The Ideal Storage for IBM® Power Systems™

December 22

2014

IBM FlashSystem™ unleashes the performance of Power Systems servers

This paper illustrates the complementary synergy achieved by combining IBM Power Systems server technology with IBM FlashSystem storage technology, resulting in ultra-high performance solutions.

By:

William Bostic – IBM FlashSystem Portfolio Manager

Agustin Mena III – IBM Senior Technical Staff Member, Power Systems Performance

David Siegel – IBM Senior Technical Staff Member, Integrated Solutions Segment Manager

Brian Twichell – IBM Senior Technical Staff Member, FlashSystem Integration Architect

Olga Yiparaki – IBM Senior Technical Staff Member, Storage Systems Performance

© Copyright IBM Corp., 2014

1. Introduction

The revolution in IT driven by Cloud, Analytics, and Mobile and Social applications has led to an explosion in the amount of data that is retained, updated, and available for analysis. In order to effectively manage and process this data, powerful computing infrastructure is required. Selecting the right servers and storage (*infrastructure that matters*) is crucial to the success of data-intensive environments.

The speed with which servers and storage can process data is typically characterized in terms of three dimensions: response time, bandwidth, and throughput. Response time is simply the amount of time required to read, write, or process a set of data. Typical units for response time are seconds (in the case of an online transaction), milliseconds (in the case of a hard disk drive read or write operation), and microseconds or nanoseconds (in the case of access to flash and DRAM respectively). Bandwidth is the ability to move or process large amounts of data and is typically characterized in units of gigabytes per second. Throughput is the rate at which several independent operations can be completed; it is typically represented in units such as transactions per second or I/Os per second (IOPS).

Different business environments have different requirements in terms of which of these three dimensions are most important. Analytics environments usually need to be able to process large datasets, so high bandwidth and scalability are required. On the other hand, mobile, social, and conventional transaction-processing environments often emphasize response time.

Clearly, for a server or storage system to be versatile enough to excel across this spectrum of environments, it must be strong in all three dimensions. But, what is not immediately apparent is that, because servers and storage are interconnected, they operate together, and their capabilities need to be well-matched across the three dimensions in order to optimally perform. After all, there's little point in investing in a powerful server if it's bottlenecked by the storage to which it's attached.

This paper describes how IBM Power Systems servers and FlashSystem storage combine synergistically to deliver leading-edge capabilities along each of the three dimensions, thus making the combination an outstanding infrastructure choice for data-intensive organizations. The remainder of the paper is structured as follows: Section 2 describes the evolution of server performance, concentrating on the improvements in performance seen in the POWER® processor over time. Section 3 discusses the evolution of storage performance, focusing on the order of magnitude improvements in response time, throughput, and density achieved with FlashSystem technology. Section 4 covers specific elements of the synergy between Power Systems servers and FlashSystem storage. Section 5 provides concrete examples of business applications that are accelerated by the Power Systems/FlashSystem combination. Finally, section 6 provides conclusions.

© Copyright IBM Corp., 2014

2. The Evolution of Server Performance

With the advent of multi-core and multi-threaded processors the amount of processing capacity available, even in small and midrange servers, is at least doubling with each new platform generation (Figure 1). The secret to leveraging that capability is to keep all those cores busy doing productive work. Whenever hardware threads are idling, stalling, or spinning for whatever reason, processor cycles are wasted. Until flash, storage performance has not kept up with the rate and pace of server advancement. As a result, an ever-increasing number of processor cycles is wasted waiting on I/O.

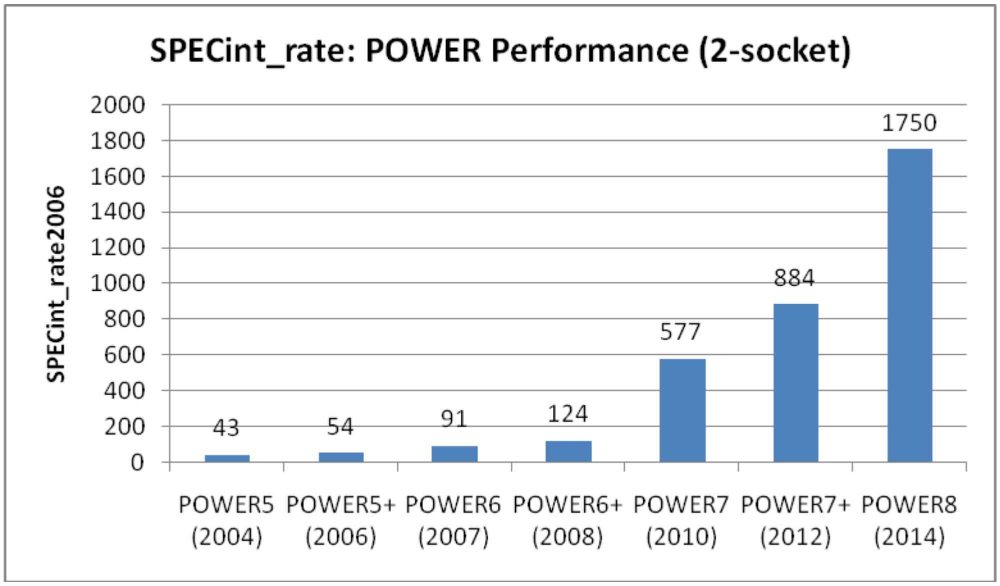

Figure 1 – Relative performance across generations of POWER processors

The best system solutions are a synergistic integration of complementary technologies. To achieve the most leverage of all the technologies employed, it pays to begin by understanding the behavior and needs of the application, as well as the business requirements. Then one is able to make informed decisions regarding the selection of complementary server and storage technology.

The Value Proposition of Power Systems

The foundation behind the Power Systems platform is the POWER processor. The POWER core is designed for industrial strength performance with strong hardware threads, delivering leadership SAP[i], integer and floating point[ii], and application serving performance[iii]. Therefore, Power Systems servers are about strong cores with scalable performance. This scalable performance becomes evident in multi-threaded workloads that take advantage of the simultaneous multithreading (SMT) capabilities of the POWER processor. On POWER8™ systems, it is typical for those workloads to see up to a 2X improvement in throughput running with eight threads, compared to a single thread (Figure 2).

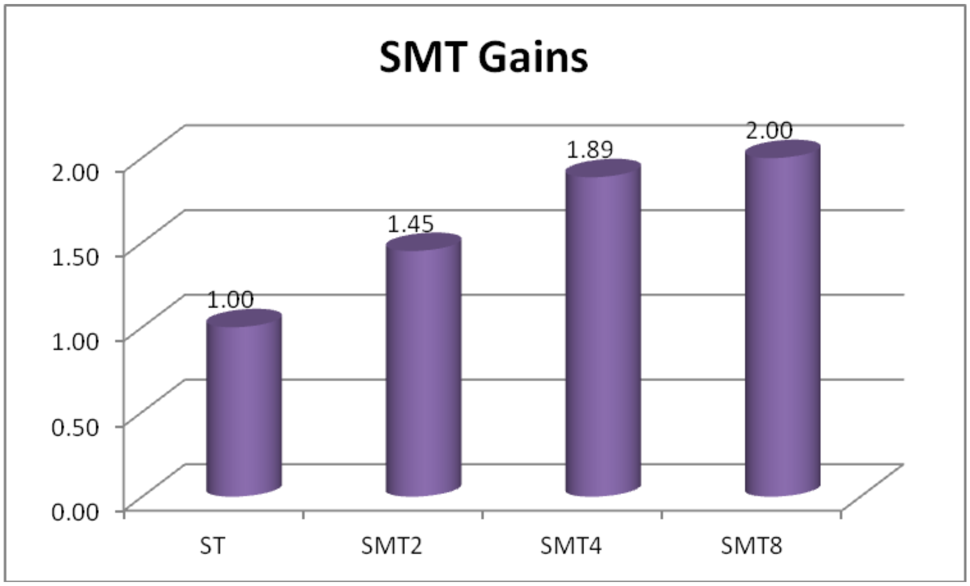

Figure 2 – Throughput comparison from a single thread (ST) per core to eight threads per core

To keep pace with all that compute capability, Power Systems at least doubled memory and I/O bandwidth with each new generation of the processor (Figure 3 and Figure 4).

