

IBM DS8880/DS8880F – HPFE Gen2 storage system performance

Silverton Consulting, Inc. StorInt™ Briefing



Introduction

Better IO performance, more IO/sec and more IO bandwidth are never-ending aspirations for most data centers today. Since the solid-state drive (SSD) or flash revolution in enterprise storage, IO/sec performance has increased significantly, and response times are much lower (in the sub-millisecond range). However, IO bandwidth has seen little improvement.

In fact, it's only over the last two years that all-flash systems have begun to break sequential IO bandwidth records in storage benchmarks. Prior to 2015, all-disk systems were still top performers in sequential read-and-write IO performance.

Recently, IBM introduced a new line of High-Performance Flash Enclosure (HPFE Gen2) storage modules for their IBM DS8880F all-flash and IBM DS8880 hybrid storage systems, which have supplied a noteworthy boost in both IO/sec and, more importantly, IO bandwidth. Many of the new HPFE Gen2 hardware changes specifically address IO bandwidth improvements.

Industries and applications that need high bandwidth

A number of industries, verticals and applications can benefit from higher sequential IO bandwidth. For example:

- **Banking, securities and other financial services** – these organizations need to provide ever faster fraud detection, stock analysis and tick/trend analysis. For example, on Black Friday (November 25) in 2016, online purchases added up to ~\$3.3 billion.¹ Ensuring the validity of credit card transactions is a significant financial undertaking that requires banks and credit card servicing organizations to monitor payment transaction logs in real time. Similar growth in securities activity over the past decade has required other financial institutions to monitor stock price changes and international/national news feeds in real time.
- **Hyper-scalar and mobile application services** – these companies need increased IO bandwidth to serve more and more people as they come online. For example, Facebook users now number 1.8 billion,² and other social media apps are trying to catch up. Downloading and serving photo, audio and video media from a significant and growing portion of the population often consumes all available IO bandwidth, and more bandwidth will be needed in the future to keep up with user growth.

¹See <https://techcrunch.com/2016/11/25/black-friday-online-sales-to-hit-a-record-breaking-3-billion-over-1-billion-from-mobile/> as of February 21, 2017.

² See <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/> as of February 7, 2017.

- **Government, science and healthcare** – these organizations have huge data problems, as the amount of experimental data seems to increase exponentially with each new project. For example, the Square Kilometer Array will transfer data at ~16 Tbps, which is 10,000 times faster than any other telescope.³ Healthcare imaging IO bandwidth needs are also increasing, as the costs of digital X-rays, MRIs/CAT scans, DNA sequencing,⁴ etc., all decrease over time while image resolution improves.

Increasing flash storage bandwidth

The main limit to flash/SSD storage bandwidth has been data access speed and lack of parallel active data paths to the NAND storage chips. Each additional data path requires more software and hardware, resulting in higher cost and slower time to market. Further, industry trends to increase the number of bits stored per NAND chip (increasing storage capacity) also make it harder to add more data paths/bits.

Major storage vendors like IBM can take actions both inside and outside the flash storage module to improve IO bandwidth. IBM, for one, has taken several steps to improve bandwidth performance with their HPFE Gen2 and DS8880F/DS8880 storage systems.



DS8880F/DS8880 - HPFE Gen2 storage

The IBM DS8880F all-flash and IBM DS8880 hybrid disk-flash enterprise storage systems offer substantial high availability and disaster recovery capabilities while also providing excellent IO performance for database-intensive applications in both open and z/OS environments. The current all-flash DS8880F and hybrid DS8880 use HPFE Gen2 for NAND storage.

The DS8880F comes in three versions:

- **Business-class DS8884F** uses two IBM Power Systems S822 with 6-core POWER8 processors per S822, 256GB of DRAM, 32 FC or FICON ports and from 6.4TB to 153.6TB of flash capacity.
- **Enterprise-class DS8886F** uses two IBM Power Systems S824 with 24-core POWER8 processors per S824, 2TB of DRAM, 128 FC or FICON ports and from 6.4TB to 614.4TB of flash capacity.

³ See https://www.skatelescope.org/wp-content/uploads/2011/05/SKA_Factsheet-for-Industry_May2011_web.pdf as of February 7, 2017.

⁴ See <http://www.cnn.com/2015/12/10/unlocking-my-genome-was-it-worth-it.html> as of February 7, 2017.

