

Accelerating Database Workloads Using HPE MSA Storage with Built-In Flash

Evaluation report prepared under contract with Hewlett Packard Enterprise (HPE)



Executive Summary

Solid state storage is increasingly deployed in all sizes of datacenters, from the small and medium business to the large enterprise. It comes in many forms including hybrid arrays, direct attached drives, PCIe flash and accelerators that fit somewhere in between servers and storage. With the price of flash continuing to drop, hybrid storage is becoming more compelling to the small and medium business for critical computing applications such as databases and day-to-day operational computing.

In August 2016, HPE introduced the HPE MSA 2042 Storage system with 800 GB of built-in flash. HPE MSA is already the industry's #1 entry SAN storage array¹ with nearly a half million systems installed worldwide.

The MSA 2042 includes built-in flash, read caching, automated performance tiering (read and write) functions, and data resiliency features for a \$10K starting price, based on US street pricing² at the time this evaluation was published.

To demonstrate the benefits of using flash for read cache acceleration and performance tiering, HPE commissioned Demartek to evaluate the MSA 2042 with 800 GB of SSD capacity using a performance and read-intensive database workload based on Microsoft SQL

Server. These results were then compared to the all-HDD MSA 2040.

Demartek compared several performance metrics including bandwidth, IOPs, number of database transactions per second, and I/O latency for a complete picture of the array's suitability to support the scale of workload that a small to mid-sized business might experience. The purpose of this evaluation was to illustrate the potential benefit of adding flash to the array for read cache and read/write acceleration.

Key Findings

- > The all-HDD MSA array was able to support a transactional database workload with 30 virtual users while delivering an average of 650 database transactions per second.
- > With the MSA 2042 configuration, enabling performance tiering on the array increased transaction counts by 60% and using SSD capacity for read cache acceleration increased by 35%.
- > Average I/O response times for the MSA 2042 dropped from nearly 20 milliseconds to 4 milliseconds with the use of SSD capacity for read caching. When we employed SSD performance tiering, I/O latency dropped to below 2 milliseconds, yielding an 80% improvement in database response times.
- > Storage bandwidth and IOPS likewise saw increases in performance, each growing by more than 2.5 times when employing SSD capacity for read caching and more than 4.5 times with performance tiering.

¹ IDC Worldwide Quarterly Disk Storage Systems Tracker, Q1 2016, external disk subsystems, entry storage category by Factory Revenue

² Estimated US street price for 1.2TB configuration with hardware and software

The HPE MSA 2042 Storage

The HPE MSA 2042 Storage array is a dual-controller, two-rack unit with support for up to 199 SFF drives or 96 LFF drives. The MSA 2042 supports 8 Gb and 16 Gb Fibre Channel, 1 Gb and 10 Gb iSCSI and/or 12 Gb SAS host protocols. Each base MSA 2042 array comes with 800 GB of flash capacity in addition to an all-inclusive software suite and simple-to-use management tools designed for IT generalists.

The all-inclusive software enables the flexible use of the 800 GB of flash capacity to either accelerate reads (by using the flash for read caching) or to configure a full read/write performance tier. The built-in performance tiering engine dynamically moves “hot” data to flash and less frequently accessed data to lower-cost media to make the most of the system’s SSD capacity while accelerating application reads and writes. The caching and performance tiering capabilities are included without additional software licensing cost, making the HPE MSA 2042 an affordable solution for accelerating performance-hungry applications and databases.

It is also important to note that all of the benefits discussed in this paper with respect to the MSA 2042 are available to current MSA 2040 customers via a low-cost addition of SSD capacity and the Advanced Data Services Software Suite to their existing system. New all-inclusive software pricing allows MSA 2040 customers to upgrade to the MSA 2042 configuration tested in this evaluation for a 40% or greater cost savings as compared to a full array refresh, based on current US list pricing.

