



Manipulating Big Data Using No SQL

Prof. Shivkumar Anadur

Department of Computer Science & Engineering, BKIT, Bhalki

(Corresponding author: Prof. Shivkumar Anadur)

(Received 28 September, 2016 Accepted 29 October, 2016)

(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: With the rise of Big Data, the utilization of NoSQL (Not just SQL) innovation is rising quickly among web organizations and different endeavors. Advantages incorporate straightforwardness of configuration, level scaling and better control over accessibility. NoSQL database esare progressively viewed as a practical other option to social databases, as more associations perceive that its composition less information model is a superior strategy for taking care of the huge volumes of organized, semi organized and unstructured information, being caught and handled today. For instance NoSQL databases are frequently used to gather and store online networking information. This anticipate means to present the ideas driving NoSQL, gives an audit of applicable writing, highlights the distinctive NoSQL database sorts, and give contentions to and against embracing NoSQL. A little model application has been produced to evaluate the expressed NoSQL advantages and outline the contrasts between the SQL and NoSQLx approaches. The last segment of the paper offers a few conclusions and suggestions for further research to develop our exploration work.

I. INTRODUCTION

1) Existing System

Social databases arrange information in tables, which are comprised of lines and sections. Tables can't have copy lines since this makes ambiguities amid inquiries; to keep this every table has an essential key section that extraordinarily recognizes each record.

Major Drawbacks of the Existing System

- Normalization can affect execution, as extra table joins might be required amid information recovery
- Multiple table legacy requires numerous join operations to acquire all the applicable qualities of a given item.
- SQL databases require a tables diagram to be pronounced before embeddings information

II. PROPOSED SYSTEM

In this anticipate, we propose a model of Document database outline utilizing NOSQL. Archive databases store information as records that are xJSONx like key quality sets; they are like columns in social databases. Records as a rule contain various key worth combines, and keys may xhold xother key qualities. Records are put away in accumulations; these are gatherings of related reports thatx have shared normalx lists.

Accumulations are like tables in social databases. Figure beneath delineates the library lists report database outline. Record databases have an adaptable

blueprint; SQL zatabases requirez a tables pattern to be proclaimed before embeddings information, however accumulations don't force archive structure. This adaptability empowers reports to coordinate the information fields of any element, notwithstanding when the information is generously changed.

Whilst planning database diagrams key choices rotate around the archives structure and how applications speak to the information connections. The connections between document databases can be spoken to utilizing either the installed or reference approaches.

Installed records store related information in a solitary report; this de standardized information model permits frameworks to question and redesign related information in a solitary database operation. Figure beneath delineates that film item records have implanted fields that contain its executive and wholesaler data. The installed model is utilized when there are one-to-numerous connections substances; tyke reports dependably show up inside the guardian record. Inserting conveys better execution for read operations; however reports may develop after creationx and this can affect compose execution.

Major Advantages of the Proposed System

- Documents as a rule contain various key quality matches, and keys may hold other key qualities.
- Document databases have an adaptable outline.

This adaptability empowers records to coordinate the information fields of any element, notwithstanding when the information is generously changed.

- NoSQL has extraordinary execution advantages over Relational Databases and is profoundly prescribed for huge information sets.

III. DESIGN AND METHODOLOGY

Introduction. Frameworks configuration is the procedure of characterizing the engineering, parts, modules, interfaces, and information for a framework to fulfill determined prerequisites. Frameworks configuration could consider it to be the utilization of frameworks hypothesis to item advancement. There is some cover with the controls of frameworks examination, frameworks design and frameworks building. In the event that the more extensive point of item improvement "mixes the viewpoint of promoting, plan, and assembling into a solitary way to deal with item advancement," then outline is the demonstration of taking the advertising data and making the configuration of the item to be made. Frameworks outline is in this manner the procedure of characterizing and creating frameworks to fulfill determined necessities of the client. Until the 1990s frameworks outline had a urgent and regarded part in the information handling industry. In the 1990s institutionalization of equipment and programming brought about the capacity to assemble secluded frameworks. The expanding significance of programming running on non specific stages has upgraded the order of programming building. Object-situated examination and configuration techniques are turning into the most broadly utilized strategies for PC frameworks outline [citation needed] The UML has turned into the standard dialect in article arranged investigation and design.[citation needed] It is generally utilized for displaying programming frameworks and is progressively utilized for high planning non-programming frameworks and organizations.[citation needed] Framework outline is a standout amongst the most vital periods of programming improvement process. The reason for the configuration is to arrange the arrangement of an issue indicated by the necessity documentation. As it were the initial phase in the answer for the issue is the outline of the task. The configuration of the framework is maybe the most basic component influencing the nature of the product. The goal of the configuration stage is to create general outline of the product. It expects to make sense of the modules that ought to be in the framework to satisfy all the framework necessities in an effective way. The configuration will contain the particular of every one of these modules, their collaboration with different

modules and the fancied yield from every module. The yield of the configuration procedure is a portrayal of the product engineering.

