



USE CASE

American brokerage and banking company replaces a RAM-based cache on a relational database with Aerospike to successfully achieve Speed at Scale

Customers are increasingly dependent on mobile applications to interact with their financial institutions. According to *Bank of America Trends in Consumer Mobility Report* for 2014, the majority of Americans are using mobile banking applications to check their balances, transfer funds, pay bills, and make deposits. Online brokerages estimate that 15-20% of all banking transactions now take place on mobile devices. †

To help customers make smarter personal finance decisions, it is imperative for modern banks to provide access to accurate data, including stock quotes, trade order status, and balances. Delays in information to customers or lack of visibility into a customer's accounts for the bank could have serious financial implications for both the customer and the bank, including the institution's ability to correctly assess risk and maintain compliance.

PROFILE OF A FINANCIAL SERVICES LEADER

ACCESS MULTI-SERVICE ACCOUNTS
>10 M CUSTOMERS

MAINTAIN CURRENT STOCK PRICES AND BALANCES ON
300 M POSITIONS

500 M TRANSACTIONS PROCESSED DAILY

2 M UPDATES PER TRADING DAY



Eliminate a RAM-based cache on an RDBMS to cost-effectively deliver mission-critical services and remove barriers to growth.

CHALLENGE

To scale without barriers, provide superlative customer experience, and continue to introduce new, engaging mobile applications, this very large financial institution had to rethink its infrastructure. Continuing with a relational database and cache-based solution as the intraday system of record would require scaling from 150 servers to more than 1,000 servers. This was not a practical strategy in terms of time, labor, and operating cost. More important for the business, reliance on nightly batch processing from the intraday system to the master DB2 (book of record) was expensive, and still didn't solve the data inconsistencies between stored and active data.

A new solution had to address:

- The company's decision to continue to leverage its legacy mainframe database (DB2), which was the compliant system of record for more than 10 million customer accounts.
- The requirement to process 250 million transactions and 2 million updates a day, and the ability to update stock prices or show balances on 300 million positions in near real time.
- The ability to create enough compute capacity to eliminate data inconsistencies.
- The elimination of frequent system crashes due to overloading the RAM-based cache and the subsequent restarts, which frequently took around an hour.
- The mandate for a cost-effective solution that would address expectations for 1,000% data growth as it executed on its mobile strategy.

The company's IT department explored a number of options and quickly realized that Aerospike was the only solution that could provide the required performance. Rather than use expensive yet unreliable RAM, the bank preferred Aerospike's ability to leverage solid state drives (SSD), also known as Enterprise Flash. Enterprise Flash pushes down costs and increases reliability because it has much greater storage density than RAM (thus requiring fewer servers) and immediate data persistence.

Competitors included Mongo DB, Gigaspaces, and Gemfire, which were eliminated from consideration for a variety of reasons.

- Mongo DB couldn't provide low read latency at a high write load. It was also not flash-optimized so it could not significantly reduce server count.
- Gigaspaces is an application-centric cache solution and guarantees vary based on the database layer choice. Its architecture ties applications to the distribution layer, making it impossible to efficiently access data from multiple clients.
- Gemfire needed at least 4x more nodes and could not deliver predictable performance with low latency at high throughput.

