Technical white paper

HP SmartCache technology

Table of contents

Abstract

As an IT professional, you are seeking cost effective ways to increase the performance of applications. With CPU core counts growing and addressable memory space increasing, storage can become the main bottleneck for virtualized and data-intensive environments. HP ProLiant Gen8 servers address application performance with Dynamic Workload Acceleration, an architectural innovation that employs HP SmartCache technology for accelerating access to multi-terabytes of data in either a Direct-Attached Storage (DAS) or Storage Area Network (SAN) environment.

This paper describes HP SmartCache technology—how converging solid-state and hard disk drive technologies with intelligent control can offer high performance and high capacity without incurring the cost burden of an all-solid state drive (SSD) configuration.

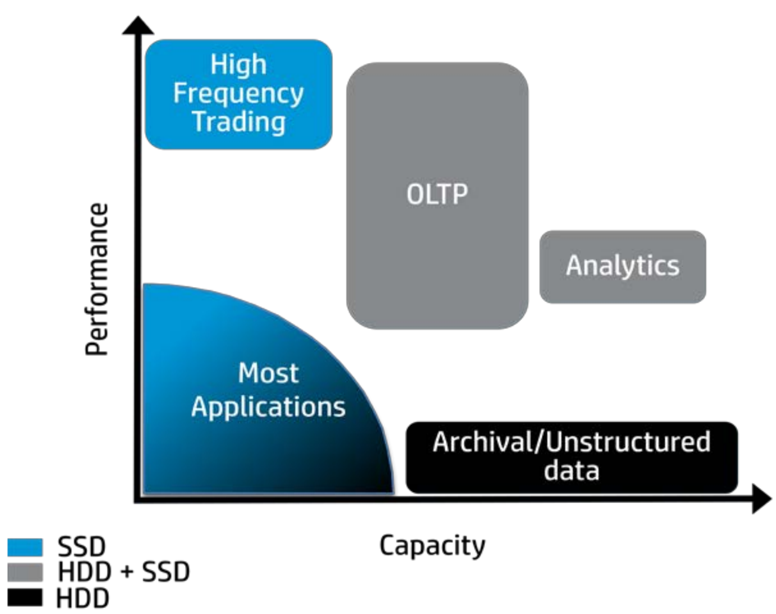
Introduction

Storage needs for enterprise systems are advancing rapidly in both performance and capacity and most applications require a balance of those attributes. However, certain applications emphasize one attribute over another. For instance, applications that process many read and write requests to the storage subsystem make performance a priority. For applications that manage archival data that grows continuously, capacity is vital.

Two storage technologies commonly address application needs. The SSD, with its very low latency, is the answer for performance-conscious applications. The hard disk drive (HDD), with its long-term reliability and economical per-GB cost, meets large capacity needs. Figure 1 illustrates the suitability of SSDs and HDDs for meeting the needs of different applications.

- High frequency trading applications require very lower latency. An all-SSD solution would provide the performance this environment requires.
- Online transaction processing (OLTP) and analytics push the boundaries of both performance and capacity growth. A mix of SSDs and HDDs best meets those needs.
- Archival and smartphone applications require unstructured data growth, a need suitably addressed with high capacity HDDs.

Figure 1. Different applications have varying storage needs.

