ESG Lab Review

NexGen Storage: Delivering Quality of Service with VMware vSphere 6 and Virtual Volumes

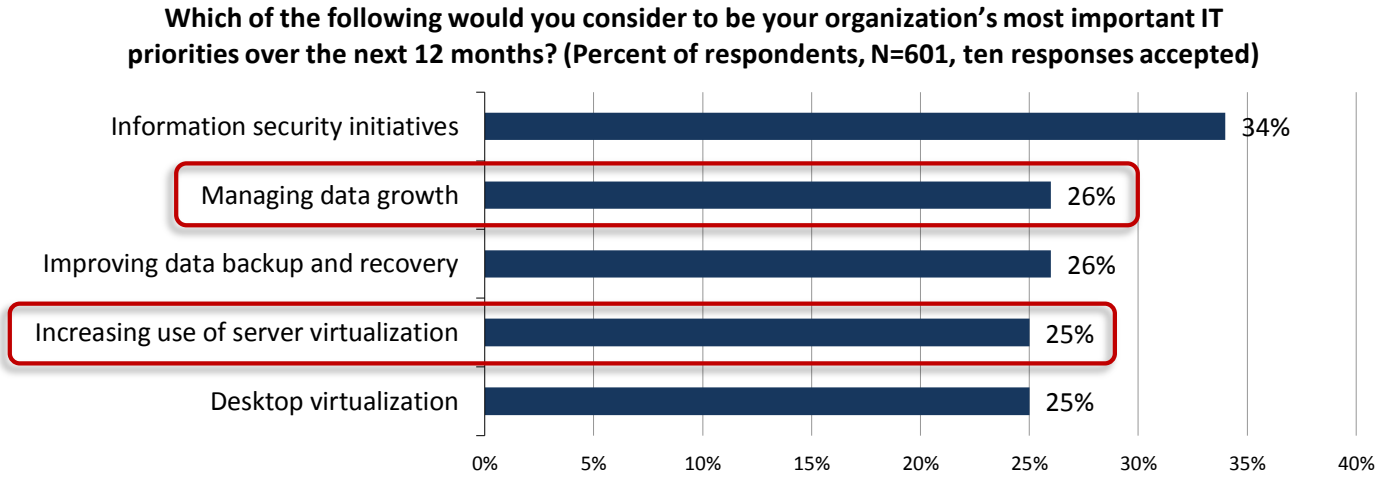
Date: May 2015 **Author:** Mike Leone, ESG Lab Analyst

Abstract: *This ESG Lab review documents testing of NexGen Storage N5 Hybrid Flash Arrays with a focus on how the vendor's QoS technology can deliver predictable and prioritized performance in ever-growing, virtualized environments. NexGen's integration with VMware Virtual Volumes was also previewed to understand how combining quality of service with storage-aware VMs allows organizations to manage virtualized environments on a per-VM basis.*

Background/Challenges

Virtualization has been one of the biggest trends the IT industry has ever seen. Offering benefits like improved efficiency, agility, and cost savings just to name a few, it is not surprising that many organizations have done everything in their power to take advantage of virtualization technology. But this widespread adoption has created a new challenge for IT: managing a rapidly growing virtualized IT infrastructure segregated by physical resources (i.e., storage). Storage administrators are tasked with provisioning and managing the underlying storage in large, heterogeneous storage environments, while the virtualization administrator is responsible for managing the VMs and applications that run on the storage. The dependency from the virtualization administrator on the storage administrator has made it increasingly difficult to deliver accurate service levels because of lengthy request cycles. This problem is exacerbated by constant data growth, which further highlights the fundamental flaw in how virtualized environments are traditionally designed and managed today. Recent ESG research shows that this challenge is something many organizations are looking to take head on. Managing data growth and increasing the use of server virtualization were two of the top five most-cited responses when organizations were asked what their most important IT priorities are over the next 12 months.[1]

Figure 1. Top Five Most Important IT Priorities Over the Next 12 Months



Source: Enterprise Strategy Group, 2015.

With customers already virtualizing a significant amount of their infrastructures to save on operational expenses and increase IT efficiency, the focus has shifted to management. Management not just in terms of VM sprawl, but from a performance standpoint. The lack of application-awareness in the traditional operational model has never been more apparent and is particularly problematic in constantly growing, shared, virtualized infrastructures. By leveraging technology like VMware vSphere 6 with virtual volumes, organizations will have the ability to manage their growing applications by accurately and efficiently delivering storage service levels based on its business value.

[1] Source: ESG Research Report, *2015 IT Spending Intentions Survey*, February 2015.

VMware Virtual Volumes

Virtual Volumes (VVOLs) make up VMware's new integration and management framework in VMware vSphere 6 that virtualizes SAN and NAS-attached storage and turns it into VM-aware storage. This creates a more efficient operational model that is focused on the applications, rather than the underlying storage infrastructure and its rigid restraints around LUNs and storage volumes. VVOLs turn the virtual disk into the primary data management unit, making it possible to execute storage operations with VM granularity by creating a more flexible policy-based approach for managing storage on a per-VM basis. The delivery of storage service levels to specific applications is simplified even further with policy-driven automation that allows VMs to be adjusted in real time to meet changing requirements.

