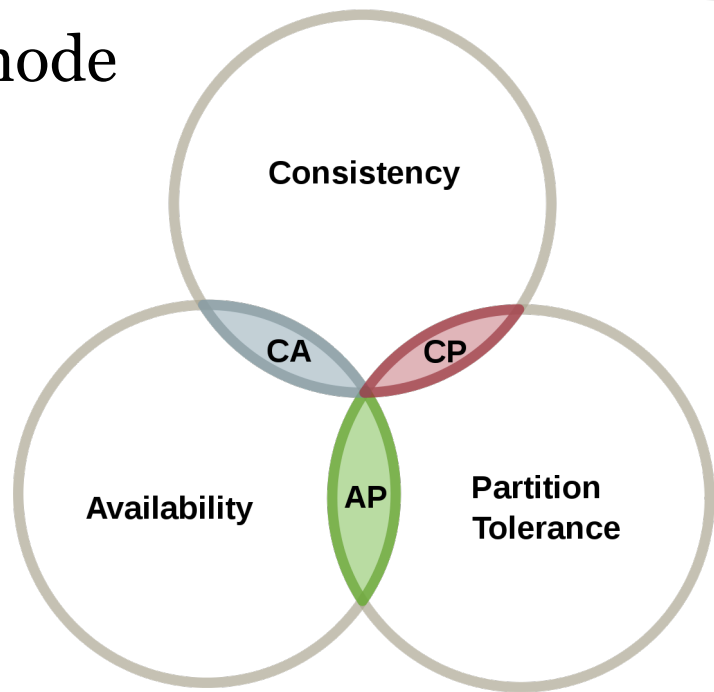


CAP Theorem - Eric Brewer

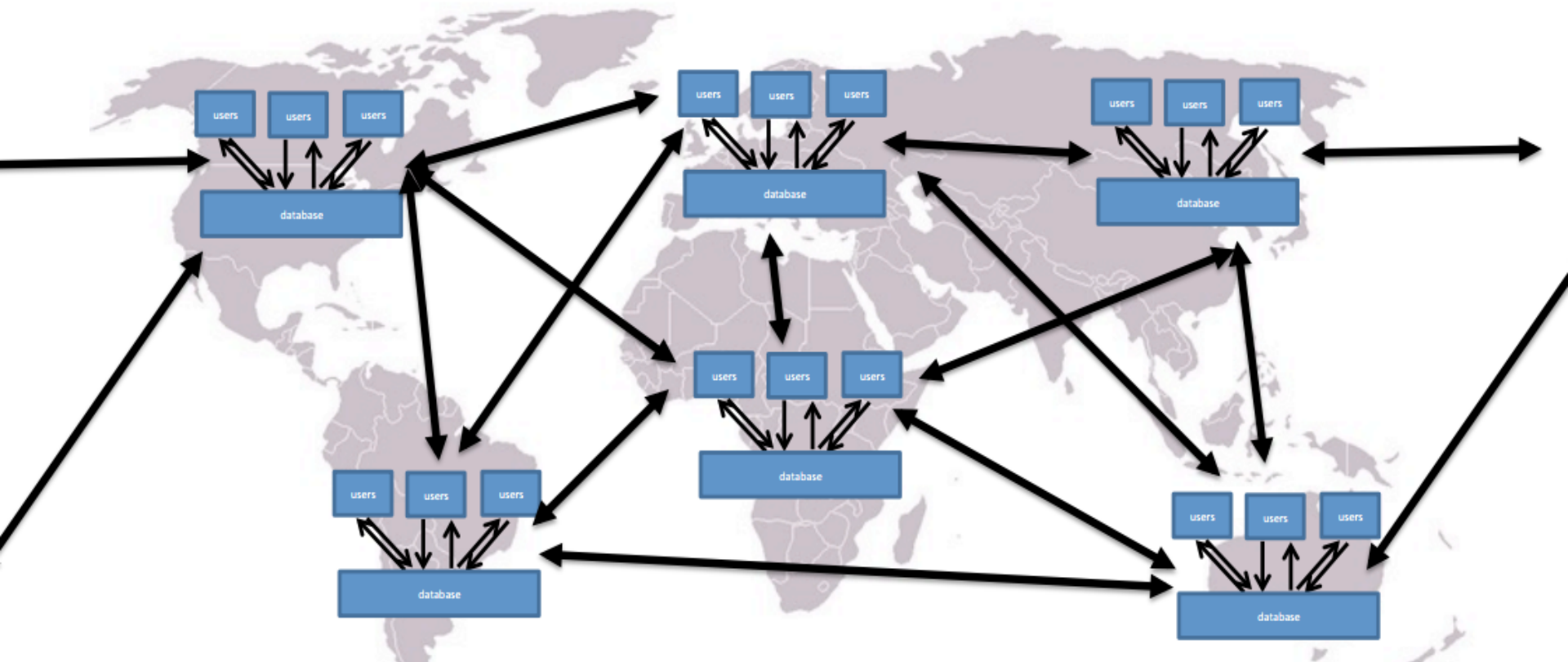
JD

Overview

- **Consistency**
 - Data appears the same on every node
- **Availability**
 - Every requests succeeds or fails
- **Partition Tolerance**
 - Systems deal with network faults
