SOLUTION BRIEF The Aerospike[®] Real-time Data Platform

The limits of what's possible with data can always be surpassed.

Overview

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

The Aerospike Real-time Data Platform enables organizations to act in real time across billions of transactions using massive parallelism and a hybrid memory storage model to ensure the smallest possible server footprint. It ingests and acts on streaming data at the edge and can combine edge data with data from systems of record, third party sources, data warehouses, or data lakes for operational, transactional, or analytical workloads - all in real time.

Highlights

- Predictable real-time performance at petabyte scale
- Cluster node reduction up to 80%
- No-code integrations with Spark, Kafka, Presto/Trino, Pulsar, JMS
- Fault and disaster tolerant
- Multi-site, multi-cloud clustering and data replication
- Real-time transactions with strong consistency for maximum accuracy



Figure 1: The Aerospike Real-time Data Platform architecture

The Aerospike Database real-time data engine

The Aerospike Database is the engine that drives the Aerospike Real-time Data Platform. It is responsible for all of the core cluster management and data services of the platform and is now a true multi-model database. It supports key-value and document database functions as well as graph database and time-series capabilities.



- Direct SSD device access
- Highly Parallelized
- Large Block Writes to SSD
- SSD vendor-optimized
- Continuous, non-disruptive defrag



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 3 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: Aerospike Data Model Implementations

Real-time performance with tiered storage

Aerospike has a patented Hybrid Memory Architecture[™] (HMA), which places data on solid-state drives (SSDs) and indexes in DRAM. 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 natively talk to hardware, not an API layer. Aerospike treats SSDs as a large parallel memory space, not a file system.



"Free is even too expensive compared to Aerospike." Jason Yanowitz, EVP and CTO, Signal

Figure 3: Node reduction up to 80% using HMA vs. In-memory

Multi-model data operations

SIGNA

The Aerospike Database can support simultaneous data models to simplify operations to suit multiple uses. Key value for fast lookups; document and object/JSON store for native data types; transactional and highly scalable graph data model for relationship traversals, and SQL access for querying.

✓ EROSPIKE

Modern Distributed Data Services

The data services of the Aerospike Real-time Data Platform are the internal processes of the Aerospike Database 6 that enable global operations, enterprise grade availability, sub-millisecond performance, and from gigabyte to terabyte scale.

Transaction Models

Aerospike's transaction engine supports strongly consistent transactions needed for modern applications including for financial services, supply chain, and commerce.

$\mathbf{E}_{\mathbf{Q}}$ Query and Search

Aerospike utilizes technologies such as set indexes, secondary indexes and Aerospike Expressions to power real-time query performance on data at the edge, in a System of Record, and across data centers and clouds. Massively parallel secondary indexes enable complex querying and SQL queries. Aerospike Expressions enable highly efficient value-based searches closer to the data.

ູ່ 🌐 ິ Dynamic Data Distribution

Aerospike provides dynamic data distribution via its shared-nothing cluster architecture, where all nodes are peers and there is no single point of failure. Using the Aerospike Smart Partitions[™], data is distributed evenly across all nodes in the cluster for maximum performance and scale.



Figure 5: Aerospike Data Services

😔 Change Data Notification

Aerospike's Change Notification feature enables outside databases and message queues to be updated in parallel with writes from database clients. This is a key usability feature when Aerospike is used as the database of record.

Smart Cluster Management

Aerospike's dynamic cluster management handles node membership, network fault detection, and ensures a single consistent view of current cluster members.



Figure 6: Some Aerospike customer milestones



Figure 7: Aerospike Real-time Data Platform benefits at a glance

✓ EROSPIKE

Aerospike Connect

No-code integrations with your data stack

Aerospike has created a series of connector products that provide tight integration with important components of a real-time data stack. Aerospike Connect products provide a direct and easy way to create real-time data pipelines with Spark, Kafka, Pulsar and more. These integrations require no special coding by developers and provide seamless SQL access and massively parallel data ingest for maximum throughput.



Figure 8: Aerospike Connect no-code integrations for streaming and SQL access



©2023 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.

aerospike.com