Evaluation Report: Supporting Microsoft Exchange on the Lenovo S3200 Hybrid Array

Evaluation report prepared under contract with Lenovo

Executive Summary

Love it or hate it, businesses rely on email. It's a mission critical application and outages can cost businesses a lot in reduced productivity and in revenue lost through customer frustration. Email systems see a lot of use, and like many business infrastructure systems, they may not get much attention until they stop working correctly or aren't meeting business needs. The decision to keep an email service in-house, or contract it out to a third party, is based on cost along with ease and flexibility of management. Small and medium sized businesses are frequently quite cost conscious and demand an affordable, well-performing hardware platform to run an in-house email system.

Lenovo satisfies this demand with the Lenovo S3200 storage area network (SAN) hybrid array. The S3200 array is a hard disk drive (HDD) array that can be upgraded with a customer configurable number of solid state drives (SSD). Base-model all hard disk drive (HDD) arrays may be enhanced to provide SSD read caching or an SSD performance tier to accelerate I/O performance. Lenovo commissioned Demartek to evaluate the S3200 storage array as backend storage for Microsoft Exchange Server for a small-to-medium sized business.

We modelled Microsoft Exchange with Microsoft Jetstress 2013, first on an all-HDD array as a baseline, and again with SSD read caching and performance tier upgrades. We found that with the I/O load configured for a heavy usage simulation, the Lenovo S3200 array supported 500 mailboxes in an all-HDD configuration with database latencies of 17.5 milliseconds or less. This latency was reduced by 75% while increasing bandwidth up to a factor of 4.4X when we implemented SSD performance tiering.

The Lenovo S3200

The Lenovo S3200 Storage Array tested in this evaluation (Figure 1) is a dual controller, 2 rack unit, 24 small form factor drive storage array (this array is also available with 12 large form factor drives). The S3200 model supports 8Gb and 16Gb Fibre Channel along with 1Gb and 10Gb iSCSI. Each array can support seven expansion units for up to 384 TB of total storage capacity. The base configuration of 24 HDDs is upgradable by replacing some HDDs with flash storage.

Figure 1 – Lenovo S3200

The Lenovo SAN Manager storage operating system supports SSDs in either a read cache or a tiering capacity—called Lenovo Intelligent Real Time Tiering. Read caching places the most frequently accessed data onto solid state storage for quick retrieval whereas Intelligent Real Time Tiering™ dynamically stages high priority data on SDDs, with lower priority data on HDDs to accelerate both reads and writes (data is moved every five seconds). These capabilities enable applications to gain the benefits of flash performance without the need to buy an expensive all-flash array.

Lenovo provided Demartek with the Lenovo S3200 array with twenty 900 GB 10K RPM HDDs and four 400 GB SSDs for this evaluation. We set up a twenty HDD baseline test case. Ten drives were provisioned to each controller as a RAID 6 disk group. Later, we added either a single SSD for creating read cache disk groups or two SSDs in a RAID 1 disk group as performance tiers to each controller.

Microsoft Exchange Jetstress 2013

Microsoft Exchange is one of the dominant email server applications. 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 anxiously waiting for an important email to arrive), 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, including email database reads and writes plus logging. Its configuration includes the number and size of mailboxes, number of mailbox

databases and logs, and intensity of use. 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 certify 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

At Demartek, we prefer 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. We consider Jetstress an exception to this rule. Jetstress doesn't deploy a full Microsoft Exchange package, but we feel it does a good job simulating a bona fide Exchange email system. This is because Jetstress must be run on the same physical hardware that is intended to be used for the production email system as well as the same underlying software engine used by a full blown installation of Microsoft Exchange.

Workload Definition and Evaluation Objectives

We envisioned a hypothetical small-to-medium sized business with 500 mailboxes. The storage requirements were sized to support these mailboxes up to 2 GB in size. For this size mail server, we had Jetstress build six mailbox databases on six 200 GB volumes equally distributed across the array's two controller. Corresponding log volumes were created for each database[1]. These volumes were spread out equally across the array's two storage controllers.

Jetstress can exercise any percentage of the storage provisioned for its databases. We configured it to use all space equally, meaning there would be no idle spots on the volumes. Most email systems in real-life won't do this; it's a worst-case drive use scenario and has the potential to be cache unfriendly as it prevents a "hot spot", where some data is accessed more often than other data, but we wanted to make this a difficult workload

[1] Jetstress logging simulates write I/O to database log partitions, but since there is no "real" data that would ever need to be recovered in the modelled email databases these partitions are quite small. Our log partitions were 10GB in size and only a fraction of that space was consumed.

for the storage system. We also configured Jetstress to perform .5 IOPS per mailbox, which is considered a heavy load for an Exchange 2013 email server.

