ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 9, September 2016

NOSQL DATABASES

Anushka Harshad Patil, Tanvi Girish Samel, Darshan Haresh Vira

Abstract: **Technology is now on the forefront of modern era, it is growing rapidly and have become even more composite in terms of 3Vs (volume, variety and velocity). These 3Vs are three defining properties or dimensions of big data and have grown so large that handling and exploiting them using conventional data management tools like relational database management systems (RDBMS) or conventional search engines is another difficult task. To handle this issue, precisely designed DBMS alternatives such as NoSQL, NewSQL and Search-based systems are supplemented to conventional data management tools. In this research paper, we are discussing about NoSQL with an intention to help consumers and organizations to gain detailed information about various NoSQL database approaches and how different it is from SQL.**

***Index Terms*—Big Data, CAP, NewSQL Database, NoSQL Database.**

I. INTRODUCTION

NoSQL is an abbreviation for Not Only SQL or non SQL. It is pronounced as noseequel. It is a derivation of the current database systems that provides additional support to the type of data which is ever increasing, for example the data of a product collected from a social media on how it is performing. NoSQL is a non-relational database management system, fast information retrieval database and is portable.[1] NoSQL databases are non-relational, open source, distributed database. Non-relational database does not organize its data in related tables i.e., data is stored in a non-normalized way. NoSQL databases are open source; therefore, anyone can access its source code freely, update it according to his/her needs and compile it. NoSQL implements the concept of replication as it supports distributed data which operates on the data on different machines. As social networking sites and the sites demanding user to sign in have increased, the demand of new technology that can deal with this ever increasing amounts of data has lead the hunt of various new technologies and one of the most preferred one is NoSQL. NoSQL is comparatively faster than relational databases. Earlier in relational databases, we were using Query language to fetch as well as to store data; for NoSQL we store large data entities using documents in XML (eXtensible Markup Language) formats. XML language is basically used to store structured data in a human readable form.

II. CHARACTERISTICS OF NOSQL DATABASES

In computer science, ACID (Atomicity, Consistency, Isolation, and Durability) is a set of properties of database transactions. In the context of databases, a single logical operation on the data is called a transaction. For example, a transfer of funds from one bank account to another, even involving multiple changes such as debiting one account and crediting another, is a single transaction. However, scaling out of ACID-compliant systems has shown to be a problem. Conflicts are arising between the different aspects of high availability in distributed systems that are not fully solvable - known as the CAP- theorem:

A. Consistency

Every read operation on client will return the recent write data.

B. Availability

All clients can always find at least one copy of the requested data, even if some of the machines in a cluster are down.

C. Partition Tolerance

The total system keeps its characteristic even when being deployed on different servers, transparent to the client.

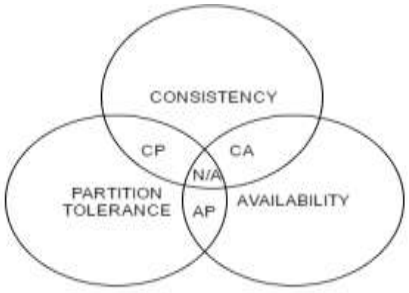

Fig 2.1 Characteristics of NoSQL Databases

The CAP-Theorem postulates that only two of the three different aspects of scaling out are can be achieved fully at the same time.[4] Many of the NoSQL databases have loosened up the requirements on Consistency in order to achieve better Availability and Partitioning.

III. CLASSIFICATION OF NOSQL DATABASES

The NoSQL databases are classified into multiple different types based on their overall structure. We have discussed few of them in this section and they are:

A. Key Value Pair database

A key-value store, or key-value database, is a data storage made for storing, retrieving, and managing associative arrays, a data structure more commonly known today as a dictionary or hash.[3][1] Dictionaries contain a collection of objects, or records, which in turn have many different fields within them, each containing data. If user wants to retrieve any data from

the database user can enter the key value and access the records efficiently. Some examples of key value pair databases are Memcached- Key value stores, Redis- Data structure server, etc.

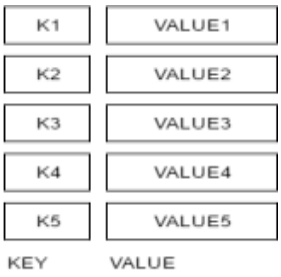

Fig 3.1 Key Value Store

B. Document-oriented database

A document-oriented database, or document store, is a computer program designed for storing, retrieving, and managing document-oriented information.[2][3] Document-oriented databases are one of the main categories of NoSQL databases. Unlike the simple key-value stores described above, the value column in document databases contains semi-structured data-specifically attribute name/value pairs. A single column can house hundreds of such attributes, and the number and type of attributes recorded can vary from row to row. Also, unlike simple key-value stores, both keys and values are fully searchable in document databases. Some examples of document-oriented databases are MongoDB, CouchDB, etc.

