© 2017 IJSRSET | Volume 3 | Issue 7 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology
National Level Technical Conference on Advanced Computing Technologies- n'CACT'17, Department of Computer Science & Engineering, Ammini College of Engineering, Kannampariyaram, Mankara, Palakkad, Kerala, India
In association with
International Journal of Scientific Research in Science and Technology

Performance Evaluation of MongoDB and CouchDB Databases

Shyama M Nair[1], Rinu Roy[2], Dr Surekha Mariam Varghese[3]

[1]Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
[2]Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
[3]Professor and HOD, Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

The rise of new type of databases have seen over the last few years, known as NoSQL databases. NoSQL databases are challenging the dominance of relational databases. Scaling a relational databases on powerful servers are expensive and difficult to handle and Nosql databases are designed to expand horizontally. Also Nosql database is schema less and data can be inserted without any predefined schema. They represent a broad category of databases which allow large quantities of unstructured and semi structured data to be stored and managed. Additionally they are designed to handle high levels of reads and writes while scaling horizontally. There are various open sourced and licenced document oriented NoSQL databases , but all have different mechanism to store data in document format. However it is extremely diligent to decide which is to be used and when. So there is need for performance evaluation of various document oriented databases. This work comprises about a detailed comparative study between MongoDB and CouchDB, two leading document oriented databases.

Keywords: NoSQL, MongoDB, CouchDB , RDBMS.

I. INTRODUCTION

NoSQL, an abbreviation of 'not only sql' describes a wide variety of database technologies came into exist in order to overcome the shortcomings of RDBMS and the demands of modern software development. A NoSQL database provides a mechanism for storage and retrieval of data that is modelled in means other than the tabular relations used in relational databases. They provide various advantages over traditional relational databases.Schema agnostic,scalability, performance and high availability are a few features of NoSQL databases.Nosql databases can be broadly classified into four types.They are key-value store, document store, column store and graph based. Among them a document oriented database is designed for storing, retrieving and managing document-oriented data and and there is an additional level of key-value indexing that allows much more efficient queries. The central concept of document-oriented database is the notion of a document.

Choosing right database for right application plays a very critical role as it is the factor that constitutes the platform for analyzing the performance of the application that is under consideration.Some applications might need consistency and availability,some might need availability and partition tolerance and so on. This leads to a complicated situation to choose one database from the many options that needs a good domain knowledge.There are various metrics that are to be considered for performing comparative performance analysis.The metrics include both the qualitative metrics as well as quantitative metrics. Some of the commonly used qualitative metrics are Persistence, Replication, High Availability, Transactions, Rack-locality awareness, Implementation Language, Influences / sponsors , License type and the quantitative measures include size and performance measurements.

Two leading Nosql document oriented databases - MongoDB and CouchDB are used for performance analysis. The metrics that we have taken into consideration are quantitative measures. For comparing the insertion rate (processing time), read / write operations of the MongoDB and CouchDB, a framework written in Javascript, NodeJS, with performance measuring tool Apache JMeter, is used. Some performance evaluation tests have been carried out.Though the database sizes used for the analysis are comparatively smaller, a clear difference in various factors of comparison has been observed.The environment used for conducting these tests was same for both MongoDB and CouchDB.

The remainder of this paper is organized as follows. Section 2 describes related work. Section 3 introduces proposed work which is followed by section 4 with benchmarking result and analysis. Finally, we conclude and explain our future work.

II. RELATED WORKS

Tudorica, Bogdan George, and Cristian Bucur[2]compares various NoSQL systems. The NoSQL database focused to offer high performance and high availability. Although the SQL and the NoSQL databases are having some shared features, some of their behaviors are not similar in given instances. This paper is trying to comment on the various NoSQL (Not only Structured Query Language) systems and to make a comparison (using multiple criteria) between them. The NoSQL databases were created as a mean to offer high performance and high availability at the price of loosing the ACID (Atomic, Consistent, Isolated, Durable) trait of the traditional databases in exchange with keeping a weaker BASE (Basic Availability, Soft state, Eventual consistency) feature.

Enqing Tang and Yushun Fan[3] compares performance between five NoSQL databases(Redis, MongoDB, CouchBase, Cassandra, HBase) by using a measurement tool - YSCB (Yahoo! Cloud Serving Benchmark)and explain the experimental results by analyzing each database's data model and mechanism.

Hecht, R., &Jablonski, S. [4]presents a survey on security issues in big data and NoSQL databases and it also evaluate underlying techniques of Nosql databases. Data encryption is lacked in most of the NoSQL databases. To have a more secure database it is essential to encrypt sensitive fields in the database. Due to the high volume, variety and velocity of big data, traditional security models have difficulties in dealing with large scale of data."Use the right tool for the job" is the propagated ideology of the NoSQL community, because every NoSQL database is specialized on certain use cases.

