Evaluation Report: Database Acceleration with HP 3PAR StoreServ 7450 All-flash Storage

Evaluation report prepared under contract with HP

Executive Summary

Solid state storage is transforming the entire computing industry and becoming the de facto standard for high-performance storage systems. Increasing demands of database workloads, growing numbers of virtual machines and more powerful processors are driving demand for ever-faster storage systems, especially those that can deliver that performance with significantly lower latency than in the past.

The HP 3PAR StoreServ 7450 All-flash storage system provides the high performance and low latency that meet today's datacenter needs.

HP commissioned Demartek to run database workloads that show the performance of the HP 3PAR StoreServ 7450 All-flash Storage system. Not only did the 3PAR StoreServ 7450 show excellent low latency and fast performance for a typical OLTP database workload, Demartek also found that the 3PAR StoreServ 7450 was not very heavily taxed with a single OLTP database accessing the array. As a result, we proceeded to run a combination of database workloads including two online transaction processing (OLTP) workloads and a data warehouse workload to see how well this storage system would handle a fairly heavy, mixed workload. We deployed a mixture of 16Gb and 8Gb Fibre Channel infrastructure in this series of tests.

Key Findings

For a combination of OLTP and data warehousing workloads from three different host servers, we found that the HP 3PAR StoreServ 7450 All-flash storage system easily handled a peak of 300,000 IOPs, sustained nearly 100,000 IOPS, and achieved 4000 MBPS while maintaining low average latency while all three workloads were running. These combined workloads, as viewed by the storage system, were 93% read and 7% write. The average read block size was 36.4 KB and the average write block size was 8.8 KB.

In our opinion, the HP 3PAR StoreServ 7450 All-flash storage system works well for database workloads because it provides low latency in addition to higher raw performance.

The Need for All-flash Storage

Today's database applications face a variety of challenges brought on by seemingly insatiable demands on server and storage infrastructure. At the same time, new technologies are being introduced that offer possible solutions to meet these growing challenges.

A number of factors contribute to the need for all-flash storage, including higher performance and lower latency demands from database applications, increasing numbers of virtualized servers and the availability of high-speed networks such as 8Gb and 16Gb Fibre Channel.

All-flash storage systems provide not only higher raw performance in terms of IOPS and bandwidth, but have the advantage of operating with lower latency. In many cases, especially OLTP environments, low latency is becoming more important than the IOPS and bandwidth performance.

HP 3PAR StoreServ 7450 All-flash Storage Architecture

The HP 3PAR StoreServ 7450 All-flash storage system provides high-speed performance with low latency for mission-critical applications such as database applications. It provides high-performance and the same enterprise-grade levels of resiliency (up to 4 nodes), rich Tier 1 data services, efficiency, data mobility and disaster tolerance that are available on other HP 3PAR StoreServ platforms. It does this with a flash-optimized architecture that reduces the performance bottlenecks that can choke general-purpose disk arrays that have been retrofitted by stuffing them with solid-state disks (SSDs), and includes enterprise-grade features that are not always available with some of the newer all-flash offerings from others.

HP 3PAR StoreServ 7450 storage system has a number of design optimizations to leverage performance from flash media while keeping the costs down. Some of them are described below briefly:

Adaptive Read and Write

Adaptive Read and Write is a software innovation that enables a more granular approach than with spinning media by matching host I/O size reads and writes to flash media to avoid unnecessary data reads and writes. This significantly reduces latency and optimizes back end performance to enable more applications to be consolidated.

Autonomic Cache Offload

Autonomic Cache Offload is another new flash software-based optimization that eliminates cache bottlenecks by automatically changing the frequency at which data is offloaded from cache to flash media, based on utilization rate and without requiring any user intervention. This ensures consistently high performance levels as you scale workloads to hundreds of thousands of IOPS.

Another important aspect of the cache offload algorithm is the decision around which cache data should be flushed to the back end, and which should not be. HP 3PAR StoreServ Storage keeps track of read cache hits and keeps hot data in cache itself, thereby lowering latencies of frequently accessed data. In addition, for handling flash, flusher threads have been added to 3PAR cache management so it can perform more operations in parallel.

