CAP Twelve Years Later: How the “Rules” Have Changed

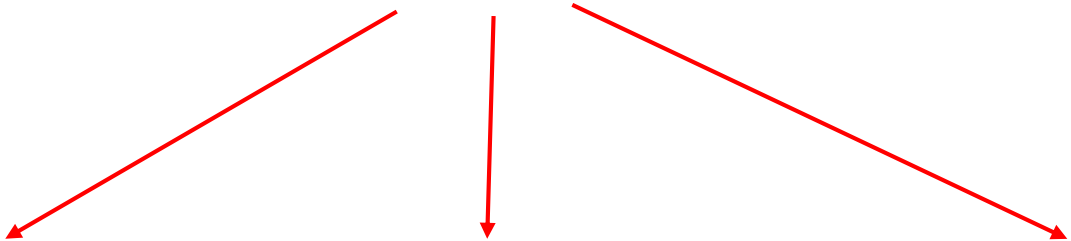
Presenters:
- Ayushi Bansal
- Krithiga Murugavel

Brewer, Eric. (2012). CAP Twelve years later: How the "Rules" have Changed. Computer. 45. 23-29. 10.1109/MC.2012.37.

Outline

- CAP Theorem
- ACID vs CAP
- Why 2 of 3 misleading?
- CAP confusion
- Managing Partition
- Partition Recovery

CAP



Consistency

Every read receives the most recent write or an error.

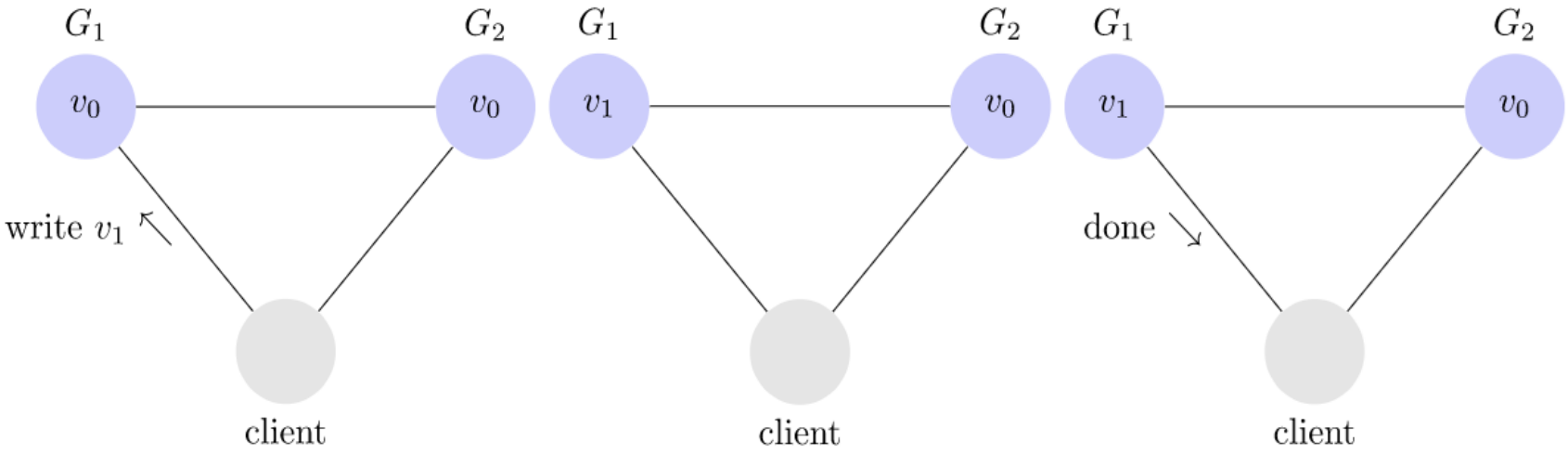
Availability

A guarantee that every request receives a response.

