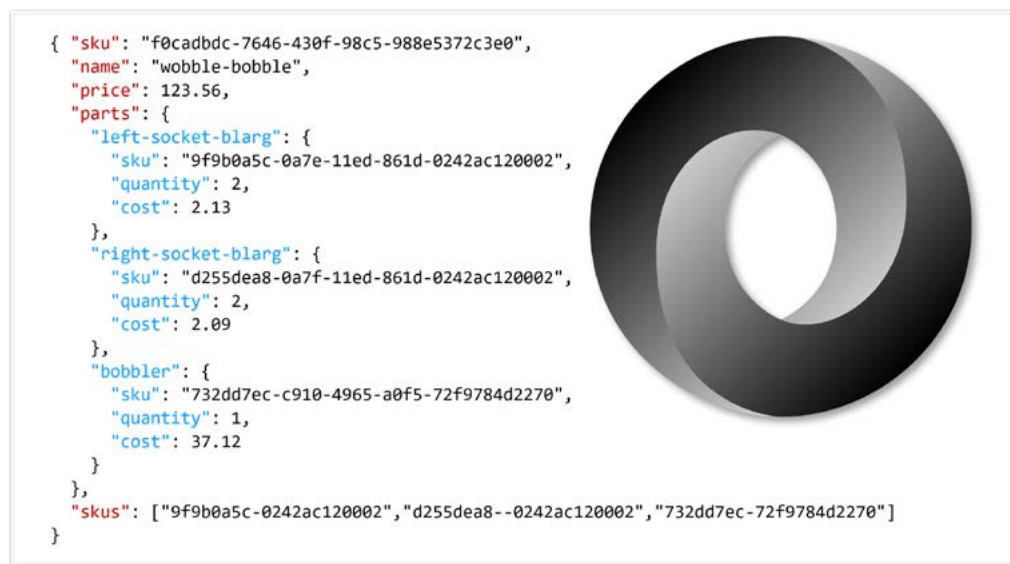# Aerospike® Document Database

SOLUTION BRIEF

High performance document database powered by the Aerospike Real-time Data Platform

Aerospike adds real-time document database capabilities to the Aerospike Database that can scale from gigabytes to petabytes. It features support for JSON document models and Java programming models via the Aerospike Document API, including JSONPath query support to store, search, and manage large scale document datasets and workloads.

The Aerospike Database has long supported storing and accessing document data programmatically through the Collection Data Types introduced in Aerospike Database 4. Aerospike Database 6 now has an integrated Document API, JSONPath query capabilities, and support for secondary indexes even on nested elements of rich documents for unprecedented real-time performance and unlimited scale.


```
{ "sku": "f0cadbdc-7646-430f-98c5-988e5372c3e0",
  "name": "wobble-bobble",
  "price": 123.56,
  "parts": {
      "left-socket-blarg": {
          "sku": "9f9b0a5c-0a7e-11ed-861d-0242ac120002",
          "quantity": 2,
          "cost": 2.13
      },
      "right-socket-blarg": {
          "sku": "d255dea8-0a7f-11ed-861d-0242ac120002",
          "quantity": 2,
          "cost": 2.09
      },
      "bobbler": {
          "sku": "732dd7ec-c910-4965-a0f5-72f9784d2270",
          "quantity": 1,
          "cost": 37.12
      }
  },
  "skus": ["9f9b0a5c-0242ac120002","d255dea8--0242ac120002","732dd7ec-72f9784d2270"]
}
```

Figure 1: High speed access to JSON document data

Key Features

Aerospike is a complete gigabyte-to-petabyte scale Document Database solution for real-time enterprise analytics and data applications:

- Real-time performance at unlimited scale for document data using binary representation of JSON (MessagePack).

- JSONPath based Document API offers a convenient way to access and modify specific elements within a document.

- Accelerated query performance using Aerospike capabilities like massive parallelism, predicate pushdown, and secondary indexes.

- Enables globally distributed document stores with local optimizations using Cross Datacenter Replication (XDR).

- Index nested elements for documents for greater scale and performance.

- Takes advantage of Aerospike's enterprise grade security including:
    - TLS for secure connectivity
    - LDAP authentication
    - PKI support for password-less authentication over TLS.
    - Avoid excessive usage via server quotas.

- Deploy in the cloud (AWS, Azure and GCP) or on premises.

Aerospike Document Database Overview

For a true real-time document database, the underlying data platform must provide for fast ingest, efficient storage, and powerful query capabilities while providing fast response times at any scale.

Document ingest, storage, and query

The Aerospike Database supports fast ingestion rates so that surges in real-time data feeds do not overwhelm the system or result in data loss. In Aerospike 6, batch operations for read, write, delete, and UDF operations are supported so that ingest can achieve very high throughput rates.

Aerospike's CDT (Collection Data Type) API allows document structures such as JSON to be modeled by composing list and map types. The resulting aggregate CDT structures are stored and transferred efficiently using the binary MessagePack format. This structured approach allows swifter traversal and modification than text based representations.

Aerospike handles concurrent queries of large datasets efficiently via indexes and granular control over parallel processing of queries. Aerospike uses a JSONPath-based Document API as a convenient way to access and modify specific elements within a document.

Document Indexing and other performance and scalability features

Reliably fast response time for read and write operations at any scale and any read-write workload mix is required to meet the real-time SLAs. Aerospike delivers through:

- Fast and uniform hash-based data distribution to all nodes for optimal resource utilization
- Hybrid Memory Architecture^{tm} utilizes DRAM, SSDs, and flash technologies to store data and indexes
- Optimized processing of writes and garbage collection for predictable response
- One-hop access to all data from the application
- Primary and secondary indexes for fast access
- Async and background processing modes for greater efficiency

Use Cases

- Web apps: tiered web applications interoperate using JSON data. Many will follow a three tiered approach with a REST based, JSON yielding middle tier
- Analyzing real-time FinServ JSON datasets
- Market data providers provide vital retail finance data such as credit check reports or employment history as JSON.

- Processing real-time device data from IoT architectures where edge devices send message payload as JSON data
- Integrating real-time social media feed data from platforms such as YouTube, Twitter and LinkedIn who expose data to developers via APIs that use JSON as the default data format

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 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.

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