

IBM FlashSystem storage speeds financial services

Industry spotlight



Contents

- 2 Storage performance matters
 - 2 Financial services need speed
 - 2 Batch process windows are narrowing
 - 3 Security and resiliency count, too
- 4 Flash accelerates financial applications
 - 4 IBM FlashSystem storage is fast, efficient, and reliable
 - 5 FlashSystem arrays fit the financial marketplace
- 6 FlashSystem storage can accelerate profits
- 8 For more information

Storage performance matters

The financial services industry is diverse, encompassing businesses ranging from traditional hometown banking, through investment and wealth management, to global securities exchanges. Though their business models and operational environments vary widely, all members of the financial services industry share something in common: the need for speed.

Financial services need speed

Core banking systems are getting faster as they evolve from pure systems of record to systems of engagement, with new online and mobile access rates increasing dramatically. On the securities side of the business, equities sales occur in milliseconds, with ultra-high-speed algorithmic trading on the rise as tick data response times decline. Throughout the financial sector these trends lead to fierce competition where the performance of information technology (IT) infrastructure makes the difference between the firms that capture market share and profits and those that do not. System response times (latency) and scalability are of critical importance to applications in the financial environment.



There are other unique storage performance requirements for financial services applications. For example, to avoid risk, equity trading applications must record every transaction to non-volatile media. By the very nature of this activity, transactions must be strictly ordered, occur at a fixed point in time, and never be lost. These requirements are enforced through logging mechanisms in databases, message queues, and custom-built applications. For each of these cases, a single threaded persistent write process usually determines the performance of the whole application. Therefore, that performance is ultimately limited by the latency of the storage device.

Beyond operational transaction processing, risk and market assessment requirements of financial services enterprises have also fostered the industry-wide adoption of online analytical processing (OLAP) tools, further emphasizing the requirement for very fast IT systems.

Batch process windows are narrowing

Batch processing can be an efficient method of handling a variety of operations performed by financial services enterprises. Many tasks, such as regulatory filings or customer account updates, need to be completed by a certain time every day or large penalties and customer dissatisfaction result. However, shrinking settlement cycles, expansion of after-hours stock trading, and increasingly stringent regulations are shortening the windows for batch operations.

Add this to ever-rising activity volumes, and the pressures on batch operations in the financial sector are compounded. When the batch process starts approaching its deadline, actions must be taken to reduce the amount of time the batch requires, even though more data must be processed.

For a given batch job, there is a set number of storage system input/output (I/O) requests that must be completed. In a traditional storage system, mechanical hard disk drive movement takes up time during each random I/O request. For enterprise disks, this time is roughly 2 - 5 milliseconds (ms). Every random I/O that the batch has to perform will take this amount of time to complete on average. Disk-based storage area network (SAN) arrays use dozens or hundreds of individual disks to enable many of these I/O requests to be serviced in parallel. However, there is a limit to the parallelism that a particular batch job can leverage. If a batch job can handle 50 parallel I/Os at 5 ms each, then the storage will need to supply 10,000 I/Os per second (IOPS), which requires at least 50 disks, assuming no caching or other factors. For a batch that requires one hundred million I/Os, the disk portion of the batch will take at least 10,000 seconds (100,000,000 I/Os / 10,000 IOPS), or about three hours, depending on caching and optimization techniques. In this situation there are just three ways to significantly reduce the batch time:

- Rewrite the batch application to increase I/O parallelism and add more disks.
- Reduce the number of I/Os the batch job requires.
- Reduce the time that each I/O takes to complete.

The first two of these options may require rewriting the application, reducing the number of features the batch supports, or eliminating customers of the batch. The number of I/Os that the batch requires can sometimes be reduced by adding more memory to the server; however, this becomes impractical when the amount of data that the batch job processes grows large.

For a batch job that processes a significant amount of data, the last option, reducing the time required for each I/O, is the most efficient and cost-effective solution to improve the performance of the batch.

Response time of the storage device is the primary factor that can limit financial system performance.

