Evaluation Report: Accelerating SQL Server Database Performance with the Lenovo Storage S3200 SAN Array

Evaluation report prepared under contract with Lenovo

Executive Summary

Even with the price of flash continuing to drop, enterprise all-flash arrays still remain unaffordable for many businesses. However, many common business workloads don't actually demand the level of intense performance that all-flash storage can deliver and, in fact, may never fully capitalize on it. Shrewd storage architects will consider workload requirements along with expected returns on investment of hardware purchases. In many cases, a little flash can go a long way and a hybrid storage array may be a better choice to satisfy business requirements than a much pricier all-flash device.

The Lenovo Storage S3200 SAN array brings enterprise-class performance and features to small and medium businesses at an affordable price. A customer has the option to replace some of the array's hard disk drives (HDD) with solid state drives (SSD) to create a customized read caching or performance tiering solution tailored to a business' specific I/O requirements.

Demartek evaluated the performance of a Microsoft SQL Server database workload backed by a Lenovo Storage S3200 Fibre Channel array in an all-HDD configuration. We then repeated the workload with the SSD read caching and SSD performance tiering options installed and configured. Several performance metrics including bandwidth, IOPs, number of database transactions per second, and I/O latency were compared for a complete picture of the S3200's suitability to support a workload similar to what a small-to-medium size business would be likely to run.

We found that the array would support a transactional database workload with 30 virtual users while delivering an average of 650 database transactions per second. When we applied read caching to the array the transaction count increased by 35%, and then by 60% with SSD tiering. At the same time, average I/O response time dropped from nearly 20 milliseconds to 4 milliseconds with read caching, and to below 2 milliseconds when we employed SSD performance tiering.

Storage bandwidth and IOPS likewise saw increases in performance, each growing by more than 2.5 times when employing read caching and more than 4.5 times with SSD performance tiering.

The Lenovo Storage S3200

Lenovo's S3200 Storage Array (Figure 1) is a dual controller, 2 rack unit storage array supporting either 24 small form factor (SSF) drives or twelve 3.5 inch drives. The S3200 array supports 8 Gb and 16 Gb Fibre Channel as well as 1 Gb and 10 Gb iSCSI. Each array can support seven expansion units for a maximum of 768 TB of total storage capacity.

Figure 1 - Lenovo S3200

The base configuration of 24 HDDs is upgradable by replacing some HDDs with flash storage. The Lenovo SAN Manager controls the addition of SSDs to an array through the creation of an SSD read-cache, or an SSD performance tier where active data is migrated from HDD to SSD every five seconds through Intelligent Real-time Tiering™ software.

This evaluation was performed on a Lenovo Storage S3200 array with 20 900GB 10K RPM HDDs and four 400GB SSDs. We set up a 20 HDD baseline test case, with 10 drives provisioned to each controller in RAID 6 disk groups. To evaluate read caching, we added a single SSD to each controller and assigned them as read cache disk groups. For measuring the effect of tiering, two SSDs in a RAID 1 disk group were added to each controller. These configurations resulted in a read cache that was roughly 5% of the usable storage space and an SSD performance tier of about 10% of the drive space[1].

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

[1] The performance tier was 10% by total SSD capacity. RAID 1 makes it effectively 5% in usable space as with the read cache configuration. However, read requests do receive the advantage of two mirrors to read from instead of a single drive which was expected to provide some additional benefit over read caching for a read-heavy workload.

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 in addition to I/O resources. These results more closely resemble a real customer environment. This is unlike benchmarks that use synthetic workloads which 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 used a transactional database workload to perform real transactions similar to those that might be executed by database application users, plus 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 modelled brokerage firm in turn interacts with financial markets to execute orders on behalf of the customers and updates relevant account information. Most of this workload's transactions are read I/O, with a relatively small amount of write transactions to the database.

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 or full database system performance. We chose this workload because it is similar to standardized real-world database workloads widely recognized by the IT industry. Data and results published in this report are not comparable to, and cannot be compared with, database benchmarking 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 transactions. The database was populated with 3,500GB of data and a 200GB log.
We envisioned a medium-sized business of about 300 employees and hypothesized that perhaps ten percent of those would simultaneously access the database at any one time.

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), 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 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 times the I/O request size determines the amount of bandwidth 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 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. 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 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, such as modelled by this OLTP workload 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 bandwidth. Before flash storage became commonplace in the datacenter, storage I/O latencies of 10 to 20 milliseconds were generally acceptable for many applications. 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 Lenovo Storage S3200 array, we were particularly interested in seeing how the storage response times would react to the addition of a small amount of flash.

Our interest in this analysis is the user/application experience of a transactional database application running with Lenovo Storage S3200 array, so we chose to conduct the measurements from the application host with Windows Perfmon. This allowed us to capture the aggregate effect of all contributors to I/O response time include storage, switching, cables, operating system, and the application software stack. Values will be at least slightly better at the storage target, but that's not representative of the end user's experience. We included database transactions per second as an additional metric as it is easily recorded by SQL Server counters in Perfmon.