Leavitt.N suggested that Big data is considered to be large volume of structured and unstructured data. Hence such a large scale of data cannot be effectively managed or exploited using conventional data management tools[5]. To handle this problem, specifically designed alternative database; such as -NoSQL and Search-based systems can be used. The author provided some advanced features of NoSQL databases and showcased how NoSQL databases can live upto their expectations when these new conditions are encountered.

Moniruzzaman, A. B. M., and Syed Akhter Hossain[1]presents classification, characteristics and evaluation of NoSQL databases in Big Data Analytics. This paper provides an understanding of the pros and cons of various NoSQL database approaches; also provides a overview of the non-relational NoSQL databases.

Strauch, Christof, Ultra-Large Scale Sites, and Walter Kriha [6]provides an overview of the motives and rationales,common concepts, techniques and patterns as well as several classes of NoSQL databases (key-/value-stores, documentdatabases, column-oriented databases). There are lot of parameters taken into consideration that includes both the qualitative[7] and quantitative features. In this work we present the qualitative features alone and we propose a system for a streaming application that uses both these databases in order to compare the performances when they encounter various types of queries. This makes sense as we compare two document oriented NoSQL databases in the same environment.With the advent of Big Data, many schema-less, structure-less data were growing prodigiously. So, the effective storage and processing of such data were not possible with the existing RDBMS. The looming of NoSQL databases proved to be one of the best solutions for handling these kind of schema-less

data. This work comprises about the various characteristics of NOSQL databases.

III. PROPOSED WORK

In this section, we work on comparing MongoDB and CouchDB - the two leading document oriented databases taken under experimentation.The reason for considering these two databases are to understand for which application, which database suits well.Web applications is taken as a reference application which involves NodeJS.

NodeJS is an open source, cross-platform runtime environment for developing server-side web-based applications. The application is developed using NodeJS and Express. Express is a framework for building web applications on top of Node js. It simplifies the server creation process that is already available in Node. Node allows us to use JavaScript as our server-side language. MongoDB and CouchDB are databases. This is the place where you store information for your web websites(or applications). CRUD is an acronym for Create, Read, Update and Delete. It is a set of operations we get servers to execute (POST, GET, PUT and DELETE respectively).

This is what each operation does:

- Create (POST) - Make something
- Read (GET)_- Get something
- Update (PUT) - Change something
- Delete (DELETE)- Remove something

The resultant JSON file from the application are loaded into MongoDB and CouchDB.

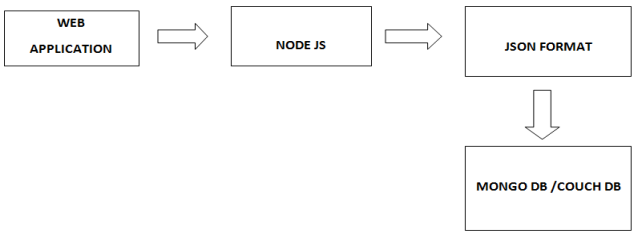

Figure 1: Proposed Framework

The System Architecture in Figure 1 shows that data is accessed through the NodeJS platform. Once the information is captured, they appear in the form of JSON documents.These JSON documents are stored in MongoDB andCouchDB to carry out the performance comparison.

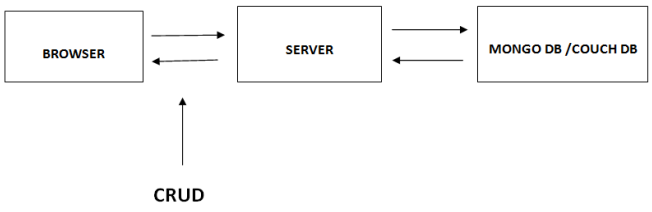

Figure 2: CRUD operations

In Figure 2 CRUD, Express and MongoDB/CouchDB are combine together into a single diagram.

IV.RESULTS AND DISCUSSIONS

We have selected few quatitative features for listing out the comparisons between the two databases. A comparison between the time for insertion, deletion, updation and retrieval are done. The following graphs show a comparative analysis between the two databases.

