SAN Acceleration Using Nexenta Connect™ View Edition with Third-Party SAN Storage

NEXENTA OFFICE OF CTO
ILYA GRAFUTKO

Table of Contents

VDI Performance 3

NexentaConnect and Storage Attached Network 3

Getting Started 4

EMC SAN over 8Gb FC Interconnect: Test Methodology and Results 6

- SAN I/O Offload – Provisioning 8
- SAN Acceleration – Steady-State Workloads 11
- User Experience 13
- Login VSI – Desktop Density with and without NexentaConnect 14

NexentaStor SAN over 10GE iSCSI: Test Methodology 15

- SAN Offload – Provisioning and Recompose 15

Why These Results Matter 17

Conclusion 17

VDI Performance

For years, storage has been a major challenge for VDI propagation and an obstacle to deployments — and this challenge continues today. Several factors contribute to increased requirements for backend enterprise-grade SAN. These include highly bursty, write-intensive workloads; small blocksize; random I/O generated simultaneously by multiple virtual desktops; and boot and login storms. To meet user expectations for performance and latency (response time) of virtual desktops (the user experience must be the same as or better than that of conventional desktops and mobile devices), many storage administrators employ an outdated approach: overprovisioning the SAN for all peak conditions and investing in backend storage at a, prohibitively high premium.

Nexenta's Virtual Storage Appliance (VSA) provides I/O acceleration to offload a significant portion of the I/O from the SAN (physical) back end, thus eliminating the need to overprovision.

The principle objective of this white paper is to measure and fully qualify the SAN acceleration capabilities of the NexentaConnect™ for View Edition (NexentaConnect) with multiple SAN storage arrays on the back end. This paper is organized into two parts.

The first part documents the performance results obtained at the VMware Lab in collaboration with End-User Computing (EUC) at VMware. The results (listed in detail and analyzed below) include significant IO offloading from the physical SAN during provisioning I/O storms, 5x reduction in the SAN load during periods of extreme stress, improved I/O latency and responsiveness in user sessions, and increased desktop density levels under Login VSI benchmarks.

The second part of the paper shows the performance runs and data obtained at Nexenta's own lab in Santa Clara, CA and further confirms the SAN offloading capabilities of the product under provisioning storms, massive recompose operations, and Login VSI medium-workload, industry-standard tests.

NexentaConnect and Storage Attached Network

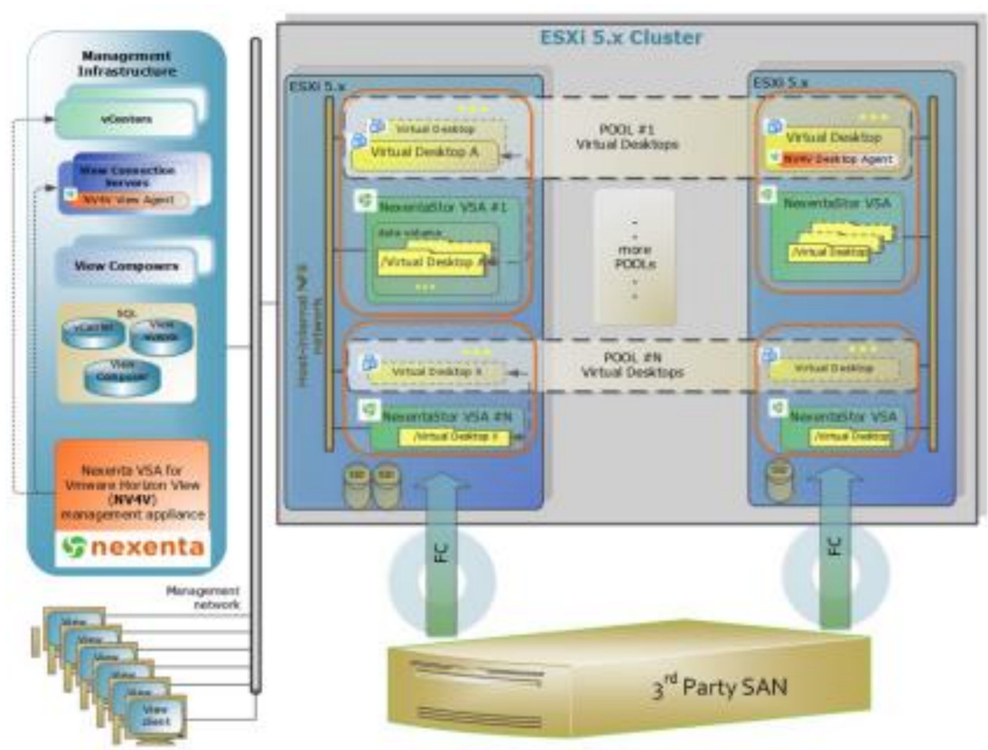
NexentaConnect, currently at version 2.1.1, combines a powerful automation engine and an enterprise-grade virtual storage appliance (VSA). NexentaConnect provides storage and SAN acceleration by leveraging ZFS and offering a range of features that benefit VDI. These features include inline compression and inline data deduplication, write journaling/logging to low-latency flash devices (SSDs), advanced level-1 (RAM), and level-2 (SSD) caching.