Multi-tenant I/O Processing

New multi-tenant I/O processing innovations enable performance optimization for mixed workloads and VDI deployments by breaking large I/O into 32 KB sub-I/O blocks. This prevents small read I/O chunks from getting held up behind large I/O requests, therefore assuring reduced latency. Breaking large sequential read I/O into sub-I/O chunks distributes these reads and lowers the possibility of smaller transactional read I/O operations of get held up because of a previous large read I/O operation, therefore ensuring consistently low latency for transactional I/O, even in mixed workload scenarios.

With the multiple workloads (OLTP and Decision Support) running on 7450 in this evaluation, enhancements like multi-tenant I/O processing help 7450 to deliver consistently good latency for the OLTP application while deliver good bandwidth for the Decision Support application.

Specialized ASIC for Mixed Workload Support and Zero Detection

In the HP 3PAR StoreServ Architecture, transactional and throughput-intensive workloads are able to run on the same storage resources without contention because of the HP 3PAR StoreServ Gen4 ASIC, which offloads some of the work of the controllers, allowing them to process smaller I/O without inordinate delay due to large block I/O. This is important, because precious CPU cycles are not wasted for data movement; instead the CPU cycles are available for delivering advanced Tier 1 data services.

The HP 3PAR StoreServ ASIC features an efficient, silicon-based zero detection mechanism. This hardware feature removes allocated but unused space without impacting performance. Every block of flash storage that gets reclaimed immediately becomes available for other applications requiring space. This ensures that existing flash storage is utilized efficiently, possibly delaying the purchase of additional capacity.

In addition, with zero detection built into the ASIC, a stream of zeroes that may be present in a write I/O can be eliminated before being written to flash. In the world of flash, where there is a penalty for every single write (in terms of media endurance), this write elimination helps elongate the life of flash-based media.

System-wide Striping

Data and I/O for each volume are striped widely across all system resources, including CPUs, ports, cache, and drives. This wide striping enables the system to deliver

accelerated performance levels (with all resources supporting each volume) while avoiding any single point of contention. Even a small volume can leverage the performance of hundreds of flash devices and all the system's controller nodes for optimal performance.

Because the system autonomically manages this system-wide load balancing, no extra time or complexity is required to create or maintain an optimally configured system. With system-wide striping, data is distributed across all SSDs in a granular fashion, keeping the command queues low on any individual SSD.

System-wide Sparing

The HP 3PAR StoreServ Architecture reserves spare chunklets in all flash media. In contrast to traditional architectures that enforce the need to reserve dedicated spares that sit idle, the HP 3PAR StoreServ Architecture uses every single flash device, reserving spare chunklets in each of them. This ensures a balanced load and wearing across all flash media. Should there be a media failure, system-wide sparing also helps protect against performance degradation by enabling many-to-many rebuild, resulting in faster rebuilds.

Moreover, with Adaptive Sparing, 3PAR StoreServ is able to release the spare chunklets present in each SSD media to the SSD itself, allowing SSDs to use that spare capacity as additional over-provisioned capacity. This increase in over-provisioned capacity enhances endurance of SSDs within a 3PAR StoreServ array

Media Wear Gauge

The HP 3PAR Operating System monitors the wear of each flash device and allows users to see the wear level of each device at any given time. This keeps users informed as to the amount of media wear taking place and helps them replace SSDs in a planned fashion.

Data Warehouse Workload Descriptions

Database Workloads

In order to show the performance of the storage system under a mixed workload, Demartek ran a combination of online transaction processing and data warehouse workloads with the configuration shown below. For these tests, we ran the TPC Benchmark™ E (TPC-E) and TPC Benchmark™ H (TPC-H), which were used to generate real-world workloads for measuring the performance of the storage system, but not for the purpose of publishing official benchmark results.

TPC-E is a variation of an On-Line Transaction Processing (OLTP) workload that 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 and some writes to its database.