For this evaluation, HPE provided Demartek with an MSA 2042 configuration that included twenty (20) 900 GB, 10 RPM HDDs and a total of four (4) 400 GB SSDs for testing. Two SSDs come standard with the MSA 2042, and two additional SSDs were provided to test performance tiering. With this test configuration, Demartek was able to set up a 20 HDD baseline test case, with ten drives provisioned to each controller in a RAID 6 configuration.

To evaluate read caching, a single SSD was added to each controller and those SSDs were assigned as read cache type in the virtual disk group. For measuring the effect of the alternative performance tiering capability, two SSDs in a RAID 1 disk group were added to each controller.

Transactional Database Workload Description

Real vs. Synthetic Workloads

The workload employed in this test used a real database (Microsoft SQL Server) with database tables, indexes, etc., to perform actual database transactions. When using real database workloads, the I/O rate will vary as the workload progresses because the database performs operations that consume varying amounts of CPU and memory resources in addition to I/O resources. These results more closely resemble a real customer environment. This is unlike benchmarks that use synthetic workloads that perform the same I/O operations repeatedly, resulting in relatively steady I/O rates which, although potentially faster, do not resemble real customer environments.

The OLTP Database Workload

Demartek ran a transactional database workload to measure the performance of the MSA 2042 storage system. This workload performs real transactions that might be executed by database application users as well as background transactions from automated processes. The workload modeled a financial brokerage firm with customers who generate transactions related to trades, account inquiries, and market research. The brokerage firm in turn interacts with financial markets to execute orders on behalf of the customers and updates relevant account information. This workload consists of a mixture of mostly reads with some writes to its database.

It was not the intention of this exercise to produce database benchmarking results for publication and, in fact, the database was limited to a very low amount of system memory to force storage I/O at the expense of the database transactions. This was done to

demonstrate storage performance rather than server performance. Data and results published in this report are not comparable to any database performance results published in any other report or forum.

Workload Definition and Evaluation Objectives

- > The OLTP workload used was read-heavy, with about 5% of its I/O consisting of write requests. The database was populated with 3,500 GB of data and a 200 GB log.
- > Demartek provisioned 30 virtual users to generate a sustained workload on the HDD-only configuration with a storage latency of 18 milliseconds for a performance baseline. (This is typically considered the high end of acceptable response times for traditional spinning hard drives.)
- > This usage level seemed a reasonable simultaneous use case for a hypothetical medium sized business.
- > Demartek ran the benchmark for 24 hours in each configuration (hybrid flash and HDD-only) to ensure a steady state for I/O while supplying ample time for the SSD read cache to warm fully and for the tiering algorithm to migrate hot data to the flash tier.

Performance Metrics

Key metrics for storage system performance analysis are I/Os per second (IOPS), bandwidth, and latency or response time. These metrics are defined as follows:

- > **IOPS** – a measure of the total I/O operations (reads and writes) issued by the application servers.
- > **Bandwidth** – a measure of the data transfer rate, or I/O throughput, measured in bytes per second or megabytes per second (MBPS).
- > **Latency** – a measure of the time taken to complete an I/O request, also known as response time. This is frequently measured in

milliseconds (one thousandth of a second). Latency is introduced into the SAN at many points, including the server and HBA, SAN switching, and at the storage target(s) and media.

It is important to consider all three metrics when evaluating the performance of storage systems because all three contribute to how the storage will support an application.

IOPS drive bandwidth. The number of IOPS multiplied by the I/O request size determines the amount of bandwidth delivered. The database application used for this evaluation performs predominantly 8 kilobyte I/O.

Latency is important even though it doesn't necessarily have a direct effect on IOPS and bandwidth. It can have a very significant effect on application performance and user experience. Unlike IOPS and bandwidth—where more is better—with latency, the goal is to keep it as low as possible. The impacts of latency vary with the workload deployed. Some applications have a greater tolerance for higher latencies, while other applications are negatively impacted by even small increases in latency.