The NexentaConnect virtual storage appliance (VSA) is a fully featured, enterprise grade NAS server that is fine-tuned specifically for VDI workloads. NexentaConnect is also a sophisticated VMware Horizon View provisioning tool that auto-provisions the storage layer (or VSA) specifically for VDI, based on user-specified parameters including:

- Type of desktop pool deployed (stateless or persistent, linked or full clone)
- Local storage vs. remote SAN via iSCSI or FC
- Additional user options selected via the NexentaConnect Deployment Wizard

In addition to these features, NexentaConnect includes built-in performance benchmarks and calibration capabilities (accessed via easy-to-use GUI wizards) that allow VMware Horizon View administrators to test the resulting pool of virtual desktops and establish the actual performance level of users.

Figure 1: High-Level Architecture of NexentaConnect and FC/ISCSI SAN.



Getting Started

You can download NexentaConnect at http://nexenta.com/vdi. The product includes the following virtual machine (OVF) templates:

- NexentaConnect Management Appliance that provides centralized management GUI to administer and monitor multiple VDI environments (including multiple vCenter and View Connection servers, multiple ESXi clusters and multiple desktop pools, etc.)
- Nexenta Virtual Storage Appliance (VSA) – enterprise-grade NexentaStor fine-tuned and customized specifically for the VDI environment.

During testing, the VMware management environment was running VMware ESXi™ 5.0u1, VMware vSphere® 5.1U1, VMware Horizon View 5.2, Windows 7 x64 virtual machines, and the NexentaConnect Management Appliance version 2.1.1.

The VMware Lab environment specifications are detailed in Table 1.

Table 1: VMware Lab Test Environment Specifications.

COMPONENT	DESCRIPTION
Server Specs – Single Host Test, Unconstrained SAN	Single Host – Dell PowerEdge R715m, 24 CPU cores x 2.3 GHz / 132 GB RAM
	Processor type – AMD Opteron™ Processor 6176 SE
	SAN – EMC VNX-5300
	Desktops – 80 virtual machines with 1 vCPU / 1 GB RAM each
	Nexenta VSA – 2 vCPU / 12GB RAM

COMPONENT	DESCRIPTION
Server Specs – Cluster Test, Constrained SAN	Two (2) hosts – Dell PowerEdge R715m, 24 CPU cores x 2.3 GHz / 132 GB RAM
	Processor type – AMD Opteron Processor 6176 SE
	Two (2) SAN load-generator virtual machines – 2 vCPU / 4GB RAM • Iometer – 60/40 sequential/random, 70/30 write/read, 1 worker per virtual machine
	Nexenta VSA – 3 vCPU / 32GB RAM
Software	Nexenta VSA for VMware Horizon View (NV4V) 2.1.1
	VMware vSphere 5.0U1
	VMware Horizon View 5.2
	Login VSI 4
Storage Specs	SAN – EMC Storage Array VNX-5300 • RAID5 on top of 9 x 600 GB SAS HDD
	Nexenta VSA – local virtual storage appliance • Stripe (RAID0) on top of 2 TB FC LUN
	One (1) NV4V volume per VSA exported as NFS for the VDI deployment over a dedicated vSwitch (10GbE and Jumbo Frames enabled)

The Nexenta Lab management environment was running ESXi 5.0u1, vSphere 5.1U1, Horizon View 5.2, Windows 7 x64 virtual machines, and the NexentaConnect Management Appliance version 2.1.1.

The Nexenta Lab environment specifications are detailed in Table 2.

Table 2: Nexenta Lab Test Environment Specifications.

COMPONENT	DESCRIPTION
Server Specs – Cluster Test, Constrained SAN	Two (2) hosts – Supermicro X8DTH-i/6/iF/6F,
	12 CPU cores x 2.4 GHz / 196 GB RAM
	Processor type – Intel Xeon E5645
Software	Nexenta VSA for VMware Horizon View (NV4V) 2.1.1
	VMware vSphere 5.0U1
	VMware Horizon View 5.2
	Login VSI 3
Storage Specs	SAN – over 10GbE iSCSI – NexentaStor on top of Supermicro • RAID0 on top of 10 x 300 GB SAS drives
	Nexenta VSA – local virtual storage appliance • Stripe (RAID0) on top of 2 TB FC LUNs
	One (1) NV4V volume per VSA exported as NFS for the VDI deployment over a dedicated vSwitch (10GbE and Jumbo Frames enabled)

EMC SAN over 8Gb FC Interconnect: Test Methodology and Results

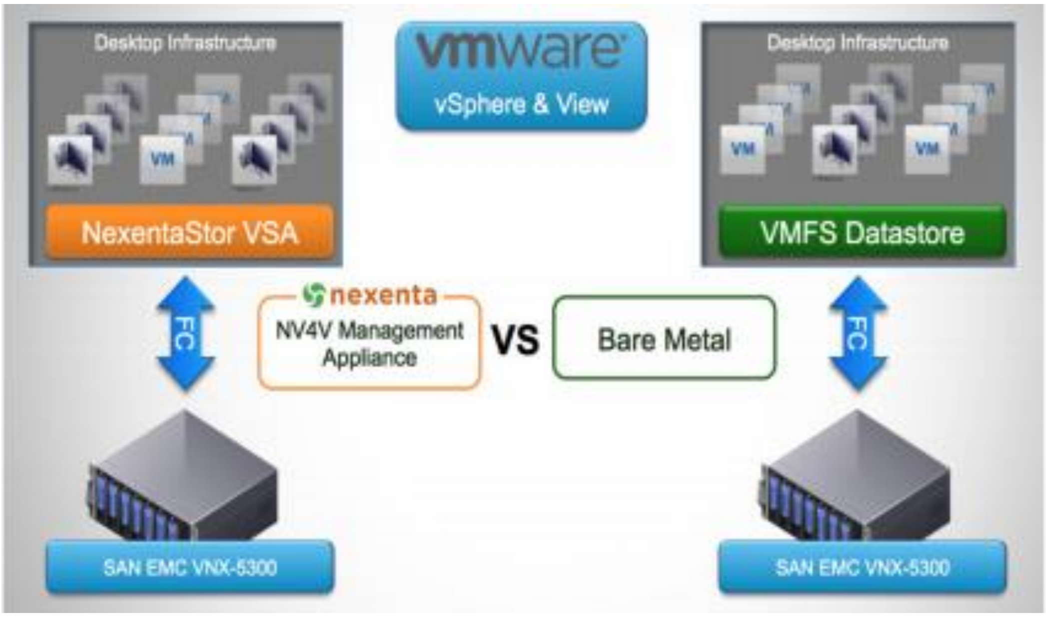
The VMware Lab test goals were to examine the performance and user experience of NexentaConnect vs. bare-metal VDI deployment on the EMC SAN, and to answer this question:

"Why would a Nexenta VSA that is obviously taking a chunk of the host's local (CPU and memory) resources for itself — resources that could be utilized in other ways by the virtual desktops — provide any performance benefits?"

The question may seem counterintuitive at first; as we take a deeper dive into the performance data the answer becomes apparent.

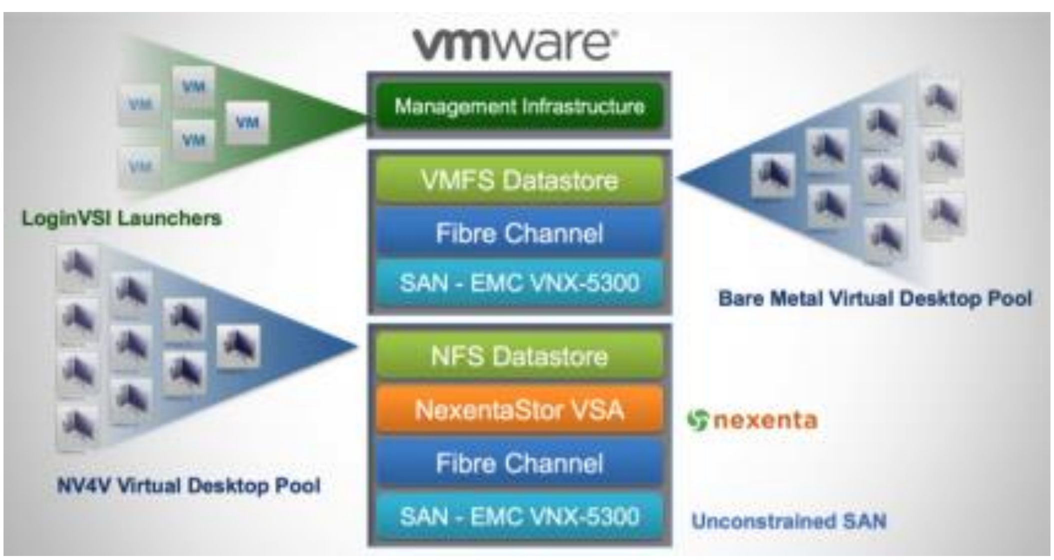
The lab started out by setting a baseline on top of a typical bare-metal VDI deployment. For this typical deployment, a third-party storage array serves as a back-end server and provides LUNs over FC to the ESXi cluster. The virtual machine desktop pool sits on top of the VMFS datastores created from the FC LUNs. The results from this baseline were compared to a NexentaConnect VDI deployment, where, similarly, a third-party storage array serves as a back-end server and provides LUNs over FC to the ESXi cluster. However, unlike a bare-metal scenario, Nexenta VSA sits between the VMFS datastore and the pool of virtual desktops, by utilizing the FC storage and providing a local NFS share for the desktop pool as illustrated in Figure 2 below.

Figure 2: EMC VNX: NexentaConnect vs. Bare Metal.



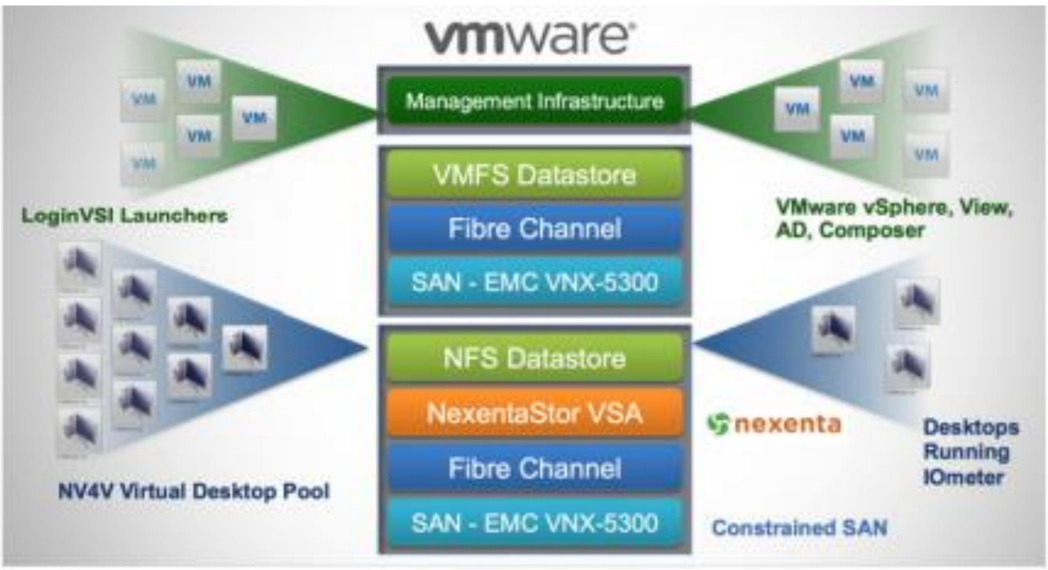
The VMware Lab tests for the Single-Host scenario focused on establishing the effect NexentaConnect had on the user experience when the SAN was performing at high levels and was not constrained by other competing workloads.

