

# The Aerospike® Difference

## SOLUTION BRIEF

## How the Aerospike Real-time Data Platform technology uniquely helps your business

### Overview

The intensely competitive global economy is forcing firms to deliver highly personalized products and services in real time, around the clock, and around the globe, placing incredible demands on existing data management infrastructures. Existing data platform technologies include relational databases, first-generation NoSQL databases, and even mainframes – but they each have distinct shortcomings.

Relational databases, even distributed ones, don't have the high performance (low latency millisecond responsiveness coupled with millions of data transactions per second) that NoSQL technologies do to meet this need. First-generation NoSQL technologies rely on volatile DRAM memory to get high performance but require high server counts to hold large amounts of data. Mainframes offer speed with accuracy and uptime but cannot be distributed globally at the Edge where users access data and are costly to operate.

Aerospike solves these challenges with its highly patented Hybrid Memory Architecture™, which delivers DRAM-like high performance with SSDs. As a result, Aerospike technology provides high performance, strong consistency, and availability while being globally distributed, reducing server footprint by up to 80% while your business grows - unlike any other data platform in the world.

Companies have built their foundation for the future with Aerospike. Some examples include:

- [PayPal](#) slashing its fraud exposure 30x.
- [Wayfair](#) increased its average cart size at checkout by double-digit percentages.
- [Top 3 Global Brokerage](#) processing trades five times faster.

### Highlights

- SSDs treated like DRAM
- Tiered storage options for optimal densities with high performance
- Node reduction by up to 80 percent
- Predictable performance from terabytes to petabytes
- Direct path to data (single hop)
- Hotspots avoided with intelligent auto-sharding
- Node add/remove/update without disruption
- High availability upon site failures without intervention
- Intercloud data replication with fine-grained control
- Real-time transactions without data loss
- Many use cases for the Aerospike Real-time Data Platform

## The Aerospike architecture is unique

The Aerospike Real-time Data Platform is a shared-nothing, multi-threaded, multimodal data platform designed to operate efficiently on a cluster of server nodes, exploiting modern hardware and network technologies to drive reliably fast performance, often at sub-millisecond speeds across petabytes of data.

