Demartek

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

Adding a small amount of flash storage to MSA storage can provide significant performance improvements

Executive Summary

Solid state storage shows up in many places in the datacenter today, from all-flash arrays to direct attached drives and PCIe flash to accelerators that fit somewhere in between servers and storage. With the price of the flash continuing to drop, it is becoming increasingly available to the small and medium business for critical computing applications and day-to-day operational computing.

The HPE MSA 2050 Storage array delivers enterprise-class storage at a cost that may be attractive to small and medium businesses. These arrays can offer improved performance by combining hard disk drives (HDD) with solid state drives (SSD) for read caching or performance tiering. HPE commissioned Demartek to evaluate the performance of a Microsoft SQL Server database workload on the MSA 2050 Storage Fibre Channel array in an all-HDD configuration and then repeat the workload with the SSD read caching and SSD performance tiering options installed and configured.

We compared several performance metrics including throughput, IOPs, number of database transactions per second, and I/O latency for a complete picture of the MSA 2050 Storage's suitability to support the scale of workload a small to medium size business might experience.

Key Findings

> **Latency:** Compared to the HDD-only configuration, the *read cache configuration* achieved **70% lower average latency** and the *performance tier* achieved **84% lower average latency** for the logical storage volumes as measured by the operating system.

> **Throughput:** Compared to the HDD-only configuration, both the *read cache configuration* and the *performance tier* achieved more than **3x higher average throughput.**

> **IOPS:** Total storage IOPS for the *read cache configuration* and *performance tier configuration* were more than **3x the HDD-only configuration.**

> Adding a small amount of flash storage to MSA storage can provide significant performance improvements

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

Demartek
May 2018

The HPE MSA 2050 Storage

The HPE MSA 2050 Storage array is a dual-controller, two rack-unit (2U) storage system with support for up to 192 small form factor (SFF or 2.5-inch) drives or 96 large form factor (LFF or 3.5-inch) drives. The MSA 2050 supports 8 Gb and 16 Gb Fibre Channel, 1 Gb and 10 Gb iSCSI and/or 12 Gb SAS host protocols. The MSA model 2052 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 MSA base configuration of 24 drive bays can be populated with HDDs or SSDs. The base enclosure can be expanded with additional SFF or LFF enclosures. The MSA's Advanced Virtualization features manages the addition of SSDs to an array by enabling the creation of an SSD read-cache or configuring it as a full R/W tier using the MSA Advanced Data Services SW License (included with the model 2052, available as an added feature with the model 2050). Both features use a real-time sub-LUN tiering engine which migrates active data between HDDs and SSDs every five seconds.

For this evaluation, HPE provided Demartek with an MSA 2052 array with 20 1.2 TB 10K RPM HDDs and four 800 GB SSDs. With this hardware, we were able to set up a 20 HDD baseline test case, with 10 drives provisioned to each controller in a RAID 6 configuration. Then to evaluate read caching, we added a single SSD to each controller and assigned those SSDs as read cache type in the virtual disk group. For measuring the effect of the optional Performance Tiering license, 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 for measuring the performance of the 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 models 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 is read-heavy, with about 5% of its I/O consisting of write requests. The database was populated with 3,500GB of data and a 200GB log. We 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 time for traditional spinning hard drives.

This usage level seemed a reasonable simultaneous use case for a hypothetical medium sized business. We ran the benchmark for 24 hours in each configuration 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), throughput, 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.

> **Throughput** – 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 throughput. The number of IOPS times the I/O request size determines the amount of throughput delivered. The database application used for this evaluation performs predominantly 8 kilobyte I/Os.

Latency is important even though it doesn't necessarily have a direct effect on IOPS and throughput. It can have a very significant effect on application performance and user experience. Unlike IOPS and throughput, 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 throughput streaming or sequential workloads may be able to tolerate higher level 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 has been bringing down I/O response times as well as driving up IOPS and throughput. Before flash storage became commonplace in the datacenter, storage I/O latencies of 10 to 20 milliseconds were generally acceptable for many applications, and is why we chose 20 milliseconds as the upper limit for latency in the HDD-only configuration. Latencies lower than 2 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 option to add flash to the MSA 2050 Storage array, we were particularly interested in seeing how the storage response times would react to the addition of a small amount of flash.