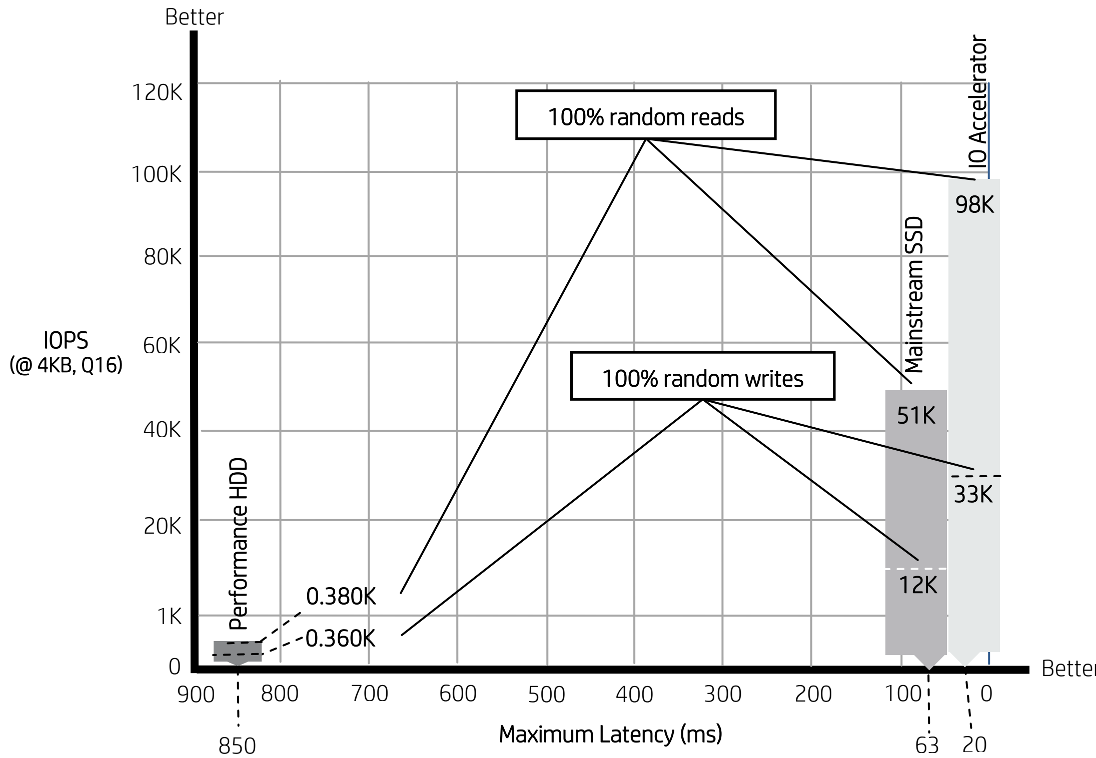


Comparing storage technology performance

SSDs provide significant performance advantages over traditional HDDs. Figure 2 illustrates the relative performance in Input/output Operations Per Second (IOPS) and latency for a performance HDD, a mainstream SSD, and an IO Accelerator. The HP IO Accelerator is an interface card containing an SSD, flash memory, and a custom controller that significantly improves data throughput with external storage devices.

As shown in Figure 2, for each vertical bar, the top number indicates the IOPS number for 100% random reads and the bottom number indicates the IOPS number for the 100% random writes. The chart illustrates how the mainstream SSD and HP IO Accelerator vastly outperform the HDD and are capable of reading and writing data much faster.

Figure 2. IOPS and latency performance of storage hardware.



Benchmark IO Meter 2006

SSD technology is the foundation for both the HP IO Accelerator and SATA SSD; however, their performance capabilities differ. The underlying SSD implementation in the HP IO Accelerator provides faster read and write performance:

- The HP IO Accelerator's read performance is almost double that of the mainstream SATA SSD.
- The HP IO Accelerator's write performance is roughly five times that of the SSD.

HP IO Accelerators and SSDs obviously address performance requirements but at a cost that may be prohibitive for applications requiring high capacity and scalable performance.

What about hybrid drives?

Hybrid drives combine a small SSD with traditional HDD technology in an integrated package and are a good option in the consumer space. However, in the server space where you have many HDDs, hybrid drives lack the flexibility to scale SSD technology differently based on application needs. If an application is referencing data from just one drive out of dozens in a hybrid drive configuration, performance will suffer from the small SSD capacity on the single drive and the unused SSD capacity on the other hybrid disks becomes a stranded resource. With a larger SSD dedicated for caching, performance scales better as capacity or the amount of hard disk drives is increased.

Why caching?

So you may ask, if SSDs are superior in performance, why not use HP IO Accelerators and SSDs in all server configurations? The answer is that you need to weigh requirements for capacity and performance against the increased costs for those abilities. As the cost of SSDs is trending downward, SSD technology is making inroads into the enterprise server market to address the low latency requirements for certain applications. However, even with the performance benefit and lower cost, SSD technology will not replace HDD technology in the near future since HDDs remain an economical and durable (long-term) storage medium. Also, SSDs support a finite number of write cycles before they begin to degrade, although their durability is improving and the capacity is increasing to where SSDs are more viable for enterprise applications.