Security and resiliency count, too

In addition to intense application performance requirements, resiliency of IT systems is critical. The need to electronically create legally binding agreements for tremendous monetary sums produces unique requirements for infrastructure redundancy and real-time position management. For example, in the equities market, a firm that is not in control of its trades and positions is left without a view of its risk and is thus at the mercy of counterparties. In the banking sector, credit card fraud now happens in real time and at massive scale. All of these factors can result in dramatic costs from settlement failures, fines for missing regulatory timing requirements, and complaints from customers. To avoid these failures, data in financial applications not only needs to be fast, it must be secure.



Flash accelerates financial applications

To combat the relatively lengthy response times of disk drives—which are slowed by the mechanical movements of spinning platters and waving armatures—enterprises in the financial sector deploy expensive storage arrays with large battery-backed caches to support their revenue-generating applications. These cached SAN systems have complex architectures and offer a variety of advanced features to help manage point-in-time copies, data movement, RAID level controls, and cache sharing between competing host applications. All of these features come with a latency price because every I/O operation must flow through complex logic. This can add multiple milliseconds of latency. The additional overhead is trivial when compared to back-end disk speeds, but when the bottom line of your business depends on the latency of your storage, every microsecond counts.

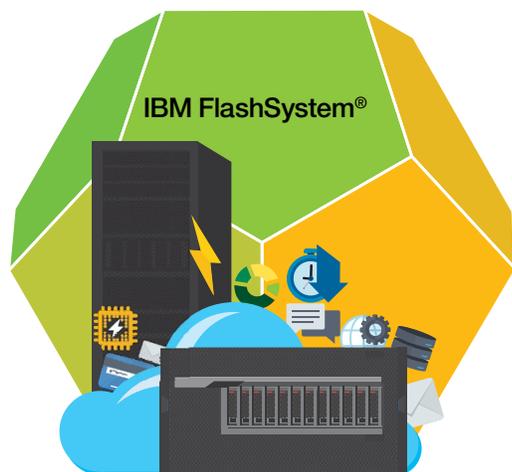
For example, if a host process requires a persistent write, has 50 microseconds of server and network latency, and 200 microseconds of storage latency, by adding these system latencies and dividing by one million microseconds in a second, you can see that a maximum of 4,000 transactions per second is possible for a given process. If the application is recording securities tick data, customer transactions, or persistent middleware messages, then this is the limit of the performance that the application can achieve. In many applications, there are several of these processes that must be completed serially before a transaction can be considered to be executed. In the competitive securities arena many opportunities exist for only a moment, and reducing the time that capital is delayed between a decision and its execution is critical. Customers will rapidly be lost to the competition if execution time is poor.

With the performance gap widening between processors and hard drive-based storage systems, flash-based storage is becoming more and more attractive. Because flash systems rely on memory chips for data storage, they offer unprecedented access times that narrow the gap between processor speeds and storage speeds. Financial services enterprises have used solid state storage systems to resolve I/O performance problems for over three decades. These systems have become increasingly sophisticated, higher performing, and lower cost, which sends a clear message.

There is no better tool than flash for improving financial application performance.

IBM FlashSystem storage is fast, efficient, and reliable

Flash technology has transformed storage and given enterprises the ability to extract extraordinary value from complex data sets. IBM has invested one billion dollars and established worldwide Flash Centers of Competency to help customers design and implement flash-based systems. IBM® FlashSystem® arrays provide industry-leading performance, reliability, and ultra-low latency. The all-flash storage systems offer a full spectrum of enterprise-grade management and feature-rich storage services and provide multiple options for addressing the ultra-low latency requirements of securities transaction systems, removing large batch application performance bottlenecks, and increasing the effectiveness of financial industry computing environments.



Several key differentiators set IBM FlashSystem apart from other flash storage platforms:

- IBM FlashSystem architecture is designed with IBM MicroLatency™ technology to speed response times, delivering data reads and writes in the hundred-microsecond range. This low latency helps financial enterprises complete more transactions faster, gain competitive advantage, and increase revenues.
- Among IBM FlashSystem engineering objectives is a focus on extreme performance. In addition to a passion for low latency, IBM FlashSystem engineers also optimized the platform for scalability, aggregate IOPS, and bandwidth. The resulting extreme performance helps ensure that as financial workloads increase, FlashSystem storage continues to scale performance while minimizing latency degradation. Whether supporting a single application that needs to handle high numbers of concurrent users or multiple applications with diverse workloads, FlashSystem extreme performance can translate into performance scalability and better business results.