Fig 3.2 Document Store NoSQL Database

C. Graph database

Graph databases supplant relational tables with structured relational graphs of interconnected key-value pairings.[3] They are identical to object-oriented databases as the graphs are represented as an object-oriented network of nodes (conceptual objects), node relationships ("edges") and properties (object attributes expressed as key-value pairs). [2]

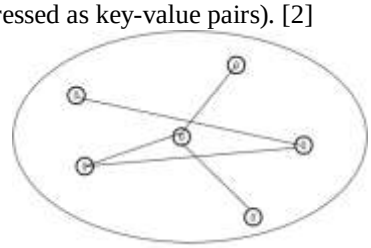

Fig 3.3 Structure of Graph database

They are the only type of NoSQL types discussed above that concern themselves with relations, and their focus on visual representation of information makes them more user-friendly than other NoSQL DMS.

IV. Adoption of NoSQL Database

According to the Couchbase Survey [2] conducted in the year 2012, it is inferred that approximately half of more than 1,300 appellant invested in NoSQL projects in first half of the year. It stated that the reasons given for migrating to NoSQL was lack of scalability and high latency/low performance of traditional model. Almost 40% of overall appellants say that NoSQL is very important or critical to their daily operations, with another 37% indicating it is becoming more important.

V. Implementation of noSQL database

For implementing a NoSQL Database at professional level, the firm begins at a small-scale trial of a NoSQL database, which helps to develop an understanding of the technology in a facile way.[5] NoSQL systems exhibit the ability to store and index arbitrarily big data sets while enabling a large amount of concurrent user requests. Since NoSQL databases are open-source, the implementation and maintenance is handled at low cost. NoSQL modifies with the changing technology, thus the firm can also innovate more rapidly and deliver surpassing customer experience at low cost.

VI. Advantages of nosql databases

1) It provides high scalability and availability
2) It is an open source, thus it saves measurable amount of money by deploying the open source database.
3) The object-oriented approach in NoSQL helps to increase the flexibility of the model.
4) It is cost efficient.
5) Horizontally distributed scale-out architecture helps to distribute the database across multiple geographic regions.

*It provides f*ast and quick result.

VII. Disadvantages of nosql databases

1) Still a developing technology, there are many important features that are yet to be implemented
2) It requires lot of technical skills for both installation and maintenance of the database.
3) Analytical and business intelligence tools fail at certain level to provide connectivity to NoSQL.
4) There are few experienced database developers for NoSQL. Most of the NoSQL developers are in a learning mode.

VIII. Conclusion

The rapidly increasing range of values typically of a large data set has pushed SQL-like centralized databases to their

limits. This led to the development of horizontally scaled out, distributed NoSQL databases. The aim of this paper was to give a thorough overview and introduction to the NoSQL databases, which serves as an effective alternative to the predominant relational database management systems. In this paper we have explored various aspects of NoSQL databases however; there are lots of enhancements to be done in future like polyglot persistence, etc.

REFERENCES

[1] SilvanWeber, "NoSQLDatabases", "http://www.christof-strauch.de/nosqldbs.pdf".
[2] International Journal of Database Theory and Application Vol.6, No.4, " NoSQL Database: New Era of Databases for Big data Analytics-
[3] International Journal of Advanced Research in Computer Science and Software Engineering, "SQL and NoSQL Databases ", https://www.ijarcsse.com/docs/papers/8_August2012/Volume_2_issue_8/V2I800154.pdf
[4] Couchbase Survey from web: http://www.couchbase.com/press-releases/couchbase-survey-shows-accelerated-adoption-nosql-2012."
[5] NoSQL Databases Explained from web: https://www.mongodb.com/nosql-explained.

Anushka Harshad Patil currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai. She has written another paper titled CONTENT FILTERING USING ARITIFICIAL INTELLIGENCE which is under review of IJARCET for issue 9, October 2016.

. **Tanvi Girish Samel** currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai. She has written another paper titled CONTENT FILTERING USING ARITIFICIAL INTELLIGENCE which is under review of IJARCET for issue 9, October 2016. She has completed Oracle Certified Java Professional examination.

Darshan Haresh Vira currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai. He has written another paper titled CONTENT FILTERING USING ARITIFICIAL INTELLIGENCE which is under review of IJARCET for issue 9, October 2016. He has completed Oracle Certified Java Professional examination.