Comparative Study of MongoDB and NoSQL Databases

Sandhya Singh
Scholar of Computer Science
Bhabha Institute of Technology

Anamika Tiwari
Faculty of Computer Science
Bhabha Institute of Technology

Abstract –

With the data winding up radically and the structure of data ending up being dynamically versatile, MongoDB has displaced the social database in various applications. In numerous applications it has outflanked the customary SQL databases, and subsequently in this paper there would display about the progressions from SQL no NoSQL database likewise drawing out it's preferences and burdens.

Keywords -- *NOSQL Database, MongoDB, Comparison between SQL and NOSQL databases.*

1. INTRODUCTION

MongoDB was discharged in 2009, and is written in C++. It is the most prominent NoSQL database framework accessible. As a report situated database, MongoDB holds a course of action of accumulations, which have equivalent ramifications as the table of connection database. [2] An accumulation of MongoDB contains numerous records and does not have a predefined outline. Every record is an arrangement of fields, which is remarkably distinguished by the ID field, and it additionally has an adaptable structure. This paper starts by a simple survey on current NoSQL database systems and then it emphasizes on the major difference between current day SQL systems and the one with MongoDB. The best approach to achieve the change from SQL to MongoDB and after that security issue looked with MongoDB that may have been missing in SQL framework is referenced to sum things up. In spite of the fact that, MongoDB is new however, it is an open source and is proficient. It is likewise being utilized in numerous cutting-edge database tasks and standards.

2. SURVEYS ON NOSQL DATABASE

2.1 Reasons

Many organizations came together, which had more demands and aspiration from the database technology in the given aspects [3]:

2.1.1 Efficient big data storage

There are enormous applications, for example, the web indexes which need a proficient database to meet the capacity prerequisite and which can react to around billions of traffic.

2.1.2 Incremented consistency, regularity, unity

With the regularly augmenting number of back to back and nonstop demands and information the database needs to help information development and after that at the apex ensure a not coming up short administration.

2.1.3 Confined capacity

The present relational databases have a very confined data capacity and they are unable to support the massive data like in a search engine.

2.1.4 Lower management and operational price

As the information develops generally, the database costs that contains equipment costs, programming costs and working costs have augmented. In this way, a decrease in these costs is truly necessary.

2.1.5 Slow reading and writing

A social database has a well-characterized rationale intricacy. As the information increases; it is increasingly helpless against gridlocks and different issues that prompts less effectiveness.

2.1.6 Expansion difficult

Multi table correlation mechanism, which is present in relational database.

2.2 Features of NOSQL database

2.2.1 Features of NOSQL

It reads and writes data fast. It is simpler to expand and it also supports massive storage with a low price.

2.2.2 Data model

It has a key worth information model which means there is a worth which prompts key, the structure is basic with an expansion in question speed. It has a segment situated database that is a table in any case, it doesn't bolster table affiliation that is a join. Each furthermore, every section is the record of the database. There is a database in which there are accumulations and in accumulations sections and in segments archives are there.

2.2.3 CAP Theorem

It says that there could be three properties of a common information framework that is normality, versatility and accessibility, and resistance that is there in system allotments.

2.3 Mainstream NOSQL database

2.3.1 Key Value databases

A standout amongst the most productive key worth databases is Flare. It was made by the Green.jp in Japan. It has increased adaptability. Flare included a hub server that is before the information servers so

that information servers could be taken care of at the back consequently, clients can include or erase information servers with the assistance of flare. Another productive key worth database is the TC in particular Tokyo Cabinet. It is an extremely superior motor. Tokyo Cabinet moreover guarantees superior exhibitions of read and compose in the meantime; it depends on information solidness system. Its fundamental disadvantage is that when there is gigantic information, the simultaneous compose execution drops at an extraordinary degree.

2.3.2 Column oriented database

It oversees superior of investigating information and business knowledge preparing. It incorporates Cassandra and Hyper table. Cassandra is a database of Facebook that is an open source. Its exceptionally symmetric in nature. It overseas enormous measure of organized information and gives exceptionally solid administrations with least measure of disappointment. It doesn't have a fully-fledged social database model however it has a model that is less complex and has authority over information design and arrangement. It was Cassandra that actualized the Inbox search highlight of Facebook. It has a key worth store. Key here maps to values, which are generally conveyed into section families. Here, segments must be attached to keys, which are as of now indicated, and which infers every single key has a distinctive number of segments in a specific family. It additionally has super sections, a segment whose qualities are segments that is in a perfect world an arranged exhibit of segments connected cooperatively. The values from segment key for every family are put away together. This characterizes Cassandra as large crossover information chief between a section situated DBMS and a line arranged store.