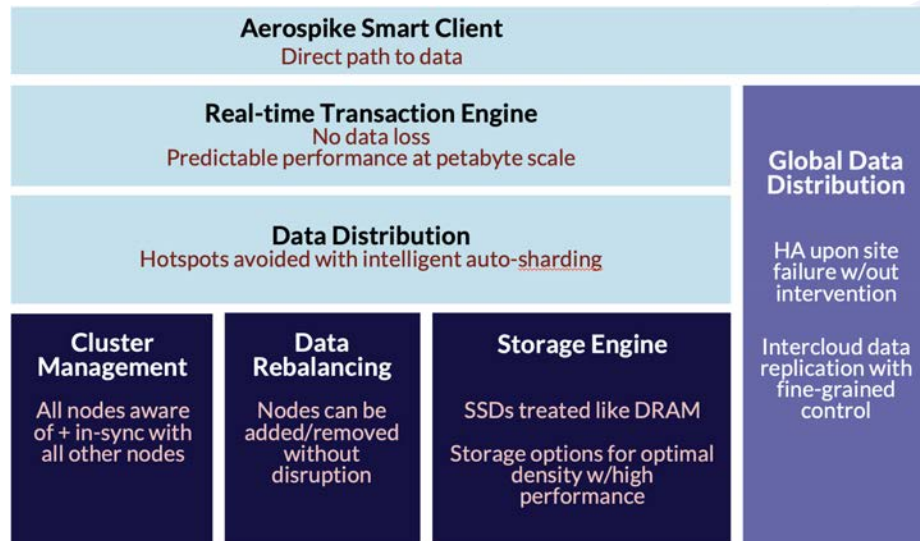


Figure 1: The Aerospike architecture key components and capabilities

## SSDs treated like DRAM

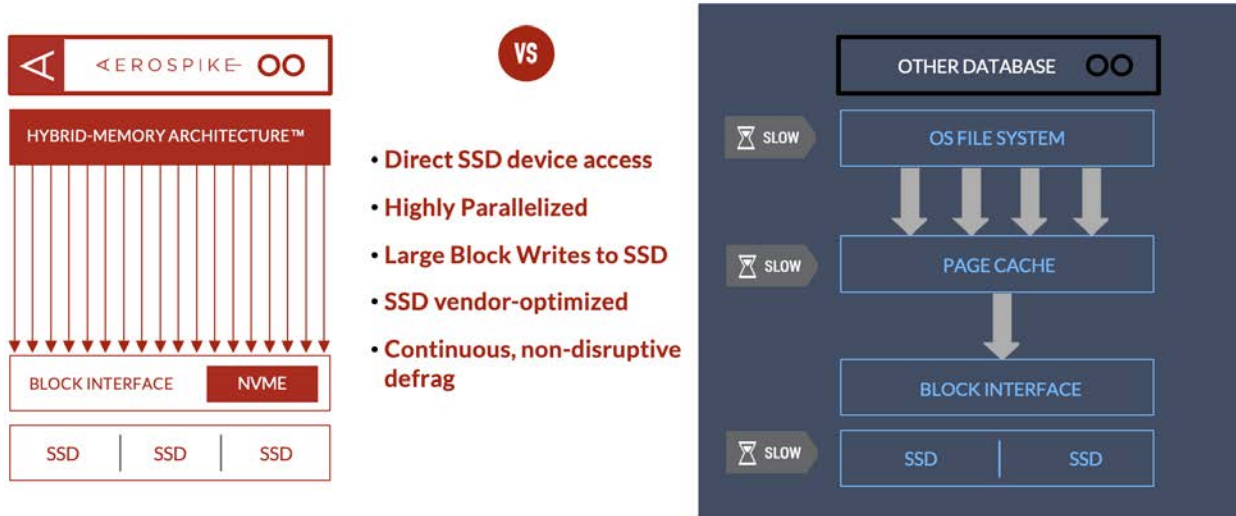
Aerospike has a patented Hybrid Memory Architecture™ (HMA), which places data on solid-state drives (SSDs) and indexes-only in DRAM. In addition, Aerospike treats SSDs as raw devices, writing data in large blocks using a highly efficient custom file format that avoids wear-leveling issues common with other providers.

Net-net: you can get the performance of DRAM for the price and reliability of SSDs.

How was Aerospike able to do this? Simply put, Aerospike is software that was written in C to talk to hardware, not an API layer natively. Aerospike treats SSDs as a large parallel memory space, not a file system.

Per Figure 2 below, Aerospike has:

- Direct SSD device access
- High parallelism (i.e., is not single-threaded)
- Conducts large block writes to the SSD - which is how Flash prefers writes (by dividing an SSD into multiple partitions)
- On-going technology partnerships with top vendors such as Intel, Samsung, and Micron for testing
- Continuous, non-disruptive defragmentation



*Software written in C to natively talk to hardware, not an API layer*

Figure 2: Aerospike exploits SSDs in ways other systems don't

## Tiered storage options for optimal densities with high performance

Administrators can configure Aerospike's storage usage to suit their cost, performance, availability, and scalability objectives. Options include keeping indexes and user data all in traditional memory (DRAM), all in Persistent Memory (PMem), all on SSDs, or in hybrid configurations (with indexes in DRAM or PMem and user data on SSDs).