- **Choose 2 of the 3 above**



But Why Should I Care?



Consistency

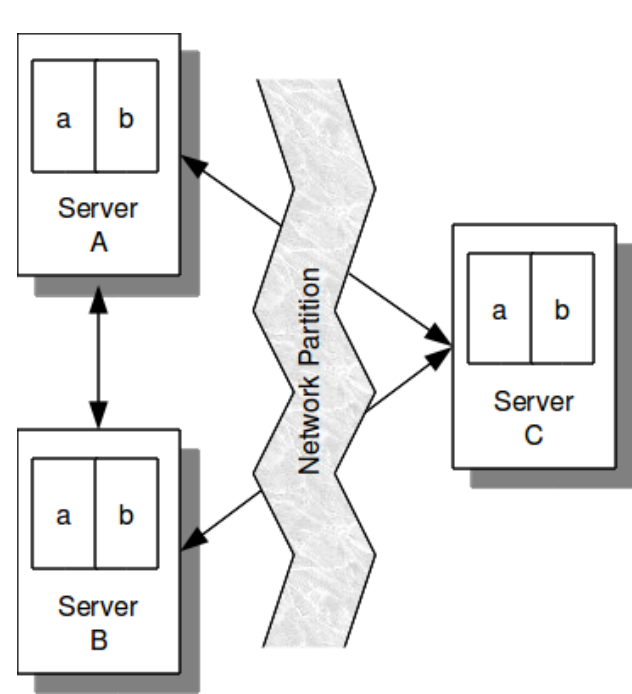


- C = single-copy consistency
- update a majority nodes before advancing
- Several levels of consistency
 - Sequential consistency
 - Atomic consistency
 - Causal consistency
 - Eventual consistency

