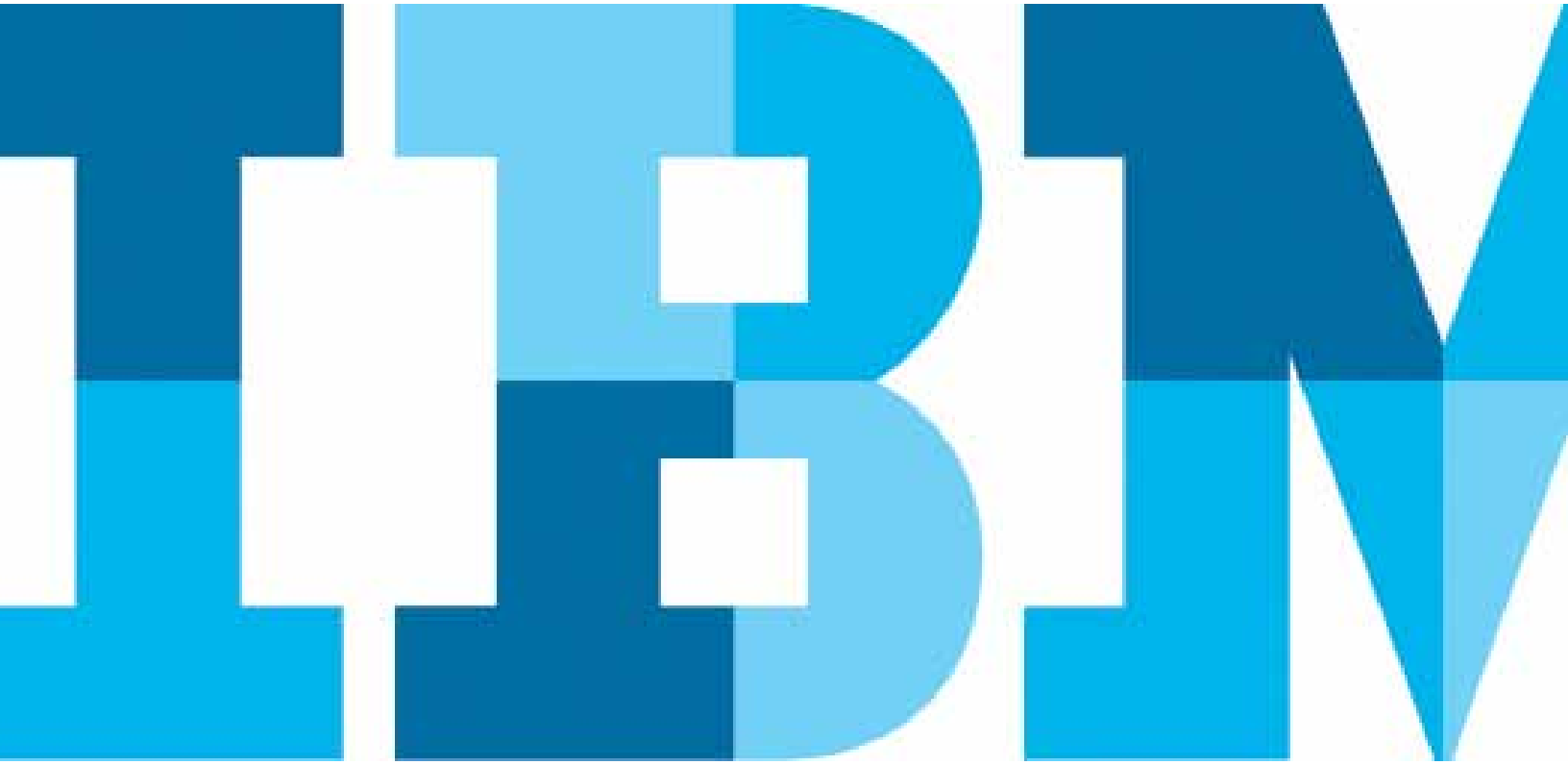
IBM Systems and Technology Group
Thought Leadership White Paper

April 2013

Utilizing IBM FlashSystem in the securities industry



IBM

Executive summary

The securities industry is one of the most performance sensitive sectors of the economy. Electronic exchanges, trading and clearing systems and algorithmic trading are growing as tick data response times decline. This has created an environment of fierce competition where the performance of IT infrastructure makes the difference between the firms that capture market share and profits and those that do not. System response times (latency) are of critical importance to applications in this environment.

