Demartek

Accelerating Exchange Server 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

For many businesses, email is often considered a mission-critical application. In addition to facilitating intra-company communications, many small to medium businesses rely on email as a method of communicating directly with customers, be it primary contact, confirmation of orders, support requests, or for other customer service functions. Adding the inevitable personal use and a seemingly never-ending assortment of spam email such as phishing, ransomware and more, email systems can be quite busy. It's vital for business continuity to know that corporate email remains operational and efficient 7x24.

The HPE MSA 2050 Storage is positioned in the market as a high-performance hybrid storage system. Upgradeable with a customer configurable number of solid state drives (SSD), a base-model all hard disk drive (HDD) array can be enhanced to provide SSD read caching or an SSD performance tier to accelerate array performance. HPE commissioned Demartek to evaluate the MSA 2050 Storage array as backend storage for Microsoft Exchange Server, configured as it might be deployed by a typical small or medium sized business.

We ran Microsoft Exchange Jetstress 2013 on an all-HDD array as a baseline and again with SSD read caching and performance tier upgrades to simulate the Microsoft Exchange I/O load and database I/O patterns. We considered the Jetstress output reports along with Windows Perfmon data to evaluate the suitability of the MSA 2050 Storage in these three configurations for this purpose.

We found that with the I/O load configured for a heavy usage simulation, the MSA array supported 500 mailboxes of size 4GB each in an all-HDD configuration with acceptable maximum database latencies of 16 milliseconds. However, SSD tiering produced an impressive improvement on this baseline. The MSA array accelerated performance with the addition of SSD real-time tiering by reducing latency 79% while increasing write throughput by a factor of 4.2X.

Key Findings

> **Latency:** Compared to the HDD-only configuration, the *read cache configuration* achieved **53% lower average latency** and the *performance tier* achieved **79% lower average latency** for the logical storage volumes as measured by the operating system.

> **Throughput:** Compared to the HDD-only configuration, the *read cache configuration* achieved **2x higher average throughput** and the *performance tier* achieved **more than 4x higher average throughput**.

> **IOPS:** Total storage IOPS for the *read cache configuration* were more than **2x the HDD-only configuration** and the *performance tier* configuration achieved **nearly 5x the IOPS** of the HDD-only configuration.

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

Accelerating Exchange Server 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 2050 array with 20 1.2TB 10K RPM HDDs and four 800 GB SSDs. With this hardware, we were able to set up a 20 HDD baseline test case, with ten 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.

Microsoft Exchange Jetstress 2013

Microsoft Exchange is one of the dominant email server applications in use today. Many small and medium sized businesses that do not offload their email to third parties run their own Exchange servers. Email is a latency sensitive application (just ask anyone waiting for an important email), and therefore properly designing and scaling storage is critical to an acceptable user experience.

Microsoft Exchange Jetstress 2013 models the I/O loading and patterns that would be generated by an Exchange 2013 server and Exchange 2016 server, including email database reads and writes as well as logging. Its configuration includes the number and size of mailboxes, number of mailbox databases and logs, and intensity of use, among other things. Jetstress uses the same Extensible Storage Engine files used by Microsoft Exchange to ensure that the Jetstress modelled performance is representative of a bona fide Exchange Server of the same version. The key metric used by Jetstress to pass or fail an email system is latency. Jetstress marks a test a failure if any Exchange database experiences an average transactional latency of 20 milliseconds or higher.

Microsoft recommends Exchange Jetstress be used to validate email systems before putting those systems into production.

Real vs. Synthetic Workloads

Demartek prefers to test storage systems with real workloads that users are likely to run on their systems as opposed to synthetic workloads that tightly control I/O. While Jetstress does not deploy a full Microsoft Exchange package, it is our opinion that it does a satisfactory job of simulating an operational Exchange environment. It achieves this by requiring the same physical hardware and configuration that would be used for the production email system as well as the same underlying software engine used by a full installation of Microsoft Exchange.