Results and Analysis

With only HDDs installed, the Lenovo Storage S3200 SAN array supported the 30 user OLTP workload with an average I/O request latency of about 18 milliseconds, which fell beneath a 20 millisecond ceiling we set for acceptable response times. The database maintained a consistent 650 database transactions per second. We considered that our baseline (Figure 2).

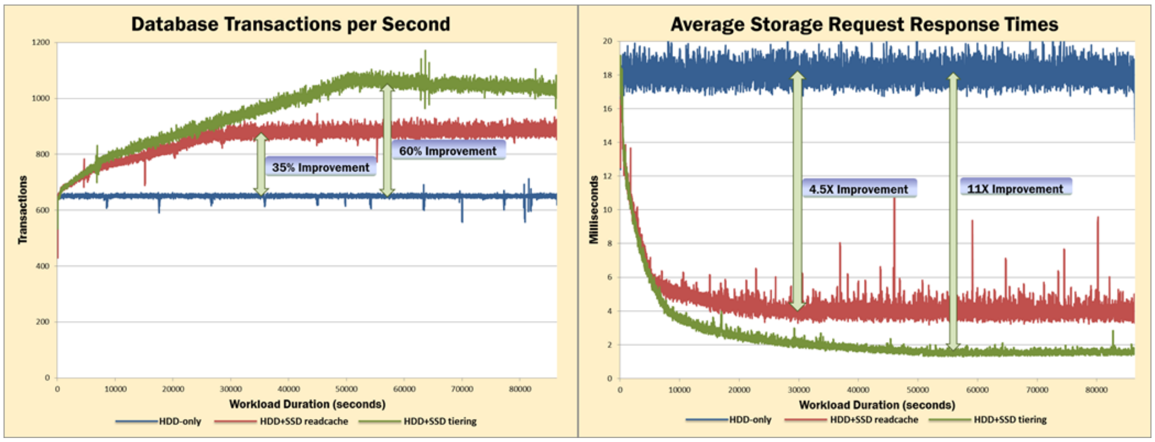

Figure 2 - Database transactions and response times

It's no surprise that adding flash into the mix improved performance significantly. Read caching resulted in 35% increase in transactions per second—it took about eight hours to fully warm the cache—while I/O response time went down by a factor of 4.5X, to 4 milliseconds. In simplified layman's terms, more work in less time.

Using the SSDs as a performance tier was even better. Tiering accelerates writes as well as reads, but this wasn't a write-intense workload. The amount of usable SSD space remained the same as with caching, but two drives in a RAID 1 configuration undoubtable benefitted read I/O by providing dual paths to the data instead of single one. The database transaction count increased by 60% over the baseline while latency dropped to below 2 milliseconds. Warmup for tiering was proportionally similar to caching at about eight hours to reach the same transaction level as the read cache, plus an additional six and a half until peak performance.

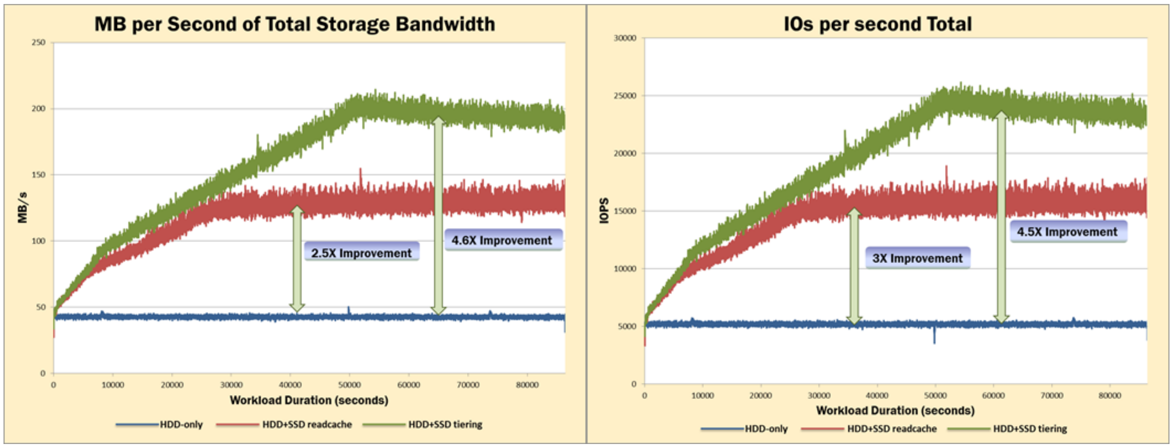

Figure 3 - Bandwidth and IOPS

With database memory limited, increased transactional performance had to be the result of improvements on the Lenovo Storage S3200 SAN array. Bandwidth increased 2.5 times and 4.6 times over the baseline through SSD read caching and tiering respectively. In terms of megabytes per second, bandwidth went up from a baseline of 42 MB/s to a steady-state average of 127 MB/s caching and 193 MB/s tiering. IOPS likewise improved from about 5,175 to 15,450 and again to 23,500 as we employed caching and tiering.