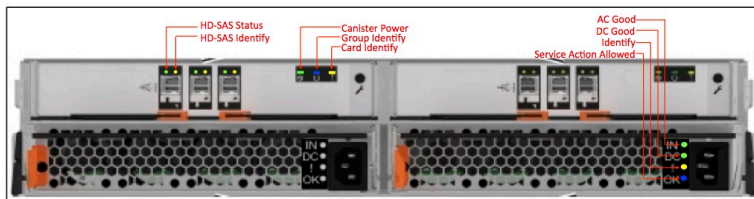
- **Analytics-class DS8888F** uses two IBM Power Systems E850 with 48-core POWER8 processors per E850, 2TB of DRAM, 128 FC or FICON ports and from 6.4TB to 1.2PB of flash capacity.

The HPFE Gen2 storage module is a 4U enclosure that can hold 16, 32 or 48 2.5" flash storage cards. HPFE Gen2 flash cards come with 400 GB, 800 GB, 1.6 TB or 3.2 TB of NAND storage. With 48 3.2 TB flash cards, a single HPFE Gen2 storage module can support 153.6 TB of flash storage.



By making use of the higher density (3.2 TB) flash storage cards available today, IBM can offer customers more flash capacity per floor tile and a lower cost (\$/GB) for HPFE Gen2 flash storage.

To increase IO/sec and IO bandwidth performance, IBM upgraded the connection between the DS8880F and the HPFE Gen2 from PCIe Gen 2 to PCIe Gen 3. PCIe Gen 3 provides twice the bandwidth per path than PCIe Gen 2 (1000 MB/s vs. 500 MB/s).



card.

In addition, the HPFE Gen2 RAID controller now uses eight 6Gbps SAS connections to connect to flash card storage, increasing the IO/sec and IO bandwidth available to each flash storage

The RAID controller field-programmable gate array (FPGA) in the prior-generation HPFE has also been upgraded to an application-specific integrated circuit (ASIC), which has improved both IO/sec and IO bandwidth performance as well as added RAID 6 to the previously supported RAID 5, 10 and 0. Now RAID 6 can be used to further optimize the performance/capacity of the HPFE Gen2 for different application workloads.

DS8880F - HPFE Gen2 performance results

The Storage Performance Council (SPC)-1 benchmark is an industry-standard, independently audited block storage benchmark used to compare storage system IO performance. SPC-1 benchmark results were released last year for a DS8888 with HPFE Gen2 storage.⁵

⁵ See http://www.storageperformance.org/results/results_spc1_v3/spc1_v3_active-a31002 as of February 9, 2017.

The DS8888 – HPFE Gen2 combination reached a maximum of 1,500,187 SPC-1 IOPS™, with an SPC-1 Overall Response Time of 0.336 ms for the benchmark. The configuration used for the benchmark was an IBM DS8888 storage system with eight HPFE Gen2 storage modules populated with 320 400 GB flash cards. The host connection employed 32 16Gbps FC links. Benchmark results are shown in Figure 1.

