

No sql db are auto sharded.. but someone has to shard it.. how will we do it?

Key value: base don key

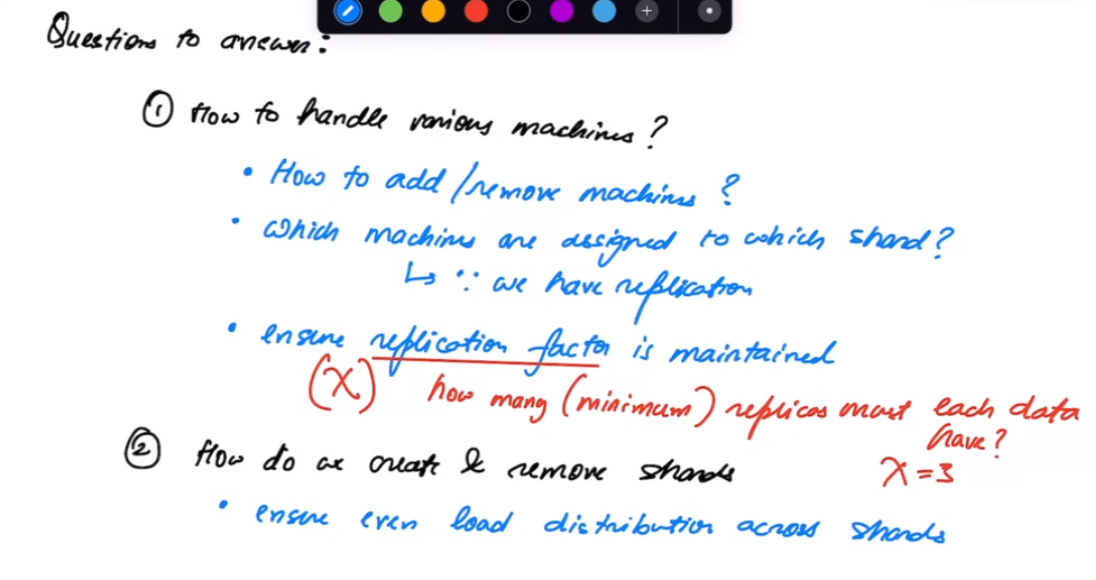
Doc db: shared based on doc id.

When we are sharding it, we have 10, 100, 1000 machines, we have to manage, decide how to add remove machines. Which machines are assigned to which shard.

Can be multiple machines can be in same shard: Replication. Data inside a single shard has to be replicated to multiple times.. if machine fails. We have replica..

We alos have to ensure the replication factor is maintained.. RF = how many minimum replicas must each data have.. given by term X (chai) every price of data is replicated to 3 times at least.. can be 4, 5.

Even load distribution across shards.



To scale a db horizon we need machines…

Q; how do we create and remove Shards:

1. So we can maintain even load distribution across shard..

Create a Dedicated machine Databse Orchestrator – will handle all the above… ORC will figure how to manage machines. If have 100 machine, o bought 25 new machines. Orc will manage that. Manage diff shards. Works as DB cluster coordinator – DB orchestor.

**CH : statefull routing..**

Db = always statefull..

Ch: reduces no of movement.

Here these are db

