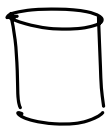
Agenda.

- → Sharding
- → Replication
- → CAP Theorem.
- → PACELC Theorem.

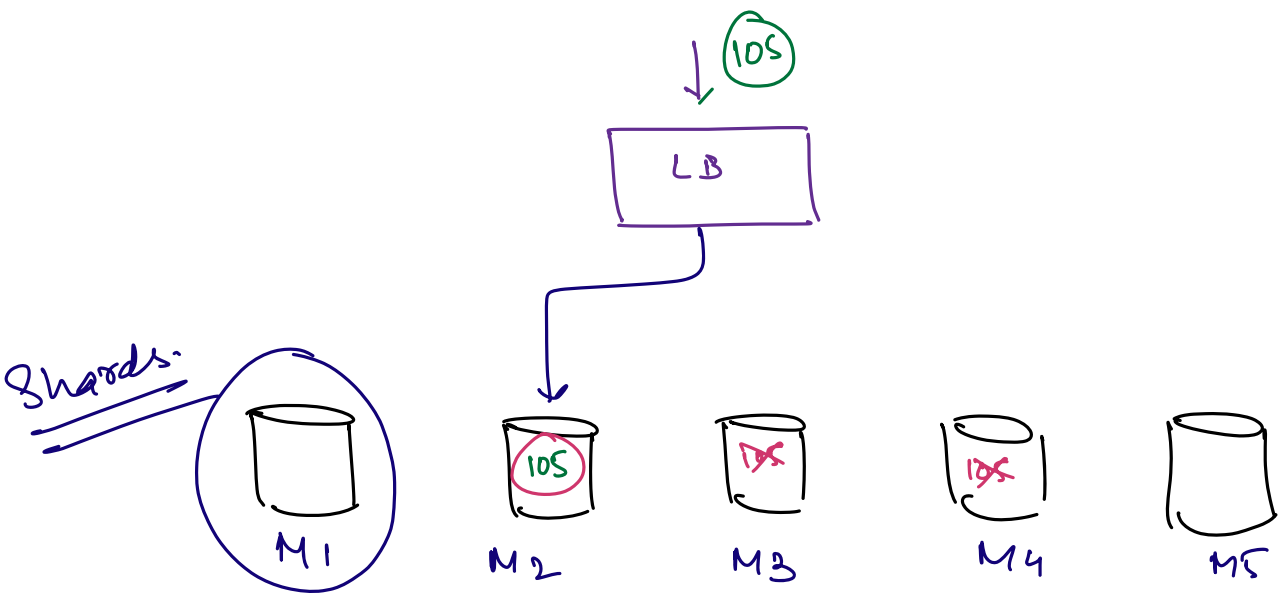
SHARDING.



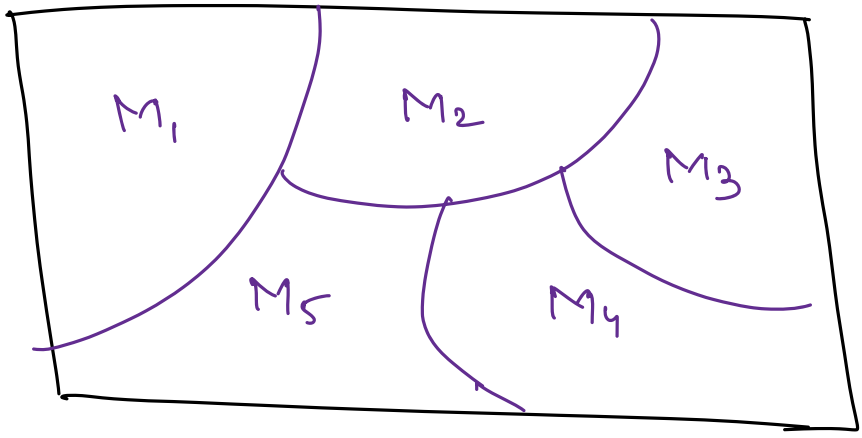
→ 16 TB.

PB
↓
1000TB.

=> All of my data may not fit into single m/c, hence we need to distribute our data across multiple machines, this is called as Sharding.