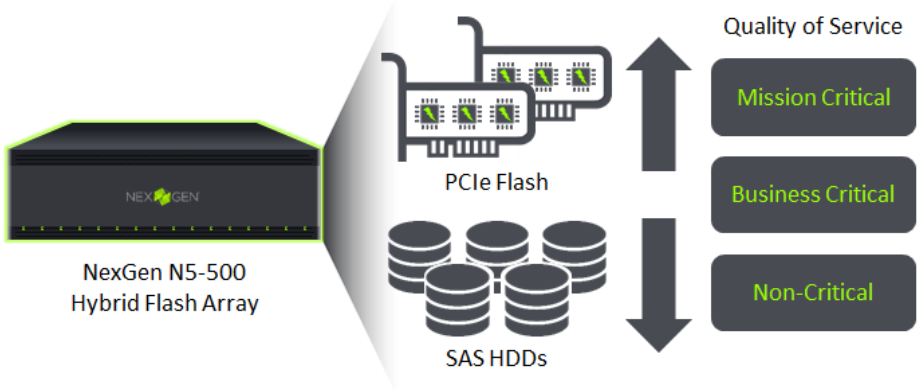
NexGen N5 Hybrid Flash Arrays

The NexGen N5 is a hybrid flash array that delivers cost-effective and predictably fast performance for mixed-application workloads in consolidated and virtual server environments. The NexGen architecture uses a combination of PCIe flash, cost-effective SAS disk drives, and intelligent software to reduce costs while providing predictable and prioritized performance for mission critical applications. NexGen offers differentiation across three areas:

- *PCIe Flash* – By putting the flash capability on the PCIe bus within the storage array instead of in a SSD behind a raid controller, NexGen eliminates the potential bottleneck that can occur when storage controllers and protocols are in the data path between the application and the flash memory. This enables higher levels of performance and lower response times without sacrificing hard drive capacity.
- *Dynamic Prioritization* – NexGen provides quality of service, with performance targets for minimum IOPS, throughput, and not-to-exceed latency for each volume that can be monitored in real time. This ensures that performance is predictable by workload. NexGen uses its QoS engine to enable dynamic prioritization by moving data between PCI-e flash and disk-based storage in real time. This ensures that applications with a higher QoS target get a higher percentage of data in flash, while those with a lower QoS setting have more blocks on disk. Snapshots and replication can also be applied to each of the volumes, enabling granular, policy-driven data protection.
- *Policy Based Management* – NexGen volumes can be defined across five different performance policies, from mission critical to non-critical, with service level minimums attached to each designation. These preconfigured policies, along with the ability to modify VM-storage QoS policies on the fly, eliminate the need for back-end storage vMotion and simplify administration. Additionally, NexGen is integrated with VMware vCenter via a NexGen VASA provider and a VMware vSphere plug-in.

NexGen and VMware VVOLs

Figure 2. NexGen N5 Hybrid Flash Arrays



NexGen has taken the benefits of VVOLs, specifically the simplicity of provisioning and managing storage on a per-VM basis, to the next level by applying NexGen's industry-proven storage quality of service (QoS) and service levels. This creates a solution where the storage QoS on the back-end is tied directly to the front-end of the virtualized applications. With NexGen's simple policy management being extended to VVOLs, NexGen QoS policies can easily be mapped to VM storage polices and applied to VMs directly in the vCenter interface. The ability to surface NexGen QoS policies within vCenter and apply them to a single storage container allows for the acceleration of VMs without the need for storage vMotion. This is a key differentiator for NexGen as they are currently the only VVOL certified storage array that eliminates the need for storage vMotion between back-end storage pools when changing a QoS policy on the fly. This results in improved VM density and consolidation. By enabling organizations to assign NexGen's granular storage policies to individual VMs, organizations can now benefit from a more predictable application and VM experience with consistent performance levels.