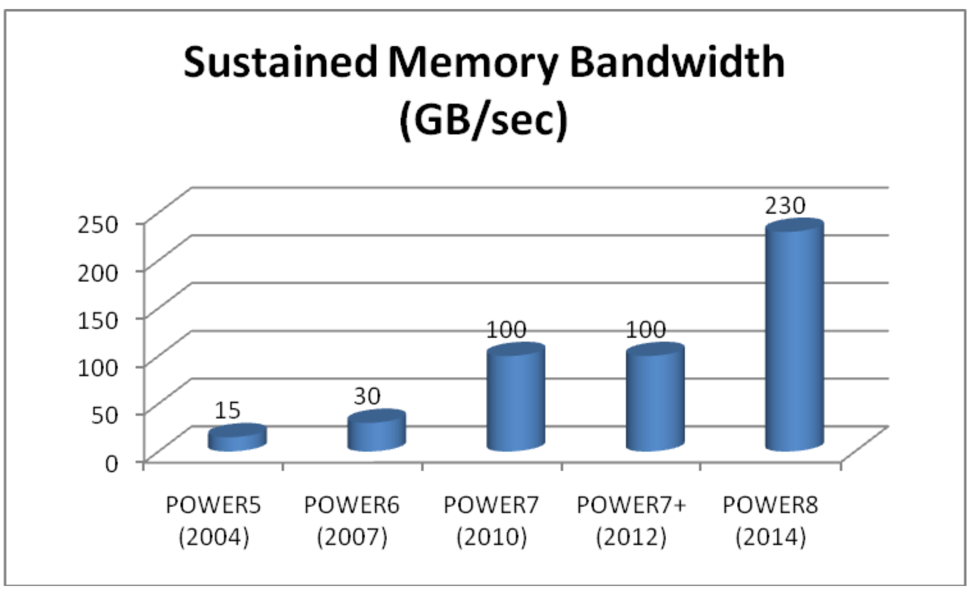

Figure 3 - Memory bandwidth improvement across generations of POWER processors

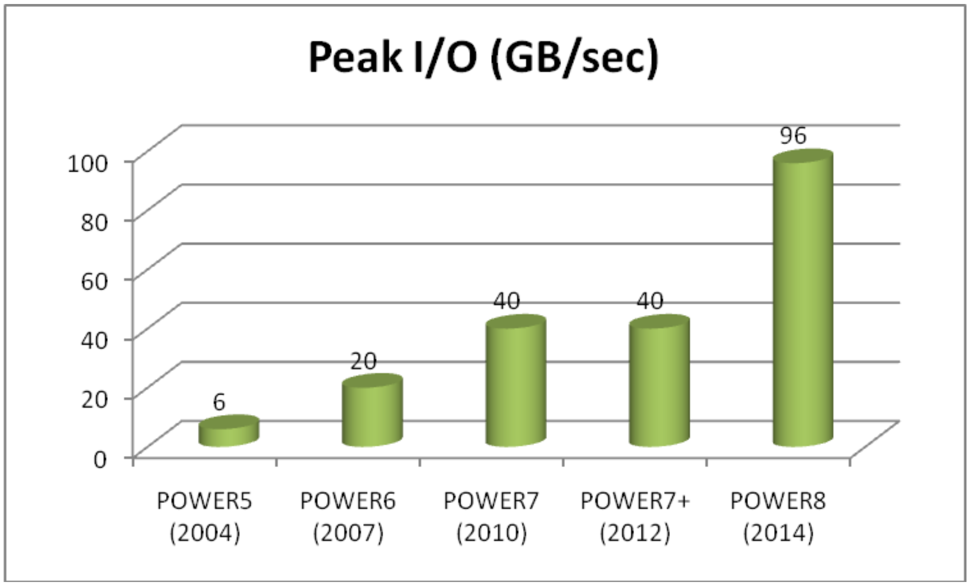

Figure 4 – I/O bandwidth improvement across generations of POWER processors

Such a design lends itself to mission-critical production environments where low latency transactions, high throughput, and massive scaling are important. Typical production environments include online transaction processing (OLTP), batch, analytics, business applications, and cloud. All of these tend to be data hungry applications and running them on high performance Power Systems servers amplifies the need for high performance storage.

3. The Evolution of Storage Performance

Storage solutions leveraging traditional spinning disk technology have not kept pace with the rapid performance growth enjoyed by the server industry. This situation has changed with the advent of flash technology, first with solid state disks (SSD) and PCIe flash adapter-based offerings, and more so now with high capacity performance optimized solid state arrays (SSA). Other strengths of SSAs include robust fault tolerance, concurrent maintenance, support for multiple interface types (most notably Fibre Channel), and integration of other advanced features such as encryption. For the purposes of this discussion we will focus on addressing the performance needs of a high performance server such as Power Systems. The following sections depict the impressive storage performance journey from disk-based storage to flash storage, culminating with the high performance FlashSystem 840.

History and Evolution of Key Performance Metrics

How did IBM get to today's performance, showcased in this paper? The path to today's impressive capabilities required relentless technological innovation, starting with the seminal RAMAC in 1956, and taking advantage of everything storage media could support at its time.

Looking at some of the key performance metrics such as bandwidth (Figure 5), we must take into account the physical dimensions of each system in order to normalize the throughput capabilities of each. Starting with the RAMAC, and considering the innovations every decade, the bandwidth has changed so significantly that the graph must be presented in a logarithmic scale.