- IBM FlashSystem storage is optimized to provide macro efficiency through compact physical space requirements, low energy consumption, and greater utilization of existing resources. The arrays are some of the highest-density solutions on the market, offering dozens of terabytes of usable storage capacity in only a few rack units of space. While providing over one million IOPS, they only draw about 600 watts of power, making them extremely power efficient.
- A key IBM FlashSystem value is enterprise reliability. The system employs two RAID dimensions—patented IBM Variable Stripe RAID™ technology at the flash module level as well as system-level RAID—resulting in more data protection levels than are available from competing systems. FlashSystem design enables rapid servicing because all hot-swappable and redundant components, including flash modules, power supplies, fans, batteries, and canisters, are accessible from the front or back of the system. In addition, software and firmware updates can be completed with the system up and running.

Finally, the IBM FlashSystem family offers a wide range of advanced storage services such as snapshots, data compression, and replication. For financial customers with data that requires an extra layer of protection for adherence to internal or regulatory requirements, FlashSystem products support AES 256 hardware-based data-at-rest encryption.

FlashSystem arrays fit the financial marketplace

There are two broad categories of financial services applications where IBM FlashSystem arrays can be most effectively used:

- Applications where the best performance would come from running in-server memory but the volatility of memory (risk of data loss) cannot be tolerated.
- Applications where such large capacities are needed that server memory is either not large enough or too costly.

IBM FlashSystem arrays are designed to offer a solution to both of these cases.

Applications that require persistent writes do so because they cannot tolerate data loss. The inability to tolerate data loss is the primary reason that an external storage device is needed; otherwise, if performance were the only concern, the application could potentially operate entirely out of server memory. In order to achieve maximum uptime during market hours, storage that receives the persistent write must support clustering technologies and be capable of being shared between servers. To minimize the possibility of data loss, the storage system should also be mirrored as part of the solution.

IBM FlashSystem arrays meet all of the requirements of this environment and feature extremely low latency. They are external arrays that can present volumes to multiple servers. FlashSystem models are standard block storage devices that can be deployed with a wide array of clustering solutions. If FlashSystem arrays are deployed, storage latency can be cut from milliseconds in the cache of traditional SAN arrays to microseconds—even faster than SSD-based storage arrays. Adding the server and network latency, total transaction latency typically still remains well below one millisecond! This might enable over 20,000 transactions per second to be handled, compared to 4,000 previously. IBM FlashSystem products can support multiple servers generating financial transaction workloads, with a million IOPS or more per array.

IBM FlashSystem storage solutions are ideal for financial industry customers who have large I/O bound batch processes. For the batch process described earlier in this paper, cutting

the response time of the storage from 5 ms to less than 1 ms increases the IOPS workload from 10,000 IOPS to 50,000 IOPS. This cuts the runtime of the batch process from nearly 3 hours to 30 minutes! And this performance increase can be accomplished without making any changes to the code.

FlashSystem storage can accelerate profits

FlashSystem storage offers increased performance to some of the most important applications in the financial industry. By dramatically reducing storage latency, FlashSystem products can improve overall application performance without requiring costly code rewrites. Extreme performance, IBM MicroLatency technology, macro efficiency, and enterprise-grade reliability and serviceability make IBM FlashSystem storage a powerful and cost effective tool for accelerating critical financial transactions, securities trading, and banking applications.



For more information

To learn more about FlashSystem storage, please contact your IBM representative or IBM Business Partner, or visit the following website: ibm.com/systems/storage/flash/flash-array.html

Additionally, IBM Global Financing can help you acquire the IT solutions that your business needs in the most cost-effective and strategic way possible. For credit-qualified clients we can customize an IT financing solution to suit your business requirements, enable effective cash management, and improve your total cost of ownership. IBM Global Financing is your smartest choice to fund critical IT investments and propel your business forward. For more information, visit: ibm.com/financing



© Copyright IBM Corporation 2015

Route 100
Somers, NY 10589

Produced in the United States of America
October 2015

IBM, the IBM logo, ibm.com, IBM FlashSystem, MicroLatency, and Variable Stripe RAID are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at ibm.com/legal/copytrade.shtml

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.



Please Recycle