Workload Definition and Evaluation Objectives

Microsoft Exchange Jetstress 2013 exercises the storage system with read and write I/O to simulate email database I/O. We envisioned a hypothetical small-to-medium sized business with 500 mailboxes and sized the storage to support each mailbox up to 4 GB in size. To support this scale of mail server, we configured Jetstress to build 6 mailbox databases on 200 GB volumes, with corresponding log volumes. We spread these volumes equally across the two storage controllers of the array.

Accelerating Exchange Server with the HPE MSA 2050 Storage with Built-in Flash

Jetstress can be configured to exercise all the storage space given it, and we did so. This has the potential to be somewhat cache unfriendly as it reduces hot spots on the drives, but we wanted to make this a difficult workload for the storage. We also configured Jetstress to perform .5 IOPS per mailbox, which is considered a heavy load for an Exchange 2013 email server.

We performed an initial 30 minute tuning run to allow Jetstress to build the database and identify a thread count, then a 2 hour execution at the proposed thread count followed by a 24 hour test to make sure that the system reached a steady I/O state and to allow the read caching and Performance Tier scenarios time to fully migrate data to SSD (this is also the test method that Microsoft requires for publishing results through the ESRP program, though it was not our intention to publish outside of this report). The data and analysis included in this report is from the 24-hour test run. The mailbox databases were rebuilt prior to beginning each new test.

The purpose of this test was to show performance comparisons in a typical SMB data center, not the maximum load. The environment with 500 mailboxes represents a common configuration to demonstrate the benefit of adding SSD as read cache or SSD as a tier in the array.

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.

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. Jetstress sets an upper limit of 20 milliseconds for database reads and writes. Unlike IOPS and throughput, where more is better, the goal with latency is to keep it as low as possible.

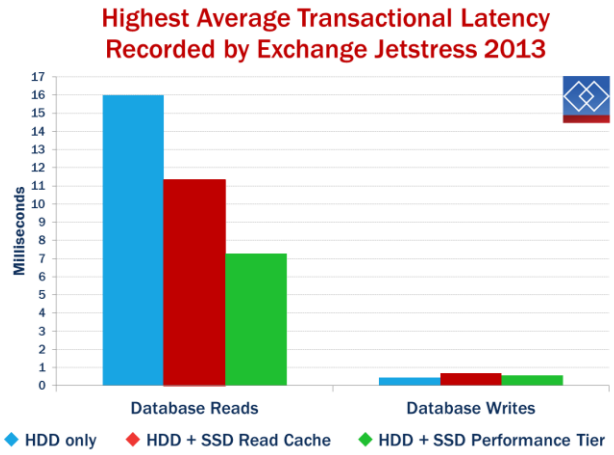
Jetstress collects its own performance metrics, which are viewable through Windows Perfmon. We converted these to CSV files for our analysis. Jetstress also performs its own analysis to score test runs which we used to confirm that the tests were within the Jetstress limits for a successful run and validation of the storage system.