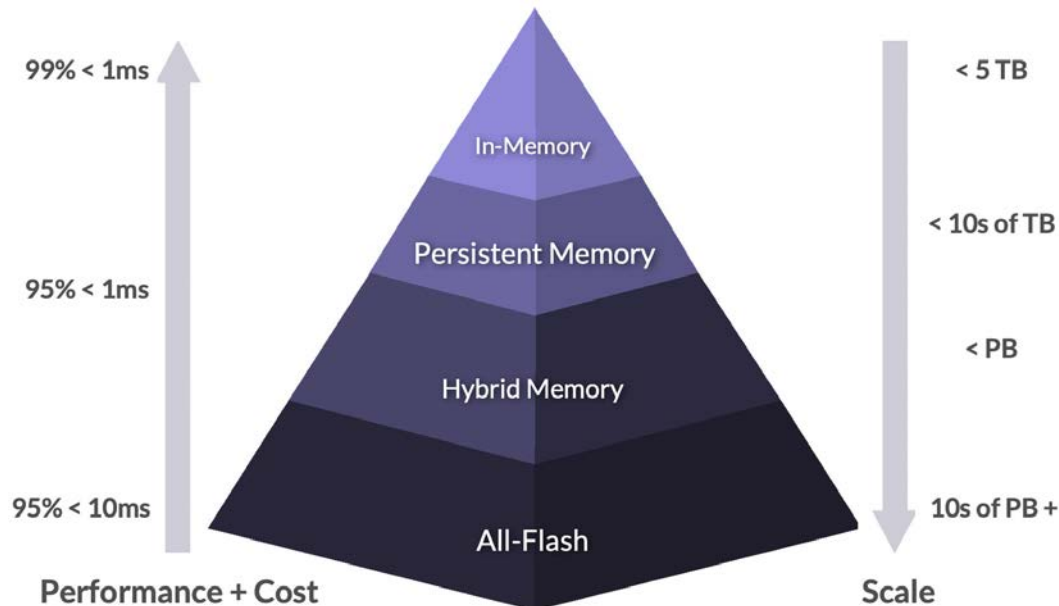


Figure 3: Tiered storage options with Aerospike provide optimal densities with high performance.

There is an advantage to having a single data platform that can execute in real time for various data set sizes: one codebase; no reprogramming; scale as you grow.

## Node reduction by up to 80 percent

In-memory DRAM can only store a fraction of the hot data that Aerospike Hybrid Memory Architecture™ SSDs store per node. Figure 4 illustrates this node reduction. It compares an In-Memory system using AWS Memory Optimized instance nodes versus an Aerospike HMA system using Storage Optimized instances. The critical difference is in the addressable space: In-memory only has 122 GiB/node, whereas Storage Optimized nodes have 3.2 terabytes of addressable space per node - a whopping difference.

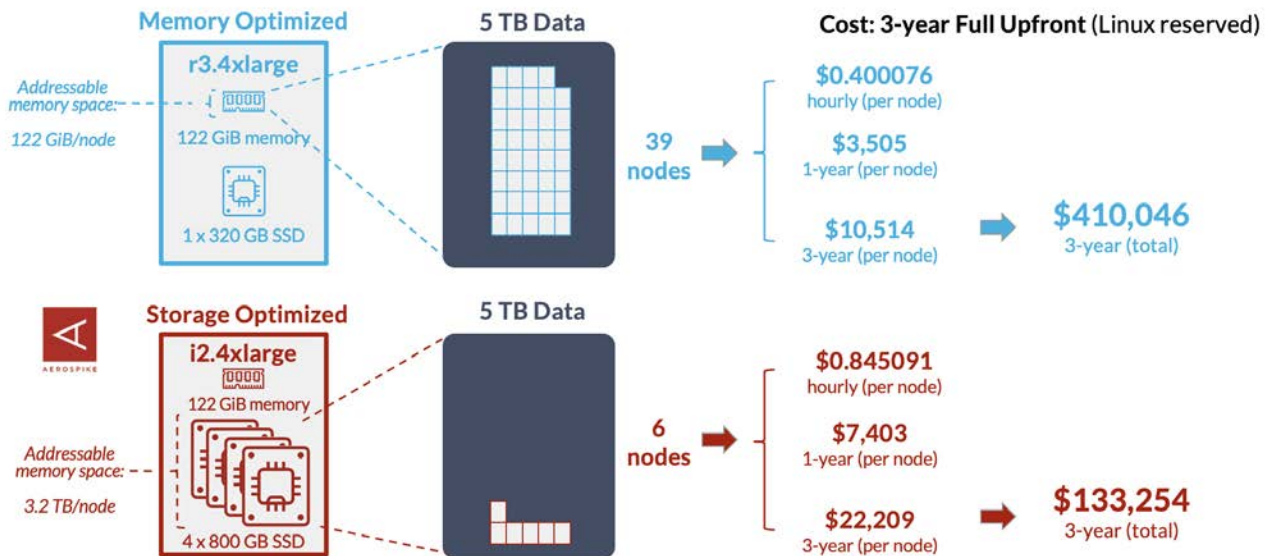


Figure 4: Node reduction by up to 80% using Aerospike HMA vs. In-memory (illustrated on AWS instances).

