Oracle NoSQL Database
Enterprise Edition, Version 20.1

ORACLE®

NOSQL DATABASE

KEY BUSINESS BENEFITS

Multi-region deployment

Hybrid and multi-cloud deployment

High throughput

Bounded latency

Near-linear scalability

High availability

Short time to deployment

Smart topology management

Online elastic configuration

Enterprise grade software and support

Oracle NoSQL Database is a multi-model, multi-region database, designed to provide a highly-available, scalable, performant, flexible, and reliable data management solution to meet today's most demanding workloads. It can be deployed in on-premise data centers, or in private, public, hybrid and multi-clouds. It is well-suited for high volume and velocity workloads, like Internet of Things, customer 360, online contextual advertising, fraud detection, mobile application, user personalization, and online gaming. Developers can use a single application interface to quickly build applications that run in on-premise and cloud environments.

Applications send network requests against an Oracle NoSQL Database data store to perform data operations. With multi-master replication, data can be globally distributed and automatically replicated in real-time across different regions. Data can be modeled as fixed-schema tables, documents, key-value pairs, and large objects. Different data models interoperate with each other through a single programming interface.

Oracle NoSQL Database is a sharded, shared-nothing system which distributes data uniformly across multiple shards in a NoSQL database cluster, based on the hashed value of the primary keys. An Oracle NoSQL Database data store is a collection of storage nodes, each of which hosts one or more replication nodes. Data is automatically populated across these replication nodes by internal replication mechanisms to ensure high availability and rapid failover in the event of a node failure.
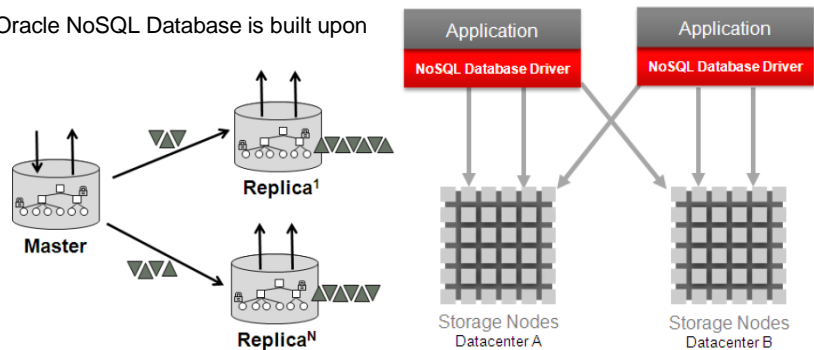
Oracle NoSQL Database provides Java, Python, Node.js, Go, C#, and C drivers and REST API for application development. It is also readily integrated with other Oracle products and open source applications to simplify and streamline application development. Oracle NoSQL Database is dual-licensed and available as an open-source Apache licensed Community Edition as well as a commercially licensed Enterprise Edition.

Architecture

The Oracle NoSQL Database is built upon

the proven Oracle Berkeley DB Java Edition high-availability storage engine, which is in widespread use in enterprises across industries. In addition, it adds a layer of services for use in distributed environments.

The resulting solution provides distributed, highly available key/value storage that is well suited to large-volume, latency-sensitive applications.

High Availability and No Single Point of Failure

Each shard in the Oracle NoSQL Database provides dynamically elected leader nodes (masters) and multi-replica database replication. Transactional data is delivered to all replica nodes in the shard with flexible durability policies per transaction.   In the event the master replica node of a shard fails, a PAXOS-based automated fail-over election provides a new shard master with minimal affects to write latency. This allows for scalability, high-availability, and low latency read and write operations.

High Performance

The Oracle NoSQL Database is network topology and latency aware. The Oracle NoSQL Database Driver working in conjunction with highly scalable, fault tolerant, high throughput storage engine enables a more granular distribution of resources and processing, which reduces the incidence of hot spots and provides greater performance on commodity based hardware.

Transparent Load Balancing

The Oracle NoSQL Database Driver partitions the data in real time and evenly distributes it across the storage nodes. It is network topology and latency-aware, routing read and write operations to the most appropriate storage node in order to optimize load distribution and performance.

Configurable Smart Topology

System administrators indicate how much capacity is available on a given storage node, allowing more capable storage nodes to host multiple replication nodes. Once the system knows about the capacity for the storage nodes in a configuration, it automatically allocates replication nodes intelligently. This results better load balancing for the system, better use of system resources and minimizing system impact in the

event of storage node failure. Smart Topology also supports Data Centers, ensuring that a full set of replicas is initially allocated to each data center.

Elastic Configuration

The Oracle NoSQL Database includes a topology planning feature, with which an administrator can now modify the configuration of a NoSQL database, while the database is still online. This allows the administrator to:

- **Increase Data Distribution**: by increasing number of shards in the cluster, this increases the write throughput.

- **Increase Replication Factor:** by assigning additional replication nodes to each shard, which increases read throughput and system availability.

