Demartek Evaluation – Accelerate Business Results with Seagate EXOS™ 15E900 and 10E2400 Hard Drives

Executive Summary

In today's world of digital transformation, organizations require high performance compute and hard disk drive (HDD) options that can match a broad spectrum of workloads. Equally essential in this world of flat and shrinking budgets is the application of the appropriate cost/performance HDDs to the workload for the best return on investment. Seagate provides an interesting twist on HDD technology that contributes some significant performance benefits in real world applications.

Every application has its own I/O demands, whether highly sequential, large block archiving and backups, random transactional processing, latency sensitive real time data analysis, or something altogether different. Storage and compute systems have to meet or exceed service expectations to keep businesses competitive. Where performance is essential, the Seagate Exos 15E900 and Exos 10E2400 SAS drives represent the peak in HDD technology. These newest additions to the HDD market promise performance improvements in throughput and response times for new storage deployments or as drop-in replacements. In keeping with evolving standards, these drives come in legacy native 512-byte sector format (512B) and the newer four kilobyte sector size in a 512-byte emulation mode (512e) for easy integration into existing enterprise infrastructure.

This new generation of enterprise HDDs is also available with the right amount of flash per drive for enhanced I/O caching. Critical application workloads will benefit from a significant performance advantage without breaking the bank.

Seagate commissioned Demartek to independently evaluate the new Exos 15E900 HDDs in a variety of use cases to help IT experts make informed technology choices. Many workloads, such as OLTP, Webserver, Virtual Desktop, or other random read/write tasks, as shown within the test results below, may be best served by enterprise HDDs that can offer an extra performance boost with a flash cache.

Key Findings

	ST9000MP006 (512B)	ST9000MP006 (512B) with Advanced Write Cache	ST9000MP0146 (512e/4K)
OLTP IOPS	6614	6601	19016
Bootstorm max IOPS	8975	8680	12355
Random Write IOPS	4591	7828	8017
Webserver IOPS	1777	1772	3446
SNIA™ Hot Band IOPS at 5 ms	620	2575	4534
OLTP I/O Response	8 ms	8 ms	1.7 ms
Bootstorm Completion	449 secs	402 secs	268 secs

Figure 1 – Some of the performance gains from the Seagate Exos 15E900 hard drive with Enhanced Cache

Accelerate Business Results with Seagate EXOS™ 15E900 and 10E2400 Hard Drives

Seagate Exos 15E900 Hard drives

There are two models of this sixth generation 15K-RPM Seagate HDD, available in multiple capacities: the 512B with Advanced Write Cache™ and the 512e/4K[1] drive with a flash enhanced read cache trademarked Seagate TurboBoost™ along with write caching as a standard feature. Supporting 12Gbps SAS, and at the peak of the spinning drive technology and performance, these drives are marketed to businesses that demand 24x7 enterprise-level performance and reliability. Also included in this analysis is the 512B drive without flash acceleration. This drive serves as a baseline for performance comparison and, in some instances, may even be the appropriate drive platform for certain workloads.

	ST9000MP006-A	ST9000MP006-B	ST9000MP0146
Capacity	900 GB	900 GB	900 GB
Sector Size	512B Native	512B Native	4K Native / 512B Emulated
Advanced Write Cache™	No	Yes	Yes
TurboBoost™ Read Cache	No	No	Yes

Figure 2 – Seagate Exos 15E900 SAS drives considered in this evaluation.

The addition of flash caching offers particularly demanding application workloads I/O acceleration beyond the limits that physics places on purely mechanical drives. Read caches tend to perform best where some portion of data is retrieved with regularity, often referred to as "hot" data. Write caches are most effective where randomly written data can be temporarily placed in cache, to be flushed to media at a later time. Extremely heavy write workloads will eventually fill a write cache and trigger continuous flushing to media. Workloads of this nature do not tend to benefit much from cache and can be reasonably served by mechanical drives alone.

Seagate Exos 10E2400 Hard drives

There are two models of the ninth generation 10K-RPM Seagate HDD, available in multiple capacities: the 512B and the 512e/4K[2] with Advanced Write Cache™ and flash enhanced read cache trademarked Seagate TurboBoost™.