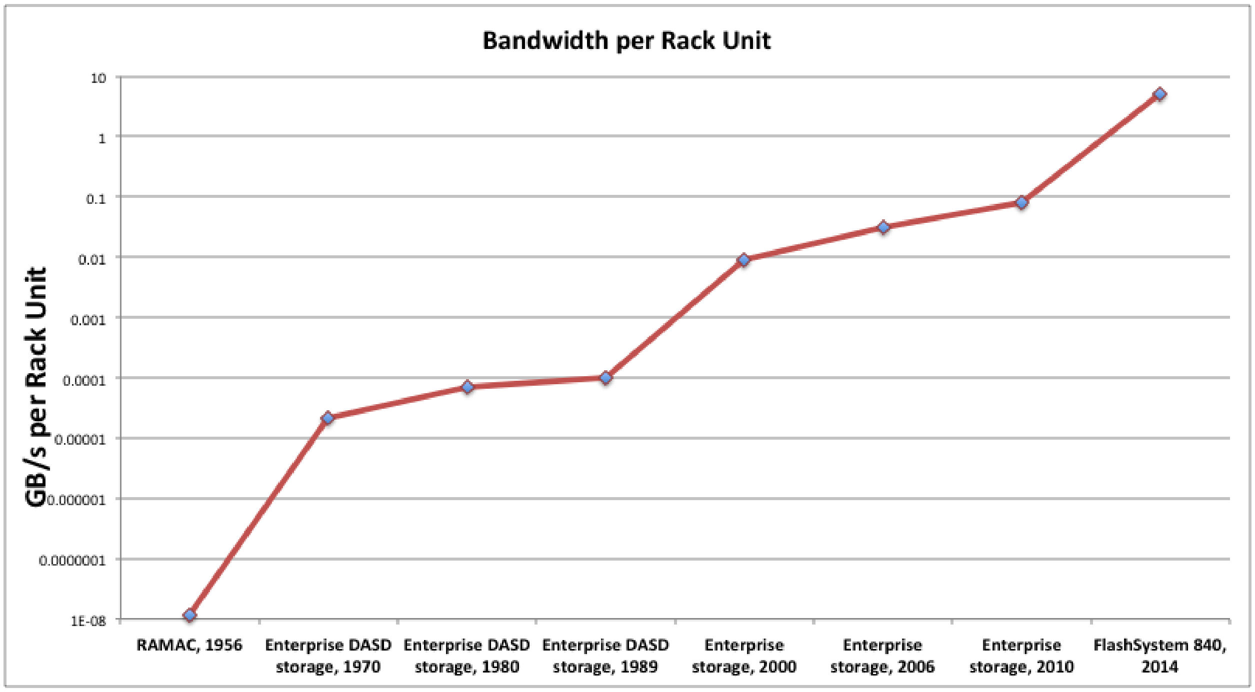

Figure 5 - Bandwidth GB/s per Rack Unit

© Copyright IBM Corp., 2014

In fact, the physical dimensions of these systems has changed (Figure 6) from the size of a refrigerator, approximately, to multiple frames (up to approximately four refrigerators in the case of state-of-the-art enterprise disk storage systems), and then down to two Rack Units (about the size of two pizza boxes) for today's FlashSystem 840, replacing multiple refrigerators with higher performance.

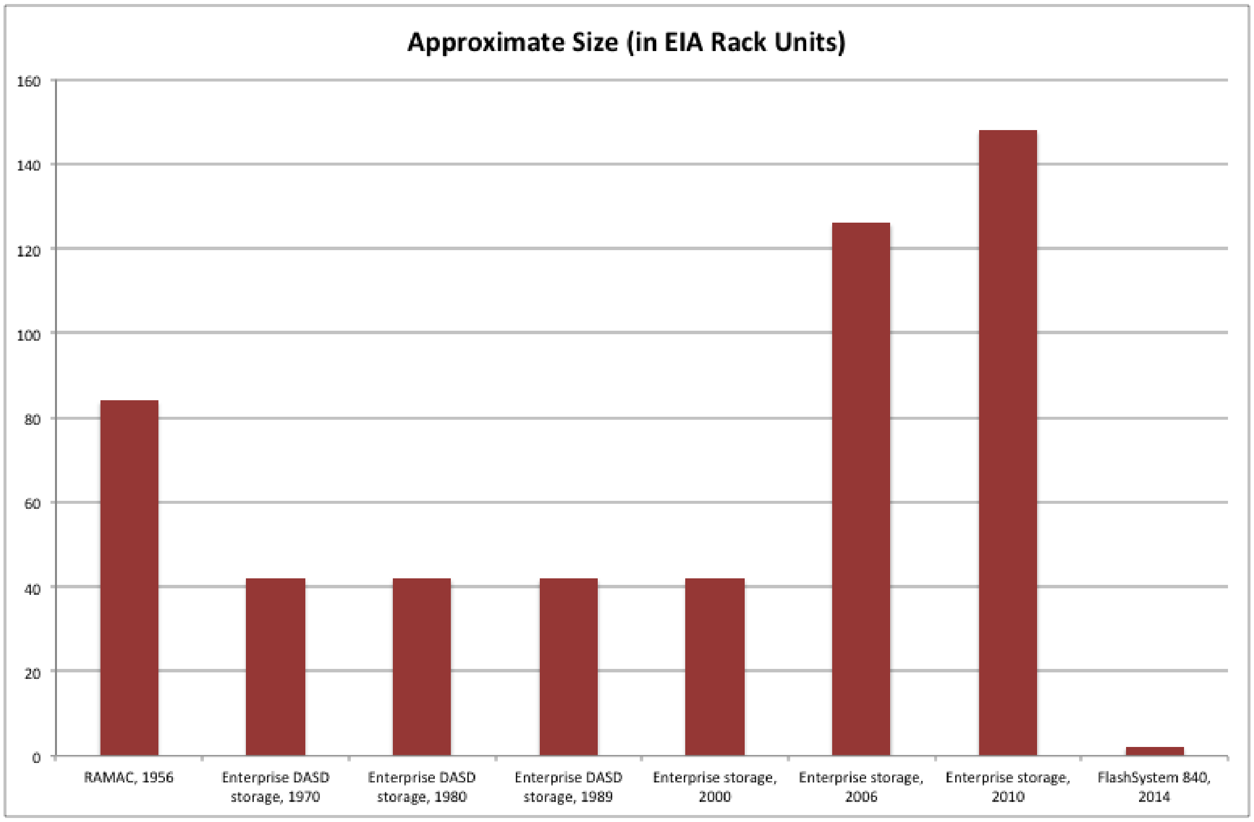

Figure 6 – System storage size

IBM's storage technology advances translate into tangible and significant performance capabilities that satisfy the needs of today's most demanding applications, with over a million IOPS. Additionally, the significant size reduction enables a dramatically more compact and higher performing solution with far greater energy efficiency to operate and cool, improving every dimension of today's data center storage solution.

Access Time Evolution

A key storage metric is the minimum time required to access a unit of data (such as a 4 KB page) from the storage media. For many years, it was a challenge to keep up with CPU and server performance. This has required significant innovation in storage system design, including seminal IBM technologies such as the introduction of cache and significant improvements of cache utilization using smart cache algorithms, Easy Tier® technology, and much more.

In addition to the impressive storage controller design improvements from 1956 to 2014, hard drive technology improved dramatically. Large single platters were replaced by faster and smaller drives, with the advent of RAID. Drive density, arm movement, rotational speed, and overall drive performance advanced by improving the mechanics and the use of buffers. Additionally, hard drive algorithms such as Rotational Position Optimization improved hard drive performance yet again, especially at high I/O rates and when the access pattern was not entirely random. However, despite significant improvements, hard drive technology did not keep pace with the demands for storage system performance.

Storage controller design innovations have marked significant turning points in the evolution of storage. All of them address a key need: to improve performance, especially in cases where the access pattern is not entirely random, so the system can anticipate near-future data accesses.

It is easy to gloss over the advancements that IBM made along the way, though diving into the details is beyond the scope of this paper. Worth noting is that in 2000, IBM received the National Medal of Technology and Innovation, recognizing 40 years of innovation in hard drives and storage system design.

However, the thorny problem of entirely random access remained. The minimum access time to data from the permanent media continued to be a challenge, until flash entered the market.

Figure 7 showcases the march toward zero, starting with the RAMAC and ending with today's FlashSystem 840 whose read latency is about 135 microseconds. Due to the dramatic improvements made, the graph requires a logarithmic scale to showcase the evolution.