- **Rebalance Data Store:** by modifying the capacity of a storage node(s), the system can be rebalanced, re-allocating replication nodes to the available storage nodes, as appropriate. The topology rebalance command allows the administrator to move replication nodes and/or partitions from over utilized nodes onto underutilized storage nodes or vice versa.

Easy Administration and Enhanced System Monitoring

The Oracle NoSQL Database provides an administration service, which can be accessed from command-line interface (CLI) interface. This service supports core functionality such as the ability to configure, start, stop and monitor a storage node, without requiring manual effort with configuration files, shell scripts, or explicit database operations.

In addition it also allows Java Management Extensions (JMX agents to be available for monitoring. This allows management clients to poll information about the status, performance metrics and operational parameters of the storage node and its managed services

Arbiters

The ability to reduce hardware requirements using fewer replicas per shard instance.

Online Rolling Upgrade

Upgrade and patching is an important part of any software support cycle. The Oracle NoSQL Database provides facilities to perform a rolling upgrade, allowing a system administrator to upgrade all of the nodes in the Oracle NoSQL Database cluster while the database continues to remain online and available to clients.

Multi-Region Table

Oracle NoSQL Database multi-region architecture enables the creation of Multi-Region Tables in multiple KVStore clusters and maintains consistent data across these clusters. The multi-region deployment may be geographically distributed. A Multi-Region Table is a global, logical table that is deployed and maintained in different regions. It is a read-anywhere and write-anywhere table that lives in multiple regions. All Multi-Region

Tables deployed in different regions are synchronized via NoSQL Streams. Each region KVStore may have different topology, different hardware configuration, different operations and security management.

Multi-Zone Deployment

Oracle NoSQL Database supports the definition of multiple zones from within the topology deployment planner. It leverages the definition of these zones internally to intelligently allocate replication of processes and data, ensuring optimal reliability during hardware, network & power related failure scenarios.

There are two types of Zones: Primary zones contain nodes that can be served as masters or replicas and are typically connected by fast interconnects. Secondary zones contain nodes which can only be served as replicas. Secondary zones can be used to provide low latency read access to data at a distant location, or to offload read-only workloads, like analytics, report generation, and data exchange for improved workload management. The Oracle NoSQL Database allows users to continue business operations in the event of zone failures. This allows for any planned maintenance that results in the taking of one or more zones offline without impacting business operations. Additionally, with the zone affinity feature it's possible to place master nodes in primary zones that are in close network proximity to the user applications. This helps to get predictable write latencies.

Single Application Programming Interface

HTTP proxy is a new middle tier component that sits between the client applications and the NoSQL Database server. The HTTP protocols are identical for on-premise Oracle NoSQL Database and Oracle NoSQL Database Cloud Service. Client applications can connect and move between both products easily. With a single common application programming interface, developers can easily build applications that run and interoperate in a hybrid cloud environment.

Table Data Model

A tabular data structure is available, which simplifies application data modeling by leveraging existing schema design core concepts. Table model is layered on top of the distributed key-value structure, inheriting all its advantages and simplifying application design even further by enabling seamless integration with familiar SQL-based applications.

Native JSON Data Type

JSON is a first-class citizen making it easy to store data that doesn't confirm to rigid schema. Only valid JSON documents can be stored providing automatic JSON document validation. JSON documents stored in JSON columns are converted to an internal binary (optimized) format that allows quick read access to document elements. The ability to create JSON indexes on JSON column allows developers access to the nested attributes embedded deep within a JSON document.

Secondary Index

Primary key only based indexing limits number of low latency access paths. Sometime application needs a few non-primary-key based paths to support the whole solution for the real-time system. Being able to define secondary index on any value field dramatically improves performance for queries.

SQL for NoSQL

Oracle NoSQL Database provides a SQL-like interface to Oracle NoSQL Database that can be used from a command line interface, scripts, or from the Oracle NoSQL Database Java Table Driver. The SQL for Oracle NoSQL Database data model supports flat relational data, hierarchical typed (schema-full) data, and schema-less JSON data. SQL for Oracle NoSQL Database is designed to handle all such data in a seamless fashion without any "impedance mismatch" among the different sub models. SQL Path Expression allows to navigate inside complex values and select their nested values using different types of step operations. For JSOC docs, it is possible to "introspect" the JSON docs, in case you don't know what is in there.

JSON Indexing and Query

Create indexes on JSON columns to access the nested JSON attributes efficiently. Query your JSON data type with familiar SQL queries. This powerful feature gives developers the ability to use SQL to query schemaless JSON data. NoSQL now offers the flexibility of rich query over schemaless data along -side more structured queries.

Partial JSON Update

Developers can update (change, add, remove) a part of JSON document. This update happens on the server side eliminating the need for *read-modify-write* cycle, is atomic and thread safe.

GeoJSON Support

Data can be stored in GeoJSON format to represent geographical features, properties, and boundaries. Geometry types supported are Point, LineString, Polygon, MultiPoint, MultiLineString, MultiPolygon, and GeometryCollection. Search functions support queries on geographical data that have a relationship based on a certain geometry. Indexes can be created for optimal search performance.

