

## Analysis of Nosql Database State-of-The-Art Techniques and their Security Issues

Harpreet Kaur<sup>a</sup>

A

Computer Science Engineering Dept, CU, Mohali

**Article History:** Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

**Abstract:** NOSql database systems are extremely optimized for performing retrieval and adjoining operations on large quantity of data as compared to relational models which are relatively inefficient. They are used majorly for real-time applications and statistically analyzing the growing amount of data. NoSQL databases emerging in market claim to outperform SQL databases. In Present time of technology, every person wants to save and secure its data so that no one can check their information without their permission .However, there are multifarious security issues which are yet to be resolved. In this paper, we are discussing and reviewing about the Nosql databases and their most popular security issues link (Cassandra and Mongo DB).

**Keywords:** Nosql database, Security issues, Cassandra, MongoDB.

### 1. Introduction

#### NoSQL

Database represent Not Only SQL database or likewise called as Non Relational or appropriated database that implies NOSQL database are not the counter foundation database. These are archive based record having the dynamic pattern for amorphous information. In recent times, there are many different number of companies which have adopted the Non relational database (NOSQL) as these provide number of advantages to secure our data and files like email, social sites, and multimedia. NOSQL database are basically used for large data sets .NOSQL database is a open source and clusterfriendly which means these Kdatabase's normally work as the simple distribute database

TYPES OF NOSQL :- NOSQL database basically having following categories:-

TYPES OF NOSQL :- NOSQL database basically having following categories:-

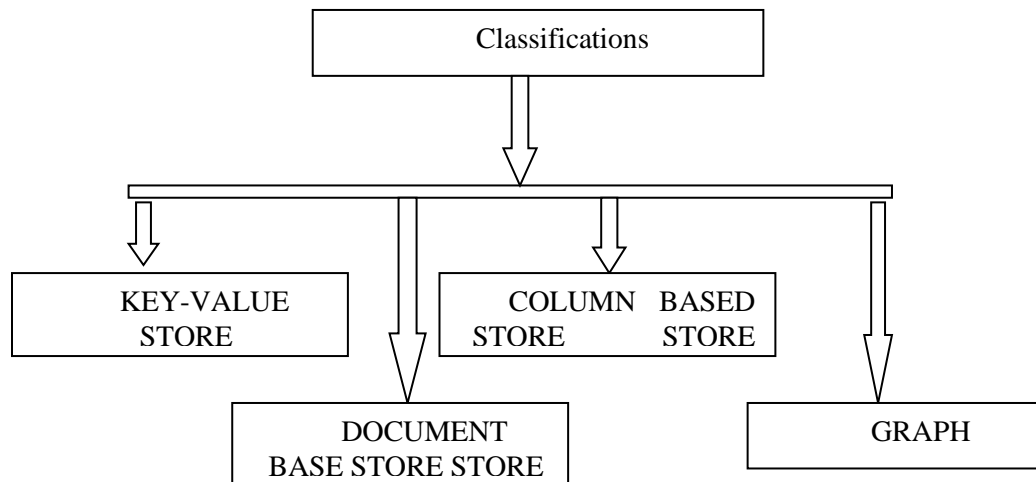


Figure 1:-Types Of NOSql database system

1.1 Key-Value Store:- Although this database is simple one among others stores,yet it is a most prevailing data store. it is used to greatly handling database for its simultaneous access. The data,which stored in it has an unique key to identify data from this store .This concept make it esay to understand and usefulness.