Supporting 12Gbps SAS, and at the peak of the spinning drive technology and performance, these drives are marketed to businesses with mission-critical servers and external storage arrays. This drive series is aimed at database and online transaction process (OLTP) applications. The 10K HDD drive models may also be a cost-effective alternative to higher-priced performance solutions.

	Exos 10E2400 (512B)	Exos 10E2400 (4K/512B)
Capacity	1.2 TB	2.4 TB
Sector Size	512B Native	4K Native / 512B Emulated
Enhanced Caching features	No	Yes

Figure 3 – Seagate Exos 10E2400 SAS drives (not considered in this evaluation).

The 4K/512e models of the 10K drive include Enhanced Caching features which offer the user additional performance benefits over the 512 Native models. Included on the Exos 10E2400 (and the Exos 15E900) is eMLC NAND flash memory, I/O Acceleration and Response Time Optimization and Intelligent NAND Endurance management.

[1] The ST900MP0146 512e drive comes standard with a flash enhanced read and write cache (TurboBoost and Advanced Write Cache). The standard 512e model can switch to 4Kn in a matter of seconds using the industry standard Fast Format.

[2] All 512e/4K model drives come standard with Enhanced Caching features including I/O Acceleration and Response Time Optimization, NAND Flash and Intelligent NAND Endurance Management.

Accelerate Business Results with Seagate EXOS™ 15E900 and 10E2400 Hard Drives

The Test Platform

Demartek was provided twenty-four 512B Exos 15E900 SAS drives for the evaluation of baseline drive performance. Twenty-four 512e/4K drives were also provided for measuring the effect of Seagate TurboBoost on common I/O patterns. Each set of twenty-four drives were placed into a twenty-four bay JBOD chassis with 12Gbps SAS interfaces for direct attachment to a server.

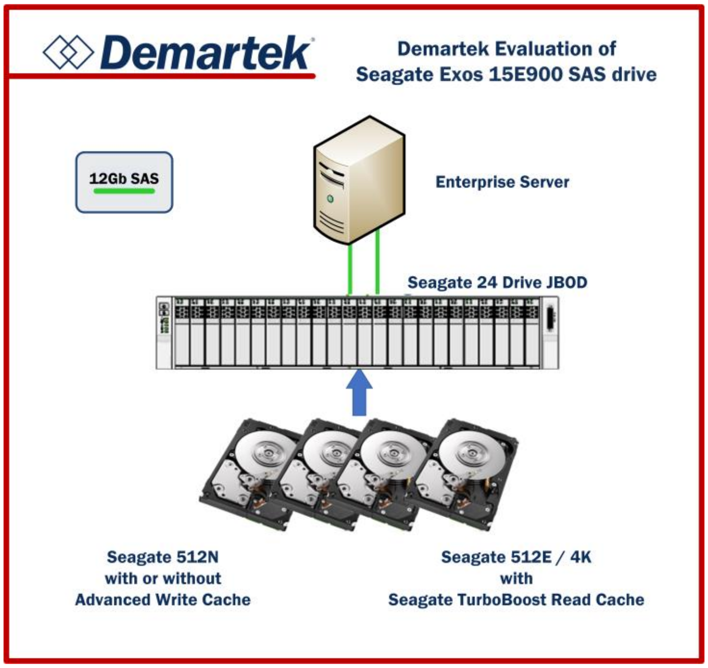

Figure 4 - Evaluation testbed

A modern multi-processor enterprise server with the compute power to handle high-performance workloads and host large virtualization environments was deployed to provide processing power and memory for this evaluation. A 12Gbps SAS HBA connected the JBOD to the server through two SAS channels. SAS MPIO was enabled in Windows and VMware ESXi to guarantee that I/O would not be limited by a single channel of SAS bandwidth. This server included over a terabyte of DDR4 RAM, but most workloads in this evaluation required only a fraction of that amount of memory. Microsoft Windows Server 2016 and VMware ESXi 6 were installed, as applicable, to support Windows based or virtualized workloads.