Partition Tolerance

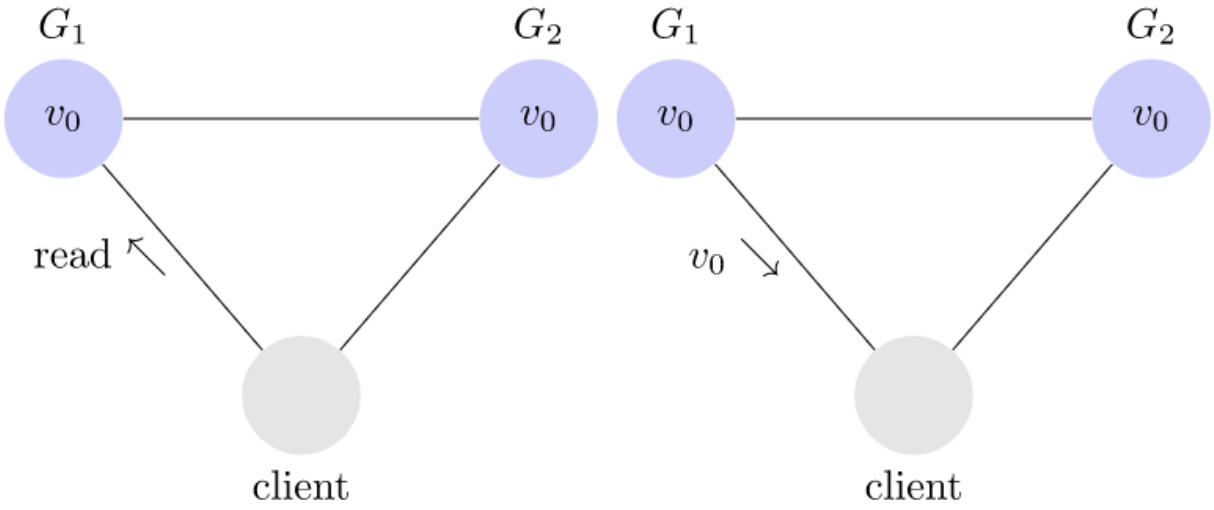
The system continues to operate despite an arbitrary number of messages being dropped (or delayed) by the network between nodes

