storage efficiency savings.



	Capacity
Capacity Used	84 GB
Dedupe Savings	3296 GB
Compression Savings	880 GB

Because of the synthetic nature of the data used to perform these tests, these results are not typical of real-world savings. In addition, although thin provisioning was used for each volume, thin provisioning is not a storage-reduction technology and therefore was not reported on. Table 6 lists the efficiency results from the testing after steady state but before patching.

Table 6) Efficiency results for each FlexVol volume.

Capacity Used	Total Savings	Dedupe Savings	Compression Savings
84GB	4176GB	3296GB	880GB

6.3 Boot Storm Test

This section describes test objectives and methodology and provides results from boot storm testing. See Appendix: Workload Profilesfor specific I/O profile data for this workload.

Test Objectives and Methodology

The objective of this test was to determine how long it would take to boot 1,500 virtual desktops, which might occur, for example, after maintenance activities and server host failures.

This test was performed by powering on all 1,500 VMs from within XenDesktop and observing when the status of all VMs in Citrix XenDesktop changed to `Registered`.

Table 7 lists the boot storm data that was gathered.

Table 7) Results for full-clone boot storm.

Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
~10 min	1.35ms	102,437	54,724	3.1GBps	1.5GBps	100%	52%

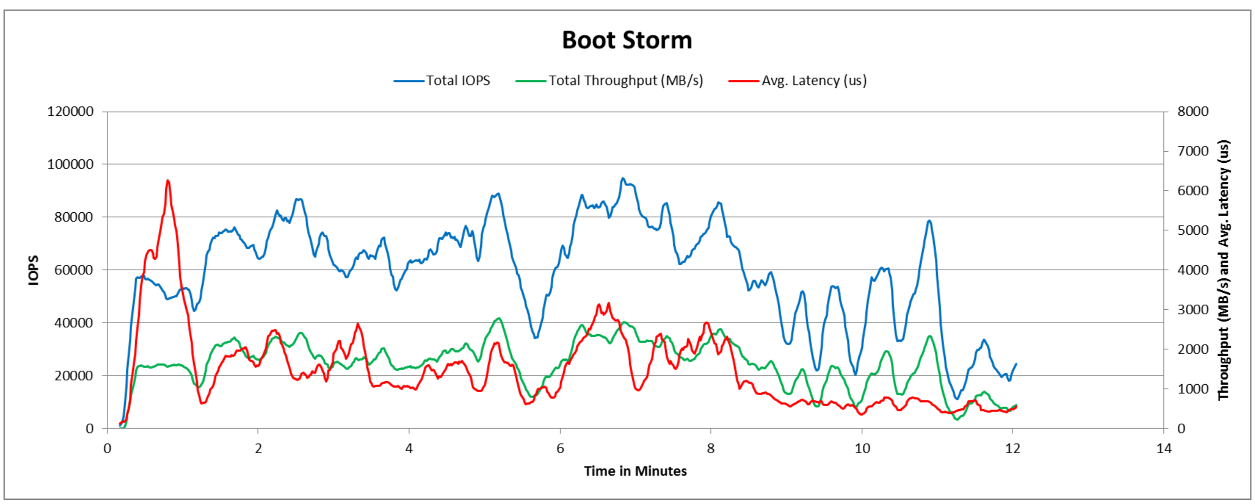
Note: All desktops had the status `Registered` in Citrix XenDesktop.

Note: CPU and latency measurements are based on the average across both nodes of the cluster. IOPS and throughput are based on a combined total of each.

Throughput, IOPS, and Latency

Figure 9 shows throughput and IOPS for a full-clone boot storm.

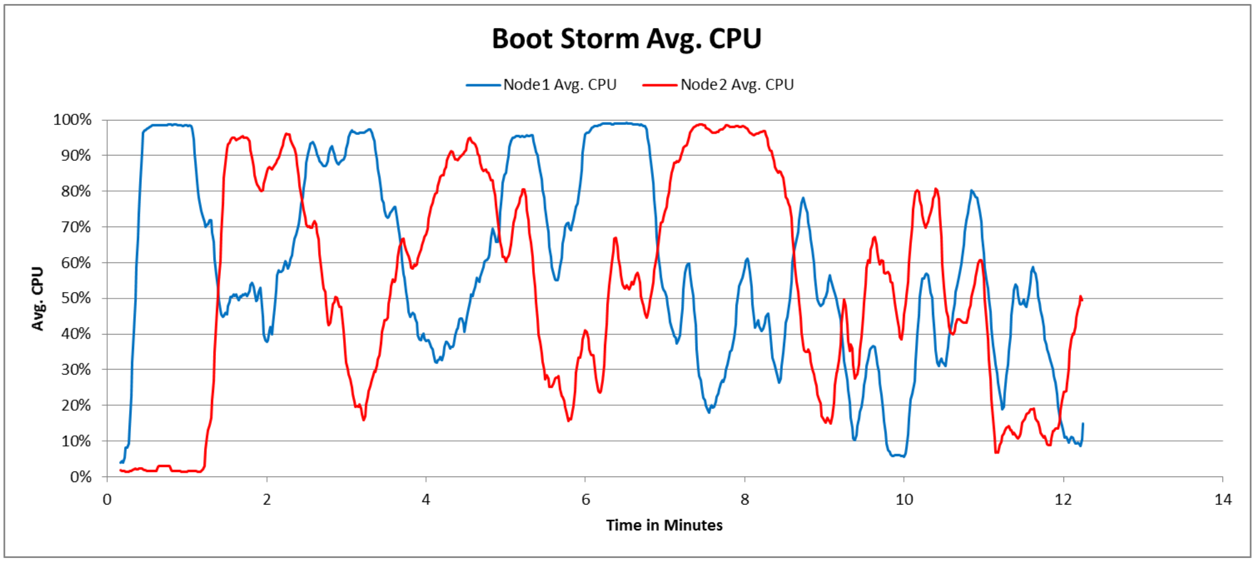
Figure 9) Throughput and IOPS for full-clone boot storm.



Storage Controller CPU Utilization

Figure 10 shows storage controller CPU utilization for a full-clone boot storm.

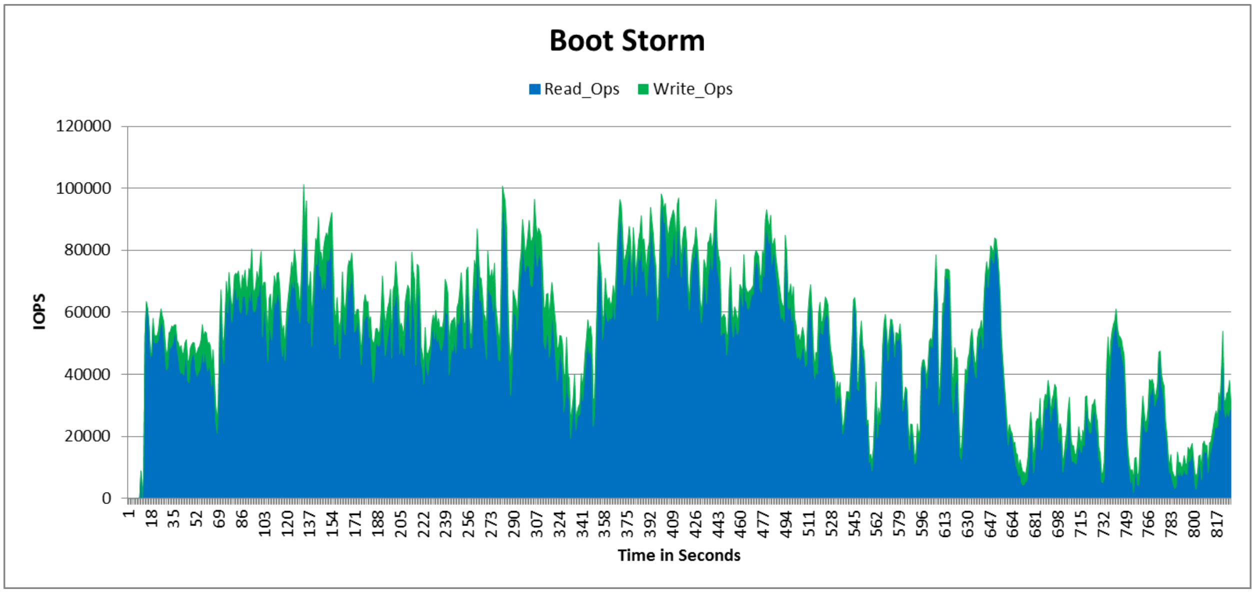
Figure 10) Storage controller CPU utilization for full-clone boot storm.



Read/Write IOPS

Figure 11 shows read and write IOPS for a full-clone boot storm.

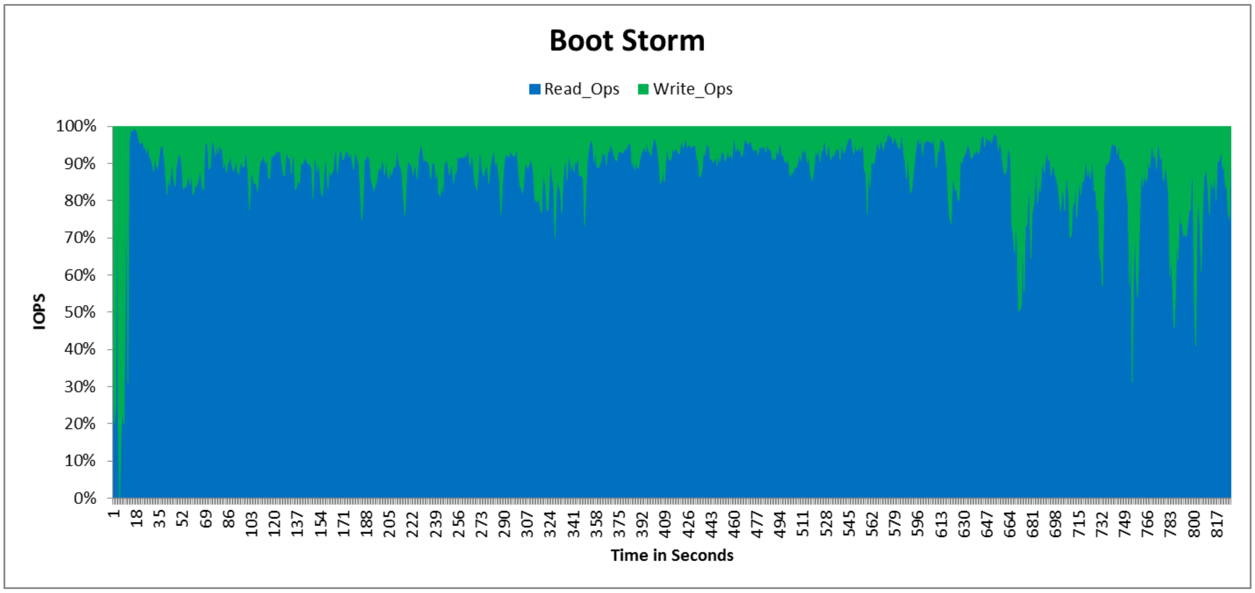
Figure 11) Read/write IOPS for full-clone boot storm.



Read/Write Ratio

Figure 12 shows the read and write ratio for a full-clone boot storm.

Figure 12) Read/write ratio for full-clone boot storm.



Customer Impact (Test Conclusions)

During the boot of 1,500 persistent desktops, the storage controller reached about the maximum number of concurrent boot operations. However, booting more desktops might take longer as utilization increases. The focus of this test was not on client latency but on restoring the users' desktops as quickly as possible.

6.4 Boot Storm Test During Storage Failover

This section describes test objectives and methodology and provides results from boot storm testing during storage controller failover. See Appendix: Workload Profiles for specific I/O profile data for this workload.

Test Objectives and Methodology

The objective of this test was to determine how long it would take to boot 1,500 virtual desktops if the storage controller had a problem and was failed over. This test used the same methodologies and process that were used in section 6.3, "Boot Storm Test."

Table 8 shows the data that was gathered for the boot storm during storage failover.

Table 8) Results for full-clone boot storm during storage failover.

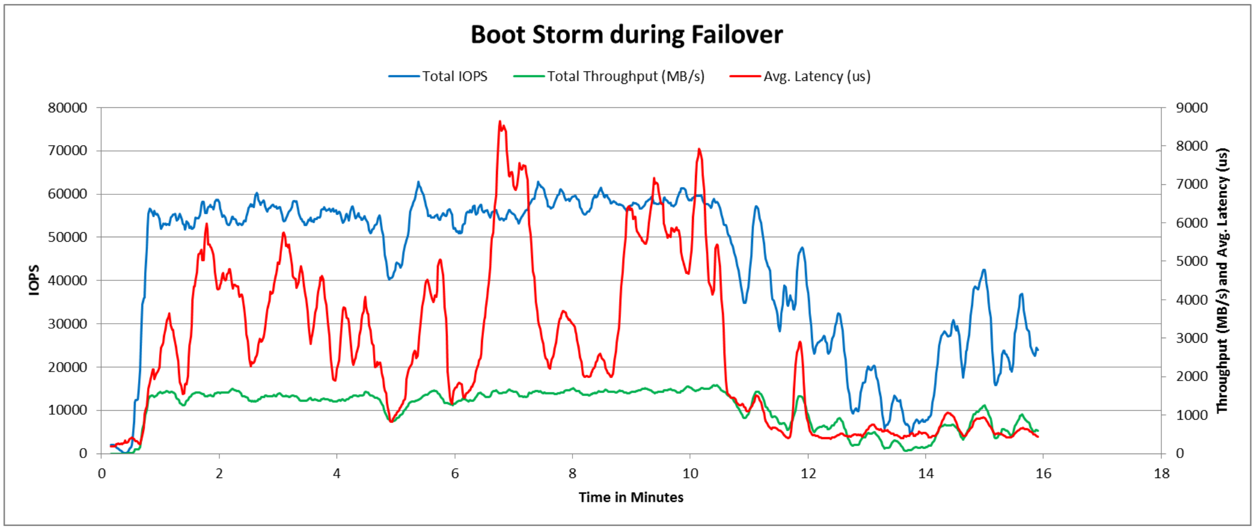
Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
~16 min	2.7ms	67,769	43,981	1.9GBps	1.2GBps	100%	78%

Note: All desktops had the status of `Registered` in Citrix XenDesktop.