Accelerating Exchange Server with the HPE MSA 2050 Storage with Built-in Flash

Results and Analysis

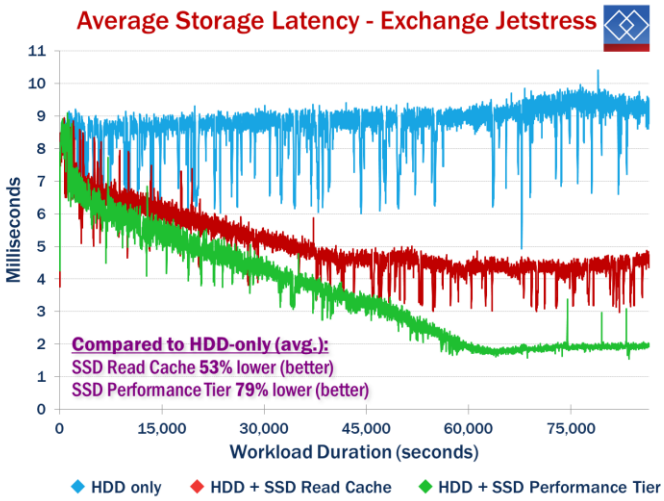
Exchange Jetstress Latency

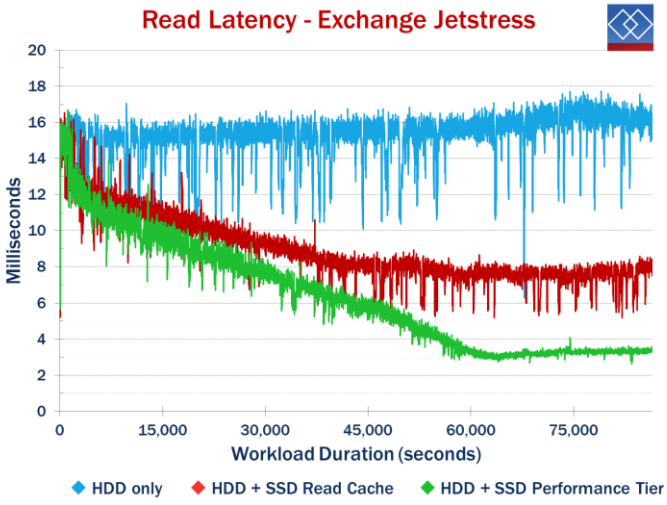
Unlike some other applications, where a business can set its own standards for acceptable performance, Microsoft Exchange Jetstress expects to see database read response times below 20 milliseconds or it will fail a configuration. Our test of the HPE MSA 2050 Storage, simulating 500 mailboxes on a heavily loaded email system, met the Jetstress response time requirements for all configurations: HDD-only, HDD plus an SSD read cache, and HDD plus SSD performance tiering.

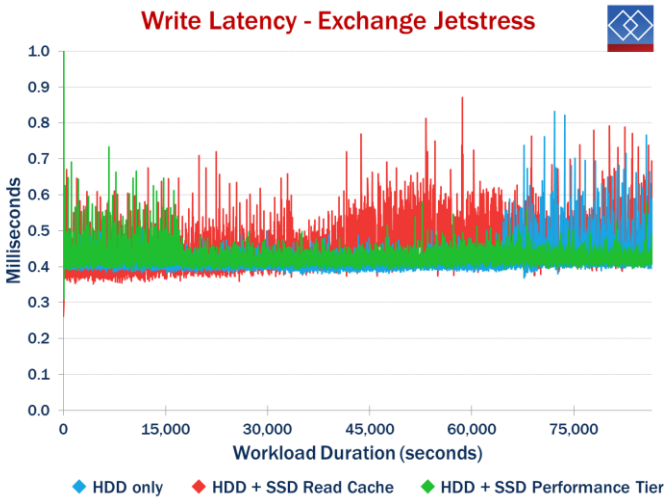


The highest latencies measured by the Exchange Jetstress application were consistently generated by read transactions. These values are the total of all system contributions to latency, including the storage device and SAN infrastructure plus server and application latencies. Scores above 20 milliseconds would prevent Jetstress from validating the system for Exchange 2013 with the number of mailboxes and IOPS we configured. As shown above, the highest response time, at 16 milliseconds for the HDD-only configuration, fell within acceptable limits. The flash accelerated configurations were even better. This tells us that the MSA 2050 Storage is a suitable Exchange Server 2013 storage device for 500 rather busy mailboxes, even without any flash upgrades.

Logical Volume Latency (measured by O.S.)







Accelerating Exchange Server with the HPE MSA 2050 Storage with Built-in Flash

The data measured by the operating system represents the average response times of I/O on the data and log volumes provisioned from the HPE MSA 2050 Storage, not including application-introduced latency. Again, we can see that read I/O latency clearly received the largest benefit from flash upgrades in the storage. The baseline HDD-only configuration performed I/O transactions with a maximum average read I/O response time of 9.5 milliseconds. With the addition of SSDs for read caching and Performance Tier performance tiering, latency drops by 53% and 79%, respectively, to roughly 4.5 milliseconds and 2 milliseconds.

