Evaluation Report: All-Flash HP 3PAR StoreServ 7450 Storage System and Generation 5 (Gen 5) 16 Gb/s Fibre Channel

Evaluation report prepared under contract with HP

Executive Summary

Solid State storage is becoming increasingly common in the computing industry with hybrid and all-flash array deployments taking the place of traditional HDD storage for high performance workloads. The marriage of Gen 5 16 Gb/s Fibre Channel infrastructure (switches, host HBAs) with all-flash arrays—also equipped with Gen 5 16 Gb/s FC targets—is yet another enhancement for high performance flash storage.

HP 3PAR StoreServ 7450 All-Flash Storage system delivers high performance at low latency to meet today's datacenter requirements. To augment this HP recently made Gen 5 16 Gb/s FC targets available on new 3PAR StoreServ storage systems and as an upgrade to existing systems.

HP commissioned Demartek to evaluate the performance of the All-Flash HP 3PAR StoreServ 7450 Storage system in an end-to-end Gen 5 16 Gb/s FC SAN and compare it to an end-to-end 8 Gb/s FC environment. Demartek ran an industry-standard, online transaction processing (OLTP) workload in an 8 Gb/s end-to-end FC configuration and then simply replaced the 8 Gb/s FC infrastructure with Gen 5 16 Gb/s FC, measuring the resulting difference in performance.

Key Findings

A common OLTP workload can saturate 8 Gb/s host FC bandwidth without fully utilizing all of the compute or storage resources. By replacing the 8 Gb/s FC components with 16 Gb/s FC on the HP 3PAR StoreServ array, FC switch, and on the host HBAs, we saw storage I/O bandwidth and IOPs increase by 35% without making any application-level changes to an already fairly well performing workload (Figure 1). Latency improved by a factor of 2.5X or better, remaining significantly and consistently under a millisecond. (These findings are presented in greater detail later in this document.)

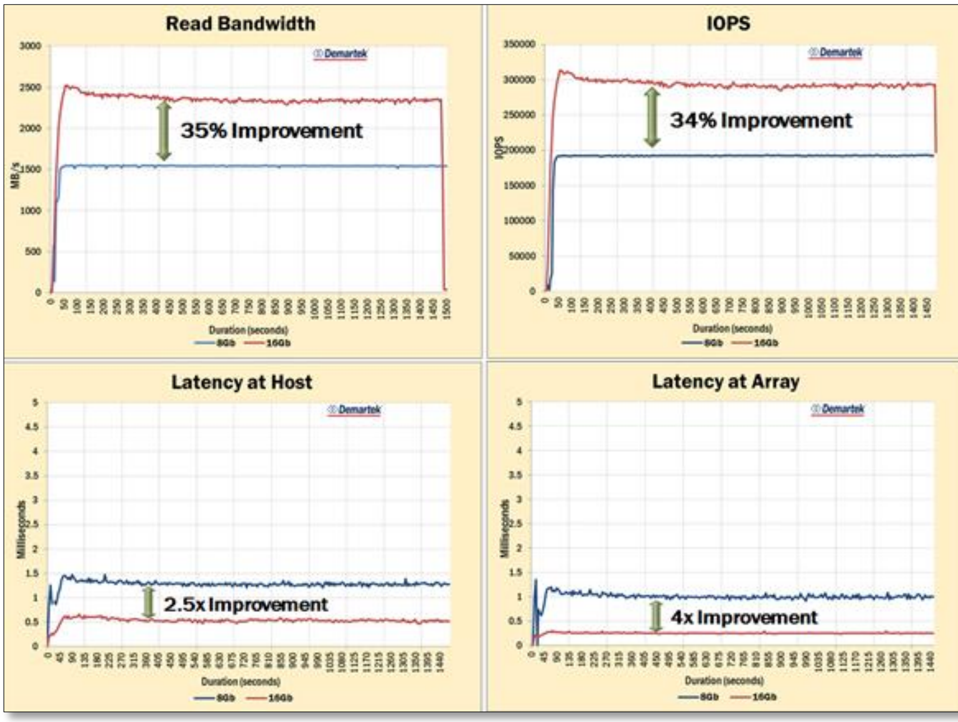

Figure 1 – 16 Gb/s FC observed improvements

We noted that if we chose to do so, workload intensity could have been increased considerably without saturating the Gen 5 16 Gb/s FC SAN.