## Predictable performance from terabytes to petabytes

Your apps have to perform consistently well no matter the circumstances - lest they damage your brand. Enter Aerospike's Hybrid Memory Architecture™, designed to scale seamlessly and perform linearly at scale because it only needs to access indexes from volatile memory, not the data itself - unlike In-memory systems. No "cache misses," no associated spikes in latency. Aerospike Hybrid Memory Architecture has low variation and low response times versus In-memory systems (see Figure 5).

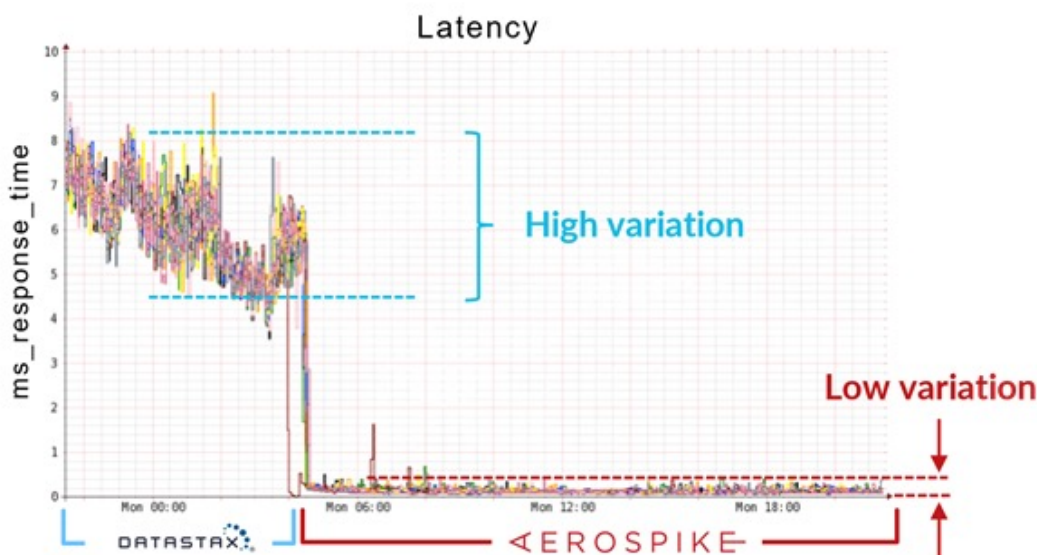


Figure 5: Latency variation comparison of an In-memory system (DataStax) vs. Aerospike



## Direct path to data (single hop)

Aerospike's Smart Client™ is a first-class participant in the architecture and data fabric and communicates with every node. Thus, every node in the Aerospike system knows what data every other node has and can reach data in a single hop. The result is consistent, predictable access times, whereas other technologies often require multiple jumps from node to node to find data - increasing latency.

## Hotspots avoided with intelligent auto-sharding

To avoid any single node receiving a disproportionate number of requests (thus making the system susceptible to either slowdowns or, worse, crashes), Aerospike automatically employs a deterministic, uniform data distribution scheme. As a result, Aerospike conducts load balancing continuously and automatically on all servers, even when scaling up or down or with cluster reconfigurations. The benefit: new use cases don't require re-tuning, and DevOps teams can reuse the same scheme and algorithms.

