

Accelerating Data Economics: IBM FlashSystem* and Intel® QuickAssist Technology



Greater value from data assets

Conventional wisdom suggests that the rapid adoption of data reduction technologies within enterprise data centers is being driven by the collision of dramatic data growth and flat IT spending. But this isn't entirely true. In fact, extremely large data volumes can easily be handled by IBM* tape storage systems at a cost of around two-tenths of a cent per GB per month.¹

Rather, one of the real driving forces attracting enterprises to explore various data reduction strategies is the need to gain and maintain competitive advantage by deriving greater value from data assets. Simply storing data is not enough. Making faster, more informed business decisions; staying ahead of the competition; creating better customer experiences; implementing the best possible fraud protection and data security – these and many similar requirements are the drivers of the new cognitive era of business.

Because of their unmatched ability to help transform the burgeoning volume and velocity of data assets into business value, flash-based storage deployments are accelerating around the world. IBM FlashSystem* all-flash storage arrays are market leaders in

this IT industry sector. Flash storage economics benefit greatly from data reduction, but flash loses significant value when implementation of technologies such as data compression increases system response times (latency) and reduces inputs/outputs per second (IOPS) performance.

Intel® QuickAssist Technology™ offers an elegant solution with minimal trade-offs between performance and cost. It enables real-time data compression of data in motion that reduces the physical amount of storage capacity needed while also accelerating the speed at which data is compressed, transmitted, and stored. Together, FlashSystem and Intel QuickAssist Technology help the cognitive business derive the greatest value from data assets while reducing storage costs with minimal system performance impact.

The value of data reduction

Market research by IT analyst firm Enterprise Strategy Group (ESG) confirms the popularity of data reduction technology. As noted in Figure 1, data reduction ranks second on a list of storage “must have” features that include everything from high availability, through flash storage, to asynchronous replication.²

Originally, data reduction technologies such as data compression were developed to reduce the capacity requirements and costs of archived information, which meant the technologies could be applied when convenient and out of the data path. But more and more, enterprises are looking to take advantage of data reduction to deliver immediate storage savings for active production data sets. Even for these use cases, many compression solutions still do their work “post-process” and do not perform well in random input/output (I/O)-intensive primary storage environments. The best data compression solutions perform compression inline in real-time and provide excellent results for all I/O patterns, with minimal impact to storage latency and performance.

Data Compression

Data compression reduces the number of bits needed to represent data. Compressing data can save storage

capacity, speed file transfer, and decrease costs for storage hardware and network bandwidth.

Data compression is performed by applying formulas or algorithms that replace longer bit strings with shorter ones using a dictionary for the conversion between them. Compression can also be achieved by inserting a reference or pointer to a string of bits the compression engine has already seen. With the technology available today, very high compression rates can be achieved, depending on the type of data and application workload. Data compression provides the best results in traditional database and online transaction processing (OLTP) environments, as well as with many data warehouse and analytics workloads.

IBM FlashSystem* V9000

Building on decades of storage leadership, IBM offers a comprehensive

portfolio of integrated, flash-optimized storage solutions that can propel organizations into the future of IT and thus of business itself. These proven, easily-integrated flash storage platforms accelerate critical applications for faster decision making, and deliver enterprise-grade reliability and new cost efficiencies across the entire business environment for a faster return on investment. IBM flash storage solutions can provide enterprises with the storage performance needed to compete, innovate, and grow.

IBM FlashSystem V9000 offers the advantages of software-defined storage at the speed of flash memory. This all-flash storage platform combines the high performance, ultra-low latency, superior efficiency, and extreme reliability of IBM FlashCore* technology with a rich set of virtualization and storage features, including IBM Real-time Compression™, dynamic tiering, thin provisioning, data copy services, and high-availability configurations.

