USING AEROSPIKE AS A REAL-TIME SYSTEM OF RECORD

Highlights

Strong consistency:
Data correctness is guaranteed, no data loss nor stale reads. Commit-to-device support for enhanced durability.

Scalability:
The All Flash and Hybrid Memory architectures allow the Aerospike database to scale to petabytes of data.

Speed:
Low latency is maintained at high scale, allowing Aerospike customers to make better decisions in real-time.

Superior Uptime and Availability:
Provides high availability and a demonstrated uptime of five 9s or more.

Ease of Deployment and Management:
Integral component of an end-to-end platform which is easy to deploy and manage in the cloud or on-premise.

Low TCO:
Fueled by a hybrid-memory architecture and compression, Aerospike provides significantly lower TCO than first-generation NoSQL and relational databases.

Overview

In this age of digital transformation, organizations are now required to make lightning-fast decisions powering applications like fraud detection, digital payments, ad bidding and recommendation engines. However, this cannot be accomplished with legacy data platforms as they cannot handle real-time data – let alone at scale, but it can be achieved using high-speed, extremely scalable and reliable modern data platforms. These modern data platforms consist of System of Engagement (SOE) databases that capture data from edge and near-edge devices, and System of Record (SOR) databases. SOR databases store both real-time data from the SOE database and historical data and act as a single source of truth. SOR databases, that host Real-time Analytics and Machine Learning (ML) applications also need to support real-time updates of the models used by these applications. This way the updates can be immediately synchronized with the SOEs for real-time decisioning.

In addition, SORs are also required to integrate data from legacy systems and to provide storage for petabytes of data.

Aerospike Hyperscale Data and Analytics Platform

Aerospike provides unmatched performance at scale for Systems of Record by utilizing an end-to-end platform (Figure 1). The platform consists of the:

1 - **Aerospike Edge (SOE) Database** - Used for real-time decisioning based on local transactional data plus historical data pulled dynamically from the SOR.
2 - **Aerospike Real-time System of Record (SOR) Database** - Stores transactional and historical data and pushes data as needed to the SOEs also powering ML and AI-based applications.
3 - **Aerospike Query and Reporting Database** – Stores historical data primarily for reporting and visualization purposes, integrated via Aerospike Connect for Spark.
4 – **Legacy Data Store** – While not an Aerospike technology, it is an important data source for the SOR, integrated via the Aerospike Connect products.
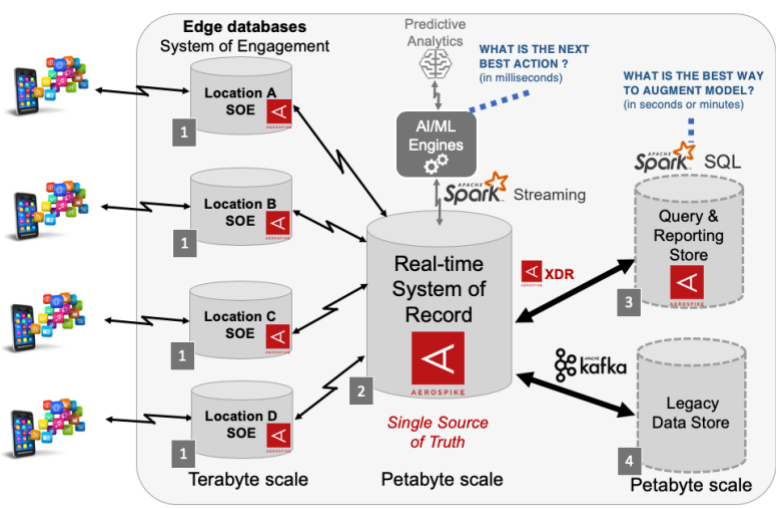
Aerospike Edge Databases:

Edge databases can reside in different geographic locations and are connected via high-speed data transfer with the SOR. This can be accomplished by utilizing different built-in features such as Aerospike Cross Datacenter Replication (XDR), and Aerospike Connect for Kafka.

