**Partitioning**

- Splitting a subset of data within the same machine

**Partition Type**

- Horizontal partition: Row level or document level

- Vertical partition: Column level or table level

**Sharding**

- Distributing data across multiple machines

- Breaking up large table into partitions and storing each partition on separate servers is called sharding

**Sharding Types**

- Automatic and Manual

**Example**

- **Geo based** sharding (based on user location) (static location is choosen location when account was created)

- **Range based** sharding (based on first character of user name which creates 26 bukets) | may result in uneven shard thus leading to hotspot (one larger shard)

- **Hash based** sharding (creating a hash which distributes data evenly acorss all partitions)

**Advantages**

- sharding allow scalability

- small data on each shard results faster performance

- reliability and accessibility (if one shard is down it won’t impact other shard)

- Node run on commodity hardware (lower price of hardware)

**Disadvantages**

- May create hotspot (bigger shard) which is expensive to manage (operational complexity)

- cross shard queries are expensive

- not all data is amenable to sharding (foreign key relation need to be in single shard)

- once sharding is setup, it can be difficult to undo (higher cost when compared with single RDBMS)