1.2 Document Base Store:- These are thosed element which are made by tagged element (e.g Mongoddb, couchDB, rave



Figure 2:- Types of nosql database[18]

1.3 Column Based Store:- In this we are storage contains only one data column in the database (e.g Cassandra, HBase).

1.4 Graph:- Normally the simple graph it's also having thenodes and edge for store the data (e.g neo4j).

NEED of NOSQL:- In the present time there is easy way to access the data and capture the information to the other useable party such as Facebook, Google+ and other networking sites which are use the to sending the data between the different user or third party . The availbilty of downtime so its remains on. In the Nosql database there is less partion tolerance that's why server having multiple group but these are not communicated each other for data securing purpose[1.2]. These are also be use in the reservation system and banking systems. These databases are very flexible and open source so data is available for every person or any user local or remote[3]. it also have not the particular schema that's the reason we are using the NOSQL database in this present time[4,5] .In Present time every single web application is based on the NOSQL database as it maintain huge data properly.

The remainder of this paper is formed as follows.Second portion , present brief writing overview. Security Isssues are explained in in portion 3.The last part Section 4 is dedicated to summarize.

## 2. Literature survey

Md. Razu Ahmed,et.al,(2018),made literature review on the NoSQL Database for Big Data processing,Out of 18 selected articles after comparison, 8 articles provided required information on the focused area[6].Structural and transactional issues are mentioned with the NoSQL databases for Big Data processing in this paper aiming to provide few guidelines that could become a helping hand for the advent of the Big Data processing.

Hossain Shahriar,et.al(2017), within their paper they highlights the man differences between SQL and Nosql database ,vulnerability that's are associated among databases such as in MongoDB as well as in Cassandra , despite these they described the various types of attacks related to these databases especially to MOOCs online databases[7]. This research work end up with the result that configuration of the data is vitally important to maintain its security.

Rabi Prasad Padhy and Deepti Panigraphy(2015), this research work showcase the Data models ,distinctiveness features and structural design. Furthermore, categorize the database on the CAP rule and certain guidelines to entrepreneurs while selecting particular database to store their sensitive database[8]. They also depicts several security challenges within the paper.

Biswajeet Sethi,et.al,(2014),proposed in this paper,The advent of cloud computing and internet which drives NoSQL movement urging horizontal scaling of relational database. Moreover, NoSQL database is classified based upon CAP theorem[9]. Various data models of NoSQL are compared in this paper to extract robustness and constrains of those models.

Lior Okman,et.al(2011),explained the concept of need of rising demand for storing a huge sensitive data on different database platforms. They put emphasis mainly on two databases: Monogodb, Cassandra and concluded that encryption causes major problem for these datatypes[10]. Also, authentication scheme need to be improved for future uses of these data types between clients and servers.

### 3. Security issue in NoSQL

Every one went to isdata will be secure every time but some time we know that their will be some security issue will be create or our secure data will be show to other.same thing will be happened in the NOSQL database so in his papar we are discuss the this security issue or solve the issue[11,12]. In the NOSQL database we are analyze the two problems in the most popular NOSQLdatabase namely:-(1) MongoDB,(2) Cassandra.

**3.1 Cassandra:-** Cassandra is always run on the top of thenodes these are the combination of the graph and it is the type of column family basically we know that Cassandra will be store the data only one column.or it will be distribute the large amount of data in the commodity servers or available service with no single point of view[13]. So, at this time Cassandra will be design to survive these type of failures. On account of one segment information it won't bolster the full social information model or it will gave the customer a basic information layout. Cassandra will be basically support the search feature of inbox of facebook SO that why we are 200 millions of facebook user will be use the system continuously.

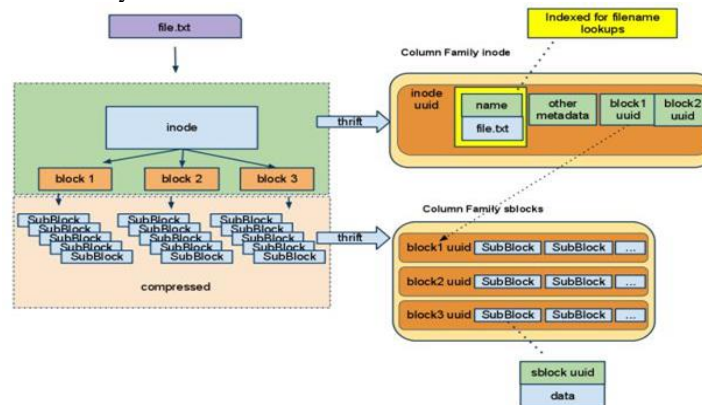


Figure 3:- Cassandra user interface [19]