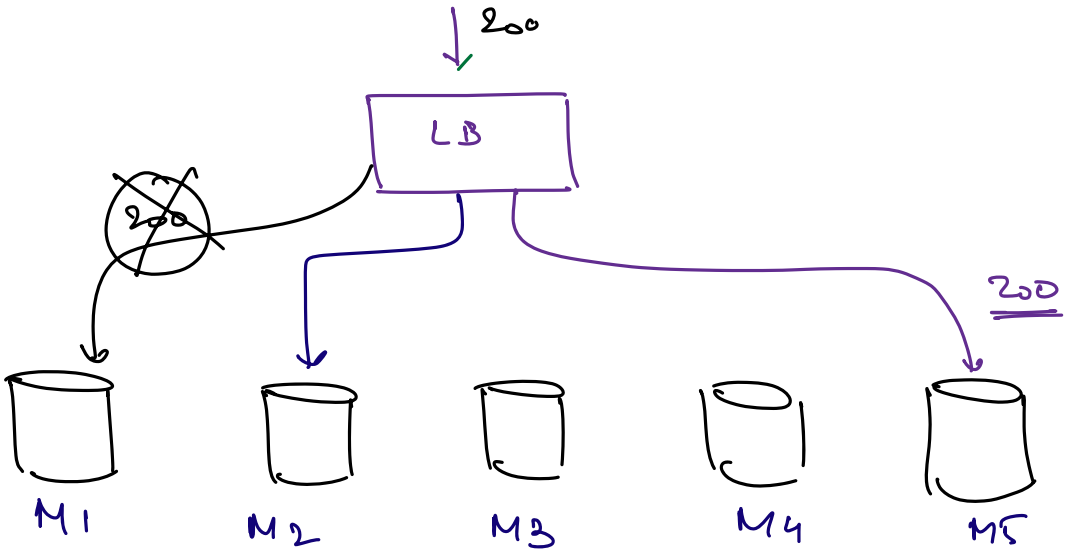


Sharding will lead to creation of multiple DB machines in DB cluster.