2.3.3 Document Database

It doesn't concern read or composing in the meantime and their execution. They help in overseeing huge information and viable inquiry execution. It includes MongoDB [1], CouchDB. CouchDB is a flexible database that can tolerate faults. It has some drawbacks that its reading and writing concurrent is less. MongoDB on the other hand is a nonrelational database and has a very high access on massive data. It has a very efficient query language and supports complicated data types. It is a cluster of nodes where not every node is the same that is it lack symmetry.

3. DIFFERENCE BETWEEN ORACLE (SQL) AND MONGODB (NOSQL)

The correlation is done, in light of the fact that Oracle is the most prominent framework utilized and has been around for a long time while, MongoDB is generally new, yet at the same time has been utilized by numerous undertakings and items.

3.1 Main Difference

The rule differentiation is that NoSQL is a class of database engines that don't support SQL remembering the true objective to achieve execution or unflinching quality components that are opposite

with their versatility of SQL. [7] Utilizing NoSQL forces some engineering imperatives that are depending of what one is doing with the information. It is uncommon for a whole existing SQL-driven application to be moved in its integrity over to NoSQL still some algorithms like the mentioned above can be used to translate a database in SQL to MongoDB. As mentioned, the above paper, MongoDB doesn't store data in fixed tables but have a rather flexible outline as compared to SQL. Also, MongoDB is document oriented schema-less whereas Oracle as we all know is a relational database. The input data size for MongoDB is 16MB contrasted to 4KB of Oracle, which is indeed too small to be modern. The major difference is that MongoDB is open source and free whereas Oracle is a service one has to pay for, but is the mainstream and have been for long.

3.2 Syntax Difference

Oracle Database utilizes regular SQL dialect. It utilizes information control proclamations as SELECT, UPDATE, DELETE. On the other hand, MongoDB utilizes functions for the operations of including new records, upgrading and erasing existing records. In MongoDB there is no relation between two tables So any table can be dropped without worrying about the fact that it may create exceptions and other errors. In MongoDB, the formation of the table is made on the first insertion operation. As a result, of the adaptability of the MongoDB accumulation, the structure doesn't need to be altered. In Oracle databases have restrictions or conditions on tables So when one is attempting to add another record, they must be mindful so as not to damage any of the limitations. Then again, MongoDB is more adaptable with the information. One doesn't have any imperatives with respect to the information from the collections. This is the significant contrast between a SQL Database and a NoSQL one. Using collections, data is stored in NoSQL databases. The collections of MongoDB have no imperative regarding the data contained in them. There is no settled structure in the collections. The areas may have various kinds of data. NoSQL varies from SQL as in SQL information type for a section is fixed and in it just the predefined esteem can be entered, generally a mistake emerges. In MongoDB, to locate a specific information discover work is utilized and here in MongoDB, there aren't any tables. The data is put away using accumulations of data. This accumulation doesn't have a modified structure. The things put away in these collections can have distinctive fields. MongoDB uses the concept of class, i.e. all the functions are member functions of a class and are called using '.' operator of C++. Example, db.tt.find().sort('field_name'). To completely understand MongoDB, one has to go through all the documentations provided with it, and is available on docs.mogodb.com.

3.3 Time Comparison

While switching to MongoDB, one not only has to consider time of execution but it plays a major part in the decision of switching and can be one of the deciding factors. MongoDB is more productive while embedding a lot of information. It takes too little too long to embed less number of records. [7] Prophet

Database deals amazingly wonderful with little proportion of data, yet when there is a talk on broad wholes, progressively important then 10000 records, the time spent on this activity is more noteworthy than MongoDB by a noteworthy sum. The story is the equivalent for refreshing a record, while Oracle takes a great deal of time, refreshing few or numerous records, MongoDB takes no time contrasted with Oracle refreshing records regardless of what number of are refreshed the normal time to refresh a record is around 1ms. The time taken in erasing a record is colossal, MongoDB beating Oracle. While time taken ascents exponentially in Oracle it stays same for MongoDB, regardless of what is the measure of records that need to be erased.

4. TRANSFERRING SQL DATABASE TO MONGODB

The use of a MongoDB instead of SQL database may bring about change in many aspects. The steps to transfer are very easy but cautiousness is needed while doing so. Thereby, there are changes initiated in the model.

4.1 Challenges

4.1.1 Model transformation Challenge

In view of the nonappearance of specific gadgets, a significant proportion of model change frameworks relies upon database's head's involvement. People consistently plot the physical model of MongoDB physically in perspective on the current social database. It is uncommonly gravely structured when facing with an incredible structure of associations. Additionally, as one builds the quantity of related table, execution might be expanded yet information repetition is instigated. Subsequently, we have to make an exchange off between quantities of relations included.

4.1.2 Data Migration challenge

Presently, information relocation apparatuses are utilizing extremely basic movement methodologies, such as moving every table of relational database to a particular collection of MongoDB.