Throughput, IOPS, and Latency

Figure 13 shows throughput, IOPS, and latency for a full-clone boot storm during storage failover.

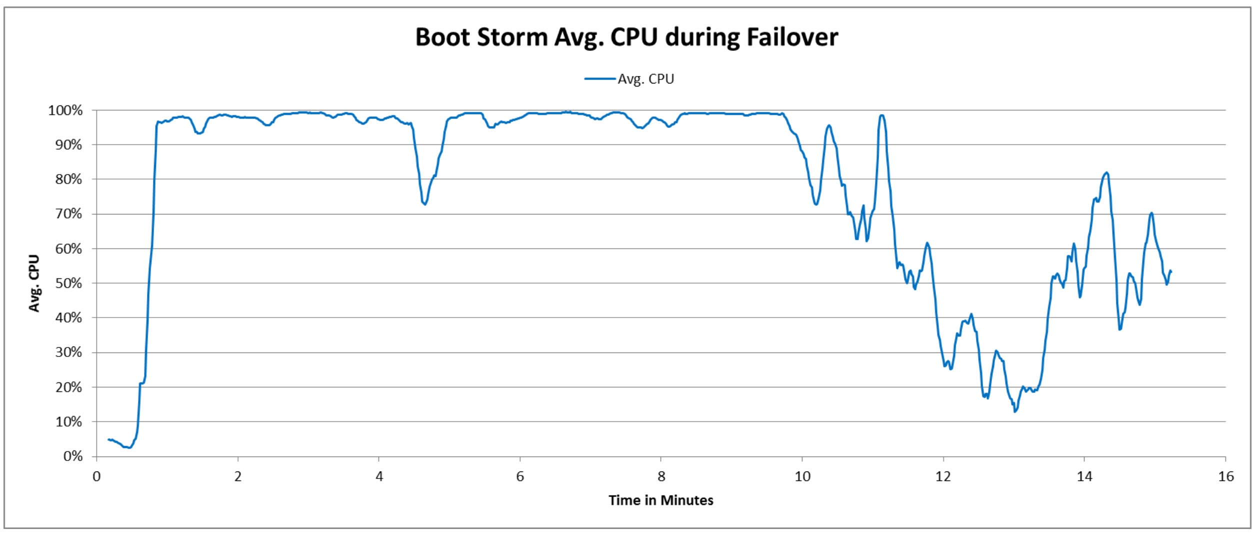
Figure 13) Throughput, IOPS, and latency for full-clone boot storm during storage failover.



Storage Controller CPU Utilization

Figure 14 shows storage controller CPU utilization for a full-clone boot storm during storage failover.

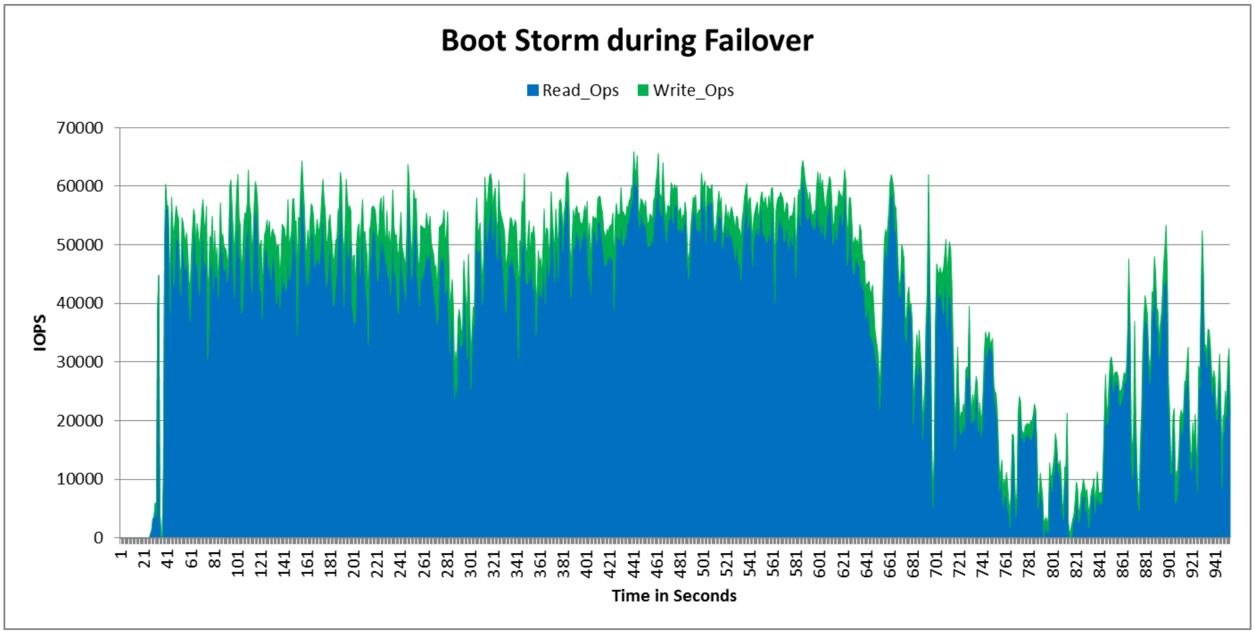
Figure 14) Storage controller CPU utilization for full-clone boot storm during storage failover.



Read/Write IOPS

Figure 15 shows read and write IOPS for a full-clone boot storm during storage failover.

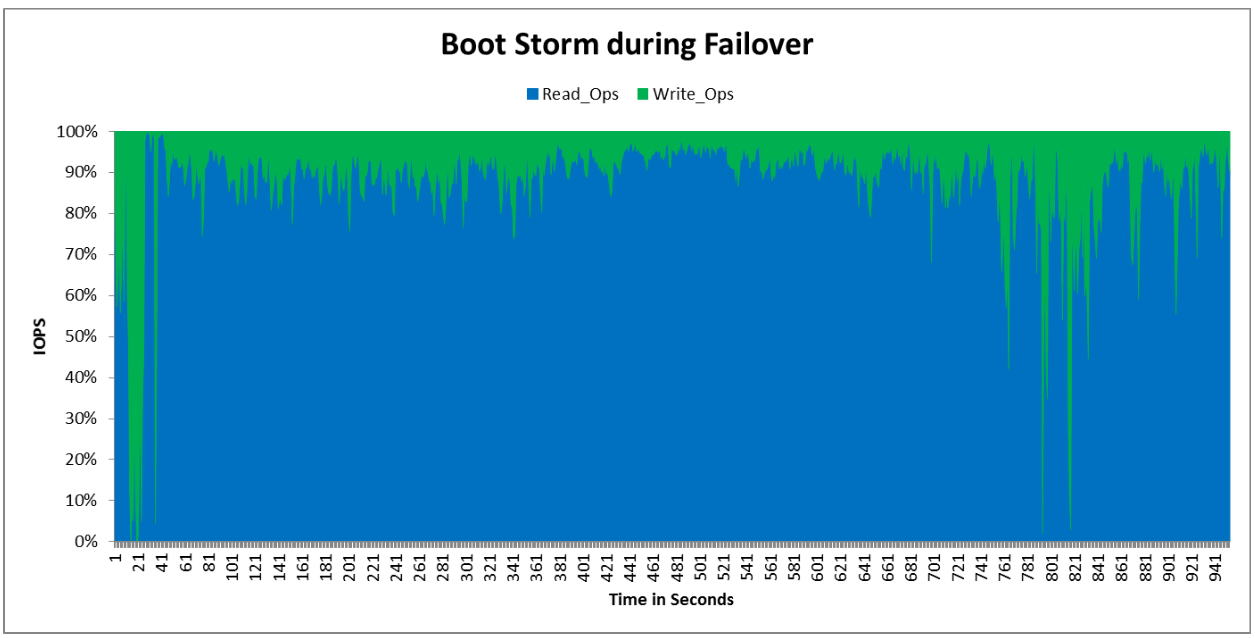
Figure 15) Read/write IOPS for full-clone boot storm during storage failover.



Read/Write Ratio

Figure 16 shows read and write ratios for a full-clone boot storm during storage failover.

Figure 16) Read/write ratio for full-clone boot storm during storage failover.



Customer Impact (Test Conclusions)

During the boot of 1,500 persistent desktops with storage failed over, the storage controller was able to boot 1,500 desktops on one node in approximately 16 minutes. Tests were conducted to measure the impact of using Citrix XenDesktop to boot the desktops. The focus of this test, however, was not on client latency but on restoring the users' desktops as quickly as possible.

6.5 Steady-State Login VSI Test

This section describes test objectives and methodology and provides results from steady-state Login VSI testing. See the Appendix: Workload Profiles for specific I/O profile data for this workload.

Test Objectives and Methodology

The objective of this test was to run a Login VSI office worker workload to determine how the storage controller performed and evaluate the end-user experience. This Login VSI workload first had the users log in to their desktops and begin working. The login phase occurred over a 48-minute period.

Three different login scenarios were included because each has a different I/O profile. We measured storage performance, login time, and VSImax. VSImax is a Login VSI value that represents the maximum number of users who can be deployed on a given platform. VSImax was not reached in any of the Login VSI tests. The following sections define the login scenarios.

Monday Morning Login and Workload Test

In this scenario, 1,500 users logged in after the VMs had already been logged into once, the profile had been created, and the desktop had been rebooted. During this type of login, user and profile data, application binaries, and libraries had to be read from a disk because they were not already contained in the VM memory. Table 9 shows the results for full-clone Monday morning login and workload. See the Appendix: Workload Profiles for specific I/O profile data for this workload.

Table 9) Results for full-clone Monday morning login and workload.

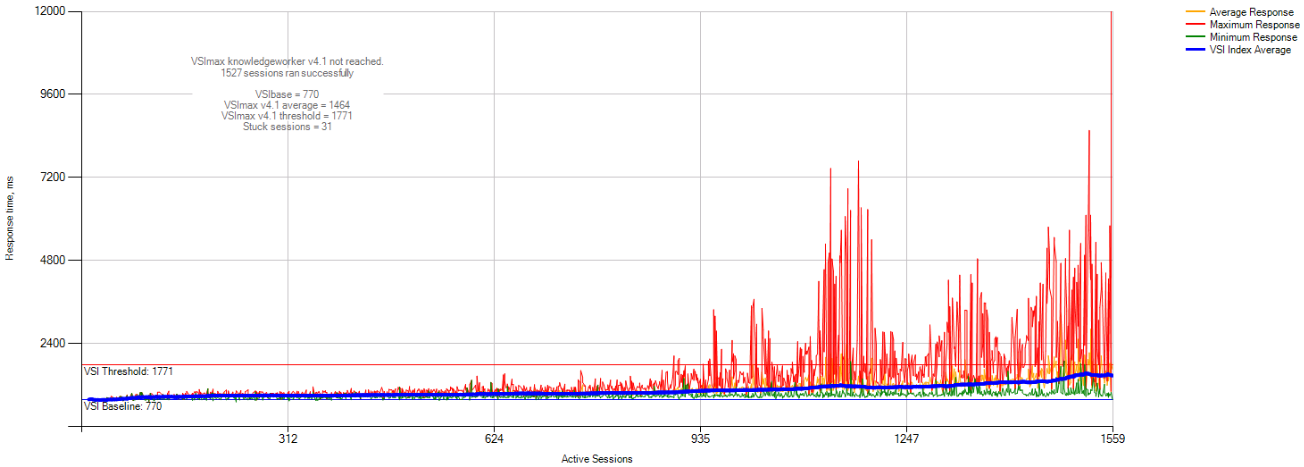
Desktop Login Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
7.8 sec/VM	0.46ms	60,533	14,877	0.7GBps	0.3GBps	88%	25%

Note: CPU and latency measurements are based on the average across both nodes of the cluster. IOPS and throughput are based on a combined total of each.

Login VSI VSImax Results

Because the Login VSI VSImax was not reached, more VMs could be deployed on this infrastructure. Figure shows the VSImax results for Monday morning login and workload.

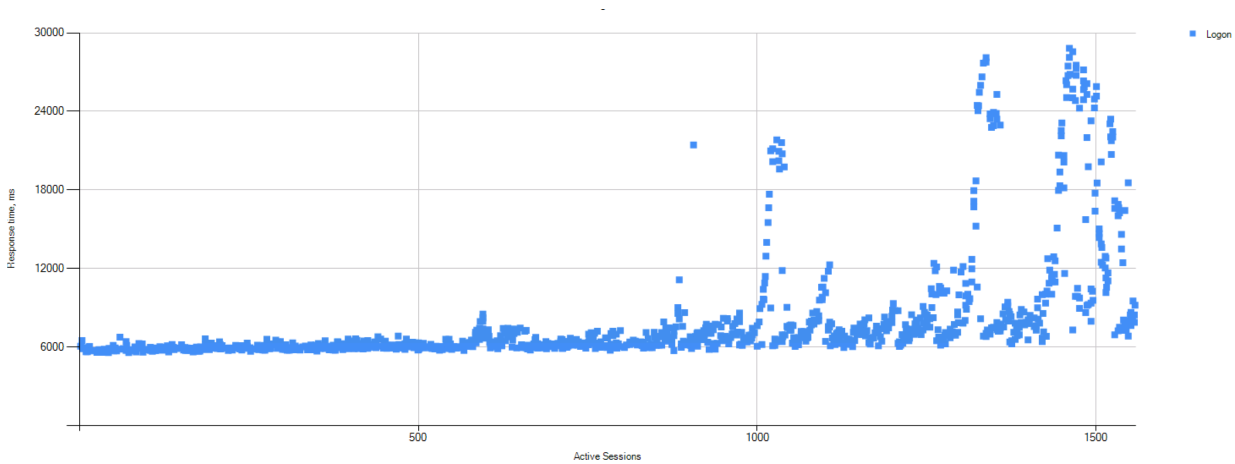
Figure 17) VSImax results for full-clone Monday morning login and workload.



Desktop Login Time

Average desktop login time was 7.8 seconds, which is considered an excellent login time. Figure 18 shows a scatterplot of the Monday morning login times.