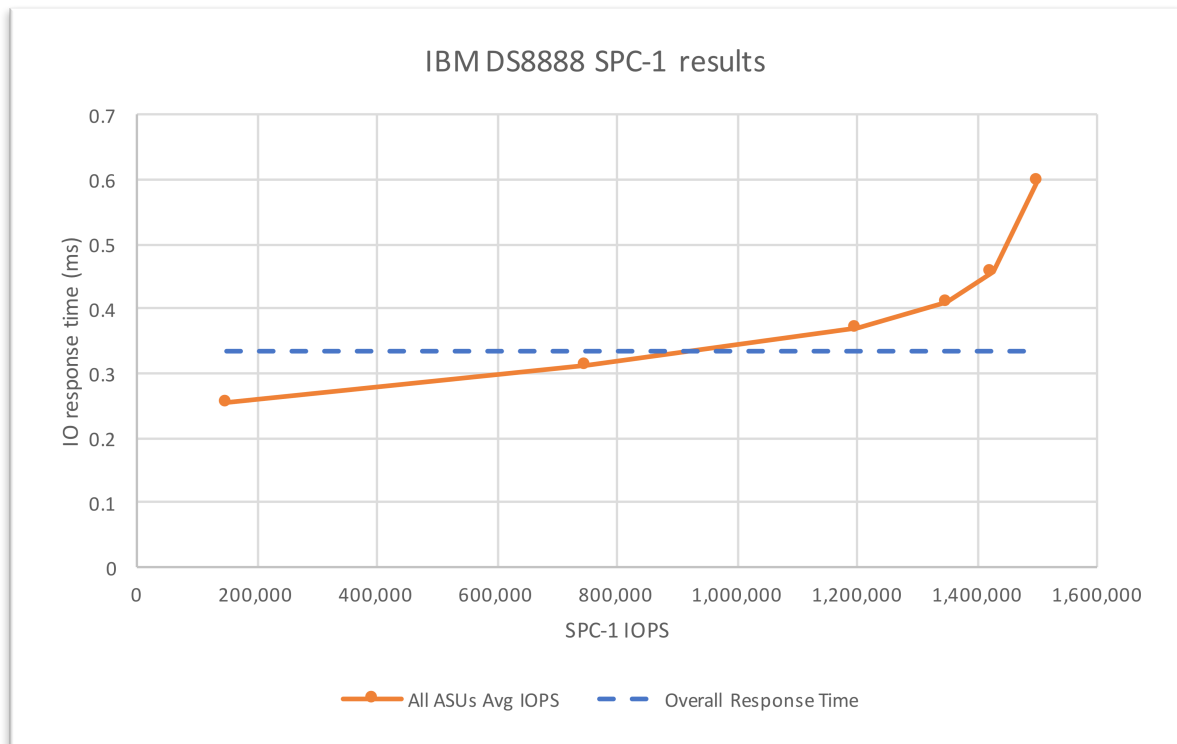


Figure 1 IBM DS8888/HPFE Gen2 SPC-1 results

Although the discussions above don't focus on IO/sec enhancements, most bandwidth changes also improved IO/sec performance. In Figure 1, at maximum IO/sec rate, the DS8888 – HPFE Gen2 system completed its IO at a response time of under 0.6 ms.

The SPC-1™ Overall Response Time is a metric derived from the average IO completion time for a number of benchmark test sequences that range from 10% to 100% of the maximum IO/sec achieved. An SPC-1 Overall Response Time of 0.336 ms means that most IO activity for these benchmark runs completed in under 0.336 ms.

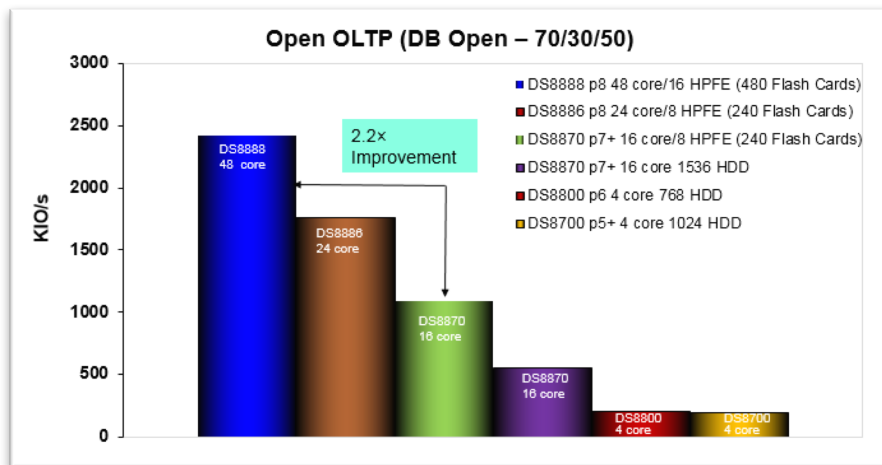


Figure 2 compares current- and previous-generation DS8880 – HPFE storage subsystem performance. The current-generation DS8888 with HPFE Gen2 provides ~2.2X the IO/sec than the DS8870 with HPFE Gen1. (Note that this is an open system database OLTP

Figure 2 DB Open OLTP performance comparisons

workload with a 70:30 read:write ratio and a 50% read cache hit rate.)⁶

Figures 3 and 4 offer similar comparisons of sequential IO bandwidth.