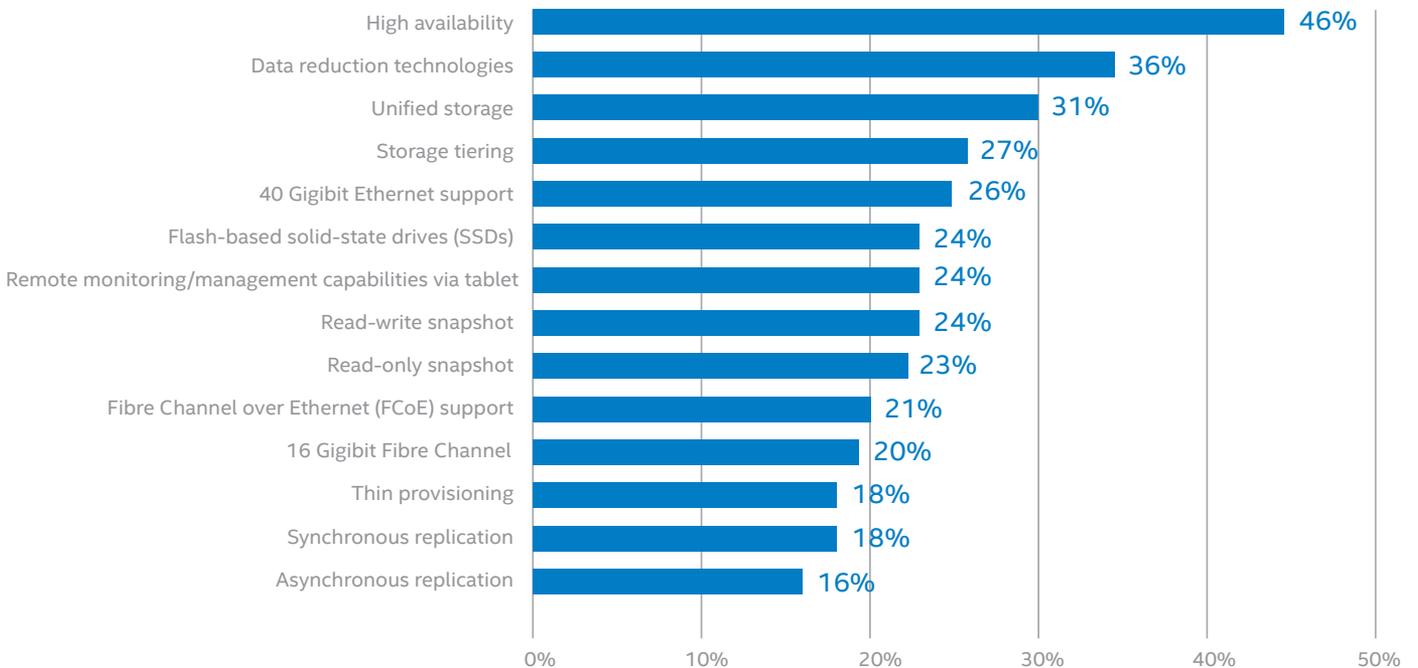


Figure 1. Must-have IT technologies.



FlashSystem V9000 is designed as a comprehensive storage solution for all active data sets. By accelerating the full range of applications and infrastructures, FlashSystem V9000 can help organizations reduce costs, increase revenue, and improve customer satisfaction. It includes powerful IBM Real-time Compression based on over 70 patents that can deliver flash for less than the cost of conventional enterprise storage.

FlashSystem V9000 can function as a feature-rich, software-defined storage layer that virtualizes all managed storage. In this capacity, it acts as the virtualization layer between the host and other external storage systems, providing flexibility and extending functionality to the virtualized external storage capacity. Up to 32PB of external storage can be managed by a single FlashSystem V9000 array, and because the storage is virtualized, volumes can be non-disruptively moved between external and internal storage capacity. This functionality enables very agile integration into existing storage environments with seamless data migration between FlashSystem V9000 and legacy storage systems.

Flash storage solutions such as FlashSystem V9000 reap great

benefits from data reduction due to the historically higher purchase price per unit of storage capacity for flash compared to other storage media. But data reduction also brings challenges, because unless done effectively, it can negatively impact storage performance, latency, and resource utilization.

IBM Random Access Compression Engine

The data compression technology embedded in IBM FlashSystem V9000 addresses all the requirements of primary storage data reduction, including performance. It does so by using a purpose-built technology called IBM Random Access Compression Engine (RACE) integrated with Intel QuickAssist Technology. RACE operates on active primary data and is executed as data is written to the backend storage media. This eliminates the need to reserve both time and storage space for post-process compression and maximizes the amount of data that can benefit from compression. The real-time data compression of data in motion produced by the RACE/Intel QuickAssist Technology combination dramatically cuts the cost of flash storage without sacrificing the benefits of IBM FlashSystem microsecond latency.

