Database studypoints 3 assignment

What is sharding in mongoDB?

Sharding is used to distribute data across multiple servers, known as shards, to provide scalability and high availability. It helps to increase the capacity of a MongoDB system to handle large amounts of data and high traffic loads.

What are the different components required to implement sharding?

Config Servers:

These are the MongoDB instances that store metadata and configuration information for the entire sharded cluster.

Shard Servers:

These are the MongoDB instances that store a portion of the sharded data.

Query Routers:

These are the MongoDB instances that act as a proxy between the client applications and the shard servers.

Explain architecture of sharding in mongoDB?

MongoDB uses the shard key to distribute a collection's documents across shards. MongoDB splits the data into “chunks”, by dividing the span of shard key values into non-overlapping ranges. MongoDB then attempts to distribute those chunks evenly among the shards in the cluster.

Provide implementation of map and reduce function

Since MapReduce was deprecated, we used an aggregate instead to find the top 10 students based on their exam score:

```
db.students.aggregate([
  // Project the name and exam score of each student
  { $project: { name: 1, examScore: { $arrayElemAt: ["$scores.score", 0]
} } },

  // Sort the documents in descending order by exam score
  { $sort: { examScore: -1 } },

  // Limit the results to the top 10 documents
  { $limit: 10 }
])
```

Provide execution command for running MapReduce or the aggregate way of doing the same

See above code

Provide top 10 recorded out of the sorted result. (hint: use sort on the result returned by MapReduce or the aggregate way of doing the same)