International Journal of Allied Practice, Research and Review
Website: www.ijaprr.com (ISSN 2350-1294)

Data Management Shift using NOSQL

Owais Noor Trumboo[1], Nikita Bakshi[2] and Ravindra Bangar[3]
M.tech Scholar and Assistant Professor
Department of Computer Science & Engineering, Universal Institute of Engineering & Technology, Lalru, Patiala, Punjab, India
Assistant Professor, Pacific University, Udaipur, Rajasthan, India

Abstract - Database is an essential part of modern applications to perform data management operations. Experts have proposed many database models for data management and the relational model is most widely used and acceptable but in present era we need new data management requirements for modern applications that are not supported by traditional databases. To handle these new data management requirements we need some new database approaches like NOSQL approach which is open source, distributed and supports horizontal scaling. Modern applications based on social networking and business intelligence changes the data store strategy to a great extent. The traditional databases are not efficient for handling the Big Data that is the management of data in terms of volume, variety and velocity of data.

Keywords- NOSQL, SQL, Big Data, JSON documents.

I. Introduction

Applications change radically over the last 15 to 20 years based on many dimensions like the number of users, the production of huge amount of data or in short terms Big data etc. There are three mega concepts accountable for the adoption of NOSQL databases- Big Data, Big Users and Cloud Computing. Relational databases supports all the essential characteristics of data management like simplicity, flexibility, compatibility and scalability but fails to provide options like handling huge volumes of unstructured data when implemented in cloud environment. Present computing demands fast enough applications poor application response results in selection of some other options by users. For best user experience the selection of right database is necessary that is very fast, scalable and robust. NOSQL databases have proven the efficient databases for applications that read and write swiftly massive amount of data. Choosing a right NOSQL database for our applications is also the tricky aspect. In this research I am going to analyze the three popular NOSQL database solutions Cassandra, Mongo DB and Couchbase and how these databases behave in terms of interactive web applications.

Discussion on Cassandra, Mongo DB and Couchbase

- **Cassandra:** A distributed columnar key-value database. Read or write of data can be started from any cluster node. No single point of failure is found and supports seamless horizontal scaling. If there is a problem in any node in the cluster another node replaces it.

- **Mongo DB:** Is a schema free, document oriented NOSQL database technology and uses BSON format to store data. A BSON is simply a JSON document in binary format and allows faster and easy integration of data in most applications. Horizontal scalability is also the all feature of this kind of NOSQL database technology.

- **Couch Base:** Is a famous document database for interactive web and mobile applications. Couch base also uses JSON to store documents with low latency read and write operations. No single point of failure is found in the Couch Base architecture. The Couch Base cluster is very easy to scale-out and is based on live cluster topology changes means non application downtime during upgrading the hardware, software or database.

Key Database Criteria for Interactive Applications

There are some important factors to keep in mind when choosing a database for interactive applications:

Scalability:

When website traffic increase suddenly and database is not responding because not having enough capacity, we need to scale database swiftly and without any changes in application. Also having possibility to decrease and optimize the resources when system is at rest. Database scaling needs to be very simple operation means not dealing with complicated database concepts or any change in application.

Performance:

For better performance the database needs to support low latencies regardless of the data size and load. The read and write latency of NOSQL databases is very low because of sharing of data across nodes in clusters that is also the demand of interactive applications.

Availability:

Availability is one of the immense database concept required in modern applications. For high availability system should be able to support online upgrades, for maintenance it is very easy to remove a node without affecting the whole cluster and provide options for backups and disaster recovery in case the whole data center goes down.

Architecture:

Traditional databases are based on rigid schema and for any change in application we need to change database schema as well on the other side NOSQL databases support flexible schema and very simple query language.

The present digital world becomes more complex with Big Data means the volume, velocity and variety of data. The Big Data is a global system and is considered a data collection that increases so large and can't be managed efficiently by traditional databases management systems. The modern

application like social networking, business intelligence has changed the storage system to a great extent. To tackle these problems we have a new kind of databases commonly known as NOSQL databases.

The universe is full with smart devices that generate a huge volume of data and enterprises are dealing with massive amount of data. The data generated by these devices is very essential for an enterprise to compete with other business vendors. With this data enterprises can detect customer behavior and can also analyze various markets with best business potential. In present Big Data environment we are dealing with un-structured data that is very difficult to manage with traditional database technologies and evolve a new data management approach known as NOSQL database technology that can manage terabytes, petabytes and even Exabyte's of data. There are several challenges that enterprises face while dealing with Big Data like:

✓ The security of Big Data which contains very sensitive information, personal information and intellectual property.