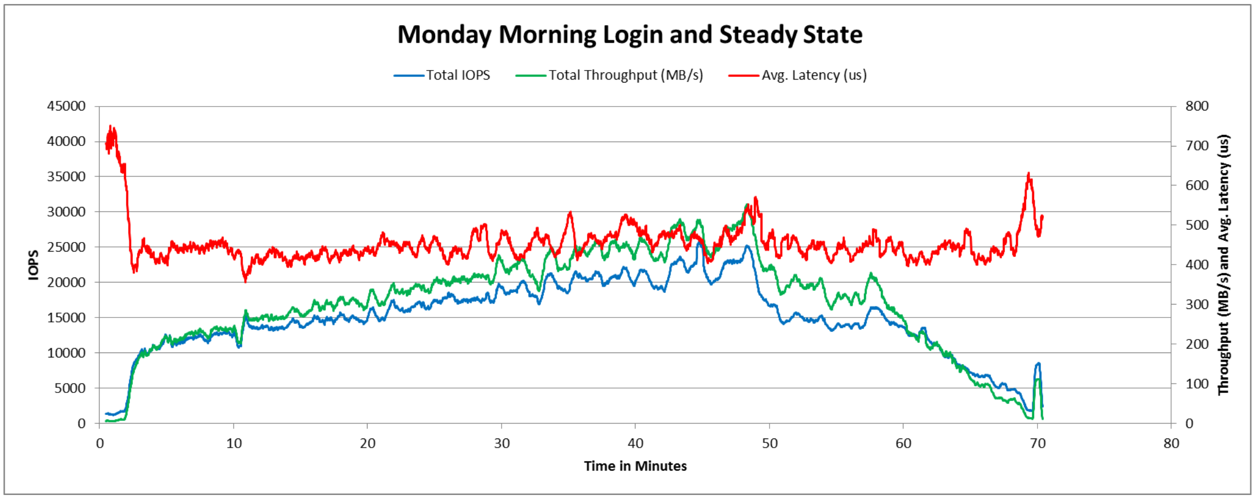
Figure 18) Scatterplot of full-clone Monday morning login times.



Throughput, IOPS, and Latency

Figure 19 shows throughput, IOPS, and latency for full-clone Monday morning login and workload.

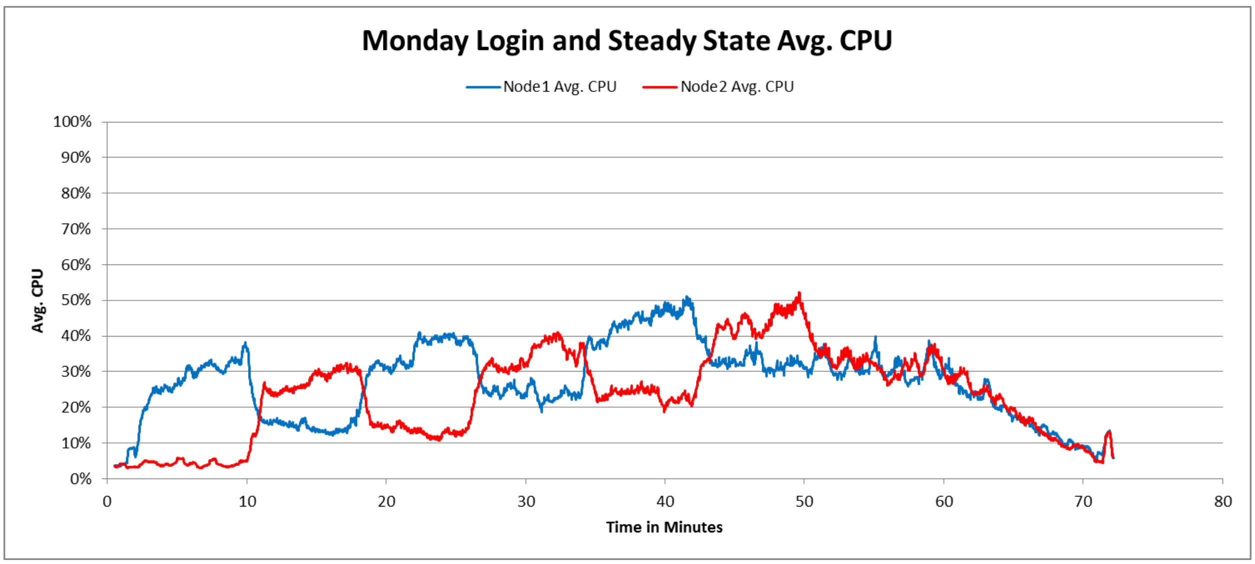
Figure 19) Throughput, IOPS, and latency for full-clone Monday morning login and workload.



Storage Controller CPU Utilization

Figure 20 shows storage controller CPU utilization for full-clone Monday morning login and workload.

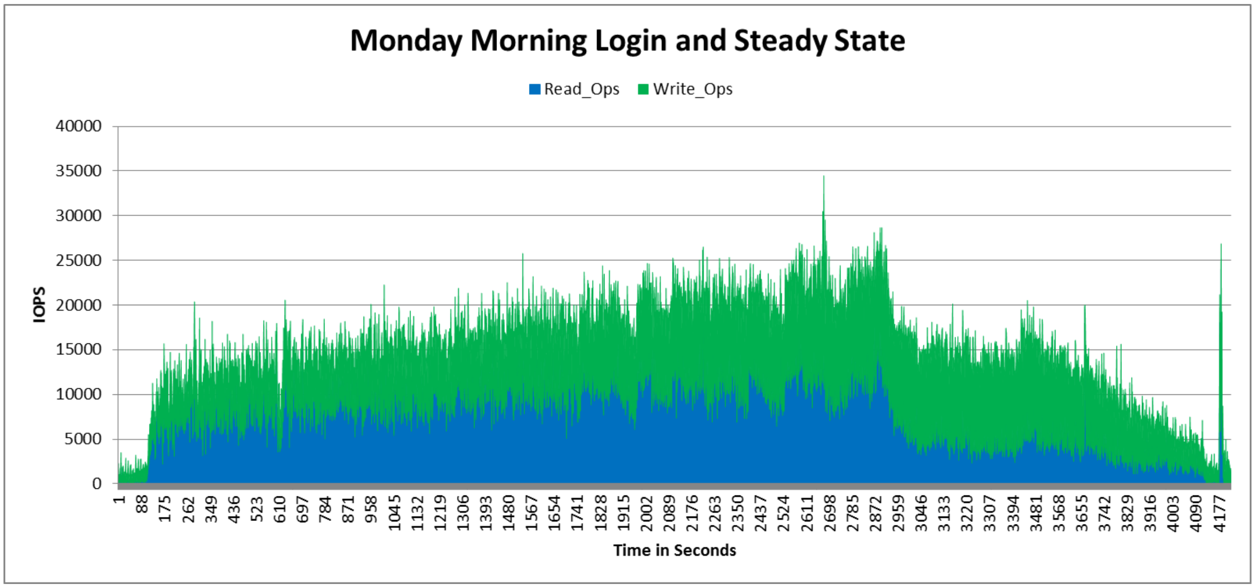
Figure 20) Storage controller CPU utilization for full-clone Monday morning login and workload.



Read/Write IOPS

Figure 21 shows read and write IOPS for full-clone Monday morning login and steady state.

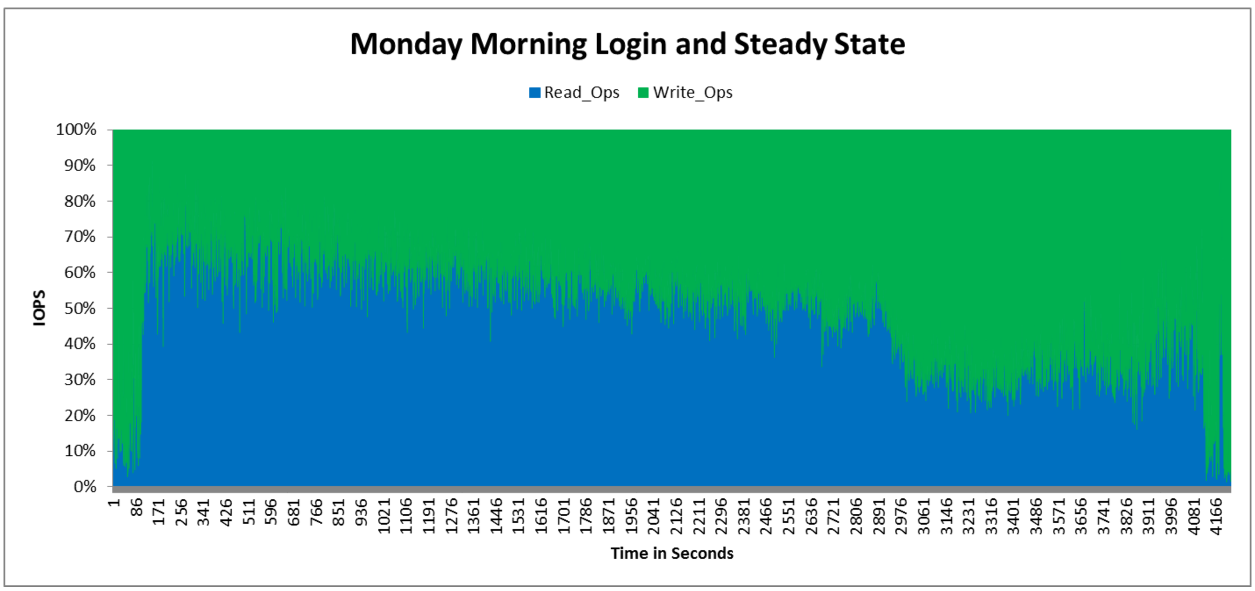
Figure 21) Read/write IOPS for full-clone Monday morning login and workload.



Read/Write Ratio

Figure 22 shows read and write ratios for full-clone Monday morning login and workload.

Figure 22) Read/write ratio for full-clone Monday morning login and workload.



Customer Impact (Test Conclusions)

During the Monday morning login test, the storage controller performed very well. The CPU utilization was not high during this test, latencies were under 1ms, and desktop performance was excellent. These results suggest that it might be possible to double the storage controller workload to 3,000 users or more and still maintain excellent end-user performance. The Monday morning login during storage failover test described in the following section reinforces that point.

Monday Morning Login and Workload During Storage Failover Test

In this scenario, 1,500 users logged in after the VMs had already been logged into once, the profile had been created, and the desktops had been rebooted, but during a storage failover event. During this type of login, user and profile data, application binaries, and libraries had to be read from a disk because they were not already contained in the VM memory. Table 10 lists the results for Monday morning login and workload during storage failover. See the Appendix: Workload Profiles for specific I/O profile data for this workload.

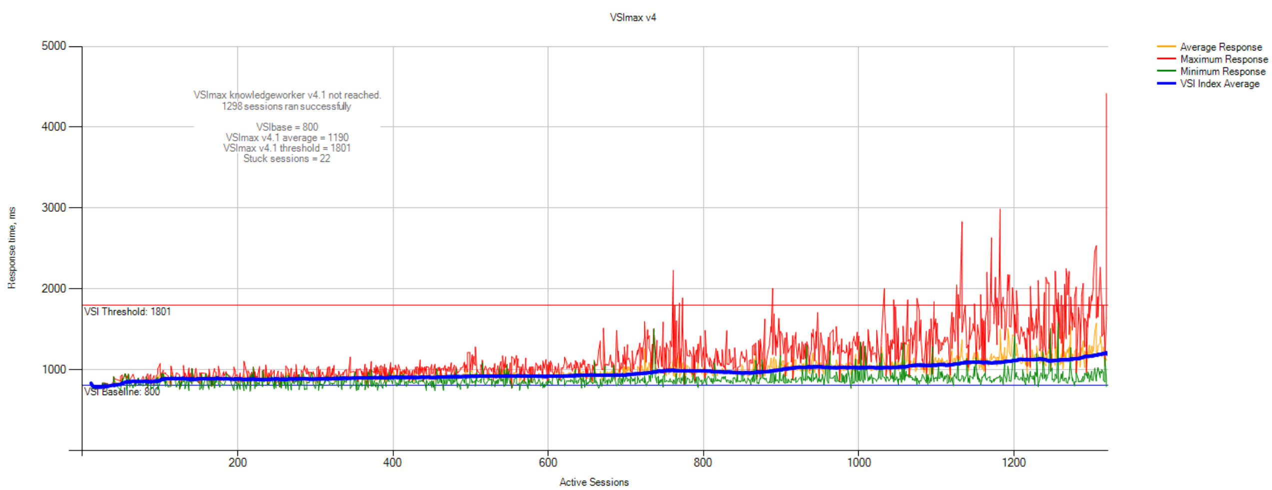
Table 10) Results for full-clone Monday morning login and workload during storage failover.

Desktop Login Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
6.95 sec/VM	0.64ms	36,573	12,119	0.6GBps	0.2GBps	99%	41%

Login VSI VSImax Results

Because the Login VSI VSImax v4.1 limit was not reached, more VMs could be deployed on this infrastructure. Figure 23 shows the VSImax results for Monday morning login and workload during storage failover.

Figure 23) VSImax results for full-clone Monday morning login and workload during storage failover.

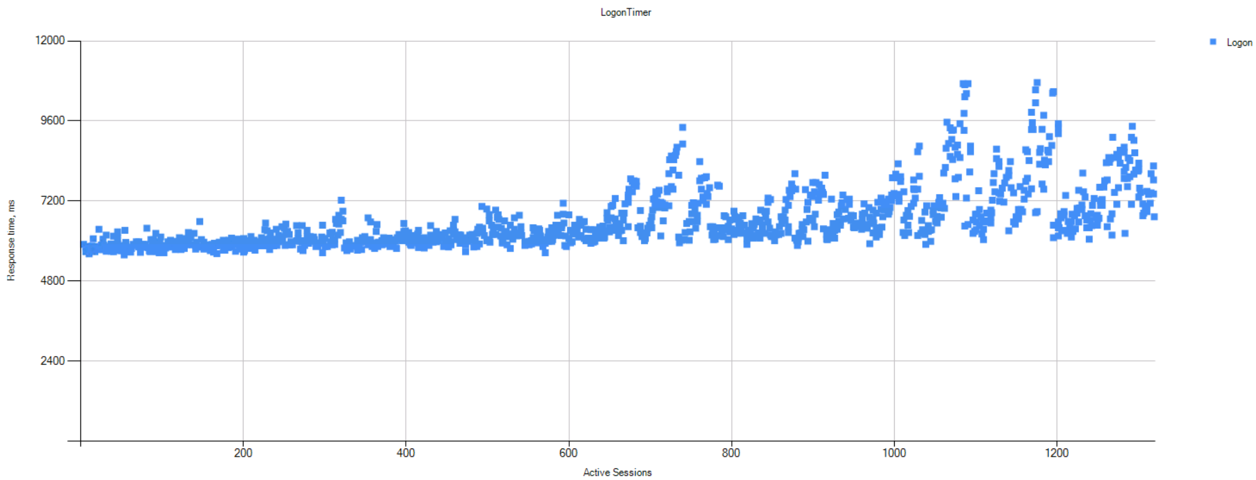


Desktop Login Time

Average desktop login time was 6.95 seconds, which is considered an excellent login time, especially during a failover situation.