Figure 3: Single Host Lab Comparison of NexentaConnect vs. Bare Metal on Top of Unconstrained SAN.



The VMware Lab tests for the Cluster scenario (illustrated in Figure 4), consisting of two ESXi hosts, focused on establishing the impact of NexentaConnect on SAN that was heavily loaded by competing I/Os. These tests used several virtual machines running Iometer workloads that simulated a typical set of busy VDI users. The RAID group used on the back-end SAN used only 9 spindles, limiting the amount of IOPS available to roughly 8000 on the bare-metal baseline tests. The Iometer virtual machines (that is, Iometer clients running inside each virtual desktop) generated continuous high levels of I/O, consuming roughly 50% of the target LUN's available IOPS for the workload pattern used. Such conditions reflect real-world use, as VDI pilots expand more quickly than expected, or as users exert higher-than-anticipated load levels.

Figure 4: Cluster Lab Comparison of NexentaConnect vs. Bare Metal on Top of Constrained SAN.



SAN I/O Offload – Provisioning

The VMware View Composer™ provisioning operation generates very high levels of IO. The lab conducted a test provisioning 80 linked clone virtual machines on a single host to compare I/O levels passed down to the LUN on both the bare-metal and NexentaConnect configurations.

Figure 5 below illustrates the disk read/write throughput during the View Composer provisioning operation. The highlighted values represent the read/write rate averages over a one-hour period; because the actual provisioning cycle took half an hour, the figures have been adjusted accordingly. The read rate of 47 MBps and write rate of 19 MBps represent the adjusted bandwidth levels hitting the SAN during the actual provisioning activities.

The test ran with Content Based Read Cache (CBRC) both enabled and disabled. It was observed that CBRC helped to dramatically reduce read I/O by 50% on the bare-metal configuration, but offered minimal incremental value when used in tandem with NexentaConnect.

Figure 5: Provisioning Bare Metal without CBRC.

In contrast to the read I/O rates of 47 MBps achieved on the bare-metal configuration, the Nexenta NexentaConnect showed a 38x offload, passing only 1.2 MBps down to the physical LUN (Figure 6). These results demonstrate the superior read caching offered by the ZFS-based virtual appliance. Even compared to the CBRC-enabled, bare-metal configuration, the Nexenta NexentaConnect still offers roughly 19x offload.

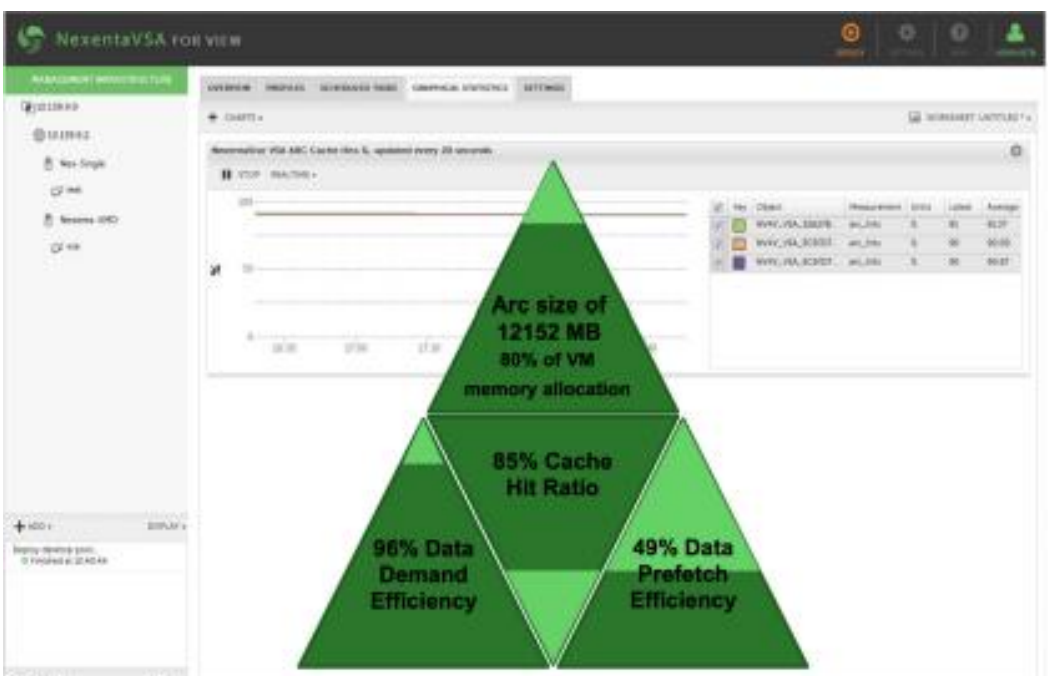
On the write side, we observed 5.7 MBps or a 2.5x write I/O offload from the SAN. This represents a 60% reduction in writes as compared to the bare-metal configuration. NexentaConnect uses Lempel-Ziv family LZJB lossless data compression, which replaces repeated occurrences of data with references to a single copy and improves I/O throughput. The deployment of 80 linked-clone desktops revealed a compression rate of 1.86x, resulting in a partial write performance boost.