RACE technology is based on over 40 patents that actually are not directly related to data compression. Rather, they define how to make industry-standard Lempel-Ziv (LZ) compression of primary storage operate in real-time and allow random access. When a host sends a write request, it is acknowledged by the write cache of the system and then staged to the storage. As part of its staging, it passes through the compression engine and is then stored in compressed format onto the storage media. Writes are acknowledged immediately after they are received by the write cache, with compression occurring as part of the staging to internal or external physical storage.

Over the years, IBM has introduced a series of real-time compression algorithms and solutions used in a wide range of technologies:

- The LTO-DC algorithm used in IBM Linear Tape Open (LTO) tape drives
- The Streaming Lossless Data Compression (SLDC) algorithm used in IBM Enterprise-class TS1130 tape drives
- The Adaptive Lossless Data Compression (ALDC) used by the IBM Information Archive for its disk pool collections
- The RACE used inside IBM FlashSystem V9000 and other IBM storage systems.



Figure 2. IBM FlashSystem V9000

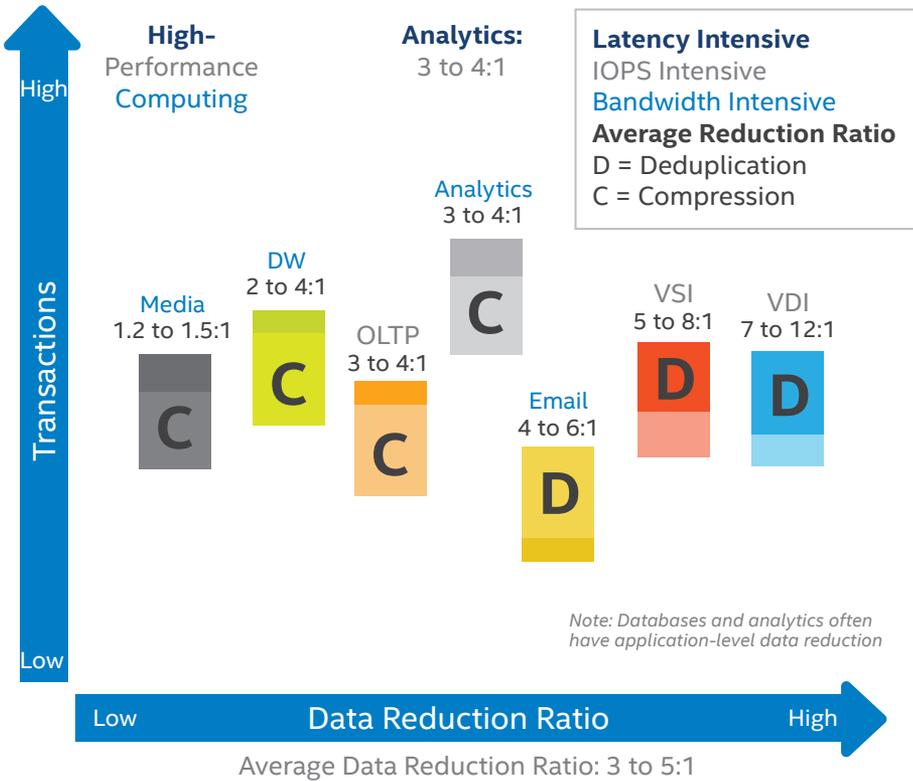


Figure 3. Varying benefits of data reduction

Certain data and file types benefit more than others from data compression, as illustrated in Figure 3. Most general-purpose data volumes compress well, such as home directories, CAD/CAM, oil and gas data, and log data. The expected compression ratios for these data types range from 50 to 60 percent. File systems that contain audio, video files, and compressed files are not good candidates for compression. Conversely, it is common to observe high compression ratios of up to 80 percent in database volumes. The proliferation of application virtualization has increased the use of storage space, with more virtual server images and backups kept online. Data compression works well on these highly-variable, mixed application workloads, with compression ratios reaching up to 75 percent.