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 HPE MSA 2050 Storage 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.

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

Results and Analysis

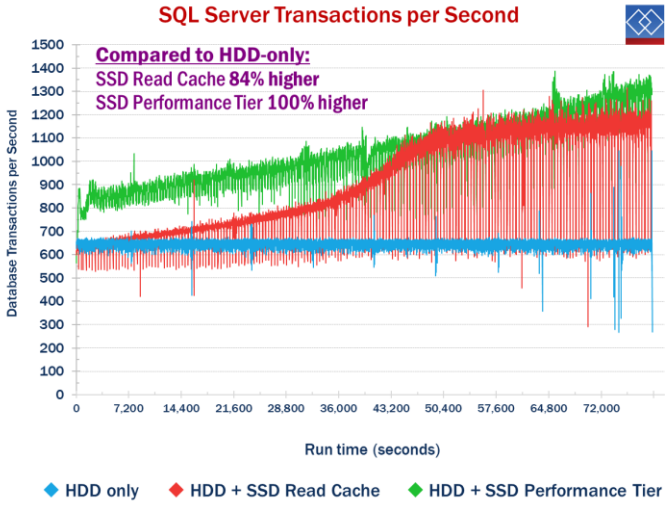
Every business is going to be different with respect to the performance demands its applications place on the underlying compute infrastructure.

Database Transactions per Second

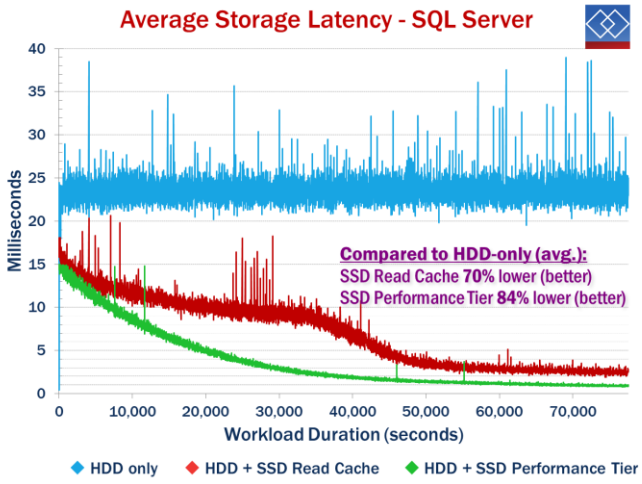
With our test parameters set, we saw that the MSA 2050 Storage was able to support an average of 649 database transactions per second at approximately 23 milliseconds of I/O response time with the all-HDD configuration.

With the addition of read caching to the array, we saw the database ***transaction count increase by 84% percent*** for the same number of virtual users and the I/O latency reduce by 70% to less than three milliseconds.

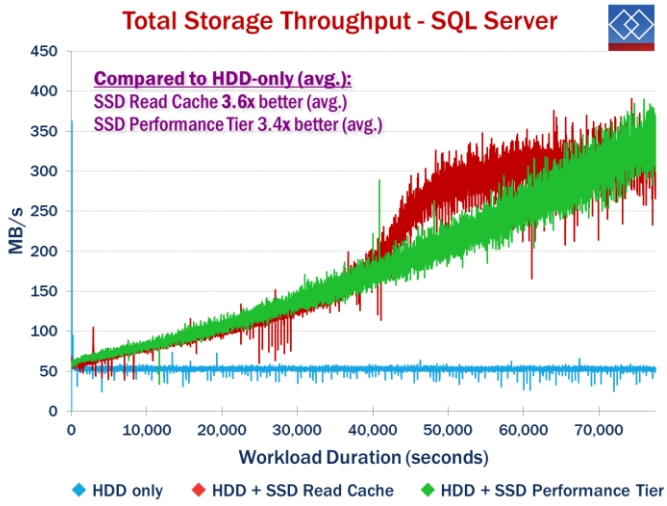
This transactional SQL Server workload responded best when an SSD performance tier was added. Tiering provides benefits by accelerating write transactions with flash as well as reads. However, with writes making up a very small portion of this workload, it seems likely that the read transactions were able to not only take advantage of the flash, but also benefitted from the RAID 1 configuration of the SSD performance tier by have two drives to read from. Database ***transactions doubled*** over the baseline while latency went down by 84% to approximately one millisecond.