Writing to a Server

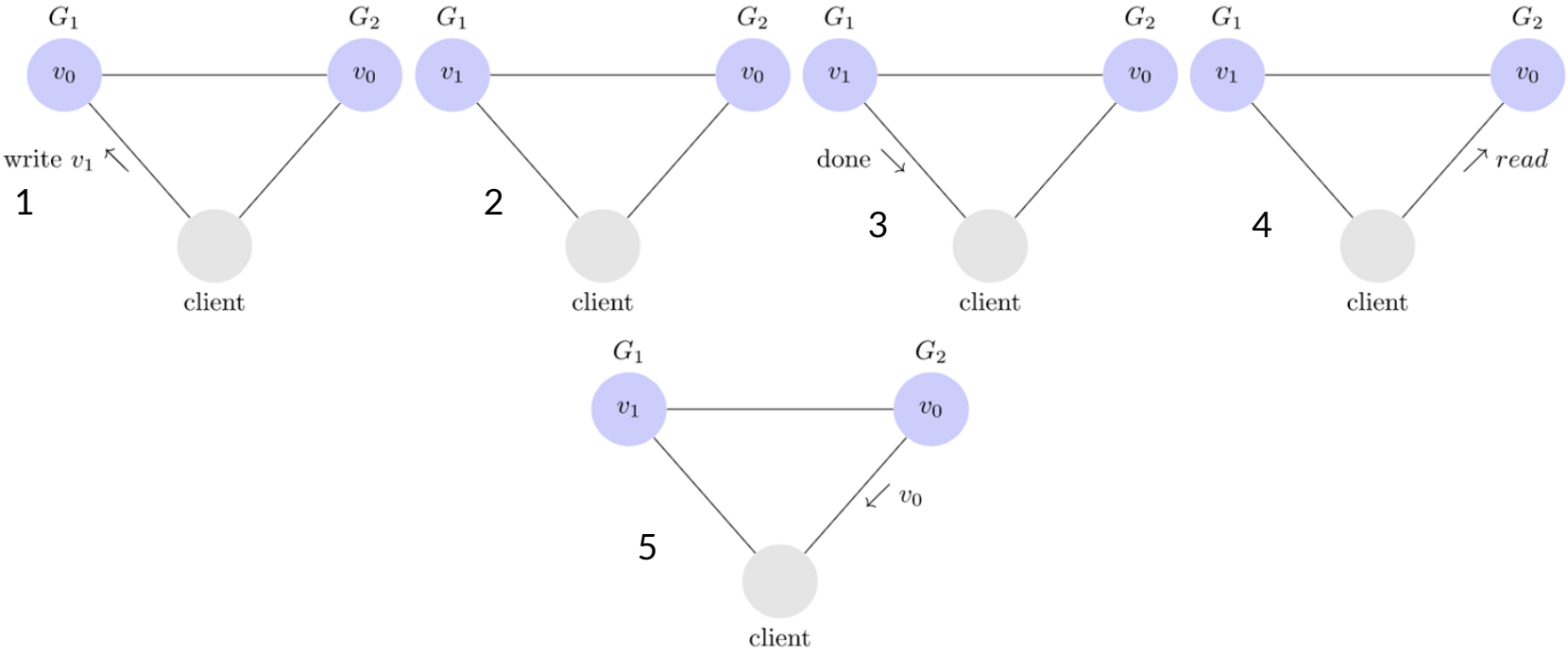


https://mwhittaker.github.io/blog/an_illustrated_proof_of_the_cap_theorem/

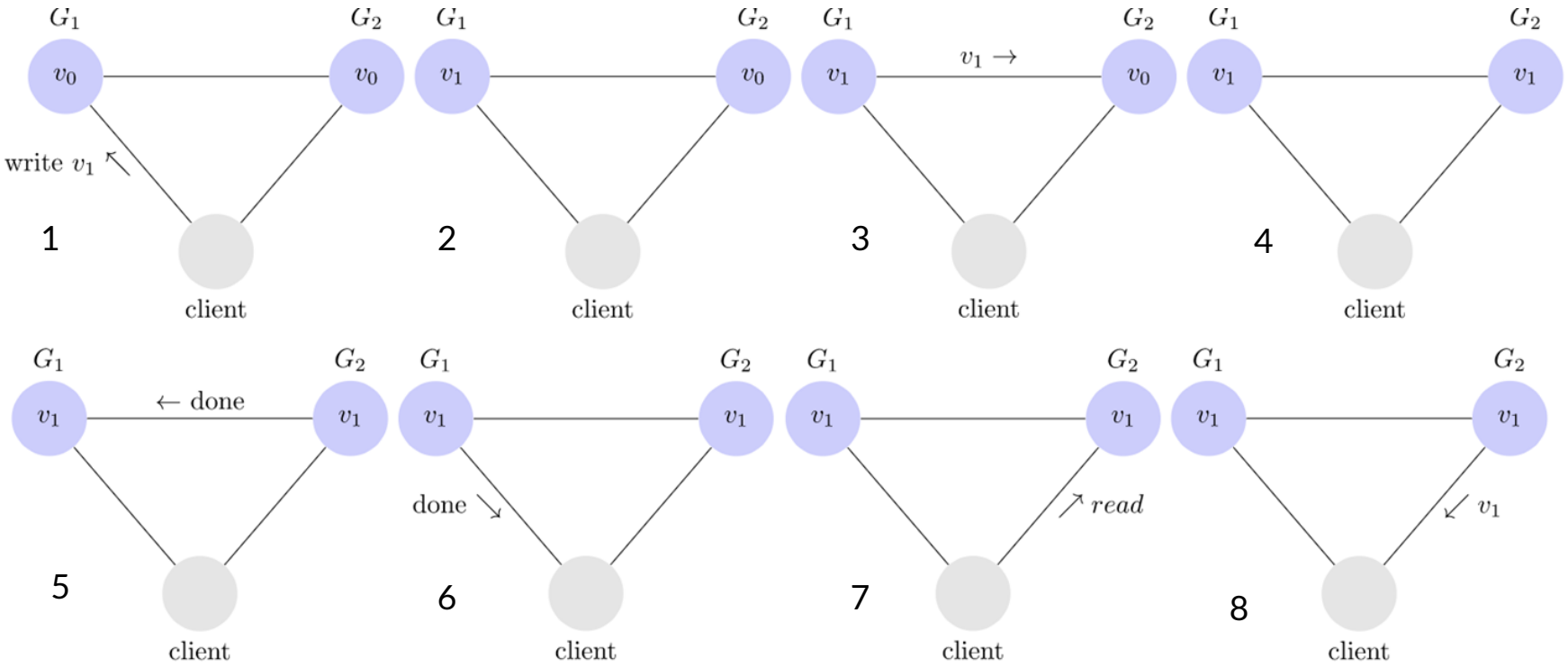
Reading from a Server



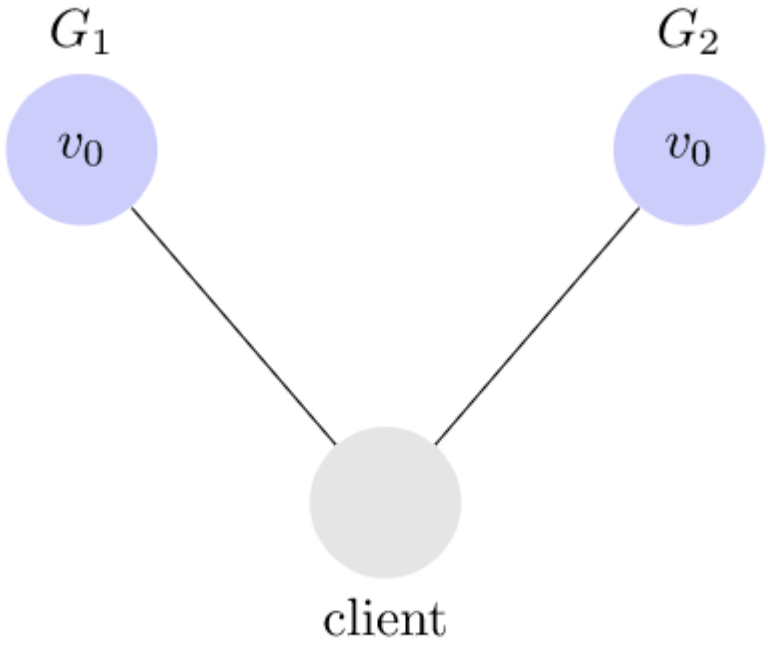