High bandwidth streaming (sequential) workloads may be able to tolerate higher levels of I/O response times, particularly where read-ahead buffering is employed. Data warehousing and video streaming are examples of applications where this may be true. Highly transactional workloads are more sensitive, particularly in cases where database transactions are time sensitive and have dependencies on prior transaction results. Applications performing real-time trend analysis, like weather forecasting or stock trading (similar to the model used in this evaluation), or applications that process lots of data fit into this second category.

Flash storage brings down I/O response times and drives up IOPS and bandwidth. Before flash storage became commonplace in the datacenter, storage I/O latencies of 10 to 20 milliseconds were generally acceptable for many applications, and that is why we chose 20 milliseconds as the upper limit for latency in

the HDD-only configuration. Latencies lower than two milliseconds are almost unachievable on spinning hard disk drives, simply because of the time it takes to perform the mechanical motions of the platters and heads.

With the MSA 2042 storage array, we were particularly interested in seeing how the storage response times would react to the use of flash for both read acceleration and for performance tiering.

Performance metrics can be taken at many points in the compute environment. Taking measurements at the host provides a complete picture of the user experience, including all of the additive effects of the downstream components of the compute system, such as the array, switching, cables, adapters, physical server and operating system, and application contributions. Our interest in this analysis is the user/application experience of a transactional database application running with the HPE MSA 2042, so we chose to conduct the measurements from the application host. Since we measured performance at the host, we included database transactions per second as an additional metric related to the user experience.

Performance measurements were taken with Windows Perfmon.

Results and Analysis

Every business is going to be different with respect to the performance demands that its applications place on the underlying compute infrastructure. The evaluation performed here makes assumptions on acceptable parameters such as number of simultaneous users (30) and the ceiling for acceptable response times (20 milliseconds). With these baseline parameters set, we saw that the MSA 2042 storage array was able to support 650 database transactions per second at slightly under 20 milliseconds of I/O response time.

Read Caching Results for MSA 2042

- > **Database Transactions per second: +35%**
- > **Latency: 4.5x Improvement**

> **Throughput: 2.5x Improvement**

> **IOPS: 3x Improvement**

Performance Tiering Results for MSA 2042

> **Database Transactions per second: +60%**

> **Latency: 11x Improvement**

> **Throughput: 4.6x Improvement**

> **IOPS: 4.5x Improvement**

With the addition of read caching to the array, we saw the transaction count increase by 35% for the same number of virtual users and the I/O latency reduce by a factor of 4.5, down to four milliseconds as compared to an HDD-only MSA configuration. This means that not only was more real work accomplished, but the user experience for this application was significantly improved as I/O service requests were completed in a fourth of the time.

This transactional SQL Server workload responded even better when the SSD performance tier was enabled. Tiering provides benefits by using flash to accelerate write transactions as well as reads. However, with writes making up a very small portion of this workload, it seems likely that the read transactions were able to not only take advantage of the flash capacity, but also benefitted from the RAID 1 configuration of the SSD performance tier since there were two flash drives to read from. As a result, when using the MSA 2042 with an 800 GB SSD tier, database transactions increased 60% over the baseline HDD-only configuration, while latency went down to below two milliseconds.