The configuration stage is trailed by two sub stages

- High Level Design
- Detailed Level Design

System Architecture Diagram. The underneath figure demonstrates a general piece chart depicting the exercises performed by this anticipate.

The whole engineering has been executed in nine modules which we will find in abnormal state outline and low level configuration in later sections.

This is a model application created utilizing server side programming model and it will be conveyed in one of the openly accessible server like Tomcat or Wild Fly. This model gives a constant introduction on the use of Mongo DB in live applications and its design related settings. This application can be gotten to by any present day web program through sending a HTTP solicitation to the predetermined URL. Application created utilizing server side programming model and it will be sent in one of the uninhibitedly accessible server like Tomcat or WildFly. This model gives a constant presentation on the utilization of Mongo DB in live applications and its arrangement related settings. This application can be gotten to by any present day web program this model gives a constant introduction on the utilization of Mongo DB in live applications and its setup related settings. This application can be gotten to by any cutting edge web program.

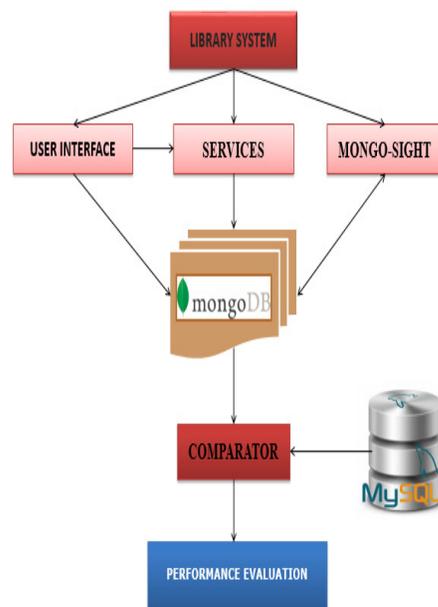


Fig. 1. System Architecture.

This prototype gives a real time exposure on the usage of Mongo DB in live applications and its configuration related settings. This application can be accessed by any modern web browser through sending a HTTP request to the specified URL. Application developed using server side programming model and it will be deployed in one of the freely available server like Tomcat or Wild Fly. This prototype gives a real time exposure on the usage of Mongo DB in live applications and its configuration related settings.

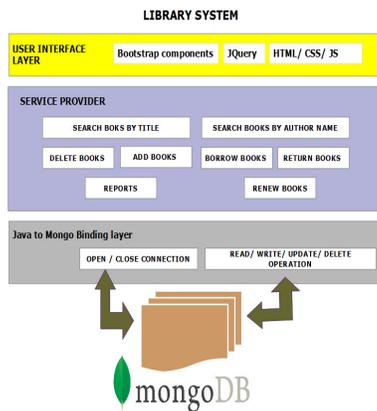


Fig. 2. Library System Architecture.

The server side programming model and it will be sent in one of the openly accessible server like Tomcat or WildFly. This model gives an ongoing introduction on the utilization of Mongo DB in live applications and its design related settings. This application can be gotten to by any current web program through sending a HTTP solicitation to the predefined URL. This model gives a continuous presentation on the use of Mongo DB in live applications and its setup related settings. This application can be gotten to by any cutting edge web program through sending a HTTP solicitation to the predetermined URL.

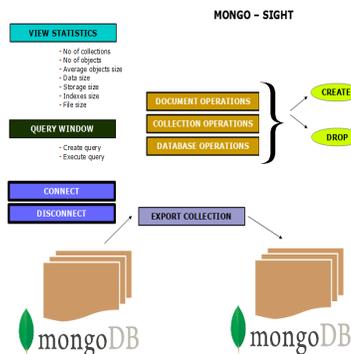


Fig. 3. Mongo Sight.

