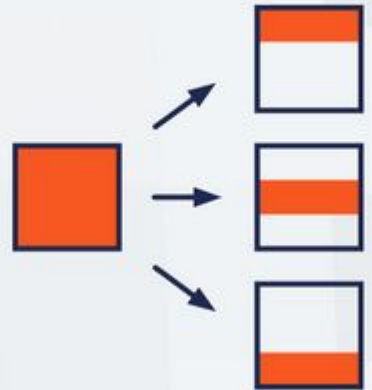


Advantages and Disadvantages of Database Sharding

- Sanjana Sunil (20171027)
- Hitesh Kumar (2019201039)
- Rishabh Daga (2018101015)



Advantages



Orders

Place	Item
Hyderabad	Sony TV
Mumbai	Samsung Phone
Delhi	Samsung TV
Delhi	Acer Laptop
Hyderabad	Casio Watch



Hyderabad Orders

Place	Item
Hyderabad	Sony TV
Hyderabad	Casio Watch



Mumbai Orders

Place	Item
Mumbai	Samsung Phone



Delhi Orders

Place	Item
Delhi	Samsung TV
Delhi	Acer Laptop

Advantages



- **Facilitation of horizontal scaling:** We can now keep the tables in different machines, as opposed to earlier.
- **Separate logical instances:** Shards could be geographically distributed to be closer to customers.
- **Reduction of search space in some cases:** If we want all orders placed in Hyderabad, we don't have to look at all the rows now!

Advantages



- **Load balancing:** Several smaller tables leading to better search load distribution and better performance. Parallelization can also be done.
- **More resilient to outages:** Entire Orders table going down vs. outage in a single shard.

Disadvantages



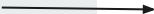
Orders

Place	Item
Hyderabad	Sony TV
Mumbai	Samsung Phone
Delhi	Samsung TV
Delhi	Acer Laptop
Hyderabad	Casio Watch



Hyderabad Orders

Place	Item
Hyderabad	Sony TV
Hyderabad	Casio Watch



Mumbai Orders

Place	Item
Mumbai	Samsung Phone



Delhi Orders

Place	Item
Delhi	Samsung TV
Delhi	Acer Laptop

Disadvantages



- **Implementation complexity:** More code to handle sharding logic. Incorrectly done sharding can cause issues like loss of data. Need to maintain extra information about shards, basis of sharding etc. What if an order is changed from Hyderabad to Mumbai?
- **Operational complexity:** More difficult to backup, change schema, add indexes, add columns etc.
- **Hard to reverse:** Once sharded, it can be hard to go back to unsharded architecture, merge shards etc.

Disadvantages



- **Network latency:** Increased latency due to reliance on communication between nodes.
- **Database hotspots:** What if Delhi has significantly more traffic than Hyderabad? Imbalance, and resharding might have to be done due to slowdowns and crashes.
- **Some searches might perform more slowly:** Queries requiring access to most shards might have worse performance. Join operations can also be trickier since multiple servers might have to be checked.

Thank you!