Figure 3 shows that the sequential read IO bandwidth (GB/s) of the current-generation DS8888 with HPFE Gen2 improved 2.6X over the previous-generation DS8870 with HPFE Gen1 storage.⁷

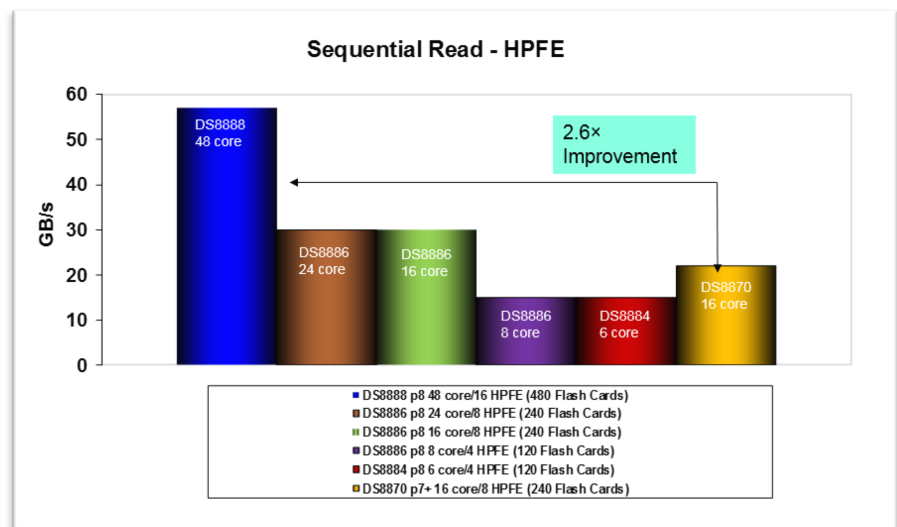


Figure 3 Maximum read IO bandwidth comparisons

⁶ The data in Figure 2 was supplied by IBM. The data comes from internal IBM performance testing with another online transaction processing (OLTP)-like benchmark.

⁷ The data in Figure 3 was supplied by IBM.

As shown in Figure 4, the current-generation DS8888 and HPFE Gen2 improved sequential write bandwidth 2.4X over the previous-generation DS8870 and HPFE Gen1 storage.⁸

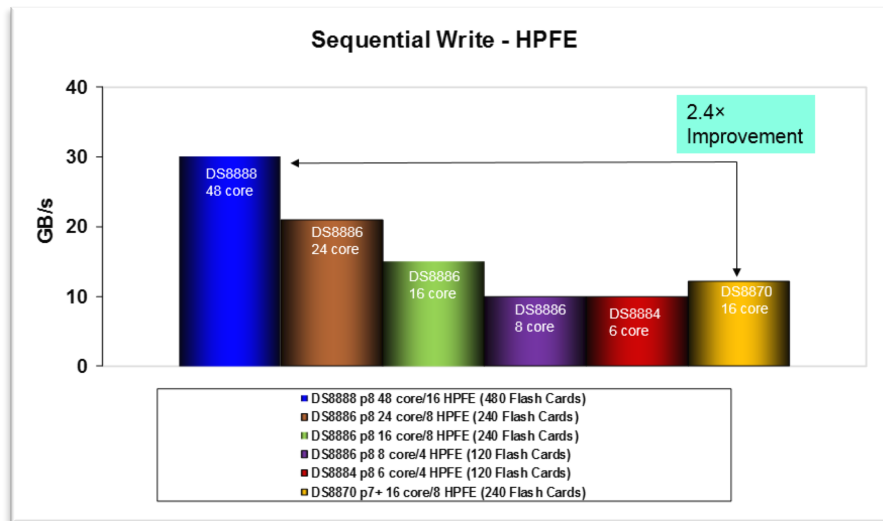


Figure 4 Maximum write IO bandwidth comparisons

Summary

The new HPFE Gen2 has significantly elevated the IO performance of the IBM DS8880F all-flash and IBM DS8880 hybrid family of storage systems. Even though IO bandwidth improvements are rare in the industry today, the new HPFE Gen2 – DS8888F/DS8888 storage combination has increased IO bandwidth by 2.4X to 2.6X over prior-generation IBM storage for write and read sequential IO respectively. This same storage combination has also improved open system OLTP database IO performance by 2.2X.

In addition, the latest HPFE Gen2's higher density flash storage cards, together with DS8880/DS8880F storage, offer improved flash storage density at an economical (\$/GB) cost.

If you need high sequential IO bandwidth, better DB OLTP IO/sec or more flash storage for your infrastructure dollar, you can't go wrong with the IBM DS8880F/DS8880 family of storage systems with new HPFE Gen2 flash enclosures.

Silverton Consulting, Inc., is a U.S.-based Storage, Strategy & Systems consulting firm offering products and services to the data storage community.



⁸ The data in Figure 4 was supplied by IBM.

Disclaimer: This document was developed with International Business Machines Corporation (IBM) funding. Although the document uses publicly available material from various sources, including IBM, it does not necessarily reflect the positions of such sources on the issues addressed.