✓ There may be delay in speed and availability requirements of Big Data when security mechanisms are operated.

✓ How Big Data is always safe and secure irrespective of where it is stored.

II. Problem Formulation

"Data Management Shift using NOSQL" explores the need to move towards NOSQL environment. In present era we need scale-out database architectures for data management so that we can deploy virtual machines instead of purchasing very costly servers. The term Big Data is not only considered size of data but also the variety and velocity of data. The data volume in an organization grow enormously and there is a requirement of data management for data generated from various fields like medical data, financial data etc. Today the volume of data in most organizations is in the terabyte range but soon can transfer to petabytes or hexabytes. Today in enterprises management of data is not only for structured category but also needs to manage un-structured and semi-structured type of data. The NOSQL databases are the best solution for new data management challenges.

III. Objectives

- To manage huge amounts of data (Big Data)
- To deploy Scale-out architecture.
- To deploy database for Cloud Platform.
- Faster execution of Data.
- To make high availability of Data that is the prime concern for any corporate.

IV. Methodology

Existing Methodology

We are using database concept from decades to perform different data operations. Experts have proposed different models and the most popular one is Relational model. The famous products are SQL Server, MySQL and Oracle that are full filling requirements of all types of applications. The traditional Databases are also known as SQL databases because of query language. The relational database are used from very long time in database driven applications but in present era of computing traditional databases are facing number of challenges like managing huge quantity of data popularly known as Big Data.

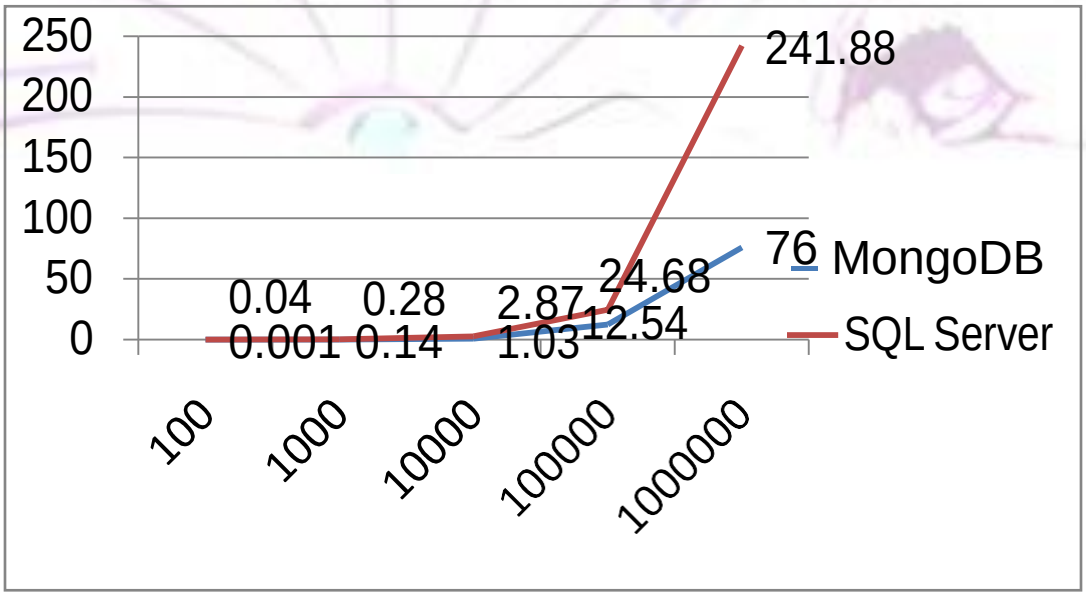
Proposed Methodology

In proposed system we are deploying NOSQL database to overcome the Big Data problem. In this research I am deploying SQL server that is very famous relational database and MongoDB that is very famous document oriented NOSQL database to check the performance based on query processing to analyze how MongoDB suits better for handling Big Data. The operations performed are Read, Write, Select and Delete.