An Inconsistent System



A Consistent System

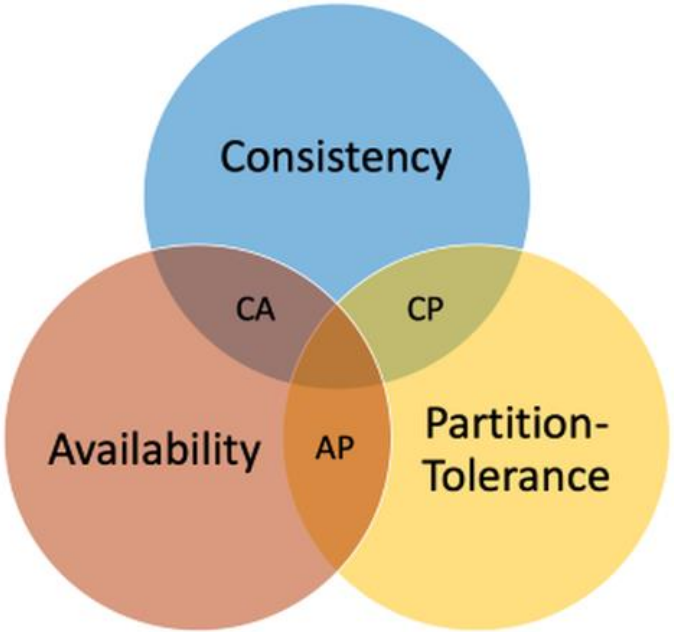


Partition



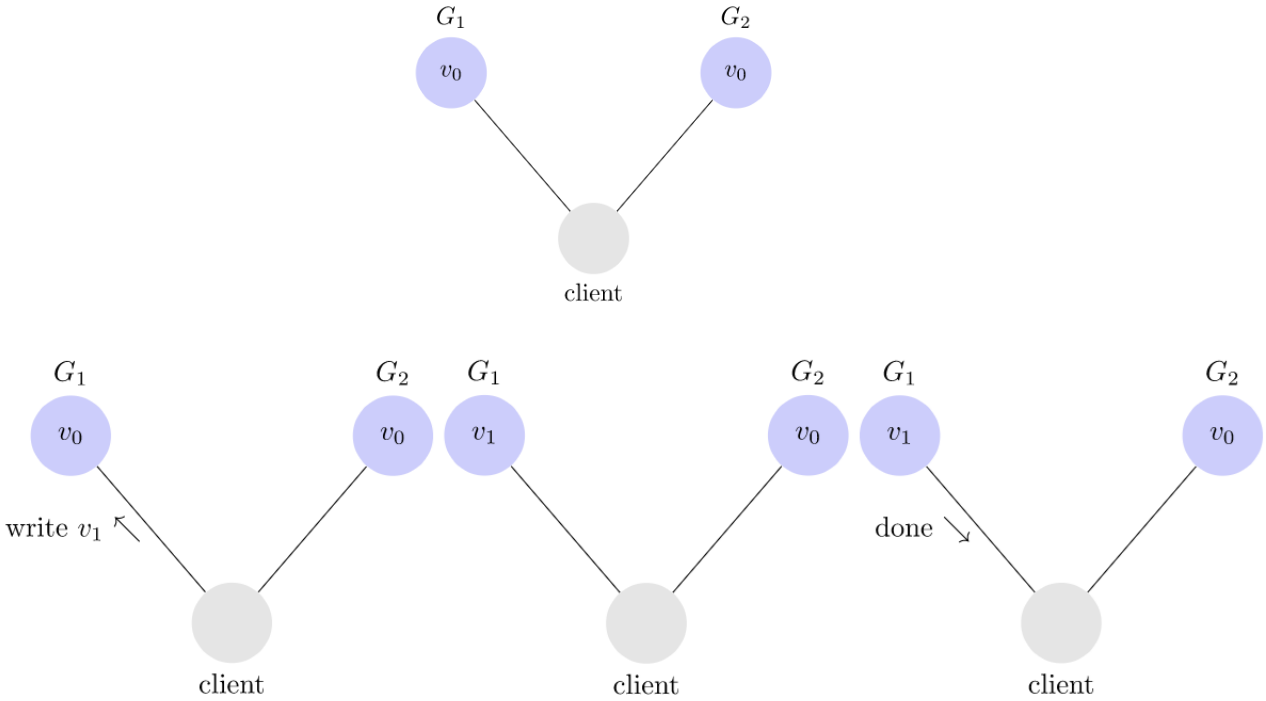
CAP THEOREM

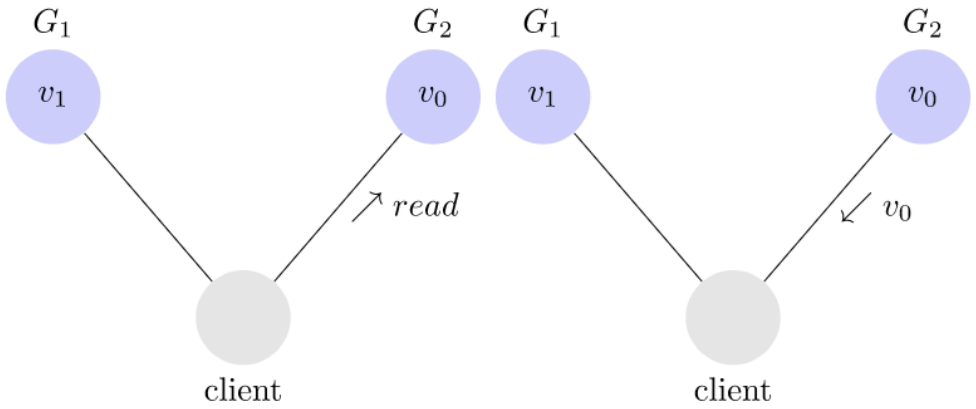
Any networked shared-data system can have at most two of three desirable properties.



https://www.easttexassoftware.com/post/cap-theorem-and-distributed-systems

Choosing A over C



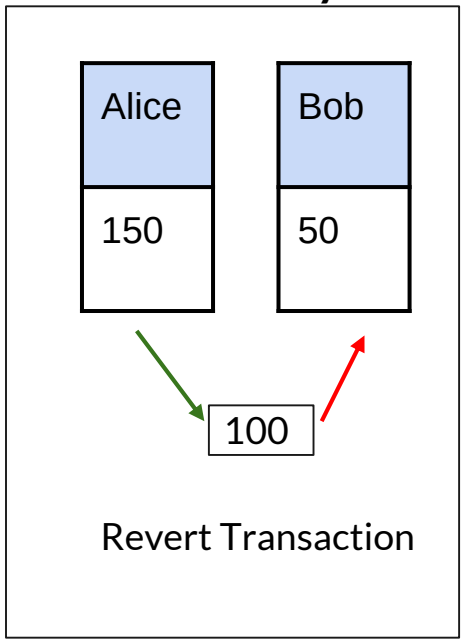


Hence proved that we cannot have both Consistency, Availability and Partition Tolerance .