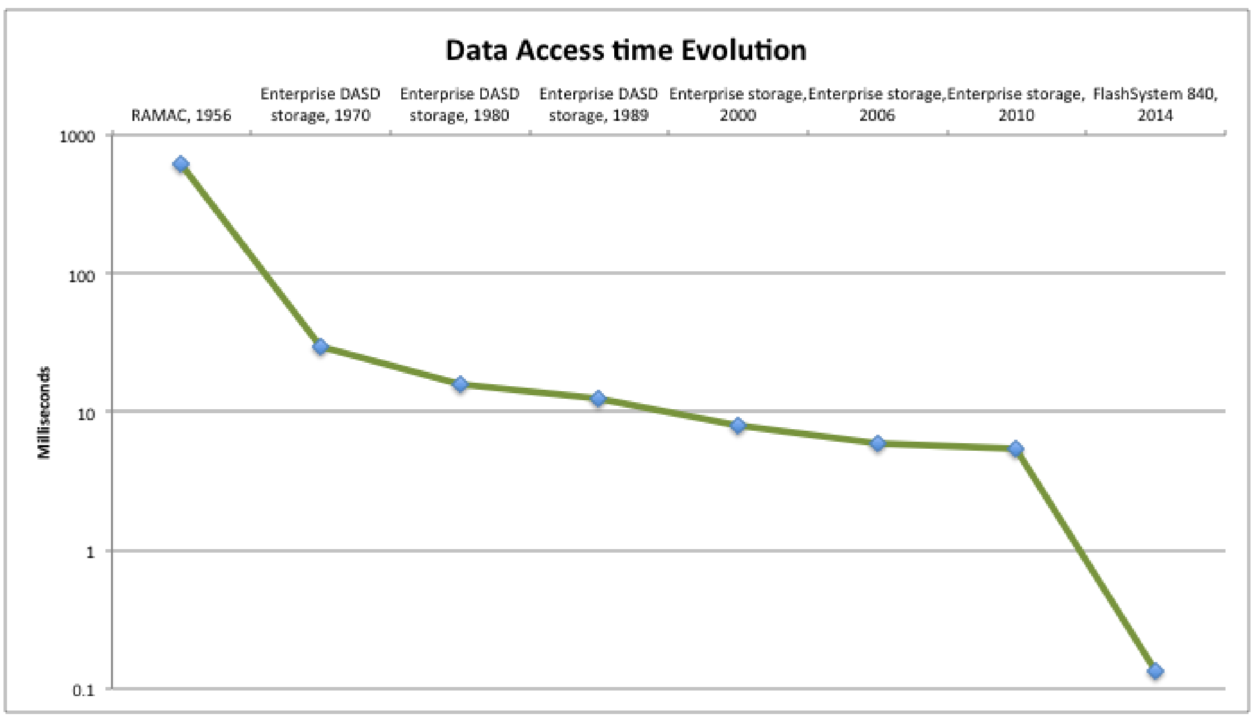

Figure 7 – Data access time evolution

IOPS Evolution

Another key metric is the maximum IOPS the system can serve. It is a measure of work units that can be served in parallel to users, and it's a measure of OLTP application performance. Figure 8 shows the evolution of this metric normalized per Rack Unit (RU). Note the massive improvement the high performance solid state array delivers over traditional hybrid enterprise disk systems.

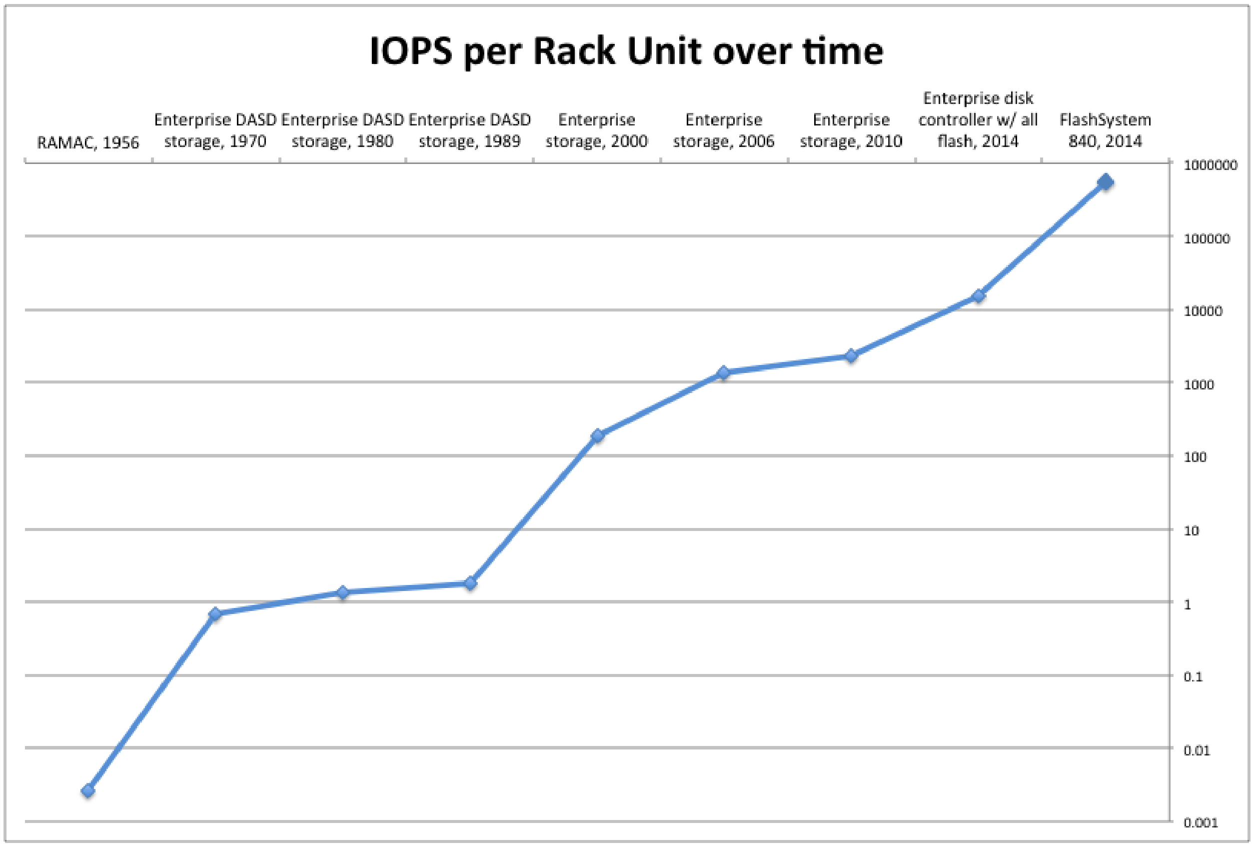

Figure 8 – IOPS per Rack Unit over time

The Value Proposition of FlashSystem

IBM FlashSystem 840 delivers breakthrough performance compared to other SSA products on the market today. FlashSystem 840 delivers super low response time with IBM MicroLatency™ technology and massive IOPS per rack unit. The FlashSystem design excels in real world mixed workloads at high demand rates, at high capacity utilization, and with long-term sustainable rates, well beyond just the initial burst performance that most systems can achieve.

Figure 9 shows a response time vs. IOPS range of random read/write (r/w) workloads. The blue line on the bottom is 100% random read, while the yellow line on the left is 100% random write. The lines in between are varying percentages of reads versus writes. While the end points illustrate the maximum pure read or pure write performance, it is the mixed r/w lines in between that are most interesting,

© Copyright IBM Corp., 2014

because customer workloads are invariably mixed I/O. This graph illustrates how well FlashSystem design can satisfy high I/O rates while keeping response time well under 1 millisecond. It is this long flat curve that makes FlashSystem an ideal storage solution for a wide range of real-world customer workloads.