Using a high performing all-SSD configuration for an application may be viable for applications with modest capacity requirements. Still, it is difficult to ignore HDDs for their high capacity and greater durability at a very economical price, particularly in enterprise systems that may require thousands of storage devices.

HP SmartCache technology eases your decisions by providing an incremental, economical way to introduce SSD technology into your environment. HP SmartCache provides SSD performance benefits while simultaneously preserving investments in previously purchased hard disk assets.

How does caching work?

Caching works by intelligently placing data in a lower latency device so that responses to future requests for that data can occur much quicker. If an application requests data that is in the cache (called a “cache hit”) then the lower latency device can service that transaction. Otherwise, a “cache miss” occurs and the data must be accessed from the original, slower device. As more cache hits occur, overall performance improves.

Caches are very common in computers. Processors use faster memory caches to speed up instruction execution as they fetch data from memory DIMMs. HDDs contain caches (also known as buffers) that allow data to be queued to and from the magnetic media. Host operating systems utilize caches to improve application performance. Caching has proven itself in many areas of computing because of the nature of how applications access data. Caching improves performance while maintaining a cost-effective platform. As storage capacities continue to expand, SSDs bring caching to storage systems to provide a cost-effective performance boost, especially to server configurations.

HP SmartCache architecture

HP SmartCache combines different technologies and device types to close the cost/performance gap. SmartCache uses a caching architecture where a copy of the data resides on the hard disk drive as well as on a lower latency device that is used for caching. The basic HP SmartCache architecture is comprised of the following three elements:

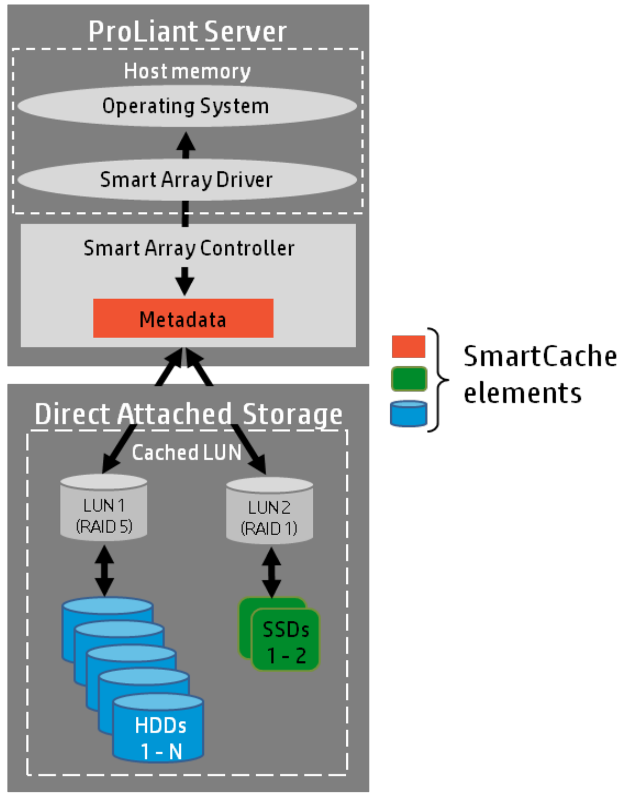
- Bulk storage: The first element is the bulk storage device, which can be either HDDs or connections to SAN storage.
- Accelerator: The second element, the accelerator, is a faster/lower latency device that caches data. The capacity of the accelerator is less than the capacity of the bulk storage device.
- Metadata: The final element is metadata, information held in a relatively small storage area that maps the location of information residing on the accelerator and bulk storage devices.

The HP SmartCache architecture is flexible and supports numerous device types for bulk storage, accelerator, and metadata. The HP SmartCache architecture can adapt to accommodate the evolution of storage devices. Initially available HP SmartCache solutions are designed for DAS and SAN systems.

HP SmartCache for Direct-Attached Storage

The direct-attached HP SmartCache solution includes the three elements of the HP SmartCache architecture; HDDs serving as bulk storage, SSDs as the accelerator, and Flash-Backed Write Cache (FBWC[1]) memory for storing the metadata (Figure 3). For this implementation, the SmartCache control layer resides in the firmware of the onboard Smart Array Controller of the HP ProLiant Gen8 server, below the operating system and driver. This allows caching for devices connected to a single array controller.