Parent-Child Joins

Oracle NoSQL Database includes support for a special kind of join among tables that belong to the same table hierarchy. This is implemented with a NESTED TABLES clause that is semantically equivalent to the LEFT-OUTER-JOINS defined by standard SQL and supported by all RDBMS implementations. The Left Outer Join creates a result set containing pairs of matching rows from the left and right tables and you would see a similar behavior in Oracle NoSQL.

Aggregation Functions

Aggregate functions in Oracle NoSQL Database iterate over the rows, evaluate an expression for each row, and aggregate the returned values into a single value. Syntactically Aggregate functions appear in the SELECT clause. Supported Aggregate functions supported are: sum, count, avg, min, max.

Simple and Easy to Use APIs in Multiple Programming Languages

The Oracle NoSQL Database includes **Java, C, Python, C# and Node.js APIs**. These simple APIs allow the application developer to perform CRUD operations on the Oracle NoSQL Database. C and Java drivers, also include Avro support, so that developers can serialize key-value records and de-serialize key-value records interchangeably between C and Java applications.

Full Text Search (FTS)

Gives users the ability to perform very secure fast text and indexed searches on data stored in Oracle NoSQL Database. FTS combines the TABLE interface with ElasticSearch (ES) for a powerful way to find documents that satisfy a query. This provides a high performant, secure full-text search of Tables stored in Oracle NoSQL Database.

Streams Processing

Based on Reactive Streams, streams processing in Oracle NoSQL Database allow for a notification service that permits a user to subscribe to all logical changes (table row puts and deletes) made to an Oracle NoSQL Database store. Applications can be alerted to these changes which allows for asynchronous monitoring of database changes.

Time-To-Live

Allows for data to be stored for a specified period of time and then deleted automatically which is a critical requirement for sensor data capture in an Internet Of Things (IoT) service.

Oracle Database Integration via External Tables

Support for external table allows fetching from the Oracle NoSQL Database data from Oracle database using SQL statements such as Select, Select Count(*) etc. Once data from the Oracle NoSQL Database is exposed through external tables, one can access the data via standard JDBC drivers and/or visualize it through enterprise Business Intelligence tools.

Oracle Big Data SQL and Hive Integration

Oracle Big Data SQL is a common SQL access layer to data stored in Hadoop, HDFS, Hive and the Oracle NoSQL Database. This allows customers to run queries on the Oracle NoSQL Database from a Hive or Oracle Database. Users can also run MapReduce jobs against data stored in the Oracle NoSQL Database that's configured

for secure access. The latest release also supports both primitive and complex data types.

Integration with Other Oracle Products

The integration of Oracle NoSQL Database with OEM primarily takes the form of an EM plug-in. The plug-in allows monitoring through Enterprise Manager of NoSQL Database store components, their availability, performance metrics, and operational parameters.

With Oracle SQL Developer integration it's possible to view (Read-only) the data that's stored in Oracle NoSQL Database.

Oracle Event Processing (OEP) provides read access to the Oracle NoSQL Database via the Oracle NoSQL Database cartridge. Once the cartridge is configured, CQL queries can be used to query the data.

The Oracle Semantic Graph has developed a Jena Adapter for the Oracle NoSQL Database to store large volumes of RDF data (as triplets/quadruplets). This adapter enables fast access to graph data stored in the Oracle NoSQL Database via SPARQL queries. An integration with Oracle Coherence has been provided that allows the Oracle NoSQL Database to be used as a cache for Oracle Coherence applications, also allowing applications to directly access cached data from the Oracle NoSQL Database.

Large Object Support

Stream based APIs are provided in the product to read and write Large Objects (LOBs) such as audio and video files, without having to materialize the value in its entirety in memory. This permits low latency operations across mixed workloads of objects of varying sizes.

Apache Hadoop Integration

KVAvroInputFormat and KVInputFormat classes are available to read data from the Oracle NoSQL Database natively into Hadoop Map/Reduce jobs. One use for this class is to read the Oracle NoSQL Database records into Oracle Loader for Hadoop.

Import/Export Capabilities

Data can be moved to and from Oracle NoSQL Database using a simple, data exchange format.

Enterprise Security

OS- OS-independent, cluster-wide password-based user authentication and Oracle Wallet integration, enables greater protection from unauthorized access to sensitive data. Additionally, session-level Secure Sockets Layer (SSL) encryption and network port restrictions deliver greater protection from network intrusion. Oracle NoSQL Database now can use Kerberos integration for external authentication. This allows for Oracle NoSQL Database to be easily integrated with customers' existing application which are already protected by Kerberos.

Commercial Grade Software and Support

The Oracle NoSQL Database overcomes a significant limitation faced by many enterprises considering the implementation of NoSQL databases—the need for full supportability. The Oracle NoSQL Database is a commercial product fully supported by Oracle. This gives organizations the confidence and reduces the risk they need to deploy the Oracle NoSQL Database in the production environments they depend on to manage their business-critical data.