AEROSPIKE

The Aerospike Real-time Data Platform

SOLUTION BRIEF

Advantages over Cassandra/DataStax

On the surface, Cassandra seems like a great choice for NoSQL projects. However, if you plan to scale your data or use complex workloads, it is likely that Cassandra will not be able to meet your latency and stability needs. It is our experience that Cassandra will fail you at the worst possible time - just as your business is experiencing rapid growth.

Unlike Cassandra, the Aerospike Real-time Data Platform powers leading, innovative businesses like Wayfair, The Trade Desk and LexisNexis Risk Solutions to act in real time across billions of transactions with predictable performance up to petabyte scale, all with 99.999% uptime. Best yet, Aerospike customers typically reduce their server footprint by 80% even as their business grows.

Aerospike Benchmark Results vs. Cassandra

AEROSPIKE DELIVERS

(as tested using Yahoo Cloud Serving Benchmark)

15.4x better transaction throughput

14.8x better 99th percentile read latency

12.8x better 99th percentile update latency

In less than 10% of Cassandra's cost

Why Cassandra/DataStax users switch to Aerospike

"With our previous database (Cassandra), we were operating at a request duration that would hurt our customers, because they couldn't make a risk decision in the allocated time window that they need to... Aerospike helps us to be way under that range and to return in almost all cases with a very low latency so that all risk decisions can be made."

Matthias Baumhof, Vice President of Worldwide Engineering, ThreatMetrix

You should contact us to obtain a proof of concept or at least Trial Aerospike for Free if any of the following five concerns apply to you:

- You are struggling with "server sprawl" and worried about TCO
- Your peak workloads are disrupting SLAs
- You need persistence with high performance
- Your operations team keeps growing and you are concerned about costs
- You are struggling to acquire and retain Cassandra expertise

All of the above are non-issues with Aerospike. Sound too good to be true? Read about the unique technology and patents we bring to bear to achieve our extraordinary results in our comprehensive whitepaper, "Five Signs You Have Outgrown Cassandra".

AEROSPIKE

How Aerospike's Real-time Data Platform Achieves Such Outstanding Results

Act in Real Time

Aerospike was designed and built from the ground up to take advantage of modern computing architectures. This allows large and rapidly growing businesses to act in real time and rely on Aerospike as their foundation for the future.

"One thing I love about Aerospike is that it's very capable of taking advantage of any hardware you throw at it. This might be your network devices, this might be your SSDs or your NVMe storage, it might be your RAM."

Henry Snow, VP, Infrastructure, Nielsen Marketing Cloud, Nielsen

"Before Aerospike, we were spending more and more of our time on the care and feeding of Cassandra, and less and less time on the building of new product offerings. With Aerospike, we've now cleared the roadmap and we're just focused on adding new functionality to our platform for our customers."

Jason Yanowitz, Vice President of Engineering, Signal

Predictable Performance into Petabyte Scale+

Aerospike was designed with ultra-fast and predictable performance at scale in mind. It features a patented Hybrid Memory Architecture (HMA) that delivers exceptional runtime performances for mixed read/write workloads at scale, often at a savings of 60% or more in operational expenses (OPEX) when compared to Cassandra and other alternatives.

"I have to be able to persist millions of events per second, that's millions of IO's per second. The only solution that can get close with today's technology hardware and software is Aerospike running on HPE Persistent Memory DIMMs, and that is about 280,000 sustained read and write operations per second – which is about 2,000 percent more than anything else out there."

Theresa Melvin, Chief Architect of AI-driven Big Data Solutions, HPE

"During our initial deployment to Aerospike and our shaking of it out, we pushed the limit up to 8 million transactions per second and saw the p50 at 10 microseconds. Which was absolutely stunning to us, and almost a thousand times faster than what we were seeing before that (with Cassandra). (In addition) the data was more reliable, and what we were putting in there was actually what would come back out."

Jason Yanowitz, Vice President of Engineering, Signal

99.999% Uptime

Aerospike customers experience five nines uptime with strongly consistent, globally distributed data. Leveraging a shared-nothing architecture, Aerospike is built with a unique master-based cluster algorithm, where if you lose a node, you have another copy at-the-ready. Unlike other systems, Aerospike writes synchronously across all copies of the data.

This provides huge benefits because when a node fails or is removed from the cluster, any node that has a secondary copy of the partition can be instantly promoted to be the master of that partition without the typical delays imposed by a consensus algorithm. Combined with Cross Datacenter Replication (XDR) being used to ensure availability across regions, Aerospike is hardened against human error and natural disasters.