5. LIMITATIONS OF SQL DATABASE

The inspiration driving moving data from social database to NoSQL database is to procure a prevalent execution. Thusly, there is a need to consider the obstacles of the current social database. From that point forward, the upsides of NoSQL database can be taken to overcome the insufficiencies of social database. MongoDB doesn't bolster joins, however it has chain of importance and settled tables.

5.1 Conversion

All SQL systems maintain a log, so as to store about the commands executed and implemented by the user. In this paper, all these tags are extracted from the log and divided into categories such as.

1. **Frequent Joins:** It implies two tables as often as possible execute join operation and the query speed is not sufficient.

2. **Big Size:** Data is large

3. **Frequent Modify:** Data is modified frequently
4. **Frequent insert:** New data is added frequently

Characterization labels and the relationship of social database creates "Action labels". The fundamental frequency motivation behind creating action labels is to utilize installing or denotation amid the manikin alteration and route the nesting heading. The ER model of the database in relational form is used to form an outline of database in MongoDB. In spite, of the fact that MongoDB does not bolster join operation, it can store organized information by references. Consequently, Reference in MongoDB could be utilized to depict Relationship in social database. As mentioned earlier, sometime recently, a considerable measure of movement devices does that way. The main distinction is that every report will have an id-field to extraordinarily distinguish it. Three algorithms were specified to convert from relational database to MongoDB, which are: Conversion of ER model of relational database to physical model of MongoDB. But as there is a need of action tags, an algorithm to generate action tags from the ER model, and a third algorithm to transfer the data is used.

6. SECURITY ISSUES IN NOSQL DATABASE

A very disastrous yard bird is NoSQL databases have compromised with the security and bunco game sentiency for performance improvement. Here, is a glance of the proficiency that are implemented for security in NoSQL databases that are namely Cassandra and MongoDB. The briny welfare of using the NoSQL is that unlike relational database, it can manage all the unstructured data such as documents, multimedia system, e-mail etc. It has a very simple data exemplar with no complexity, easy interrogation language. The con or the most unreliable feature of the NoSQL database is its lack of security and regularity. It is not well equipped with management of data with I and integrity constants like foreign key handling.

6.1 Security in MongoDB
1. In MongoDB the data files are unencrypted in nature and it does not provide an implementation to encrypt the files automatically.
2. It uses a binary wire level protocol which uses TCP port 27017 by default. It is at high verge of injection attacks because it uses mostly java script language. Thus, any attacker can easily acquire the passwords of the data files of user in a particular database.
3. No auditing is allowed here.
4. In data communication no encryption is there.

The main drawback in both of the NoSQL database was the lack of encoding, weak authentication between the customer and the waiter, defence -of-service and the vulnerability to injection in the database. To decrease the injection in the database it should be verified that the application does

reasonable input validation. The data at rest that is unencrypted can be protected with the OS level mechanism[3].

7. CONCLUSION

NoSQL database in many respects, varying from velocity to flexibility, is quicker and better than SQL database. A current SQL database can also be easily transmitted to a MongoDB (NoSQL), which is why many businesses shift their projects to use MongoDB instead of the traditional SQL database. The Erwin HAWK tool can be used to choose a suitable NoSQL database and replace the traditional relational database. NoSQL is also called the future of data economy. In most queries an Oracle database faces, MongoDB is not only quicker, it is also more versatile and can easily store big information. [4]The primary up-votes to move to MongoDB are flexibility and velocity. Also, as stated previously, with little difficulty, one can migrate from SQL-based Oracle to MongoDB. Taking everything into consideration, on the off chance that he or she can rely on MongoDB if you need to use a fast, adaptable database. In case the speed of the database is not a major concern, and one should depend on the great Oracle Database on the off chance that relationships are required between the tables and the accumulations. There are also certain disadvantages of NoSQL database, the main problem being addressing the safety problem, which is still in its nascent state in the case of NoSQL and requires time to evolve. It is very rare to use MongoDB in cloud and other such apps for the same reason and some other reason. As the next best thing for the future, one can readily migrate to NoSQL database as it is an open-source and free, and also its more flexible. But some significant concerns such as safety still need a fixation, without which NoSQL will always be a step behind SQL, as safety is one of the main factors for implementing any database using SQL and now safety is one of the greatest problems with all the scams and abilities that hackers have obtained.

8. REFERENCES

1. MongoDB Documentation https://docs.mongodb.com
2. Tianyu Jia, Xiaomeng Zhao, Zheng Wang, Dahan Gong and Guiguang Ding - Model Transformation and Data Migration from Relational Database to MongoDB IEEE 2016.
3. Lior Okman, Nurit Gal-Oz, Yaron Gonen, Ehud Gudes, Jenny Abramov - Security Issues in NoSQL Databases IEEE2011
4. Zhu Wei-ping, Li Ming-xin- Using MongoDB to Implement Textbook Management System instead of MySQL IEEE2011

****