IBM FlashSystem® 900 Enhancements

Silverton Consulting, Inc. StorInt™ Briefing

Introduction

IBM recently enhanced its flagship FlashSystem 900 with both new hardware and new software to bring more value to its customers. IBM FlashSystem® 900 has always been at the forefront of NAND technology adoption and since the beginning has offered extremely low latency response times with high IO rate performance. The most recent FlashSystem 900 enhancements continue this tradition of excellence, using next-generation NAND technology and adding other features that provide more economical storage to users without compromising performance.

IBM FlashSystem 900 enhancements

The new FlashSystem 900 adds a number of hardware and software functionality enhancements:

- **3D triple-level cell (TLC) NAND** to boost system capacity using denser (GB/sq.in.), more economical ($/GB) flash storage; and
- **Hardware-based inline data compression** to increase effective storage capacity for typical application data.

FlashSystem 900 now offers the following software enhancements:

- **Compresstimator and capacity visualization tools** to better estimate, monitor and manage FlashSystem 900 storage capacity;
- **IBM Security Key Lifecycle Manager (SKLM) support** to centralize key management and facilitate the adoption of FlashSystem 900 data-at-rest security;
- **New FlashSystem 900 dashboard** to give admins better visualization of performance and other characteristics;
- **New IBM service log collection process** to provide proactive service for FlashSystem 900 customers; and
- **IBM Spectrum Control™ support** to supply an on-prem and in-cloud single management UI for multi-vendor and multi-product storage infrastructure.

Capacity enhancements

NAND technology

Many of the new FlashSystem 900 capacity improvements come from next-generation NAND technology adoption. The previous generation of FlashSystem 900 used 2D multi-level (2-bit) cell (MLC) NAND technology. The new generation now adds 3D TLC (3-bit) NAND technology.

3D NAND and TLC technology provide an immediate increase in raw NAND storage capacity for the same chip footprint. As such, each FlashSystem 900 MicroLatency® flash module now stores more data. [A similar transition occurred when FlashSystem 900 changed from 1-bit single-level cell (SLC) to MLC NAND.]

The downsides to moving to TLC NAND technology include the potential for decreased endurance (writes per cell) and the possibility of increased write and read times. However, the new FlashSystem 900 TLC flash modules include IBM proprietary enhanced endurance technology and have been optimized for real-world mixed workloads, which eliminates the potential endurance degradation. Another solution to decreased endurance is to increase the amount of material in a TLC NAND cell.

Specifically, the change from 2D (or planar) to 3D (or vertical) NAND has increased the material used for each cell, which improves endurance. Just as a high-rise building can offer a larger number of apartments within a set footprint than a single-level building, 3D NAND uses multiple layers to stack many more (and larger) NAND cells, supplying more storage with better endurance within the same chip footprint.

Most NAND suppliers view 3D NAND as the next generation of technology, which can scale up in density for a long time to come without undue degradation in endurance. Thus, while current-generation 3D NAND uses 32 layers of NAND cells, NAND suppliers are already sampling 64-layer NAND, and 96-layer 3D NAND is on the horizon.

In addition, the prior-generation FlashSystem 900 MicroLatency modules only offered 1.2TB, 2.9TB or 5.7TB per module for a maximum usable capacity of 57TB per system. With the new 3D TLC NAND flash modules, FlashSystem 900 now offers 3.6TB, 8.5TB and 18TB flash modules. With the 3.6TB 3D TLC flash modules, the system supports up to 36.1TB of usable capacity; with the 8.5TB modules, the system supports up to 85.5TB of usable capacity; and with the 18TB modules, the system supports up to 180TB of usable capacity. As a result, customers gain (~3X) increased capacity flexibility in configuring new FlashSystem 900 solutions.

It is also important to note that FlashSystem 900 customers cannot mix different flash media types (MLC and TLC) or flash module capacity levels in the same system. That is, all the occupied FlashSystem 900 drive slots must have the same Micro-

Latency modules. In addition, older-generation FlashSystem 900s cannot be updated in the field with the new flash modules.

Hardware inline data compression

Hardware inline data compression is the other major FlashSystem 900 hardware enhancement that improves system capacity. *Inline* means compression is done during data transfer from the host to the flash module storage. For the new FlashSystem 900, hardware inline compression is done by the flash module hardware. *Hardware compression* means that the compression algorithm is executed by dedicated electronic componentry.

This hardware inline data compression eliminates redundant information inside a block of data being written to flash modules, improves flash endurance (as less data is written to new flash module NAND) and increases **effective (compressed) capacity** (customers can store more data per flash module).

The use of hardware-based inline data compression requires new terminology to describe system capacity:

- **Raw capacity** is the capacity of each flash module multiplied by the number of flash modules in a system;
- **Usable capacity** is the amount of user data that can be stored uncompressed on a system after data protection is applied to the storage;
- **Effective capacity** is the typical (average) amount of user data that can be stored on a system after data compression; and
- **Maximum reported capacity** is the maximum amount of the compressible user data stored on a system after data compression.

IBM sells new FlashSystem 900 systems on an effective capacity basis, which uses an average of 2:1 compression for the 3.6TB and 8.5TB flash modules and 1:1 compression for the 18TB flash module. As a result, the effective capacity is 72.2TB for the 3.2TB flash module, 171TB for the 8.5TB flash module and 180TB for the 18TB flash module. But with the 18TB flash module, if customers use compressible data, they could easily exceed 180TB in effective capacity.

While software data compression has been available for a decade or more, hardware data compression is uncommon in storage today due to the complexity and cost of providing data compression hardware. Indeed, IBM is the first company to offer hardware inline compression with 3D TLC NAND flash storage. Hardware data compression helps to minimize the latency or response time overhead in IO operations.