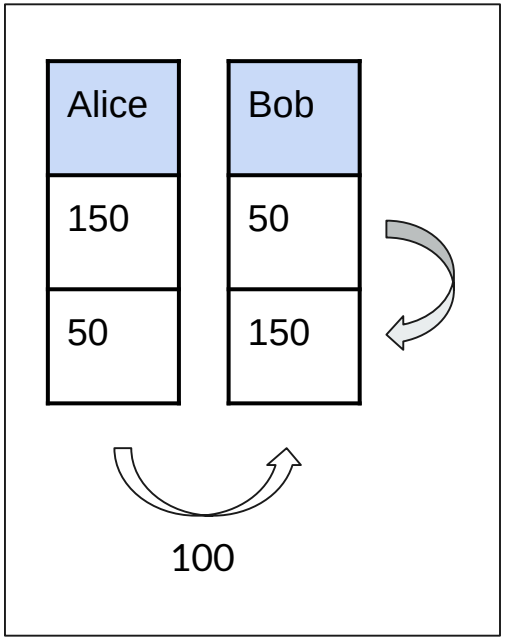
ACID

Alice	Bob
150	50

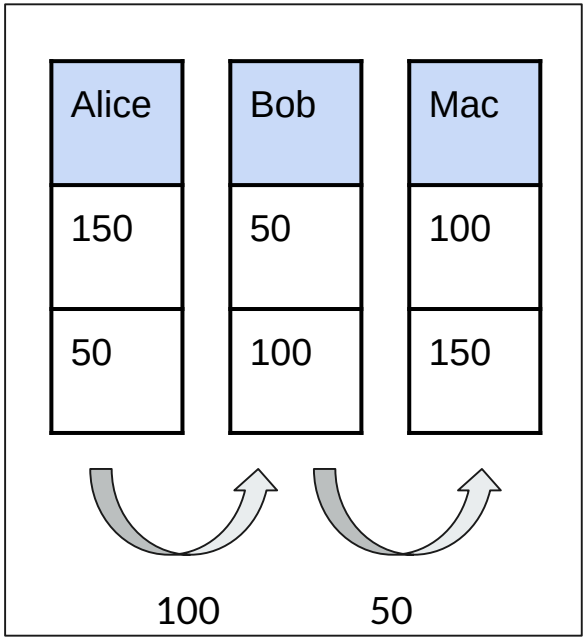
Atomicity



Consistent



Isolation



ACID VS CAP

- **Consistency** in **C**AP refers to Single copy consistency and **Consistency** in A**C**ID means the transaction preserves all the database rules
- A stands for **Atomicity** in **ACID** and **Availability** in **CAP.**

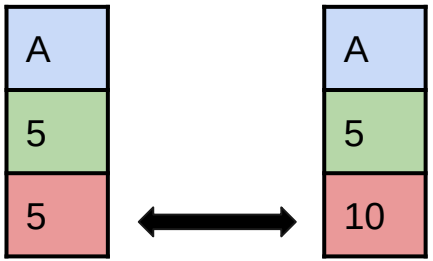
ACID

Alice	Bob
10	40
5	40

→ ACID Inconsistency

Integrity constraint: A+B = 50

CAP



CAP Inconsistency

Why is 2 of 3 misleading?

- Partition are Rare
- Granularity in making choice between C and A based on:
  - ❑ Applications
  - ❑ Users Involved
  - ❑ Data
- All three properties are more continuous than binary

CAP Confusion

- Hidden cost of forfeiting consistency.
- Can a designer choose not to have partitions?
- When User cannot reach the service at all there is no choice between C and A except when part of the service runs on the client.

CAP-LATENCY CONNECTION

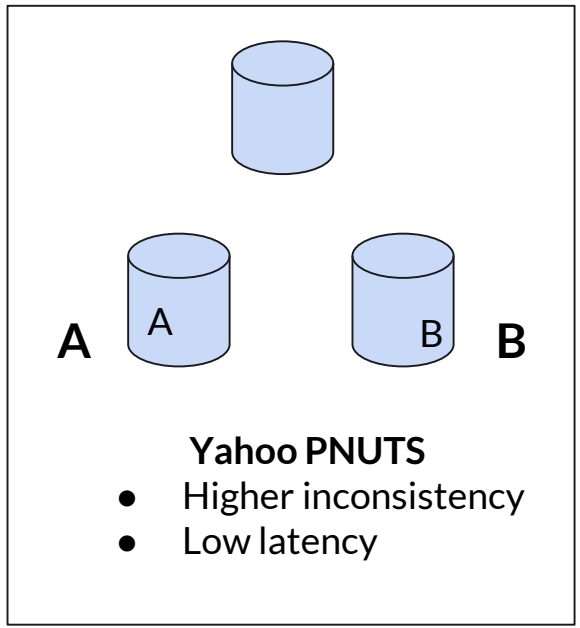
Latency: The delay from input into the system to desired outcome