In addition to the intense performance requirements of the securities industry, resiliency of IT systems is critical. The need to electronically create legally binding agreements for tremendous monetary sums places unique requirements for redundancy and real-time position management on infrastructure. A firm that is not in control of its trades and positions is left without a view of its risk and is at the mercy of counterparties. This can result in dramatic costs from settlement failures, fines for missing regulatory timing requirements and customer complaints. To avoid this situation, data in securities applications not only needs to be fast, it must be persistent.

Improve performance with flash

With the performance gap between processors and hard drive-based storage systems widening, solid state storage is entering the limelight. Since, solid state systems rely on memory chips for data storage, they offer unprecedented access times that narrow the gap between the processor speeds and storage speeds. Companies have used solid state storage systems to resolve input/output (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 for improving I/O performance.

Understanding solid state storage

Solid state storage as drives or disks (SSD) and arrays are a proven technology. Solid state storage is non-volatile storage that use RAM or Flash as the primary storage media. Solid state storage stores and accesses data directly on chips, which results in storage speeds far greater than conventional magnetic storage devices.

Flash storage has the following characteristics:

Lowest possible access times

- Flash storage, such as the IBM® FlashSystem™ 720, have access times that start at 0.1 milliseconds for reads (50 times faster than a hard disk of 5 milliseconds) and integrate write buffers for access times as low as 0.025 ms for writes.

High bandwidth

- The enterprise solid state storage market includes products such as the IBM FlashSystem 820 which can support over 5 GB per second of random data throughput and across 20 TB of two-dimensional RAID-protected capacity

High I/Os per second (IOPS)

- Flashstorage offer extraordinarily high random I/O performance because of their low access times and high bandwidth. This performance initially scales to over 500,000 I/O requests per second and scales linearly with each system.

Low price for performance

- Flash arrays provide excellent price/performance and feature lower price/capacity than high performance spinning disks that are optimized for performance.

High availability

- Flash arrays are inherently more reliable than hard disk drive based systems because their data path does not require moving parts.

Non-volatile

- Enterprise flash arrays offer non-volatile solutions. With some smaller capacities of RAM within the arrays to offer extremely low latencies beyond native Flash, internal batteries are used to persist any table or uncommitted user data in the event of full A/C power loss. Flash, however, is inherently non-volatile.
- Flash arrays are an excellent solution for I/O bottlenecks, particularly those bottlenecks caused by the high access times of traditional disk-based storage systems.

Problem

Latency sensitive transactional systems

There are unique storage performance requirements for companies offering products and services to the securities industry. Applications in this space demand that every transaction is recorded to non-volatile media to avoid risk. *Latency of the storage device is the primary factor that can limit the entire system performance.* By the very nature of the securities industry, transactions must be strictly ordered, occur at a fixed point in time and not 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 determines the performance of the whole application. The performance is bound by the latency of the storage device. This is not an equation of most scalable IOPS or bandwidth.

To combat the latency of mechanical disk drives, expensive arrays with large battery backed caches are deployed for latency-bound 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 complex circuitry must be traversed for

every IO operation. This can add 200 microseconds up to multiple milliseconds of latency. The additional overhead is trivial when compared to the 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, 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 financial arena, many opportunities exist for only a moment, and reducing the time that capital is in limbo between a decision and its execution is critical. Customers will rapidly be lost to the competition if execution time is poor.

Solution

IBM FlashSystem 720

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 was the only concern, the application could potentially operate entirely out of server memory. In order to meet 100 percent uptime during market hours, storage that receives the persistent write must be capable of being shared between servers and support clustering technologies. To eliminate the possibility of any data loss the storage system should also be mirrored as part of the solution.

IBM FlashSystem 720 meets all of the requirements of this environment with extremely low latency. It is an external array that can present volumes to multiple servers. FlashSystem 720 is a SCSI-III device and can be deployed with a wide array of clustering solutions. A typical deployment of FlashSystem in this environment involves clustered servers and mirrored FlashSystem units for full availability and performance through any maintenance. This architecture, shown in figure 1, allows applications to leverage the performance of low latency storage while tolerating server, storage or power failure.