Figure shows a scatterplot of the Monday morning login times during storage failover.

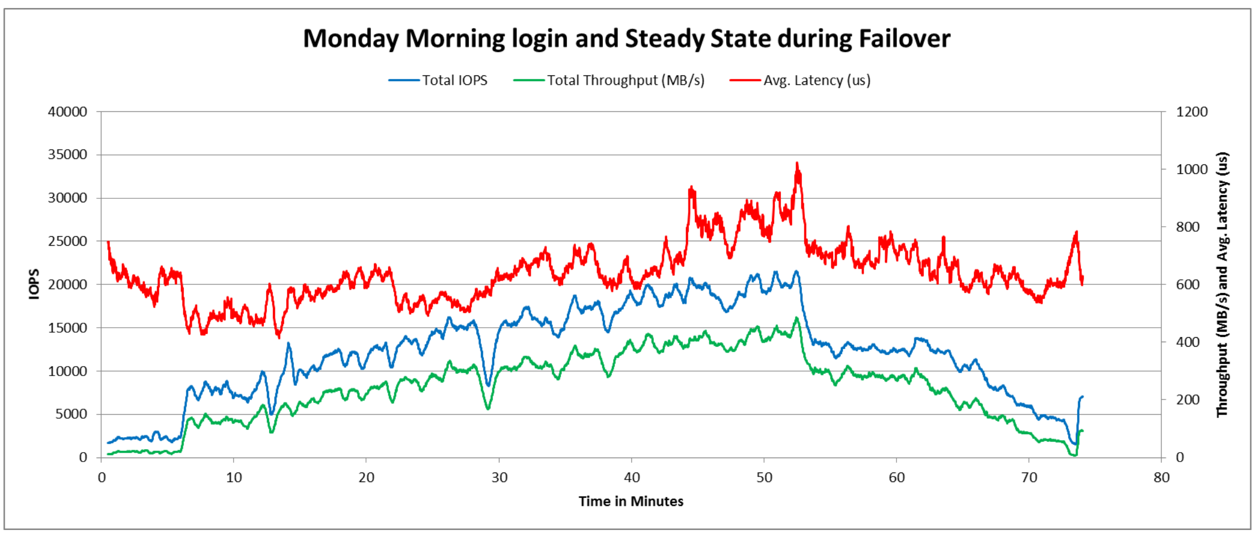
Figure 24) Scatterplot of full-clone Monday morning login times during storage failover.



Throughput, IOPS, and Latency

Figure 25 shows throughput, IOPS, and latency for full-clone Monday morning login and workload during storage failover.

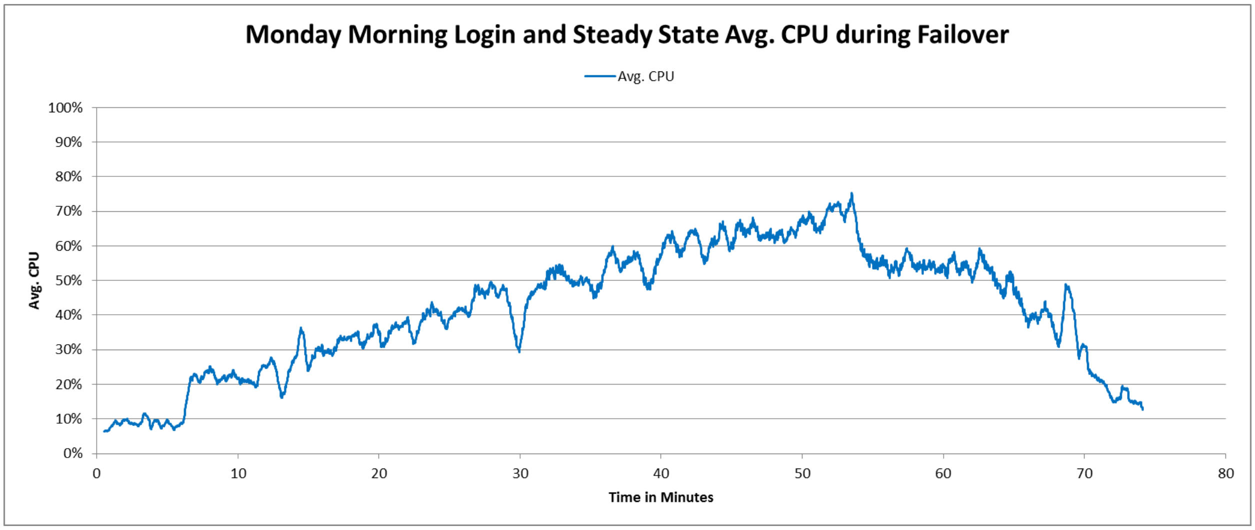
Figure 25) Throughput, IOPS, and latency for full-clone Monday morning login and workload during storage failover.



Storage Controller CPU Utilization

Figure 26 shows storage controller CPU utilization for full-clone Monday morning login and workload during storage failover.

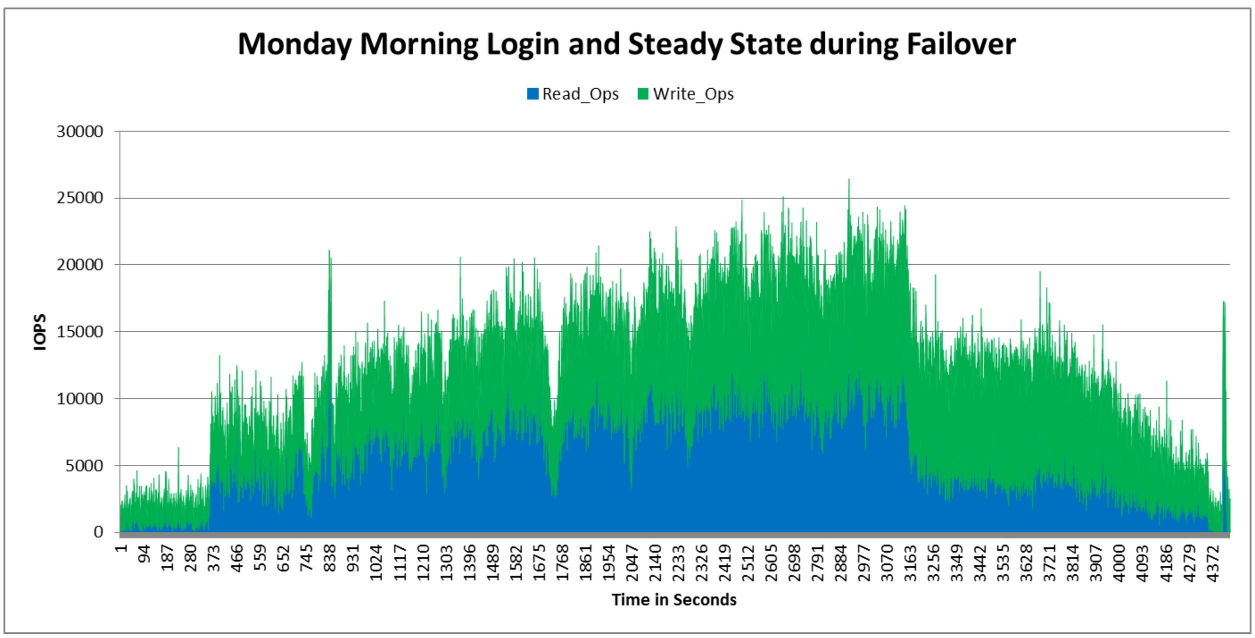
Figure 26) Storage controller CPU utilization for full-clone Monday morning login and workload during storage failover.



Read/Write IOPS

Figure 27 shows read and write IOPS for full-clone Monday morning login and workload during storage failover.

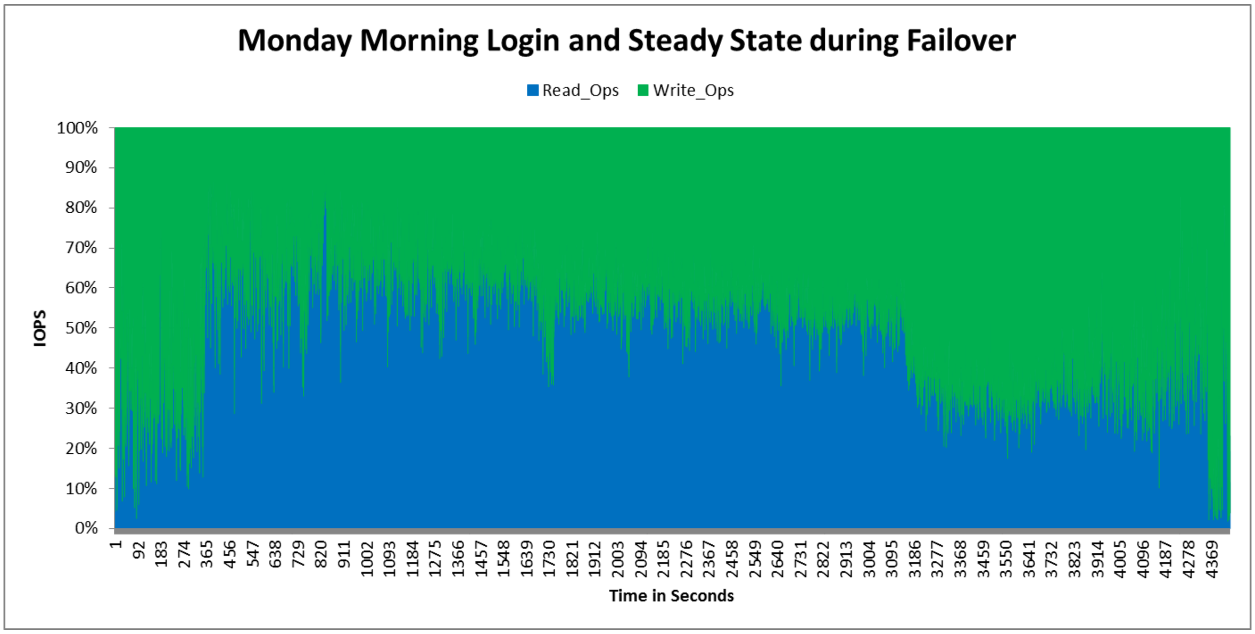
Figure 27) Read/write IOPS for full-clone Monday morning login and workload during storage failover.



Read/Write Ratio

Figure 28 shows read and write ratios for full-clone Monday morning login and workload during storage failover.

Figure 28) Read/write ratio for full-clone Monday morning login and workload during storage failover.



Customer Impact (Test Conclusions)

During the Monday morning login test during storage failover, the storage controller performed very well. The CPU utilization averaged less than 50%, latencies were under 1ms, and desktop performance was excellent. These results suggest that for this type of workload it might be possible to double the storage controller workload to 3,000 users total (1,500 per node) with excellent end-user performance and with the ability to tolerate a storage failover.

Tuesday Morning Login and Workload Test

In this scenario, 1,500 users logged in to virtual desktops that had been logged into previously and that had not been power-cycled. In this situation, VMs reduced the impact on storage by retaining user and profile data, application binaries, and libraries in memory. Table 11 lists the results for Tuesday morning login and workload. See the Appendix: Workload Profiles for specific I/O profile data for this workload.

Table 11) Results for full-clone Tuesday morning login and workload.

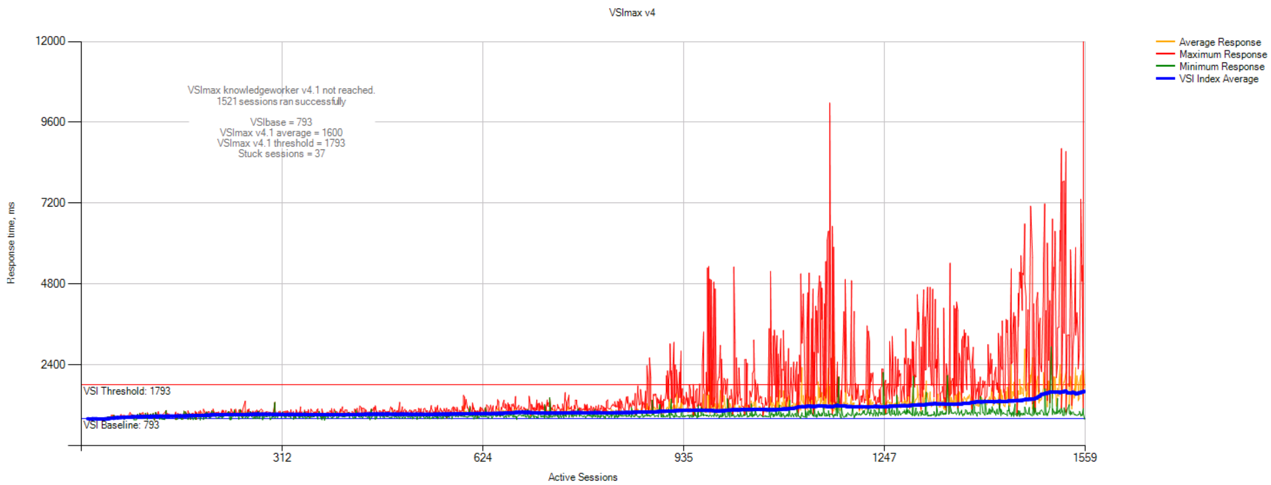
Desktop Login Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
7.03 sec/VM	0.47ms	26,237	8,157	0.5GBps	0.2GBps	79%	18%

Note: CPU and latency measurements are based on the average across both nodes of the cluster. IOPS and throughput are based on a combined total of each.

Login VSI VSImax Results

Because the Login VSI VSImax v4.1 was not reached, more VMs could be deployed on this infrastructure. Figure 29 shows the VSImax results for Tuesday morning login and workload.