V. Results

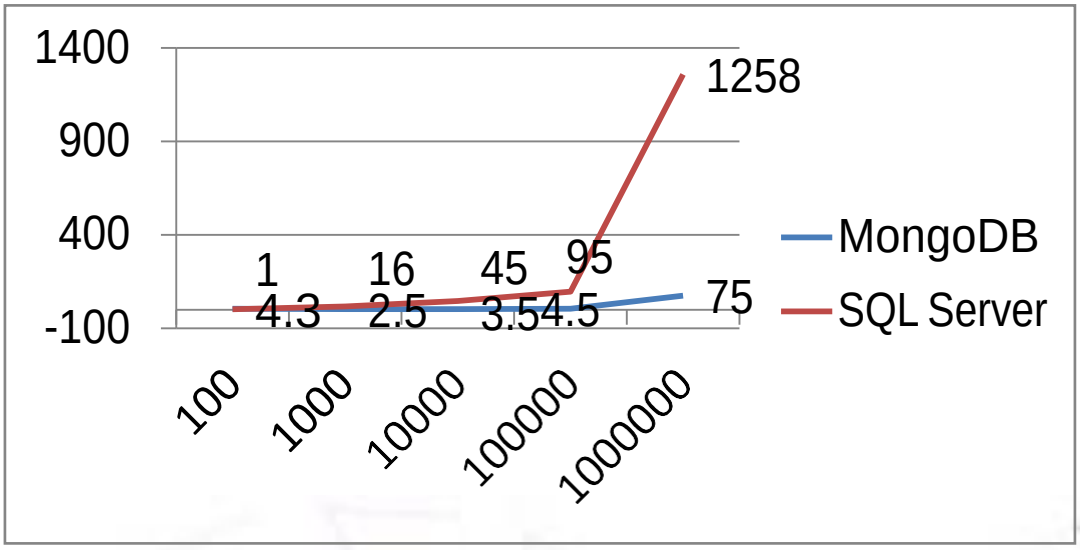
INSERT Operation:

Insert operation is performed on both the databases at scales of 100, 1,000, 10,000, 100,000 and 1,000,000 records for the table Customer in SQL and collection in MongoDB. After performing insert operation at five scales MongoDB performs operations with higher speed as compared to Relational Database. The results are described below:



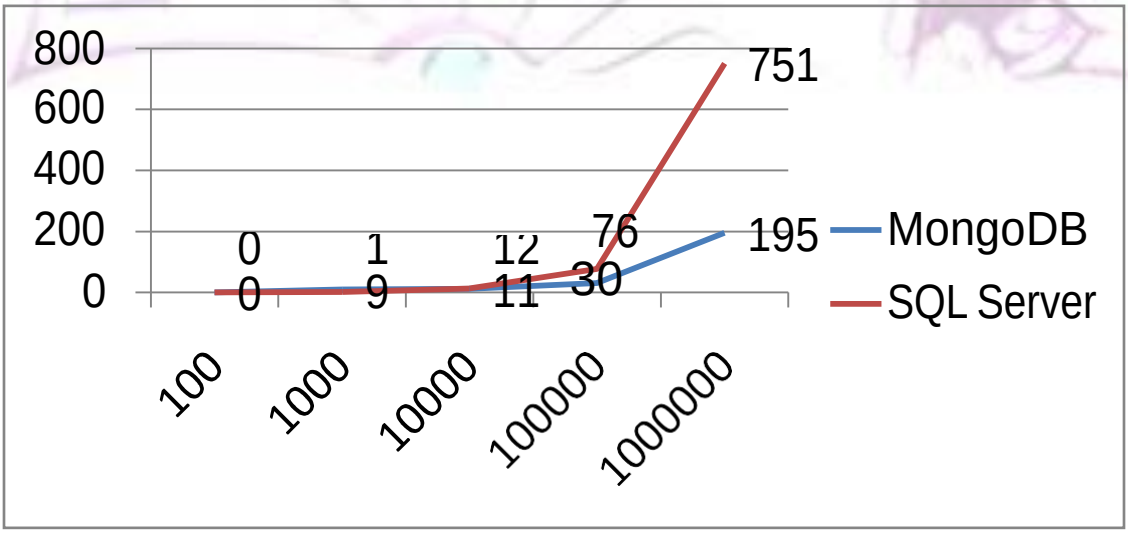
DELETE Operation:

Delete operation is performed on both the databases at scales of 100, 1,000, 10,000, 100,000 and 1,000,000 records for the table Customer in SQL and collection in MongoDB. After performing Deleteoperation at five scales MongoDB performs operations with higher speed as compared to Relational Database. The results are described below:



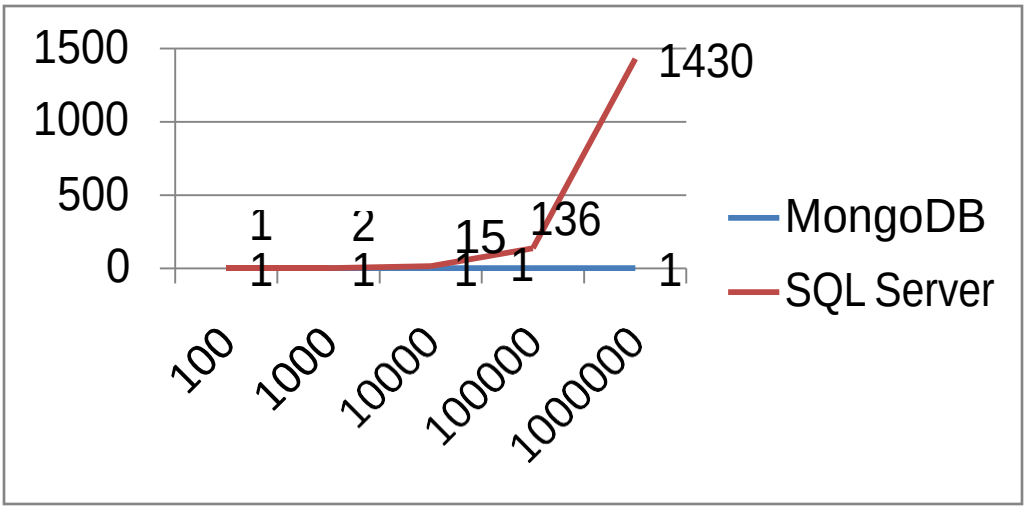
UPDATE Operation:

Update operation is performed on both the databases at scales of 100, 1,000, 10,000, 100,000 and 1,000,000 records for the table Customer in SQL and collection in MongoDB. After performing Updateoperation at five scales MongoDB performs operations with higher speed as compared to Relational Database. The results are described below:



SELECT Operation:

After performing this operation the MongoDB take less time than Relational SQL Server. Most of these tests are performed 100 times and sometimes 10 times and the results are described below.



VI. Conclusion

As for the practical work is concerned the result is very respectable in terms of deployment of NOSQL databases specifically the MongoDB. The modern application requirements need databases that support Big Data, Big Users, Internet of Things and Cloud platform and as for the traditional database technologies are concerned they are not fulfilling the above mentioned challenges.

VII. References

[1] A B Moniruzzaman and Syed AkhtarHossain, "NOSQL Database: New Era of Databases for Big Data Analytics-Classification, Comparison and Characteristics", International Journal of Database Theory and application.

[2] Aaron Schram and Kenneth M. Anderson, "MySQL to NOSQL: Data Modelling challenges In Supporting Scalability".

[3] AmeyaNayak, Anil Poriya, and DikshayPoojary, "Types of NOSQL Databases and its comparison with the Relational Databases", International Journal of Applied Information Systems, Volume 5 No. 4, March 2013, www.ijais.org.

[4] AnkitaBhatewara and KalyaniWaghmare, "Improving Network Scalability using NoSql Database", International Journal of Advanced Computer Research, Volume-2, Number-4, Issue-6, December-2012.

[5] Chad DeLoatch and Scott Blindt, "NOSQL Databases: Scalable Cloud and Enterprise Solutions", August 2,2012.

[6] Chieh Ming Wu, Yin Fu Huang, John Lee, "Comparisons between MongoDB and MS-SQL Databases on the TWC Website", American Journal of Software Engineering and Applications, Volume 4, No 3, April 2015.

[7] Clarence J M Tauro, Aravindh S and Shreeharsha A.B, "Comparative Study of the New Generation, Agile, Scalable, High Performance NOSQL Databases", International Journal of Computer Applications, volume 48-No. 20, June 2012.

[8] David Taniar, "High Performance Database Processing", 26thIEEE International Conference on Advanced Information Networking and Applications, 2012.

[9] DrK.Chitra and B.Jeevarani, "Study on Basically Available, Scalable, and Eventually Consistent NOSQL Databases", Volume3, Issue7, July 2013, www.ijarcsse.com.

[10] Felix Gessert, Wolfram Wingerath, Steffen Friedrich, Norbert Ritter, "NoSQL database systems: a survey and decision guidance", Springer, November 2016.

[11] GuoYubin, Zhang Liankuan, Lin Fengren, Li Ximing, "A Solution for Privacy-Preserving Data Manipulation And Query on NOSQL database", Journal of Computers, Vol 8, No. 6, June 2013.

[12] Hanen Abbes, FaiezGargouri, "Big Data Integration: a MongoDB Database and Modular Ontologies based Approach", Elsevier, September 2016.

[13] InduArora and andDrAnu Gupta, "Cloud Databases: A Paradigm Shift in Databases", International Journal of Computer Science Issues, Vol 9, Issue 4, No. 3, July 2012, www.IJCSI.com

[14] IoannisKonstantinou, Evangelos Angelou, Christina Boumpouka, DimitriosTsoumakos, NectariosKoziris, "On the Elasticity of NOSQL Databases over Cloud Management Platforms (extended version)", Computing Systems Laboratory, School of Electrical and Computer Engineering National Technical University of Athens.