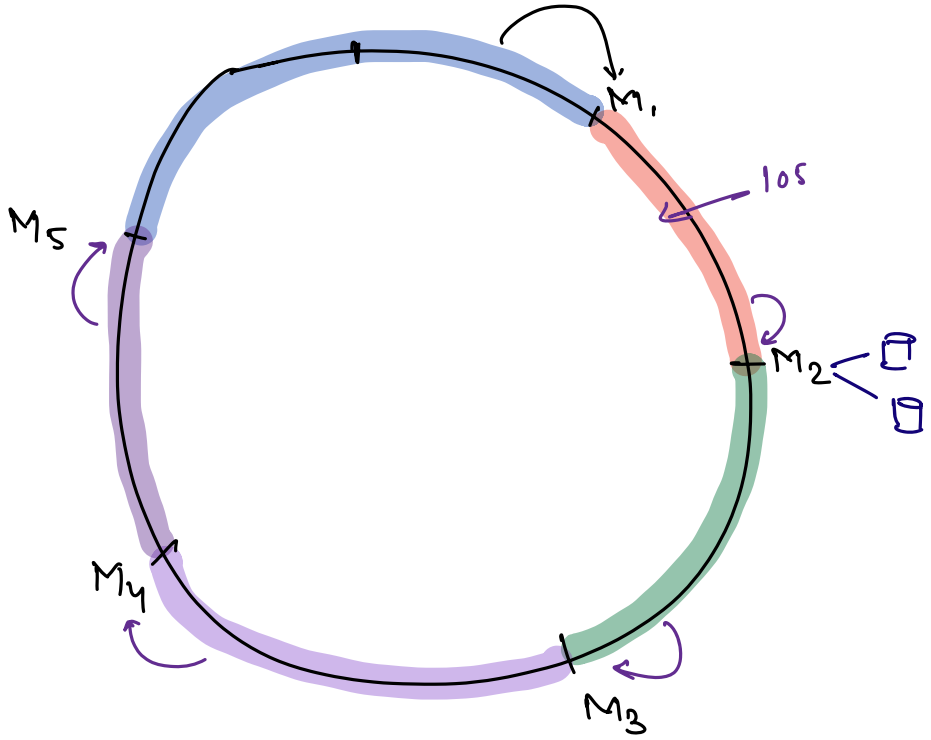


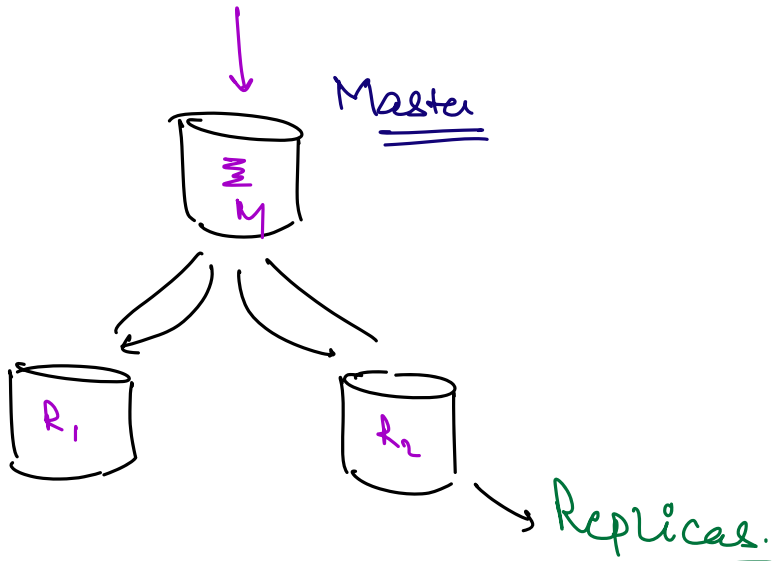
Machines.

SHARDING ≡ Shards which are mutually exclusive + collectively exhaustive.



⇒ Consistent Hashing





Master Slave Architecture

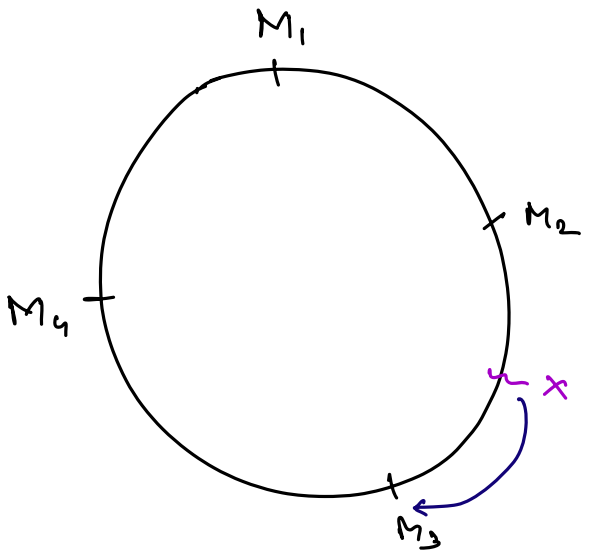
⇒ id using which we are going to shard the data ⇒ SHARDING KEY

⇒

Sharding key - [user_id]

↓

hash(user_id) = ⊗



↓

Trade-Off

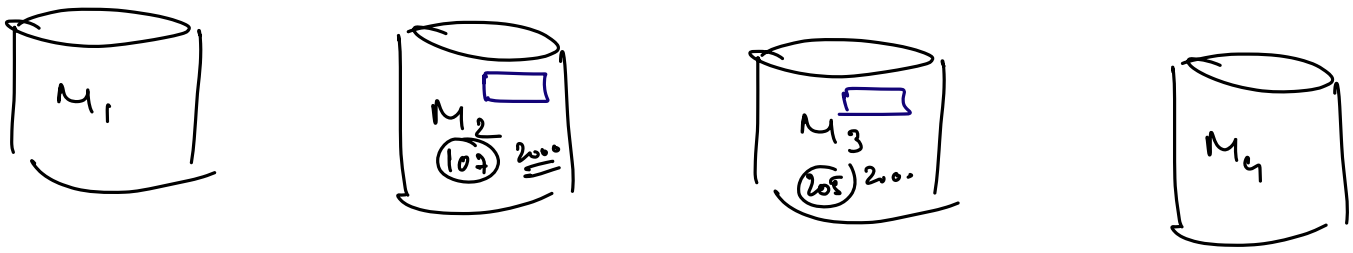
Inter Shard Query

vs

Intra Shard Query

Amazon Products DB.