Figure 3. HP SmartCache can be applied to direct attached storage.



HP SmartCache offers flexibility in creating logical disk volumes from hard disk drives:

- The accelerator or cache volume is designed to support any RAID configuration supported by the Smart Array controller.
- Each logical disk volume can have its own cache volume or none at all.
- Cache volumes can be created and assigned dynamically without adversely impacting applications running on the server.

Only SSDs can be used for cache volumes, and a cache volume may be assigned only to a single logical disk volume. The HP SmartCache solution consumes a portion of the FBWC memory module on the Smart Array controller for metadata. To ensure sufficient storage for accelerator and metadata, we recommend:

- 1 or 2 GB of FBWC memory

or

- 1 GB of metadata space for every terabyte of accelerator space

When using the HP SmartCache solution for direct-attached storage, legacy cache is still present and operational and utilizes the remaining space in the FBWC memory. When using HP SmartCache, we recommend setting the legacy cache for 100 percent write operation. This allows write-back support on the Smart Array controller, which accelerates write performance. This also allows SSDs configured as HP SmartCache volumes to provide much larger read cache.

Since the HP SmartCache solution is implemented within the Smart Array firmware, this solution can be used on any operating system supported by ProLiant servers. The HP SmartCache solution operates transparently, without any dependencies on operating system, device driver software, file system type, or applications.

The initial release of the HP SmartCache solution supports write-through caching. When an application writes data to the disk, the Smart Array controller writes the data to the HDDs and if the data is also in the cache volume, the data is written to the SSD.

[1] FBWC: Flash Backed Write Cache, an integrated persistent memory device used for caching operations conducted within the Smart Array controller's firmware. Writes are posted to the FBWC, acknowledged immediately, and later are completed to the actual hard disk drives for improved latency of write operations.

HP SmartCache for SAN Storage

Future HP SmartCache solutions will support external storage configurations, including SAN storage products such as the HP 3PAR StoreServe solutions. HP SmartCache will also support additional accelerators like the HP IO Accelerator. For SAN implementation, the HP SmartCache control layer moves into the host memory with the operating system above the device drivers and beneath the file system. This allows caching to be independent of the underlying hardware controller. With HP SmartCache running in the host operating system, this allows support of various devices such as both bulk storage devices as well as accelerators. For example, cache devices may be either HP IO Accelerators or SSDs connected to Smart Array controllers. The initial versions of HP SmartCache for SAN Storage utilize host memory for metadata to support both read-only and write-through caches.

The HP SmartCache architecture adapts as future technology changes. Initial releases of HP SmartCache for SAN will support Windows, Linux and VMware operating systems.

Management and analytics

To configure HP SmartCache, use the HP Smart Array configuration utilities: the Array Configuration Utility (ACU) or its command-line version (ACUCLI).These utilities will also incorporate analytics to provide more visibility into the disk storage subsystem to show what is happening within HP SmartCache, as well as detailed information on both the physical and logical disks. In addition, integration with HP Insight Control comes standard, providing information and alerts about HP SmartCache status.

Future capabilities may allow customers to profile their application and then model the HP SmartCache performance through a simulator allowing various accelerator substitutions by capacity and performance capability.