The Workloads

Drive vendors produce multiple drive models to meet the different I/O patterns that result from different application engines. For this reason, Demartek ran five workloads on each new generation 15K HDD and commented on the best use cases. These workloads included both synthetic I/O generation tools, which are useful for creating finely tuned but artificial I/O patterns, and bona fide "real-world" applications which do not lend themselves to as much predictability, but are a better approximation of the I/O that occurs where it really matters—within a customer's application setting. Summaries of each workload and deployment parameters follow.

Iometer

Four random I/O patterns were defined and scripted. These workloads included read-only and write-only I/O plus 50:50, 70:30 read-write mixes with four kilobyte block sized I/O requests. Fully random workloads such as these, combined with a large storage target (a single RAID 10 volume consuming all 24 drives) and short execution durations intentionally made read caching extremely difficult, highlighting instead the 15K RPM performance of each drive model to determine if the caching algorithms employed affected baseline drive performance. However, we did expect to see some improvement in write I/O from the Advanced Write Cache.

These Iometer workloads were run on a Windows Server 2016 environment against raw drive volumes, to eliminate filesystem effects on the test results.

SNIA Emerald™ Hot Band Workload

The Storage Networking Industry Association Emerald™ benchmark is an industry standard I/O load generation workload. Included in the Emerald™ benchmark script is a workload specifically designed to exercise a system's caching or tiering capability by grouping some reads and writes to discrete locations on storage—the "hot band." The read/write ratio of the workload is 65:35 with about sixty percent of the workload being random I/O.

In compliance with the SNIA Emerald™ specification requirements, a little more than 50% of the total RAID 5 capacity of the JBOD was deployed as a single raw volume on Windows Server 2016. For the 512e/4K drive configuration, this volume was pre-conditioned through an initial multi-hour execution of the hot-band script prior to recording performance metrics to allow the TurboBoost cache to fill with hot data. The benchmark was then executed with an increasing number of threads performing I/O on the volume. Metrics were collected until performance bounding trending could be identified.

The Hot Band workload is the second of two fully synthetic workloads executed for this analysis, and as such, results should be considered similarly to those of the Iometer analysis. The SNIA Emerald™ benchmark can be downloaded from the SNIA Emerald™ Program website at http://www.snia.org/emerald.

Microsoft IIS Read-Only Webserver

As the majority of web traffic is composed of read I/O, Demartek deployed a one terabyte read-only webserver with 600,000 html pages on a single 1.2TB RAID 5 volume. Page sizes varied from 15 kilobytes to 4 megabytes, with an average page size of 2 megabytes. Browsing traffic was generated with Neotys Neoload™, scaling from five to twenty virtual users requesting random pages. Each page request was made as soon as the prior request was satisfied, to drive the I/O load per user as high as possible.

Eighty percent of all page requests were directed to the same 180GB of files to test read caching. The remaining twenty percent were equally distributed among the other 820GB of files.

Microsoft SQL Server OLTP

Demartek deployed a one terabyte OLTP database, including logs and indexes, for this real-time read-write test case. A load generation program was tuned to a transactional intensity level that kept baseline I/O latency below 10 milliseconds. Like most OLTP database applications, this workload was intentionally read-heavy, with a read-to-write ratio close to 10:1.

Two RAID groups were provisioned from the 24 drive JBOD. Twenty drives were configured as a RAID 5 drive group, from which three 300GB volumes were deployed for data files, and a four drive RAID 10 group from which a single 225GB volume was deployed to store logs. The database was intentionally limited to eight gigabytes of system memory to force I/O to the drives instead of caching data in RAM. This reduces the transactional rate in favor of higher drive IOPS and bandwidth.

ESXi VDI Bootstorm

Booting virtual desktops creates yet another pattern of I/O because OS files are read while system and log files are created or updated. Bootstorms—large numbers of virtual machines starting near simultaneously—are not desired events, but can occur when a work shift begins and may place a heavy load on processing, memory, and storage resources. For this test case, sixty Microsoft Windows 10 desktops were deployed on a RAID 5 datastore provisioned from all twenty-four drives. All desktops were booted in sequence with a 3 second wait time between power-ons.

The VMware ESXi 6 hypervisor was installed on the server for this use case. Metrics were collected from the hypervisor with the 'esxtop' utility, and each virtual desktop was designed to report statistics to a test harness as the boot process progressed.