Figure 6: Provisioning NexentaConnect without CBR

C.

Read performance is directly related to caching efficiency. The NexentaConnect uses Adaptive Replacement Cache (ARC) technology that allows it to aggressively cache, proving efficient in a VDI environment. Figure 7 illustrates that ARC is heavily utilized, using 12 GB or 80% of RAM. The figure also shows a very high hit ratio of 85%, implying that over three-fourths of our data requests are already in memory. Additionally, 50% of the data is pre-fetched or already in the cache, and the Data Demand results show that there is 96% chance that requests will hit the data already read.

Figure 7: NexentaConnect VSA Caching Efficiency.



SAN Acceleration – Steady-State Workloads

Table 3 summarizes the I/O throughputs obtained during Login VSI tests. The Bare-Metal LUN column shows the load levels on the physical disk during testing without the NexentaConnect acceleration layer. The bare-metal testing revealed instability and data loss in multiple tests; some values were interpolated. The NexentaConnect LUN column shows the load levels on the physical disk during testing using the NexentaConnect acceleration layer. The NexentaConnect NFS Datastore column shows the I/O load levels observed on the NFS datastore published by the Nexenta VSA up to ESX during these same tests — these represent the levels of service provided to the virtual machines consuming storage, whereas the NexentaConnect LUN column shows the level of consumption of the physical SAN by the Nexenta VSA itself.

Table 3: I/O Throughout During Login VSI Tests.

I/O TO DISK	NV4V NFS DATASTORE	NV4V LUN	BARE-METAL LUN
Average Read	7.3 MBps / 350 IOPS	3MBps / 247 IOPS	0.86 MBps / 79 IOPS
Average Write	11.8 MBps / 403 IOPS	10.9 MBps / 188 IOPS	5.3 MBps / 259 IOPS
Adjusted Latency*	37 ms read / 143 ms write 22 ms read / 13 ms write	52 ms read / 32 ms write 56 ms read / 37 ms write	209 ms read / 629 ms write 47 ms read / 86 ms write

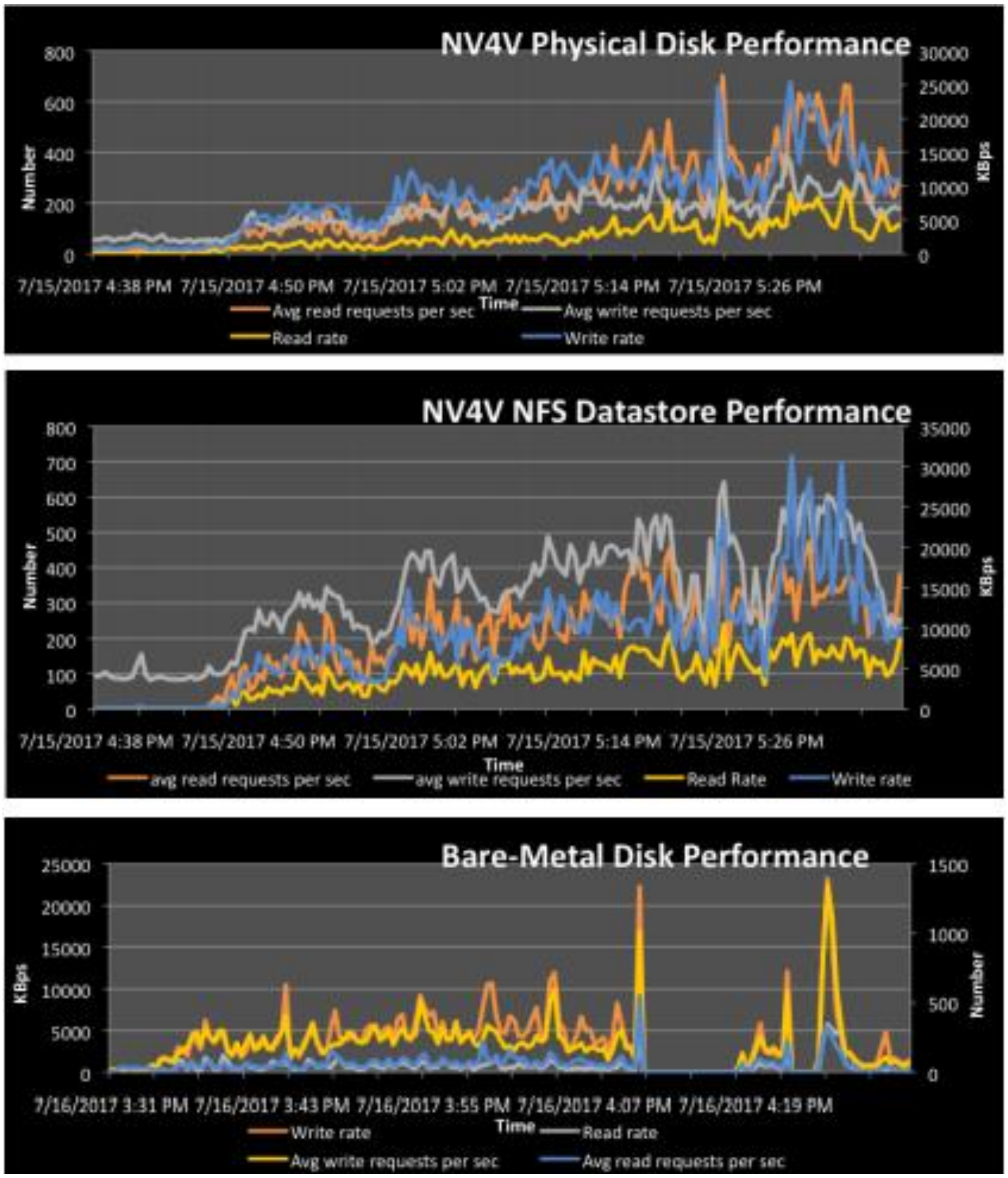
*Adjusted Latency represents the values before the Login VSImax was reached.