Software compression, on the other hand, always adds latency for read (decompression) operations. Depending on the (inline or post-process) implementation, software compression also adds latency for write (compression) operations.

- For **post-processed software compression**, systems must write the data uncompressed onto the backend and then later read the uncompressed data back in, compress it and write it back out to storage. Post-process software compression multiplies the number of IOs that take place for the same customer data and essentially triple processes each byte of data. It also adds to controller overhead for each data block write and further reduces flash write endurance.
- For **inline software compression**, systems can avoid the multiple writes and additional backend read but they still must add processing overhead to de-staging, which can easily become a bottleneck at high write IO rates.

Hardware compression eliminates all the read and write overhead inherent in software compression. Moreover, as data is never written on NAND in uncompressed format, hardware compression improves flash endurance.

Any kind of data compression works well for most applications with well-behaved compressible data but may not work as well for some applications with uncompressible data. The most common examples of uncompressible data are JPEG image files and MP3 audio files, which are already heavily compressed.

Furthermore, hardware inline compression is always on for the new FlashSystem 900. However, as discussed earlier, IBM assumes a 2:1 compression ratio for the smaller flash modules and a 1:1 compression ratio for 18TB flash modules. In other words, for 18TB modules, IBM assumes no compression.

To help customers make the most informed decisions and storage capacity estimates during planning, IBM offers a downloadable command line **Compresstimator** tool.[1] The tool runs on AIX, HP-UX, IBM I through VIOS, ESXi, Linux (RHEL, SUSE, Ubuntu and CentOS) and Windows Server hosts. Customers can use the tool to estimate the expected data compression ratio for FlashSystem 900 using their own application data. In fact, the tool can be run in the customer environment prior to ordering a FlashSystem 900 so customers can identify candidate applications for use with hardware compression.

[1] See http://www14.software.ibm.com/webapp/set2/sas/f/comprestimator/home.html.

Finally, hardware inline data compression on the new FlashSystem 900 comes at no additional cost beyond the cost of the new flash modules. For customers using the 3.6 and 8.5TB flash modules, this benefit decreases storage capacity cost ($/GB) even more than from the transition from 2D MLC to 3D TLC alone.

Data services enhancements

In addition to the capacity changes above, several other software and service enhancements to the new FlashSystem 900 are worth noting: IBM SKLM, FlashSystem 900 performance and capacity dashboard, IBM periodic collection of service information, and IBM Spectrum Control.

IBM SKLM is a software offering that centralizes and automates management of many security features in the enterprise today. In addition to local key management, the new generation FlashSystem 900 now supports IBM SKLM for data-at-rest encryption key management. IBM SKLM also offers OASIS Key Management Interoperability Protocol (KMIP), a standards-based security interface that provides centralized key management for non-IBM vendor security capabilities within the enterprise.

The FlashSystem 900 performance and capacity dashboard is a new display panel and ribbon that shows real-time performance and capacity metrics for FlashSystem 900s in the data center. Administrators can use the new dashboard to see real-time FlashSystem 900 IO performance and capacity consumption, as well as view system performance and capacity trends over time.

IBM periodic collection of service information is a new approach to supporting storage in the field. Previously, FlashSystem 900 would call IBM service to offload status and diagnostic information only when it detected a serious alert condition or failure, making the IBM service more reactive to events than proactive to trends. Now IBM service is also gathering status and diagnostic information constantly from all enabled systems deployed worldwide. Using this new information, IBM service can broaden its view of FlashSystem 900 system operations across the field and gain a better understanding of the health of any specific system. Combining this new data with sophisticated intelligent service analytics promises to improve FlashSystem 900 availability, reliability and performance.

IBM Spectrum Control is an on-prem or in-cloud, centralized, cross-vendor management, automation and analytics solution for storage infrastructure. IBM FlashSystem 900 adds support for Spectrum Control, which means admins can now manage all the data center's IBM and non-IBM storage infrastructure from one management console. Spectrum Control provides management and monitoring capabilities for storage systems and can be used to automate provisioning, improve tiering and manage replication.

IBM FlashSystem 900 IO performance

Although 3D TLC NAND in flash modules has similar or possibly equivalent endurance as planar MLC NAND, it can take a little longer to write and read data. Further, hardware compression may add a minimal amount of overhead to IO operations. On the other hand, flash module hardware compression can reduce the data that's read or written to NAND storage.

Nonetheless, at the latency levels of the previous-generation FlashSystem 900, 3D TLC NAND and hardware inline compression may add some overhead to write IO latency and no change to read overhead. (Results may vary depending on data compressibility.)

Even with the slight degradation in write IO, the new FlashSystem 900 retains the same data throughput and IO rate as the previous generation. And in some cases, data compression, by reading and writing less data, may actually help increase IOPS and data throughput.

Summary

FlashSystem 900 continues to rapidly adopt the latest NAND and other hardware, software and support technologies to improve system capabilities and lower customer costs, all while maintaining FlashSystem 900's lead in IO performance.

The shift to 3D TLC NAND storage capacity allows FlashSystem 900 storage customers to consolidate more applications and use larger data sets. The addition of hardware inline compression makes this increase in raw capacity even more appealing for customers with compressible data.

The new support for IBM SKLM, performance and capacity dashboards, periodic information collection and IBM Spectrum Control make it even easier for customers to integrate, automate and manage highly secure FlashSystem 900 in their data center environments. In the long run, periodic service data collection combined with planned cognitive service analytics should give IBM the information needed to improve system availability, reliability and performance for customers worldwide.

In the end, the higher flash module capacity, hardware inline compression and functional enhancements of the new FlashSystem 900 are great improvements to the premier performing all-flash array in the industry today.

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.