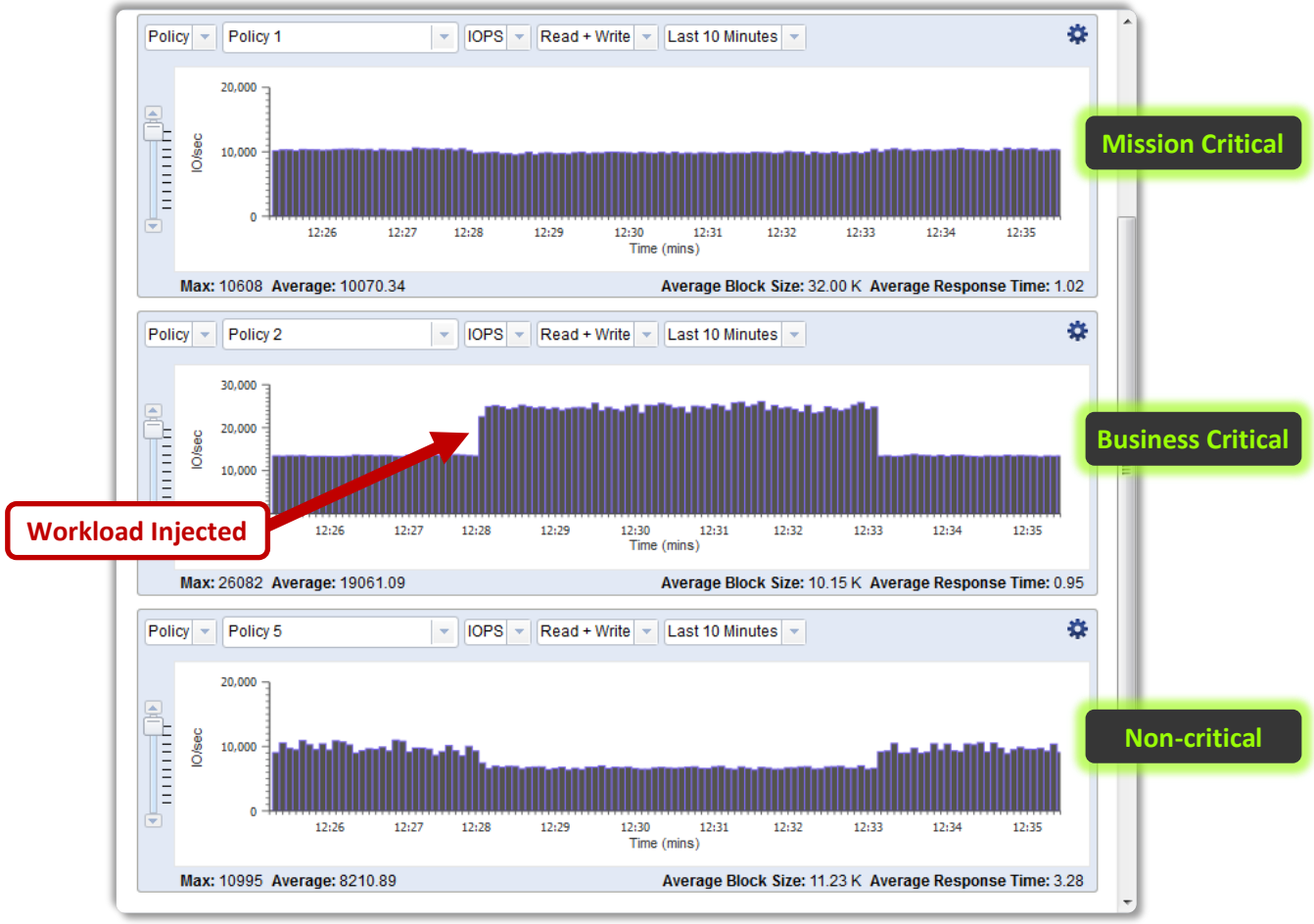
NexGen Quality of Service

NexGen's QoS comes standard with every NexGen N5 Hybrid Flash Array. The QoS engine is used to ensure minimum application performance for a specific volume by prioritizing I/O within the storage system. This functionality guarantees application performance by monitoring the response time, throughput, and IOPS of the system and isolating individual workloads on the array. There are three QoS service levels: mission critical, business critical, and non-critical. Five performance policies are predefined and fall within these three service levels. Each of the policies have differing IOPS, bandwidth, and latency targets. Based on the current workload, available resources, and user-defined service levels, the order in which I/O requests are processed is adjusted in real time to meet application performance requirements. When faced with a spike in activity from a low-priority application or a hardware failure, the NexGen software dynamically adjusts I/O requests and caching algorithms to achieve desired performance levels.

Predictable and Prioritized Performance

ESG Lab tested the functionality of NexGen's QoS with a goal of showing how predictable performance can be achieved for mission critical applications, even in demanding virtualized, mixed workload environments. A scenario was simulated that consisted of running three virtualized application workloads and injecting a fourth to see the impact it would have on the other three. The industry-standard Iometer tool was used to simulate the four application workloads, all of which can be found in many of today's businesses. The workloads included Microsoft Exchange, an OLTP database, VDI steady state, and a file server workload, which was used for the injection. Each workload was assigned a NexGen QoS policy. Microsoft Exchange was assigned mission-critical, the OLTP database was assigned business-critical, VDI steady state was assigned non-critical, and the injected file server workload was also assigned business-critical. Using the NexGen user interface, the total IOPS for each workload were monitored and after steady state was achieved for the three primary applications, the file server workload was injected. The results of the injection are shown in Figure 3.

