

MongoDB Architecture Guide

Introduction

Working with data is very challenging for developers because of high demands of productivity, unable to extract insights because of the rapid growth of data in applications, monolithic databases. NoSQL databases address this, but sacrifices other features of a relational database.

The Document Model

Mongo replaces the tables of relational databases with documents that map to the way developers think and code. They store data as a JSON and model the data for the application. Some advantages of this model are:

- Intuitive: faster and easier for developers
- Flexible schema: dynamically adapt to change
- Universal: JSON documents are everywhere

Working With Document Data

The MongoDB Query API allows you to federate queries across databases with a connection string. The MongoDB Aggregation Pipeline allows you to transform and analyze data. It has ACID transactions and strong data consistency. It has features like Change Streams that detects and notifies applications of a modification in the database.

Distributed Architecture: Scalable, Resilient, and Mission Critical

Availability and data protection with replica sets

Replica sets are created for resilience. The election process is based on the implementation of the Raft consensus protocol. Write concern ensures the write operations to be in other replicas of the cluster. Replica sets are also used to scale read operations by routing queries to the data closest to the user.

Scale Up, Out, and Across Storage Tiers

With native sharding, MongoDB can scale out the database to handle write-intensive workloads. You can change the shard key without impacting the system. MongoDB native sharding options are:

- Ranged sharding
- Hashed sharding
- Zoned sharding

Privacy and Security

MongoDB implements capabilities to protect the data. This includes:

- Authentication
- Authorization
- Auditing
- Network isolation
- Encryption everywhere

One Platform for All Your Workloads

Key characteristics of MongoDB's application data platform include:

- A data plane that support different application types
- A unified and consistent experience for developers, data analysts, data scientists, and critical supporting functions
- Global, multi-cloud data distribution to support data residency requirements and provide deployment flexibility
- Transparent data movement between services and automated data life-cycle management

Real-Time Analytics

Atlas Database can deploy a read-only analytics node for resource-intensive analytics queries.

Atlas Search

It's easy to build fast and have full-text search capabilities. Search indexes are kept in sync.

Atlas Data Lake

It's an on-demand query service that analyzes data in cloud object storage using the Query API.

Atlas Charts

It's a data visualization service that supports richly structured JSON data.

Realm Sync

It provides bidirectional data sync between Atlas Database clusters and Realm.

GraphQL

Enabling GraphQL can increase performance by querying to a single endpoint to get the data you need.

Event-Driven Architecture

Atlas Triggers automatically execute functions in real time in response to an event.