"When we switched over to using Aerospike, we saw immediate improvements (over Cassandra) on a bunch of axes. One was, the data was more reliable, and what we were putting in there was actually what would come back out. Secondly, we saw huge performance improvements. Our p99s went from 3,900 milliseconds to 23 milliseconds."

Jason Yanowitz, Vice President of Engineering, Signal

"Cross Datacenter Replication is vital to our organization. We need to have redundant data centers. We need our user objects to be available in multiple facilities. . . . The ability to replicate data across regions is something that Aerospike provides that very (few) other NoSQL databases do with ease."

Henry Snow, VP Infrastructure, Nielsen Marketing Cloud, Nielsen

Reduce Footprint Even While Growing Business

Everyone knows that a smaller footprint means less moving parts and a lower total cost of ownership. What sets Aerospike apart is that customers typically reduce their server footprint by 80% even as business and data grows. How?

Aerospike is a native C implementation and therefore does not experience Java runtime inefficiencies. Instead it takes advantage of ultra-fast key lookups in DRAM and by reading and writing in parallel to all devices so it fully utilizes all of the IOPs and disk slots available before running out of CPU.

Customers like The Trade Desk and Signal chose Aerospike over Cassandra in part because of its resource efficiency and smaller footprint.

"With Cassandra there's a lot more configuration and tuning out of the box. Aerospike? Pretty much change a few things and you're good to go."

Ken Bakunas, NoSQL Data Architect, Wayfair

"In order to get the throughput that we needed [with Cassandra], we needed to scale the number of machines to a high number of machines with a lot of CPU compared to the disk they had. Aerospike gave us another alternative."

Matt Cochran, Director of Engineering, The Trade Desk

CASE STUDIES

Customers and Partners on Aerospike over Cassandra

Signal is the leading SaaS provider of data onboarding and real-time identity resolution. They looked to Aerospike to replace their existing Cassandra datastore and speed up their business.

Reason they migrated to Aerospike from Cassandra

Signal, being an identity resolution platform, was looking to replace its existing data store, which was becoming increasingly expensive, unreliable, and nonperforming – affecting the bottom line. One of the biggest problems they were running into was large and unpredictable latency response as well as uptime, both of which were affecting every element of their business processes. They were experiencing more frequent and more severe issues and incidents, all related to an unreliable data store.

Signal experienced the following results (and more) from Aerospike

- TCO Reduction of 68% over three years
- Server count reduced from 450 to 60
- Performance improved 100x at the 99th percentile
- Business processes executing in 1/10th the time – or better
- Time freed up to focus on more strategic, forward-looking projects

SIGNAL®
A TransUnion® Company

Read the full case study

Wayfair is one of the world's leading home furnishing platforms. The Wayfair platform leverages Aerospike for customer scoring and segmentation, tracking events online, "listening" to customer activity for marketing decisions, onsite advertising, and recommendation engines.

Reason they migrated to Aerospike from Cassandra

Wayfair has undergone immense and rapid growth; 15.2 million active customers, more than 14 million products on offer, and the successful creation of a retail holiday that pushed its systems to the limit. "Way Day", as it's affectionately known, was full of highs and lows, and showed the Distributed Systems team it needed a higher performance database than Cassandra. The demand for both on-premises datacenters and multiple in the cloud was telling

Results Wayfair has Experienced with Aerospike

The Wayfair platform leverages Aerospike for customer scoring and segmentation, tracking events online, "listening" to customer activity for marketing decisions, onsite advertising, and recommendation engines. In the process of implementing Aerospike, Wayfair was able to cut its server footprint to 1/8th of what it was previously, keeping up its commitment to a superior customer experience.

- Server reduction from 60 servers down to 7 servers
- Leverage the cloud to deal with usage spikes
- Aerospike cost less than Cassandra despite being much faster and providing much more value

Watch Ken Bakunas presentation "Moving Sofas in Millisecond Time"

The Trade Desk is an advertising technology platform. They represent the buyers in an ad exchange where, when impressions are shown on digital media, they facilitate the transaction by buying the right impressions for customers. With around 10 million queries a second, the company has to condense, read, and analyze millions of data points, and store the most relevant data to be quickly accessible.