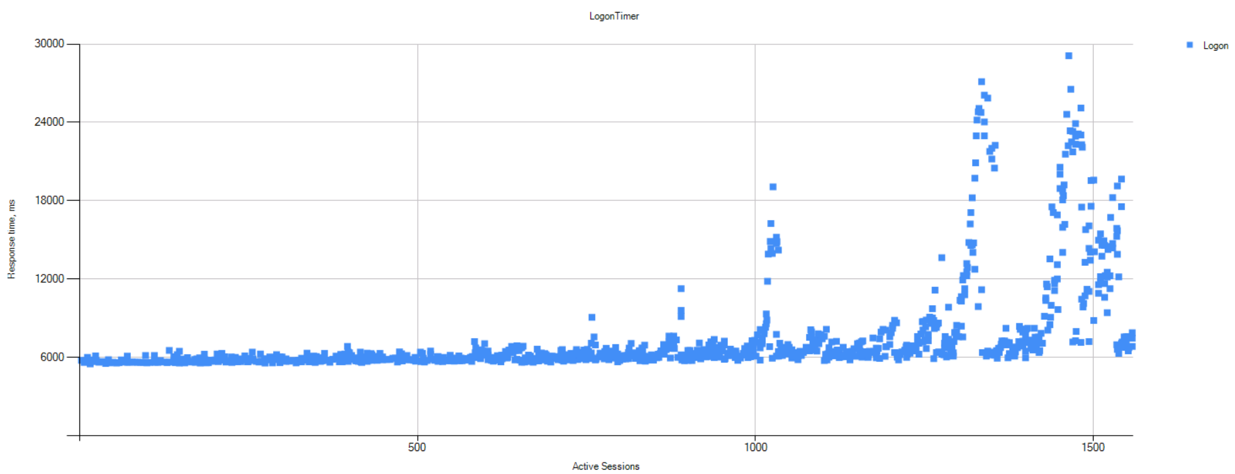
Figure 29) VSImax results for full-clone Tuesday morning login and workload.



Desktop Login Time

Average desktop login time was 7.03 seconds, which is considered an excellent login time. Figure 30 shows a scatterplot of the Tuesday morning login times.

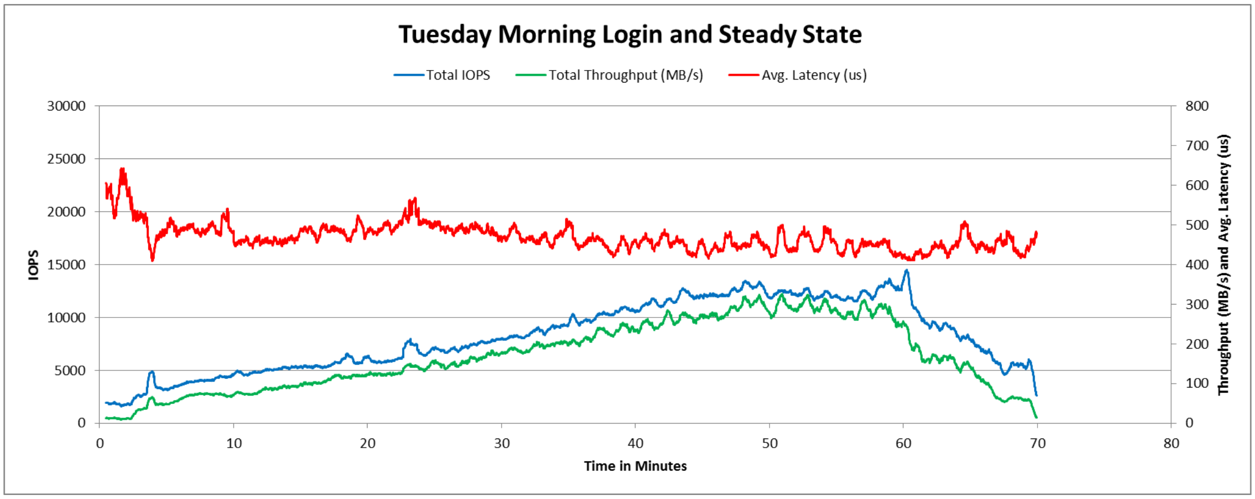
Figure 30) Scatterplot of full-clone Tuesday morning login times.



Throughput, IOPS, and Latency

Figure 31 shows throughput, IOPS, and latency for full-clone Tuesday morning login and workload.

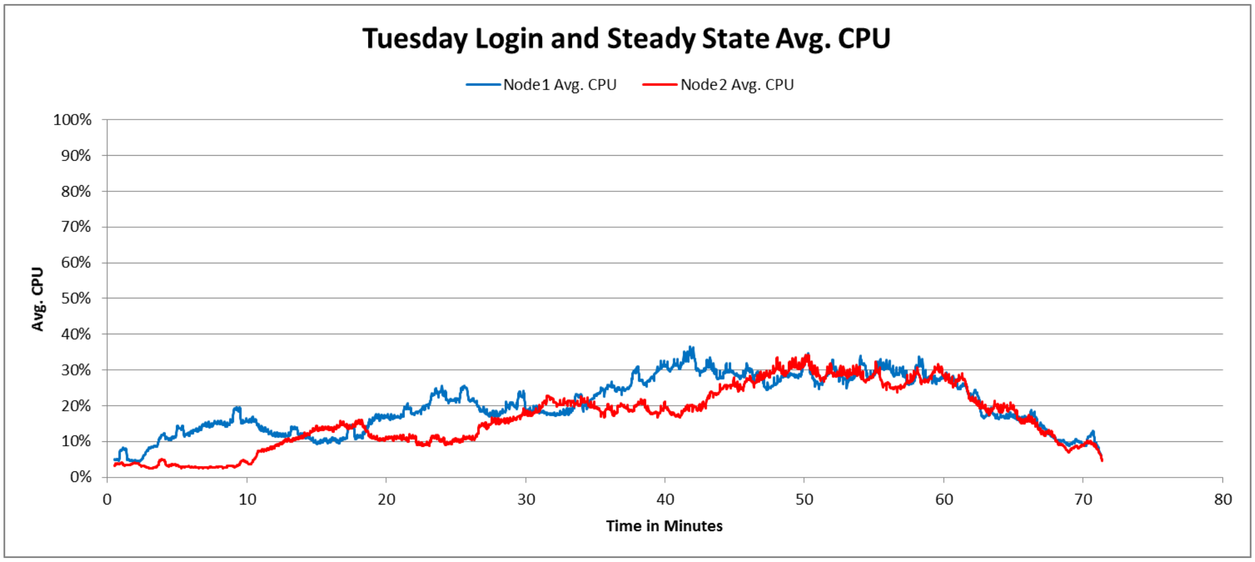
Figure 31) Throughput, IOPS, and latency for full-clone Tuesday morning login and workload.



Storage Controller CPU Utilization

Figure 32 shows storage controller CPU utilization for full-clone Tuesday morning login and workload.

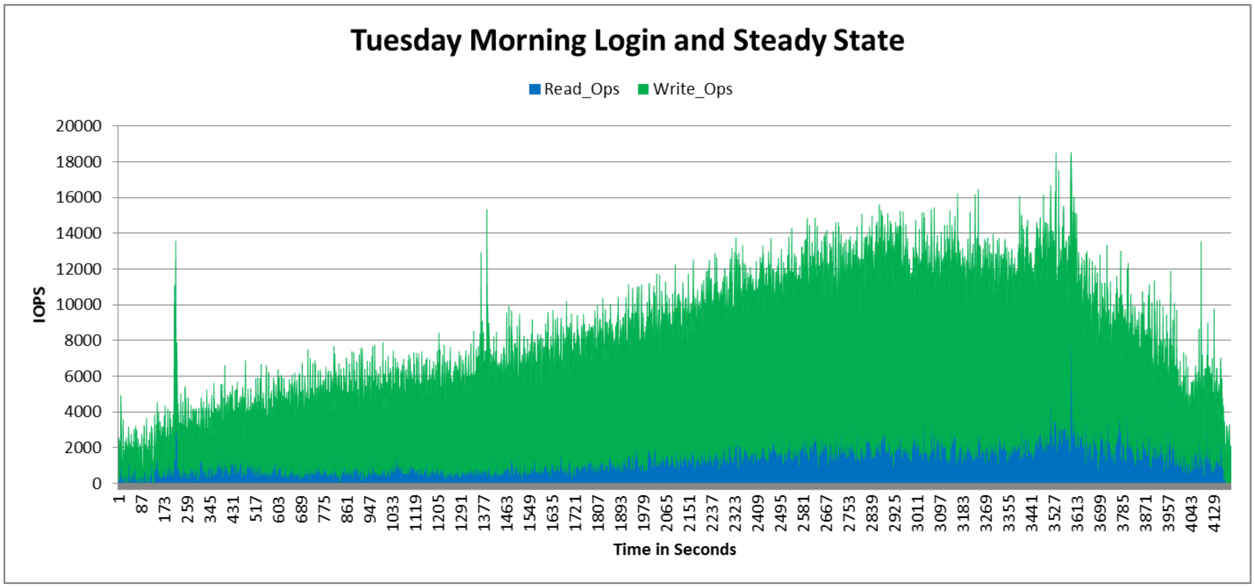
Figure 32) Storage controller CPU utilization for full-clone Tuesday morning login and workload.



Read/Write IOPS

Figure 33 shows read and write IOPS for full-clone Tuesday morning login and workload.

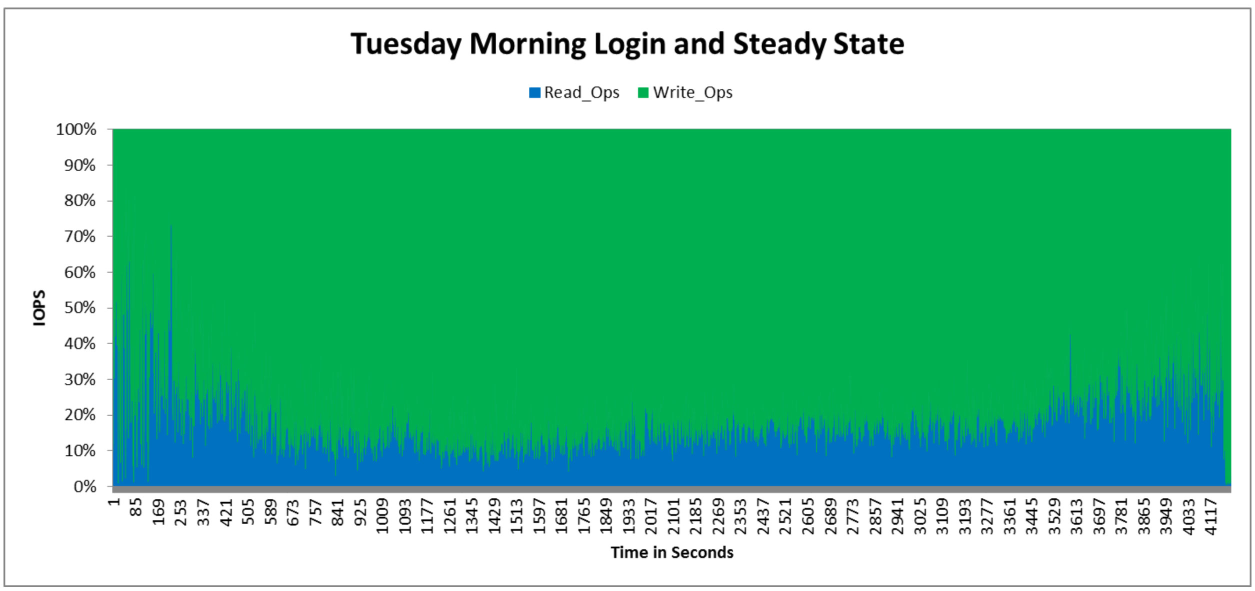
Figure 33) Read/write IOPS for full-clone Tuesday morning login and workload.



Read/Write Ratio

Figure 34 shows read and write ratios for full-clone Tuesday morning login and workload.

Figure 34) Read/write ratio for full-clone Tuesday morning login and workload.



Customer Impact (Test Conclusions)

During the Tuesday morning login test, the storage controller performed very well. The CPU utilization was not high during this test, latencies were under 1ms, and desktop performance was excellent. These results suggest that it might be possible to double the storage controller workload to 3,000 users or more and still maintain excellent end-user performance. The Tuesday morning login during storage failover test described in the following section reinforces that point.

Tuesday Morning Login and Workload During Storage Failover Test

In this scenario, 1,500 users logged in to virtual desktops that had been logged into previously and that had not been power-cycled, and the storage controller was failed over. In this situation, VMs retained user and profile data, application binaries, and libraries in memory, which reduced the impact on storage. Table 12 lists the results for Tuesday morning login and workload during storage failover. See the Appendix: Workload Profiles for specific I/O profile data for this workload.

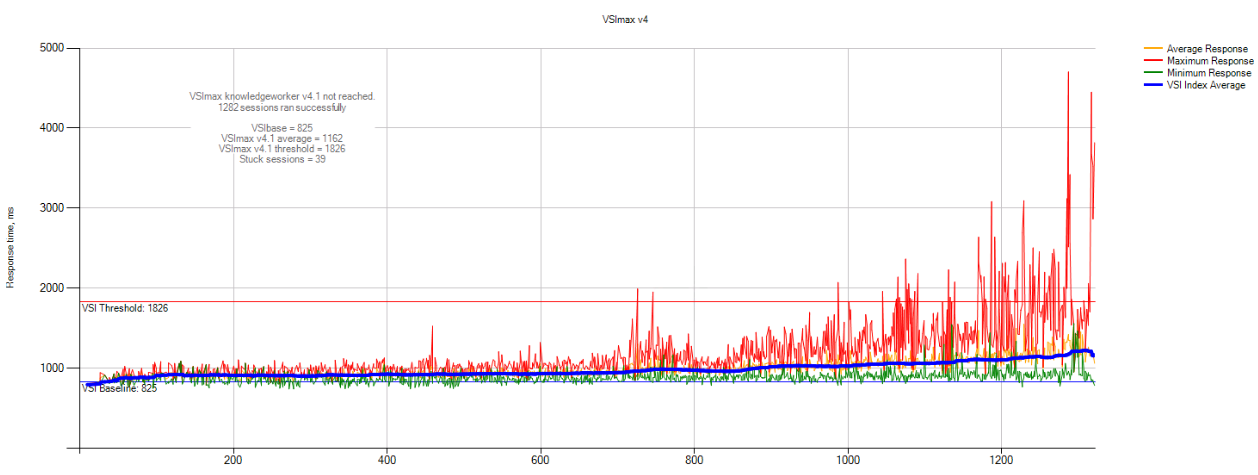
Table 12) Results for full-clone Tuesday morning login and workload during storage failover.