Figure 3. NexGen QoS for Predictable and Prioritized Performance

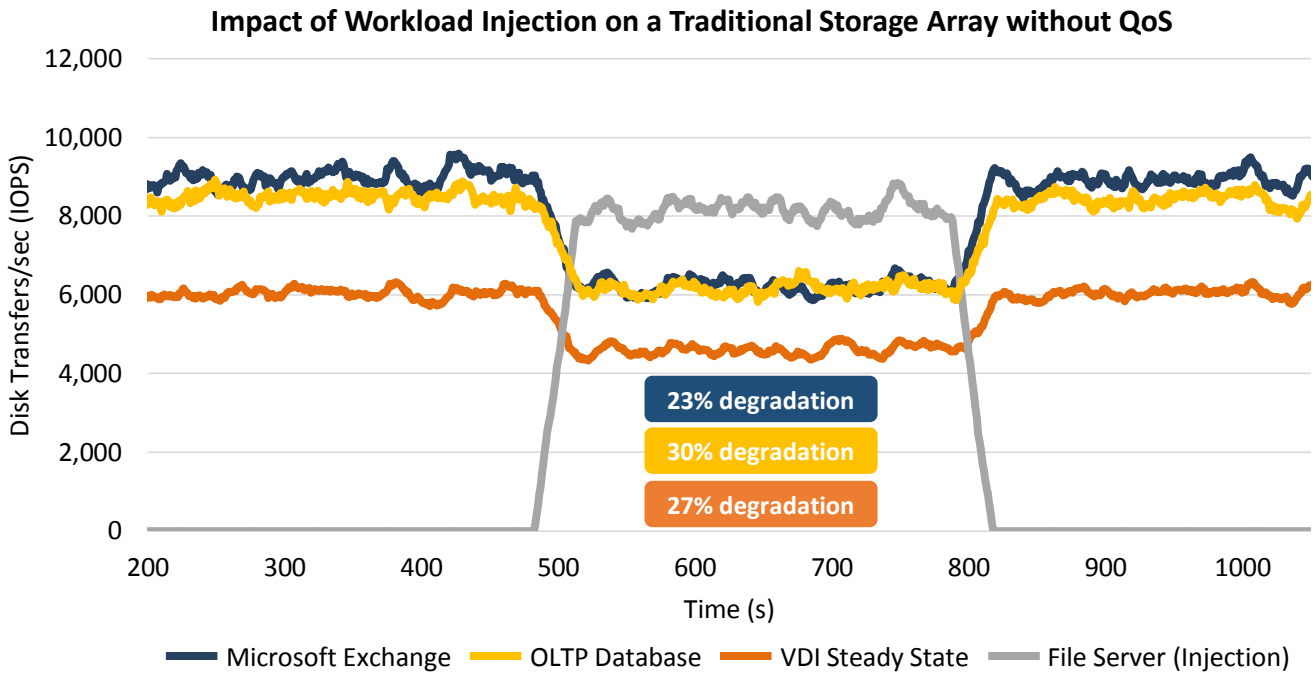


What the Results Mean

- After achieving steady state performance for the mission-critical Microsoft Exchange workload, business-critical OLTP database workload, and non-critical VDI steady state workload, the business-critical file server workload was injected.
- As expected, the performance for the business-critical tier doubled once the injection was initiated. This was representative of two workloads being run on this tier: the OLTP database and the file server workload.
- By injecting the file server workload in the business critical tier, it was given priority over the non-critical VDI steady-state workload. This led to a 31% reduction in IOPS for the VDI workload, a reduction that was controlled and predictable, meeting the non-critical policy performance minimums.
- The mission-critical Microsoft Exchange workload saw no degradation in response time and a minimal 5% reduction in IOPS, easily remaining within the mission-critical performance policy goals.

After witnessing NexGen's QoS capabilities, ESG Lab tested the same scenario on a traditional hybrid storage array without QoS. As shown in Figure 4, when the file server workload was injected, the performance of all three primary workloads was impacted. The Microsoft Exchange workload decreased by an average of 30%; the OLTP database workload decreased by an average of 30%; and the VDI steady state workload decreased by an average of 27%.

Figure 4. Traditional Storage Array Handling a Workload Injection

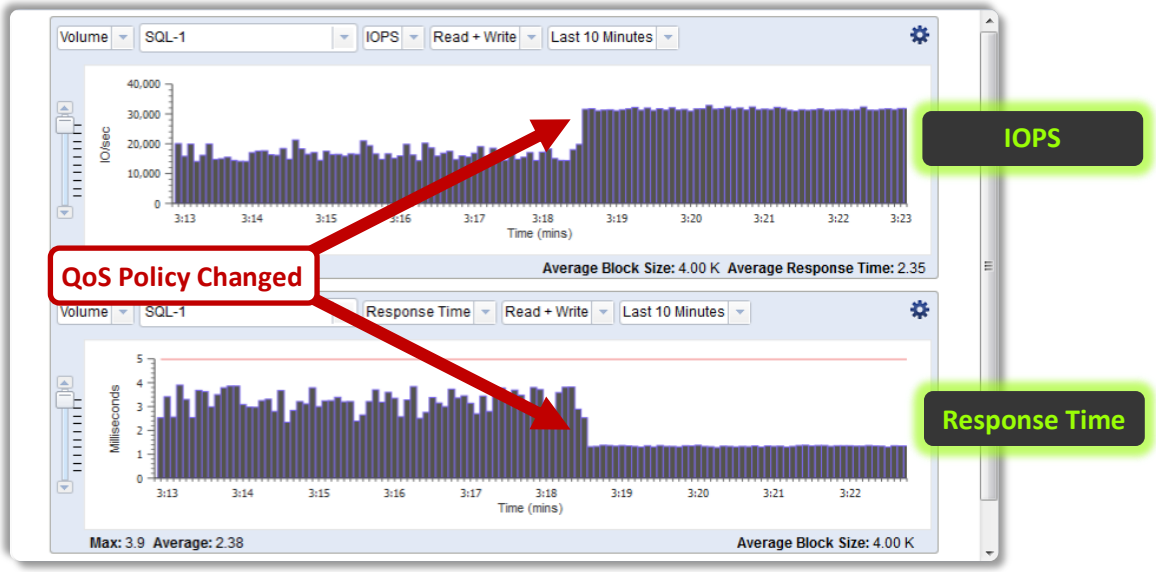


Software-Defined Performance

The next phase of ESG Lab testing consisted of validating the flexibility of storage QoS to adjust performance in real time for actively running applications. This is symbolic of what commonly occurs in virtualized environments with live VM migrations. In VMware environments, when an application requires more resources to keep up with demands or meet higher SLAs, automatic optimization will occur that requires the physical location of a VM to be relocated using vSphere storage vMotion. While the process occurs, organizations experience no application downtime and no disruption of business operations. With NexGen, QoS policy changes take effect immediately because of its dynamic data placement technology, which moves data in and out of PCIe flash depending on the assigned QoS policy and its priority without the need for traditional storage vMotion between performance pools