This model gives a constant introduction on the use of Mongo DB in live applications and its arrangement related settings. This application can be gotten to by any advanced web program through sending a HTTP solicitation to the predefined URL. This application is a revolutionary tool that allows the user of mongo db to visualize the data stored in it and also allows him/her to perform various operations on it like

- Creating and closing connections
- Creating and deleting documents
- Creating and deleting collections
- Executing Mongo DB queries
- Finding and filtering the records
- Exporting the data into other machines

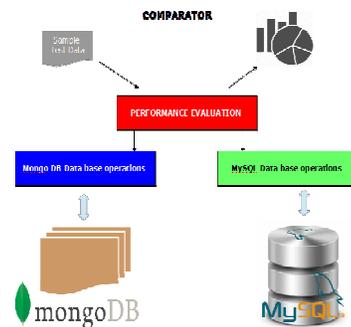


Fig. 4.

This application proves that Mongo DB is better than MySQL in terms of both read and write operations on some sample records. This prototype gives a real time exposure on the usage of Mongo DB in live applications and its configuration related settings. This application can be accessed by any modern web browser through sending a HTTP request to the specified URL.

High Level Design

The System plan stage includes two sub stages

- High Level Design
- Low Level Design

In the abnormal state outline, the proposed practical and non useful prerequisites of the product are portrayed. General answer for the design is created which can deal with those necessities. This part includes the accompanying thought.

•Design thought

Data stream chart

Design thought. There are a few configuration thought issues that should be tended to or determined before getting down planning a complete answer for the framework

Assumptions and conditions. The primary suppositions and conditions distinguished are as per the following:

- JDK must be introduced in all the machines where the interface layer is sent furthermore where the guide lessen program has been executed.
 - The application servers like either the JBOSS or the Apache Tomcat will must be bolstered by the host machines
 - There should not be any firewall or different motors that keeps the remote solicitations from the interface gateway.
 - The RHEL box on which the hadoop group will be designed will be up and running all an ideal opportunity to maintain a strategic distance from the blackouts.
- Mode of operation of a System.** The system is capable of operating in two different modes

Data Flow Diagram 1

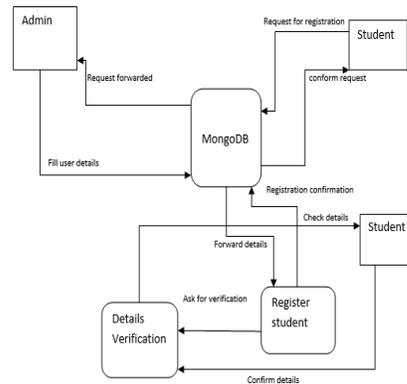


Fig. 5. Library Admin Account.

Data Flow Diagram2

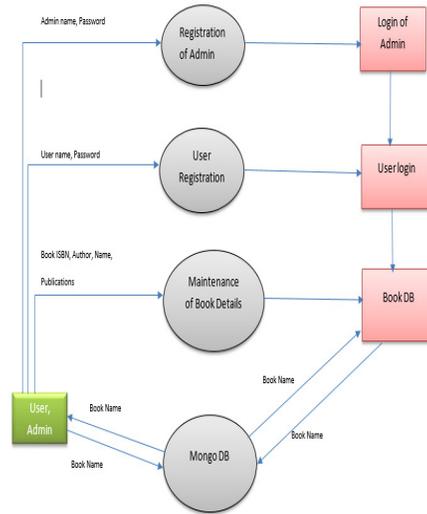


Fig. 6.

NOSQL Prototype: Library Application	This is a prototype application developed using server side programming model and it will be deployed in one of the freely available server like Tomcat or WildFly. This prototype gives a real time exposure on the usage of Mongo DB in live applications and its configuration related settings. This application can be accessed by any modern web browser through sending a HTTP request to the specified URL.
Mongo Sight	This application is a revolutionary tool that allows the user of mongo db to visualize the data stored in it and also allows him/her to perform various operations on it like <ul style="list-style-type: none"> - Creating and closing connections - Creating and deleting documents - Creating and deleting collections - Executing Mongo DB queries - Finding and filtering the records - Exporting the data into other machines etc
Comparison - MySQL vs NOSQL	This application proves that Mongo DB is better than MySQL in terms of both read and write operations on some sample records.

