

PayPal Puts Data at the Heart of its Fraud Strategy with Aerospike



Extreme Scale Computing for Real-time Decisioning without Breaking the Bank

About PayPal

PayPal (www.paypal.com) is the world's largest online money transfer, billing and payments system. It owns the PayPal, Venmo, iZettle, Xoom, Braintree, and Paydiant brands. By leveraging technology to make financial services and commerce more convenient, affordable, and secure, the PayPal platform is empowering more than 325 million consumers and merchants in more than 200 markets to join and thrive in the global economy. In 2019, PayPal's [total payment volume](#) amounted to more than 712 billion U.S. dollars, representing a 23 percent growth from 2018.

Benefits:

- Meet 99.95% of fraud calculation SLAs
- 10x improvement in fraud calculation data
- TCO reduction of over 80% in 5 years
- Each node supports up to 4X as much data per node.
- 5X throughput improvement scaling

CHALLENGES

PayPal's fraud rate is between 0.17% and 0.18%* of revenue. While this number is well below the industry average of 1.86%*, it still represents over 1 billion dollars a year for the firm. Needing to be able to more quickly process and analyze its data to identify emerging fraud patterns in real-time, PayPal sought to quickly build a real-time decisioning platform that was highly effective while minimizing end-user friction. The platform merged big data with fast data such as customer enrollment, payment, invoicing preferences, and profile data.

PayPal determined that the massive amounts of data it was collecting was dragging down its existing database. It needed a cost-effective strategy as it began to scale horizontally – while also ensuring that performance and uptime didn't suffer. Data scale challenges included hundreds of petabytes, and a requirement to grow from 3.5 to 7 million transactions a second.

WHY AEROSPIKE

PayPal began its partnership with Aerospike in 2015 to take advantage of its Hybrid Memory Architecture™ to scale with higher storage density and fast SSDs. Aerospike is designed with a Hybrid Memory Architecture™ (HMA) that delivered a fundamentally different approach to traditional architectures that enabled PayPal to overcome its fraud prevention data challenges.

“PayPal is innovating deep analytics to rapidly respond to emerging fraud patterns, then deploying into an event-driven, fast data, in-memory architecture to accelerate detection, reduce losses and achieve near-continuous availability.”

- Mikhail Kourjanski PHD, Lead Data Architect, Risk and Compliance Management Platform, PayPal

REQUIREMENTS MET WITH AEROSPIKE

When their legacy NoSQL database couldn't deliver results without a deterioration in performance, PayPal turned to Aerospike. With its high-performance NoSQL database, Aerospike is built explicitly to run on flash and Persistent Memory (PMem), not DRAM, while consistently delivering speed at scale. Here's what Aerospike was able to do for PayPal:

- Meet 99.95% of fraud calculation SLAs.
- Cost effectively scales up and scales out in order to handle extreme growth in data and objects.
- Improve query performance and data consistency.
- Maintain consistent high availability in a 24/7 environment necessary for digital payments (99.99 percent uptime).
- Enable its fraud solution to scale up and out at an annual projected data growth of 32 percent.
- Slash data load times, while providing reliable access to fresh data.
- Enable more than eight million executions per second across database environments, which includes RDBMS and NoSQL.

INITIAL BENEFITS REALIZED WITH AEROSPIKE (2015)

Based on PayPal's initial projected fraud data size of 50 TB in 2015 with its legacy in-memory system, these were the benefits with the Aerospike® Hybrid Memory Architecture™:

- 30X reduction in the number of fraud transactions missed
- 15X reduction in server footprint
- \$9m hardware cost savings projected from \$12.5m down to \$3.5m
- 5X throughput improvement scaling from 200K to 1M transactions per second

CONTINUED GROWTH BRINGS ADDITIONAL CHALLENGES

PayPal’s fraud decisioning platform comes at a significant cost in terms of the number of servers and staff required to manage them. They are challenged with over 4,000 database instances, 100 petabytes of data, and 32% data storage growth each year. In order to keep database nodes, count down, PayPal looked to continually increase the density per node. Given advancements in SSD technology and Aerospike’s Hybrid Memory Architecture being built to optimize SSDs, Aerospike was poised to capitalize. However, as the storage density in terms of number of keys per node started to exceed 1 Billion, re-index and rebalance times started to grow.

“Prior to Aerospike, we were using another in-memory data store and we are running into challenges in terms of the cost of scaling. We needed to seamlessly leverage both the memory and disk in such a way that it can guarantee a consistent performance. We moved to Aerospike for its hybrid memory architecture to leverage next generation memory and SSDs to their fullest advantage.”

- Sai Devabhaktuni, Sr Director of Engineering, PayPal

AEROSPIKE ON INTEL OPTANE PERSISTENT MEMORY (2020)

Intel® Optane™ DC Persistent Memory, with its high-density design, delivers a unique combination of affordable large capacity and persistence (non-volatility). Aerospike’s [Hybrid Memory Architecture™](#) is engineered to make advanced use of Intel® Optane PMem for real-time computing. A multiyear collaboration between the two companies enables optimizations for higher performance than SSDs, lower cost per GB than DRAM, higher node densities, and data indexes which persist over system restarts.

SUBSEQUENT BENEFITS WITH AEROSPIKE ON INTEL OPTANE PERSISTENT MEMORY (2020)

PayPal challenged itself to look at next generation memory and storage differently and began evaluating 2nd Gen Intel® Xeon® Scalable processors and Intel® Optane™ PMem to increase their overall storage density and reduce their DRAM costs with Aerospike. PayPal now has a total 2000 Aerospike servers, including 200 high-density servers with Intel Optane Persistent Memory and expects continued growth.

By doing so, they were able to lower TCO through DRAM reduction with extremely fast application response and low latency. Additional benefits include:

- 4X increase in storage/node (3.2 – 12TB)
- 2X reduction in servers
- 30% reduction in costs/cluster
- 12X reduction in re-indexing time

Comparison Evaluation

	AEROSPIKE INTEL SSD	AEROSPIKE INTEL PMEM
Max Keys/Node	~2B	~10B
Max Usable Storage/Node	3.2 TB	12.8 TB
Nodes / Cluster	20	10
Reindexing Time	59 minutes	4 minutes
Cost/Cluster Group	~\$1.3M	~\$0.9M

NEXT STEPS IN MASTERING SCALE

Fraud prevention is an important area of investment for PayPal. The company has successfully used real-time decisioning with robust fraud prevention models for more than a decade. However, fraudsters are constantly changing their patterns and uncovering new ways to take advantage of the system. As a result, PayPal must continuously find ways to improve fraud detection accuracy and decrease fraud detection time.

Deploying Aerospike's Hybrid Memory Architecture™, a NoSQL data platform with Intel® Optane™ PMem, the PayPal team is able to deliver superior speed, horizontal scalability, reliability and lower TCO.

Powered By:



Resources:

Sources:

[LexisNexis® Risk Solutions 2019 True Cost of Fraud Study](#)

[The Science Behind a Delightful User Experience](#)

[Hybrid Memory Architecture at PayPal \(2018\)](#)

[PayPal: Aerospike + Persistent Memory = Voila! \(2020\)](#)