In this fig we want to described In the inbox of facebook or email or other message site we are follow always on the top of nodes of graph that will help to user to use the system data continuously every time[14] . Cassandra will be providing the key-store unable reliability. keys guide to charge that are bunch into segment relations these segment are combined uniquely to the predetermined keys or these segment having the various information in each extraordinary single section or makes the half and half information the board framework contrast between the vertical oriented or the horizontal oriented store.

#### 3.1.1 Feature of CASSANDRA

3.1.1.1 Symmetric Cassandra:-In which the system software having identical not inside its having no data in the system. It's having not one purpose of disappointment or it is simple punishment

3.1.1.2 Consistent Hashing (Distributed Hash Tables): -Consistent confusion also known as Distributed Hash Tables since it will give Hash Table in way that the expansion or evacuation of information openings doesn't vary the correspondence tonality[15]. It's also having the selectivity decision because it will provide the work on the basis of its capacity. It will work smoothly to the end of the downtime of data.

3.1.1.3 Flexible Partitioning and High Availability:- NOSQLdatabase is a cluster friendly so that why the reasons its having flexible partition or high availability in the data security system .it ill provided the user to partition their data in different replica placement or it is always placed on the next node of the ring.

3.1.1.4 The Client interface Cassandra:- In which systemprovided the RPC Language based on the basis of client interface .because it is not a easy to use so that why the creater assign this RPC interface language to secure our data.

3.1.1.5 Data model(Row , column, super column):-Everysingle topic is based on the data model because data model is the roll-model of the NOSQL database or it the combination of the three required feature like Row, Column,Supercolumn Row is the combination of the column or super columnso in which data is stored is own different file or file will be store in the Row major order[16]. Column is the transient addition information or it will contain just name, timestamp or a value. in which timestamp will be assign to readability of program. Section is the transient addition information.Super Column is a segment whose qualities are columns, that is, an (arranged) affiliated exhibit of sections.

3.2 MongoDB:- In this type of database we will write in the c++language because it's schema free oriented language It allow the data into two part quarry-able or index-able. That will be use in the social sites like email,

facebook etc. In this time hacker will be hack the user account or use the personal data without the permission of user or we know that in present time ransomware attacke

3.2.1 Feature of MongoDB:- There are some daily usefeature of mongodb database that will help the user

3.2.2. Data model :- The data model is the basic need of the database or every programing languageA MongoDB is the set of collection . Or this collection is the basic term or basic need of table but it will not follow the table database because NOSQL database use the Documentation database in the programming. A record include composite arrangement, for example, list or even text or each and every archive having their login id (user customer id)

3.2.3 API MongoDB:- API( Application programmer interface ) in This type of database we the having own query language called the MongoDB query language. And also having are use the RESTFUL.API.REST(Representation State Transfer) is an engineering style for planning net-worked utilization of the MongoDB.It depends on a homeless, customer host, hoard capable interchanges convention (e.g., the HTTP convention). Relaxing applications use HTTP solicitations to post, scan information and erase information.