With this architecture a typical server and storage network latency of 50 microseconds exists. Returning to the example figure 1, if FlashSystem 720 is deployed the storage latency can be cut from the 200 - 2,000 microseconds of the cache in traditional SAN arrays to 25 microseconds. Adding the servers and network latency, the total transaction latency is less than 50 microseconds! This allows 20,000 transactions per second to be handled, compared to 4,000 previously. IBM FlashSystem products can handle multiple servers generating this workload, up to 500,000 IOPS.

Problem

End of day batch process

A batch process can be an efficient method of handling a variety of operations required of a securities company's IT organization. There are many tasks that need to be completed by a certain time every day (for instance, regulatory filings or customer account updates) or large penalties and customer dissatisfaction result. Shrinking settlement cycles and expansion of after-hours trading are reducing the windows for batch operations. Taking this with the ever-rising trading volumes and the pressures on batch operations in the securities industry 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.

Figure 1: Typical IBM FlashSystem architecture for critical persistent write applications.

For a given batch job there is a set number of IO requests that a server needs to complete. On a disk system mechanical movements of the disk take time to fulfill a random IO request. For an enterprise disk this time is roughly 5 - 10 ms. Every IO that the batch has to perform will take this amount of time to complete. SAN disk arrays use multiple disks to enable many of these IO 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 IOs at 5 ms each, then the storage will need to supply 10,000 IOPS, which requires at least 50 disks. For a batch that requires one hundred million IOs, the disk portion of the batch will take at least 100,000,000/ 10,000 IO/s = 10,000 seconds, or about three hours. If the total batch job is in process and the CPU utilization is low, then it is most likely disk bound. In this situation there are just three ways to significantly reduce the batch time:

- Rewrite the batch to allow a greater degree of IO parallelism and complement with a greater number of disks
- Reduce the number of IOs the batch job requires
- Reduce the time that each IO 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 IOs that the batch requires could be reduced by adding more memory to the server; however, this becomes impractical when the amount of data that the batch requires grows large. For a batch job that requires a significant amount of data the last option, reducing the time required for each IO, is the most effective solution to improve the performance of the batch.

Solution

IBM FlashSystem 820

IBM FlashSystem 820 is a purpose-built flash storage system designed from the chip up to meet enterprise requirements for performance, scalability and reliability. The system takes advantage of the strengths of each technology:

- Blazing fast write performance (RAM Buffers)
- Fast reads (Flash)
- High density (Flash)
- Low power consumption (Flash)
- Lower cost per capacity (Flash)

FlashSystem 820 benefits from its fast storage backplane and IO controllers. The system includes RAM+flash buffering algorithms optimized to take advantage of our ultra low latency DDR RAM and the massively parallel array of flash memory with 25 microsecond write and 110 microseconds read performance. FlashSystem 820 has the lowest latency and highest capacity density than any other flash storage array on the market.

FlashSystem 820 meets the needs of customers who have large read-intensive applications that need high IO and low latency. With its high capacity of 20 TB in a single 1U chassis, it is ideally suited for large IO bound batch processes. For the batch process that was presented above, 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 disk portion of the batch process from nearly 3 hours to 30 minutes! And it is accomplished without making any changes to the code.

Figure 2: IBM FlashSystem 820

In addition, supporting a workload of 50,000 IOPS with disks would require at least 250 disks to work in parallel and any application rewrites necessary to use these trays of disks. FlashSystem 820, besides being much smaller and using much less power, is a fraction of the price of monolithic storage arrays with 250 disks.

Conclusion

Flash storage arrays offer increased performance to some of the most important applications in the securities industry. They improve performance without requiring costly rewrites of applications by dramatically reducing the latency of the storage. There are two broad categories of applications where solid state storage arrays are leveraged: applications where the best performance would come from running in server memory but the volatility (risk of data loss) of memory cannot be tolerated and applications where such large capacities are needed that server memory is either not large enough or too costly. IBM FlashSystem 720 and FlashSystem 820 arrays are designed to offer a solution to both of these cases. They are proven solutions deployed globally to accelerate critical financial exchange, securities trading and bank applications.

For more information

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

IBM Corporation
Systems and Technology Group
Route 100
Somers, NY 10589

Produced in the United States of America
April 2013

Please Recycle

TSW03230-USEN-00