↳ 100TB. ⇒ Sharding.

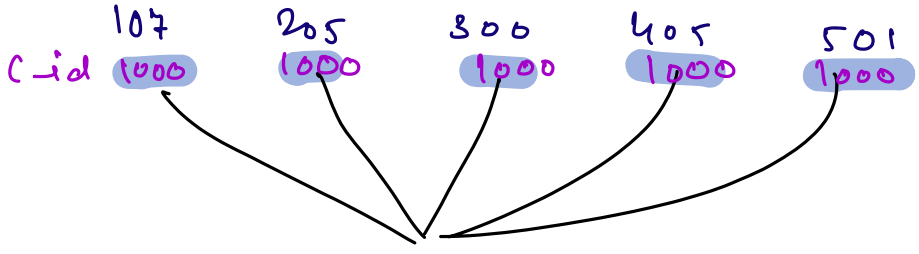


Query: Get all the products of a particular Category.

~~A)~~ Product_id ⇒ Inter Shard Query.

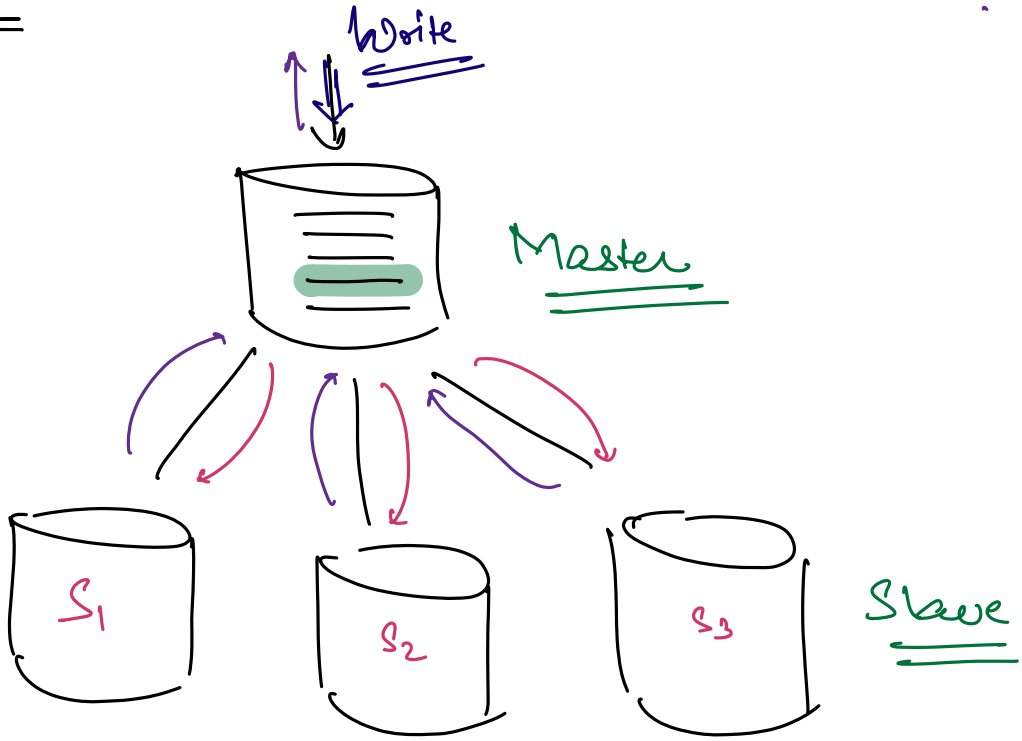
B) Category_id ✓

C) City_id



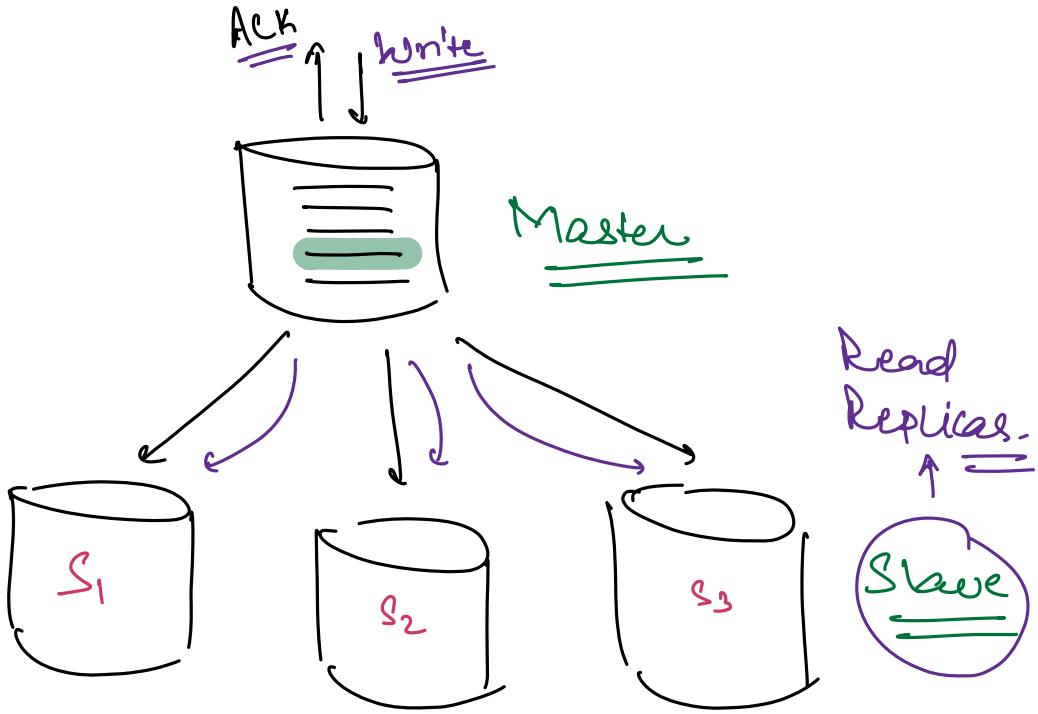
Men's Shirt.

Replication.



1) Sync
   - Slow writes.
   - Immediate Consistency.

2) Async
   - Faster Writes
   - Eventual Consistency.



⇒ Quorum Approach.

↳ 50% of the m/c.