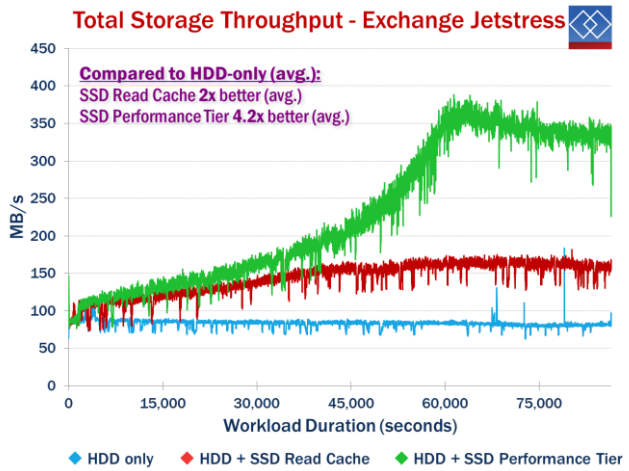
We can also see that database writes were never an issue, with every configuration maintaining a very low latency; the average being well below 1 millisecond.

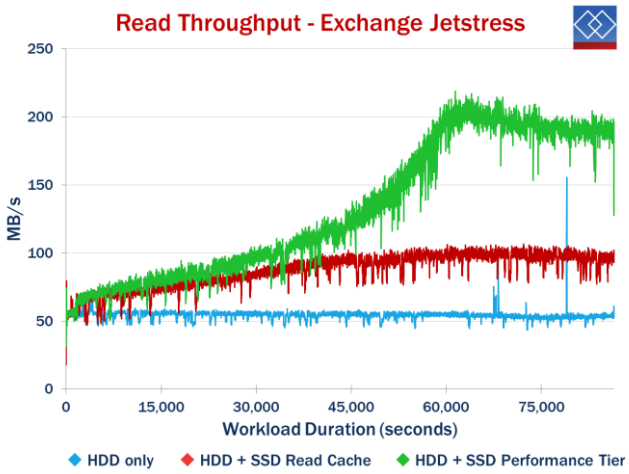
Logical Volume Throughput

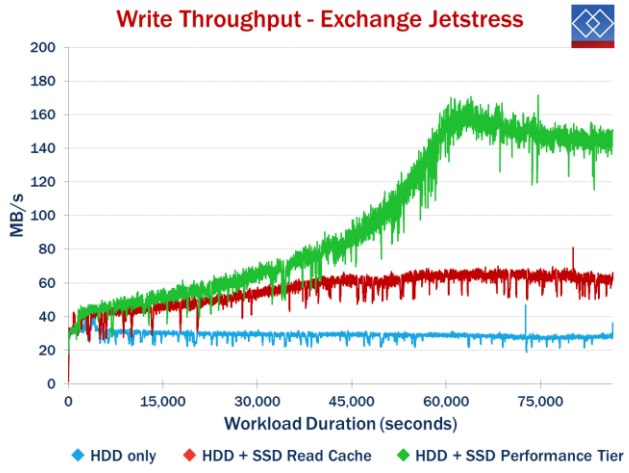
The throughput charts demonstrate the corresponding increase in read/write throughput delivered by the MSA 2050 Storage as we employed the flash upgrades. The workload is cache-unfriendly by definition. The Jetstress parameters were configured to minimize hot spots on the drives by spreading I/O across all provisioned drive space, and we can see that in the data. The abruptly halted cache warming curve and moderate performance gains of the HDD plus read cache configuration point to a strong degree of randomness in the I/O pattern, which can be hard to cache when the drive space in use is greater than the amount of cache provisioned.

In spite of this, read caching still provided approximately 2 times throughput gain over a baseline of for both reads and writes. While the read caching provides a moderate improvement, the real improvements for this workload come from the Performance Tier performance tiering solution.