High Availability

↓

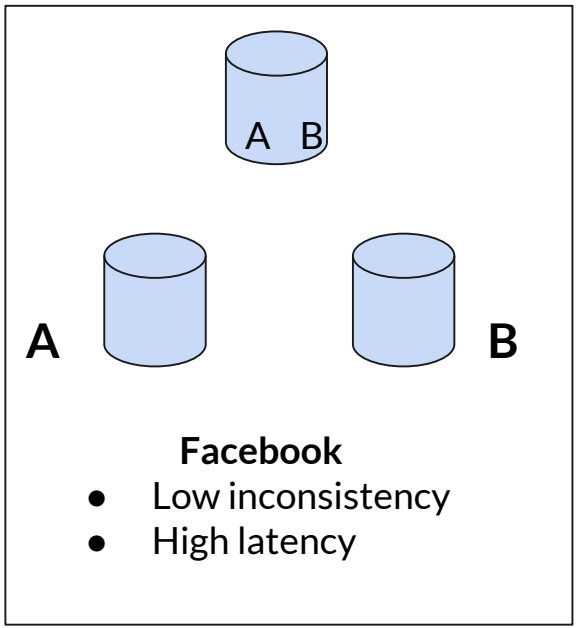
Data Replication

Consistency vs Latency



Yahoo PNUTS

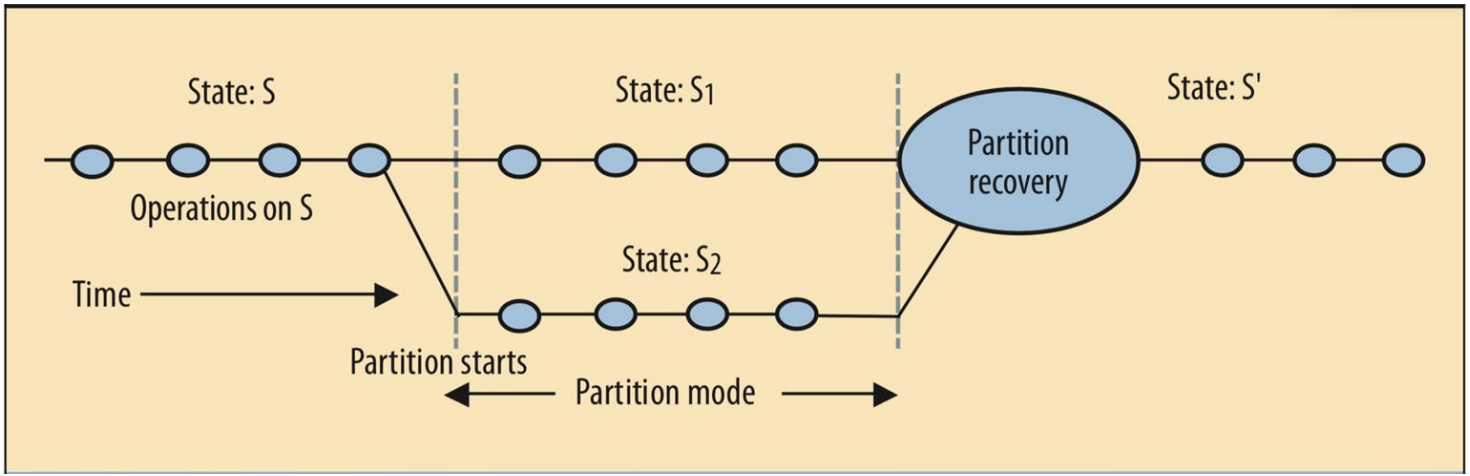
- Higher inconsistency
- Low latency



Facebook

- Low inconsistency
- High latency

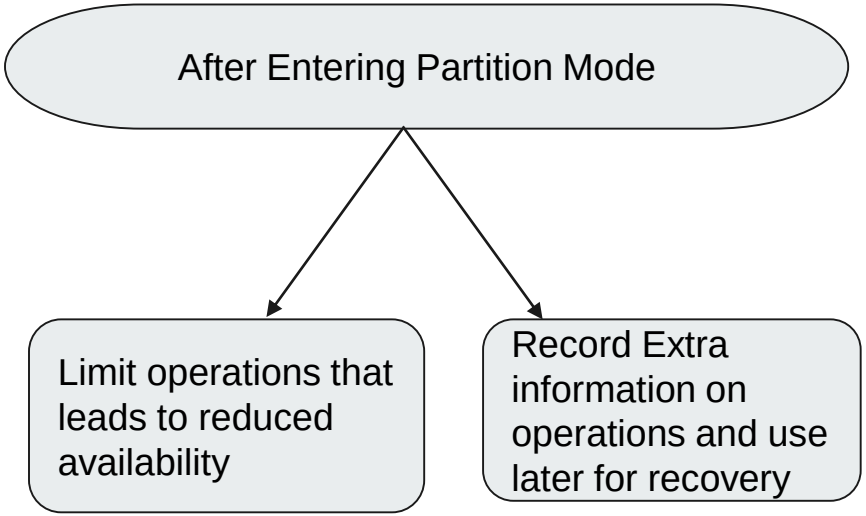
PARTITION MODE



- Detect Partition
- Enter partition Mode, and decide between Consistency and Availability
- Recover process to restore consistency and compensate for mistake

THE PARTITION DECISION

- Decision made based on the invariants.