3.2.4 STRUCTURALDESIGN:-The Architecture of mongodb culsterwill be unlike to the architecture of Cassandra database cluster in which basic difference is no node will be same or store the different date in each ring .as show in fig

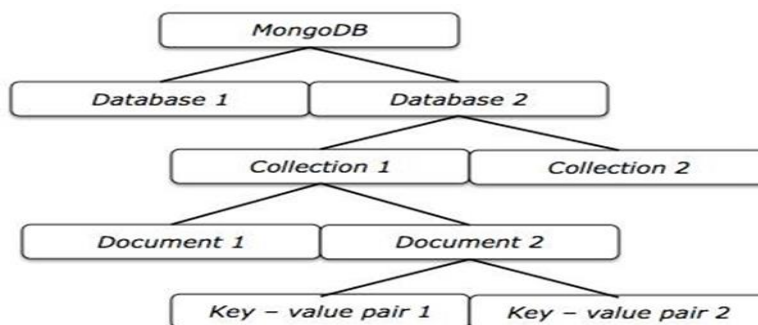


Figure 4:- Architechture of mongodb[12]

### 3.3 Security feature of Cassandra and MondoDB

There are number of security issue will be arise in the front of user to safe her private data and the Cassandra or mongodb database having same type of security issue in the database so that why we are discussed both are together in the papar

3.3.1 Cassandra and MongoDB Data Files :- The Cassandra andMongoDB will not be provide the user to save her data in the storage that means everyone will be access the files in the systm anyone without any permission so it is big issue in the Cassandra system but if we are use the operating system level mechanisms so no unauthorized user not access the data or prevent our data to othr user

3.3.2 Client interfaces :- The client interfacing is the main security feature of the database . in the Cassandra we are uses the Apache Thriftframework for user client communications interfacing to prevent or safe our inbox data to other hacker or unkown person because it support the client interface(login) or also required a user login password to safe her data . and these both username and password will be access by the network system

3.3.3 Inter-cluster communication: - In the inter-clustercommunication, the present steady 0.9 bough can be hold up call for encryption or Authentication in the between group correspondence[17]. It ought to be empowered, and customer endorsements based or it ought to gave be arranged to all bunch part sand client customer of database An interface called fat customer is utilized to stack mass information into Cassandra from various hubs in the group. In the bury bunch correspondence we are not be needed for safety or verification is accommodated fat customers forever phase.

3.3.4 Cassandra Query Language:- CQL is an simply as the SQL-likequery language, compatible with the JDBC (java database contivity ) or API(application programmer interface ). CQL syntax is analogous as much as necessary to SQL since e it is familiar with SQL syntax or it should be at home with CQL

3.3.5 Denial of Service problem:- In the Denial of administration weare utilized a string for each customer model for the system code of database system in the nosql.inasmuch as initial we are install a link needed for the Cassandra server to begin another string on each between the links the Cassandra and mongodb venture prescribes using a type of interconnection conglomerate. In the event that an assailant can attempt to keep the Cassandra server from accepting as the new client customer interconnection by inducing the Cassandra or MongoDB server to assign every one of its assets to counterfeit attachment endeavors or ensure our own information from the assailant.

3.3.6 Authentication:- In Cassandra and MongoDB given an AUTHENTICATE Interface between the user. And in which we are discussed two understanding example for understand the implementation of the database. The first is default (without anyone else) execution is one that spins off the necessity to confirm to the record. Supplementary second is given execution is Simple or typical Authenticator group that permits gear up of clients customer and passwords through a level Java resources folder of database.

3.3.7 Authorization:- The Authorization is the significant piece of the database it will given a solitary arrangement of strategy to restoring a lot of consents for a gave authenticated client customer and given some various leveled rundown of asset names. Similarly as the to the approval code that we are utilize the PC framework. The Cassandra and MongoDB gives two kinds of usage of IAuthority. Formost is a called the go through execution in which it gives consistently permits full authorizations, paying little heed to the client customer , and the other is same as level Java properties document to permit to coordinating the consents to usernames.

3.3.8 Auditing:- MongoDB and the Cassandra both are not given the easiness to her customer for reviewing activities acted in the database in the database When we are make another database (namespace) in this time mongo will compose a line in the log about information record will be made or origination , however after the information documents are allotted in the database, the same old thing shows up in the log for any ensuing inclusions, bring up to-date or questions.