The performance tiering algorithm and five second data migration interval seem to be quite effective in getting active data onto the flash tier for this workload. We recorded an overall average of 4.2 times improvement in throughput over the HDD-only baseline. If we consider this along with the impressive improvement in latency, there seems to be a great deal of headroom in which we could add additional mailboxes or potentially another application.



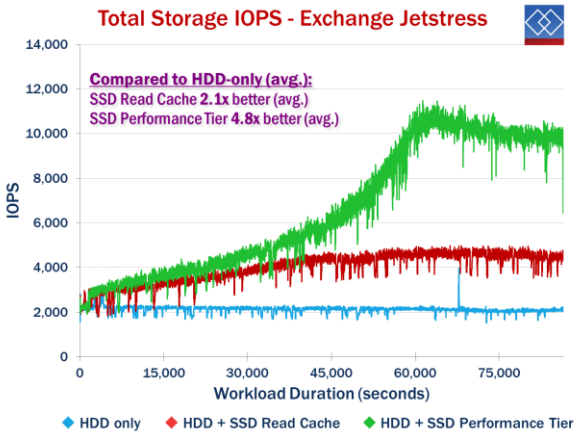


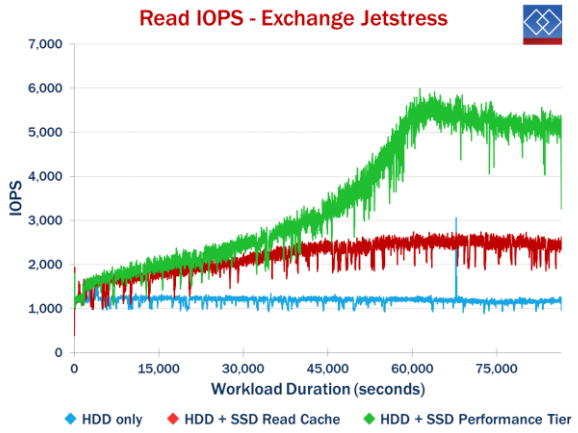


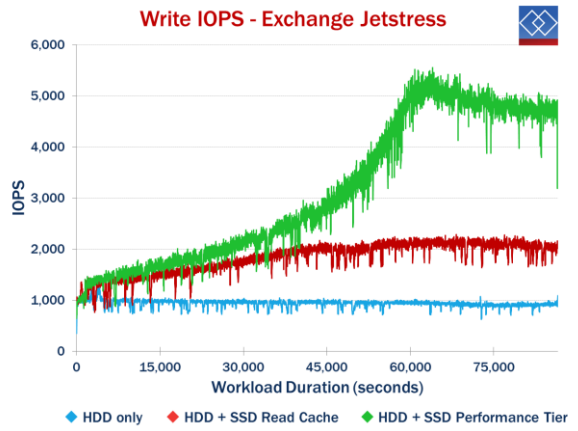
Accelerating Exchange Server with the HPE MSA 2050 Storage with Built-in Flash

Logical Volume IOPS

As with the throughput, the logical volume IOPS showed similar performance improvement with the read cache and the performance tier. It should be clear that adding a relatively small amount of SSD to the storage system makes a significant difference.