ESG Lab simulated two virtualized SQL Server application workloads that were assigned the non-critical QoS storage policy. After reaching steady state, one of the VM storage policies was changed from non-critical to mission critical. IOPS and response time were monitored throughout the test from within the NexGen GUI. As shown in Figure 5, the change to a policy with higher prioritization yields a noticeable improvement in both consistency and performance.

Figure 5. Real-time Performance Flexibility



Improved VM Density and Increased ROI

The next phase of ESG Lab testing consisted of auditing performance benchmarks run by NexGen Storage with a goal of quantifying the VM density and return on investment benefits that can be expected when leveraging NexGen's QoS. NexGen used an industry-standard VM workload generator to simulate a group of applications running in a virtualized IT infrastructure. One group consisted of eight different sub-workloads, including an Exchange database server, the DVD Store database application with web applications servers, Olio database and web application servers, and hypervisor automation workloads. The overall workload was increased by adding additional groups until 70% of the I/O response times exceeded 30ms.[2]

Two scenarios were tested and both leveraged 40 configured LUNs (11.4TB) on a NexGen N5 hybrid flash array. The first scenario consisted of strategically placing LUNs across QoS policies to ensure predictable performance minimums and an overall performance-optimized experience. The second scenario consisted of placing all LUNs within the same QoS policy, which served as a simulation of a traditional storage array by effectively disabling QoS. Table 1 highlights the comparative results of both test scenarios. Also included is a comparison to a leading mid-range storage array.

Table 1. Comparing VM Workload Performance Results

Storage	# of VMs	NexGen Feature Leveraged for Result
Leading mid-range array	56	
NexGen N5-500 without QoS	240	PCIe Flash Integration
NexGen N5-500 with QoS	296	PCIe Flash Integration plus QoS

What the Numbers Mean

- When comparing NexGen against a leading mid-range array, ESG Lab witnessed an increase of over 4x in VM density as a result of NexGen's PCIe flash integration.
- With the unique architectural approach of PCIe flash in the storage array, along with QoS to effectively leverage it, the VM density performance improved to over 5x that of a leading mid-range array.
- It should be noted that aside from assigning LUNs to QoS policies, no other optimizations were done, leading ESG Lab to believe that even greater VM density and lower cost per VM are highly likely.

[2] For more detailed information about the test bed, benchmark, and results, visit *NexGen N5 Performance in a Virtualized Environment*.

Why This Matters

Storage performance is a challenge in virtualized environments: Multiple applications with different workloads and different performance requirements all access the same underlying storage. With the user experience being top of mind for most organizations, poor application performance, especially for mission critical applications, is simply unacceptable. Because of that, organizations have resisted moving these tier-1 applications to virtual servers for fear that users will have a poor application experience, leading to a loss of customers and therefore a loss of revenue.

ESG Lab confirmed NexGen N5 hybrid flash arrays with QoS provide organizations with predictable and prioritized performance for applications running in virtualized environments. Testing showed the performance of a simulated mission critical application go unhindered as other applications were simultaneously running on the same storage array and an additional application workload was injected into the mix. When a similar scenario was attempted with a traditional hybrid storage array, all application workloads were impacted and performance decreased by as much as 30%.

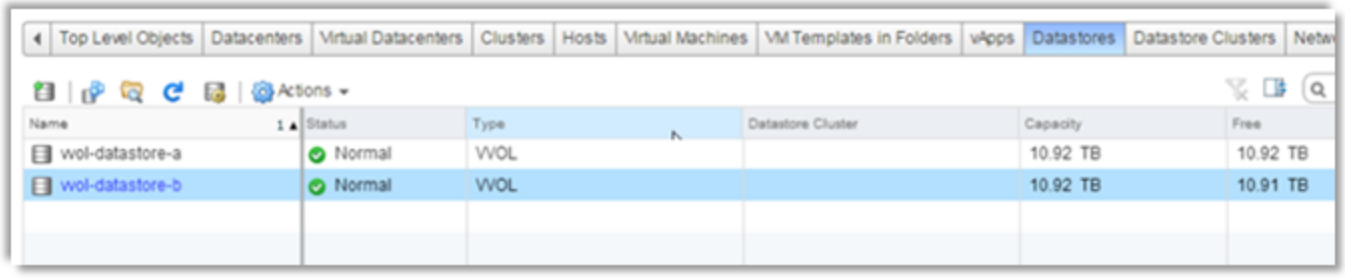
ESG Lab validated the real-time performance flexibility of QoS by changing a policy while an application workload was running. IOPS performance immediately increased because of the higher performance thresholds while response times decreased and became more predictable. Audited testing also showed how the PCIe flash-first architecture of NexGen N5 storage arrays delivers significant performance gains and improved VM density over traditional storage arrays. Additionally, it was demonstrated that utilizing Nexgen QoS could not only help consolidate more virtualized applications, but also deliver a significant cost savings per VM when compared with a storage solution that did not leverage QoS.