In our opinion, upgrading Fibre Channel components to Gen 5 16 Gb/s FC has immediate payoff in environments where the rather significant investment in high-performance storage has already been made or is planned.

Gen 5 (16 Gb/s) Fibre Channel Targets and All-Flash arrays – An Obvious Choice

Current Gen 5 16 Gb/s FC HBAs not only provide double the available bandwidth and IOPS compared to 8 Gb/s FC HBAs, but have advancements in other areas. The latency of a 16 Gb/s FC HBA is generally less than half of the latency of an 8 Gb/s FC HBA because of the newer ASICs, firmware, and chipsets on the adapter. If storage targets also support 16 Gb/s FC, latency in the low hundreds of microseconds is not out of the question. In addition, HP Gen 5 16 Gb/s FC HBAs have new Emulex ExpressLane™ quality of service (QoS) features, T10 PI integrity offload, and support Brocade ClearLink physical layer diagnostics. Gen 5 16 Gb/s FC switches and HBAs have been around since 2011, but 16 Gb/s FC storage targets have only recently become an option in enterprise arrays.

The price of flash storage has dropped considerably (HP delivers all-flash storage at a net price of about $2 per gigabyte usable at the time of writing), and when deployed properly, it can be well worth the investment. The impressive bandwidth and IOPs of all-flash arrays, plus the very low latencies of the media, have been well established for years and tend to be the reasons customers make investments in flash storage. The HP 3PAR StoreServ 7450 Storage system is advertised as capable of several hundred thousands of IOPs, but more importantly, at sub-millisecond latency. Our testing data, shown later in this paper, confirms this latency claim in a real-world workload along with a bandwidth of multiple gigabytes per second.

With this type of raw performance available out of the box, adding Gen 5 16 Gb/s FC storage targets to all-flash arrays has seemed like a logical progression for some time and the industry is catching up. Bandwidth and IOPs comparable or better to that delivered by 8 Gb/s FC are obtainable across fewer FC ports in servers, in the switch, and at the array. In instances where I/O requests saturate or nearly saturate the available bandwidth, this upgrade can bring an immediate reduction in latency while at the same time increasing the amount of work done by servers and storage across the SAN.

In essence, Gen 5 16 Gb/s FC allows the infrastructure to exchange lots of small pipes for fewer, larger ones. Management is simplified while QoS and Return on Investment (ROI) generally go up for the entire infrastructure.

Transactional Database Workload Description

The TPC-E Database Workload

In order to show the performance of the storage system in both 8 Gb/s FC and Gen 5 16 Gb/s FC end-to-end configurations, Demartek ran the TPC Benchmark™ E (TPC-E) to generate a real-world workload for measuring the performance of the storage system. It was not the intention of this exercise to produce TPC-E scores for publication; in fact, the database was intentionally memory-constrained to drive storage I/O rather than maximizing database transactions.

TPC-E is a variation of an 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 with some writes to its database.

Real vs. Synthetic Workloads

The workload 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 rates 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.

Workload Definitions

Our goal was to find a workload intensity that would just saturate a dual port 8 Gb/s FC host HBA without driving latency up significantly. A workload definition running with 100 virtual users was sufficient to take storage bandwidth to 1600MB/s (the combined bandwidth of both ports of the 8 Gb/s HBA), while keeping the latency at the higher end of acceptability for the application. The 3PAR StoreServ 7450 was initially configured with 8 ports of 8 Gb/s FC to keep the performance bottleneck at the server HBA and not at the storage system. All 8 Gb/s FC components would then be swapped, port for port, with Gen 5 16 Gb/s FC.