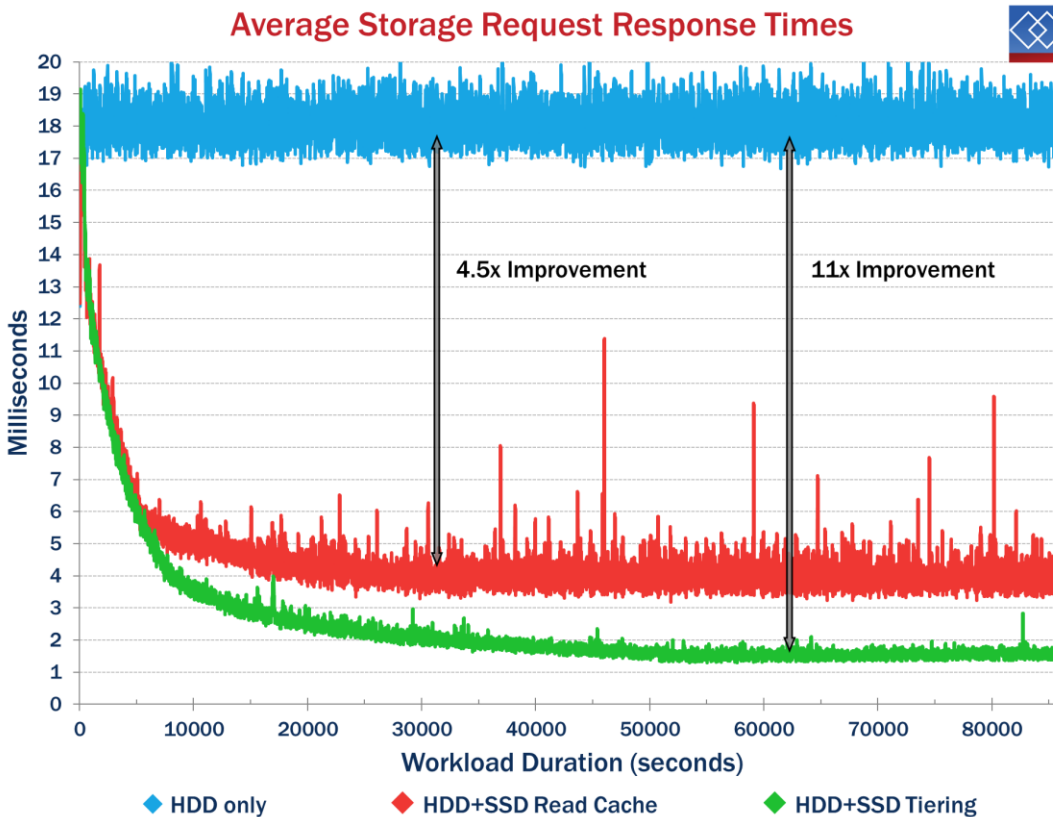
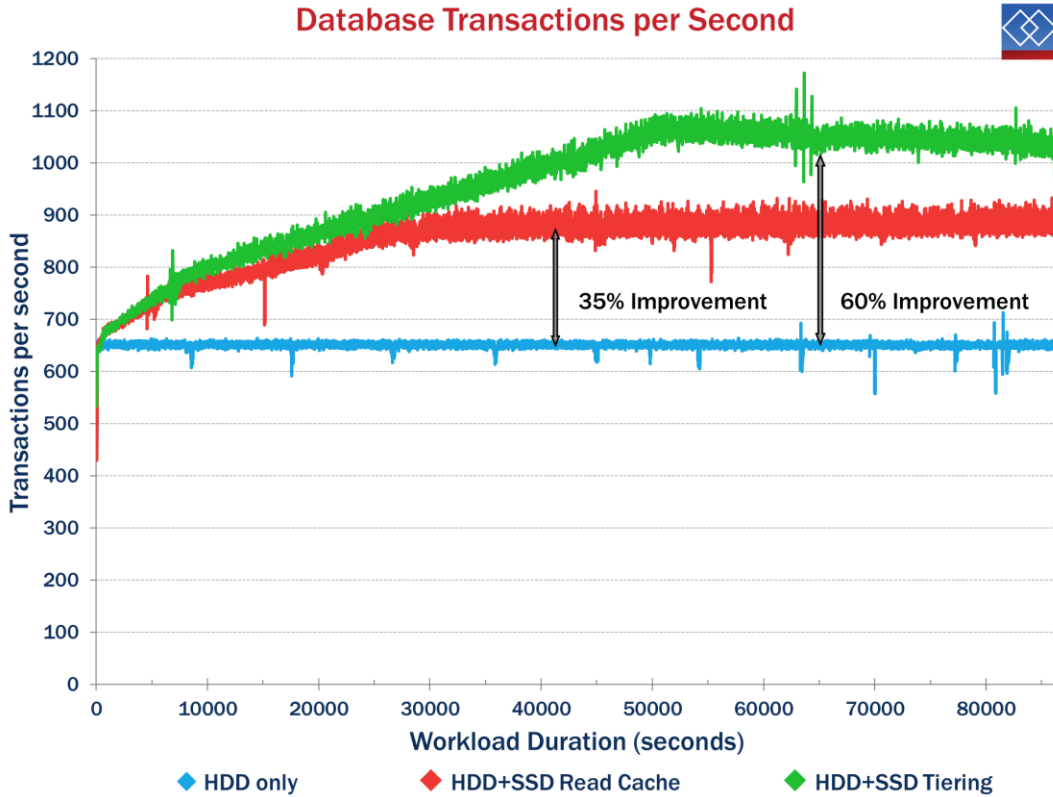
These performance improvements obviously came with corresponding increases in bandwidth and IOPS. We saw bandwidth increase by a factor of 2.5 with read caching and 4.6 with SSD tiering, from a baseline of 42 MB/s to 127 MB/s up to an average of 193 MB/s for the respective scenarios.