Reason they migrated to Aerospike from Cassandra

The Trade Desk has a vast pool of data stored in long term storage. They used Cassandra for their cold store, but the data structures available using Cassandra to get the high write throughput that they needed weren't as effective for some of the read cases that they had. Their biggest challenge with Cassandra was the need for a high ratio of CPU-to-data. In order to get the level of writes they needed they had to use compression and tombstoning. There was a lot of CPU-utilization needed relative to the size of the data they were working on. To get the necessary throughput, they had to scale to a huge number of machines with a lot of CPU compared to the disk that they had. Their Cassandra solution simply couldn't scale to meet their business requirements. They needed to create an offline cold store to store unused data and optimize for real-time systems.

Results The Trade Desk has Experienced with Aerospike

As a result of using Aerospike, The Trade Deck achieves hyperscale for real-time bidding leveraging both hot cache and cold store. Combining a hot cache with a cold store provided The Trade Desk with more bid opportunities and a more efficient infrastructure. Aerospike runs at the edge as a cache for real-time bidding—which receives 11 million queries per second (800 billion queries per day) and also as a system of record on AWS to manage peak loads of 20 million writes per second in its "cold storage" of user profiles.

theTradeDesk

Read the complete customer story: The Trade Desk achieves hyperscale for real-time bidding

The Trade Desk has a lot more flexibility in how they approach organizing their data. They can use their data in different ways now. In addition they have:

- Record-level compression while reducing machines needed to 60
- Flexible data organization with one key to represent many different dimensions so they get back only the data needed for a given use case.
- Rapid access to cold store – thaw data in 8 milliseconds for real-time bidding
- Record-based model with all data put into one record – not nearly as CPU intensive
- Reduced in-datacenter footprint - write 30MM key-value tuples/sec into 1 PB cold cache.

ThreatMetrix, a LexisNexis Risk solutions company, is the world's largest Digital Identity Network. The company does authentication and fraud prevention for online transactions, providing real time decisions for customers in the e-commerce and financial services spaces that need to make a decision in milliseconds for fraud avoidance.

Reason ThreatMetrix Replaced Cassandra with Aerospike

ThreatMetrix's main challenge was performance. They were not able to provide the authentication and fraud prevention information their customers needed in the time needed, reliably and at scale. Latency and response times were not where they needed to be and it was affecting customers. This was preventing them from scaling their business effectively. Examples included customers having to wait too long to make a decision to avoid fraud and website transaction times taking too long and transactions being abandoned.

Results ThreatMetrix has Experienced with Aerospike

Aerospike allows ThreatMetrix to perform more complex risk calculations in less time and at the same time take full advantage of their hardware. Matthias Baumhof summarizes their experience by saying, "We are simply in a latency game and Aerospike is the best in the latency game." Other results of moving to Aerospike include:

- Replaced 96 Node Cassandra deployment with just 28 nodes of Aerospike while improving SLAs
- Now able to handle over 130 million transactions a day
- Reduced latency from 120 milliseconds down to 30 milliseconds
- Manage real-time customer trust decisions in less than 300 milliseconds while virtually eliminating false positives and greatly enhancing fraud detection

LexisNexis® RISK SOLUTIONS | ThreatMetrix®

Watch their presentation, "Replacing Cassandra: A Digital Transformation for the World's Largest Digital Identity Network"