Performance Metrics

The three key metrics of system performance evaluations are Input/Outputs per second (IOPS), bandwidth, and latency. IOPS and bandwidth have a direct relationship, where IOPS times I/O block size equals bandwidth in kilobytes, megabytes, or gigabytes per second. Latency, measured in milliseconds, is the measure of time taken to complete an I/O request and is indirectly related to IOPS and bandwidth. Lower latencies may enable higher IOPS in some workloads. More importantly, latency impacts the performance of applications submitting I/O requests. Some applications, particularly those with high bandwidth sequential I/O, can tolerate a high degree of latency whereas transactional applications performing real-time processing often do not.

Accelerate Business Results with Seagate EXOS™ 15E900 and 10E2400 Hard Drives

As storage performance increases, it is not uncommon to see other system metrics, such as CPU usage, improve as well due to the increased amount of data available for processing.

Test Results

Iometer

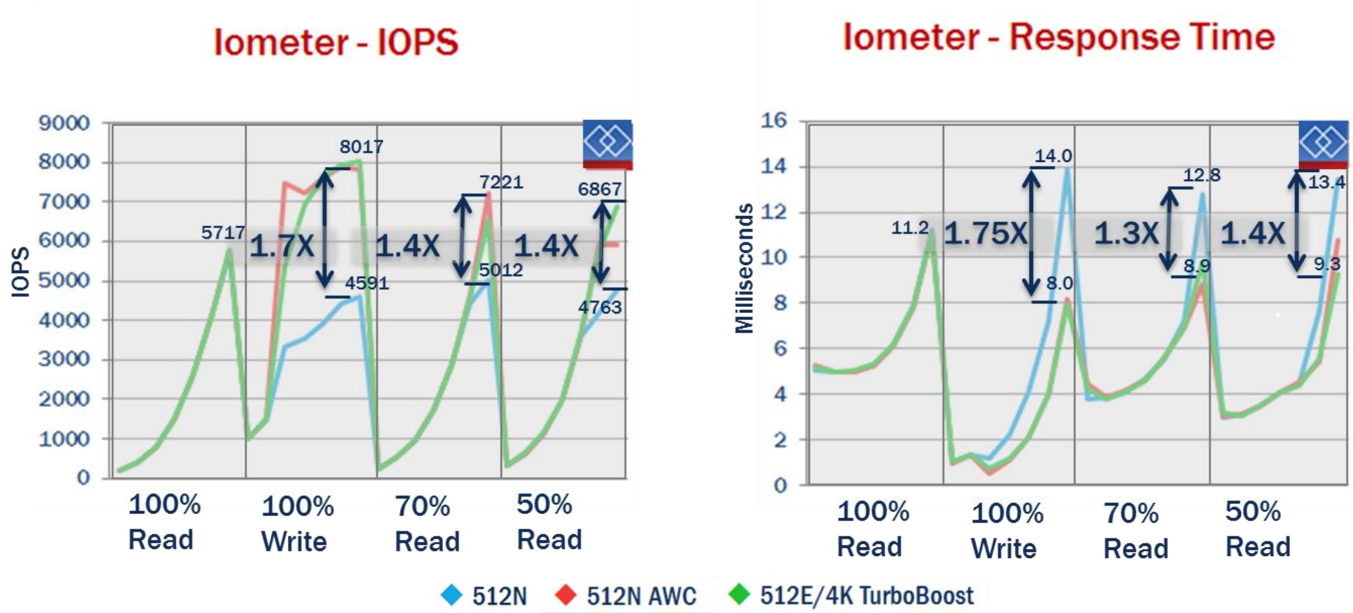

Figure 5 – Iometer results for 4K block size 100% random I/O

Iometer was run on Windows Server 2016 against a single RAID 10 raw volume consuming all drives in the JBOD. The charts represented in Figure 4 demonstrate completely random 4KB block size I/Os. Queue depths scaled from one outstanding I/O up to sixty-four. Fully random I/O, particularly against very large storage targets, presents little opportunity for read caching and all drive models appear to perform at near parity regarding random read performance. The number of random read IOPS, with a maximum of nearly 6000, is in line with expected performance of 15K RPM SAS drives in a 24-drive array. Latency begins at an acceptable five milliseconds but increases sharply at higher queue depths.