Summary and Conclusion

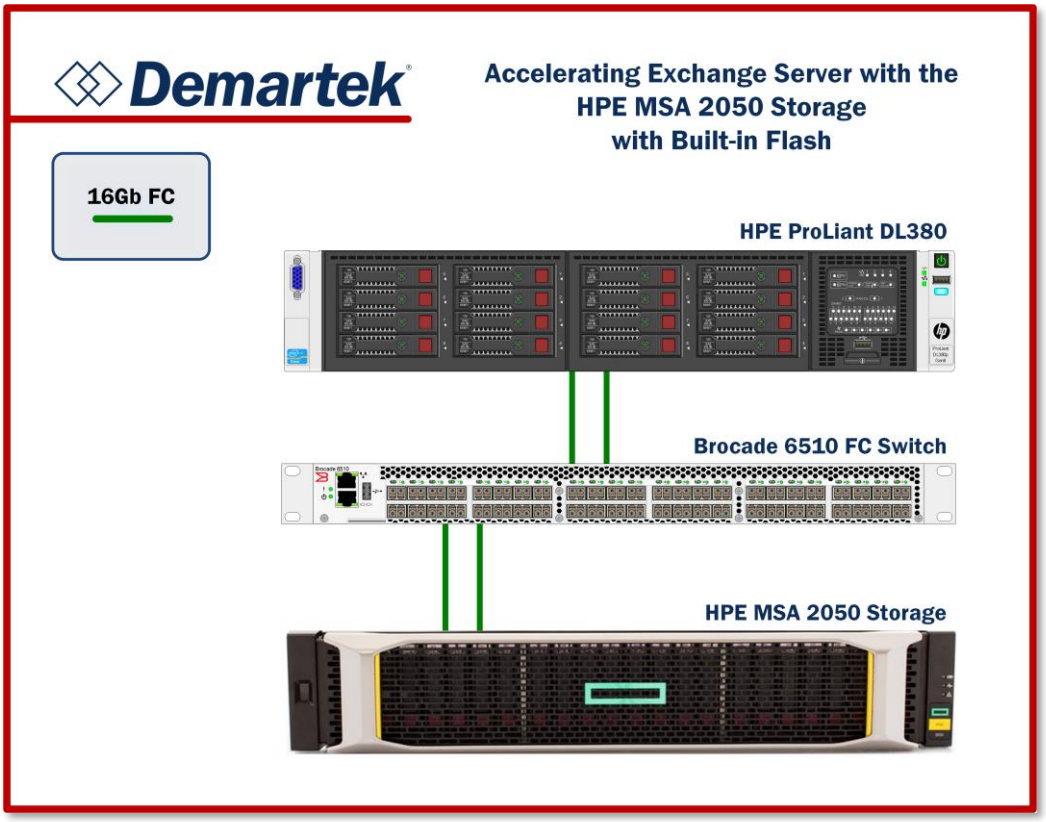
MSA hybrid arrays with real-time tiering clearly improved performance in terms of reduced latency and increased read and write throughput for this test case. Based on Jetstress modelling, the MSA 2050 Storage easily supports an email application with 500 very busy Microsoft Exchange 2013 mailbox users. While the addition of SSDs to improve performance is not required, we would advise businesses not to ignore the advantages that a hybrid array with real-time tiering offers over all-HDD configurations. On the other hand, the all-HDD configuration is less expensive. Deploying Performance Tier SSD tiering will significantly reduce latency for a 500-mailbox environment and provide "future-proofing" in the event that the Exchange server usage goes up. A growing business would be wise to consider where its Exchange requirements might be in two to five years.

Another consideration for a flash enhanced MSA 2050 Storage would be to support additional workloads. The environment we deployed in this evaluation did not demand anywhere near the entire drive space provided by the array. To keep latency below 20 milliseconds for the HDD-only array configuration, we are essentially short stroking the drives. This is wasteful in drive space and in return on investment of the array. SSD performance tiering improved latency by 79%, giving us the option to deploy more data in the form of additional mailboxes or an entirely different application altogether onto the array while still keeping latency within the recommended threshold for Microsoft Exchange. We would recommend that a business seeking to deploy storage in a cost-effective manner give this option serious consideration.

Accelerating Exchange Server with the HPE MSA 2050 Storage with Built-in Flash

Demartek
May 2018

Appendix A – Microsoft Exchange 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 Exchange Jetstress 2013
> Jetstress core engine v15.01.1019.000
> Exchange Db. storage engine v15.01.0225.037
> Exch. Db. Performance Lib. v15.01.0225.037

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

Exchange Jetstress Parameters

> 500 mailboxes, 4 GB per mailbox
> 40 threads, 0.5 IOPS per mailbox
> 6 Exchange databases, 2 copies each
> Each database on its own volume (drive letter)

Accelerating Exchange Server 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_Exchange_Server_Environment_2018-05.html on the Demartek website.