CASE STUDIES

Hewlett Packard Enterprise

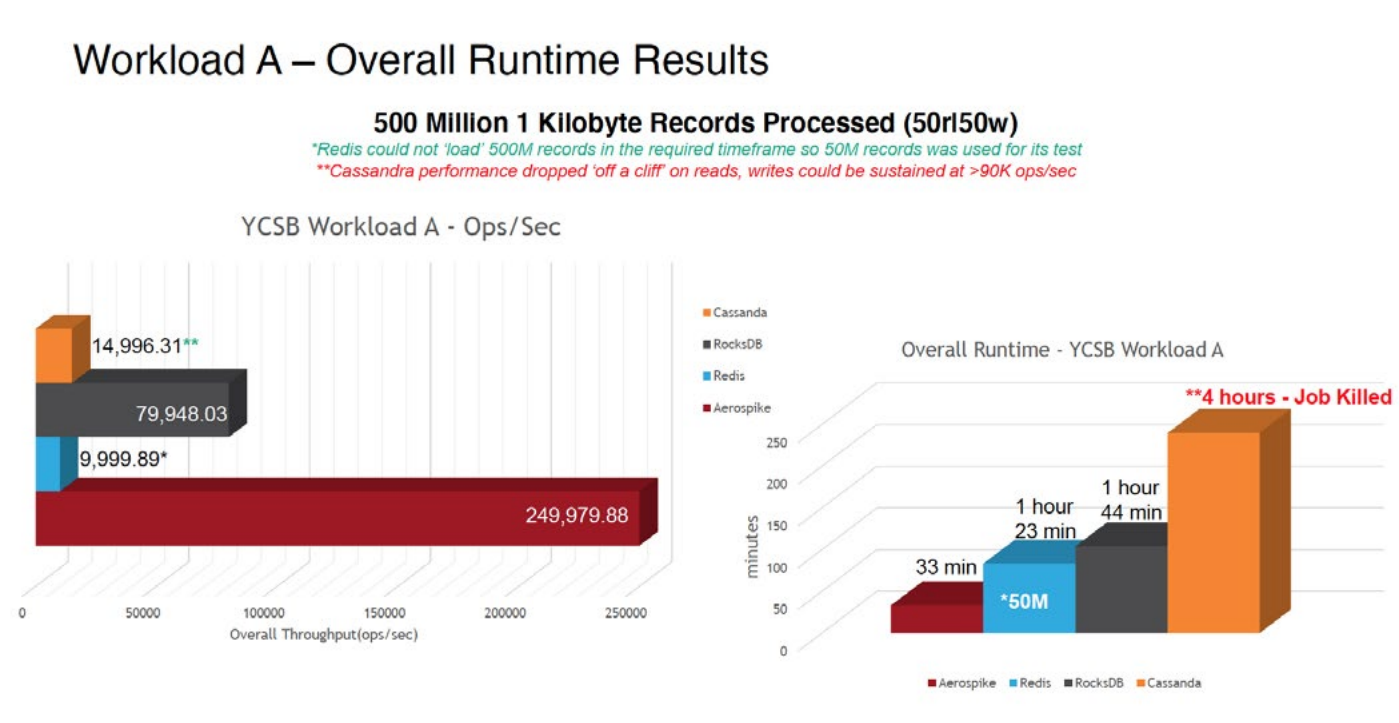
Hewlett Packard Enterprise designs and delivers systems that are capable of hyper and exascale computing to a number of industries. In order to deliver the performance their customers need to run AI leveraging their big data, HPE runs benchmark testing to develop solutions that deliver.

Reason They Ran Benchmark Tests Using Aerospike

HPE's cloud, US government and academia customers are focusing on hyper exascale computing with many other industries following closely. Companies looking to deploy AI for real-world use cases often get stuck at data ingest. Hyperscale and Extreme-scale both require an entirely new data architecture to deliver both scale and performance without breaking the bank. In order to create the "lean, mean, efficient machine" they needed to deliver hyperscale computing at an affordable price point, they tested a variety of databases.

Results of HPE's Benchmark Testing

Aerospike's performance far exceeded all competitors including Cassandra, Redis and RockDB. Cassandra performance "dropped off a cliff" on reads and writes could only be sustained at >90K ops/sec. Redis could not even load the 500M records necessary for the test in the required timeframe.



Most importantly, Theresa Melvin says, “Designs that I put together, they have to be able to write as fast as they read. A lot of times I have a 1 to 10 write-read ratio. For every one terabyte that is inserted, I have to read out 10 terabytes. So, that requires a very special type of NoSQL database, and unfortunately every single database that I have tested over 20 months failed in that regard with the exception of Aerospike.”

Learn More About the Results of HPE’s Benchmark Testing with Aerospike

Hear Theresa Melvin, Chief Architect of AI-Driven Big Data Solutions for HPE review her research and benchmark testing for real-world AI-driven use cases requiring Exa-to-Zetta scale, low latency, end-to-end automation, and ever-shrinking data center footprints.

Watch Theresa showcase the results of work with Aerospike on HPE servers with Intel Optane DC persistent memory.

AEROSPIKE

The Aerospike Real-time Data Platform enables organizations to act instantly across billions of transactions while reducing server footprint by up to 80 percent. The Aerospike multi-cloud platform powers real-time applications with predictable sub-millisecond performance up to petabyte scale with five-nines uptime and globally distributed, strongly consistent data. Applications built on the Aerospike Real-time Data Platform fight fraud, provide recommendations that dramatically increase shopping cart size, enable global digital payments, and deliver hyper-personalized user experiences to tens of millions of customers. Customers such as Airtel, Experian, European Central Bank, Nielsen, PayPal, Snap, Verizon Media and Wayfair rely on Aerospike as their data foundation for the future. Headquartered in Mountain View, California, the company also has offices in London, Bangalore and Tel Aviv.

2525 E Charleston Road, Mountain View, CA, 94043 | (408) 462-2376 | aerospike.com