

Aerospike® for Query and Search

SOLUTION BRIEF

Enterprise access for real-time data

Overview

Increasingly, organizations depend both on the real-time data that powers their operational applications as well as the need to provide access to that data for a range of machine learning and analytics uses. A common and inefficient approach is to copy operational data into a separate analytics repository - be it for direct analytics purposes, or into a separate data warehouse, lake, or lakehouse.

This is often appropriate, but there are times when organizations benefit from accessing real-time data in place, minimizing delays plus averting additional data copies and the associated downsides of such as keeping that data in compliance, up-to-date, or even consistent. Aerospike's query and search capabilities open up these analytic uses, powered by both native capabilities such as secondary indexes and expressions, as well as via industry-leading tools Elasticsearch, Presto/Trino and Starburst for SQL, and Apache Spark™.

Benefits

- **Accelerate Time to Insight:** Combine the massive parallelism of Aerospike with native secondary indexes, as well employing native SQL queries via Aerospike SQL or Presto/Trino to act upon real-time data.
- **Lower TCO:** Analyze massive datasets with a smaller storage footprint and lower total cost of ownership (TCO) than alternative approaches.
- **Easy BI Integration:** Analyze data stored in Aerospike via your favorite Business Intelligence (BI) tools, including Tableau, Qlik, and Looker.

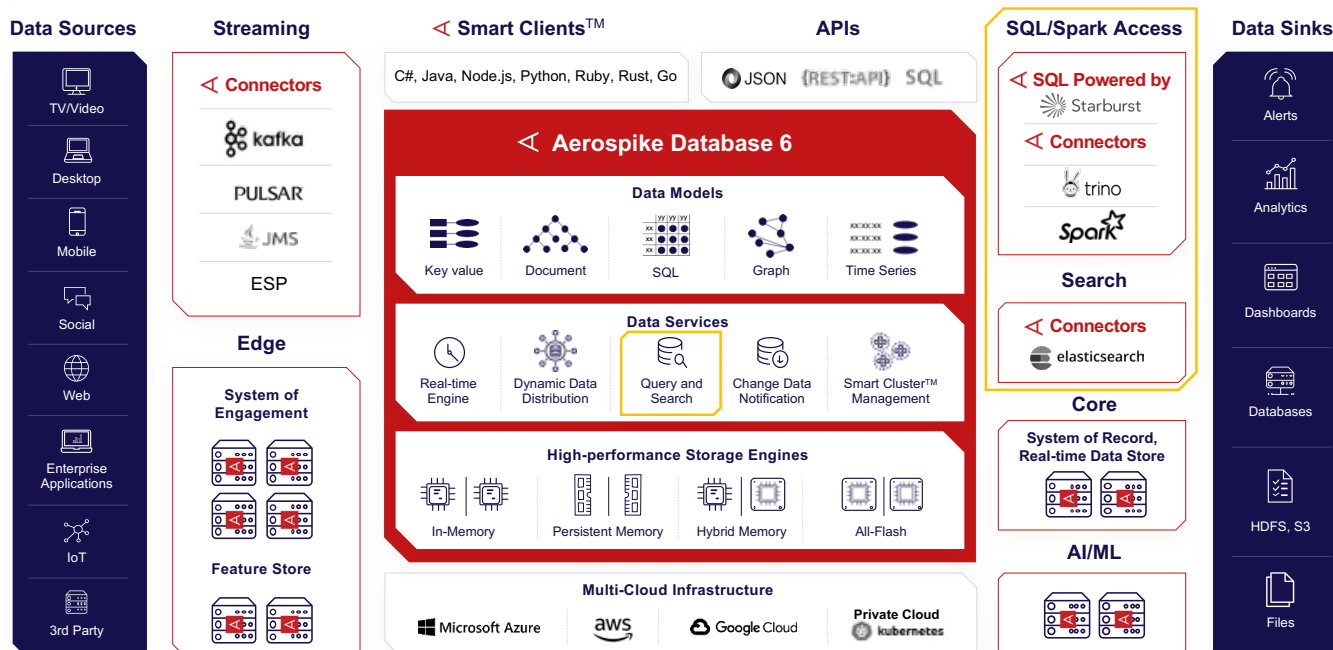


Figure 1: Query and search access patterns in the Aerospike Real-time Data Platform

Aerospike query and search access

Aerospike is a multi-model, real-time, operational database and data platform that powers mission critical applications that require constant uptime and low latency. Aerospike works with organizations' broad data infrastructures ensuring real-time data in Aerospike is readily accessible to those who need it, without impacting operational performance.