Desktop Login Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
6.65 min	0.66ms	26,187	6,942	0.5GBps	0.1GBps	96%	32%

Login VSI VSImax Results

Because the Login VSI VSImax v4.1 was not reached, more VMs could be deployed on this infrastructure. Figure 35 shows the VSImax results for Tuesday morning login and workload during storage failover.

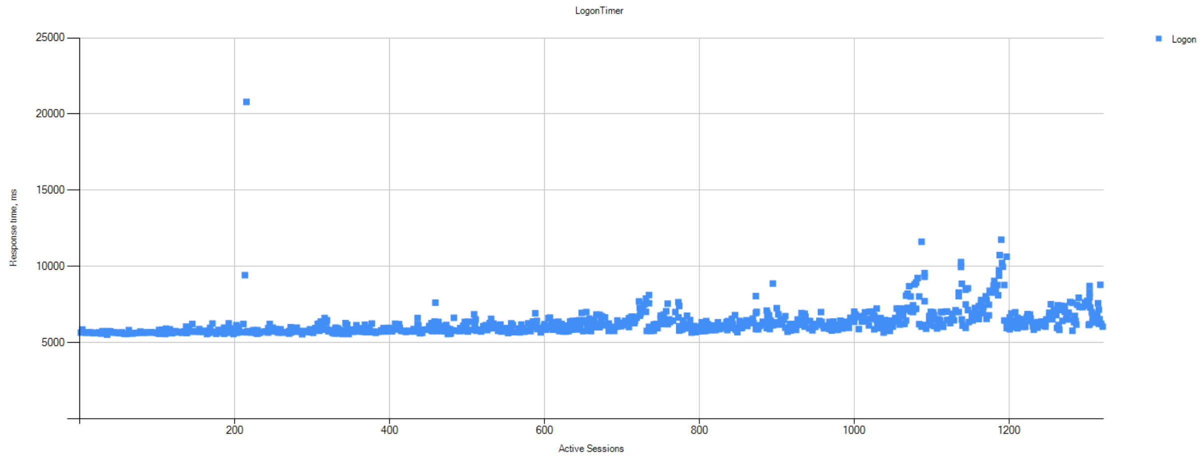
Figure 35) VSImax results for full-clone Tuesday morning login and workload during storage failover.



Desktop Login Time

Average desktop login time was 6.65 seconds, which is considered an excellent login time. Figure 36 shows a scatterplot of the Tuesday morning login times during storage failover.

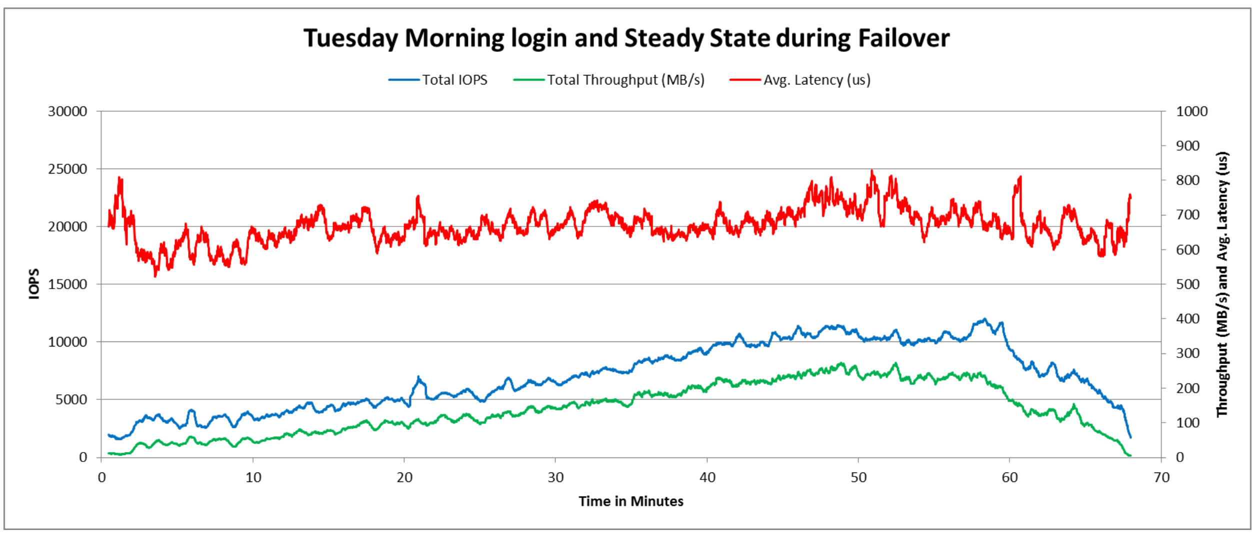
Figure 36) Scatterplot of full-clone Tuesday morning login times during storage failover.



Throughput, IOPS, and Latency

Figure 37 shows throughput, IOPS, and latency for a full-clone Tuesday morning login and workload during storage failover.

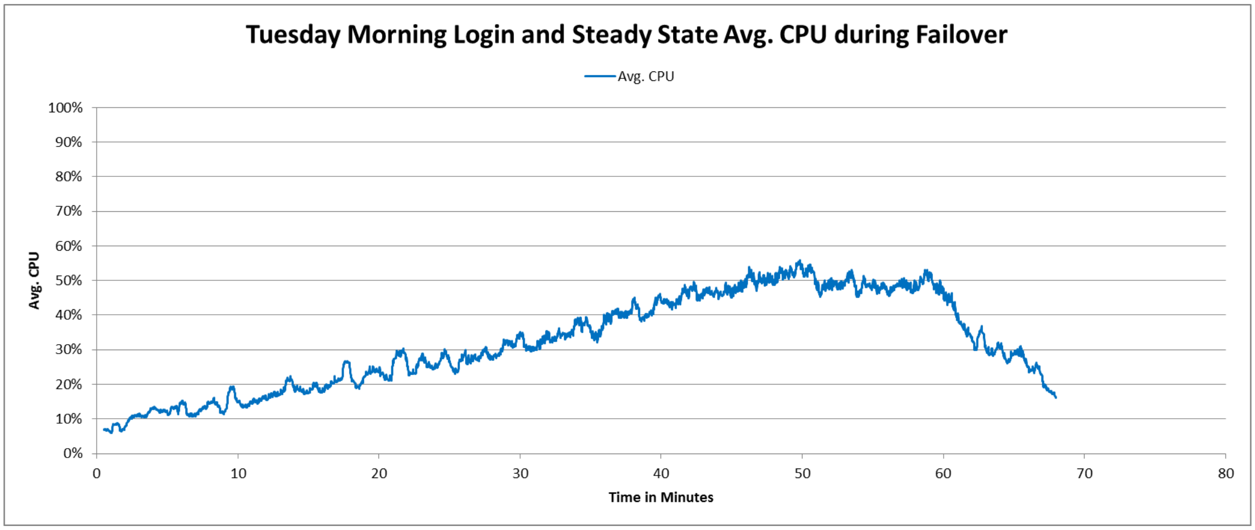
Figure 37) Throughput, IOPS, and latency for full-clone Tuesday morning login and workload during storage failover.



Storage Controller CPU Utilization

Figure 38 shows storage controller CPU utilization for a full-clone Tuesday morning login and workload during storage failover.

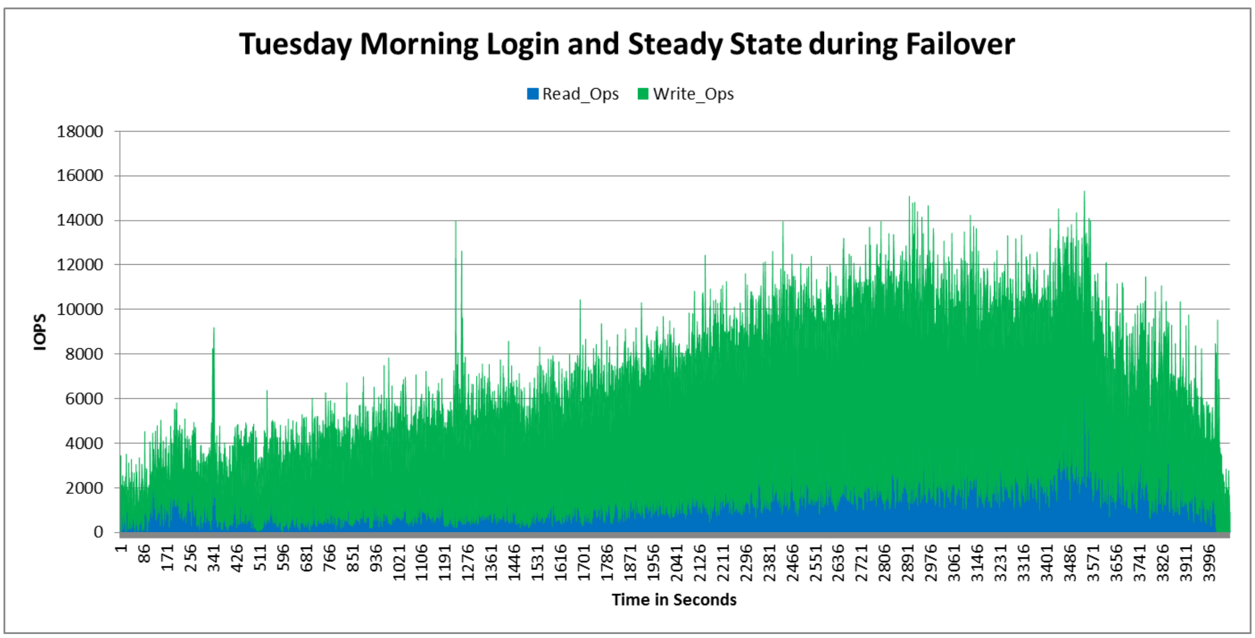
Figure 38) Storage controller CPU utilization for full-clone Tuesday morning login and workload during storage failover.



Read/Write IOPS

Figure 39 shows read and write IOPS for a full-clone Tuesday morning login and workload during storage failover.

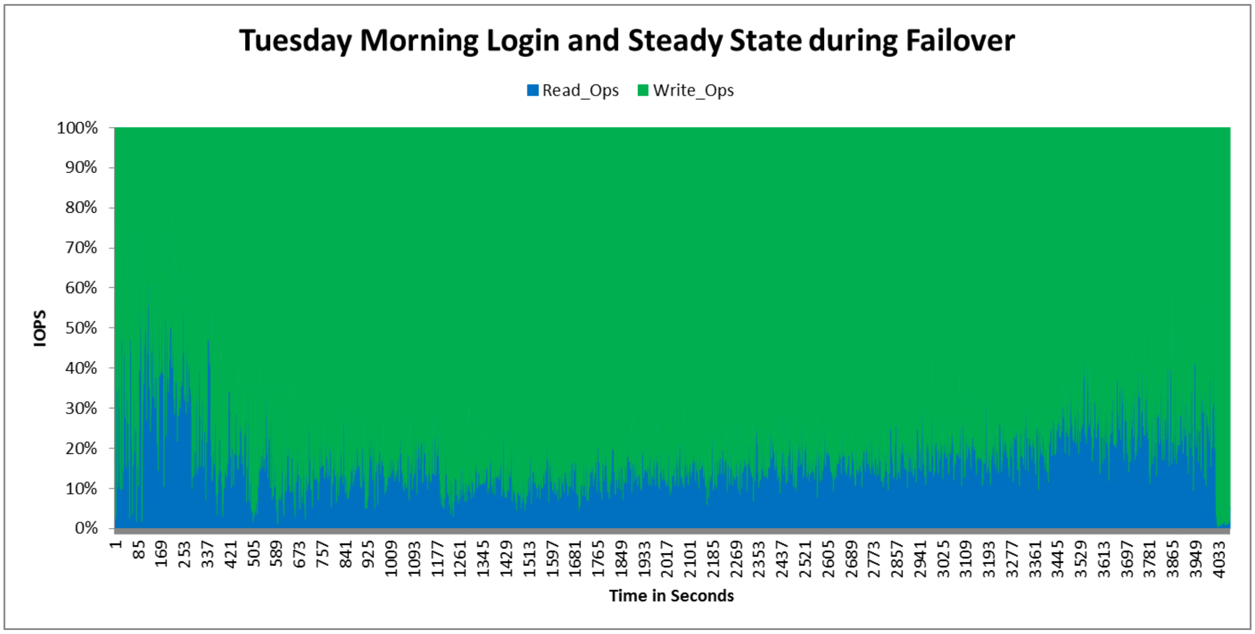
Figure 39) Read/write IOPS for full-clone Tuesday morning login and workload during storage failover.



Read/Write Ratio

Figure 40 shows the read to write ratio for a full-clone Tuesday morning login and workload during storage failover.

Figure 40) Read/write ratio for full-clone Tuesday morning login and workload during storage failover.



Customer Impact (Test Conclusions)

The purpose of this test was to demonstrate that an ordinary login and workload can be performed during a failover event. This is one of the easier workloads for the storage controller to perform.

6.6 Test for Patching 1,500 Desktops

This section describes test objectives and methodology and provides results from patch testing. See the Appendix: Workload Profiles for specific I/O profile data for this workload.

Test Objectives and Methodology

This test patched 1,500 desktops across both nodes of the storage infrastructure. To prevent the server host's CPU from becoming a bottleneck during this test, we inserted a delay of 10 seconds between each desktop's update start.

For testing, we used Windows Server Update Server (WSUS) to download and install patches to the 1,500 desktops. A total of approximately 900MB of patches was downloaded and installed on each machine. The patch update was initiated from a Windows PowerShell script that directed each VM to find available updates from the WSUS, apply the patches, and reboot the VMs. Table 13 lists the test results for patching 1,500 desktops on one node.

Table 13) Results for patching 1,500 persistent full clones on one node.

