International Journal of Allied Practice, Research and Review

Website: www.ijaprr.com (ISSN 2350-1294)

NOSQL Databases: The Need of Enterprises

Basit Maqbool Mattu
M-Tech CSE Student. (4th semester).
Department of Computer Science & Engineering
Ganpati Institute of Technology and Management,
Bilaspur Haryana India.

Abstract - The aim of this paper is to study about the growing need of the NOSQL Databases for the present organizations and the business enterprises. NOSQL Databases have some features which make them suitable for enterprises and also for social networking like performance, scalability, availability etc. Today the data is increasing in large amounts from the organizations which cannot be handled by the traditional database systems. This is known as Big Data which increases in variety and velocity also. To manage and handle this Big Data issue we need the use of the NOSQL Databases. NOSQL Databases can manage the increasing data from the enterprises. Also in this paper we will also have a look towards the various categories of the NOSQL Databases.

Keyterms: NOSQL; Big Data; CAP; BASE; ACID; RDBMS.

1. INTRODUCTION

Whenever we develop any kind of the application, we need to have the back-end associated with it. This back-end is usually known as the Databases. The Databases are the most important part of the modern applications, as our applications cannot function without the Databases. There is also data management tasks associated with the Database Systems. As we know that data is increasing in huge amounts from terabytes to petabytes and exabytes, so it is not possible for the traditional databases eg RDBMS to handle so much huge amounts of data. The volume of the data is increasing in the Business enterprises and organizations. The data to be stored and retrieved is increasing day-by-day. This data is commonly known as Big Data problem. The NOSQL Databases are open source Databases based on distributed computing which are very effective in handling the Big Data problem. Usually this Big Data not only increases in volume but also in velocity and variety also. This means that Big Data flows very swiftly and very fast in the process of storage and retrieval. The variety indicates that data can have many forms like structured, unstructured and semi-structured data. In today's world only 20% data is structured or formatted and rest of the 80% data is unformatted. Structured data is that data which is in the form of the relational tables and the unstructured and semi-structured data are those types of data which is unformatted and not in the form of

relational tables. The un-structured data are also videos, images and audio etc. Now coming back to main problem of the Enterprises that the data released from these organizations is very large so it is not possible for the traditional databases to handle this Big Data problem. Now there are many features of the NOSQL Databases which make them suitable for handling the Big Data problem of the organizations. As there are millions of the users supported by these enterprises today. So to support these millions of users organizations make use of the NOSQL Databases. Since NOSQL Databases supported the Distributed Systems also. Another main advantage of the NOSQL Databases are that it has Horizontal Scalability also. It means that we can add or remove nodes in any way as we have a need. In NOSQL Databases it is very easy to add the nodes without causing any failure in the other nodes. The NOSQL Databases follow the CAP Theorem. The traditional RDBMS Systems supported only consistency and partition tolerance. NOSQL Databases follow two attributes which are availability and partition tolerance and they give birth to a new model known as BASE (Basically Available, Soft State, Eventually Consistent). This is reverse of the model ACID which was followed by the traditional RDBMS Systems. The NOSQL Databases have a very good performance since the data is always available to the user. They support scalability which make good in competition against the traditional RDBMS Systems. The NOSQL Systems are non-relational and do not use the relational tables.

II. CATEGORIES OF THE NOSQL DATABASES

There are many categories of the NOSQL Databases. We will have a look into all the categories of the NOSQL Databases which are used nowadays. Some of these main categories are:

1. **KEY-VALUE STORE DATABASES:**This is the very simple type of the NOSQL Databases. It usually consists of the Key-value pairs. Here we have two columns of key and their values. The key is a unique. It usually refers to the data or values which is stored in the form of strings. Any of the indexing or the hashing method can be used to find the values of strings associated with the key. There are some features of the Key-value store databases which make them fast, reliable, and appropriate to use. The examples of the Key-value store NOSQL Databases are Amazon's Dynamo, Berkeley DB, Simple DB, Big Table.

