Course name: A level   **SUBJECT: DATABASE TECHNOLOGIES**

Topic: MongoDB   **DATE: 27/5/2020**
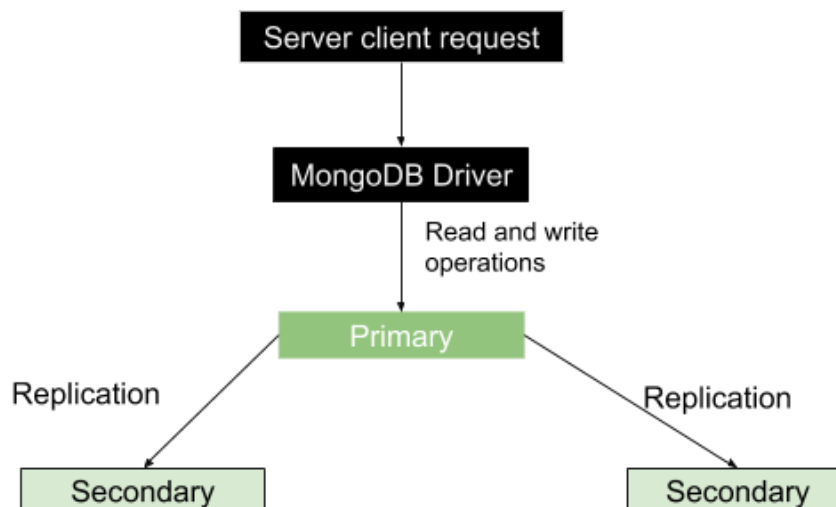
Replication

Replication is the process of synchronizing data or simply duplication (replication) of data across multiple servers. Replication provides redundancy and increases data availability with multiple copies of data on different database servers to meet any unforeseen circumstances. Replication protects a database from the loss of a single server. Replication also allows user to recover from hardware failure and service interruptions. With additional copies of the data, we may dedicate servers to disaster recovery, reporting, or backup.

Replication in MongoDB

Replication is simply the process of ensuring that the same data is available on more than one Mongo DB Server. This is sometimes required for the purpose of increasing data availability and to meet any disaster also. If there is only one MongoDB Server, and it goes down for any reason, there will be no access to the data. But if data has been replicated to another server or to multiple servers, the data from another server(s) will be accessible even if the primary server fails.

A very important aspect of replication is the possibility of load balancing. If there are many users connecting to the system, instead of having everyone connect to one system, users can be connected to multiple servers so that there is an equal distribution of the load.

In MongoDB, multiple MongDB Servers are grouped in sets called **Replica sets**. The Replica set will have a primary server which will accept all the write operation from clients. All other instances added to the set after this will be called the secondary instances which can be used primarily for all read operations.

Advantages of Replication

- To keep your data safe
- 24 by 7 (24 x 7) availability of every data
- Helps in disaster recovery and backup of data.
- Load Balancing
- No downtime for maintenance (like backups, index rebuilds, compaction)
- Read scaling (extra copies to read from)
- Replica set is transparent to the application
- Minimizes downtime for maintenance.

Disadvantages of Data Replication

- High Cost for Hardware and other infrastructure
- Redundant data is stored, so more space and server processing required
- Need connectivity with all the replicated servers

Implementation of Replication in MongoDB

MongoDB implement replication by the use of replica set. A replica set is a group of **Mongod** instances that host the same data set. In a replica, one node is primary node that receives all write operations and all other instances called secondary instances, apply operations from the primary so that they have the same data set. Replica set can have only one primary node and rest all are the secondary nodes. In short,

- Replica set is a group of two or more nodes. A minimum of 3 nodes are recommended.

- In a replica set, one node is primary node and remaining nodes are secondary.

- All data replicates from primary to secondary nodes.

- New primary nodes get elected in case there is automatic maintenance or failover

- Once the failed node is recovered, it again join the replica set and works as a secondary node.

Features of Replica Set

- A cluster of N nodes

- Any one node can be primary

- All write operations go to primary

- Automatic failover

- Automatic recovery

- Consensus election of primary

Assignment

1. What is replication? Where it is used?

2. How Replication is implemented in MongoDB?

3. What are the advantages and disadvantages of Replication?