NexGen Integration with VMware VVOLs

The final phase of testing consisted of ESG Lab assessing how VVOLs are integrated with the NexGen Storage product. This included configuring a VVOL, creating a VM, and applying a NexGen QoS policy to that VM directly through the VMware vSphere interface.

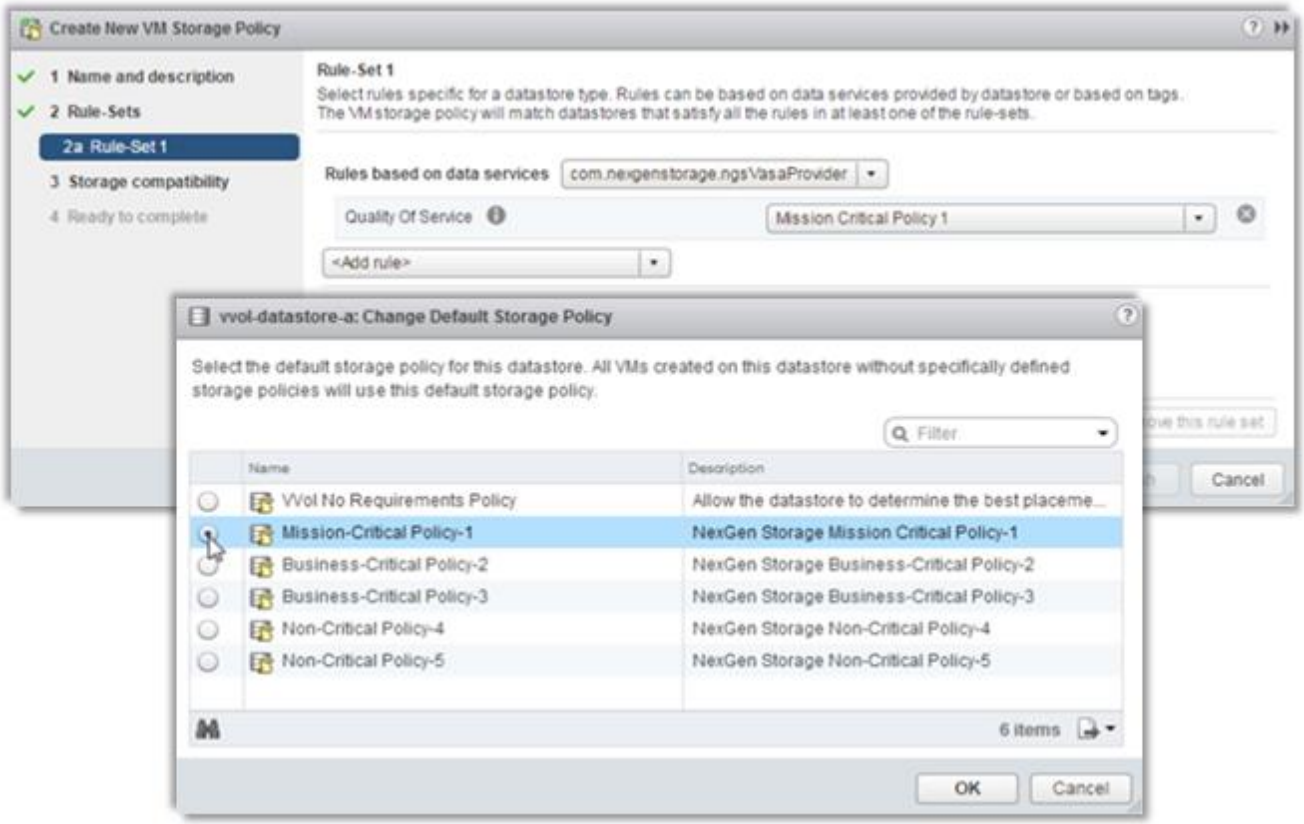
The first step was registering the NexGen VASA Provider from within vCenter. This process enabled the out-of-band communication between vCenter and the NexGen storage array. ESG Lab provided the NexGen storage array URL and array credentials to enable NexGen's capabilities to be quickly exported from the NexGen storage array to vCenter. After verifying that the registration process completed successfully, two 10.92TB VVOL datastore containers were created. A view of the newly added datastores is shown in Figure 6.

