

Replication Cont'd

Terminologies used in Replication

Heartbeat

Replication is used to keep multiple instances of the database simultaneously running usually on different systems of the same replica set. Heartbeat is the process to identify the current status of such servers(also called node servers) within the replica set. To do so, members of replica set send heartbeats (technically pings) to each other at every **two seconds**. In case, if a heartbeat does not return (i.e. no response) within 10 seconds, then the other members of replica set mark such delinquent member as an inaccessible member.

A heartbeat request is a short message that checks everyone's current state. Heartbeat is the process to know the other member's state, like who's primary, from which member they need to sync from or which node is down. The most important activity of heartbeat is to check that the primary server live and reachable by all secondary servers. In case, if the most of the secondary servers can't reach to the primary server, then the process automatically demotes the primary server as a secondary server.

Elections

Replica set members uses election to determine which set member will become primary as the Primary node is not accessible. If a primary becomes unavailable, elections allow the set to recover normal operations without manual intervention. Elections may also be called as the part of the failover process.

Elections occur anytime when the primary becomes unavailable. It is important to know that the primary is the only member in the replica set which can accept write operations. And while an election is in process, the replica set has no primary and cannot accept writes. MongoDB avoids elections unless necessary.

Election Process

During the heartbeat check, if a member can't reach the primary server then that particular member raises the **election flag** to the other members of the replica set so that other members of the replica set can't raise the same election flag within the process.

If there is no objection against the election request, the other members will vote for the member seeking election. And if the member receives the majority votes from other members, then the election is successful and it will be promoted as the primary node. In case, if the member did not receive the majority of votes then it will remain as secondary node and maybe try to become a primary node in the future.

Elections take time to complete Elections are essential for independent operation of a replica set; however,. while an election is in process, the replica set has no primary and cannot accept writes. MongoDB avoids elections unless necessary. If the network condition is healthy and most of the servers are up, then the entire elections process should happen very quickly. Two seconds is required to notify all the members that primary is down as heartbeat response has not been received yet and the election process starts immediately.

Oplog

The oplog, i.e. operations log, is a special capped collection in MongoDB that keeps a rolling record of all operations that modify the data stored in your databases.

All the database operations are applied to the primary and operations are recorded on the oplog of the primary. All the secondary members then copy and apply these operations in an asynchronous process. To maintain the current state of the database, all replica set members contain a copy of the oplog in the local.oplog.rs collection. Any secondary member may import oplog entries from any other member.

Syncing

A log of operations called oplog is maintained by MongoDB having every write (transaction/ operation) information onto the primary server. As the main aim of the replication process is to keep the same or an identical set of data on multiple servers of the Replica Set, the oplog

helps in achieving this goal. Oplog is a capped collection and is there in the local database on the primary server. The secondary servers query this collection(oplog) to obtain the operation details so that they can replicate that data.

Oplog is also maintained at every secondary server in which MongoDB stores each operation related to the replication process from the primary server. These log files allow any replica set members to use as a sync source for other members. Every time, the secondary server first fetch the information related to the pending operations from the primary member, then apply that operation to their own data set and then write down the logs about that operation into the oplog.

In case, if the secondary server goes down and is going up after some time, then it starts the syncing process the last operation is done by itself according to its oplog file. As the operation first applied to the data and then it writes to the oplog, the secondary server may redo the operations that it has already applied to its data.

Assignment

1. What is HeartBeat? How it is helpful in Election Process?
2. What is Election? When it is carried out. Explain.
3. What is oplog?