Adding write components to the workload allows the 512e drives to begin exercising their Advanced Write Cache™ feature. Unlike read caching, a write cache performs best against extremely random I/O (up to the point where the cache is completely filled and must begin flushing data to mechanical drive components) and that is seen here with a marked IOPS gain—70% more IOPS for the pure write scenario and roughly 40% improvement over baseline in mixed read-write workloads.

Either of these drive models, whether 512B or 512e, have essentially the same baseline performance. Usage that is primarily composed of cache unfriendly workloads, such as completely random, large datasets, streaming I/O, or archival processes would see similar performance with any drive model and drive selection could be logically based on price or availability. Workloads with time-critical random write operations will benefit from Seagate's Advanced Write Cache technology.

SNIA Hot Band

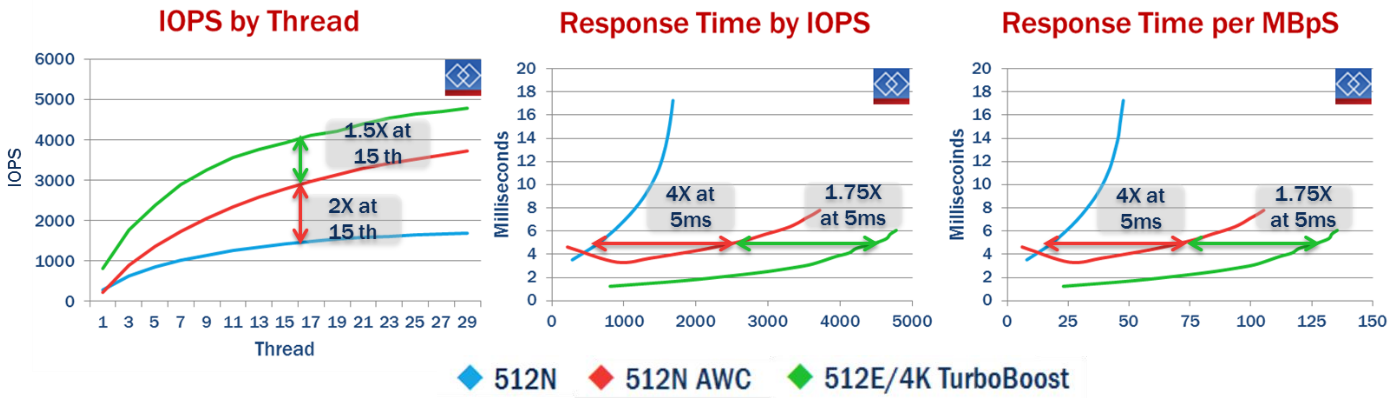

Figure 5 – SNIA Emerald Hot Band IOPS and response time performance by drive model

The SNIA Emerald™ Hot Band workload is a vdbench script that intentionally exercises read and write caches of storage devices. Most application I/O tends to be read heavy and the SNIA workload attempts to model this with a 65:35 read-to-write ratio. Limiting some I/O to narrow ranges of drive sectors, instead of spreading it evenly across the entire RAID device creates areas of hot data that cache easily with TurboBoost. Random write I/O benefits from Advanced Write Cache™ regardless of where data is written.

IOPS measurements bear this out as the 512e model drives performs more I/O than the 512B non-caching drives. At fifteen worker threads, the 512B drive with Advanced Write Cache achieved twice the IOPS of the baseline drive, while the 512e/4K drive delivered another 1.5X IOPS above that. This delta continues to increase as threads were added to the workload. It appears that the 512e trend curve will reach a steady state between 5000 and 6000 total IOPS. By offloading read and random write requests to flash, the spinning media is freed to do what it does best—efficiently serve the sequential writes that make up fifteen percent of I/O and other cache-unfriendly I/O patterns.