Figure 6. Creating VVOL Datastores



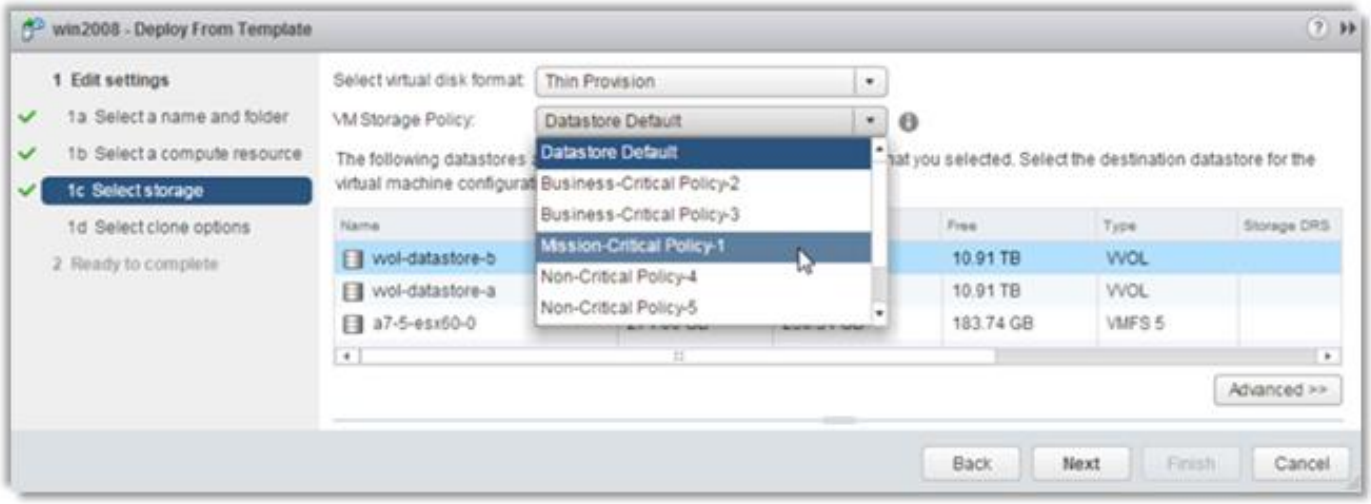
Next, ESG Lab setup storage QoS policies. Policies are created based on application requirements like performance, high availability, and data protection, all of which are visible as a result of registering the NexGen VASA Provider. After providing a name and description for a new mission critical policy, ESG Lab defined a rule set. The NexGen VASA Provider data service was selected, which enabled ESG Lab to select NexGen's mission critical QoS policy and apply it to the new mission critical VM storage policy. As shown in Figure 7, this process of mapping NexGen QoS policy levels to VM storage policies was done for each of the five NexGen QoS policy levels.

Figure 7. Assigning a NexGen QoS Policy to a VM Storage Policy



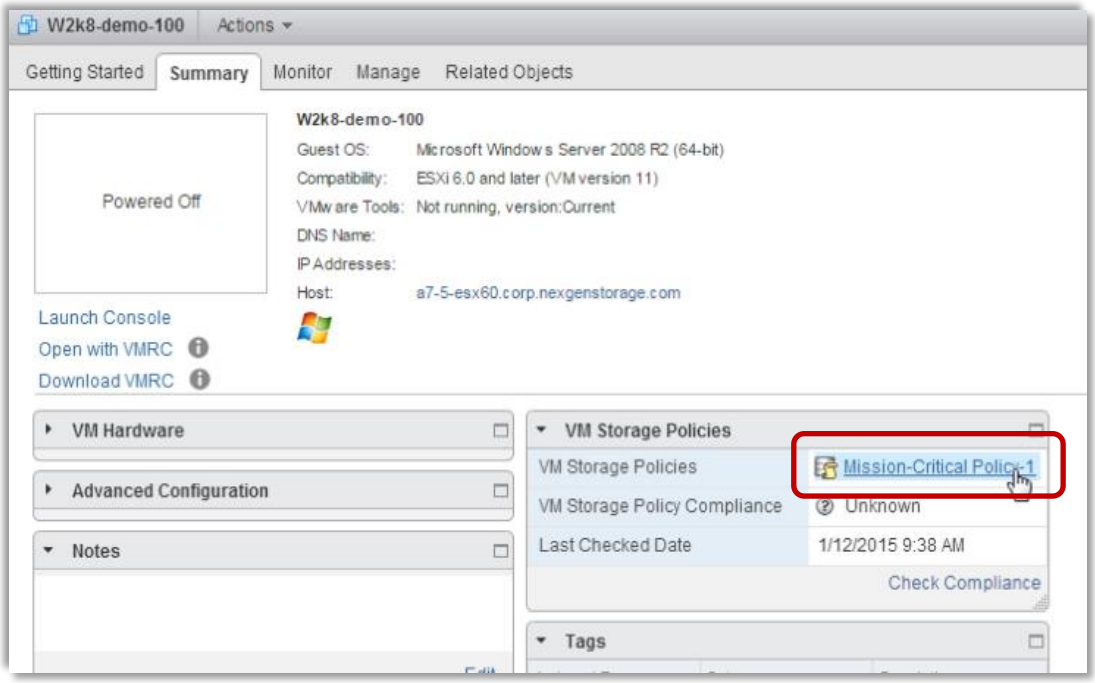
After configuring the new VM storage policies, ESG Lab deployed a VM from a template to a VVOL. As shown in Figure 8, this step required selecting a virtual disk format, a VM storage policy, and the VVOL datastore. After finalizing all the selections, the VM creation and deployment process began, which consisted of the storage array automating where the VVOL was created. The physical location of the VVOL on the array is determined based on the VM storage policy.