#### **4. Conclusion**

Thus, we have talked about the most well known security issues and highlight functionalities of the two most famous database's Cassandra or mongodb. The most well-known issue in both the databases is absence of encryption in our information documents and making a frail validation between the customer client and the servers and between some other server individuals. Simple approval without help for RBAC (Role based access control) and powerlessness to SQL infusion and Denial of Servicing is a portion of different issues which may assist programmers with attacking.

#### **References**

1. Ahmed, Md Razu, et al. "A literature review on NoSQL database for big data processing." *International Journal of Engineering & Technology* 7.2 (2018): 902-906.
2. Cassandra File System Design online. Accessed from <https://www.datastax.com/blog/2012/02/cassandra-file-system-design>
3. Dadapeer, N. M. "Indravasan, and G. Adarsh," A Survey on Security of NoSQL Databases, "." *Journal of Innovative Research in Computer and Communication Engineering* 4.4 (2016): 5249-5254.
4. David Lebron and Hossain Shahriar, "Comparing MOOC-Based Platforms: Reflection on Pedagogical Support, Framework and Learning Analytics," *Proc. of 2015 International Conference on Collaboration Technologies and Systems, Atlanta, USA, June 2015, IEEE Press*, pp. 167-174.
5. Eassa, Ahmed M., et al. "Nosql racket: A testing tool for detecting nosql injection attacks in web applications." *International Journal of Advanced Computer Science and Applications* 8.11 (2017).
6. F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. Gruber, "Bigtable: A distributed storage system for structured data," *ACM Transactions on Computer Systems (TOCS)*, Vol. 26, Issue 4, pp. 1–26, June 2008.
7. Gupta, Shrankhla, Nikhil Kumar Singh, and Deepak Singh Tomar. "Analysis of NoSQL Database Vulnerabilities." *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT)*. 2018.
8. Han, Jing, et al. "Survey on NoSQL database." *2011 6th international conference on pervasive computing and applications*. IEEE, 2011.
9. L.Yishan and S. Manoharan, "A performance comparison of SQL and NoSQL databases," *Proc. of IEEE Communications, Computers and Signal Processing (PACRIM)*, 2013, pp. 15-19.
10. Mohamed, Mohamed A., Obay G. Altrafi, and Mohammed O. Ismail. "Relational vs. nosql databases: A survey." *International Journal of Computer and Information Technology* 3.03 (2014): 598-601.
11. MongoDB Security Concept. [Online]. Accessed from <http://docs.mongodb.org/master/core/security/>
12. NoSQL Concept. [Online]. Accessed from <https://zhuanlan.zhihu.com/p/92946573>.
13. Okman, Lior, et al. "Security issues in nosql databases." *2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*. IEEE, 2011
14. Padhy, Rabi Prasad, ManasRanjan Patra, and Suresh Chandra Satapathy. "RDBMS to NoSQL: Reviewing some next-generation non-relational databases." *International Journal of Advanced Engineering Science and Technologies*, 2015
15. Saxena, Upaang, and Shelly Sachdeva. "An Insightful View on Security and Performance of NoSQL Databases." *International Conference on Recent Developments in Science, Engineering and Technology*. Springer, Singapore, 2017.
16. Sethi, Biswajeet, Samaresh Mishra, and P. Patnaik. "A study of NoSQL database." *International Journal of Engineering Research & Technology (IJERT)* 3.4 (2014): 1131-1135.
17. Shahriar, Hossain, and Hisham M. Haddad. "Security Vulnerabilities of NoSQL and SQL Databases for MOOC Applications." *International Journal of Digital Society (IJDS)* 8.1 (2017).
18. Sharma and M. Dave, "SQL and NoSQL Databases," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 8, pp. 20-27, 2012.
19. Tudorica, B. G.; Bucur, C., "A comparison between several NoSQL databases with comments and notes," *Roedunet International Conference (RoEduNet)*, 2011 10th, vol., no., pp.1,5, 23-25 June 2011.