Average I/O latency showed improvement with the addition of cache as well. Using five milliseconds as a reference point, the 512B drive supported about 620 IOPS of the SNIA Emerald™ Hot Band workload. Enabling Advanced Write Cache on the 512B drive increased that figure to 2575 IOPS. The 512e drives managed more than seven times that number of IOPS—4534 IOPS—at the same five millisecond latency.

This test case was performed under SNIA Emerald™ 2.1.1 parameters, in which roughly 50% of the total drive capacity was exercised as a raw parity RAID (RAID 5) volume on Windows Server 2016.

Accelerate Business Results with Seagate EXOS™ 15E900 and 10E2400 Hard Drives

Read-Only Webserver

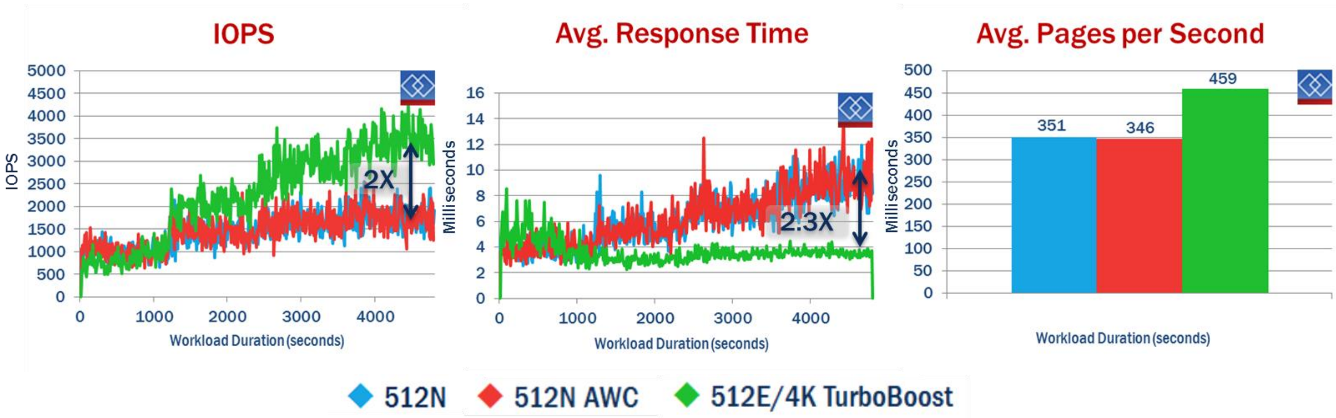

Figure 6 – Webserver IOPs and Bandwidth

As the majority of web traffic is composed of read I/O, Demartek deployed a one terabyte read-only Microsoft IIS 10 webserver with 600,000 html pages on a single 1.2TB RAID 5 volume. Page sizes varied from 15 kilobytes to 4 megabytes, with an average page size of 2 megabytes. Browsing traffic was generated with Neotys Neoload™, scaling from five to twenty virtual users requesting random pages.

This testing obviously highlights the 512e TurboBoost. When adding users to the workload, the drive model without TurboBoost experienced only modest gains in IOPS, with a corresponding increase in latency. On the other hand, the 512e drives show a significant increase in performance—doubled or better—as the read cache warmed. The maximum IOPS value of around 4000 was in fact limited by a 10GbE network bandwidth ceiling. Higher network bandwidth would likely have an even greater IOPS potential. We can also deduce by the inverse latency curve, which reaches steady-state performance of just under four milliseconds, that additional I/Os are supportable by the drives. We would expect to see latency increasing rapidly as users were added, as seen with the 512B drives, if this were not true.

I/O response time is critical to webservers. Page load time has a direct impact on reader retention for websites. Maximum permissible load time is debatable, and dependent on many factors, but certainly I/O response is one of those factors and keeping that time as low as possible will improve user experience. Obviously, retrieving as many pages as possible from a cache drives down I/O latency, as is demonstrated by the 2.3X reduction in I/O response time.

Another consideration for the application or system architect is the number of pages served over a time interval. Averaged over the duration of the test, with no consideration given to the changing user count, the 512e TurboBoost drives served 100 more pages per second. These results are no doubt biased by early performance where user demand was not as heavy, so steady-state page counts should be even higher.

