

(1sl)

Hello everyone. I want to talk about MongoDB

(2sl)

Introduction to MongoDB **(3sl)**

MongoDB is a NoSQL database which stores the data in form of key-value pairs. It is an Open Source, Document Database which provides high performance and scalability along with data modelling and data management of huge sets of data in an enterprise application.

MongoDB also provides the feature of Auto-Scaling. Since, MongoDB is a cross platform database and can be installed across different platforms like Windows, Linux etc. **(4sl)**

What is Document based storage?

A Document is nothing but a data structure with name-value pairs like in JSON. It is very easy to map any custom Object of any programming language with a MongoDB Document. **(5sl)** For example : Student object has attributes name, rollno and subjects, where subjects is a List.

We can see, Documents are actually JSON representation of custom Objects. Also, excessive JOINS can be avoided by saving data in form of Arrays and Documents(Embedded) inside a Document. **(6sl)**

Brief History of MongoDB **(7sl)**

MongoDB was developed by Eliot Horowitz and Dwight Merriman in the year 2007, when they experienced some scalability issues with the relational database while developing enterprise web applications at their company DoubleClick. According to Dwight Merriman, one of the developers of MongoDB, this name of the database was derived from the word humongous to support the idea of processing large amount of data.

In 2009, MongoDB was made as an open source project, while the company offered commercial support services. Many companies started using MongoDB for its amazing features. The New York Times newspaper used MongoDB to build a web based application to submit the photos. In 2013, the company was officially named to MongoDB Inc. **(8sl)**

Key Features of MongoDB

Apart from most of the NoSQL default features, MongoDB does bring in some more, very important and useful features : **(9sl)**

1. MongoDB provides high performance. Input/Output operations are lesser than relational databases due to support of embedded documents(data models) and Select queries are also faster as Indexes in MongoDB supports faster queries. MongoDB has a rich Query Language, supporting all the major CRUD operations. The Query Language also provides good Text Search and Aggregation features.
2. Auto Replication feature of MongoDB leads to High Availability. It provides an automatic failover mechanism, as data is restored through backup(replica) copy if server fails.
3. Sharding is a major feature of MongoDB. Horizontal Scalability is possible due to sharding.

4. MongoDB supports multiple Storage Engines. When we save data in form of documents(NoSQL) or tables(RDBMS) who saves the data? It's the Storage Engine. Storage Engines manages how data is saved in memory and on disk. **(10sl)**

Organizations that use MongoDB

Below are some of the big and notable organizations which are using MongoDB as database for most of their business applications.

- Adobe
- LinkedIn
- McAfee
- FourSquare
- eBay
- MetLife
- SAP **(11sl)**

Overview of MongoDB **(12sl)**

MongoDB consists of a set of databases. Each database again consists of Collections. Data in MongoDB is stored in collections. The below figure depicts the typical database structure in MongoDB.

- Collection is nothing but a set of MongoDB documents. These documents are equivalent to the row of data in tables in RDBMS. But, collections in MongoDB do not relate to any set schema as compared to RDBMS. Collections are a way of storing related data. Being schemaless, any type of Document can be saved in a collection, although similarity is recommended for index efficiency. Document's can have a maximum size of 4MB. **(13sl)**
- We can use namespace to logically group and nest collections. **(14sl)** For example : There can be one collection named **(15sl)** db.studytonight.users **(16sl)** to save user informations, then there can be others like **(17sl)** db.studytonight.forum.questions **(18sl)** and **(19sl)** db.studytonight.forum.answers **(20sl)** to store forum questions and answers respectively.
- If we create an Index on a namespaced collection, it will only apply on that namespace only.
- A collection is physically created as soon as the first document is created in it.
- You must be wondering why create multiple collections with different namespace, when we can keep any form or data in a single collection itself. It's because, MongoDB does not index attributes for totally unrelated documents. So it is advised to keep related data in collections.