User operations

Access the Library Portal	Only Library Admin can access this
Mongo Sight	Any users of mongo db can be using this tool
Comparison	Any users can access this

Testing. The reason for testing is to find mistakes. Testing is the procedure of attempting to find each possible shortcoming or shortcoming in a work item. It gives an approach to check the usefulness of parts, sub-gatherings, congregations and/or a completed item. It is the procedure of practicing programming with the goal of guaranteeing that the product framework lives up to its prerequisites and client desires and does not fall flat in an unsatisfactory way. The framework has been confirmed and approvedxby running the test information and live information.

Levels of Testing

Unit Testing. Unit testing is a strategy by which singular units of source code, sets of one or more PC project modules together with related control

information, use systems, and working strategies, are tried to figure out whether they are fit for use. Naturally, one can see a unit as the littlest testable part of an application. In item arranged programming a unit is frequently a whole interface, for example, a class, yet could be an individual strategy. For unit testing first we embraced the code testing technique, which analyzed the rationale of system. Amid the improvement procedure itself all the language structure mistakes and so forth got found. For this created experiment that outcome in executing each guideline in the system or module i.e. each way through system was tried. Test cases are information picked aimlessly to check each conceivable branch after all the circles. It gives an approach to check the usefulness of segments, sub-congregations, gatherings and/or a completed item. It is the procedure of practicing programming with the aim of guaranteeing that the product framework lives up to its prerequisites and client desires and does not come up short in an unsatisfactory way.

Steps	Test Action	Results
Step 1	Enter the following URL http://localhost:8080/LibraryPortal/index.jsp	Library portal loaded successfully
Step 2	Click on Register	Registration page loaded successfully
Step 3	Enter username and password and hit register button	Registration was successful
Step 4	Click on Login	Login page loaded successfully
Step 5	Enter invalid username and password	Error message shown
Step 6	Enter valid username and password	Login successful
Step 7	Click on students and try adding and deleting student students	Add and remove operations for students executed successfully
Step 8	Click on books and try adding and deleting student students	Add and remove operations for books executed successfully
Step 9	Click on transaction	Transaction

	and map few students and books	was successful
Step 10	Execute mongo sight	Mongo sight tool loaded successfully
Step 11	Try all the possible actions provided	All the operations are executing successfully
Step 12	Run the following URL http://localhost:8080/Comparison/comapre.jsp	Comparison portal loaded successfully
Step 13	Enter sample records for mysql and nosql and click on COMPARE button	Results shown successfully

User Input. In User Interface the information passage is finished by Providing test understudies and books related information as contribution for the library entrance

Error Handling. In this framework we have attempted to handle every one of the mistakes that happened while running the application. the basic blunders we saw were perusing a tuple with an ascribe set to invalid and database association getting lost. For Testing we utilized Top-Down configuration a deterioration procedure which centers as the stream of control, at last systems worry about code generation. The initial step is to consider the general parts of the current workload and break it into various free modules. The second step is to break one of these modules further into autonomous sub modules. One of the critical elements is that every level the points of interest at lower levels are covered up. So unit testing was performed first and after that framework testing.

Integration Testing. Information can be lost over an interface, one module can adversy affect the other sub capacity, when joined may not deliver the wanted capacities. Coordinated testing is the orderly testing to reveal the blunders with an interface. This testing is finished with straightforward information and created framework has run effectively with this basic information. The requirement for coordinated framework is to locate the general framework execution.

Ventures to perform reconciliation testing

Step 1xxz: Create a Test Plan

Step 2xxz: Create Test Cases and Test Data

Step 3xxz: Once the segments have been coordinated execute the experiments

Step 4: Fix the bugs if any and re test the code
 Step 5: Repeat the test cycle until the components have been successfully integrated

Test cases for integration testing

Name of the Test	Integration testing
Test plan	To check whether the system works properly when all the modules are integrated.
Test Data	Sample students and books related data

System testing. Ultimately, software is included with other system components and the set of system validation and integration tests are performed. System testing is a series of different tests whose main aim is to fully exercise the computer-based system. Although each test has a different role all work should verify that all system elements are properly integrated and formed allocated functions.

Test cases for Input-Output:

Name of the Test	System Testing
Item being tested	Over all functioning of GUI with all functions properly linked.
Sample Input	Sample students and books related data
Expected Output	All the modules like login, execution, etc
Actual Output	Application reacts to user inputs in expected manner.
Remarks	Successful

Validation Testing: At the zenith of discovery testing, programming is totally amassed is as a bundle. Interfacing blunders have been revealed and the right and last arrangement of tests, i.e., approval tests starts. Approval test is characterized with a basic definition that approval succeeds when the product capacity in a way that can be sensibly acknowledged by the client.