IOPS likewise increased from about 5,175 to 15,450 and again to 23,500 as we employed caching and then Performance Tier. These figures clearly

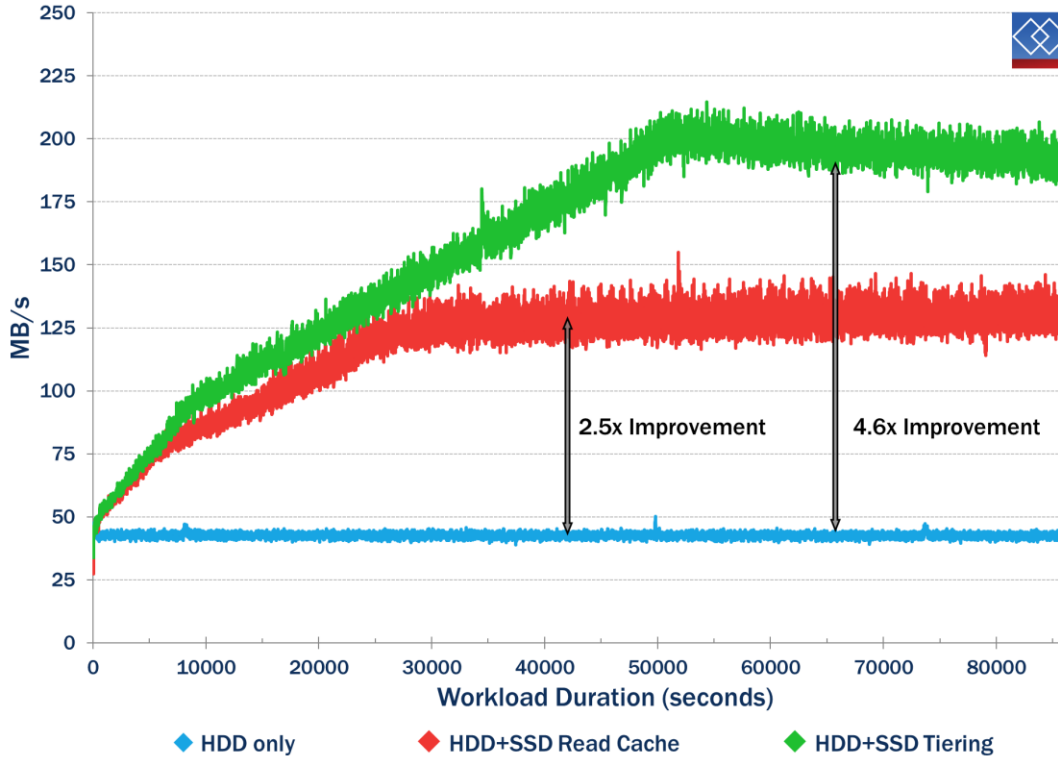
Accelerating Database Workloads Using HPE MSA Storage with Built-In Flash

display how the addition of a small amount of flash, compared to total storage and total data, has a significant effect on storage performance.