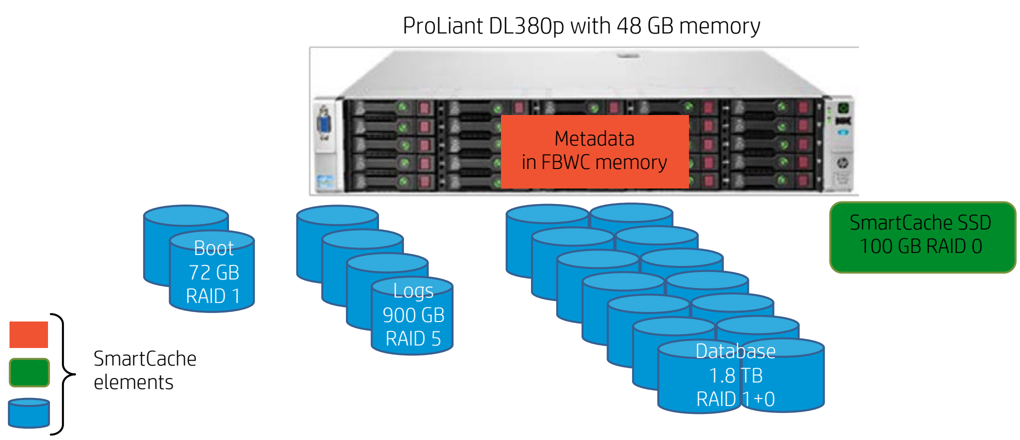
Case study: OLTP workload

Data integrity is the most important characteristic for a disk storage subsystem. HP SmartCache technology provides a reliable, enterprise class solution. Performance, which is dependent upon the application, and the price:performance ratio are the characteristics that generate interest in HP SmartCache. HP SmartCache improves performance when applications produce the following conditions:

- There is more read traffic than write traffic
- Data is referenced repeatedly, which is common in many applications

For our OLTP workload case study, we ran a Transaction Processing Performance Council type E (TPCe)-like[2] benchmark on a ProLiant DL380 server with 48 GB of memory, a Smart Array P421 controller, and a 2-GB FBWC module configured as 100% write cache.

Figure 4. This shows the configuration used for the TPCe-like case study.



[2] Modified TPCe workload as defined in presentation “Using Solid State Drives As a Mid-Tier Cache in Enterprise Database OLTP Application” by Badriddine Khessib, Microsoft for TPC Technology Conference 2010 – Singapore, (http://www.tpc.org/tpctc2010/TPCTC2010-12.pdf)

The Smart Array P421 controller disk configuration included the following:

- Boot Volume: 2 - 72GB 10K 6G SAS drives configured in RAID 1
- Database Logs: 4 - 300GB 10K 6G SAS drives configured in RAID 5
- Database: 12 - 300GB (10K 6G SAS drives configured in RAID 1+0
- HP SmartCache: 1 – 100GB SSD MLC 6G SSD drive configured in RAID 0. The HP SmartCache was configured in write-through mode and assigned to cache the database volume

The database size was roughly 1.6TB during the two-hour run. Figure 5 shows the performance of the TPCe-like workload both with HP SmartCache (red) and without HP SmartCache (blue). TPeS performance of the TPCe-like workload was significantly better when using HP SmartCache. The SSD used as a cache was less than 7% the capacity of the database. The cache-hit ratio of HP SmartCache was 84%, meaning that roughly five out of every six disk reads to the database volume were serviced by the SSD versus the HDDs. A smaller amount of HP SmartCache would have yielded a smaller cache hit percentage and lower performance gains.

Figure 5. TPCe-like workload performance with and without HP SmartCache.