One notices immediately that the bare-metal configuration is only able to drive about one-third of the physical I/O which the NexentaConnect can generate. One reason is that NexentaConnect is able to read more blocks from the LUN under identical load conditions. NexentaConnect writes are all sequential writes because of the ZFS journaling file system. This leads to less random HDD activity, less 'seek time' penalty, and consequently, higher read I/O levels. The write efficiency is clearly reflected in the reduced write latency of the Nexenta LUN vs. the bare-metal LUN — 629ms vs. 32ms. Nexenta is driving more write bandwidth while consuming less IOPS because the write I/O is dramatically more efficient — driving 2x the write bandwidth but using 28% less IOPS to do so. Looked at another way, compare the bandwidth-to-IOPS ratios (0.057978 for NexentaConnect LUN vs. 0.020463 for bare metal). Bare metal is only 35% of the NexentaConnect ratio score, meaning that NexentaConnect has almost a 3x overall write efficiency.

The Nexenta NFS provides 7.3 MBps read I/O to the virtual machines while only passing along 3 MBps to the underlying LUN. This demonstrates that 85% of the read requests are being serviced from the ARC memory cache in the Nexenta VSA. In addition, Nexenta is storing data in a compressed format with a compression rate of 1.7 as reported by Nexenta VSA, such that the uncompressed data will deliver higher bandwidth from the NFS datastore than is actually retrieved from the SAN.

The bare-metal tests also generated higher CPU load levels on the ESX host, leading to an inability to retrieve Virtual Center performance statistics. The Bare-Metal Disk performance graph (bottom illustration, Figure 8) shows a gap in data retrieval starting at 4:07PM. The NexentaConnect tests show no such instability. In fact, the overall CPU levels on the host during tests are actually reduced, even though the Nexenta VSA itself is consuming CPU resources. As a result of the reduced I/O latency, virtual machines spend less time waiting for I/O, leading to fewer Operating System cycles spent waiting for the storage stack to respond. The overall result is an improved user experience and better host stability.

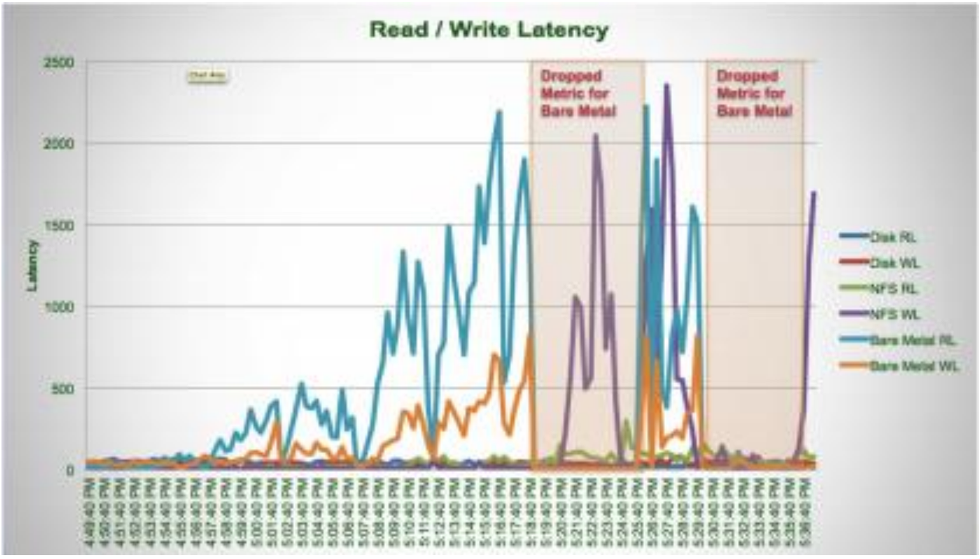
Figure 8: Read and Write Rate for Physical Disks and NFS Datastore as Reported by the Hypervisor.



Read/write latency was reported by the ESXi hypervisor. In the bare-metal scenario depicted in Figure 9, the light blue and orange lines represent the read write latency respectively. In the bare-metal data reveals an almost constant flow of latency spikes and, in some parts (5:20PM – 5:26PM), the fluctuations became so intense that the metric was dropped due to very high CPU consumption on the hypervisor. (See User Experience for more information.)

The NexentaConnect, through NFS and ZFS layers, absorbs the storm of latency spikes that we observe in the bare-metal scenario and provides a very smooth user experience.