TPC-H is an industry-standard, decision-support benchmark that simulates broad business intelligence database environments most relevant to information systems that provide organizations with answers to critical business analytics. This workload consists of almost all reads.

Real vs. Synthetic Workloads

The workloads employed in this test used a real database (Microsoft SQL Server) with database tables, indexes, etc., performing actual database transactions. When using real database workloads, 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.

These are 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.

Performance Results – Host Perspective

When measuring performance of storage systems, three key metrics are I/Os per second (IOPS), bandwidth and latency. Online transaction applications tend to generate higher numbers of IOPS with lower bandwidth. Data warehousing workloads tend to generate higher bandwidth but lower IOPS. Latency is important for both types of workloads.

Definitions

- **IOPS** – I/O's 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).

Workload Combinations

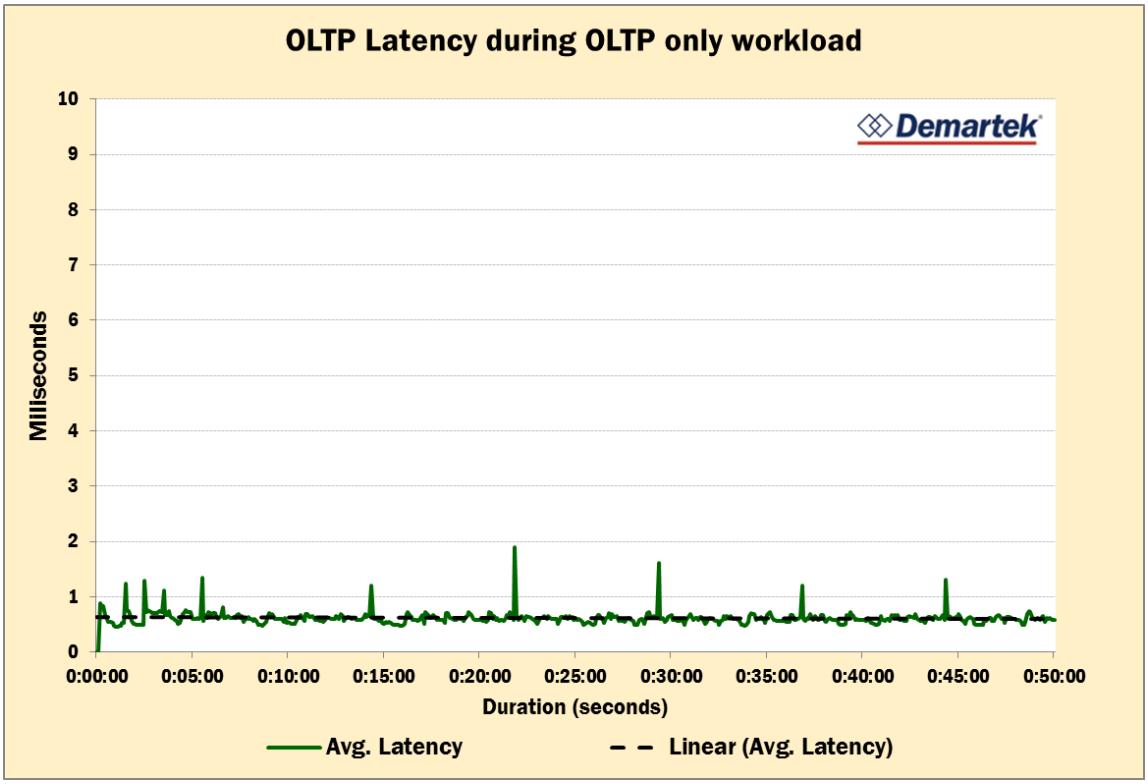
The data shown below is divided into two groups:

- Two OLTP workloads
- Two OLTP workloads plus one data warehousing workload

For all the charts shown below, the performance data is for the database volumes only, not the log volumes.