Many websites include the ability to interact with site content, adding a write component that is not reflected in these results. However, obviously the 512e TurboBoost would be available to service that additional I/O, whereas the 512B drives would be required to share precious IOPS on the mechanical platter amongst both types of I/O, reducing the number of page reads in order to service write requests.

On-Line Transactional Processing (OLTP)

Figure 7 – OLTP Workload– IOPS and latency

Demartek deployed a one terabyte OLTP database, including both logs and indexes, on Microsoft SQL Server. A load generation program set transactional intensity such that baseline I/O latency was under ten milliseconds. Like most OLTP database applications, this workload was intentionally read-heavy. The ratio of read-to-write transactions is approximately 9.7:1. What write I/O there is tends to be dependent on read request completion and much of it is sequential in nature due to database logging.

This usage scenario is almost exclusively composed of 8KB block size I/Os. This I/O profile quickly exploited the TurboBoost of 512e drives. A cache warming curve is clearly evident during the first half hour of the test phase, with a steady-state performance nearly three times that of the other drive model.

A 10-millisecond average response time is generally acceptable for application I/O relying on mechanical drives and was the baseline threshold targeted for this test case, with a steady state condition of about eight milliseconds achieved. Assuming the 6600 IOPS generated at this latency limit represent enough throughput to satisfy business requirements, the 512B drive model supports this particular application. Where stronger performance is demanded, the 512e drives drove down latency by nearly a factor of five, to less than two milliseconds while improving IOPS nearly threefold.

Typical use cases for Exos 15E900 drives include response time sensitive applications. Considering the significant boost in IOPS alongside the impressive response time, flash caching delivers impressive performance acceleration for this database application's user experience, which should not be overlooked in scenarios with tight service level requirements. The deltas measured between baseline and cached scenarios also suggest that considerably more stress could be placed upon the application engine without requiring additional hardware investment while still meeting a latency target of fewer than 10 milliseconds, providing a measure of future-proofing against increased application demand.

VMware ESXi Bootstorm

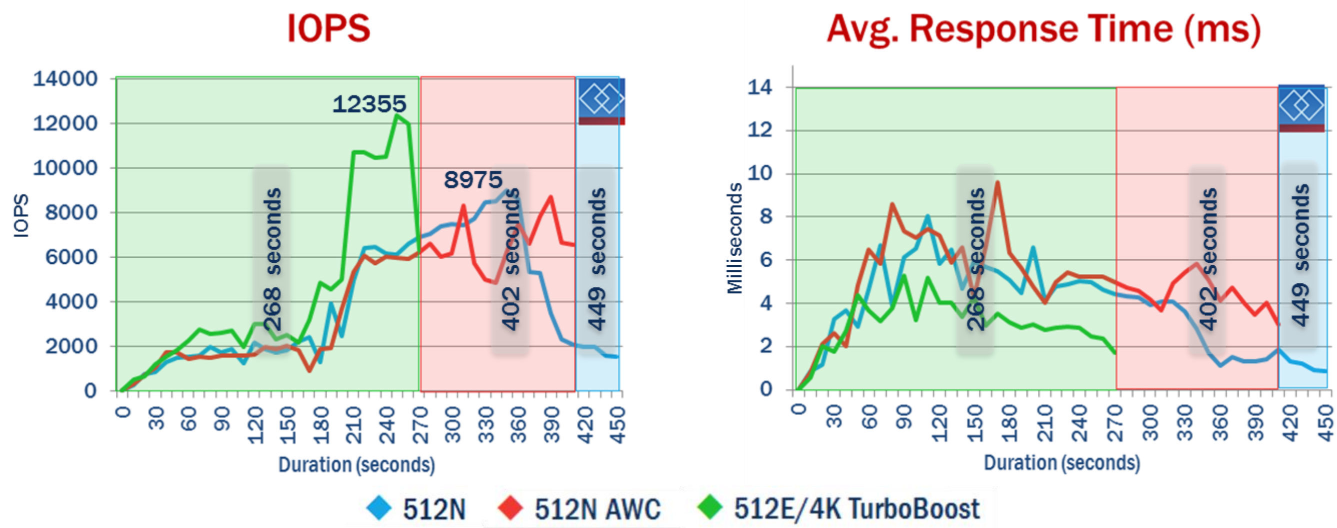

Figure 8 – VDI Bootstorm – IOPS and latency

The final use case of this evaluation is the simultaneous boot of sixty Microsoft Windows 10 virtual desktops on VMware ESXi 6. The limiting factor of this workload was expected to be storage performance. This scenario included the largest percentage of write I/O of the three real-world use cases, with a tenth or more of the I/O at any moment being write I/O. The graphs above stop recording performance for each drive type at the moment the final desktop in each test set completed booting.