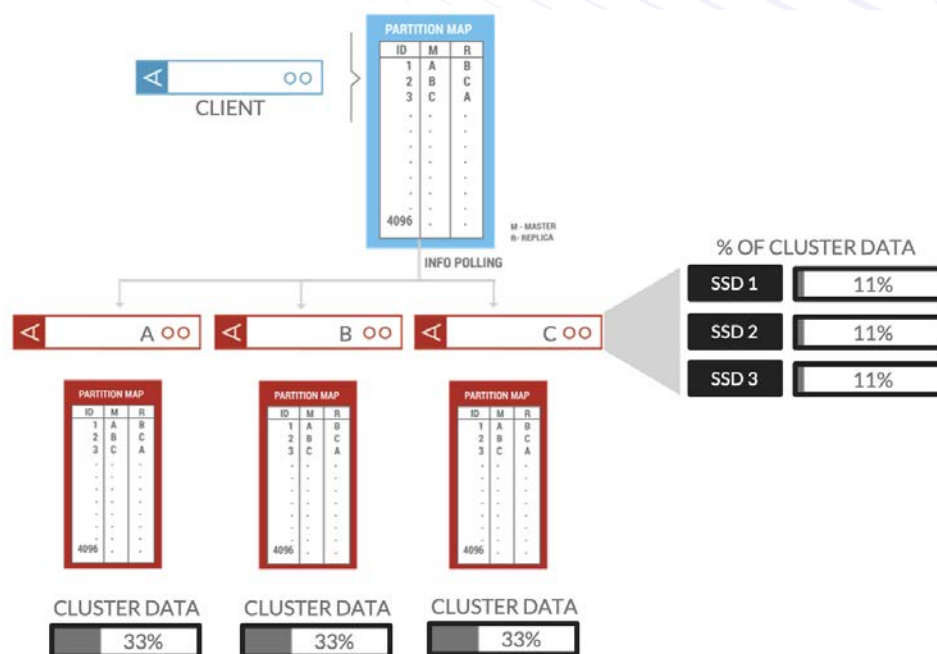


Figure 6: Aerospike avoids hotspots with intelligent auto-sharding with deterministic, uniform, algorithmic data distribution

## Node add/remove/update without disruption

Aerospike has a self-healing, auto-sharding, algorithmic cluster management system that adds, removes, or updates nodes without disruption/need to take the system down for maintenance. The result is high uptime as Aerospike has a “shared nothing” architecture, and there are no single points of failure, unlike other systems.

Figure 7 below illustrates adding a fourth node (Z), initiated by a “heartbeat” communication, rebalancing data (25% per node), and the resulting roster communicated up to the Smart Client.

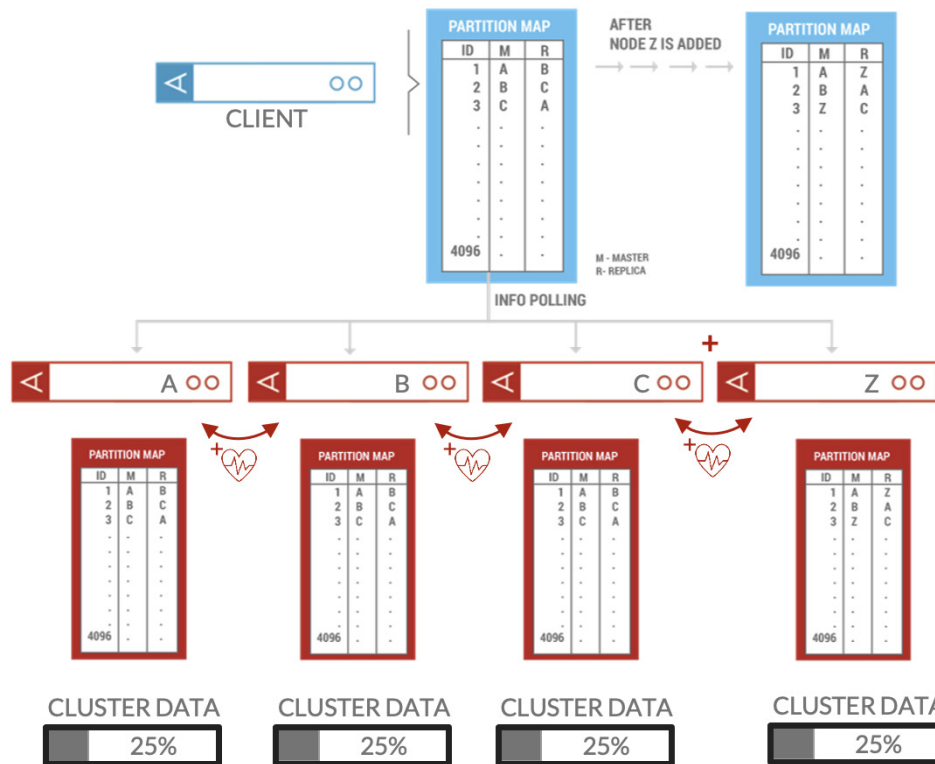


Figure 7: Aerospike's self-healing, auto-sharding, algorithmic cluster management enables nodes to be added, removed, or updated without disruption.

## High availability upon site failures without intervention

Aerospike employs a feature called “rack awareness,” which pegs copies of data (either a Master (M) or Replica (R) ) to specific racks across zones or data centers within a cluster for different hardware failure groups.

Figure 8 below exhibits how any single rack can go down, yet at least one copy of the entire data set comprised of Partition ID numbers 1, 2, and 3 are still available.