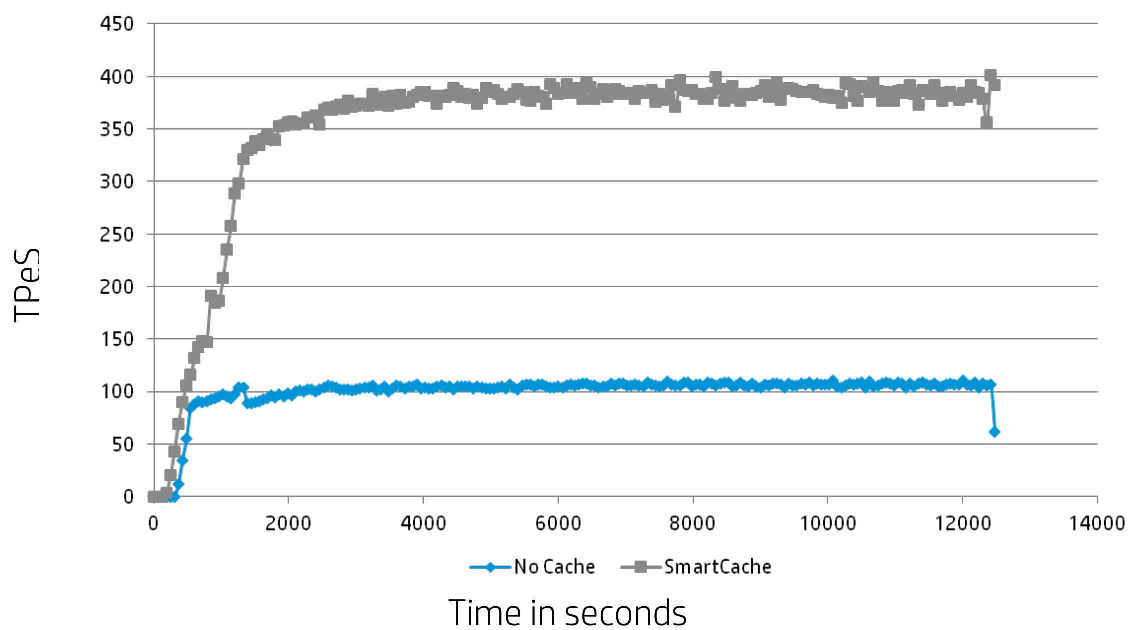


Table 1 compares the database logical volume IO statistics without and with HP SmartCache. With the addition of the HP SmartCache, performance results change dramatically simply by adding a 100-GB SSD.

Table 1. TPCe-like workload statistics

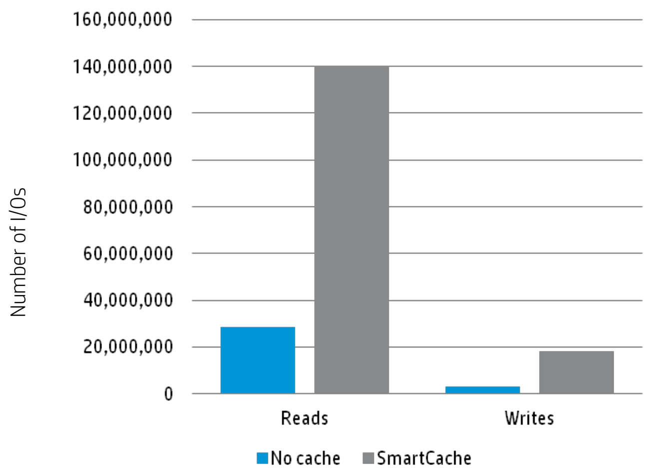
I/O Analysis	Without HP SmartCache	With HP SmartCache
Accesses	32,767,015	159,239,037 (almost 5x more)
Reads	29,114,040	140,159,417 (almost 5x more)
Writes	3,652,975	19,079,620 (5x more)
% reads	89	88

% writes	11		12	
Average IOPS, Read	4,382		19,464 (over 4x better)	
Average IOPS, Write	550		2,650 (almost 5x better)	
Duration (seconds)	6644		7201	
I/O distribution (reads)	**Count**	**Avg. latency (msec)**	**Count**	**Avg. latency (msec)**
0% @ 4K	0	0.00	0	0.00
97.6% @ 8K	28,430,311	27.97	140,159,417	6.23
0% @ 16k	71	28.21	0	0.00
<1% @ 32K	1,457	21.54	0	0.00
2.3% @ 64K	682,201	18.79	0	0.00
0% @ 128K	0	0.00	0	0.00
0% @ 256K	0	0.00	0	0.00
100% (totals)	29,114,040	27.75	140,159,417	6.23
I/O distribution (writes)	**Count**	**Avg. latency (msec)**	**Count**	**Avg. latency (msec)**
<1% @ 4K	93	10.18	98	45.39
82.9% @ 8K	3,028,218	0.13	18,015,518	89.35
16% @ 16K	585,670	0.15	1,031,087	89.27
<1% @ 32K	18,322	0.22	18,745	89.90
<1% @ 64K	13,607	0.29	11,419	90.06
<1% @ 128K	5,856	0.36	2,326	90.01
<1% @ 256K	1,209	0.46	427	90.10
100% (totals)	3,652,975	0.13	19,079,620	89.35

Taking a closer look at the data from Table 1, we see that the transaction volume through the database logical drive increased significantly with HP SmartCache. As shown in Figure 6, the disk subsystem without SmartCache was able to perform approximately 29 million reads and 3.6 million writes (blue). With HP SmartCache, the disk subsystem performance increased significantly to approximately 140 million reads and 19 million writes (gray) within roughly the same period. With