This workload was executed on both storage and SAN configurations. We chose not to tune the workload for performance at any point in the testing so as to highlight the immediate benefits of upgrading to 16 Gb/s end-to-end FC within an existing application environment.

Performance Metrics

When measuring performance of storage systems, three key metrics are I/Os per second (IOPS), bandwidth, and latency:

- **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.

Taking metrics at the host provides a picture of the full 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.

Performance metrics can also be gathered from the storage system, which allows us to exclude factors outside the array itself. This can be particularly useful when trying to troubleshoot performance issues because it can help identify where bottlenecks are occurring in the infrastructure. Bandwidth and IOPS, by their very nature, don't tend to vary much between the host and the storage system, unless technologies such as host-side caching are employed. I/O latency can and often does differ between hosts and storage systems. HP 3PAR System Reporter software provides pre-packaged and custom reporting capabilities that drill down to just about every subsystem within the 3PAR StoreServ 7450 array to pinpoint very specific performance metrics.

A Brief Commentary on Latency

Before flash storage became commonplace in the datacenter, storage I/O latencies of 10 to 20 milliseconds were generally acceptable for many applications. In fact, 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. Flash storage has been a game-changer in this area with sub-millisecond latency now the expectation for all-flash arrays. As with all technology advances, applications and user expectations have changed in response to this capability.

The impact of higher latencies depends greatly on the workload. High bandwidth streaming or very sequential workloads might be more or less unaffected, especially where read-ahead buffering grabs more data than I/O requests actually demand. Data warehousing and video streaming are two examples of these types of workloads. However, if latencies become too high, even these jobs begin to experience noticeable lags. For optimal user experiences, lower latency is always better.

Online transactional workloads can generate high numbers of IOPS and consume a respectable amount of bandwidth. Latency becomes especially important, particularly in very highly transactional workloads, when database requests are time sensitive and have a great deal of dependency on prior transaction results.

Consider applications that perform real-time trend analysis and/or process vast amounts of data. Stock trading, such as modelled by the TPC-E workload, weather forecasting, geological survey modelling, and biometric analysis are examples of workloads that can be extremely sensitive to latency. As storage systems have improved, reducing I/O latency in the process, the performance expectations for these types of applications have likewise adapted to expect very fast response times.

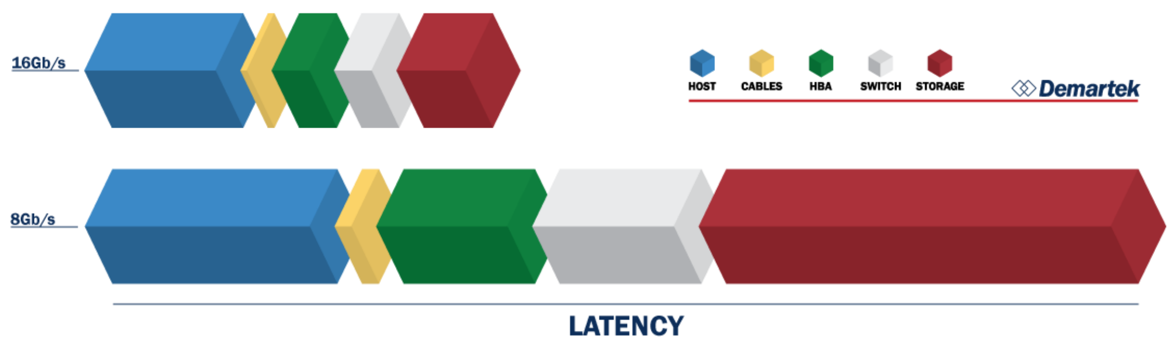

Figure 2 - Contributors to overall latency

Latency in a SAN comes from several sources (Figure 2). The primary source is the storage system itself. Total latency will never be lower than the storage system's best case scenario. However, varying amounts of latency are also introduced at the host HBA, and this can grow rapidly when IOPS or bandwidth limitations are met. As soon as host HBAs reach a saturation point, I/Os begins queuing up on the host, resulting in increased latency. Each SAN switch between the host and storage contributes to latency, typically in the low tens of microseconds. Fiber optic cables will add a few microseconds per kilometer length as well.