Latency



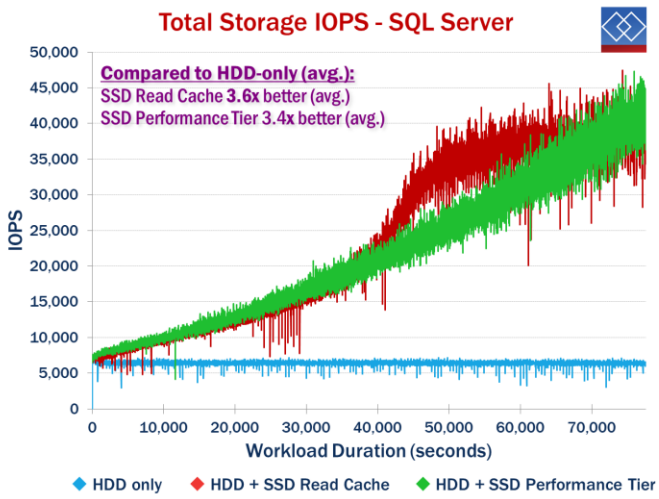
Throughput



These performance improvements obviously came with corresponding increases in throughput and IOPS as demonstrated in Figure 2. We saw throughput increase by a factor of 3.6 with read caching and 3.4 with SSD tiering, from a baseline of approximately 50 MB/s.

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

IOPS



IOPS likewise increased more than 3x for both read caching and performance tiering from an average of approximately 6400 for the HDD-only configuration.

These figures clearly 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 throughput 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.

Summary and Conclusion

High-end all-flash storage may have extreme performance figures, but it is expensive and generally unnecessary to support the application needs for small-to-medium business. The HPE MSA 2050 Storage array is intended to be affordable for small and medium sized businesses and can run enterprise applications such as a transactional database with acceptable performance (it is, of course, up to each business to determine just what the range of acceptability is). A business may find the performance delivered by a basic, HDD-only array sufficient to meet operational requirements, and if so, no more needs to be said. But HPE provides other options.

However, as storage system and server performance increases, it's not uncommon to see the service level expectations placed on application and user experience go up as well. For businesses that need more performance out of storage systems than can be delivered by spinning drives alone, HPE provides a way to take advantage of performance improvements delivered by flash in an incremental fashion that may be easier on the pocketbook, and is simple to implement.

On-array read caching, offers significant improvement in read IOPS, throughput and response times, while maintaining the redundancy of array-based flash. The size of the cache is limited only by the number of SSDs (currently up to 4TB) and a savvy administrator should be able to tune the amount of flash to the workload. We saw throughput more than double by adding about 10% of the total data capacity in SSD.

Performance tiering with HPE's proprietary Performance tiering 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 four SSDs (in two RAID 1 drive groups) netted a 60% increase in work accomplished with a tremendous reduction in response time. If we had chosen to do so, we could have set a lower threshold of acceptable response time and still scaled up the number of database users.

