Nicole Günther Guerrero

ID: 158325

MongoDB Sharding

1. **What is a sharding key? Is the choice of sharding key directly dependent to the sharding strategy? Explain and give examples.**
   A shard key is either an indexed field or an indexed compound field that exists in every document in the collection. Yes, the choice depends on the strategy that is useful, this can be either, hash based sharding, range sharding or tag aware sharding.
   - You use range sharding when MongoDB divides the data set into ranges determined by the shard key values.
   - You use hash based sharding when MongoDB computes a hash of fields value, and then uses theses hashes to create chunks.
   - You use tag aware sharding when you need to create and associate tags with ranges of the shard key.

2. **Explain how does apparently mongodb choses the number of intervals used to shard a collection in the interval-oriented strategy? In which situations would such a strategy be well adapted for sharding a collection?**
   The number of intervals depend on the attribute chosen that permits intervals, an example is the attribute age. Using the age allows us to separate a group of people by age.

3. **In which situation would the hash-based strategy be interesting for a collection to be sharded?**
   When we want to get an even distribution of the values and when we want to efficiently find or store data.

4. **Which of strategies interval or sharding would lead to a more balanced distribution of data across shards, interval, or hash?**
   Hash partitioning leads to a more balanced distribution of data across shards.

5. **What are the advantages and disadvantages of allowing access to shards directly through their server and not only through the query router?**
   Some advantages of allowing access to the shards directly from the server are:
   - It can support hedged reads to minimize latency.
   - Config servers store metadata and configuration settings for the cluter.

- The deployment of config servers and shards as replica sets provide increased availability.

Some disadvantages are:

- When allowing access directly from the server, you lose the interface that happens when a query router access as a layer between the client application and the sharded cluster.
- Load balancing is harder to obtain.

6. Give an example of a situation where tag based sharding would be an interesting option?
   Tag aware sharding can be used to optimize physical resources. One way to implement it is to increase the read and write efficiency to have a better use of the hardware and improve the overall performance of the application.

7. What happens when a new shard is added to a cluster containing already other shards with data?
   It will automatically become the shard with the lowest number of chunks for every sharded collection.

8. How would you test whether a sharded collection was an interesting solution in comparison to a centralized one?