The additive effect of these latency sources has an impact on application performance and user experience to varying degrees, depending on the application requirements. If the aggregate exceeds agreed upon QoS standards or service levels demanded by the business, a savvy storage administrator will consider all of these potential points of latency insertion. The ability to measure performance at the host, the storage system, and in between (in switches and other storage networking elements) is essential to pinpoint where action needs to be taken.

Results and Analysis of Upgrading from 8 Gb/s FC to Gen 5 16 Gb/s FC

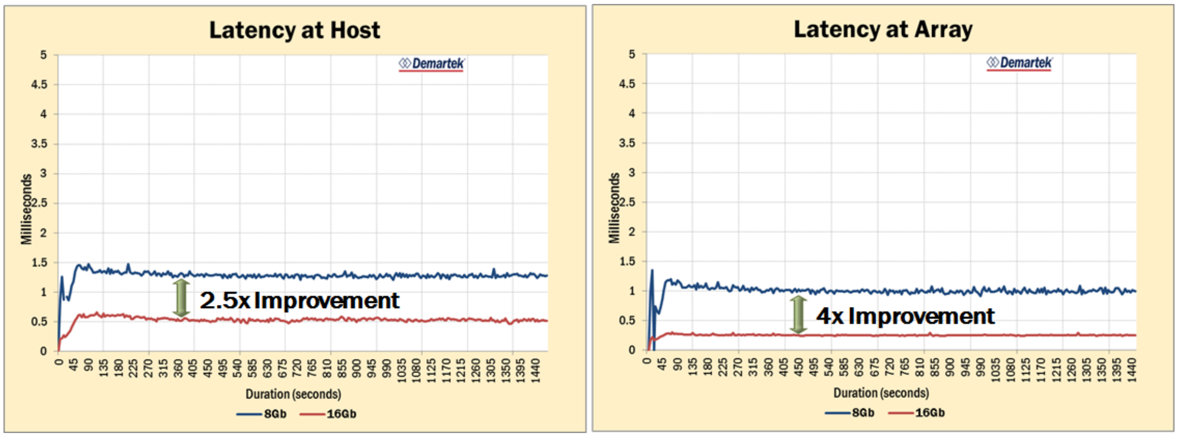

Figure 3 - Latency Results

Impressive bandwidth and IOPs numbers are easily attained on flash storage, so we'll first review the more interesting latency results (Figure 3) that can have a significant impact on transactional workload performance. As the data demonstrates, the upgrade to Gen 5 16 Gb/s FC storage targets and Gen 5 16 Gb/s FC infrastructure dropped latency at the host from over a millisecond to below half a millisecond—a 2.5X improvement.

Given this impressive decrease in latency accompanying a significant increase in the amount of work, as we'll see when we look at the bandwidth and IOPs results, it is apparent that the 8 Gb/s FC technology was impeding optimal performance of the OLTP workload to the point of introducing latency into the system under test. Examining the array-side data shows an even greater improvement. With the upstream factors such the contribution from host I/O queuing excluded from the measurement, latency is reduced by a factor of 4. Considering that the array was never an I/O bottleneck in this testing, due to having 4X the number of ports as the host, this degree of improvement highlights the technology advancements in the Gen 5 16 Gb/s FC targets compared to 8 Gb/s FC.