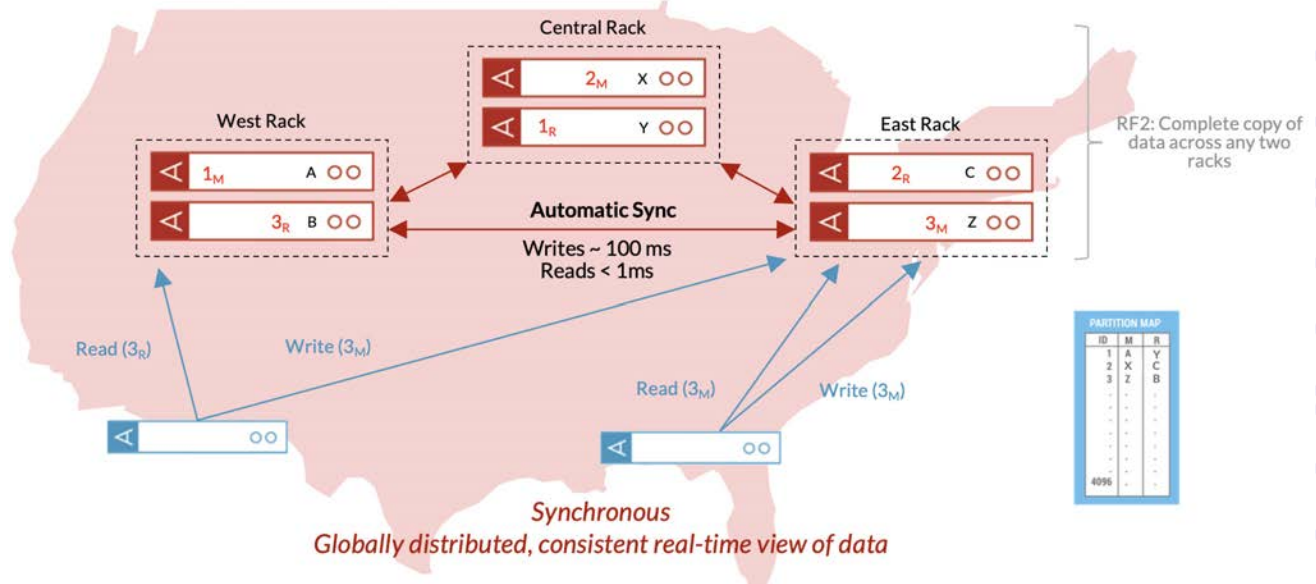


Figure 8: High availability across regions for globally distributed, strongly consistent data with Aerospike's rack awareness.

## Intercloud data replication with fine-grained control

Aerospike offers Cross-Datacenter Replication (XDR) with Expressions functionality that ships only the data you want, where you want it, asynchronously. The result is a high-performance global data hub with dynamic control for compliance needs, such as GDPR and CCPA, inventory, or logistics. In addition, Aerospike can route data captured at the Edge to non-Aerospike targets (e.g., ERP, CRM, or Inventory Systems via our Change Notification capability).