Two OLTP Workloads

OLTP workloads are transaction-oriented workloads, and one of the key metrics for these workloads is latency. The overall average latency for the two OLTP workloads shown below is 0.61 milliseconds. This measurement was taken without any data warehousing workloads running and illustrated good, consistent low latency for OLTP only workloads..



Two OLTP Workloads and One Data Warehousing Workload

Performance data was collected by the host application servers and by the storage system. In this section, we provide both perspectives of the data side-by-side. Although similar, the data shown is different, primarily because the storage system averages the data over five-minute intervals. This five-minute interval data collection has the effect of smoothing out spikes in the graph.

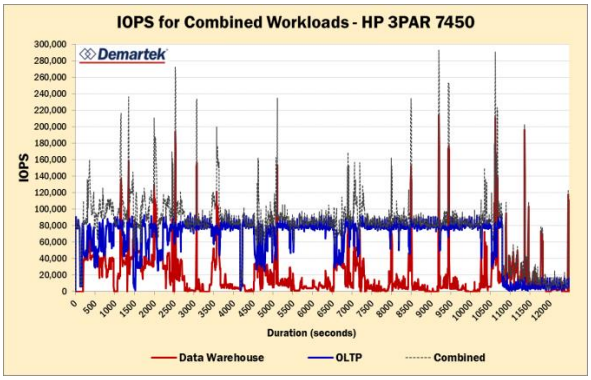
The OLTP workloads began to finish and decrease their load on the storage system at approximately 3 hours (10800 seconds).

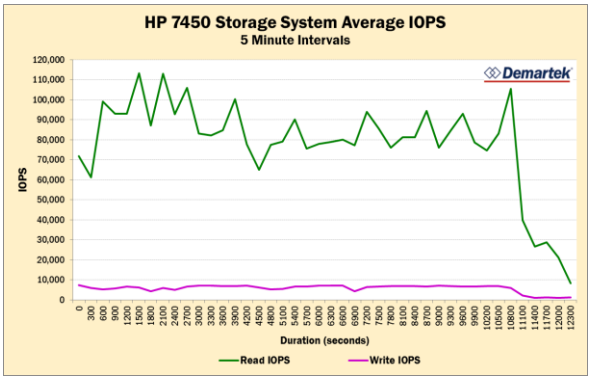
Server-side Observations

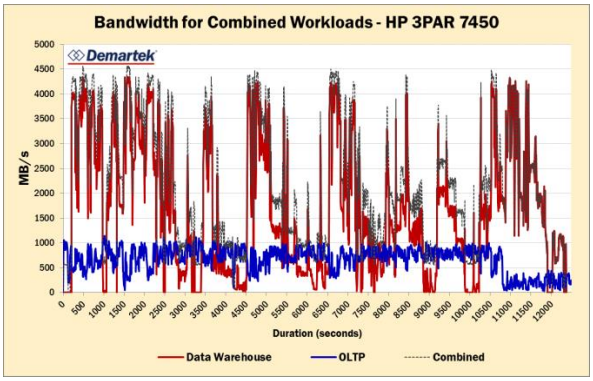
As expected, the OLTP workloads issued a higher number of IOPS than the data warehousing workload, and the data warehousing workload achieved higher bandwidth (throughput) than the OLTP workloads.

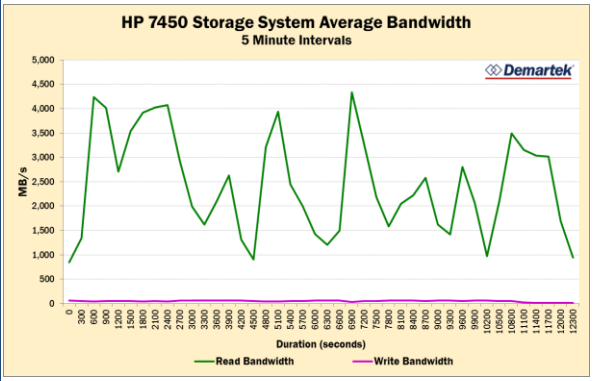
Storage-side Observations

On average, the reads were 93% of the combined IOPS, and were 98% of the combined bandwidth.