Most impressive is that this performance, high as it is, represents the worst case scenario: It is sustainable while full garbage collection is taking place in the background, because the system has been purposely stressed by enough write I/O to bring it to this state.

Furthermore, the showcased performance represents stable and consistent behavior: Each performance curve is independent of the state of the system just preceding it; whether there was relatively higher write activity preceding it or relatively lower write activity just preceding it, the performance remains consistent. This is a qualitative difference that matters to customers, because it ensures predictable performance. Other vendors are challenged to provide this high level of performance consistency.

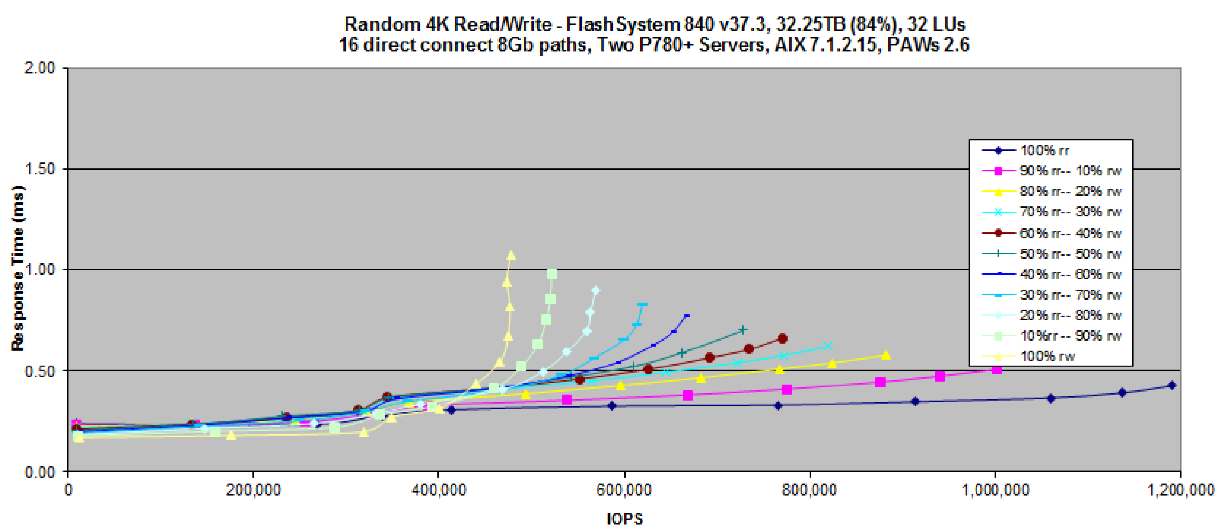

Figure 9 - FlashSystem 840 random R/W performance

Figure 10 shows the massive bandwidth the system delivers for both reads and writes.

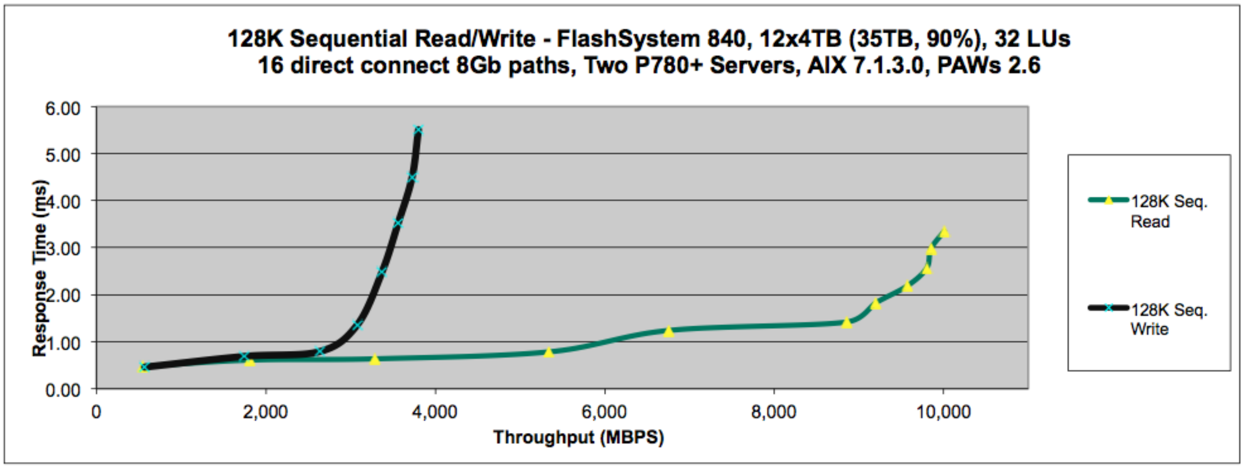

Figure 10 - FlashSystem 840 sequential R/W performance

Figure 11 shows sequential write performance over time. Note that the system delivers consistently high bandwidth even over extended operation. This represents the long-term sustainable bandwidth of the system at a typical state, after a variety of random reads and writes have taken place over the previous 24 hours. It showcases dependable, consistently high performance, not the freshly formatted bragging rights performance that competitors often showcase in marketing materials.

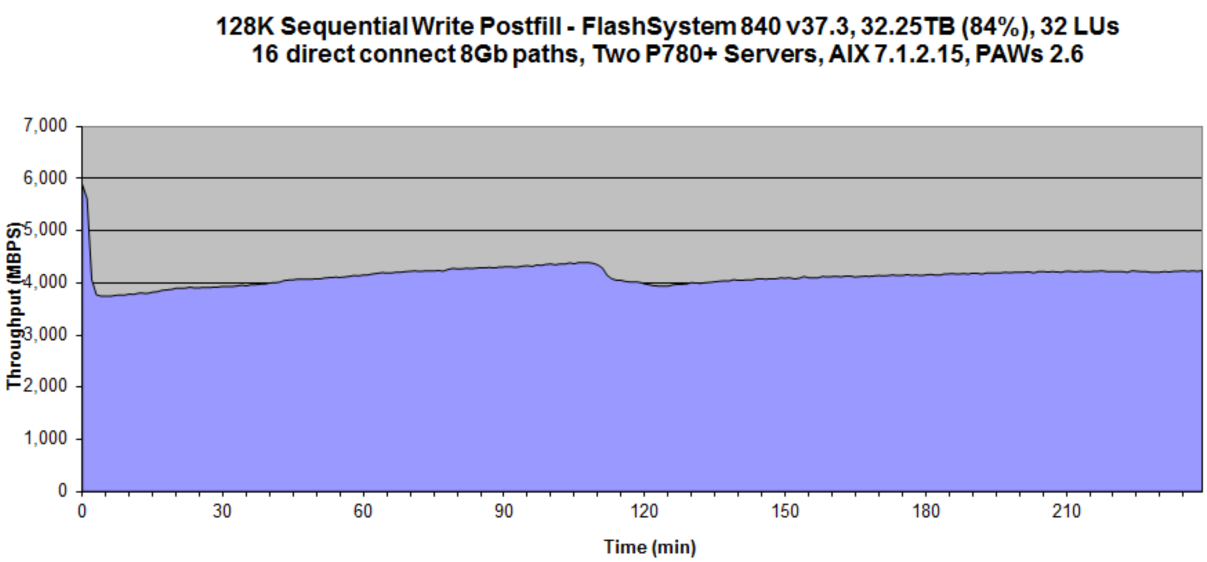

Figure 11 - FlashSystem 840 sequential write performance over time

© Copyright IBM Corp., 2014

4. Power Systems and FlashSystem Synergy

As illustrated above, Power Systems servers and FlashSystem storage are designed for outstanding performance and scalability. POWER technology delivers the highest performing core in the industry, up to 2x faster[iv] than the closest competitor. FlashSystem technology delivers consistent and predictable performance with super low latency (see Figure 9) where it matters most for real world workloads with extremely high throughput. This section illustrates the complementary attributes of each technology.