Output Testing: In the wake of performing approval testing, the following stride is yield trying of the proposed framework. Since the framework can't be valuable in the event that it doesn't create the required yield. Getting some information about the organization in which the framework is required tests the yield showed or created by the framework is required tests the yield showed or produced by the framework under thought. The yield configuration is considered in two ways, one is on screen group and the other is printed design.

The yield group on the screen is observed to be rectified as the configuration was assigned in the framework has as per the client needs. Concerning the printed copy the yield comes as per the particular asked for by the client. The yield testing does not bring about any redress in the framework.

Test information and Output. Taking different kind delicate information assumes an indispensable part in framework testing. In the wake of setting up the test information framework under study is tried utilizing the test information. While testing, blunders are again revealed and adjusted by utilizing the above strides and remedies are likewise noted for future use.

User acknowledgment Testing

Client acknowledgment testing of the framework is the key variable for the accomplishment of the framework. A framework under thought is tried for client acknowledgment by continually staying in contact with the planned framework at the season of improvement and rolling out improvement at whatever point required. This is finished as to the info screen outline and yield screen plan.

GUI Testing. GUI testing is use to guarantee the visual clarity of the framework, adaptability of the framework, ease of use of the framework. The different parts which are to be tried are:

- Relative format
- Various Links and Buttons

IV. RESULTS

This section describes the results obtained after the execution of implemented our prototype model. The following screenshots define the results or outputs that are got after step by step execution of all modules of the system.

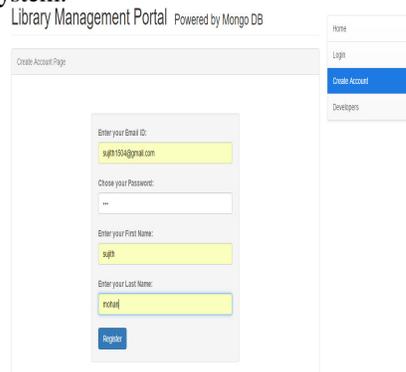


Fig. 7. Creation of Account.

In this First we are going to create Account by Giving E-mail ID, Password, First Name, Last Name.

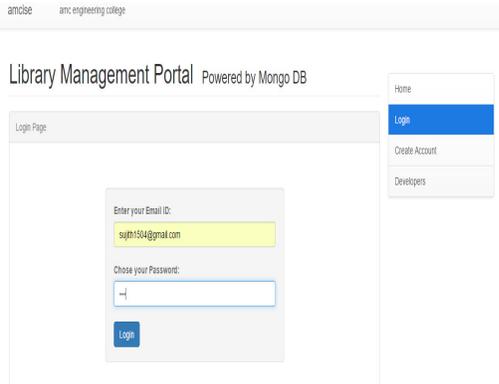


Fig. 8. Login After Creation Of Account.

After Creating the Account We will login through the E-Mail ID, Password which we have Given while Creating the Account

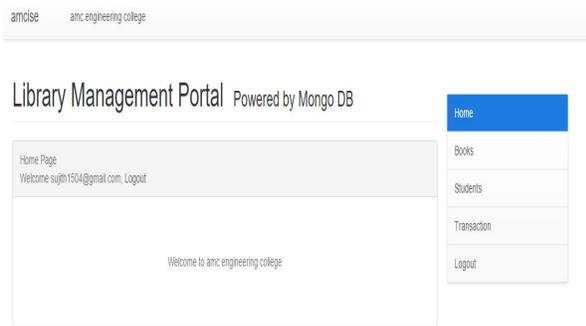


Fig. 9. Login Successful.

After login we will get the operations like Books, Students, Transactions so that we can Add Books, Students, And can do Transactions.

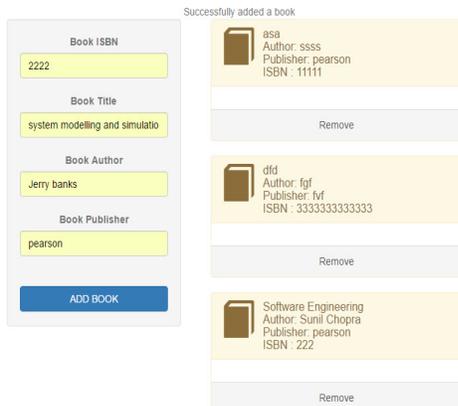


Fig.10. Adding Book.

