SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)
DEPARTMENT OF INFORMATION TECHNOLOGY

COURSE CODE: DJ19ITEL6013 **DATE:** 20/03/24

COURSE NAME: Big Data Analytics Laboratory **CLASS:** TY BTech-I1/2

EXPERIMENT 4

CO/LO: To implement NoSQL in mongodb.

AIM / OBJECTIVE:

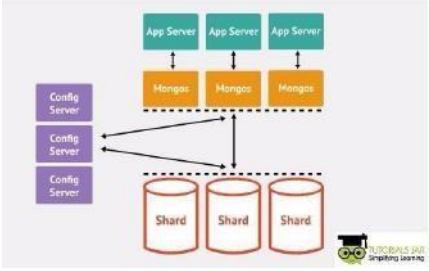
- Store and retrieve different types of data in mongodb.
- Understand the importance of NoSQL.

REQUIREMENTS: PC, Internet and VMWare software, Cloudera.

BRIEF THEORY:

MongoDB - MongoDB is a document database with the scalability and flexibility that you want with the querying and indexing that you need. Offering drivers for all major programming languages, MongoDB allows you to immediately start building your application without spending time configuring a database. MongoDB Atlas is the leading global cloud database service for modern applications. Using Atlas, developers can deploy fully managed cloud databases across AWS, Azure, or Google Cloud. Best-in-class data security and privacy standards practices means that developers can rest easy knowing that they have instant access to the availability, scalability, and compliance they require for enterprise-level application development.

Features



1) Ad-hoc queries for optimized, real-time analytics - When designing the schema of a database, it is impossible to know in advance all the queries that will be performed by end users. An ad hoc query is a short-lived command whose value depends on a variable. Each time an ad hoc query is executed, the result may be different, depending on the variables in question.
2) Indexing appropriately for better query executions - In our experience, the number one issue that many technical support teams fail to address with their users is indexing.
3) Replication for better data availability and stability - When your data only resides in a single database, it is exposed to multiple potential points of failure, such as a server crash, service interruptions, or even good old hardware failure. Any of these events would make accessing your data nearly impossible. Replication allows you to sidestep these vulnerabilities by deploying multiple servers for disaster recovery and backup. Horizontal scaling across multiple servers that house the same data (or shards of that same data) means greatly increased data availability and stability. Naturally, replication also helps with load balancing.
4) Sharding - When dealing with particularly large datasets, sharding—the process of splitting larger datasets across multiple distributed collections, or "shards"—helps the database distribute and better execute what might otherwise be problematic and cumbersome queries. Without sharding, scaling a growing web application with millions of daily users is nearly impossible.

Advantages of MongoDB over RDBMS

1) MongoDB is schema less. It is a document database in which one collection holds different documents.
2) There may be difference between number of fields, content, and size of the document from one to other.
3) Structure of a single object is clear in MongoDB.
4) There are no complex joins in MongoDB.

5) MongoDB provides the facility of deep query because it supports a powerful dynamic query on documents.
6) It is very easy to scale.
7) It uses internal memory for storing working sets and this is the reason of its fast access.

Lab Procedure

i. To see the list of databases in the system:

```
test> show dbs
admin   40.00 KiB
config  12.00 KiB
local   88.00 KiB
test>
```

ii. To select a database:

```
> use myDb
switched to db myDb
```

iii. To find out the currently selected database:

```
> db
myDb
```

iv. To delete / drop a database, make sure you have selected the database and then do this:

```
> db.dropDatabase()
{ "ok" : 1 }
```

iv. To see the collections in a database:

```
> db.createCollection("Exams");
{ "ok" : 1 }
> show collections
Exams
system.indexes
```

SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)

vi. Create a new database named "Exam":

```
> use Exan
switched to db Exan
```

vii. To read data from collection

```
> db.Exan.find().pretty()
{
        "_id" : ObjectId("622307c4f97d3432b4c4e1df"),
        "name" : "syx",
        "score" : 98
}
{
        "_id" : ObjectId("622307cef97d3432b4c4e1e0"),
        "name" : "dyx",
        "score" : 96
}
```

viii. Remove an entry from the database

```
> db.Exan.remove({name:'syx'})
WriteResult({ "nRemoved" : 1 })
```

ix. To get a count of entries in a database

```
> db.Exan.count()
1
```

x. Updating a entry

```
> db.Exan.update({name:'dyx'},{$set:{name:'shweta'}})
WriteResult({ "nMatched" : 1, "nUpserted" : 0, "nModified" : 1 })
```

xi. To drop a database from the data-set

```
> db.dropDatabase()
{ "ok" : 1 }
```

xii. Using where condition to find data

```
> db.Exam.find({"name":'cuc'}).pretty()
{
        "_id" : ObjectId("62230b28f97d3432b4c4c1c3"),
        "name" : "cuc",
        "score" : 93
}
```

SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)

xiii. Using function findOne()

```
> db.Exam.findOne()
{
        "_id" : ObjectId("622307ceF97d3432b4c4e1e0"),
        "name" : "shweta",
        "score" : 96
}
```

xiv. Using save command to fill entries in the database

```
> db.Exam.save({name:'sue',score:93})
WriteResult({ "nInserted" : 1 })
>
```

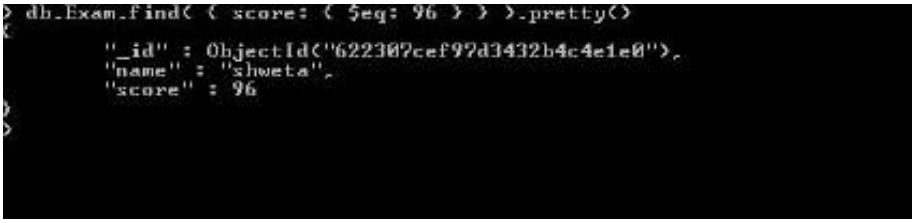
xv. To find a particular entry from database

```
Exam> db.Exam.find().pretty()
[
  { _id: ObjectId("6434fcb00071a5296f573160"), name: 'ABC', Age: 20 },
  { _id: ObjectId("6434fcbe0071a5296f573161"), name: 'ABC', Age: 20 },
  { _id: ObjectId("6434fce30071a5296f573162"), name: 'XYZ', Age: 20 }
```

```
> show roles
{
        "role" : "dbAdmin",
        "db" : "Exam",
        "isBuiltin" : true,
        "roles" : [ ],
        "inheritedRoles" : [ ]
}
{
        "role" : "dbOwner",
        "db" : "Exam",
        "isBuiltin" : true,
        "roles" : [ ],
        "inheritedRoles" : [ ]
}
{
        "role" : "read",
        "db" : "Exam",
        "isBuiltin" : true,
        "roles" : [ ],
        "inheritedRoles" : [ ]
}
{
        "role" : "readWrite",
        "db" : "Exam",
        "isBuiltin" : true,
        "roles" : [ ],
        "inheritedRoles" : [ ]
}
{
        "role" : "userAdmin",
        "db" : "Exam",
        "isBuiltin" : true,
        "roles" : [ ],
        "inheritedRoles" : [ ]
}
>
```

xvii. Use ‘equal to(==) query.

SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



CONCLUSION:

In this experiment, we learnt about NoSQL databases and their advantages over traditional SQL databases. We also learnt about MongoDB, one of the most popular NoSQL databases and studied about its features. We performed various commands under the MongoDB shell to understand how MongoDB works internally.

REFERENCES:

[1] Raj Kamal, Preeti Saxena, “Big Data Analytics”, Mc Graw Hill Education

[2] Alex Holmes “Hadoop in Practice”, Manning Press, Dreamtech Press.

[3] https://www.tutorialspoint.com/mongodb/index.htm