2. **DOCUMENT-STORE DATABASES:**This is the another category of the NOSQL Databases which usually make use of the documents to store the data. The documents can store the data in the XML format which are in the form of BSON and the JSON format. These cam hold documents of any length and store the semi-structured data. There are many types of semi-structured data which can managed by the document-store databases and can be in XML format only. Some of the kinds of data stored by Document-store Databases are messages and comments etc. The good examples of the Document-store Databases are MongoDB and CouchDB.

3. **COLUMN-STORE DATABASES:**This is also one of types of the NOSQL Databases. This can store data in the form of columns rather than rows. It makes the use of the columns which can store data in the form of the extended columns only. These databases do not make use of rows and columns as in relational databases. It can also work as the key-value pairs in which key can be any column and values can represent a table. Some of the main examples of this type of the NOSQL Databases are Cassandra, Hypertable and the Hadoop.

4. **GRAPH-STORE DATABASES**: This database uses the nodes, edges and properties to store the data in the form of the graph. The nodes can act as objects and the relationship between these objects can be represented by the edges. In these type of the Databases we can make use of any of the indexing method to search the data in the database. The stored data is represented in the form of nodes and the vertices.

The data can be retrieved from the graph using the indexing methods. The main examples of the graph-databases are Infogrid and VertexDB.

III. MAIN FEATURES OF NOSQL DATABASES

There are some important features of the NOSQL Databases which make them suitable for the need of enterprises and also for the need of the modern applications nowadays. Some of the main and important features or characteristics of the NOSQL Databases are as follows:

1. **SCALABILITY:** This is one of the important feature of the NOSQL Databases. The NOSQL Databases support the Scalability of the data in the modern applications. We need to have the scalability in the data. As the data grows very fast and number of users also increase so there must a option of the scalability. Also the proper use of resources is also very important. If we take an example of the Cassandra which is one of the NOSQL Databases, it allows us to add or remove nodes in such a way that there is no failure in the application itself. The nodes can be added in the horizontal way as we have a need without disturbing the application. From this point we come to understand the fact and reason behind the use of the NOSQL Databases. The NOSQL Databases are most reliable databases which give horizontal scalability. In the today's enterprises there are millions of users who access the data. To handle so many users and to handle so much big data we must have scalability features present in our databases. The enterprises cannot handle large number of users and data with the traditional RDBMS Systems. Our NOSQL Databases support the CAP Theorem since these are good attributes to be present in our databases.

2. **PERFORMANCE:** When we study this feature of the NOSQL Databases we say that the performance of the NOSQL Databases are measured in the terms of the latency. It has to produce low latency. Now the NOSQL Databases allow to share the data among many number of users in the cluster, so making it easy for them to communicate with each other. When the data is available to all the users, so there is also the decrease in the latency of the reads and writes.

   We can say that the NOSQL Databases provide better performance than the traditional databases which is one more need for the enterprises and organizations today.

3. **AVAILABILITY:** When we talk about this feature or characteristic of the NOSQL Databases, it gives the need of availability of the data to the users. This feature gives the potential to add, remove, backup, setup, upgrades and even detect the failure in the entire system. This allows the NOSQL Databases to become most successful for the need of the enterprises. If we take the case study of the Cassandra, in which if one node fails to perform say a write operation then another node will get the hint message giving the reason of failure of write operation of previous node. In this whole information is transferred to the next available without causing disturbance to the entire cluster. In this way in the Cassandra it can automatically repair or detect the issues of failure in the data centric system. When the failed node become active again, then the hint message will restore it with its necessary information and it will continue to perform its write operation.

4. **SIMPLE STRUCTURE:** This is again one of the main features of the NOSQL Databases. It does not have any fixed schema structure like the RDBMS. It possess very simple structure in which there is a fixed schema and if we have to make changes, then we don't need to change the existing structure. This is main characteristic of the NOSQL Databases which makes them simple to use than RDBMS.