We didn't change the number of virtual users executing the workload, which may be why the ratio of bandwidth and IOPS increases didn't result in a one-to-one increase in transactions. Clearly the storage has the capacity to support additional virtual users and still keep I/O response times to an acceptable level.



MB per Second of Total Storage Bandwidth



IOs per second Total



Executive Summary

High-end, all-flash storage arrays may demonstrate extreme performance results, but are frequently too expensive and generally unnecessary to support the application needs of small to mid-sized businesses. The HPE MSA 2042 storage array provides an affordable flash solution for businesses looking to accelerate common application workloads such as transactional databases with a limited budget.

As storage system and server performance increase, it's not uncommon to see the service level expectations placed on application and user experience go up as well. For businesses of all sizes that need more performance from their storage than can be delivered by spinning drives alone, HPE provides an affordable way to take advantage of performance improvements delivered by flash in an incremental fashion that is simple to implement.

On-array read caching offers significant improvement in read IOPS, bandwidth, and response times while maintaining the redundancy of array-based flash. Adding additional SSDs to the system beyond the two drives included in the base configuration will produce greater performance results with read-intensive workloads. For example, we saw bandwidth more than double by adding about 10% of the total data capacity in SSD.

Performance tiering with HPE's proprietary algorithm needs more SSDs to protect against data loss when accelerating writes, but we found that this workload benefitted considerably with this option. The investment of two additional SSDs above the base model (four SSDs in two RAID 1 drive groups) netted a 60% increase in work accomplished with a tremendous reduction in response time. This is a best practice recommendation from HPE, although it is not necessary to take advantage of the system's performance tiering option. Customers could set a lower threshold of acceptable response time and still scale up the number of database users.

Another benefit to hybrid storage solutions like the MSA 2042 is that, when a workload I/O profile is understood,

the amount of flash assigned to read caching or the performance tier can be matched to the amount of "hot data" in that workload. Provisioning flash beyond this point would provide minimal benefit, as the active workload is already running on the fastest drives. If the storage array is tuned in this way, the spinning drives can potentially take on additional storage demands that can be satisfied with lower cost HDDs, improving the total return on investment of the array.

