

Bases II GR 1 - Resumen 4 - 11/04/2023

Gerald Núñez Chavarría - 2021023226

Mongo DB Architecture Guide

1. Document model

Mongo DB retains the best aspects of SQL and NoSQL databases, saving the information in JSON and making a structure of documents.

The three principal benefits of the document data model are: 1- Documents in the database map to the objects in your code. JSON information is very understandable and easier to manage. 2- Mongo DB has a dynamic document schema, it allow that fields can vary from document to document. It also offers schema validation, very useful when your applications move into production because you can govern your schema without having to write controls in the application layer. 3- MongoDB stores data as JSON documents in a binary representation called BSON, to include additional types such as int, long, date...

2. Working with document data

The MongoDB Query API offers a comprehensive and expressive way to access, group, transform, and analyze data, and can federate queries across databases with a single connection string. The MongoDB Aggregation Pipeline allows for efficient data aggregation using native operations, and features such as Change Streams and strong data consistency guarantees make it easy to build reactive event-driven data pipelines. ACID transactions ensure all-or-nothing and snapshot isolation guarantees, even with MongoDB's scale-out architecture. Native drivers are provided for popular programming languages, eliminating the need for ORM abstraction layers.

3. Distributed Architecture: Scalable, Resilient, and Mission Critical

MongoDB allows for scaling out applications with always-on availability through replica sets and native sharding. Replica sets allow for up to 50 copies of data to be provisioned across separate nodes, data centers, and geographic regions for resilience and data protection. The election process for a replacement node is controlled by sophisticated algorithms that evaluate various parameters. By configuring replica sets, developers can provide tunable, multinode durability and geographic awareness. MongoDB's write concern ensures write operations propagate to a majority of replicas in a cluster, and the default durability guarantee has been elevated to w:majority in MongoDB 5.0, which provides a stronger durability guarantee. Beyond resilience, replica sets can also be used to scale read operations, intelligently routing queries to a copy of the data that is physically closest to the user.

4. Scale Up, Out, and Across Storage Tiers

MongoDB can be scaled vertically by moving to larger or smaller instance sizes and horizontally by distributing data across multiple nodes using native sharding, which allows for seamless scaling without adding complexity to the application. MongoDB's native sharding offers multiple sharding policies, including ranged, hashed, and zoned sharding, which distribute data according to query patterns or data placement requirements, providing higher scalability across a diverse set of workloads. MongoDB also offers tiered scaling through MongoDB Atlas Online Archive,

which automatically tiers aged data out of the database and into cloud object storage, enabling more economical scaling of data storage without losing access to the data.

5. Privacy and Security

The digital economy has led to increased government oversight of privacy and data security, prompting MongoDB to offer features for protecting user data. MongoDB provides strong authentication using a challenge-response mechanism based on SCRAM-256, integrates with enterprise security infrastructure, and offers role-based access control for granular user permissions. Additionally, MongoDB allows for network isolation, encryption of data in motion, at rest, and in use, and client-side field-level encryption for enhanced data protection. Unlike column encryption used in many relational databases, encryption in MongoDB is handled exclusively within MongoDB drivers on the client, making it unreadable to anyone without both client access and the keys necessary to decrypt the data.

6. One Platform for All Your Workloads

MongoDB Atlas is an application data platform that offers a set of cohesive and integrated data and database services. It provides a unified experience for developers, data analysts, data scientists, and supporting functions, along with global, multi-cloud data distribution, transparent data movement between services, and automated data life-cycle management. The platform includes Atlas Database, a general-purpose database that can deliver fast single-digit millisecond response times, and supports a flexible document data model, transactional guarantees, rich and expressive query API, and native support for vertical and horizontal scaling. Cluster auto-scale adjusts compute and storage in response to application load. Atlas Database is available in more than 80 regions across AWS, Google Cloud Platform, and Azure, and uptime is backed by a 99.995% service-level agreement.

Beyond offering fully managed MongoDB databases in the cloud, Atlas provides additional complementary services that allow organizations to support a wide range of application and real-time analytics data workloads.

Real-Time Analytics: Atlas Database allows you to deploy a read-only analytics node to serve more resource-intensive analytics queries.

Atlas Search: Atlas Search is built into MongoDB Atlas, making it easy to build fast, full-text search capabilities on top of your MongoDB data with no need to learn a different API or deploy a separate search technology.

Atlas Data Lake: Atlas Data Lake is a query service that allows users to analyze data in cloud object storage using the MongoDB Query API without the need for managing any infrastructure. The service parallelizes operations and automatically optimizes workloads by utilizing the compute closest to the data.

Realm Sync: Realm Sync enables bi-directional data synchronization between Atlas Database clusters and Realm, a lightweight, open-source mobile database, providing responsive reads and writes to users anywhere with strong data sovereignty controls for regulatory compliance.

GraphQL: When you use GraphQL alongside MongoDB's other app development features it's also simple to secure your app.

Event-Driven Architecture: Part of the broader services available to Atlas users, functions allow you to define and execute server-side logic without having to provision or manage servers, making it easy to integrate with cloud services, build APIs, and more.

7. MongoDB for Mission-Critical Applications in Your Data Center

The MongoDB Enterprise Advanced package provides software, support, certifications, and other services to help businesses run MongoDB on their own self-managed infrastructure for business or regulatory requirements. MongoDB Enterprise Server is included, which has advanced security options, comprehensive auditing, an encrypted storage engine, and an in-memory storage engine for high throughput and predictable low latency. Ops Manager simplifies administration tasks such as deployment, monitoring, backup, and scaling of MongoDB, and the MongoDB Enterprise Operator for Kubernetes automates and manages MongoDB clusters for Kubernetes users. The package also includes the MongoDB Connector for BI, allowing users to use MongoDB as a data source for SQL-based BI and analytics platforms.