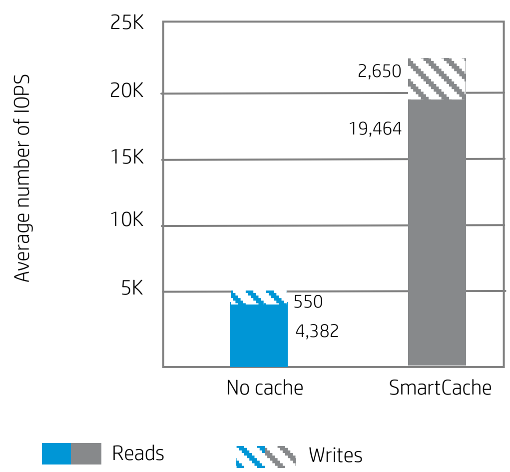
HP SmartCache, the disk subsystem was able to process approximately four times as many requests than without HP SmartCache.

Figure 6. Transaction volume through database volume.



Looking at the data another way, we see the differences in the number of input/output transactions performed per second (IOPS) with and without HP SmartCache (Figure 7). Without HP SmartCache, the disk subsystem was able to process approximately 5000 requests per second combining both read and write traffic. With HP SmartCache, the disk subsystem was able to process approximately 22,000 requests per second of both reads and writes. This shows that the disk subsystem with HP SmartCache is able to process four times as much data than the system without HP SmartCache.

Figure 7. HP SmartCache provides increased IOPS.



Furthermore, without HP SmartCache, we can see that the average read latency per request was almost 28 milliseconds and the average write latency was practically non-existent at 0.13 milliseconds (Figure 8). This workload highlights the benefit of the FBWC of the Smart Array controller. The FBWC of the Smart Array controller completes the write operation by saving the data within the FBWC memory, which is much quicker than writing the data out to the hard disk drives.

Once data is saved in the FBWC, the Smart Array controller ensures that the data is written to the hard disk drives. The primary function of the FBWC is to protect the data in the FBWC in case of unplanned power loss to the server. From an application workload perspective, the write load of the application o the Smart Array controller is not high enough nor sustained long enough to fill the capacity of the FBWC so the write IO response time is excellent.

Figure 8. Average latency per I/O.

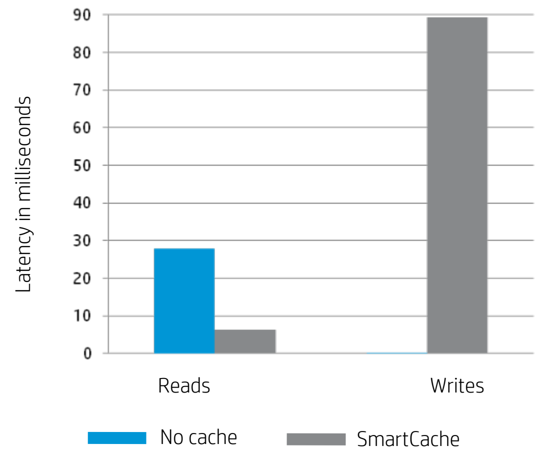
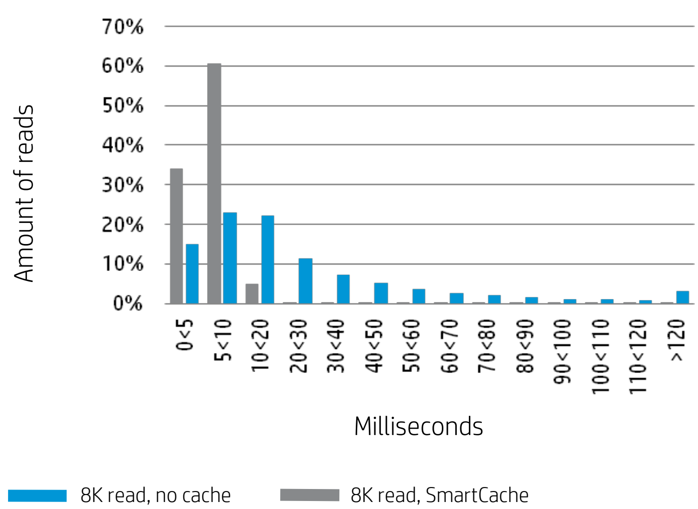


Figure 9 also shows that read latency drops from almost 28 to 6.23 milliseconds per request when HP SmartCache is enabled. This behavior is expected as more reads are satisfied by the accelerator device or SSD, which demonstrates the value of HP SmartCache. What might not be expected is the increase in write latency when HP SmartCache is enabled, which will be explained shortly.