The number of virtual users executing the workload never changed, remaining steady at thirty users in all scenarios. 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

No doubt all businesses would prefer to run every workload on high-end all-flash storage, but that can be very expensive. It's also generally unnecessary for supporting application requirements of many small-to-medium businesses. The Lenovo Storage S3200 SAN array is positioned as an affordable option that will deliver a good return on investment for these companies. It's enterprise quality hardware for enterprise applications, such as transactional databases, with performance levels demanded by successful businesses.

As technological advancements drive better and better storage system and server performance, application service level expectations go up as well. A business may find the performance delivered by a basic, HDD-only Lenovo Storage S3200 array sufficient to meet operational requirements. Businesses which need more performance than delivered by spinning drives alone can look to the S3200 for the option to add SSDs in an incremental manner, customized to the workload and easier on the pocketbook than other flash solutions.

This configurable flash acceleration can pay off in work done and user experience. SSD read caching provides a significant boost to IOPS and bandwidth while reducing I/O request latency. We saw bandwidth more than double by adding about 10% of the total data capacity in SSD. Flash performance tiering requires some redundancy to protect data written to the tier, but we found that our workload benefitted even more with this option. The extra investment in flash drives (four SSDs instead of two) netted a greater increase in work accomplished with a tremendous reduction in response time. The size of the cache or performance tier is limited only by the number of SSDs deployed.

A benefit of hybrid storage solutions like the Lenovo Storage S3200 is that when an 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 Lenovo can help here too with detailed device metrics available through the user interface. A savvy administrator can tune the number of flash drives for acceptable throughput and response time. Once optimal application performance is achieved, provisioning additional flash will only provide marginal benefit to the business, so there's no reason to incur extra expense.

Ultimately, a business must understand its workloads to determine the best storage solutions for its needs. Flexibility is an asset that shouldn't be overlooked. Businesses are advised to consider the Lenovo Storage S3200 hybrid SAN array as an affordable choice for modern applications.

Appendix A – Test Description and Environment

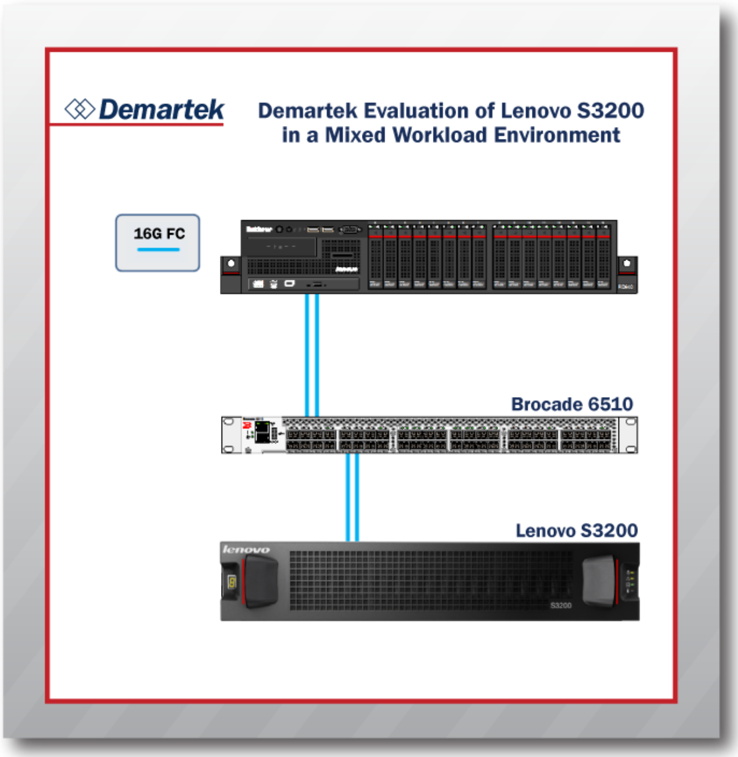

Figure 4 – Test Infrastructure

Server

- Dual processor rack server
- 2 Intel E5-2630 2.3GHz CPUs
- 16 GB RAM
- 16Gb FC dual port HBA
- Microsoft Windows Server 2012 R2
- Microsoft SQL Server 2012, Microsoft Benchcraft SQL Server testing kit

Fibre Channel Switch

- Brocade 6510 16Gb Fibre Channel Switch

Storage Array

- Lenovo Storage S3200 array
- Lenovo SAN Manager
- 20 900GB 10k RPM 6Gb SAS HDD
  - 10 drive RAID 6 per storage controller
  - 6 Data volumes – 3 per controller

- 1 Log volume
- 4 400GB SSD
  - 1 drive per storage controller for read caching
  - 2 drives RAID 1 per storage controller for tiering
- 4 16Gb FC target ports per controller (2 ports 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.