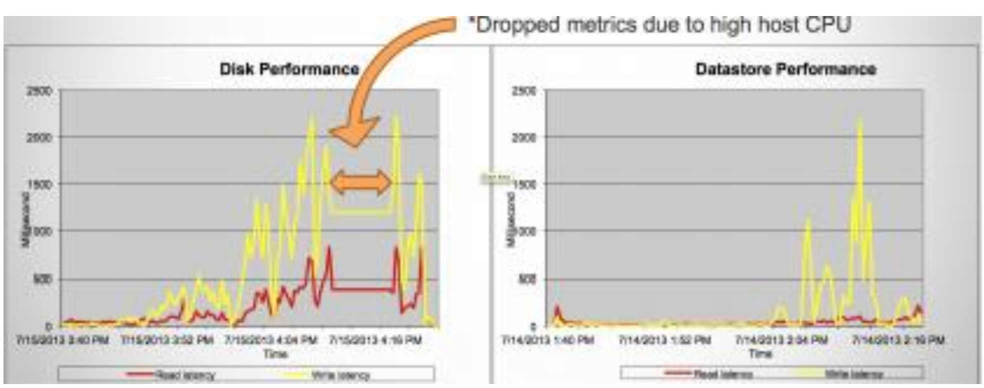
Figure 9: Read and Write Latency for Physical Disks and NFS Datastore as Reported by the Hypervisor.



User Experience

Figure 10 highlights the longer period of user satisfaction facilitated by the NexentaConnect acceleration layer. The bare-metal write I/O latency starts to ramp significantly early in the testing (see the left portion of the graph) as Login VSI adds users to the mix. By contrast, the NexentaConnect solution retains low write I/O levels much farther into the Login VSI test cycle, exhibiting elevated write I/O latencies only in the last one-third of the test cycle. This is roughly a 70% reduction in the length of time the user experience is impacted by adverse SAN conditions. Looked at another way, this represents a 70% reduction in trouble tickets called in by unhappy users due to a SAN load spike during production. NexentaConnect proves capable hiding an underperforming SAN from Horizon View users much longer than the bare-metal configuration.

Figure 10: Delayed Latency Increase Results in Better User Experience.

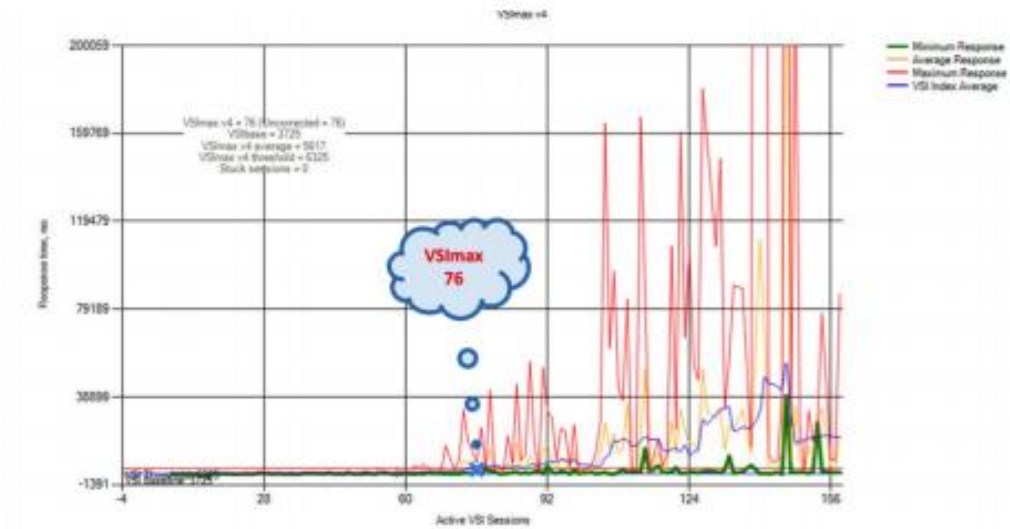


Login VSI – Desktop Density with and without NexentaConnect

Login VSI (www.loginvsi.com) is a de facto industry standard VDI benchmarking tool that validates application response times on various predefined workload options. This lab used Login VSI version 4, running a medium workload. The results measured by the VSImax value, representing the number of concurrent sessions running when the VDI environment reaches its saturation point.

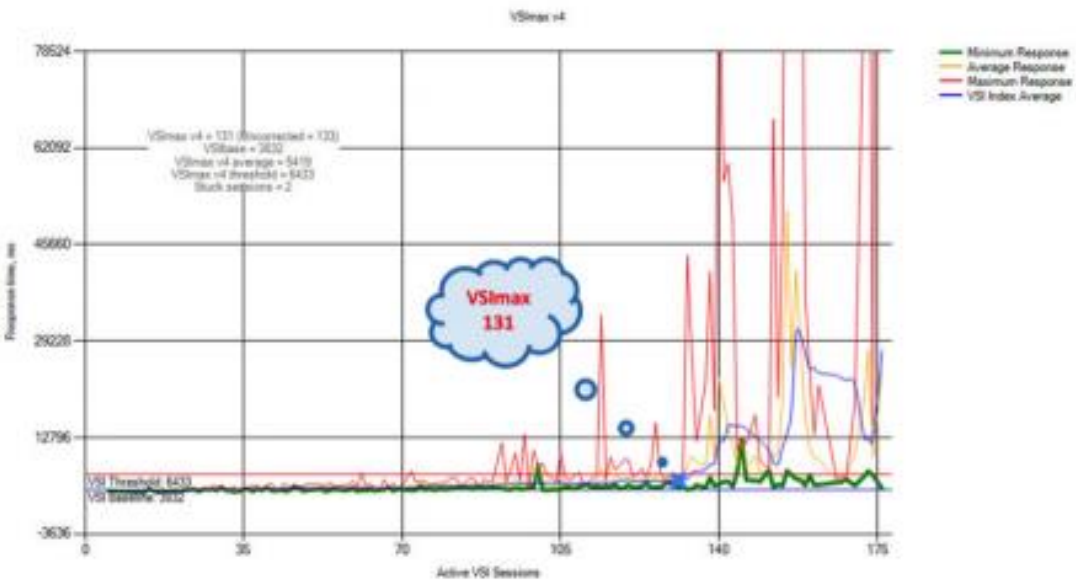
A VSImax score of 76 was all that the bare-metal configuration could achieve with the Iometer load generators running in the background (Figure 11). Without the load generators running, we observed VSImax scores in the 165 range.

