Q3. What is Sharding in MongoDB, and how does it work?

A) Sharding in MongoDB is a technique used to distribute data across multiple servers (or nodes) in a database cluster to improve performance, scalability, and manageability for large-scale applications with high data volumes and heavy read/write workloads.

Working of Sharding MongoDB:

Sharding Key: The first step is to select a "sharding key." This key is used to determine how the data will be distributed among the shards.

Config Servers: MongoDB requires a special set of servers called "config servers" to manage the metadata about sharding.

Shard Cluster: The MongoDB cluster consists of multiple shard servers and the config servers.

Data Distribution: When a new document is inserted into a sharded collection, MongoDB uses the sharding key to determine which shard should store the document.

Balancing: Over time, as data grows and usage patterns change, the distribution of data among shards may become uneven.

Queries: When querying the sharded collection, the MongoDB router (mongos) receives the query and checks the metadata in the config servers to determine which shards need to be queried to fulfill the request.