⇒ Distribute load.

# of reads >>> # of writes.

⇒ Leader Election Algo

Financial App" ⇒ Immediate Consistency.

fb | Instagram | YT ⇒ Eventual Consistency.

CAP Theorem.

Consistency : Everytime when we make a read request, we should get the latest information.

Availability : If we send a query, the system must be available to send the response back.

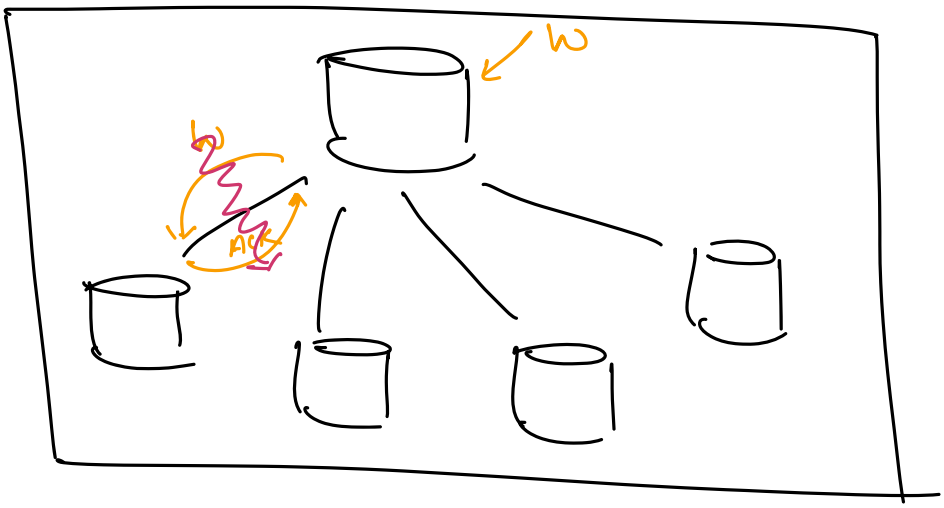
Partition Tolerant

↓

Network Partition.