In this we are Going to Add Book Through BOOK ISBN, BOOK TITLE, BOOK AUTHOR, BOOK PUBLISHER.



Fig.11. Successful Adding Book.

After Entering Book Details Like BOOK ISBN, BOOK AUTHOR, BOOK PUBLISHER in the Previous Step Book will be Added in this Step.

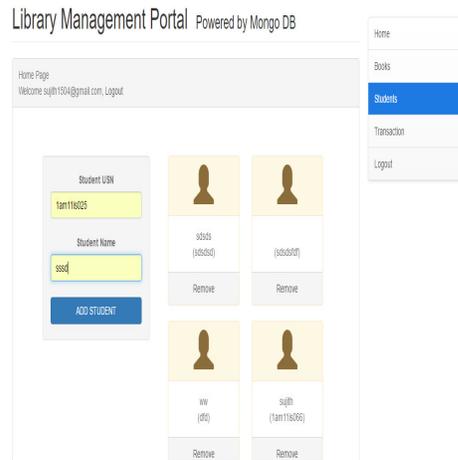


Fig. 12. Adding Students.

In this Step Students Will Be Added based On Student USN, Student NAME so that it will be Easy to Identify the Particular Students.

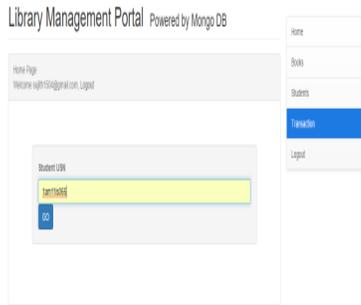


Fig. 13. Transaction.

In this if we give any Student USN we will get to know about the Details like which Book he/she has taken when is the last date to Return the Book.

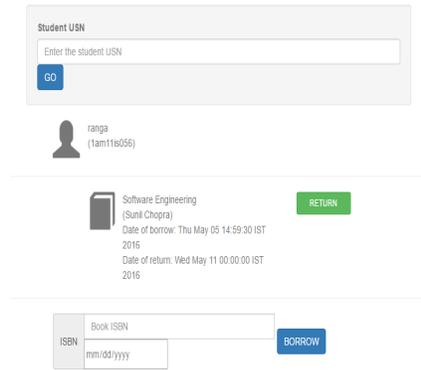


Fig. 14. Student Book Details.

So as u can see here the Student by Name Ranga has taken the book of software engineering on the date 5/5/16 will be returning the Book 11/5/16.

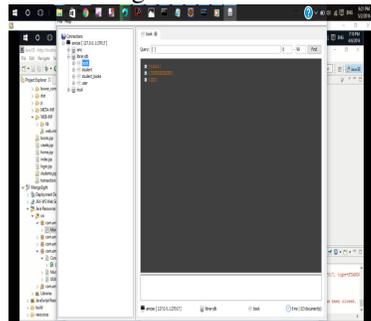


Fig. 15. Mongo Sight .

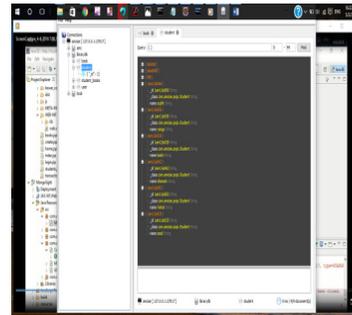


Fig. 16. Mongo Sights 2.

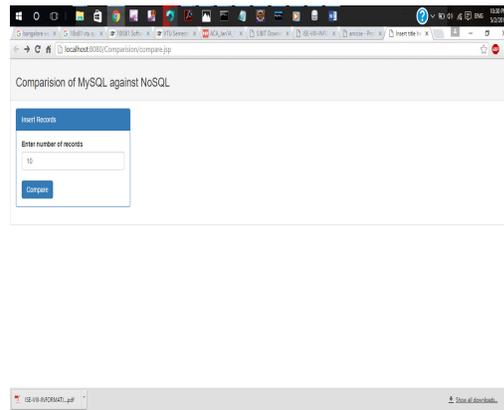


Fig. 17. Entering no of Records.

In this we will be entering the No of Records i.e. Details in this we have Entered 10 Records.

