



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

Paradigm Shift in Data Management using NOSQL

Assira Amin¹, Mudasir Ahmed Muttoo², Kirti Bhatia³

M.Tech Scholar, Department of CSE¹

Sat Kabir Institute of Technology & Management, Jhajjar, Haryana, India

Assistant Professor, Department of CSE²

SSM College of Engineering & Technology, Jammu and Kashmir, India

Assistant Professor, Department of CSE³

Sat Kabir Institute of Technology & Management, Jhajjar, Haryana, India

Abstract - Database is an essential part of modern applications to perform data management operations. We are using Relational Databases from decades for database driven applications and in present era of application development traditional database deployment results in number of data management challenges. 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. This paper demonstrates the complete deployment of MongoDB database to address the issues of traditional databases and the most important is the Big Data.

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

I. INTRODUCTION

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. Discussion

Discussion on Scalability and Availability of Cassandra, MongoDB and Couchbase NOSQL database technologies

Below are examples how scalability is achieved in Cassandra, Mongo DB and Couch base NOSQL database technologies.

- 1) **Cassandra:** in Cassandra on more capacity nodes can be added faultlessly and the new resources are automatically utilized by the cluster.
- 2) **Mongo DB:** there are a number of functions regarding scalability present in this NOSQL technology like auto-partitioning of data across servers or automatic sharing, distributed reads and writes and eventually consistent reads distributed over replicated servers.
- 3) **Couch-base:** in this type of NOSQL database all cluster nodes are identical and scales horizontally. The task of adding and removing of nodes is very simple and without any change in the application. It is also possible to scale a cluster across datacenters for faster data access because of cross datacenter replication process.

Below are examples how availability is achieved in Cassandra, MongoDB and Couch-base NOSQL database technologies.

- 1) **Cassandra:** A range of data is assigned to each and every node in the cluster. When receiving a write operation chosen to be stored in a node that has failed, automatically route the write request to another node that saves write operation with a hint. A hint is simply a message containing information of failed node and finally the hint holding node monitor the cluster for the node recovery that missed the write request. In case the failed node is online again the hint holding node will handoff the hint message to the recovered node and after the addition of new node to the cluster the workload is distributed to this new node as well.
- 2) **MongoDB:** Consists of arbiter node, a master node and multiple slaves. In case slave node fails to execute, the master node automatically distributes the workload to the rest of slave nodes and if

master node stops to execute the arbiter node chose a new master. The shard is dead in case the arbiter node fails to execute and no node instance is left in the shard.

- 3) **Couch base:** is a document oriented NOSQL database technology and maintains multiple copies of document in the cluster. In this architecture data is equally distributed across all nodes and in case there is a problem in any node in the cluster the server detects the problem and promotes replica documents on rest live nodes to active because each server is identical and serves active and replica documents.

III. The Security and Privacy of Big Data

Security of data is the prime concern for organizations and various security agencies are working on latest and emerging techniques to provide sufficient security options for enterprise data. As for the analysis the NOSQL database technologies are not providing the sufficient security and other the hand the Big Data requires the Big Security. There are some steps that by the security of Big Data can be achieved.

Cryptographic key Protection

These keys act as a gateway for the encrypted data. The data is not safe in any case if the key is unprotected and is a serious problem for enterprises that have grouped together encryption and key management solution means chance of exposing the key within the configuration file or server that stores the data encrypted because any user that access the server can access both key as well as data.

Storing the cryptographic keys on separate server or in the cloud provides the best security option for data.

Data Security at Rest

The e-commerce changes the human life to a great extent and every one use different card for online transactions. Every customer today conduct an online transaction only after seeing the symbol of trust or a certificate indicating the security of transaction so why not the same data is secured at rest in a store. Big Data that includes sensitive information also should be secured and encrypted regardless of where it is stored disk, server or the modern cloud concept.

Proper Control Access

Proper access control mechanism should be implemented in the enterprises because sometimes employees including administrators of cloud, network and server has data access beyond their role. In simple words a user with operating level access to a server does not mean he or she should have access to the Big Data stored on that server.

Trusted Application Development

After encrypted data to control the data access it is time to think about users who access the configuration files to define the access control. Not only implementing the enterprise data encryption, the security of whole environment including applications, services and specially configurations is necessary to give enterprises certain relax that the data is safe from malicious users and rascal employees.