⇒ A system is called as Partition tolerant if any event of N/w partition can be handled by the system.

⇒ N/w.



Immediate Consistency.

CAP Theorem : Out of (C) | A | P, only 2 can be picked at a time.

C A

C P

A P

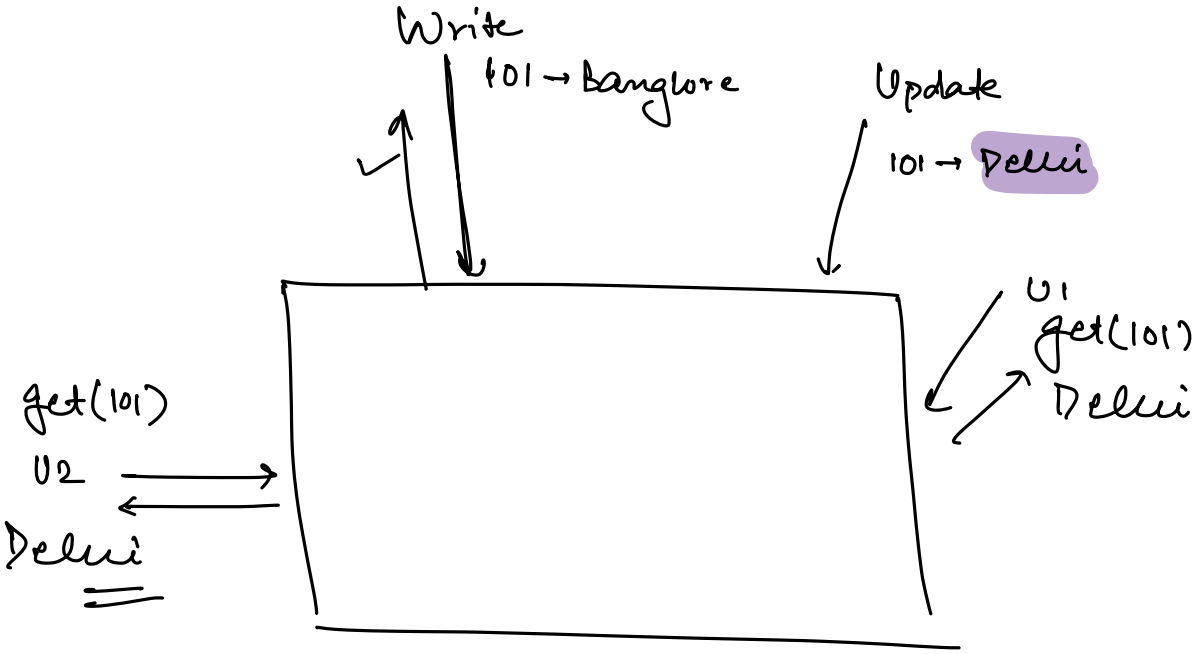
Distributed System : P ✓

C (or) A

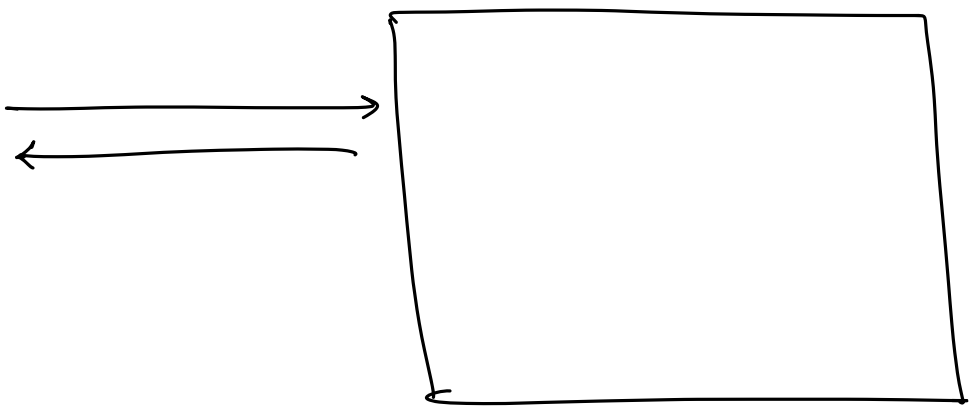
Single m/c ⇒ P ✗

Consistent | Available.