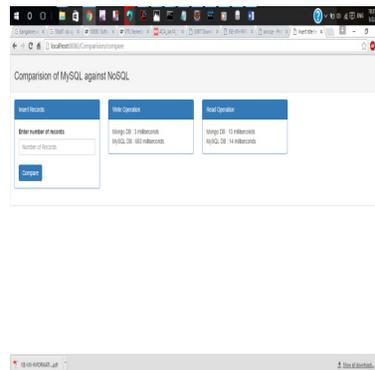


Fig. 18. Record Read Write Operation.

After Inserting the 10 Records we will Get the Read And Write Operation as Shown In Fig.

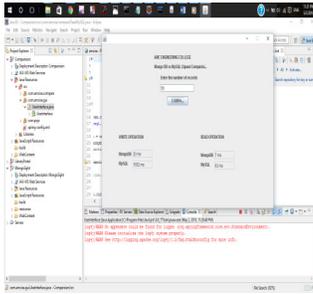


Fig. 19. Comparison.

In this we are going to compare Read and Write Operations of Mongo DB and MySQL. To check the Performance

V. CONCLUSION

The both model frameworks successfully address the libraries data stockpiling prerequisites, yet vary in their methodology. APEX requires table structure to be characterized before including information while xzMongoDBx does not expressly make accumulations report structure is characterized consequently amid the first data embed. MongoDB empowers the library to either embed product characteristics into a solitary archive or reference the data in another report. In any case, with a standardized SQL database, item characteristics may spread over various tables; thusly complex SQL joins are required, to see all the properties of an item. Likewise referential respectability decides indicate that the libraries SQL database ought to just insert records with Author_xzID into the Productxzx_Book table, if thezxAuthor_xzIDzx exists in the Author table.

The xNoSQL database is very productive and permits the library to get all film attributes utilizing one basic query. NoSQL databases make it conceivable to acknowledge awesome value from Big Data and engages organizations to be more coordinated and scalable; this helps organizations accomplish their vital objectives and generate new income streams. Most new information is unstructured and the inflexible mapping based methodology received by social databases, makes it is difficult to fuse all new data types. Sensor information can be utilized for investigating the idea of information assortment further. Future work on

utilizing NoSQL to evaluate the remaining 3V's of Big Data is required. Benchmark tests can be directed to survey the treatment of information volume and speed with the organization of NoSQL on Hadoop bunches and the utilization of continuous online networking information. The online networking information could be as content based product surveys and the model application could be stretched out to fuse opinion examination. NoSQL data bases are turning into a beneficial contrasting option to social databases, as its dynamic information model is greatly improved for overseeing large quantities of unstructured information; moreover, its construction can be modified without downtime or administration depiction.

REFERENCES

- [1]. McKinsey Global Institute, "Big Data." [Report, 2011] New York: McKinsey Global Institute.
- [2]. S. Edlich, "NoSQL Databases." [Report 2010] Chur: University of Applied Sciences HTW Chur.
- [3]. B.G.Tudorica, C. Bucur, "A comparison between several NoSQL databases with comments and notes," *Roedunet International Conference (RoEduNet), 2011 10th*, vol., no., pp.1,5, 23-25 June 2
- [4]. Readwrite, "KeyValue Database." [Image Online Available at <http://xzreadwrite.com/2009/02/12/is-the-relational-xzdatabase-doomed-z#awesmzx=~ookpfnxzVjGmGXHJ> xzx [Accessed: 13 Nov 2013].
- [5]. L. Bonnet, A. Laurent, M. Sala, B. Laurent, N. Sicard "Reduce, you say: What nosql can do for data aggregation and data bi in large repositories." In *xzxDatabase and Expert Systems Applications (DEXA), 2011 22nd International xWorkshop* xzon (pp. 483-488). IEEE.
- [6]. Microsoft, "Column Oriented Database", [Image Online] Available at: <http://xzmsdn.microsoft.com/en-us/xzlibrary/xzdn313285.aspx#sec7> [Accessed: 13 Nov 2013].
- [7]. Couchbase, "Document Database," [Image Online] Available at: <http://www.couchbase.com/why-nosql/nosql-database> [Accessed: 13 Nov 2013].
- [8]. Majestic Media, "Graph Database." [Image Online] Available at: <http://www.xzmajesticmedia.ca/xzblog/2012/06/xznew-xzfacebook-xzxguidelinesfor>.