Data warehousing workloads, because of their large queries and table scans, tend to have much higher latencies that tend to be "bursty". OLTP workloads tend to have shorter latencies because of the transactional nature of their operations.

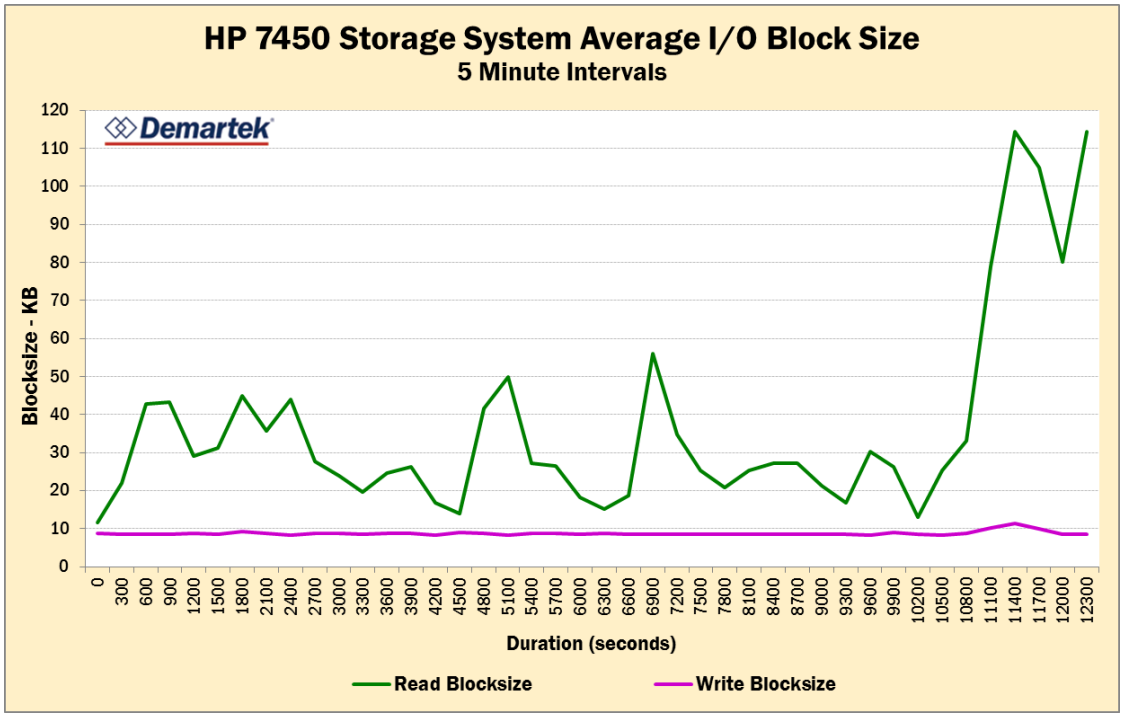
Latency is supremely important parameter for OLTP applications and often requires an entire storage system to be dedicated to run OLTP applications. The data warehousing applications are much more bandwidth hungry.

Server-side Observations	**Storage-side Observations**
As expected, data warehousing workload had the effect of increasing the overall latency for the combined workloads.	The latency for writes was fairly consistent.

The overall average latency for the combined OLTP and data warehousing workloads was 1.55 milliseconds.

HP 3PAR StoreServ 7450 is able to handle a multi-tenant environment, where response time of the OLTP application remains good even with a bursty data warehousing application running on the same system.

One final measure provided by the storage system is the average I/O block sized for reads and writes. As the OLTP workloads were winding down at about 3 hours into the run, the data warehousing workload increased its activity, noted by the larger block sizes.



Gen 5 (16Gb) Fibre Channel

Generation 5 (16 Gbps) Fibre Channel provides an increased ceiling in terms of the possible performance for each connection, doubling the throughput of the previous generation (8 Gbps). Database administrators often tune their applications to take advantage of every possible speed benefit, and Gen 5 Fibre Channel connections provide an important performance improvement opportunity. Other benefits include a reduced number of links needed to achieve the same bandwidth, reduced power consumption needed to achieve the same bandwidth and fewer cables to manage. In addition, the higher speed allows fabrics to be connected with fewer inter-switch links (ISLs), which is especially helpful in large fabrics.