The Gazzang Security Solution for Big Data

The zNcrypt of Gazzang for Big Data is a suitable tool that helps to secure the sensitive information and proprietary data present in data stores both at rest and even in motion regardless of which cloud environment is implemented and the NOSQL database technology is deployed. The Gazzang solution addresses the following aspects:

- ✓ **Transparent and unified data encryption** that helps to protect data at rest as well as in motion.
- ✓ **Very advanced key management** that helps to store keys separately from the data encrypted to ensure privacy of data fulfillment.
- ✓ **Process based access control** that helps to stop data access on the bases of specific process not based on job function and ensures that the data is accessible on to the authorized users that need it.
- ✓ **Platform independence** there are different platforms that can opt for Big stores like the different kinds of physical servers but the modern trend the cloud based servers and pricing is the important aspect for choosing a platform. Platform independence helps to protect our investment and Big Data irrespective of which cloud model and provider is deployed

IV. Results and Discussions

For results and discussions I develop a mini NOSQL DBMS project with PHP programming as front-end and MongoDB for database operations. The objectives are described below:

- The application enables faster system.
- The better accuracy.
- Complete reliable system.
- Very informative system.
- Easy Data Management Operations.

The project is developed on window platform with minimum RAM 256 MB storage capacity of 15 GB. The used software server is EasyPHP (version 14.1VC11) and NOSQL MongoDB(version 2.6.3 or above).

The Data Directory for the entire work is as below:

Collections/Table name:-Staff, Tourists, Animals, Statistics

Front-End includes:

Home Page, New Staff Form, Staff Operations, Animal Form, Tourist Form and Receipt Generation for Tourists

The results after deploying this project are described below:

- MongoDB database used in application can run on commodity server means not need of costly servers.
- The database can be hosted on cloud because supports horizontal scaling means the choice for the cloud.
- The database used is capable of storing structured/un-structured and even semi-structured data means a very good alternative for big Data.

V. Conclusion

In my research work, I did research on the NOSQL technology which is useful in the managing different data management issues that are very difficult to handle by traditional databases. Now at last, I want to give my conclusion here that NOSQL is the best database technology which can be used nowadays in many systems to manage different data management problems specially the Big Data problem. At the end of the thesis, I demonstrate the practical implementation of NOSQL database technology

VI. References

1. <http://www.wikipedia.com>
2. <http://www.oracle.com>
3. NaseerGanee, "NOSQL: The Big Data Solution", *International Journal of Advancement in Engineering Technology, Management and Applied Sciences*, Volume 1, Issue 2, July 2014.
4. NaseerGanee, "New Database Constraints and Modern Applications", *IJLTEMAS*, Volume III, Issue II, February 2014.
5. 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.
6. Paolo Atzeni, "The Relational is dead, SQL is dead, and I don't feel so good myself", Volume 42, No 2, June 2013.
7. 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*.
8. DrK.Chitra and B.Jeevarani, "Study on Basically Available, Scalable, and Eventually Consistent NOSQL Databases", Volume3, Issue7, July 2013, www.ijarcsse.com.
9. Marin FOTACHE and Dragos COGEAN, "NOSQL and SQL Databases for the Mobile Applications. Case study: MongoDBVsPostgreSQL", Volume 17, No 2/2013.

10. 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.
11. 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.
12. 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.
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.IJCSL.com
14. 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.
15. Santhosh Kumar Gajendran, "A Survey on NOSQL Databases".
16. David Taniar, "High Performance Database Processing", 26thIEEE International Conference on Advanced Information Networking and Applications, 2012.
17. AnkitaBhatewara and KalyaniWaghmare, "Improving Network Scalability using NoSql Database", *International Journal of Advanced Computer Research*, Volume-2, Number-4, Issue-6, December-2012.
18. Phyothandar Thant and Thinn Thu Naing, "Improving the Availability of NoSQLDatabases for cloud Storage".
19. Uma Bhat and ShrraddhaJadhav, "Moving Towards Non-Relational Databases", *International Journal of Computer Applications*, Volume 1-No. 13.
20. Aaron Schram and Kenneth M. Anderson, "MySQL to NOSQL: Data Modelling challenges In Supporting Scalability".
21. Chad DeLoatch and Scott Blindt, "NOSQL Databases: Scalable Cloud and Enterprise Solutions", August 2, 2012.
22. Robin Hecht and Stefan Jablonski, "NOSQL Evaluation: A Use Case Oriented Survey", *International Conference on Cloud and Service Computing*, 2011.
23. Trevor Perrier and FahadPervaiz, "NOSQL in a Mobile World: Benchmarking Embedded Mobile Databases".
24. 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*.