Also, there are fewer FC ports on the host than the storage system, and this contributed to the saturation of the host bandwidth. Bandwidth saturation has the additional effect of driving up I/O queue depth to levels that begin to affect service times at the host, and we see its impact in the data. Even so, 8 Gb/s FC appears to be a limiting factor on array performance as well. We see significant latency reduction at both ends of the SAN by upgrading the end-to-end FC infrastructure. This supports the argument for a full end-to-end FC infrastructure upgrade where possible. Focusing solely on storage or server may simply move bottlenecks to one side of the infrastructure or the other rather than eliminate them (depending on the aggregate number of ports and clients sending and receiving data from the 3PAR StoreServ 7540).

Figure 4 - Bandwidth and IOPs Results

Bandwidth and IOPs (Figure 4) provide a picture of the amount of work the storage system was doing, and where the FC infrastructure was a performance-limiting or performance-enabling factor. We focused on host-side measurements only because the storage system cannot deliver more data than the server and its applications are able to request.

A baseline measurement with end-to-end 8 Gb/s FC confirmed that the workload indeed saturated two 8 Gb/s FC ports on the host (8 target ports were available on the storage system to prevent the 3PAR StoreServ array from becoming the bottleneck). The curve flattened very quickly at 1600 MB/s, clearly demonstrating that bandwidth is limited by the 8 Gb/s FC ports installed in the host.

With no modification or tuning to the workload, the 8 Gb/s FC infrastructure was replaced with Gen 5 16 Gb/s FC storage targets, HP SN1100E Gen 5 16 Gb/s FC HBAs, and a Brocade 6510 Gen 5 16 Gb/s FC switch. Total bandwidth immediately improved by 35%. IOPS showed similar improvement, as expected.

Considerations on a Full Versus Partial SAN Upgrade from 8 Gb/s FC to 16 Gb/s FC

The purpose of this testing was to demonstrate the performance deltas of an application running on a 16 Gb/s FC SAN compared to an 8 Gb/s FC SAN, specifically with the All-

Flash HP 3PAR StoreServ 7450 Storage array as a storage target. The data presented in this paper was collected before and after making just such an upgrade. For the greatest performance gains across the board, we recommend upgrading older SANs to Gen 5 16 Gb/s FC, particularly where a significant investment in flash storage has been made. This is the best way to fully realize the return on investment made in high performance storage and servers. However, a full SAN upgrade may not be an option in some environments for different reasons, frequently related to expense or logistics. If a piecemeal upgrade is the only option, there are several considerations to keep in mind. Workload profiles and performance objectives play into possible strategies.

Without first upgrading switching to Gen 5 16 Gb/s FC, any upgrades to host and array initiators and targets will have limited benefits, since no port will be able to deliver 16 Gb/s through an 8 Gb/s switch. Fibre Channel is backward compatible to the prior two generations, meaning new switches can be added to existing SANs, and ports can be gradually migrated to them. Depending on MPIO configuration, some small performance gains may be seen from upgrading switches only. Without addressing switching in the SAN, the only benefits from host and array upgrades will be some latency and management improvements inherent in Gen 5 16 Gb/s technology.

The testing data showed that latency improvements seen on the array after upgrading to Gen 5 FC were not fully realized at the server. This suggests that upgrading server HBAs over storage targets will provide a more immediate gain in performance. This will also increase the bandwidth and IOPS potential from the host (and reduce latency from I/O queuing). A similar effect could be achieved by increasing 8 Gb/s ports, but this would defeat any goals around reducing total port counts. Host HBA upgrades do require servers to be offline.

The HP 3PAR StoreServ 7450 was not a bottleneck to begin with, but we did see that adding Gen 5 16 Gb/s FC targets resulted in a significant reduction in array latency. The array FC targets are swappable with the array live. In our testing, we configured the HP 3PAR array with 8 ports of FC, two ports per array controller. However, the HP 3PAR StoreServ 7450 supports up to 6 ports of 8 Gb/s FC targets per controller. This can be reduced to 2 ports per controller for Gen 5 16 Gb/s FC, which is a significant port reduction over a fully configured array.