IV. CASE STUDY OF FEW NOSQL DATABASES

Here under this section we will have a study of some few NOSQL Databases. These are as follows:

1. **CASSANDRA:** This is one of the most popular NOSQL Databases. It is the column based key-value store database. It has better performance, availability, and horizontal scalability. It is the most commonly used NOSQL Databases nowadays. It performs availability in a way that one node cannot perform read or write operation then it passes the hint message to the next node. When the node comes back to active state, then hint message recovers the information about the failed node and the failed node starts its operation again. This process is all automatically performed by the Cassandra. It makes the insertion and deletion of the nodes easily done. Also it makes availability in such a way that it does not disturb the whole system and recovers the failed nodes in a proper way.

2. **MONGODB:** This is another type of the NOSQL Database. It is document-based Database. It stores data in the form of XML Documents either in JSON or BSON format. It consists of master node, arbitrary node and slave nodes. Whenever there is a failure in one of slave nodes, the master node can transfer data to the other active slave nodes. Sometimes, the master node may fail, in this case the arbiter node chooses some master node on its own. This type of NOSQL Database is very good in auto-partitioning the data, distributing the data among the servers and performing consistent reads and writes.

3. **COUCHDB:** This NOSQL Database is also the document-store database. It stores data in the form of the JSON Documents. In this type of database, all the cluster nodes are same and they perform the task of availability and performance. This database gives very low read and write and removing and adding the nodes is very easy in this database. It detects the problems on its own and distributes the data equally to all the nodes.

V. CONCLUSION

After going through the study of the NOSQL Databases, I want to give my conclusion in this paper that the NOSQL Databases are better option for the Enterprises today. Since the NOSQL Databases are having some very good features different from the traditional databases. It is also possible to control the Big Data problem of the enterprises and organizations with the help of the NOSQL Databases. We can also say that today's applications have a need of NOSQL Databases. Same is the need for the big business enterprises and organizations also where the data is very large and cannot be simply managed by traditional RDBMS. The NOSQL Databases can manage the data in a very proper way and provide all the necessary features like performance, scalability and availability. At last I want to say that NOSQL Databases is an only need for all big organizations like Facebook, twitter, Google and many business intelligence enterprises. It has good future demand also and a very good scope in the fields of the research also.

VI. ACKNOWLEGEMENT

I am thankful to all the people who motivated me to do research on this topic. I am also thankful to all those research scholars who have done previously some research work on this topic and with the help of their research papers and articles I was able to write my first research paper in M Tech 4th semester on this particular topic.

VII. REFERENCES

- Ameya Nayak, Anil Poriya, Dikshay Poojary, "**Types of NOSQL Databases and its comparison with the relational databases"** March 2013, International Journal of Applied Information Systems, Volume 5, No. 4, www.ijais.org.
- Vatika Sharma and Meenu Dave, "**SQL And NOSQL Databases"** August 2012, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 8, www.ijarcsse.com.
- Dr K.Chitra and B.JeevaRani, "**Study on Basically Available, Scalable and Eventually Consistent NOSQL Databases",** July 2013, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, www.ijarcsse.com
- Trevor Perrier and Fahad Pervaiz, "**NOSQL in a mobile world: Benchmarking Embedded Mobile Databases".**
- Santhosh Kumar Gajendran " **A Survey on NOSQL Databases".**
- Uma Bhat and Sharradha Jadhav,"**Moving Towards Non-Relational Databases",** 2010, International Journal of Computer Applications Volume 1-No.13.
- A B M Moniruzzaman and Syed Akhtar Hossain, "**NOSQL Database: New Era of Databases for Big Data Analytics-Classification,Characteristics, and Comparison"** International Journal of Database theory and Application.
- Rabi Prasad Padhy, Manas Ranjan Patra, and Suresh Chandra Satapathy, "**RDBMS to NOSQL: Reviewing some Next Generation Non-Relational Databases",** International Journal of Advanced Engineering Sciences and Technologies, Vol. No. 11, Issue No. 1. http://www.wikipedia.com.
- Naseer Ganiee ," **NOSQL : The Big Data Solution",** The International Journal of Engineering Technology, Management, and the Applied Sciences, Volume 1, Issue No. 2.
- Naseer Ganiee, "**New Database Constraints and Modern Applications",** IJETMAS, Volume 3, Issue 2.