Aerospike Real-time SOR Database:

The Aerospike Real-time System of Record database is the single source of truth that propagates subsets of data to the Edge databases in real-time. Data in the SOR can be available for Real-time Analytics and ML systems. The models for these systems can be updated in real-time and the model updates can be synchronized frequently with the SOE for real-time decisioning. This reduces the time for models updates from days/hours to minutes/seconds.

Aerospike Query and Reporting Database:

Connected via high speed data transfer with the SOR, it can store petabytes of data for reporting and visualization purposes.

Figure 1. Aerospike's Hyperscale Data and Analytics Platform

Meeting Enterprise Data Management Requirements

Aerospike's Real-time System of Record database meets millisecond response times at scales of tens of petabytes for the strictest enterprise requirements:

Consistency - Aerospike provides strong consistency on primary key access that has been confirmed through Jepsen test results. Data held in Aerospike is always guaranteed to be correct in all scenarios.

Durability - Data can be replicated asynchronously across geographies and synchronously written to nodes in a cluster and to flash storage for the highest durability.

Scalability - The Aerospike Hybrid Memory Architecture™ with All Flash and Hybrid Flash options allows the Aerospike database to scale to petabytes and store both transactional as well as historical data.

Uptime and High Availability – Aerospike provides high availability and a demonstrated uptime of five 9s or more which is made possible by a unique cluster management and intelligent client technology in addition to local/remote replication.

Speed – Parallel processing and our patented Hybrid Memory Architecture designed for flash storage devices uniquely unlock flash's high performance at scale.

Real-Time Data Transfer Between SOR and SOE – Aerospike XDR enables multiple geographically dispersed data centers to stay in sync through high performance replication.

Enterprise-level Security - Aerospike supports full transport encryption, authentication, access control, exception logging, as well as in-database transparent data encryption.

Integration with Existing Data Stores and Systems - Aerospike Connect for Spark and Kafka allow SQL databases, NoSQL databases and ML-based tools to integrate seamlessly and efficiently to other data stores and systems.

Compression - Aerospike's storage compression feature provides lossless compression of records written to persistent storage.

Support for Next-generation Memory – Aerospike is the first open database supporting the new Intel® Optane™ DC persistent memory.

Rich Set of Deployment Options - Aerospike can be deployed close to the edge, in data centers and in the Cloud – Google Compute Platform, Amazon Web Services, Microsoft Azure, Alibaba Cloud and others.

Benefits

Better/Faster Decisions for Competitive Advantage - Utilizing Aerospike's technology, better real-time decisions can be made since higher volumes of historical data can be combined in real-time with transactional data. Data for the decisioning process is available in milliseconds, which cannot be achieved at scale with any other single data platform.

Ease of Deployment and Management – As a pivotal component of an end-to-end solution, Aerospike can be easily deployed and managed in the cloud and on-premises.

Can Replace Legacy SOR Databases - Due to its ease of management, uptime and high availability, it can replace Cassandra, HBase and other legacy systems that are difficult to maintain and scale.

Enables Business Growth - Aerospike's performance at scale means there is no need for customers to switch platforms as their data volumes grow.

Low TCO - Fueled by our Hybrid Memory Architecture, dynamic cluster management and compression, Aerospike provides improved performance and dramatic reduction in node count. This allows a significantly lower TCO. Signal, a Customer Identity Management provider, saved $4 million over three years by replacing their 450 node Cassandra deployment.

Typical Use Cases

Financial Services and FinTech: Edge-to-core data movement for real-time analysis for fraud prevention, risk assessment and compliance.

e-Commerce and Retail / CPG: Customer behavior data integration with clickstream integration plus product data for better personalization and improved customer experience.

Telco: Customer data integration with real-time billing plus Customer 360.

Ad Tech: Real-time clickstream data synchronization and integration with core data for more on-target ads.