Summary and Conclusion

The All-Flash HP 3PAR StoreServ 7450 Storage system provides excellent performance for OLTP workloads and has done so for a couple of years now. HP's recent upgrade to Gen 5 16 Gb/s FC targets makes it even better by uncovering the true performance potential.

Applications demanding high bandwidth and IOPs will find Gen 5 16 Gb/s FC able to meet or exceed those requirements with fewer ports. Applications with extremely time-sensitive I/O, such as the brokerage simulation used in this paper or other highly computational workloads, will benefit from the amazingly low latencies the HP 3PAR StoreServ 7450 can deliver across Gen 5 16 Gb/s FC.

Customers who already own an All-Flash HP 3PAR StoreServ 7450 Storage system with previous generation FC targets, or customers contemplating purchasing an All-Flash HP 3PAR StoreServ 7450 Array, are advised to consider upgrading to or purchasing Gen 5 16 Gb/s FC targets and deploying the storage system in a Gen 5 16 Gb/s FC infrastructure. This should pay off quickly through a greater capacity for work completed across the infrastructure, along with a higher ROI across the datacenter through port reductions on hosts and switching, while maintaining or increasing output from bottlenecked servers and virtual machines (VMs).

Appendix A – Test Description and Environment

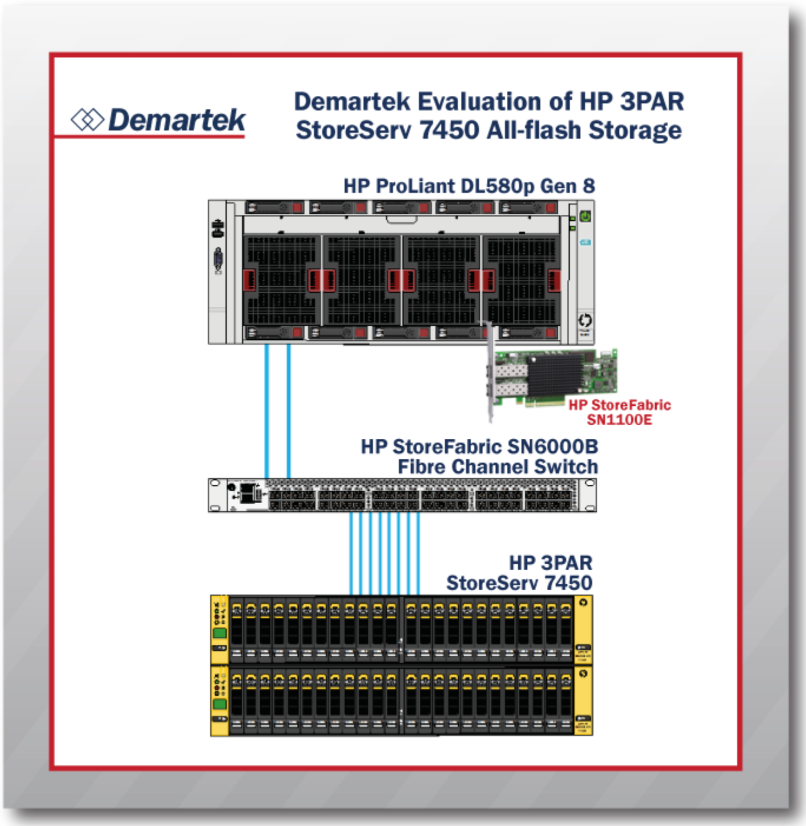

Figure 5 - Test Infrastructure

Server

- HP ProLiant DL580 G8
- 4 Intel E7-4890v2 2.8GHz CPUs
- 256 GB RAM (MS SQL Server constrained to 16GB RAM to force storage I/O)
- HP StoreFabric SN1100E 16Gb Dual Port Fibre Channel Host Bus Adapter (Generation 5 16 Gb/s Fibre Channel) or Emulex LPe12002 (8 Gb/s FC) dual port HBAs
- Microsoft Windows Server 2012 R2
- Microsoft SQL Server 2012, Microsoft TPC-E benchmark kit