Data Access Optimization

Information technology exists solely for the purpose of processing data, the fundamental unit of value driving business today. When the components of a solution complement each other in the processing of data, optimal efficiency and performance are achieved. Time to value is a competitive advantage, so it pays to exploit complementary technologies whenever possible.

A foundational aspect of computers is how they organize and manage memory. Page size is determined by the processor architecture. Modern architectures are commonly based on a 4KB page size and/or multiples of 4KB. This is true for POWER, x86, and SPARC.

It follows then that for maximum efficiency, the OS, the middleware and applications, and the storage should also be designed to complement the common unit of a memory page. It is no wonder then that 4KB is such a popular unit of data transfer across the industry.

All Power Systems operating systems operate natively on a 4KB page or multiple thereof. FlashSystem technology is also optimized around 4KB data packets and multiples thereof. Both Power Systems and FlashSystem can handle data transfer sizes outside of this, although sometimes with less efficiency.

Response Time

We've all been on the phone with a customer service representative who's apologized that "their computer is going slow." But, beyond this obvious example, what are the ramifications to business of poor application response time? Consider the following scenarios:

- A bank runs a batch job that calculates accrued interest for customer accounts. It is critical for the job to finish within a nighttime window of four hours; exceeding this window risks presenting incorrect account balances to customers the next day.
- A large retailer uses a planning application to determine when to reorder items whose inventory level falls below a certain threshold. Poor response time of this planning application exposes the retailer to the risk of depleted inventory and lost sales, or conversely, to overly-conservative inventory practices that tie up working capital and lead to excessive stock that must be sold later at a deep discount.
- An electronics supplier uses an EDI-based application to accept incoming orders. Poor response time leads to timeouts, and when orders can't be accepted reliably, customer defections.

© Copyright IBM Corp., 2014

Clearly, good application response time is critical across a spectrum of business processes. This is one reason why Power Systems are such a popular server platform. The POWER core is designed to maximize cycle efficiency. Among other things, optimization entails high performance floating point and integer arithmetic logic units and high speed / high bandwidth access to the L1/ L2/L3 caches and main memory. Furthermore, the system is designed to aggressively allocate as many CPU cycles as can be consumed by an application. The benefit of this design is that compute-hungry applications can run faster on Power Systems than on less aggressively optimized architectures. OLTP is an example where Power Systems excels in delivering consistent low application response time along with massive scalability. The key to achieving the lowest application response time is to keep the cores busy doing productive computational work. Processor cycles expended waiting for data ultimately add latency. It follows then that in the case of OLTP the system will deliver the lowest application response time when the entire database resides in main memory. This is often impractical due to the size of the database being larger than what can be contained in memory. In this common situation disk is used to store the database. Accessing disk storage introduces latency, which impacts processor cycle efficiency and ultimately application response time. Access to processor cache and main memory happens in nanoseconds. Disk latency is typically in the 5-10 millisecond range, an eternity relative to cache and memory access. FlashSystem technology with IBM MicroLatency reduces I/O latency to microseconds, conservatively 10x-30x faster compared to disk. When I/O wait time is a significant portion of your overall application response time, FlashSystem technology can unleash your system to perform at its full potential.

Figure 12 illustrates the benefit to application response time and efficiency that are conveyed by the vastly improved latency of IBM FlashSystem technology. Application processing time is typically made up of two alternating components: 1) Time spent waiting for I/O, which is the time required to access data on storage; and 2) Time spent processing the data on the server's CPU. The column on the left shows how time is divided between these two components when traditional spinning disk technology is used. Note how it is only possible to access and process one set of data in 5.2 milliseconds, because the time is dominated by the slow response time of spinning disk. The column on the right shows how time is spent when IBM FlashSystem technology is employed. There, in 5.2 milliseconds, it is possible to access and process 13 sets of data, because the latency of the FlashSystem is so much lower. In addition to reduced application response time, employing FlashSystem technology also results in significantly improved utilization of CPU resources. Note, when spinning disk is used the application only utilizes the CPU around 4% of the time; with FlashSystem technology, the task spends 50% of the time running on the CPU – a thirteen-fold improvement. Besides getting more work done, FlashSystem made the whole solution more efficient. This can translate into reduction of software licensing expense and even into server capital reductions by enabling the processing of more data with fewer cores.

© Copyright IBM Corp., 2014

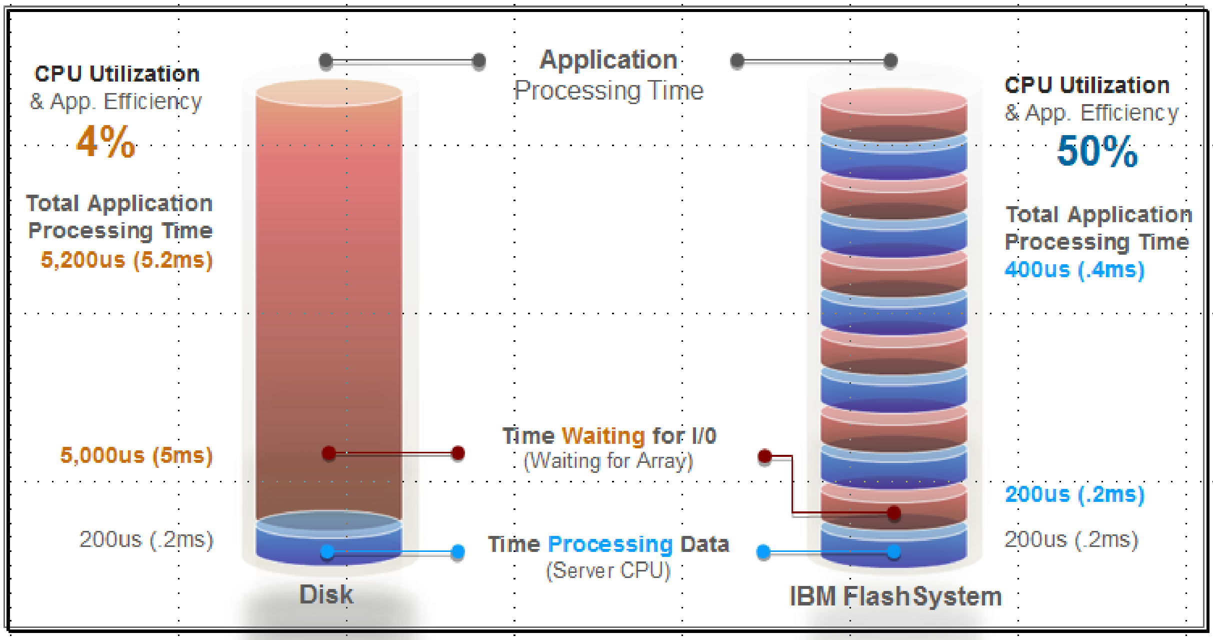

Figure 12 - Benefits of IBM MicroLatency

Bandwidth

The amount of data created by organizations is growing exponentially. For example, it is estimated that in 2012, over 2.8 zettabytes (1 zettabyte = 1 billion terabytes) of data were created; that amount is expected to double by 2015[v]. Rapidly-growing businesses in the areas of content streaming and social media serve vast amounts of data to millions of users simultaneously. Retailers analyze millions of transactions to derive the information needed to enhance individual customer relationships. With this in mind, it is clear that servers and storage need to be capable of sustaining high bandwidth; that is, they need to be able to move and process massive amounts of data quickly. Power Systems and FlashSystem both excel in bandwidth capability.