Intel® QuickAssist Technology

Intel QuickAssist Technology provides hardware-based acceleration that improves the performance of storage, networking, and data security applications across the data center. It can be used by applications to perform compute-intensive compression and security operations typically handled by CPU cores. Server, networking, big data, and storage applications use Intel QuickAssist Technology to offload servers from handling compute-intensive operations such as:

- Lossless and deterministic Deflate compression and decompression
- Symmetric cryptography functions, including cipher operations and authentication operations
- Public key functions, including RSA, Diffie-Hellman, and elliptic curve cryptography

The IBM FlashSystem V9000 platform uses the chipset-integrated capabilities of Intel QuickAssist Technology to provide hardware-based acceleration for real-time compression of data in motion. IBM allows customers to further increase their performance by offering Intel QuickAssist Technology capabilities on acceleration cards.

The Intel QuickAssist Technology implementation within FlashSystem V9000 employs four parallel compression engines (RACE) per compression acceleration card, attached to an IBM DH8 controller node. When compression is enabled, each node utilizes 38 GB of cache for this processing. Each node can be upgraded to contain a second compression acceleration card based on Intel QuickAssist Technology; this is recommended when the working data set exceeds 32 TB.

Intel QuickAssist Technology acceleration add-in cards provide Intel technology partners with a scalable, flexible, and extendable way to integrate Intel QuickAssist Technology compression capabilities into their existing product lines. Real-time data compression also increases flash endurance by using fewer write operations compared to uncompressed data. This is important because NAND Flash chips contain millions of cells that support a finite number of write cycles.



Figure 4. Compression acceleration cards based on Intel® QuickAssist Technology

Benefits of working together

Inline data compression eliminates the need to reserve both time and storage capacity for post-process compression, while maximizing the amount of data that can benefit from compression. Beyond this, Intel QuickAssist Technology offers significant advantages:⁴

- Up to 12x more compression performance
- Up to 10x more CPU cycles available for applications
- Up to 30 percent faster I/O completion
- Up to 80 percent reduction in storage requirements.

See Appendix A for system configuration details.

Intel QuickAssist Technology is deeply integrated with IBM FlashSystem V9000 real-time compression, and the combined technologies create synergies unmatched by other all-flash storage offerings. Chief among these, effective real-time compression

changes the storage economics in the data center, transforming flash into a very attractive solution for all active, and even many inactive, data sets. But the benefits don't stop there. IBM real-time compression powered by Intel QuickAssist Technology:

- Supports workloads that are not candidates for compression in other solutions. The solution allows storage administrators to regain storage capacity in existing storage systems without requiring further "clean up." This enhances the value of existing storage assets and extends their useful life spans.
- Enables compression of replicated/mirrored/remote data in addition to the primary storage volumes. This reduces storage requirements in disaster recovery and data backup environments such as IBM Metro Mirror and Global Mirror.
- Extends compression to virtualized storage systems.

- Requires no changes to the existing environment. IBM FlashSystem V9000 real-time compression is part of the storage system. It was designed with transparency in mind so it can be implemented without changes to applications, hosts, networks, fabrics, or external storage systems. The solution is not apparent to hosts, so users and applications continue to work as-is.
- Drives multiple operational cost savings. More data is stored in less space, so fewer storage enclosures are required, with the following benefits:
 - Reduced power and cooling requirements
 - Reduced software licensing for additional functions in the system
 - Immediate capacity savings because the space reduction occurs when the host writes the data, unlike other compression solutions in which some or all of the reduction is realized only after a post-process compression batch job is run