With Aerospike query and search capabilities, organizations can seamlessly integrate Aerospike data into their wider data landscape. This can save hundreds of hours that these companies would otherwise spend on custom integration work. It accelerates time to insights and facilitates new search and analytics use cases with tools that developers already use.

Key Features

- No-code integrations for best-in-class tools.
- Gigabyte-to-petabyte scale search and SQL queries.
- Aerospike secondary indexes manage memory efficiently, accelerating query results.
- Separate compute and storage needs, enabling each to scale independently to reduce costs.
- Fast and efficient value-based searches to query and manipulate record metadata and data.

How it works

Aerospike-native queries

Aerospike allows value-based queries using secondary indexes, where string and integer values are indexed and searched using filters. The Aerospike Expressions language enhances query efficiency through filters (including Boolean logic), aggregations, and more. The aggregations framework allows fast, flexible query operations in a highly parallel fashion.

SQL access

Aerospike is a real-time operational database that enables SQL access in multiple ways:



Aerospike Connect for Trino

Aerospike pre-built connector for Trino (previously called PrestoSQL), enables access to a leading open source tool for distributed SQL. As a result, Aerospike with Aerospike Connect for Trino can run federated queries across multiple databases, handle hundreds of concurrent queries, and scale to thousands of compute nodes.



Aerospike SQL powered by Starburst

Powered by Starburst Enterprise, Aerospike SQL provides an enterprise-grade implementation of the open source Trino project (formerly PrestoSQL) that is installed, managed and supported by Aerospike. It provides fast time to insight, secure SQL analytics at scale, and federated access to data in place while reducing complexity.



Aerospike Connect for Spark

Data engineers and data scientists run ETL jobs, exploratory data analytics, and AI workloads across massive data volumes on Spark. With Aerospike Connect for Spark, Aerospike is ideally suited as the massively parallel data storage for Spark.

Search

Aerospike enables developers to create powerful search applications for data in Aerospike clusters.



elasticsearch

Aerospike Connect for Elasticsearch

Enables developers to create powerful search applications for data in Aerospike clusters. A subset of the Aerospike data is moved to an Elasticsearch index. Search applications use Elasticsearch indexes and so do not impact Aerospike transactions, throughput, or latencies. Aerospike efficiently populates Elasticsearch indexes using Aerospike Change Notification / change data capture (CDC) technology and Aerospike's Cross Datacenter Replication (XDR). Search results become efficiently and rapidly enriched, pulling complete information from Aerospike to Elasticsearch also via CDC.

Use Cases



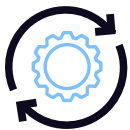
Customer insights

Incorporate user behavior including real-time purchase and social data for ML-supported personalization apps. Search user behavior based on documents and other textual data.



Analytics and Machine Learning

Provide interactive query access through business intelligence tools. Include massive data volumes with fast retrieval in machine learning (ML) models for predictions and other advanced analytics.



Products and operations

Improve the accuracy and relevance of products and services with analytics and search, including observability with textual log data and JSON files.

AEROSPIKE

Aerospike unleashes the power of real-time data to meet the demands of The Right Now Economy. Global innovators and builders choose the Aerospike real-time, multi-model, NoSQL operational database and data platform for its predictable sub-millisecond performance at unlimited scale with dramatically reduced infrastructure costs. With support for strong consistency and globally distributed, multi-cloud environments, Aerospike is an essential part of the modern data stack for Adobe, Airtel, Criteo, DBS Bank, Experian, PayPal, Snap, Sony Interactive Entertainment, The Trade Desk, and Wayfair. A global company, Aerospike is headquartered in Mountain View, California, with offices in London, Bangalore, and Tel Aviv.

©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.