JPM | DE Shaw



Consistency: Get the data of latest write operation.



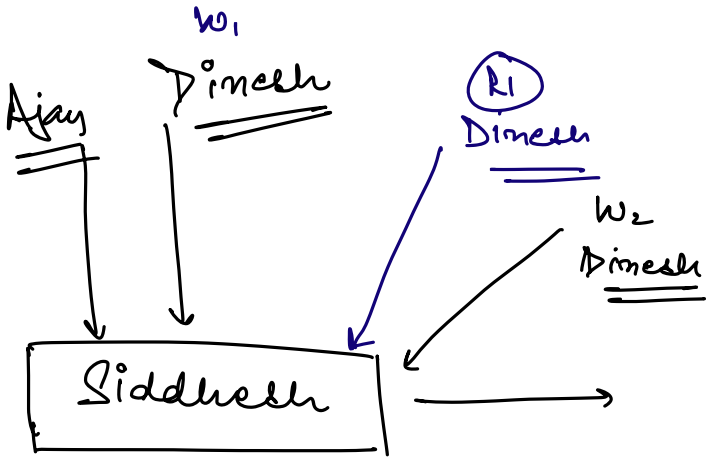
if we have a single m/c ⇒ CA

if # of m/c > 1 ⇒ P ✓

CP AP.

Siddhesh : Reminder Company.

1)



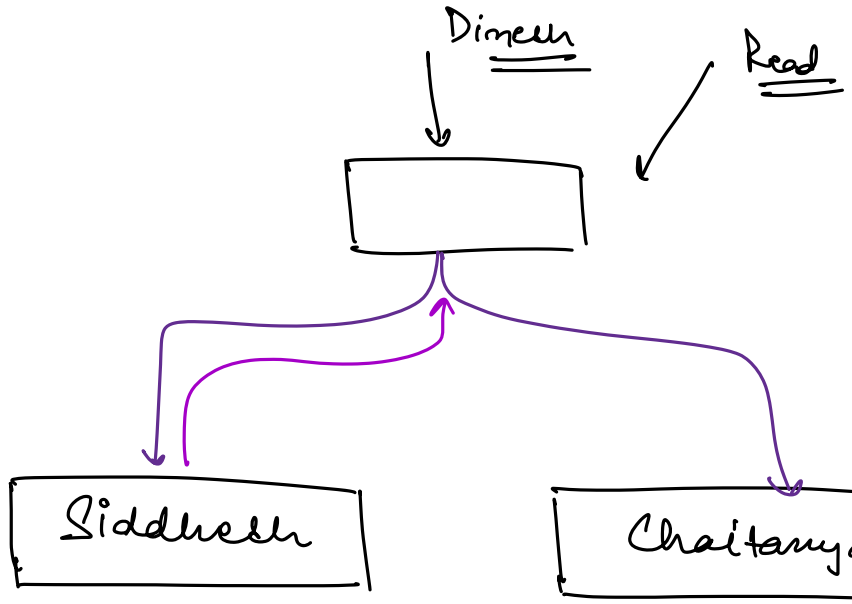
Diary

Dinesh: Date on 20th Dec with A

Ajay: Flight to Delhi on 25th Dec

Dinesh: Date on 21st Dec with B.