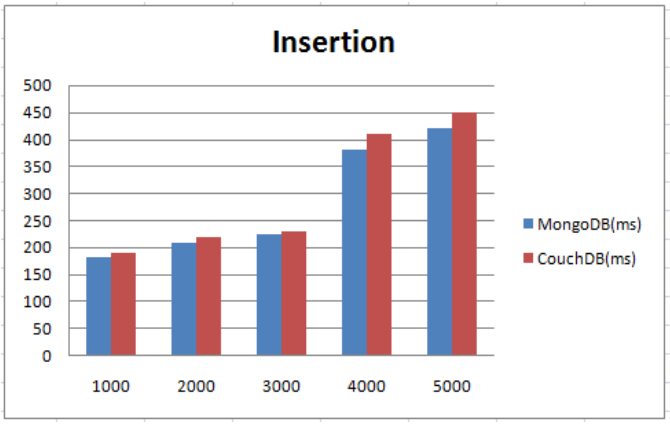

Figure 3: Insertion

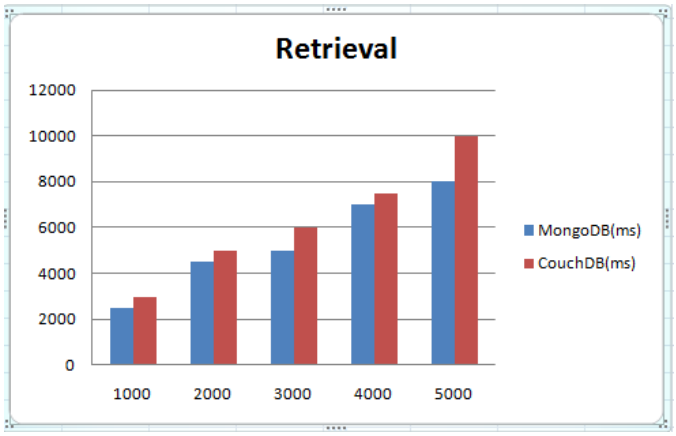

Figure 4: Retrieval

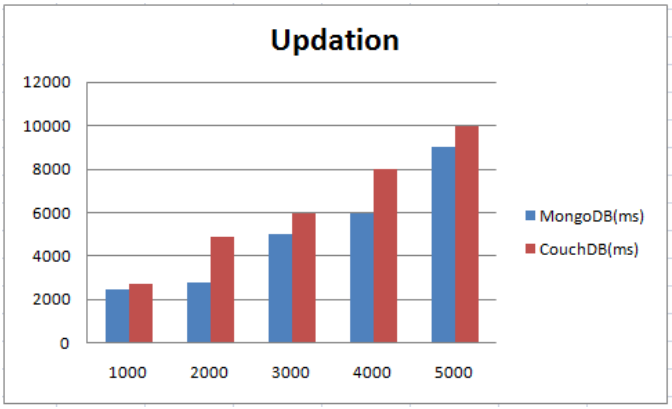

Figure 5: Updation

Figure 6: Deletion

The above graphs from Figure 3to 6 shows a comparison between the two databases under study. As discussed,we have taken a web application and performed this comparison.Hence from the graphs, we can conclude that MongoDB performs well for these kind of applications.

V. CONCLUSION

The quatitative features of both the document-store databases are analyzed in this work. It was very tough to provide a comparative analysis of NoSQL databases. Hence comparisons within NoSQL databases are performed. Further, quantitative attributes like size of the data stored in both the databases and how the databases perform when various types of queries encountered are analyzed. The results suggests that MongoDB clearly have an advantage over CouchDB. This is very much evident from the graphs shown in the above section. Since CouchDB doesn't have any options for bulk-importing JSON documents, it proves a failure model for these kind of web applications.

VI.REFERENCES

[1] Moniruzzaman, A. B. M., and Syed Akhter Hossain. "Nosql database: New era of databases for big data analytics- classification, characteristics and comparison." arXiv preprintarXiv:1307.0191 (2013).

[2] Tudorica, Bogdan George, and Cristian Bucur. "A comparison between several NoSQL databases with comments and notes." Roedunet International Conference (RoEduNet), 2011 10th. IEEE, 2011.

[3] Enqing Tang and Yushun Fan, “Performance Comparison between Five NoSQL Databases,” 7th International Conference on Cloud Computing and Big Data 2016.

[4] Hecht, R., & Jablonski, S. (2011, December). NoSQL evaluation: A usecase oriented survey.In Cloud and Service Computing (CSC), 2011 International Conference on (pp. 336-341).IEEE.

[5] Leavitt, N. (2010). Will NoSQL databases live up to their promise? Computer,43(2),12-14.

[6] Strauch, Christof, Ultra-Large Scale Sites, and Walter Kriha. "NoSQL databases." Lecture Notes, Stuttgart Media University (2011).

[7] Sundhara Kumar K B, Senthil Kumar V, Srividya, Mohanavalli.S,"Comparison of NoSQL Databases", Proceedings of National Conference on Communication and Informatics - 2016, Sri Venkateswara College of Engineering, Sriperumbudur.