Gen 5 Fibre Channel includes retimers in the optical modules and transmitter training, features that improve link performance characteristics, electronic dispersion compensation and backplane links.

Because of Fibre Channel's backward compatibility, host servers can deploy 16Gb FC HBAs and 16Gb FC switches independently, as each work with the two previous generations (8Gb and 4Gb).

Table 1 – Fibre Channel Speed Characteristics

Speed	Throughput (Mbps)	Line Rate (Gbps)	Encoding	Retimers in the module	Transmitter training
1 GFC	100	1.0625	8b/10b	No	No
2 GFC	200	2.125	8b/10b	No	No
4 GFC	400	4.25	8b/10b	No	No
8 GFC	800	8.5	8b/10b	No	No
16 GFC	1600	14.025	64b/66b	Yes	Yes

Table 2 – Fiber Optic Cable Link Distance

Speed	Multi-Mode				Single-Mode
	OM1	OM2	OM3	OM4	OS1
1 GFC	300	500	860	*	10,000
2 GFC	150	300	500	*	10,000
4 GFC	50	150	380	400	10,000
8 GFC	21	50	150	190	10,000
16 GFC	15	35	100	125	10,000

* The link distance for OM4 fiber optic cable has not been defined for these speeds.

Summary and Conclusion

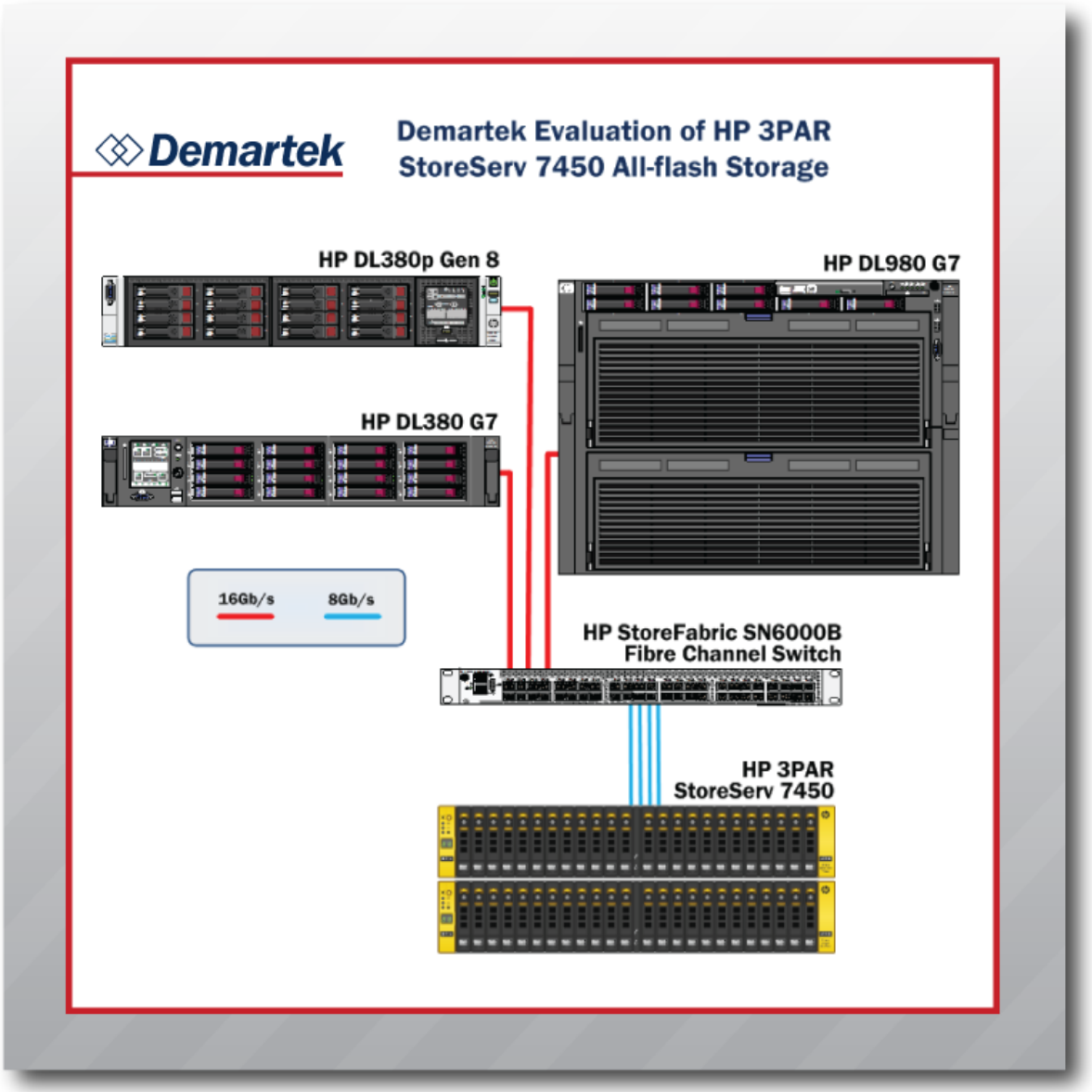
We believe that the HP 3PAR StoreServ 7450 All-flash storage system provides excellent performance for mission critical workloads. This storage system is well-suited to handle mixed workloads thereby providing a fast ROI to customers investing in an all-flash storage system.