11)



① Dinesh: 20th Dec, 8Pm mutu (Y)

Ajay: flight to delhi on 18th Dec -

=

Dinesh: 21st Dec, 2

Dinesh: 20th Dec, 8Pm mutu (Y)

Dinesh: 21st Dec, 2

Partition ✓

Availability ✓

Partition ✓

Consistency ✓ (Strong)

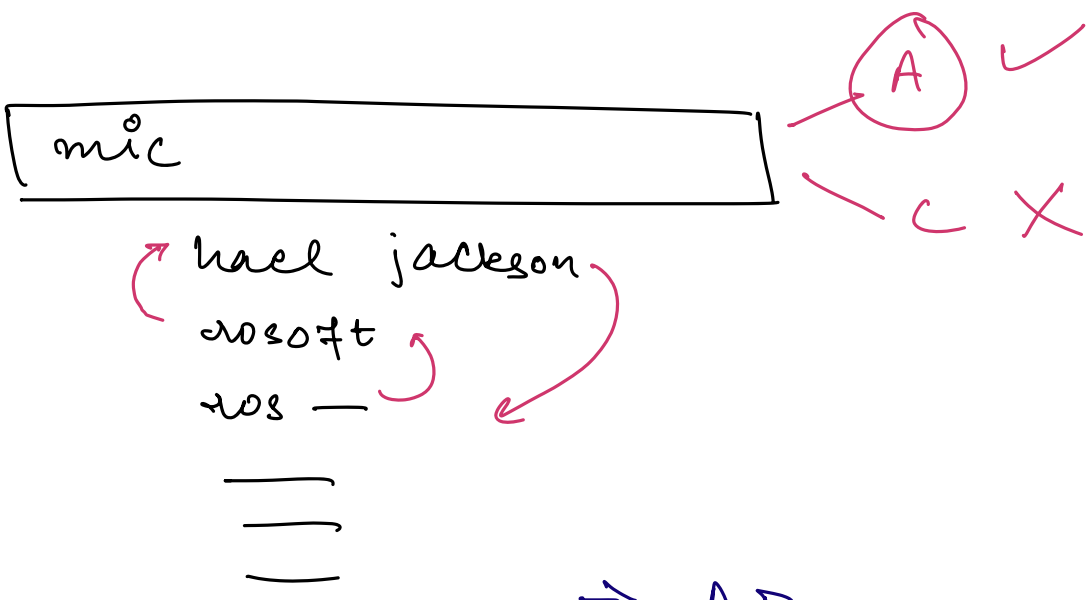
Partition ✓

Consistency

Availability.

GPay → CP

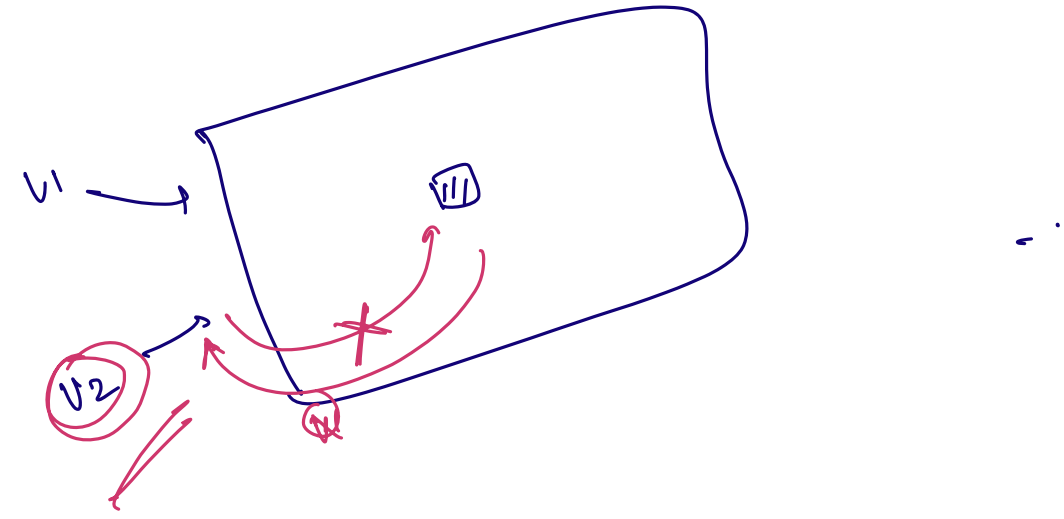
Google Typeahead Engine



⇒ AP.

Irctc

↳ 10 tatkal — 1000


V1
V2