We ran the simulation for 30 minutes initially to build the databases and settle on a thread count. After tuning, we used the Microsoft ESRP criteria to evaluate the Lenovo S3200. This meant a two hour test execution followed by a 24 hour test. It wasn't our intention to publish any results to the ESRP program, but we chose use the test method because it is recognizable and reproducible. It is also long enough to make sure that the system reached a steady I/O state by allowing the read caching and SSD tiering scenarios time to fully migrate data to flash. The data and analysis included in this report is from the 24 hour test run interval. The databases were restored prior to beginning each new test.

The purpose of this test was to show performance comparisons in typical small-to-medium business data center conditions, not to identify a maximum loading limit. 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), bandwidth, and latency or response time. These metrics are defined as follows:

- **IOPS** – I/Os per second – 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 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 times the I/O request size determines the amount of bandwidth delivered.

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, the goal with latency is to keep it as low as possible. Jetstress has an upper limit of 20 milliseconds for database reads and writes and will not validate a storage system for Microsoft Exchange if the average read or write latency of any database exceeds this.

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 Jetstress's parameters for a successful run and validation of the storage system.

Results and Analysis

Microsoft Exchange Jetstress has a strict response time limit for database reads and writes. If any database experiences average transactional latencies exceeding this limit an Exchange configuration will be failed—considered unable to support the Exchange workload being modelled. We modelled a busy, medium-sized business' Exchange environment of 500 mailboxes. The Lenovo S3200 array met the Jetstress response time requirements for all configurations: HDD-only, HDD plus an SSD read cache, and HDD with Intelligent Real Time Tiering.

The highest transactional response times recorded by our configuration of Exchange Jetstress 2013 are graphed in Figure 2. The highest average latency was 17.5 milliseconds. Unsurprisingly, this came from the HDD-only implementation and when we enhanced the array with flash latencies went down as expected. With read caching enabled the highest average latency was 11.6 milliseconds and it dropped further with SSD tiering to a high of 7.3 milliseconds. These numbers include all the additive contributors, including the storage device and SAN infrastructure plus server and application latencies. Our conclusion is that the Lenovo S3200 is a suitable Exchange Server 2013 storage device for 500 rather busy mailboxes, even without flash upgrades, though the SSD features provide a significant margin of safety.

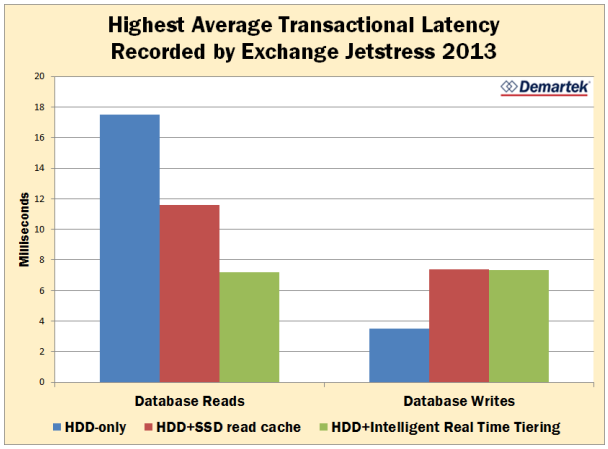

Figure 2 - Jetstress latencies

Interestingly, write I/O was faster with only HDDs in the array. This appears to be caused by the application software rather than the storage array, which will become more apparent when we examine additional performance statics.