Figure 8. Deploying a VM to a VVOL



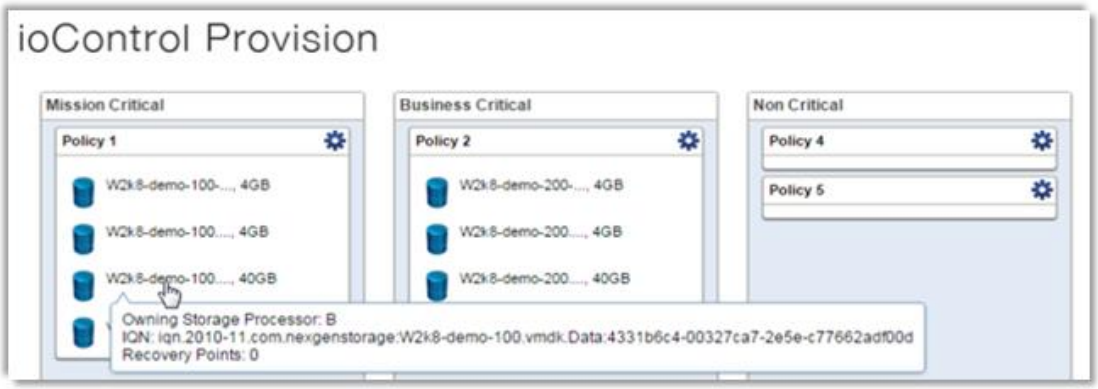
Once the VM was deployed, ESG Lab navigated to the familiar VMware interface that summarized all the information about the newly created VM, including a new VM storage policies window that displayed the selected VM storage policy for that particular VM (Figure 9).

Figure 9. VM Summary Tab



Finally, it was time to shift from the VMware user interface to the NexGen user interface. After logging in, ESG Lab saw the VMs that were created in vCenter displayed in their respective NexGen QoS buckets (see Figure 10), further validating how the NexGen QoS policies mapped to the VM storage policies that were applied to the VMs and VVOLs. It should be noted that because VMs consist of multiple data parts (VMDKs, configuration files, snapshots, etc.), they are made up of multiple VVOLs. In other words, a single VVOL does not encapsulate all of a VM's data and files.

Figure 10. NexGen User Interface Displaying VMs



Why This Matters

VMware VVOLs lay the down a perfect foundation to help organizations reach the next level of operational efficiency in a virtualized infrastructure by providing finer VM control and delivering service levels through granular provisioning and management on a per-VM basis. VVOLs hope to streamline operations by shifting control to the application, rather than the underlying infrastructure, but it should be noted that because configuring native array capabilities are passed from the storage array through the hypervisor, the underlying storage system must be well aligned and integrated with the VVOL framework. The combination of NexGen's proven storage and QoS capabilities with the application-centric VVOL framework will offer organizations the flexibility, agility, and granularity required to manage a virtualized infrastructure the way it should be: on a per-VM basis.

The Bigger Truth

With data growing at an increasing rate, organizations are looking for any way possible to consolidate their IT infrastructures. Virtualization technology has helped to accelerate that consolidation, but new challenges related to provisioning, management, and support have presented themselves. This is especially true in constantly growing, ever-changing virtualized infrastructures with a diverse storage back-end. The current operational model with virtualization has a heavy dependency on the underlying storage, as opposed to where the focus should be: on the applications that drive business value. How can organizations ever expect to keep up with the demands of a dynamic, virtualized infrastructure if the current operational model is broken?

NexGen Storage is leveraging its purpose-built software architecture to address value-driven data management priorities. By simplifying the operational model through QoS and policy-based management, organizations can better align business priorities and SLAs to their business. NexGen Storage has integrated with VMware vSphere 6 and VVOLs to offer the ability to assign NexGen's industry-proven storage QoS directly to individual VMs stored in VVOLs. This enables organizations to define business storage priorities associated with specific virtualized applications and manage those priorities on a per-VM basis.

ESG Lab confirmed the overall value of NexGen's storage QoS to deliver performance predictability and increased VM density in highly virtualized infrastructures. The performance of simulated Microsoft Exchange, OLTP database, and VDI steady-state workloads were measured before and after injecting an additional file server workload. As expected, the mission critical Microsoft Exchange workload continued to perform at nearly the same level before and after the injection, while the non-critical VDI steady-state workload slightly decreased to provide additional resources to the business critical service level for both the OLTP database and the injected file server workload. When attempting this same test using a traditional storage array, all workloads were greatly impacted. NexGen QoS delivered impressive results in terms of VM density and ROI, supporting up to 296 VMs, which proved to be a 23% improvement in VM density over a scenario that did not leverage storage QoS and a 5.3x performance improvement over a leading mid-range array.

The integration between NexGen Storage and VMware VVOLs was seamless. ESG Lab found it easy to connect the NexGen VASA Provider, generate the protocol endpoints, create VVOLs and VM storage policies, map them to NexGen storage QoS policies, and then apply them to VMs. The ability to modify a VM Storage QoS policy without the need for storage vMotion was impressive, demonstrating the ability to speed-up or slow-down VMs in real time. Though prerelease code was used for testing the NexGen/VVOL integration, ESG Lab is confident that based on the existing validation of QoS technology and its benefits, customers can expect a smooth transition to this new, lean operational model that focuses more on the application than the underlying storage resources. If you are looking to dynamically manage application performance based on well-defined QoS policies that better align customer expectations with application performance SLAs, ESG Lab suggests taking a look at NexGen Storage.