In our opinion, all-flash arrays do well with database workloads because they provide low latency in addition to higher raw performance.

Appendix – Test Description and Environment

Database workloads were run on the three servers combined to provide an intensive mixture of random and sequential I/Os on the storage system, stressing the storage system. This set of tests was designed to be representative of a typical mission-critical environment with mixed database workloads.

These tests were run with the Gen 5 (16Gb) infrastructure running at 16 Gb/sec including the Fibre Channel HBAs in the servers and the 16Gb Fibre Channel switch. The HP 3PAR StoreServ 7450 storage system used 4x 8Gb Fibre Channel host ports.



Server #1

- HP ProLiant DL380 G7
- 2x Intel® Xeon® X5690, 3.47GHz, 12 total cores, 24 logical processors
- 28GB RAM
- HP SN1000E 16Gb FC HBA (Emulex LPe16002)
- Microsoft Windows Server 2012
- Microsoft SQL Server 2012, Microsoft TPC-E benchmark kit

Server #2

- HP ProLiant DL380 Gen8
- 2x Intel® Xeon® E5-2609, 2.4GHz, 8 total cores, 24 logical processors
- 64GB RAM
- HP SN1000E 16Gb FC HBA (Emulex LPe16002)
- Microsoft Windows Server 2012
- Microsoft SQL Server 2012, Microsoft TPC-E benchmark kit

Server #3

- HP ProLiant DL980 G7
- 8x Intel® Xeon® E7-2870, 2.4GHz, 80 total cores, 160 logical processors
- 320GB RAM
- HP SN1000E 16Gb FC HBA (Emulex LPe16002)
- Microsoft Windows Server 2012
- Microsoft SQL Server 2012, HammerDB 2.14 and TPC-H benchmark kit

Fibre Channel Switch

- HP StoreFabric SN6000B Fibre Channel Switch

Storage Array

- HP 3PAR StoreServ 7450 Storage
- 3PAR Operating System 312 MU2
- 48x 400GB 6Gb SAS SSD
  - RAID5 configuration for database volumes
  - RAID10 configuration for log volumes
- 4x 8GFC host ports

Note: HP 3PAR StoreServ 7000 and StoreServ 10000 Storage arrays currently use Emulex quad-port 8GFC HBAs, configured in target mode.

The original version of this document is available at:
http://www.demartek.com/Demartek_HP_3PAR_StoreServ_FlashStorage_2014-05.html on the Demartek website.