Fibre Channel Switch

- Brocade 6510 (sold by HP as the HP StoreFabric SN6000B 16Gb 48-port/24-port active Fibre Channel Switch)

Storage Array

- 4-node HP 3PAR StoreServ 7450 Storage
- 3PAR Operating System 3.2.1 MU2
- 48 x 400GB 6Gb/s SAS SSD in 2 controller enclosures
  - 1 1500GB RAID5 data volume
  - 1 500GB RAID10 log volume
- 8x 8GFC target ports or 8x 16GFC target ports (HP OEM'ed Emulex target adapters)

Appendix B – All-Flash HP 3PAR StoreServ 7450 Storage Architecture

The All-Flash HP 3PAR StoreServ 7450 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 across the HP 3PAR StoreServ Storage family, which shares the same flash-optimized architecture and operating system. This flash-optimized architecture reduces the performance bottlenecks that can choke general-purpose disk arrays that have been retrofitted with solid-state disks (SSDs), and includes enterprise-grade features that are not always available with other all-flash offerings.

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

Adaptive Read and Write

Adaptive Read and Write is a software innovation that avoids unnecessary reads and writes to flash media by matching host read and write I/O size to flash media on a more granular level than is used with spinning medias. 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 software-based optimization for flash 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, to optimize for flash, flusher threads have been added to 3PAR cache management so the system 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 virtual desktop infrastructure (VDI) deployments by breaking large I/O into 32 KB sub-I/O blocks. 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 getting 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 (i.e. OLTP and Decision Support) running on the HP 3PAR StoreServ 7450 in this evaluation, enhancements like multi-tenant I/O processing help the array deliver consistently low latency for the OLTP application while assuring good bandwidth for the Decision Support application.

Specialized ASIC for Mixed Workload Support and Zero Detection

In the HP 3PAR Architecture, transactional and throughput-intensive workloads are able to run on the same storage resources without contention because of the system's specialized 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 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 series of zeroes that may be present in a write I/O stream 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 extend media lifespan.

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 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 Storage 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 RAID rebuild, resulting in faster rebuilds.

Moreover, with Adaptive Sparing, 3PAR StoreServ arrays are able to release the spare chunklets reserved on each solid state drive by drive manufacturers (for the purpose of sparing), allowing the SSDs to use that spare capacity as additional over-provisioned capacity. This increase in over-provisioned capacity enhances endurance of SSDs within the 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.

Appendix C – A Brief Summary of Gen 5 Fibre Channel

Gen 5 (16 Gb/s) FC provides an increased ceiling in terms of the possible performance for each connection, doubling the throughput of the previous generation (8 Gb/s FC). Database administrators often tune their applications to take advantage of every possible speed benefit, and Gen 5 FC 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 FC includes retimers in the optical modules and transmitter training, features that improve link performance characteristics, electronic dispersion compensation and backplane links.

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

Table 1 – Fibre Channel Speed Characteristics

Speed	Throughput (Mbps)	Line Rate (Gb/s)	Encoding	Retimers in the module	Transmitter training
Gen 1 (1 GFC)	100	1.0625	8b/10b	No	No
Gen 2 (2 GFC)	200	2.125	8b/10b	No	No
Gen 3 (4 GFC)	400	4.25	8b/10b	No	No
Gen 4 (8 GFC)	800	8.5	8b/10b	No	No
Gen 5 (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
Gen 1 (1 GFC)	300	500	860	*	10,000
Gen 2 (2 GFC)	150	300	500	*	10,000
Gen 3 (4 GFC)	50	150	380	400	10,000
Gen 4 (8 GFC)	21	50	150	190	10,000
Gen 5 (16 GFC)	15	35	100	125	10,000

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

The original version of this document is available at: http://www.demartek.com/Demartek_HP_3PAR_StoreServ_7450_16GFC_2015-03.html on the Demartek website.