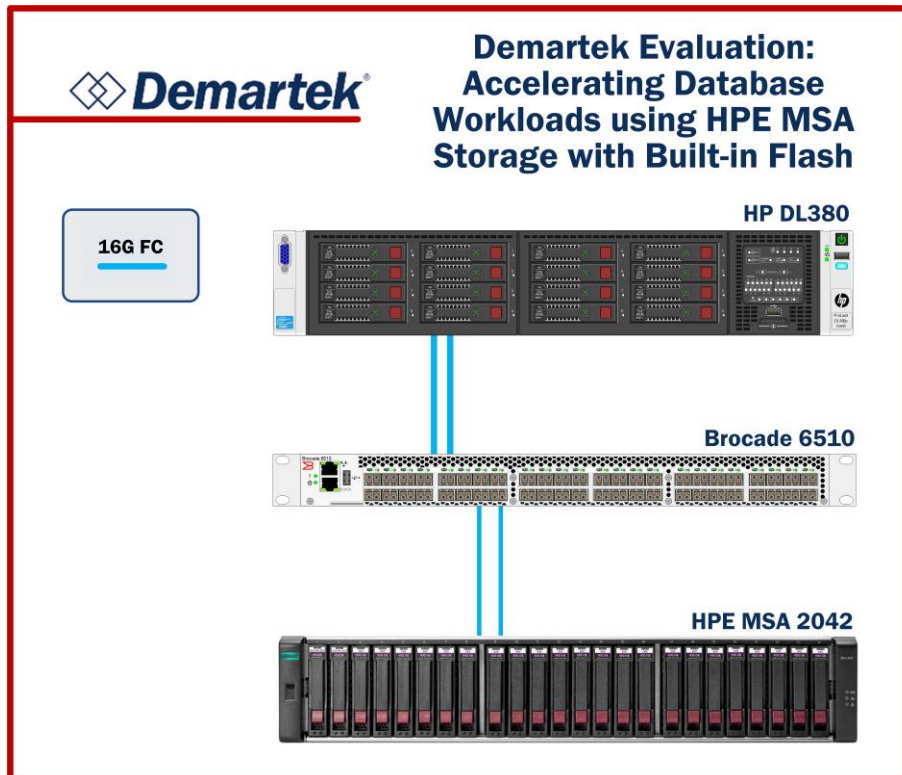
We concluded that for a read-intensive transactional workload on Microsoft SQL server, HPE MSA 2042 Storage was up to the task of supporting 30 simultaneous users without the use of flash—which seems reasonable for a small or mid-sized business. For businesses looking to gain the performance and operational benefits of flash, the HPE MSA 2042 with built-in flash and dynamic tiering capabilities is an affordable choice for database acceleration for a low starting price.

As previously stated, existing MSA 2040 customers can achieve these same results via a low-cost addition of SSD capacity and the Advanced Data Services Software Suite to their existing system.

Flash is enjoying mainstream adoption among large enterprises—it's now the rule rather than the exception—meaning that many enterprises have adopted a policy of deploying flash storage for all new production workloads by default. We now see more ways for small and mid-sized businesses to affordably deploy flash. For example, the MSA 2042 array evaluated in this report starts at \$10K US and has flash built-in as well as software needed to take advantage of that flash capacity.

This is an excellent time for small and mid-sized businesses to consider incorporating flash into their storage environments, particularly if they are looking to accelerate transactional workloads such as the real-world database workload used in this evaluation. Prices for flash continue to fall and there is greater flexibility in deployment, for example, the ability to add an SSD tier or to use flash for read caching.

Appendix A – Test Environment



Server

- > HPE ProLiant DL380 Gen 8
- > Two (2) Intel E5-2630 2.3GHz CPUs
- > 16 GB RAM
- > One (1) HPE SN1000E (16 Gb FC) dual port HBA
- > Windows Server 2012 R2
- > Microsoft SQL Server 2016, Microsoft Benchcraft

Fibre Channel Switch

- > Brocade 6510 16 Gb Fibre Channel Switch

Storage Array

- > HPE MSA 2042 storage array with 800 GB of flash built-in and software support included for advanced virtualization features necessary to take advantage of this capacity.
- > Read Cache is a standard feature. MSA Performance Tiering license (standard) was used in SSD tiering scenarios.
- > Twenty (20) 900GB 10K RPM 6 Gb SAS HDDs
 - > Ten (10) drive RAID 6 per storage controller
 - > Six (6) data volumes – 3 per controller
 - > One (1) log volume
- > Four (4) 400GB SSDs
 - > One (1) drive per storage controller for read caching
 - > Two (2) drives RAID 1 per storage controller for Performance Tier
- > Four (4) 16 Gb Fibre Channel targets per controller (2 ports active per controller)

Accelerating Database Workloads Using HPE MSA Storage with Built-In Flash

The most current version of this report is available at http://www.demartek.com/Demartek_HPE_MSA_2042_SQL_Server_Environment_2016-08.html on the Demartek website.

Intel and Xeon are registered trademarks of Intel Corporation.

Microsoft, Windows, and Windows Server are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Demartek is a registered trademark of Demartek, LLC.

All other trademarks are the property of their respective owners.