[15] Joao Ricardo Lourenco, Bruno Cabral, Paulo Carreiro, Marco Vieira, Jorge Bernardino, "Choosing the right NoSQL database for thejob: a quality attribute evaluation" Journal of Big Data, Springer, 2015.

[16] Katarina Grolinger, Wilson A Higashino1, AbhinavTiwari,Miriam AM Capretz, "Data management in cloud environments: NoSQLand NewSQL data stores"Journal of Cloud Computing, Springer, 2013.

[17] Leonardo Rocha, Fernando Vale, Elder Cirilo, Darlinton Barbosa, FernandoMourao, "A Framework for Migrating Relational Datasets to NoSQL", Elsevier, Volume 51, 2015.

[18] Liana Stanescu, Marius Brezovan, and Dumitru Dan Burdescu, "An algorithm for mapping the relational databases toMongodb – a case study",International Journal of Computer Science and Applications, January 2017,https://www.researchgate.net/publication/318599517.

[19] LiorOkman, Nurit Gal-Oz, YaronGonen, Jenny Abramov, "Security Issues in NoSQL Databases", International Joint Conference of IEEE TrustCom, 2011.

[20] Marin FOTACHE and Dragos COGEAN, "NOSQL and SQL Databases for the Mobile Applications. Case study: MongoDBVsPostgreSQL", Volume 17, No 2/2013.

[21] NadeemQaisarMehmood, Rosario Culmone, Leonardo Mostarda, "Modeling temporal aspects of sensordata for MongoDBNoSQL database", Journal of Big Data, Springer, 2017.

[22] NaseerGaniee, "New Database Constraints and Modern Applications", IJLTEMAS, Volume III, Issue II, February 2014.

[23] NaseerGaniee, "NOSQL: The Big Data Solution", International Journal of Advancement in Engineering Technology, Management and Applied Sciences, Volume 1, Issue 2, July 2014.

[24] NishthaJatana,SahilPuri,MehakAhuja,IshitaKathuria and DishantGosain, "A Survey and Comparison of Relational and Non-Relational Databases", International Journal of Engineering research and Technology, Vol. 6, Issue 6, August-2012.

[25] Oliver Schmitt and Tim A. Majchrzak, "Using Document-Based databases for Medical Information Systems in Unreliable Environments".

[26] Pankaj Deep Kaura, Gitanjali Sharma, "Scalable database management in cloud computing", Elsevier, 2015.

[27] Paolo Atzeni, "The Relational is dead, SQL is dead, and I don't feel so good myself", Volume 42, No 2, June 2013.

[28] PhyoThandar Thant and Thinn Thu Naing, "Improving the Availability of NoSQLDatabases for cloud Storage".

[29] Rabi Prasad Padhy, ManasRanjanPatra 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.

[30] Rick Cattell, "Scalable SQL and NoSQL Data Stores", December 2011.

[31] Robin Hecht and Stefan Jablonski, "NOSQL Evaluation: A Use Case Oriented Survey", International Conference on Cloud and Service Computing, 2011.

[32] RuxandraBurtica, Eleonora Maria Mocanu, MugurelIonutAndreica, Nicolaetapus, "Practical applications and evaluation of no-SQL Databases in Cloud Computing".

[33] Santhosh Kumar Gajendran, "A Survey on NOSQL Databases".

[34] Sunguk Lee, "Shared-Nothing Vs Shared-Disk Cloud Database Architecture", International Journal of Energy, Information, Communications, Vol 2, Issue 4, November 2011.

[35] Trevor Perrier and FahadPervaiz, "NOSQL in a Mobile World: Benchmarking Embedded Mobile Databases".

[36] Uma Bhat and ShrraddhaJadhav, "Moving Towards Non-Relational Databases", International Journal of Computer Applications, Volume 1-No. 13.

[37] Vatika Sharma and Meenu Dave, "NOSQL and SQL Databases", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 8, August 2012, www.ijarcsse.com.

[38] Wade L. Schulz, Brent G. Nelson, Donn K. Felker, Thomas J.S. Durant,Richard Torres, "Evaluation of relational and NoSQL database architectures to manage genomic annotations", Journal of Biomedical Informatics, Elsevier, October 2016.

[39] Wenbin Jiang, Lei Zhang,Xiaofei Liao,Hai Jin,YaqiongPeng, "A novel clustered MongoDB-based storage system for unstructured data with high availability", Springer, January 2013.

[40] Zane Bicevska, Ivo Oditis, "TowardsNoSQL-based Data Warehouse Solutions", Elsevier, 2017.