Looking at the distribution of latencies per transaction for 8K read traffic (Figure 9) shows the benefits of HP SmartCache. With HP SmartCache enabled, 95% of the read requests completed in under10 milliseconds. Without HP SmartCache, the distribution of latencies is spread out over the spectrum of response times as there is more dependency on the rotation access time of the hard disk drives. This rotational latency results in a near 28-millisecond average per read request.

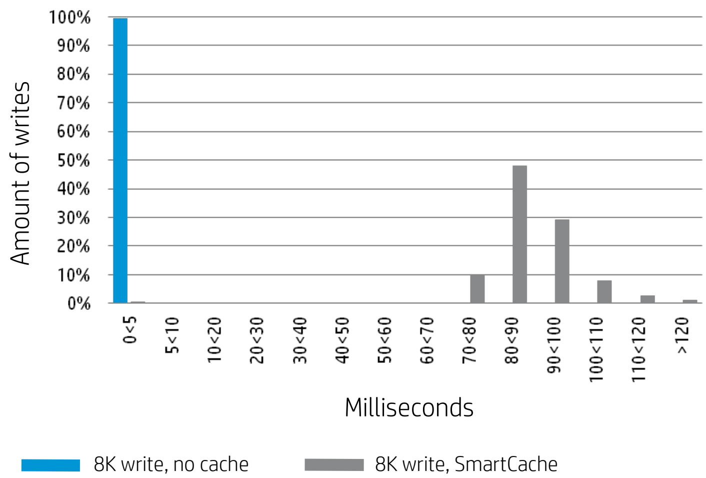
Figure 9. 8K Read Latency Distribution



Write latency jumped significantly from an average of 0.13 milliseconds without HP SmartCache to 89.35 milliseconds with HP SmartCache (Figure 10). Without HP SmartCache, the write requests were completed with low latency by the FBWC. With HP SmartCache, the increase in read requests serviced by the faster SSD results in a workload shift on the HDDs from mostly reads to mostly writes.

Note also that using HP SmartCache results in the volume of writes from the application exceeding the FBWC capacity. When the FBWC memory is full, the write latencies increase since write operations cannot complete until data is written to the HDDs. This explains the increase in write latencies when using HP SmartCache. However, the benefit of substantially increased read performance far outweighs the increase in write latency, resulting in better overall performance to the application.

Figure 10. 8K Write Latency Distribution



Conclusion

While other applications such as databases, analytics, and virtual desktops can benefit from acceleration techniques, TPCe-like workloads present the biggest challenge, and HP SmartCache succeeds. The characteristics of TPCe—more read traffic (88%) than write traffic (12%) and repetitive data requests—allow HP SmartCache to enhance performance while significantly reducing your storage costs.

Other types of workloads may not benefit from HP SmartCache, including those with the following characteristics:

- Non-repetitive I/O – Applications that do not access data multiple times over a period of time will not benefit from HP SmartCache. Examples include routine backups, as they only accesses every file once, and disk benchmark utilities, as their disk access patterns are typically random.
- Write-heavy I/O – Applications which perform predominantly write traffic versus read traffic will not benefit from HP SmartCache. While database volumes benefit from HP SmartCache, database logging volumes will not benefit from HP SmartCache because the majority of traffic to these volumes is writes to the hard disk drives.

HP SmartCache is the cost-effective, scalable, high-performance solution for accelerating data access for multi-terabyte data sets for either DAS or SAN environments.

Resources, contacts, or additional links

HP SmartCache
http://www.hp.com/go/smartcache

HP Dynamic Workload Acceleration
http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA3-9649ENW.pdf

HP PCIe IO Accelerator for ProLiant servers
http://h18006.www1.hp.com/products/servers/proliantstorage/solid-state/benefits.html?jumpid=reg_r1002_usen_c-001_title_r0007

TC1211951, February 2013