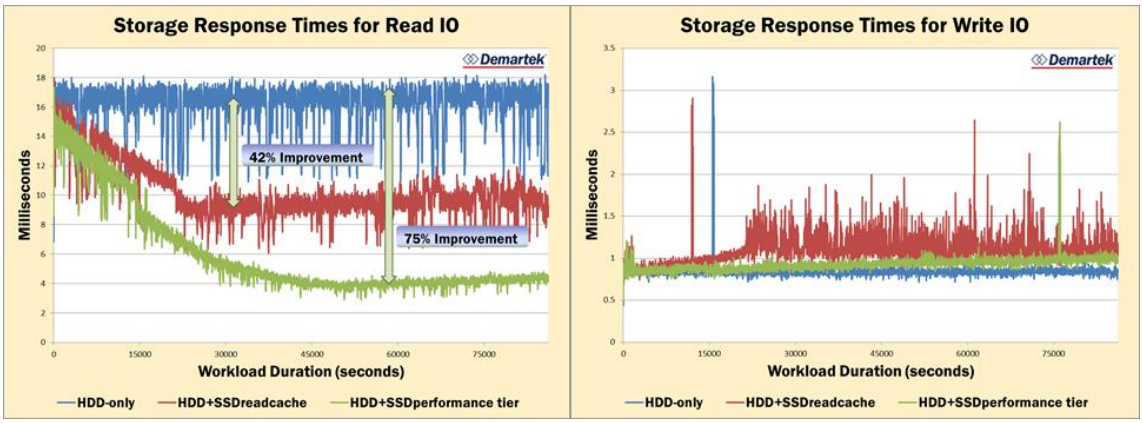

Figure 3 – Database volume response times

We also gathered performance metrics from the Windows logical volumes, excluding application induced latency. Figure 3 charts average read and write response times of I/O to these volumes for comparison with the Jetstress database latencies in Figure 2. These results pretty much mirror database transaction response times in the HDD-only setup at 17.5 milliseconds per read, but get somewhat better with flash enhanced configurations. Read caching brought response time down by 42% to about 10 milliseconds while tiering saw a 75% improvement--to 4 milliseconds. Write I/O responses were all very close to 1 millisecond regardless of the array configuration, implying the latency of the database write operations were primarily the result of how the application was processing the transactions, not because of performance limits on the array.

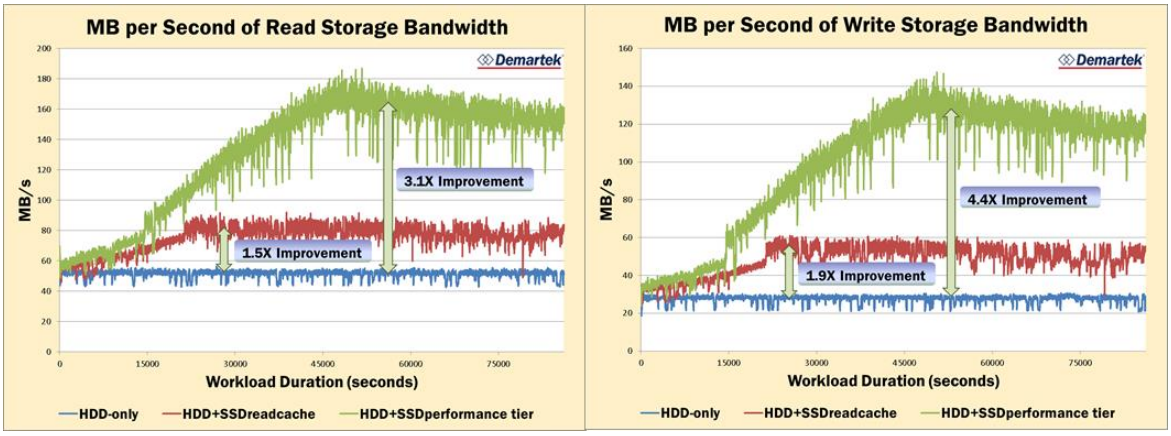

Figure 4 - Database volume bandwidths

Read and write bandwidths improved nicely from the all-HDD baseline (Figure 4). This is another benefit realized through the flash upgrades upgrade, though it has no real bearing on user experience modeled by Jetstress, which is latency sensitive rather than bandwidth dependent. What is it does demonstrate is more overall work being done by the Exchange system, above and beyond that required by the workload parameters, because of the faster storage response.

The workload is cache-unfriendly by definition, and we can clearly see that here. The abrupt halt of cache warming[2] curve and moderate performance gains over the HDD-only use case imply a strong degree of randomness in the I/O pattern, which can be hard to cache and results in a lot of the data being moved into and out of the cache. In spite of this, read caching still provided a 1.5 times read bandwidth gain over a baseline of 51 MB/s and 1.9 times the baseline write bandwidth of 28 MB/s. More cache space would likely improve performance further.

The SSD tiering algorithm and five second data migration interval seem to be quite effective in getting active data onto a flash tier for this workload. We recorded an average 3.1 times improvement in read bandwidth, up to 158MB/s and a 4.4 times improvement in write bandwidth, to 123MB/s. 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.

[2] Cache warming is an expression that refers to moving data into a new or empty cache. Performance of an application that is caching some of its data in a faster media increases steadily until the cache is full, or warm, At this point new data cannot be added to the cache unless some of the data already in cache is cleared.