Availability

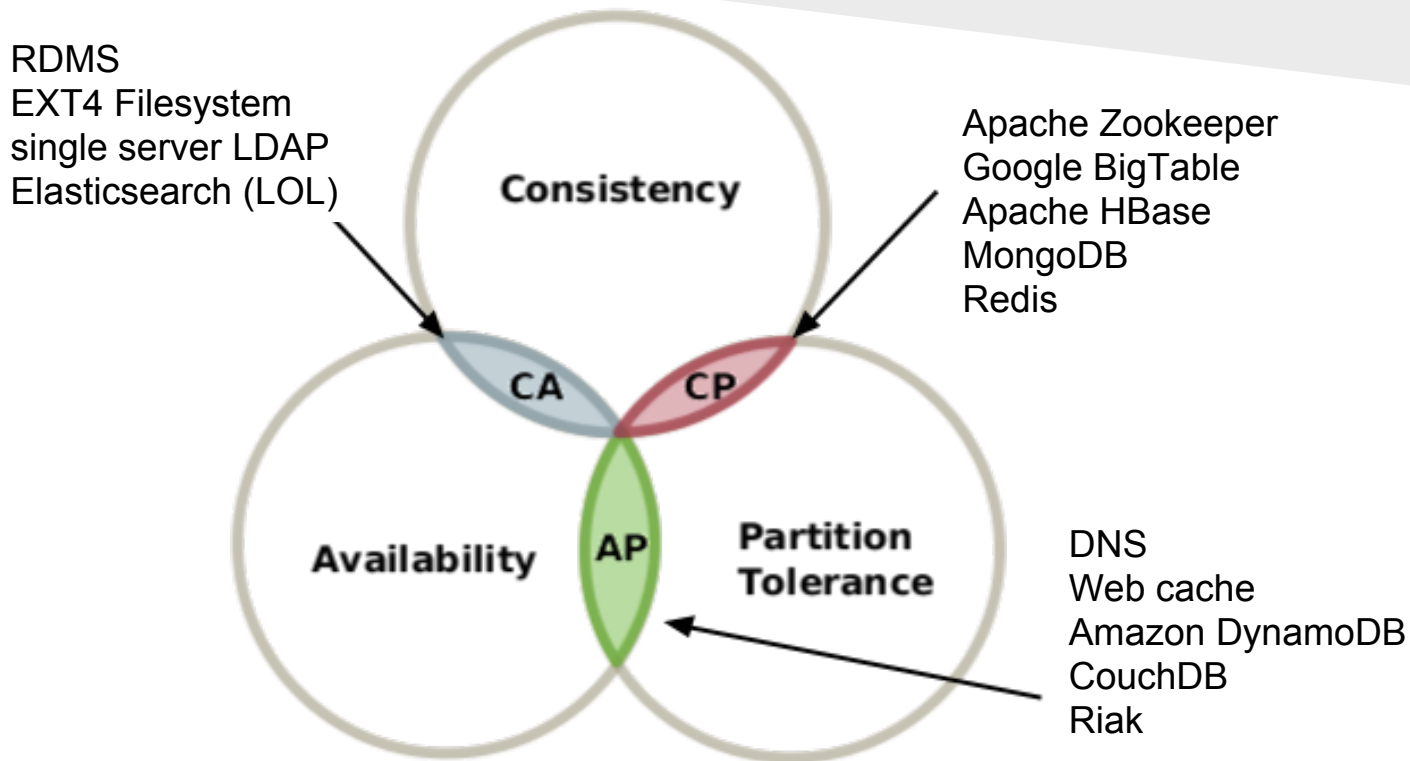
- Will all requests always succeed or fail?
- Can return stale data
- Replicate data across different machines

Partition Tolerance



- Continue working during network outages
- Concurrent requests to nodes A & C
- Time bound on communication

Real World Examples



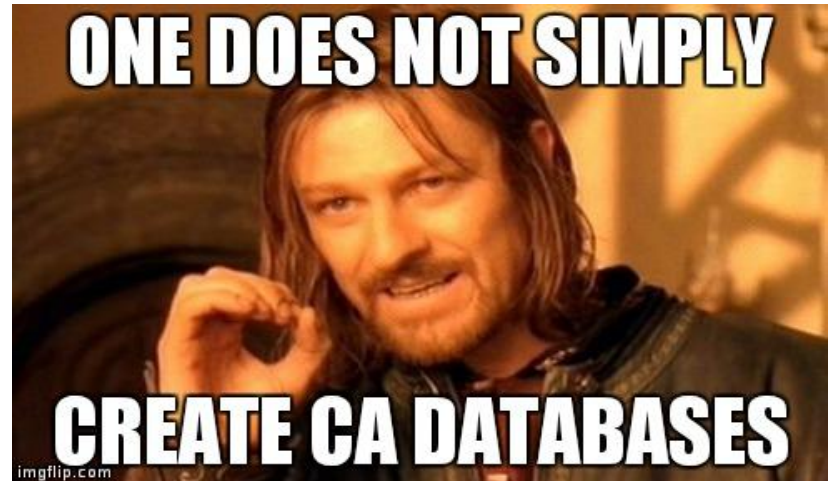
Consistency - Availability

Pros:

- Easy to use
- SQL
- Atomic read / writes

Cons:

- Network issues crash it
- Needs full replication



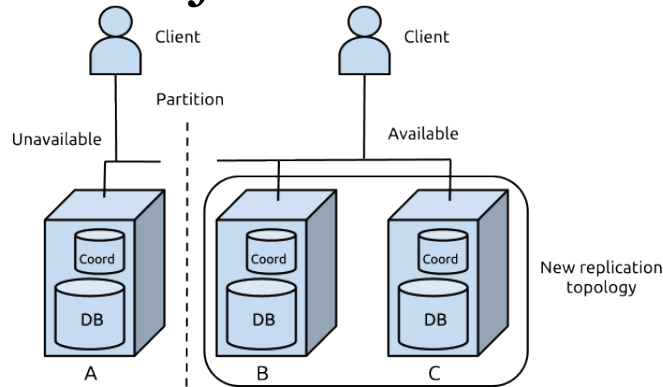
Consistency - Partition

Pros:

- Atomic read / writes
- Synchronous communication
- Master-slave setup

Cons:

- Sometimes requests fails
- Waits for a response from partitioned node
- Can reject writes



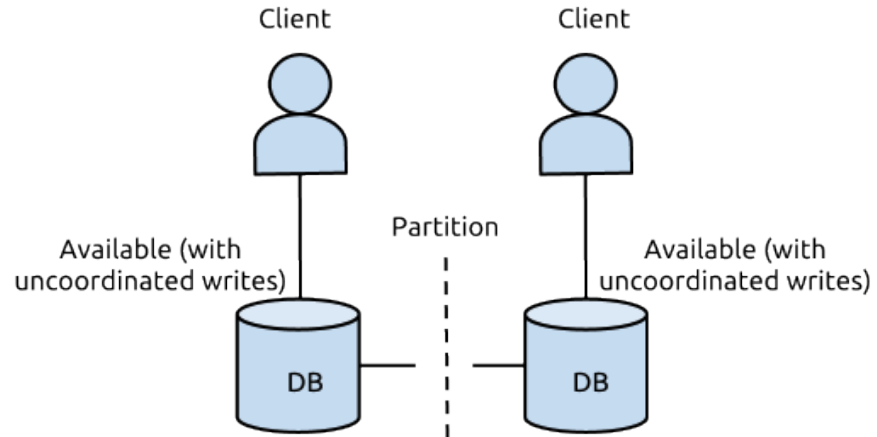
Availability - Partition

Pros:

- Accepts writes to execute later
- flexibility in data sync
- works in case of external errors
- Asynchronous communication

Cons:

- Can return stale data
- Eventual consistency
- “Split Brain”



Why not all 3?

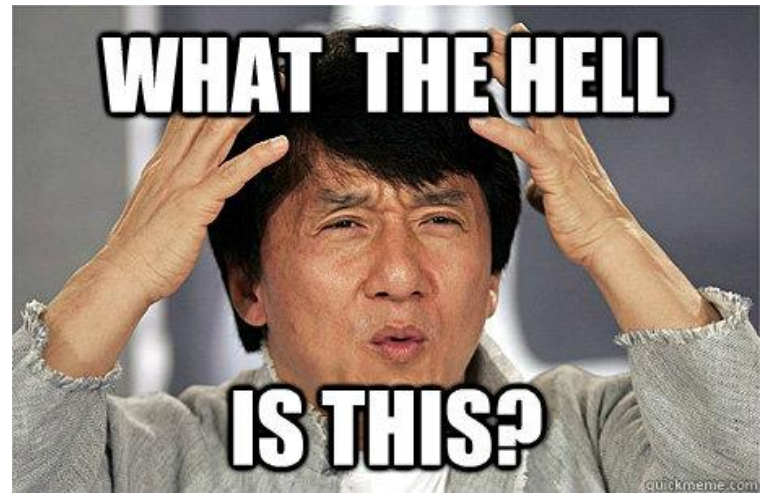
- Assume system works for partitions
 - System that allows requests while parted is available but not consistent with other nodes
 - Systems that blocks requests while parted is consistent but not available



2 out of 3 is actually 1 out of 2

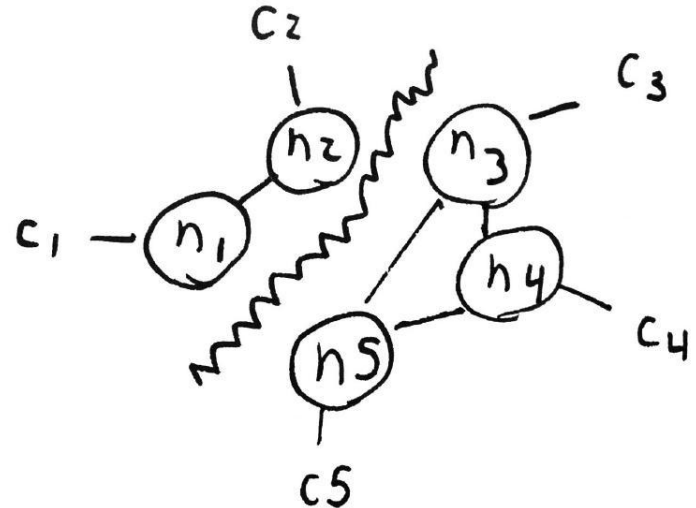
- Availability vs Consistency
- Latency can be viewed as temporary partition