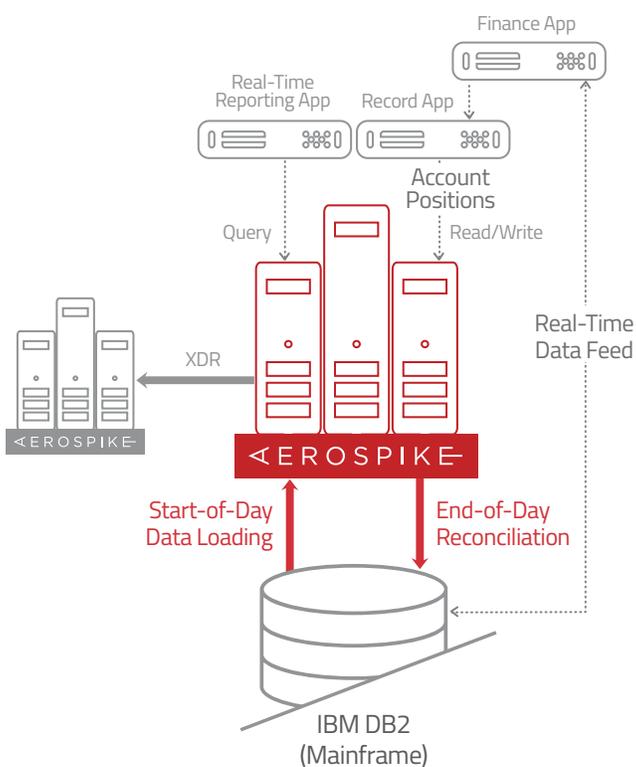
BACKGROUND

The bank's IT group was at a crossroads. The combination of an RDBMS and a RAM-based cache fronting a traditional mainframe database was unable to consistently and reliably support ever-growing workloads during trading hours. This reality had become a major impediment to the company's strategic goal of releasing a steady flow of new and updated applications to its mobile customer base.

The business mandate to seamlessly service more than 10 million customers and process more than 250 million transactions daily, while positioning itself to eventually process 1 billion transactions per day, was at risk. The company's conventional RDBMS infrastructure and cache capabilities couldn't consistently deliver the performance, scalability, and flexibility needed to meet these requirements. In its current configuration, the caching solution quickly became overloaded, rendering it unable to scale during high-volume trading scenarios or handle new applications without serious latency issues and system outages. System restarts were taking as much as an hour to complete, and given its reliance on RAM, which isn't persistent, data loss was a serious problem.

THE RESULT

Aerospike Speed at Scale: Higher throughput, lower latency, availability and one-tenth the cost of ownership



Aerospike's database technology, which is now the bank's intraday system of record, enables high throughput with low latency, horizontal scalability, availability, and operational efficiency. Key Aerospike features such as flash-optimization, key-value store, and XDR (cross data center replication) are delivering unmatched Speed at Scale with no loss of data.

Because Aerospike enabled the bank to leverage its traditional database (DB2), Aerospike was ideally suited to function as the intraday system of record. Now, the daily workload is created and managed in Aerospike, while the historical data (and a few core applications) are managed separately in the traditional RDBMS.

The diagram on the left illustrates how Aerospike is used as the intraday system of record, while DB2 still functions as the compliance related 'Book of Record' and synchronizes with Aerospike at the end of each trading day.

Because Aerospike runs natively in Flash, which is significantly faster and more dense than disc, it makes dramatically more efficient use of the hardware resources. As a result, the company reduced its hardware footprint from 150 RAM cache servers to a 10-server Aerospike cluster, while simultaneously improving performance.

As a result, the institution was able to scale applications with far fewer barriers. Aerospike enabled better performance, better reliability, and material cost savings with:

- 90% fewer servers yielding a 90% reduction in TCO
- 3x greater server capacity than RAM (4 TB → 14 TB of data)
- 4x larger object store (120 → 480 million objects)
- 5x faster throughput (200K → 1 million TPS) with sub-millisecond latency
- Predictable performance
- Unparalleled uptime and availability

Aerospike helped this bank mitigate risk, enhance customer experience, and achieve unlimited growth potential, all while improving its bottom line.

This company's deployment of Aerospike's high-performance NoSQL database delivers high throughput during the trading day with predictable low latency, and immediate consistency with no data loss. Thanks to XDR (cross data center replication), Aerospike eliminated unplanned downtime and potential data loss with no single point of failure.

The company is now able to write more applications against ever-growing data sets without negatively impacting performance. With response times in milliseconds, institutions can provide a superlative customer experience while simultaneously mitigating its own risk. By leveraging important components of the legacy infrastructure, Aerospike enables administrators to properly manage and control the system, thereby creating an audit trail that meets compliance regulations.