As depicted in Figure 4, the POWER8 processor supports peak I/O bandwidth of up to 96 GB per second, while the symmetric multi-threading capabilities of POWER8 allow up to eight software threads to be simultaneously executed on each core. This latter capability allows a larger pool of software threads to be driving I/O operations simultaneously.

AIX®, VIOS, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, Ubuntu Server, and IBM i all inherently support multi-path I/O to FlashSystem storage without any vendor-specific drivers or modules. Multi-path I/O allows operations to FlashSystem to be automatically spread over multiple SAN (storage area network) paths and adapter ports, which enhances performance while also increasing availability by automatically failing over from a bad path to a good path in the event of a SAN fault.

© Copyright IBM Corp., 2014

AIX, VIOS, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, and Ubuntu Server all support using 8 Gb and 16 Gb Fibre Channel adapters to SAN-attach FlashSystem 840 arrays.   For the recently-announced POWER8-based servers, the number of supported Fibre Channel connections ranges from 20 on the Power Systems S812L, to over 192 connections on the Power Systems E870/E880, when PCIe Gen3 I/O Expansion Drawers are attached. Thus, especially with POWER8, Power Systems have the Fibre Channel connectivity required to exploit the bandwidth and IOPS capabilities of FlashSystem technology.

Throughput

Throughput represents the quantity or amount of work the solution is able to complete in a given amount of time. There are many ways to look at it: transactions per second, number of concurrent users, and input/output operations per second. Throughput is also a measure of the scalability of a solution or its component parts. As illustrated earlier, Power Systems servers and FlashSystem storage each offer amazing throughput potential. Pairing these technologies together results in an extremely efficient and massively scalable system.   In addition to material presented earlier, Power Systems operating systems have been optimized for parallelism as described below.

AIX, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, Ubuntu Server, and IBM i all allow multiple threads to read and write the same file simultaneously. Referred to as concurrent I/O, this feature eliminates a bottleneck at the operating system level, allowing large numbers of mixed read and write I/Os directed to the same file to be concurrently executed directly against the storage. This is particularly important for database OLTP workloads which can have a large number of users simultaneously executing read and update transactions against the database.

AIX, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, Ubuntu Server, and IBM i support asynchronous I/O, which provides a means for a relatively small number of application software threads to initiate a large number of independent I/O operations without blocking. Through this mechanism, a small number of application threads can maintain a large number of outstanding I/Os, which in turn exploits the extreme scalability of FlashSystem technology. AIX, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, and Ubuntu Server driver modules support initiating and maintaining a large pool of active (i.e. not yet completed) I/O operations. For example, the Fibre Channel adapter driver in AIX can support a pool of up to 4,096 Fibre Channel commands per adapter port, while the AIX hdisk driver supports up to 256 outstanding I/O operations per hdisk. By comparison, FlashSystem 840 can simultaneously execute up to 16,384 I/O operations. The ability to maintain a large number of active I/O operations is crucial to being able to process large amounts of data rapidly, which is a requirement for high-volume transaction-processing and business intelligence scenarios.

Virtualization

Virtualization provides benefits such as sharing and improved utilization of compute and storage resources, faster server provisioning and rapid workload deployment. The majority of Power Systems server deployments employ multiple virtual machines (VM). To facilitate sharing I/O resources between VMs, Power Systems uses a virtual I/O server (VIOS). VIOS includes a rich set of features designed and

© Copyright IBM Corp., 2014

optimized to virtualize and make effective use of SAN storage. For example, Live Partition Mobility (LPM) is a popular feature customers use to move partitions (or VMs) between systems without incurring application downtime. LPM demonstrates its best performance with SAN-attached storage. VIOS has also been optimized to add very little latency (on the order of 30 microseconds) when transferring data to and from SAN-based storage, by employing Logical Remote Direct Memory Access (LRDMA) and N-Port ID Virtualization (NPIV) capable adapters. As a result, FlashSystem operating as a SAN-attached storage device is a perfect complement to the capabilities of the Power Systems VIOS, enabling customers to not only transparently exploit heavily-used virtualization features, but to also benefit from the ultra-low latency that FlashSystem affords.

5. Application Acceleration Examples

The following examples illustrate the leverage that can be achieved by combining Power Systems and FlashSystem technologies:

- IBM InfoSphere Identity Insight is an advanced security analytics offering that helps organizations identify information about individuals and groups from multiple sources and formats. It provides actionable information to quickly discover opportunities or risk in areas such as customer service, fraud detection, and threat avoidance. Identity Insight processes massive amounts of small data fragments in order to establish relationships and connections. The faster the processing the quicker opportunities, risks, and threats can be identified and addressed. Faster processing also means more requests can be handled by the system in a given amount of time, which translates directly into a better end user experience. In a recent study, solution performance improved 110x with FlashSystem as the storage choice, unleashing the true potential of the Power Systems server. Read the white paper: https://ibm.biz/BdFMs2

- FIS is a large global provider of banking and payments technology. With a long history deeply rooted in the financial services sector, FIS serves more than 14,000 institutions in over 110 countries. FIS Profile is a premier real-time, ultra-scalable core banking system. It is a fully integrated solution that includes customer, banking product, account, and transactional requirements. Loan and deposit products are supported by an extensive inventory of configurable features that are shared across all banking lines of business. With unmatched production scalability and ultrahigh availability, Profile supports hundreds of institutions, ranging from de novo startups to top-tier global banks, making it the premier choice for organizations committed to a 24/365 always-on solution. In a recent study using FlashSystem on a Power Systems server instead of a traditional disk system, batch time was reduced by 40%, application response time was reduced by 38%, online backup time improved by 75% and overall system latency improved 100x. Read the white paper: https://ibm.biz/BdFMZr

- In a recently published paper: “IBM POWER8 and IBM FlashSystem accelerate Oracle Database,” IBM demonstrated the benefit FlashSystem provides over a traditional SAN disk storage solution. Using a 6,000 customer database the flash-based system was able to reach peak performance at a lower thread count: 144 versus the 256 threads in the SAN-based solution. Additionally, the peak transactions per second (TPS) was six times higher: 13,494.63 versus 2,263 TPS. At the peak TPS, the application response time was also one-tenth —10.4 milliseconds (ms)—that of the traditional SAN (112.48 ms) per transaction. The S824 and FlashSystem 840 were also tested with the 50,000 customer database workload. Because a smaller portion of the 50,000 customer database could be cached in memory, the I/O rate increased from 98,551 to 225,726 per second; however, because of the scalable design of the FlashSystem 840, the latency per I/O remained nearly the same, in the range of 380-390 microseconds. There is no traditional SAN array that will provide such stable latency in response to a nearly ten-fold increase in workload. Read the white paper: https://ibm.biz/BdEf7J