Initially, desktops booted fairly regularly, but as the bootstorm progressed I/O began to backup. The 512e drive caches start to take very apparent effect at about three minutes into the workload to satisfy outstanding I/O requests. With a maximum recorded IOPS of 12355, the bootstorm generated the second highest throughput of any workload in this evaluation. Similarly, the 512B drives were driven to their peak recorded performance by this workload, reaching nearly 9000 IOPS.

The aggregate response time of these I/Os impacts the time needed to complete the boot process. Figure 9 summarizes the delta between the first and last boot of all sixty desktops, as backed by each drive option. These results really need no explanation. The benefit is obvious. The workload is best served by the enhanced 512e drive model.

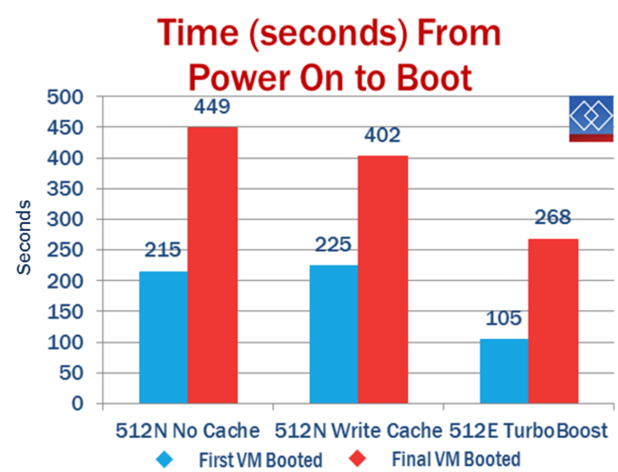

Figure 9 –Time of first and last booted virtual desktops

Accelerate Business Results with Seagate EXOS™ 15E900 and 10E2400 Hard Drives

Demartek
October 2017

Summary and Conclusion

Businesses deploy enterprise 15K and 10K SAS hard drives because performance matters. Selecting the right tool for the job at hand is critical for the system or storage architect. Seagate's Enhanced Cache feature consisting of Advanced Write Caching and TurboBoost Read Caching technologies offer a cost effective, high performance alternative to pricey all-flash systems. Not all application workloads demand the highest performance hardware to meet business requirements. For that matter, not all workloads include data patterns that cache easily, or at all.

A savvy architect will examine intended production workloads and consider current and projected data usage, I/O patterns, and user experience expectations. Equipped with this information, baseline Exos 15E900 hard drive performance may be sufficient to meet business requirements. If so, it would be illogical to incur additional expense to provision drives whose performance potential will never be exploited. On the other hand, if that same analysis shows that spinning media alone cannot deliver the I/O throughput or response times needed to meet service level requirements, the system architect is in a position to make informed storage and server recommendations, whether that be adding a caching component to traditional HDDs, or purchasing high-end all flash products.

A final consideration is ease of deployment. Some I/O acceleration solutions on the market require additional software packages for implementation, special licensing or, in extreme cases, application code, operating system, or server hardware awareness. Seagate Exos 15E900 hard drive features are completely transparent. Drives supporting Advanced Write Cache™ and TurboBoost need no preparation or activation to implement. The flash advantages are automatic and active at the moment of installation for maximum simplicity and flexibility.

With decades of experience refining drive technology, Seagate has the experience to assist customers in making these evaluations. Consumers are advised to contact Seagate sales personnel, or an OEM partner, for consultation on the right solution for their business needs.

Seagate.com/enhancedcache

Results may vary for different environments and workloads.

The most current version of this report is available at www.demartek.com/SeagateEnhancedCache on the Demartek website.