Document in MongoDB is nothing but the set of key-value pairs. These documents will have dynamic schema which means that the documents in the same collection do not need to possess the same set of fields. **(21sl)**

Since MongoDB is considered as a schema-less database, each collection can hold different type of objects. Every object in a collection is known as Document, which is represented in a JSON like (JavaScript Object Notation) structure(nothing but a list of key-value pair). Data is stored and queried in BSON, its binary representation of JSON-like data. **(22sl)**

In the above figure, the field _id represents the primary key identifier of the given document. MongoDB also stores the values in the form of arrays of value as well. Infact any type of data can be stored as values, and nothing requires to be pre-defined. In other words, you do not have to predefine the type of data to be stored, you can store anything you want. Remember, MongoDB is schema-less.

Documents are not identified by a simple ID, but by an object identifier type. The default Id is a combination of machine identifier, timestamp and process id to keep it unique, but user can changes it to anything. **(23sl)**

The value of fields in a document can be anything, including other documents, arrays, and arrays of documents, date object, a String etc. **(24sl)**

MongoDB vs SQL Databases **(25sl)**

It is a well known fact that SQL databases have ruled the world of data technologies and have been the primary source of data storage for over 4 decades. Generally the SQL databases are used, mainly for accessing relational databases.

Oracle and Microsoft SQL Server ruled the segment, but as the Web development market paced up, there came a shift towards usage of open source databases like MySQL, Postgres etc. But RDBMS was still the first choice.

Soon enough data started growing exponentially and scalability became a major issue, at that time NoSQL rolled in to save the day.

One of the cool facts about NoSQL databases is that they existed since 1960, but in the recent times they have gained more popularity especially for the scaling feature provided by the same.

The major difference between MongoDB and SQL Databases is the way they handle data. In SQL databases, data is stored in form of traditional 2 dimensional row-column structure while in MongoDB rich data document model is followed, which allows storage of any type of data.

MongoDB Documents also align with the structure of objects in modern programming languages, as they are a form of JSON. This makes it easy for developers to map the data used in the application to its associated document in the database. While in SQL Database, creating a table with columns mapped to the attributes of an object in programming language, appears a little tedious. **(26sl)**

Advantages of MongoDB **(27sl)**

Having seen the good features of MongoDB, now every developer should be able to understand why it is better to use NoSQL based database for big data transactions and for implementing a scalable model. Now, its time to leave behind the schema definitions of RDBMS and get an advantage of using schema-less databases like MongoDB. Let us see some of the vital advantages of MongoDB :

1. First and foremost, it is very easy to install and setup the MongoDB.
2. The very basic feature of MongoDB is that it is a schema-less database. No schema migrations anymore. Since MongoDB is schema-free, your code defines your schema.
3. The ability to derive a document-based data model is one of the most attractive advantages of MongoDB. Because, the way it stores the data in the form of BSON (Binary JSON), ruby hashes etc, helps to store the data in a very rich way while being capable of holding arrays and other documents.
4. The document query language supported by MongoDB plays a vital role in supporting dynamic queries.
5. Very easy to scale.
6. Due to the structuring (BSON format - key value pair) way of the data in MongoDB, no complex joins are needed.
7. Performance tuning is absolutely easy compared to any relational databases.

8. No need of mapping the application objects to the data objects.
9. Enables faster access of the data due to its nature of using the internal memory for the storage.
10. Since, it is a NOSQL database, then it is obviously secure because no sql injection can be made.
11. MongoDB can also be used as a file system, which helps in easier way of load balancing.
12. MongoDB supports, the search by regex and fields as well.
13. MongoDB can be run as windows service as well.
14. Good amount of documentation is available.
15. MongoDB does not require a VM to be run.
16. MongoDB follows regular release cycle of its newer versions.
17. The support for Sharding is one of its key feature. Sharding is the process of storing the data in different machines and MongoDB's ability to process the data, as and when the size of the data grows. This results in the horizontal scaling. With sharding, more amount of data can be written and read back as and when there is an increase in the data growth. **(28sl)**

THANK YOU FOR YOUR ATTENTION