Figure 11: Login VSI Results on Bare Metal



With NexentaConnect, VSImax scores as high as 131 were observed (Figure 12) — a 72% improvement over the VSImax score of 76 achieved by the bare-metal SAN without the Nexenta NexentaConnect acceleration layer. This VSImax score of 131 is roughly 80% of the 165 VSImax score achieved on the bare-metal SAN when no load generators were running. By contrast, with the addition of Iometer load generators running in the background the bare-metal configuration can only achieve 46% of its baseline, original performance. The VSA acceleration architecture proved able to absorb high levels of SAN compromise (50% in this case) while only passing on a 20% reduction in density and retaining a much better user experience.

Figure 12: Login VSI Results with NexentaConnect.



NexentaStor SAN over 10GE iSCSI: Test Methodology

The goal of the Nexenta Lab was to confirm and further validate the results obtained at the VMware Lab. The testing focused on the SAN offload during provisioning and recompose operations.

The Nexenta Lab used NexentaStor SAN storage on top of a Supermicro server serving iSCSI LUNs over a 10GbE network to the ESXi cluster.

SAN Offload – Provisioning and Recompose

VMware View Composer provisioning and recompose operations generate very high levels of I/O. Nexenta lab tests involved provisioning and recomposing 180 linked clone virtual machines on two ESXi hosts to compare I/O levels passed down to the LUN on both the bare-metal and NexentaConnect configurations.

Figure 13 below illustrates the disk read/write throughput during the VMware View Composer provisioning operation. The highlighted values represent the read/write rate averages over a one-hour period; because the actual provisioning cycle took 51 minutes, the figures have been adjusted upwards accordingly. The read rate of 30.3 MBps and write rate of 8.7 MBps are the adjusted bandwidth levels hitting the SAN during the bare-metal provisioning scenario. NexentaConnect demonstrates a 4000x read offload and 2.1x write offload, with a read rate of only 6.7 KBps and write rate of 4.2 MBps hitting the SAN.

The test also reported a very high ARC hit ratio at 99%, with 99% data-demand efficiency and 38% data prefetch efficiency. This explains the high read offload due to 99% of data being in ARC. Iostat showed a 87KB block size for writes, or a roughly 20x write aggregation.

Figure 13: Provisioning without CBRC.

Figure 14 below illustrates the disk read/write throughput during the VMware View Composer recompose operation. The bare-metal scenario reports a read rate of 26.6 MBps and write rate of 7.2 MBps. NexentaConnect reports a read rate of 20 KBps and write rate of 2.2 MBps for the physical disk and 276 MBps read rate, 6.1 MBps write rate for NFS datastore. The Nexenta VSA not only offers a 1360x read and 3.2x write SAN offload, but is also able to handle 10x higher read throughput than the bare-metal scenario.

The test, similar to the provisioning operation, reports an ARC hit ratio of 99%, with 99% data-demand efficiency and 85% data prefetch efficiency. Once again all the data is read from Nexenta VSA's memory. The Iostat shows a 57.2 KB block size going to the physical LUN, or a 14x write aggregation on the VSA layer.

Figure 14: Recompose without CBRC.

Why These Results Matter

SAN sizing continues to be a challenging component of VDI design — and a significant cost component as well. VSA acceleration offloads a significant portion of I/O from the SAN or NAS device, removing the pressure to size the SAN for all peak conditions. SAN sizing can be a bit more relaxed and targeted towards average consumption, because each ESX host will have a VSA amplifying the native IOPS capability of the SAN and helping ensure lower latency I/O during period of peak demand. Less expensive SAN options can be pursued from the start and with greater assurance that user experience will be protected during periods of higher than normal loads. VDI deployments that grow faster than expected can avoid degradation of user experience that may result before SAN resources can be upgraded.

Conclusion

The NexentaConnect solution delivers 3800% SAN offload during provisioning and other I/O storm scenarios. Even during steady-state workloads with lower caching efficiency, the acceleration layer amplifies IOPS between 300% and 500%. User experience is preserved under extreme conditions and I/O latency is reduced by 75% during periods of SAN saturation. Finally, the effect of acceleration reflects 72% better density during periods where SAN I/O delivery is compromised by roughly 50%.

The NexentaConnect solution creates an exceptional VDI user experience, which is increasingly important with today's competing BYOD client options. It fundamentally reduces the risk of SAN saturation and by extension the risk to user experience that can hobble a VDI deployment and lead to smaller implementations instead of mainstream adoption. NexentaConnect acceleration strategy increases cost savings by allowing the SAN to be sized for an average utilization instead of for peak conditions. The Nexenta VSA layer accommodates most of the peak traffic, leading to lower average I/O levels to the SAN and much less volatile I/O patterns.