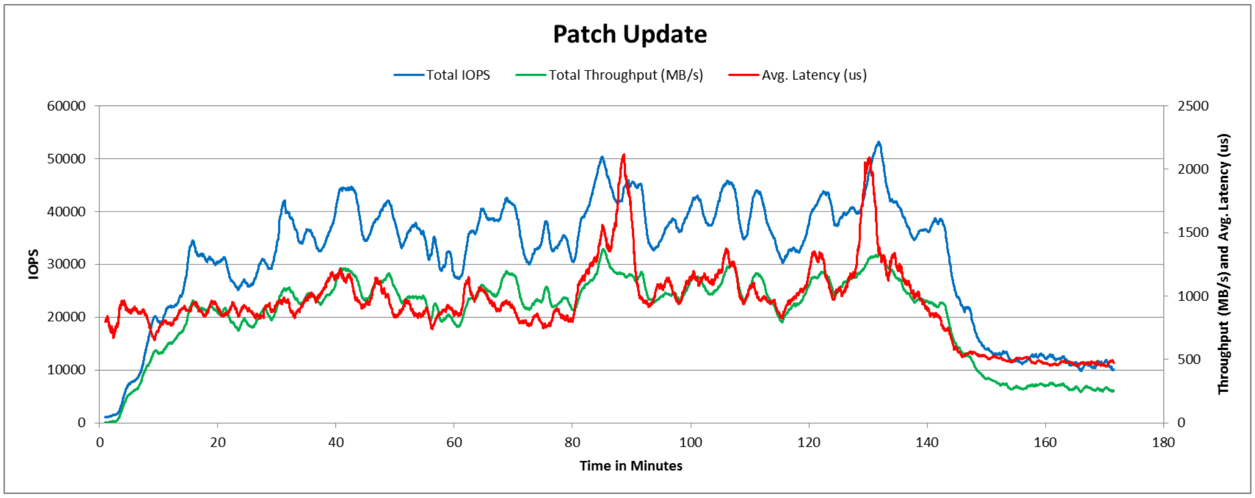
Time	Average Latency	Peak IOPS	Avg. IOPS	Peak Throughput	Avg. Throughput	Peak Storage CPU	Avg. Storage CPU
160 min	0.9ms	68,921	31,377	1.6GBps	0.9GBps	100%	60%

Note: CPU and latency measurements are based on one node of the cluster.

Throughput, IOPS, and Latency

Figure 41 shows throughput, IOPS, and latency for patching 1,500 persistent full clones.

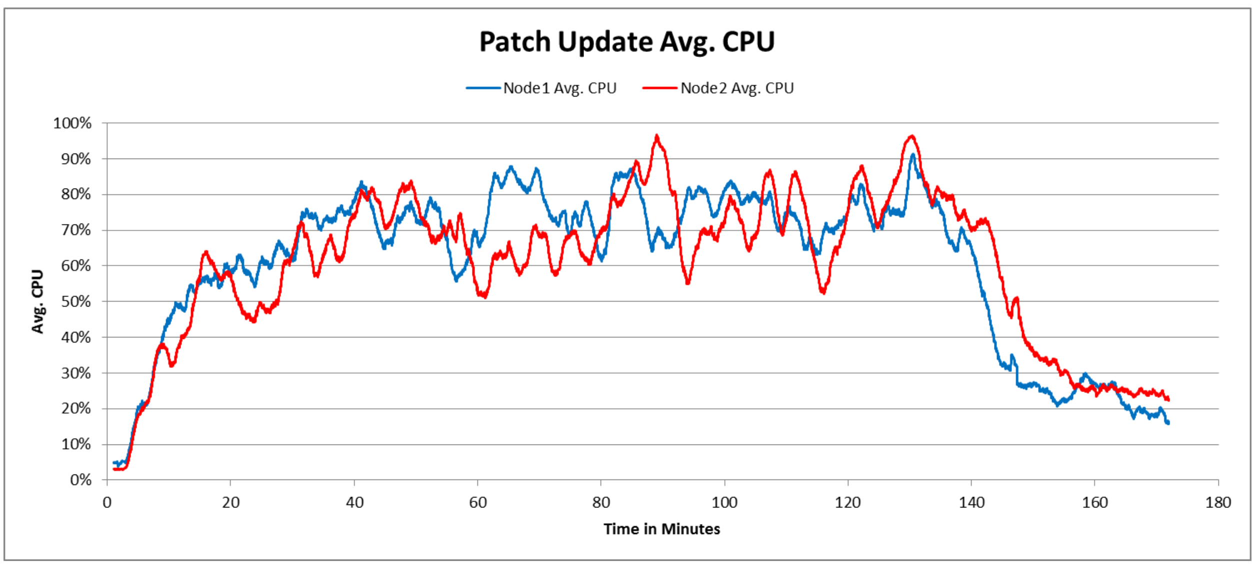
Figure 41) Throughput, IOPS, and latency for patching 1,500 persistent full clones.



Storage Controller CPU Utilization

Figure 42 shows storage controller CPU utilization for patching 1,500 persistent full clones.

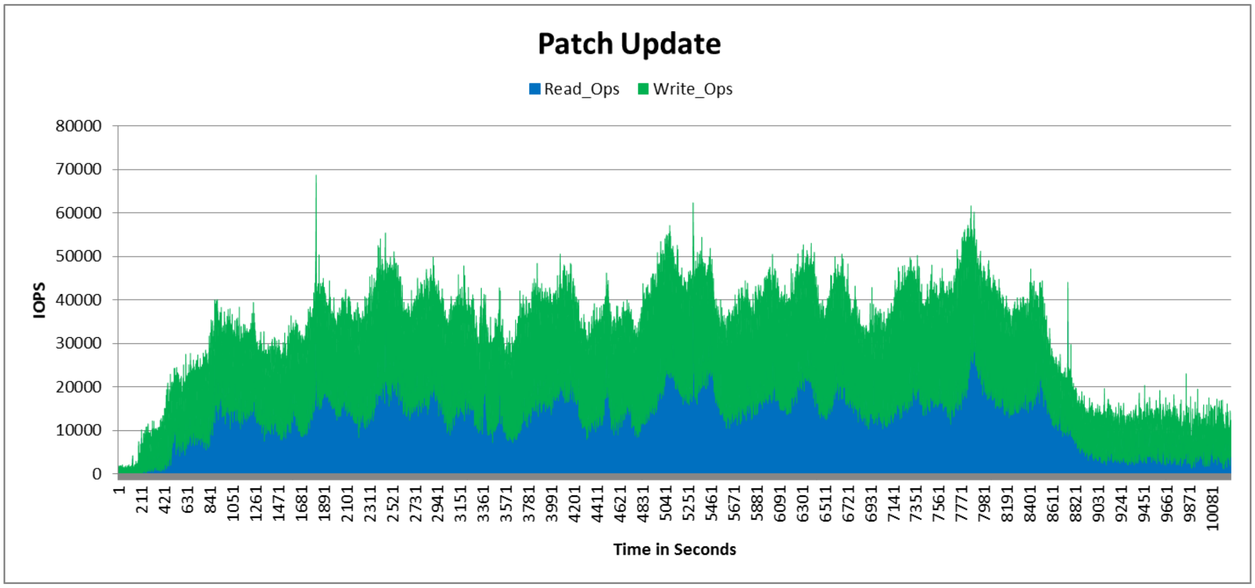
Figure 42) Storage controller CPU utilization for patching 1,500 persistent full clones.



Read/Write IOPS

Figure 43 shows read and write IOPS for patching 1,500 persistent full clones.

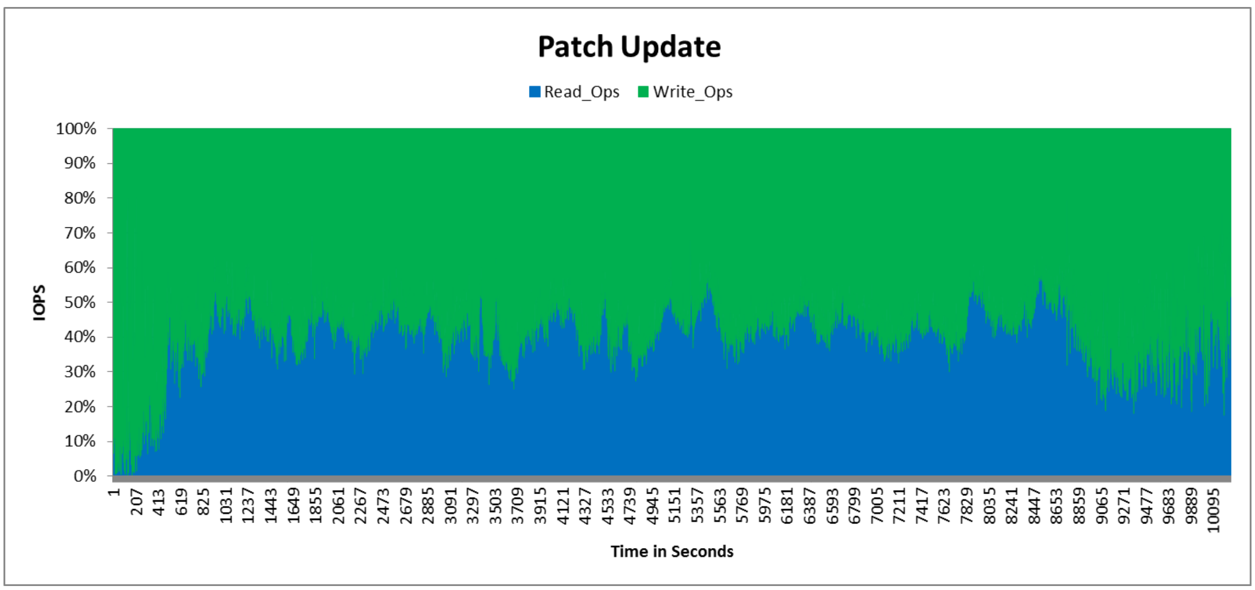
Figure 43) Read/write IOPS for patching 1,500 persistent full clones.



Read/Write Ratio

Figure 44 shows the read to write ratio for patching 1,500 persistent full clones.

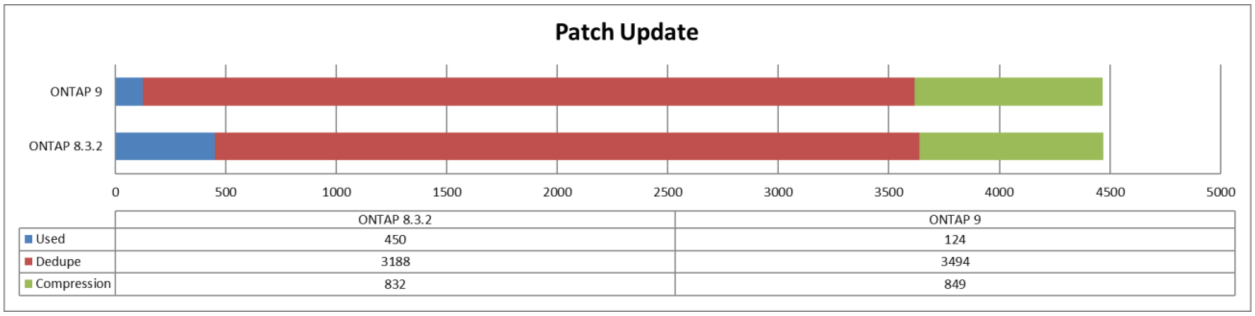
Figure 44) Read/write ratio for patching 1,500 persistent full clones.



Efficiency During Patch Update

For the patch update use case, we compared results from the previous version of ONTAP, Data ONTAP 8.3.2, and ONTAP 9. With Data ONTAP 8.3.2, it was necessary for postprocess deduplication to run in order to achieve maximum storage efficiency. Figure 45 shows that, with ONTAP 9, the inline deduplication algorithm has been significantly improved, and, along with compaction, postprocess deduplication did not give additional space savings in NetApp testing.

Figure 45) Savings between Data ONTAP 8.3.2 and ONTAP 9 (in GB) per FlexVol volume.



Customer Impact (Test Conclusions)

The patching of 1,500 virtual desktops with 900MB per VM took approximately 160 minutes to install and reboot the VM. Latency or CPU was not a concern during this test. In production environments, NetApp recommends staggering the patching over a longer period of time to reduce latency and CPU utilization.

7 Eight-Hour Test to Determine Performance Capacity

ONTAP 9 continuously monitors used performance capacity and performance capacity remaining to calculate the optimal operating point. The optimal point, as discussed in section 3.6, is the point at which the system is doing the most possible work at a given response time. It may be possible to drive a system utilization higher than the optimal point; however, doing so may come at a cost of increased latency or the inability to fail over and deliver the same consistent latency. During this monitoring, ONTAP monitors operations, optimal point operations, latency, optimal point latency, utilization, optimal point latency, and optimal point confidence factor. These point-in-time and calculated values can be seen for each controller by running the following command on the command line:

```
stat show -object resource_headroom_cpu -raw -counter ewma_hourly
```

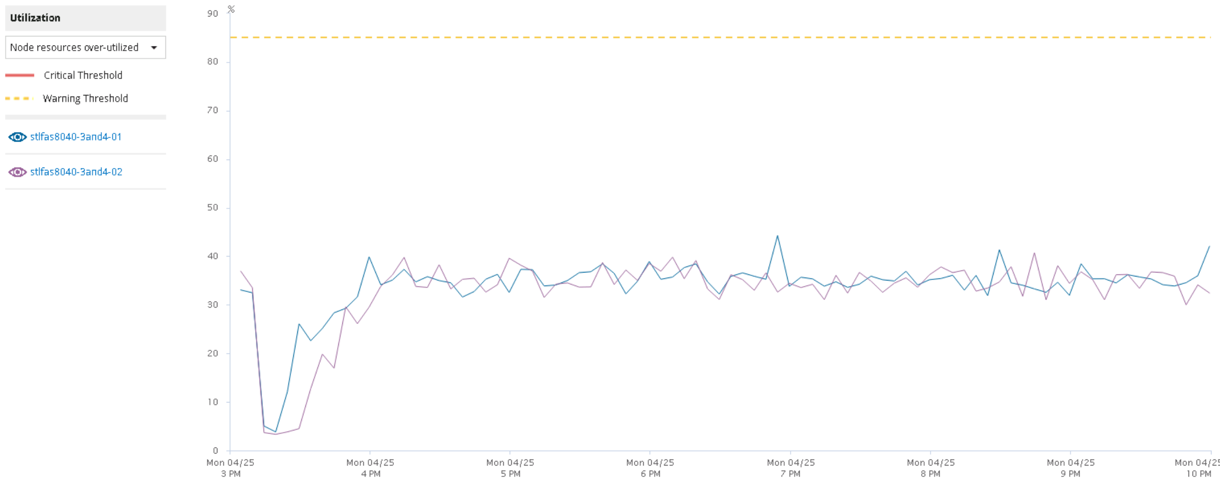
New in OnCommand Performance Manager 7 is a graphical display of system utilization, performance capacity, and remaining performance capacity. This display allows the user to graphically view the system and even estimate system performance during failover.