Summary and Conclusion

The Lenovo S3200 array is more than capable of supporting a Microsoft Exchange server with 500 very busy mailboxes. No flash acceleration is needed to provide a reasonable level of service. However, this may not be future-proof if a business expects to grow, nor may it deliver the best return on investment of an array that is capable of much more.

A little bit of flash in the Lenovo S3200 hybrid array goes a long way. We would advise businesses not to ignore the advantages that a hybrid array with real-time tiering offers over all-HDD configurations. Deploying real-time tiering will significantly reduce latency for a 500 mailbox environment and provide considerable room to expand in the event that the Exchange server usage goes up. A business would be wise to consider where its Exchange requirements might be in a few years.

A cost-conscious business may also want to consider how the Lenovo S3200 might enable it to do more with less. The environment we deployed in this evaluation did not demand anywhere near the full drive space available from this array. To keep latency below 20 milliseconds for the HDD-only array configuration, we were essentially short-stroking the drives—placing data only on the outer edges of the drive platters to minimize movement of the drive heads. This wastes space which brings down ROI. SSD tiering delivered a latency improvement of 75%, allowing the option to use more the array's capacity without negatively impacting the user experience. This new found space can be exploited by placing more data in the form of additional mailboxes, or by putting a second application on the array, while still keeping latency below the requirements of Microsoft Exchange.

We recommend that small-to-medium sized businesses, looking to support critical applications, consider the Lenovo S3200 hybrid array as an affordable option for enterprise-class performance and features.

Appendix A – Test Description and Environment

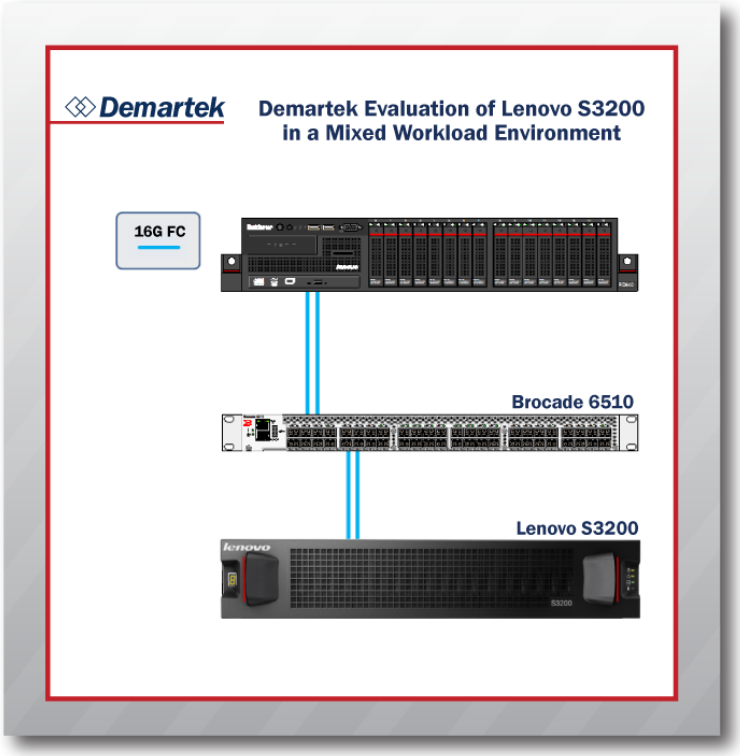

Figure 5 – Test Infrastructure

Server

- Dual processor rack server
- 2 Intel E5-2630 2.3GHz CPUs
- 16 GB RAM
- 16 Gb FC dual port HBA
- Microsoft Windows Server 2012 R2
- Microsoft Exchange Jetstress 2013, Microsoft Exchange 2013 ESE files

Fibre Channel Switch

- Brocade 6510 16Gb Fibre Channel Switch

Storage Array

- Lenovo S3200 array
- Lenovo SAN Manager Storage Operating System
- 24 900GB 10k RPM 6Gb SAS HDD
  - 10 drive RAID 6 per storage controller
  - 6 Data volumes – 3 per controller

- 6 Log volumes – 3 per controller
- 4 400GB SSD
  - 1 drive per storage controller for read caching
  - 2 drives RAID 1 per storage controller for Intelligent Real Time Tiering
- 4 16Gb FC target ports per controller (1 port active per controller)

The original version of this document is available at http://www.demartek.com/Demartek_Lenovo_S3200_Evaluation_2016-01.html on the Demartek website.