Flash for Less than the Cost of Tier 1 Disk

Enduring Economics

Real-time Compression, for active and inactive data

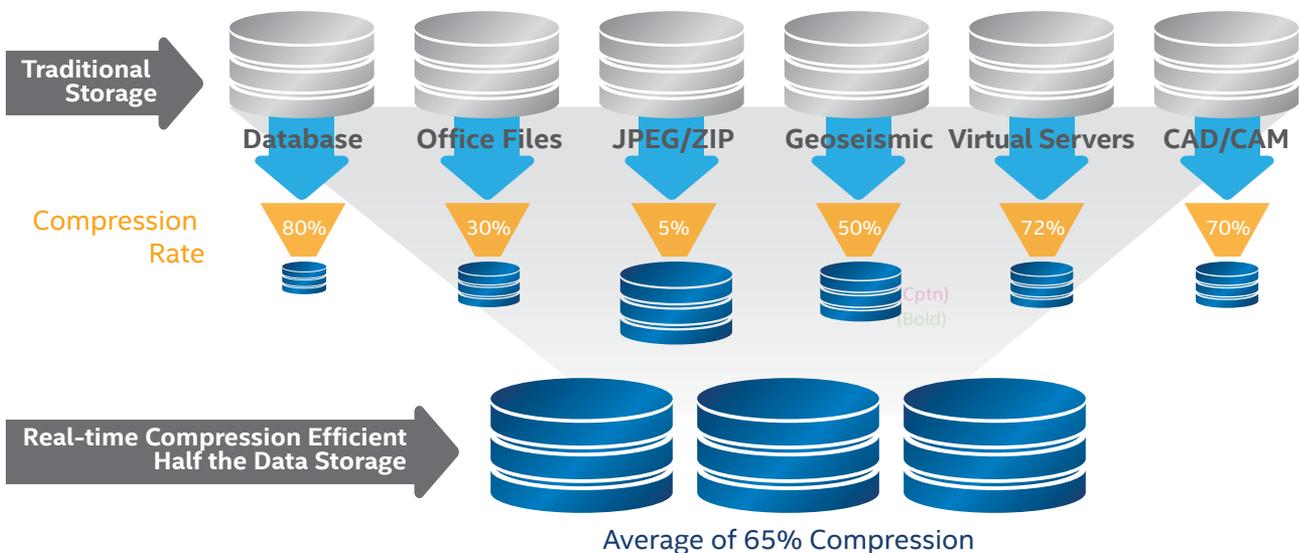


Figure 5. The enduring economics of real-time compression

Changing the dynamics of data storage

Inline, real-time data compression on primary data sets is mandatory for flash-based storage systems. Enterprises can justify extra cost to accelerate business-critical applications when doing so provides significant business advantage. But if flash is to provide the storage solution for all active data sets, positive data economics will be the main driver, as has always been the case.

IBM and Intel have responded with data compression accelerated by Intel QuickAssist Technology that makes IBM FlashSystem V9000 all-flash storage a very compelling solution for any active data sets. This solution is bringing innovation to enterprise, and changing the dynamics of storage through improved storage efficiency.



APPENDIX A: SYSTEM CONFIGURATION

IBM FlashSystem* V9000 Model	9846/8-AC2, 9846/8-AE2
Flash module configuration	4 x 1.2 TB, 6 x 1.2 TB, 8 x 1.2 TB, 10 x 1.2 TB, 12 x 1.2 TB, 6 x 2.9 TB, 8 x 2.9 TB, 10 x 2.9 TB, 12 x 2.9 TB, 6 x 5.7 TB, 8 x 5.7 TB, 10 x 5.7 TB, 12 x 5.7 TB
Maximum storage capacity	Internal storage enclosure(s): Scalable from 2.2 TB (usable) up to 456 TB (usable) with full scale-out of control enclosures. Effective internal: From 11 TB to 2.2 PB with full scale-out of control enclosures (at 80% reduction) External: Up to 32 PB usable capacity.
IBM FlashSystem V9000 host connectivity	8 x 16 Gb Fibre Channel 16 x 8/4 Gb Fibre Channel 8 x 10 Gb Fibre Channel over Ethernet (FCoE) 8 x 10Gb iSCSI
Shared SMP processor configuration	Dual eight-core Intel® Xeon® processor E5 v2 family
Processor memory	64 GB per engine
Number of Intel QuickAssist compressions acceleration cards	2



¹ Two-tenths of a cent cost of tape based on average street price of IBM near-line tape solution

² The ESG Research Report: 2015 Data Storage Market Trends, Oct 21, 2015

³ Joe Unsworth, Gartner presentation (Gartner_IBM Event wide Final.pptx) at FlashSystem 900/V9000 product launch event

⁴ According to Intel research and performance testing.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting Intel's Web site at www.intel.com.

Copyright © 2016 Intel Corporation. All rights reserved. Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.