- Epic Systems Corporation develops applications for medical groups, hospitals, and other healthcare organizations; their integrated electronic medical record (EMR) software covers all aspects of healthcare operations, including clinical, billing, and revenue, in addition to patient record access. Epic software typically exhibits high-frequency, random storage accesses with stringent latency requirements, and IBM has been working with Epic to develop host-side and storage-side solutions to meet these requirements. Extensive testing has demonstrated that the combination of POWER8 servers and IBM FlashSystem storage more than meets the performance levels Epic recommends for the backend storage supporting its software implementations—at a cost point multiple times lower than other storage alternatives. For example, when running Epic's performance evaluation suite, the combination of the IBM Power Systems S824 and a FlashSystem V840 achieved average read latency 95% lower than Epic requirements. The latencies were 90% below those measured with a high-end spinning disk array and 40% lower than those measured with a high-end array populated with SSDs. These results are even more compelling considering that the V840 supports advanced storage functions such as thin provisioning, tiering, and copy services. Read the white paper: https://ibm.biz/BdEH56

- The IBM Data Engine for NoSQL-Power Systems Edition combines Power Systems server and FlashSystem technology to provide infrastructure for NoSQL databases. NoSQL (Not-only SQL) is a rapidly growing area of database technology that is frequently employed in mobile, social, and analytics environments. The IBM Data Engine for NoSQL leverages the unique Coherent Accelerator Processor Interface (CAPI) feature of POWER8-based servers to radically reduce the software and memory overhead of running a NoSQL database, which in turn enables large-scale consolidation and reduction in total cost of ownership. Read the white paper: https://ibm.biz/BdEheR

6. Summary

IBM Power Systems servers and FlashSystem storage are leading-edge, complementary technologies, capable of processing massive amounts of data with unprecedented speed. With their highly multi-threaded design, Power Systems servers are capable of large-scale parallelism, managing thousands of independent tasks retrieving, processing, and updating data. FlashSystem, with its ability to process over 16,000 I/O operations simultaneously and over 1 million operations per second, is designed to satisfy the data storage requirements for such massively-parallel environments. POWER8-based servers are capable of driving 96 GB per second of I/O per socket, with over 192 Fibre-Channel connections (on E870/E880 with external I/O drawers), allowing them to exploit the bandwidth capability of several FlashSystem SSA's. Both Power Systems servers and FlashSystem storage are optimized for data unit sizes that are 4 KB in size or a multiple thereof. The powerful cores of POWER8-based servers, combined with FlashSystem MicroLatency, address latency-sensitive scenarios such as constrained batch-processing windows. To exploit FlashSystem capabilities, software platforms such as AIX, IBM i, Red Hat Enterprise Linux, SUSE Enterprise Linux, Ubuntu Server, and PowerVM®, provide highly-scalable I/O drivers, supporting features such as multi-path I/O, concurrent I/O, asynchronous I/O, and low-latency virtualization. Altogether, Power Systems servers and FlashSystem storage are a winning combination that delivers breakthrough business value for today's storage- and compute-intensive application environments.

Appendix - Complementary Characteristic Summary

<table>
<tr><th>Attribute</th><th>AIX</th><th>Linux</th><th>IBM i</th><th>VIOS</th><th>FlashSystem 840</th></tr>
<tr><td>I/O Bandwidth</td><td colspan="4">POWER8-based servers: Up to 96 GBps per socket</td><td>Up to 10 GBps sequential read</td></tr>
<tr><td>I/O Attachability</td><td colspan="4">Power S812L: Up to 20 Fibre-Channel connections

Power S814: Up to 24 Fibre-Channel connections

Power S822(L): Up to 32 Fibre-Channel connections

Power S824: Up to 40 Fibre-Channel connections

Power E870 and E880: Up to 64 Fibre-Channel connections in internal slots in two-building-block systems; up to 192 connections in PCIe Gen3 I/O Expansion Drawer</td><td>Up to eight 16 Gb Fibre-Channel connections

Up to sixteen 8 Gb Fibre-Channel connections</td></tr>
<tr><td>Concurrent I/O Operations</td><td>256 per hdisk; 4,096 per FC adapter port</td><td>Emulex adapters: Up to 512 per LUN; 8192 per adapter port;

Qlogic adapters: Up to 65,536 per adapter</td><td>1024 per adapter port</td><td>256 per vFC adapter driver; 512 per vSCSI adapter, less three per LUN for error recovery and two used by adapter)</td><td>16,384 per FlashSystem 840</td></tr>
</table>

Multi-path I/O	Supported	Supported	Supported	Supported	Supported
Concurrent File I/O	Supported	Supported	Supported	Supported – implemented by O.S.	Supported – implemented by O.S.
Asynchronous I/O	Supported	Supported	Supported	Supported – implemented by O.S.	Supported – implemented by O.S.

Footnotes and references

[i] IBM Power System S824 on the two-tier SAP SD standard application benchmark running SAP enhancement package 5 for the SAP ERP 6.0 application; 4 processors / 24 cores / 96 threads, POWER8; 3.52GHz, 512 GB memory, 21,212 SD benchmark users, running AIX® 7.1 and DB2® 10.5, dialog response: 0.98 seconds, line items/hour: 2,317,330, dialog steps/hour: 6.952,000 SAPS: 115,870 database response time (dialog/update): 0.011 sec / 0.019sec, CPU utilization: 99%, Certification #: 2014016. Source: http://www.sap.com/benchmark. (1.1) Fujitsu RX300 S8 on the two-tier SAP SD standard application benchmark running SAP enhancement package 5 for the SAP ERP 6.0 application; 2 processors / 24 cores / 48 threads. Intel Xeon E5-2697 processor 2.70 GHz, 256 GB memory, 10.240 SD benchmark users, running Windows Server 2012 SE and SQL Server 2012, Certification #: 2013024

[ii] IBM Power S824 (3.5 GHz, 24 core) results. PowerEdge R920 (Intel Xeon E7-8893 v2, 3.40 GHz) results. Source: http://www.spec.org

[iii] IBM WebSphere Application Server V8.5.5.2 and DB2 10.5 on IBM Power S824 result of 22,543.34 published on Apr 22, 2014. Oracle Weblogic Server Standard Edition Release 12.1.2 and Oracle Database 12c on Oracle Sun Server X4-2 result of 11,259.88 published on Sep 23, 2013. Source: http://www.spec.org

[iv] IBM Power System S824 on the two-tier SAP SD standard application benchmark running SAP enhancement package 5 for the SAP ERP 6.0 application; 4 processors / 24 cores / 96 threads, POWER8; 3.52GHz, 512 GB memory, 21,212 SD benchmark users, running AIX® 7.1 and DB2® 10.5, dialog response: 0.98 seconds, line items/hour: 2,317,330, dialog steps/hour: 6.952,000 SAPS: 115,870 database response time (dialog/update): 0.011 sec / 0.019sec, CPU utilization: 99%, Certification #: 2014016. Source: http://www.sap.com/benchmark. (1.1) Fujitsu RX300 S8 on the two-tier SAP SD standard application benchmark running SAP enhancement package 5 for the SAP ERP 6.0 application; 2 processors / 24 cores / 48 threads. Intel Xeon E5-2697 processor 2.70 GHz, 256 GB memory, 10.240 SD benchmark users, running Windows Server 2012 SE and SQL Server 2012, Certification #: 2013024

[v] Tucker, P. (May 7, 2013). Has Big Data Made Anonymity Impossible? Retrieved from http://www.technologyreview.com/news/514351/has-big-data-made-anonymity-impossible/

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both: AIX®, DB2®, Easy Tier®, FlashSystem™, IBM®, MicroLatency™, POWER®, Power Systems™, POWER8™, and PowerVM®.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.