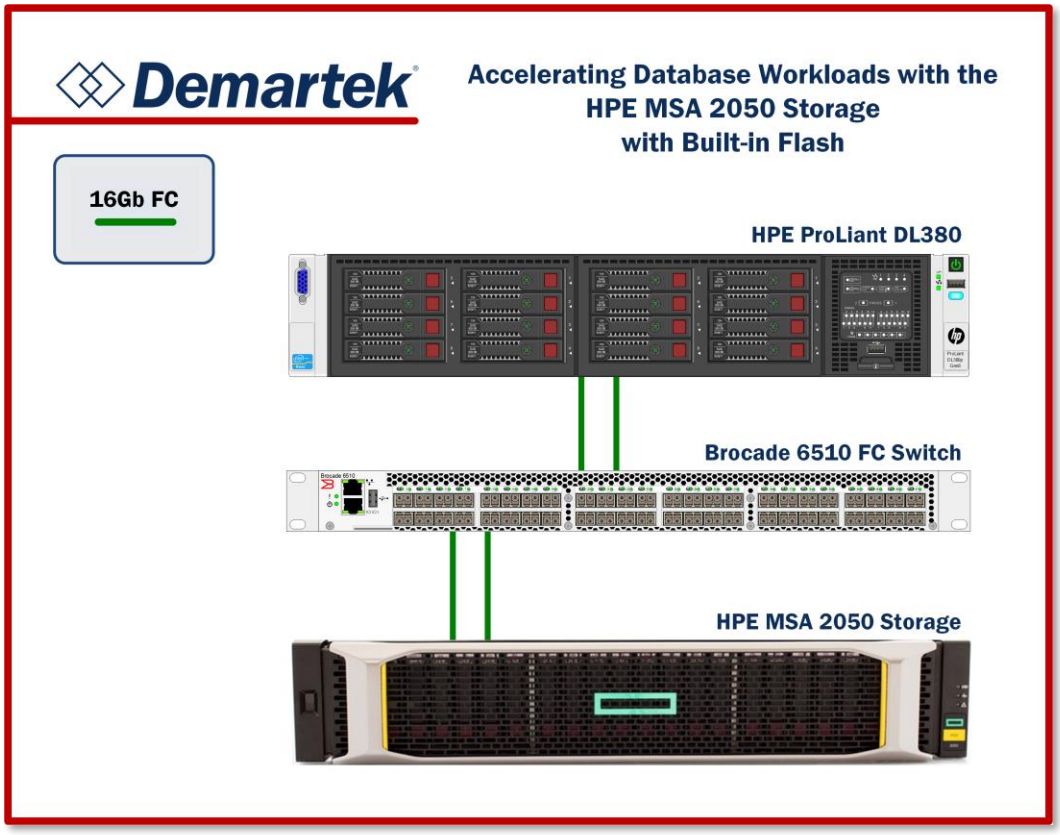
Another benefit to hybrid storage solutions like the MSA 2050 Storage is that when a workload I/O profile is understood, the amount of cache or performance tier can be configured to the amount of hot data in that workload. This requires some research on the part of the storage administrator, but HPE can help here too with detailed device level metrics available through the user interface. 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 demand that can be satisfied with slower HDDs, improving the total return on investment of the array.

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

Ultimately, a business must understand its workloads to determine the best storage solutions for its needs. We concluded that for a read-intensive transactional workload on Microsoft SQL Server, the HPE MSA 2050 Storage array was up to the task of supporting 30 simultaneous users out of the box, with no upgrades, which seems reasonable for a small or medium sized business. If that same business is interested in "getting its feet wet" with flash storage solutions, HPE has affordable upgrades that can significantly enhance application performance and usability.

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

Appendix A – Microsoft SQL Server Test Environment



Server

> HPE ProLiant DL380 Gen 8
> 2x Intel Xeon E5-2630, 2.3 GHz, 128GB RAM
> 1x HPE SN1100Q Dual-port 16Gb FC HBA
> Windows Server 2016 Datacenter
> Microsoft SQL Server 2017

Fibre Channel Switch

> Brocade 6510 16Gb FC Switch

Storage Array

> HPE MSA 2050 SAN Storage Array
> 4x 800 GB SSD
> 20x 1.2 TB HDD

Accelerating Database Workloads with the HPE MSA 2050 Storage with Built-in Flash

Demartek

May 2018

The most current version of this report is available at http://www.demartek.com/Demartek_HPE_MSA_2050_SQL_Server_Environment_2018-05.html on the Demartek website.