CP vs. AP

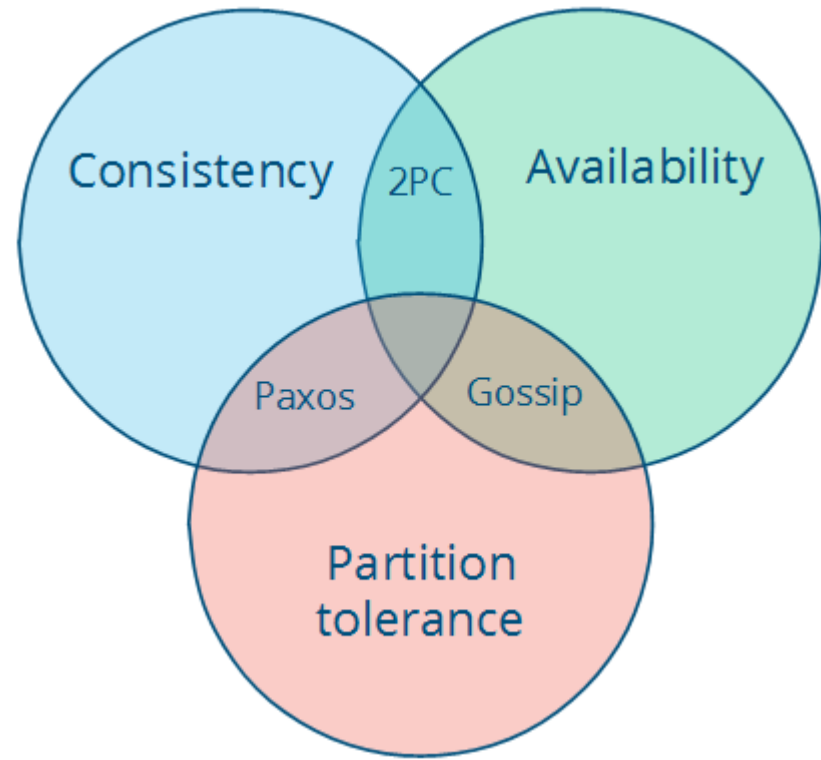


Don't Drop Partition Tolerance!!

- Allowing reads before updating all nodes: **only Availability**
- Locking all nodes before allowing reads: **only Consistency**



Communication solutions



2 Phase Commit

Master / slave replication

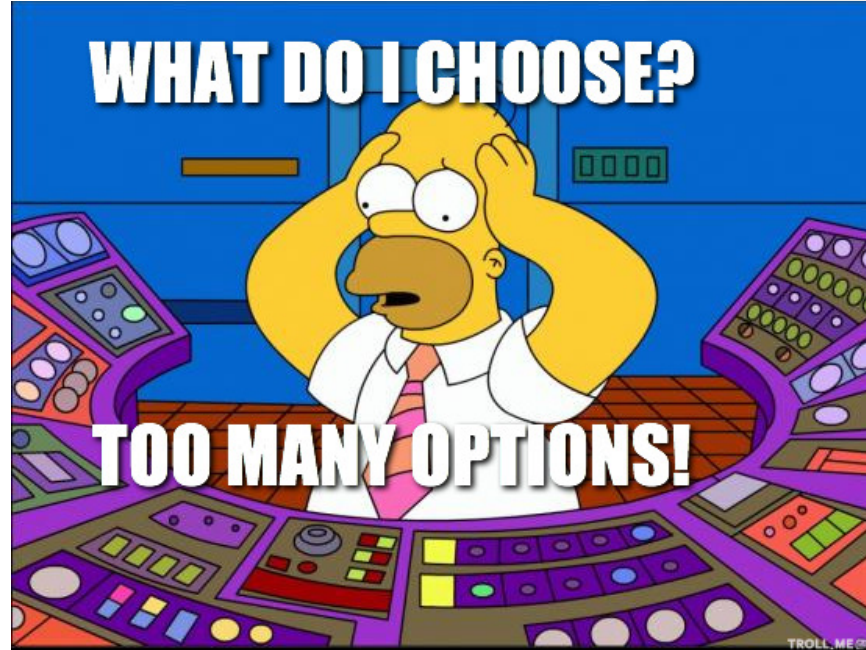
Paxos

Consensus protocol

Gossip

Read-repair

CP vs AP



Well, it depends

Questions?