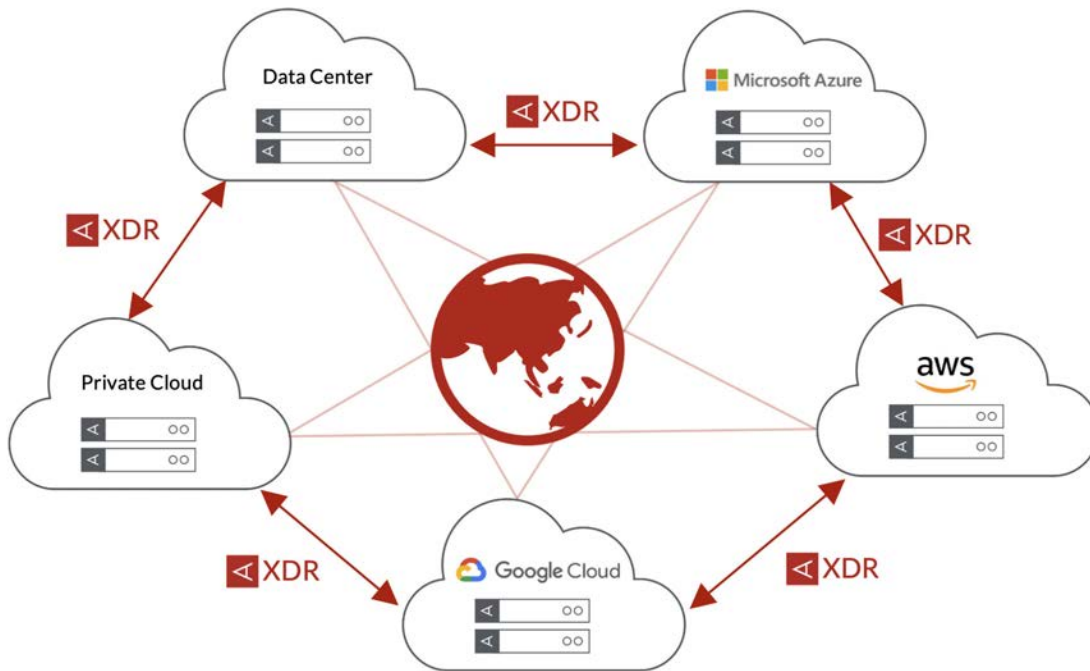


Figure 9: Aerospike XDR with fine-grained control provides true intercloud data replication, which ships only the data you want, where you want it, asynchronous, serving as a Global Data Hub.

## Real-time transactions without data loss

Aerospike provides strong consistency with high performance. For example, in case of a node failure or a network break to a node (known as a “partition”), it does not allow reads until that partition between two nodes gets healed, after which they can synchronize data, making the data consistent. Furthermore, no stale (old) nor dirty (non-consistent) reads are allowed.

Net-net: Data across Aerospike nodes are guaranteed to be consistent.

Aerospike’s high performance is maintained while making these consistency checks; an industry-independent third party ([Jepsen](#)) has validated the Aerospike Strong Consistency claims. (Note: Aerospike is the only NoSQL provider outside MongoDB to receive such validation.)

## Frequent use cases for the Aerospike Real-time Data Platform

Below is a quick synopsis of some of the deployments seen across the Aerospike customer base, cutting across industries:

<b>Caching</b> (Edge-only)	Content caching, legacy database augmentation: user session store, leaderboards, user spike management.
<b>Cache/ODS consolidation</b>	Replace caching layers built on operational data stores (ODSs), consolidating into a single real-time layer.
<b>Real-time data pipeline ingestion</b>	Ingest data from multiple sources at the Edge across clouds for IoT, customer 360, Recommendations, AdTech, Fraud.
<b>Mainframe Offloading</b>	Real-time systems that augment existing mainframes but sync with Systems of Record.
<b>Edge to System of Record</b> (Warm store)	Thaw, access, and push to the Edge in under 10ms data that would otherwise be archived for real-time applications at scale.
<b>Global real-time transaction processing</b>	Transmit strongly consistent data across regions at sub-millisecond speeds cost-effectively.
<b>Low latency store shared service</b>	Consolidate your streaming data into a single data store at petabyte scale with multi-tenancy and rate-limiting per application.



## Conclusion

When you examine your business and realize that business is won and lost nowadays in milliseconds, you have not only to ask yourself, “how do we gain competitive advantage?” but also, “how do we ensure we continuously thrive?” Therefore, your business should look to apply more data in real-time (i.e., sub-millisecond speed) and ensure it has the infrastructure to support these modern apps and the “plumbing” to interconnect it all.

Aerospike is a platform that was built from the ground up with the above requirements in mind. In addition, Aerospike is highly patented with proven benchmarks from companies like [AWS](#), Intel, and [HPE](#), with its business value validated by the likes of [Forrester Consulting](#).

Numerous industry-leading customers are building their future paths with the Aerospike Real-time Data Platform, such as [Adobe](#), [Wayfair](#), [Sony Interactive Entertainment](#), [The Trade Desk](#), [PayPal](#), and a [Top 3 Global Brokerage](#), to name a few.

We encourage you to give Aerospike a [try directly](#) or in [the cloud](#) or [reach out to us](#) now to see for yourself and learn more.

## 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, Nielsen, PayPal, Snap, Wayfair and Yahoo 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.

©2021 Aerospike, Inc. All rights reserved. Aerospike and the Aerospike logo are trademarks or registered trademarks of Aerospike. All other names and trademarks are for identification purposes and are the property of their respective owners.