MONGODB INTRODUCTION

MongoDB is a source-available, cross-platform, document-oriented database program.

Classified as a NoSQL database product, MongoDB utilizes JSON-like documents with optional schemas.

MongoDB is developed by MongoDB Inc. and current versions are licensed under the Server Side Public License.

MongoDB *is a data management platform that enables quick and easy query development and deployment of online, real-time data applications. It is a distributed, not-backed store that runs on a collection of servers and uses a JSON-like data model.*

MongoDB Replica Management allows you to easily and cost-effectively scale your MongoDB architecture. MongoDB provides a rich set of analytical tools for data profiling, load analysis, and monitoring. It can be used for a variety of purposes including data mining, Big Data, and online analytical processing.

Internet and enterprise application developers that require flexibility and scaling efficiently may consider using MongoDB. MongoDB is particularly suited to developers of varied types who are creating scalable applications using agile approaches.

Advantages and Disadvantages of MongoDB

MongoDB has both pros and cons just like other NoSQL databases.

Pros:

- *Any type of data can be stored in MongoDB, which gives users the flexibility to create as many fields in a document as they desire.*
- *Documents map to native data types in many programming languages, which provides a means of adding to data. Sharding, which involves dividing data across a cluster of machines, is also achieved by this.*
- *MongoDB includes its own file system, similar to the Hadoop Distributed File System (HDFS), called GridFS. The file system is primarily used to store files that exceed MongoDB's 16 MB per document BSON size limit.*
- *MongoDB is also compatible with Spark, Hadoop, and other data processing frameworks like SQL.*

Cons:

- *When a MongoDB master node goes down, another node will automatically become the new master. Despite the fact that it promises continuity, the automatic failover strategy is not instantaneous – it may take up to a minute. In contrast, the Cassandra NoSQL database supports multiple master nodes, so that if a master goes down, another one is ready to run a highly available database infrastructure.*

- Although MongoDB's single master node restricts how fast data can be written to the database, it also limits how much data can be written. Because of this, data writes must be recorded on the master, and new information cannot be added to the database quickly.
- MongoDB doesn't provide full referential integrity using foreign-key constraints, which could affect data consistency.
- User authentication isn't enabled by default in MongoDB databases. Because of this, there is a default setting that blocks networked connections to databases if they've not been configured by a database administrator.
- There have also been instances of ransomware attacks that forced the setting to be turned on by the database administrator.

Features of MongoDB

1. Replication: The MongoDB replica set feature is known for providing high availability. Two or more copies of data constitute a replica set. A replica-set acts as a primary or a secondary replica. Secondary replicas keep a copy of the data of the primary, preserving it in an orderly manner, as part of a replicated MongoDB system. Whenever a primary replica crashes, the replica set automatically determines which secondary should become the primary and conducts an election if necessary. Secondary replicas may additionally serve read operations, but the data is only eventually consistent by default. To resolve the election of the new primary, three standalone servers must be added as secondary servers.

2. Indexing: A MongoDB field can be indexed with primary and secondary indices or indexes. A MongoDB index stores a small portion of the data set in a form that is convenient to traverse. The index stores the value of a particular field, or set of fields, ordered by their value. In MongoDB, indexes assist in efficiently resolving queries by storing a small portion of the data set in a convenient form. A MongoDB index is similar to a typical relational database index.

3. File storage: GridFS, which uses MongoDB as a file system, can be used to balance and replicate data across multiple machines. A file can be stored in MongoDB as a grid file system. It has features similar to a file system such as load balancing and data replication.

4. Aggregation: The aggregation pipeline, the map-reduce function, and single-purpose aggregation methods are available in MongoDB. According to MongoDB's documentation, the Aggregation Pipeline provides better performance for most aggregation operations over map-reduce. With the aggregation framework, users can obtain the kind of results for which the SQL GROUP BY clause is used. The aggregation framework includes $lookup and standard deviation like statistical operators.

5. Sharding: Sharding is the splitting up of data among machines. To permit this, we refer to it as "partitioning" or "sharding." We may store more data and handle more load without upgrading our machines, by dividing data across them. MongoDB's sharding allows you to split up a collection among many machines (shards), allowing it to grow beyond resource limitations.