For this test, a Login VSI workload ran for 8 hours to determine steady-state performance and performance capacity. Each of the graphs tells a different story about the system, the workload, and the potential for additional work to be performed.

7.1 Eight-Hour Test System Utilization Graph

In this graph you can see that each node of the cluster averages around 35% utilization. This is the same number that you would see when issuing a `statistics show-periodic` from the command line.

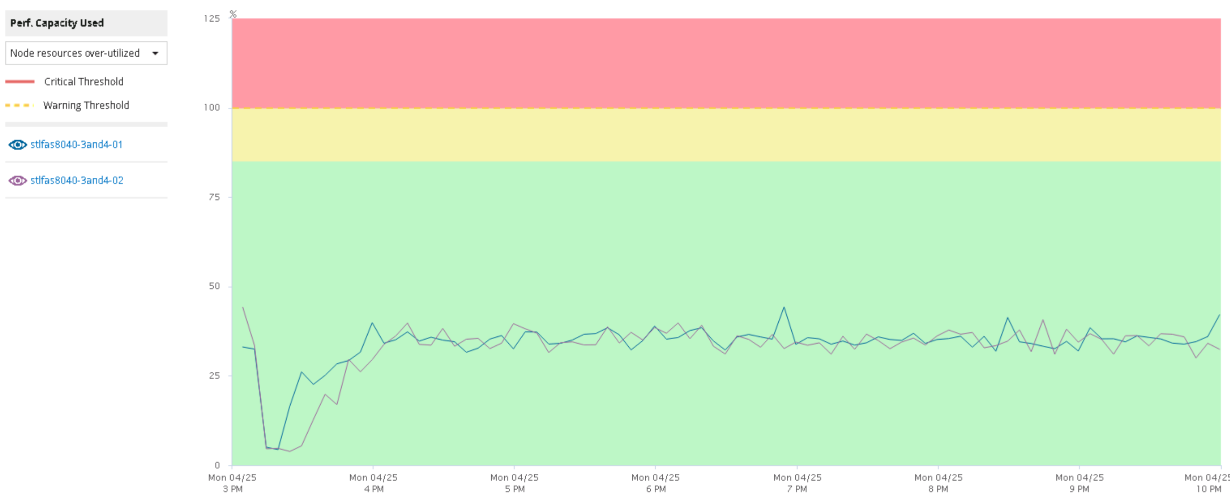
Figure 46) Eight-hour test system utilization graph.



7.2 Eight-Hour Test Performance Capacity Used

This graph is very similar to the system utilization in Figure 46, but performance capacity incorporates both utilization and latency when calculated. The performance capacity thresholds are also shown in this graph. This graph shows that the system is performing well and can safely host the 1,500 desktops that were provisioned.

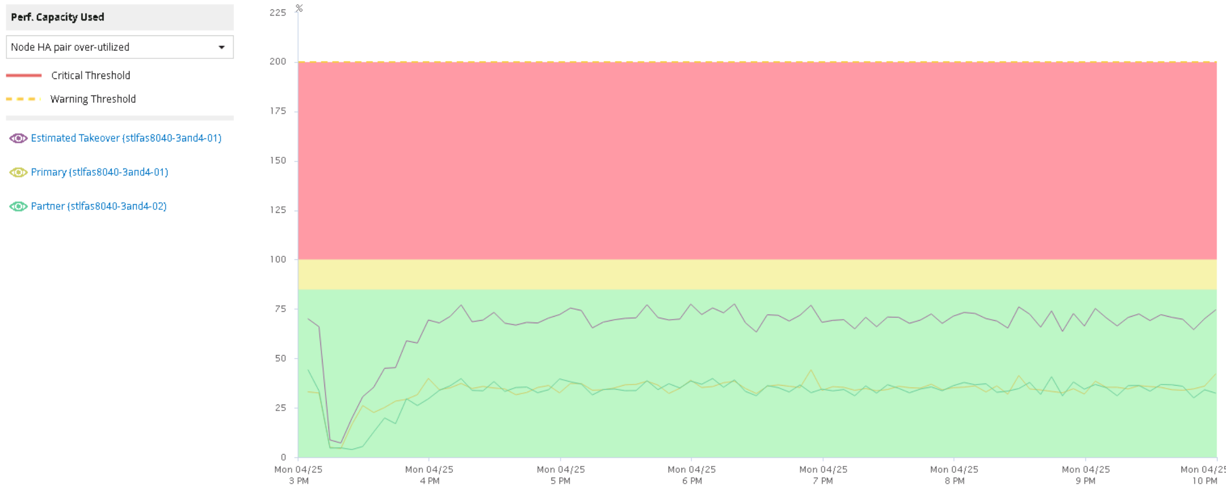
Figure 47) Eight-hour test performance capacity used graph.



7.3 Eight-Hour Test Performance Capacity Used Estimated Takeover

This graph is the most telling graph and adds estimated takeover performance to the preceding graph. It shows that given the current performance capacity of approximately 35%, takeover could be performed without overutilizing the system. The Estimated Takeover line shows that there is about 25% additional performance capacity. Using this estimation, an additional 375 desktops could be added to the 1,500-seat configuration without the system being overutilized for a total of 1,875 desktops.

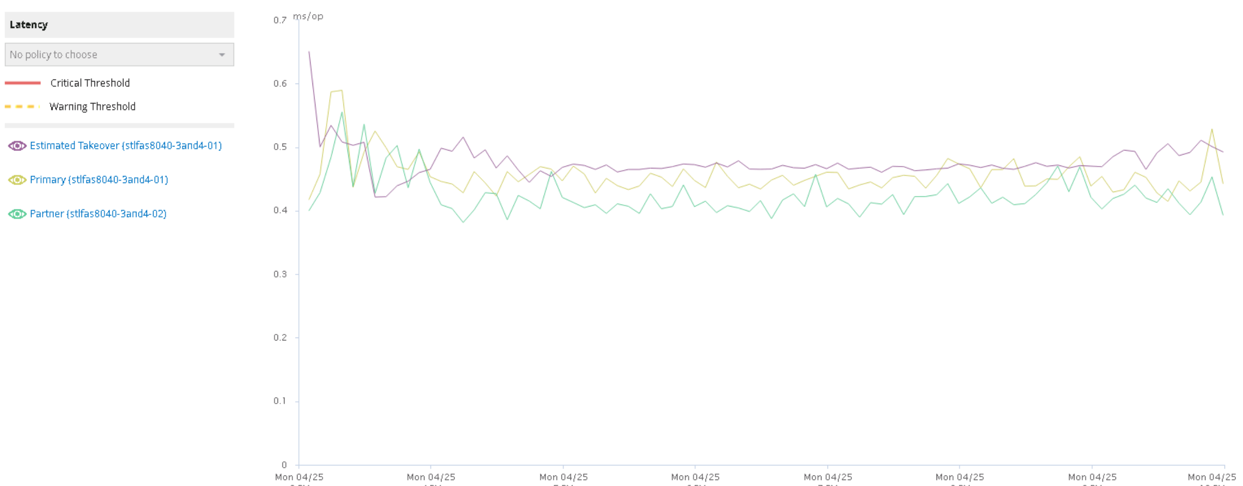
Figure 48) Eight-hour performance capacity used and estimated takeover graph.



7.4 Eight-Hour Test Latency

In Figure 48, approximately 25% or more performance capacity is available. What this graph does not show is the very low latency at that performance capacity point. Figure 49 shows that latency even at the estimated takeover line is less than 500µ. NetApp All Flash FAS with ONTAP 9 is able to deliver consistent latency that most other all-flash arrays (AFAs) in the market cannot. This suggests that even at 100% performance capacity during takeover, latencies could still be significantly less than 1ms. Therefore the system could potentially serve even more than the 1,875 VMs that were estimated earlier.

Figure 49) Eight-hour test latency and estimated latency during takeover.



Conclusion

In all desktop use cases, end-user login time, guest response time, and maintenance activities performance were excellent. The NetApp All Flash FAS system performed very well for the variety of tests. These tests, while important to demonstrating a proper end-to-end solution, do not represent a design that would take each individual component to the maximum. Thus, when designing a VDI deployment, proper sizing of each component should be performed with customer workloads in mind.

For persistent desktops, the following key findings were observed during the reference architecture testing:

- The NetApp All Flash FAS solution was able to very easily meet all IOPS requirements of the 1,500-user workload (boot, login, steady state, logout, patch storms) at an ultralow latency of approximately 1ms, delivering an excellent end-user experience.
- During all login and workload scenarios, the Login VSI VSImax was not reached.
- During boot storm testing, VMware vCenter did not throttle the boot process and produced an excellent boot time of 10 minutes for all 1,500 VMs.
- For all of the nonfailover tests, almost twice as many users could have been deployed with the same results. Only in the cases of the failed-over boot storm and initial login and workload did the CPU average over 50% of performance utilization.
- Inline deduplication, inline compression, and FlexClone technology storage efficiency saved over 24TB of storage, which translates into a savings of 50:1, or 98%.
- During the 8-hour testing to determine performance capacity, it was determined that this architecture could very easily deliver 1,875 desktops at 500µ and be able to sustain that performance during a storage failover event. Internal stress testing has shown that the AFF8040 is capable of supporting approximately 2,600 desktops at ~1ms latency during failover.

Appendix: Workload Profiles

In this section we break down the different I/O profiles for each of the different workloads in this paper. These I/O profiles were seen by vscsiStats (https://communities.vmware.com/docs/DOC-10095) at the ESXi layer. This information demonstrates that each workload is different, and each workload affects the storage controller differently.

These workload profiles can be used across different platforms as standard workload for proof of concept. Many vendors just use a single block size to replicate what VDI is. As you can see from the following profiles, there are many different I/O sizes and read/write ratios during each of the test cases. Using tools such as Login VSI allows customers to generate loads that are more representative than a single I/O size. However, if an existing VDI deployment is in use, vscsiStats can be captured and replayed to match the storage workload closely. What this methodology does not capture are the real data and how factors such as deduplication, compression, and compaction affect the workload and storage efficiency. Nor does this capture the end-to-end server and network impact of the workloads. This is where tools such as Login VSI can bring reality to a given workload.

Boot Storm

Here is a workload profile that matches what was observed from the ESXi layer.

Block Size (KB)	% Read	% Random	Workload Percentage
512	85%	100%	2%
1,024	94%	100%	2%
2,048	95%	100%	4%
4,096	77%	100%	40%
8,192	94%	100%	7%
16,384	98%	100%	18%

Block Size (KB)	% Read	% Random	Workload Percentage
32,768	99%	1%	26%
>=65,536	91%	1%	1%

Monday Morning Login and Workload

Here is a workload profile that matches what was observed from the ESXi layer.

Block Size (KB)	% Read	% Random	Workload Percentage
512	10%	100%	3%
1,024	19%	100%	2%
2,048	63%	100%	1%
4,096	19%	100%	42%
8,192	38%	100%	8%
16,384	87%	100%	16%
32,768	96%	50%	22%
49,152	15%	1%	3%
131,072	12%	1%	1%
262,144	21%	1%	1%
524,288	3%	1%	1%

Tuesday Morning Login and Workload

Here is a workload profile that matches what was observed from the ESXi layer.

Block Size (KB)	% Read	% Random	Workload Percentage
512	5%	100%	7%
1,024	4%	100%	4%
2,048	13%	100%	1%

Block Size (KB)	% Read	% Random	Workload Percentage
4,096	1%	100%	64%
8,192	2%	100%	8%
16,384	5%	100%	4%
32,768	60%	1%	4%
49,152	0%	1%	4%
65,536	18%	1%	1%
131,072	4%	1%	1%
262,144	3%	1%	1%
524,288	0%	1%	1%

Patching Desktops

Here is a workload profile that matches what was observed from the ESXi layer.

Block Size (KB)	% Read	% Random	Workload Percentage
512	2%	100%	2%
1,024	2%	100%	4%
2,048	4%	100%	3%
4,096	15%	100%	52%
8,192	8%	100%	8%
16,384	10%	100%	7%
32,768	17%	1%	6%
49,152	1%	1%	3%
65,536	5%	1%	3%
81,920	1%	1%	1%

Block Size (KB)	% Read	% Random	Workload Percentage
131,072	1%	1%	2%
262,144	1%	1%	5%
524,288	1%	1%	4%

References

The following references were used in this technical report:

- NetApp All Flash FAS Overview Data ONTAP 8.3.2
  http://www.netapp.com/us/media/tr-4505.pdf
- VMware vSphere 6 on NetApp Clustered Data ONTAP
  http://www.netapp.com/us/media/tr-4333.pdf
- TR-4138: Citrix XenDesktop on NetApp Storage Solution Guide
  http://www.netapp.com/as/media/tr-4138.pdf
- TR-4342: NetApp All Flash FAS Solution for Persistent and Nonpersistent Desktops with Citrix XenDesktop and XenApp
  http://www.netapp.com/us/media/tr-4342.pdf
- NetApp All Flash FAS Datasheet
  http://www.netapp.com/us/media/ds-3582.pdf
- Citrix XenDesktop 7.6 Blueprint
  http://www.citrix.com/content/dam/citrix/en_us/documents/products-solutions/xendesktop-deployment-blueprint.pdf?accessmode=direct
- Citrix Product Document on XenApp 7.6 and XenDesktop 7.6
  http://docs.citrix.com/en-us/xenapp-and-xendesktop/7-6.html
- FlexPod Datacenter with Citrix XenDesktop/XenApp 7.7 and VMware vSphere 6.0 for 5000 Seats
  http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/cisco_ucs_xd77esxi60u1_flexpod.html

Acknowledgements

The authors thank Eric Lacell, Scott Lane, Michael Peppers, and Rob McDonald for their contributions to this document.

Version History

Version	Date	Document Version History
Version 1.0	June 2016	Original document
Version 1.1	September 2016	Added NFS.MaxQueueDepth recommendations, performance capacity and AFBP, and 8-hour test determining performance capacity and workload profiles.

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

TR-4519-0916