- For invariants that needs to be maintained, the system prohibits or modifies the operation. Eg: Credit Card transaction.
- Suggests to build a Crosstable of operations and invariants.



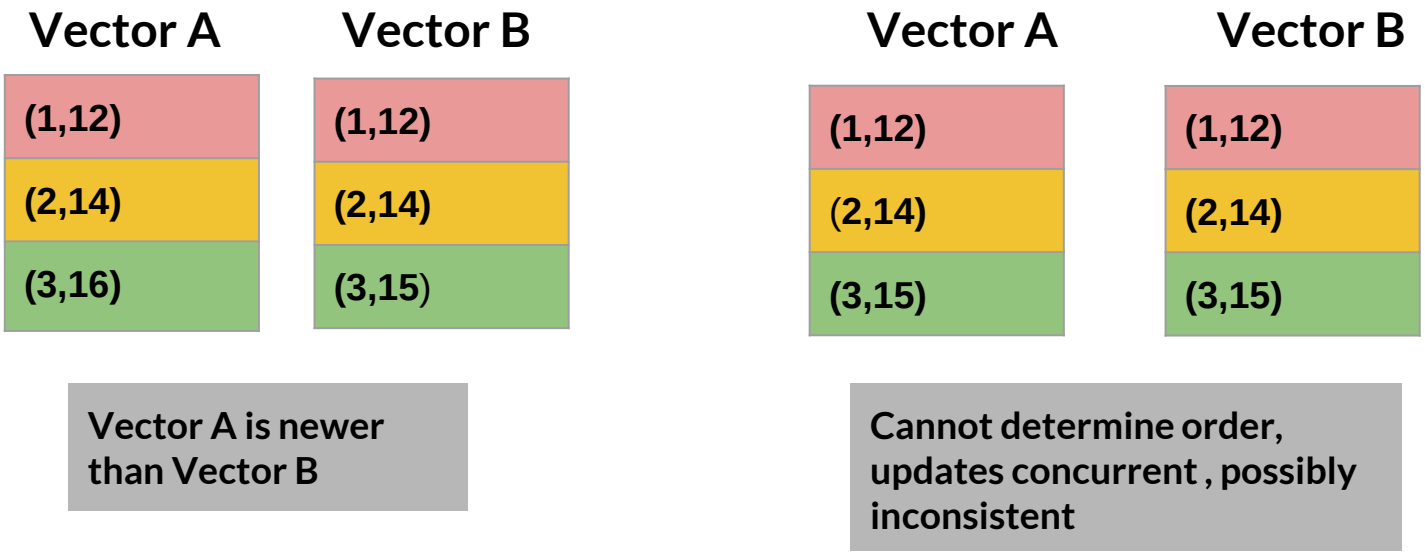
Partition Recovery

Two goals

- State on both sides must become consistent
- Handle the mistakes done during partition mode
  - Undo the mistakes
  - Compensate for mistakes

Version Vectors

- Best way to track history on both sides of partitions.
- Captures causal dependencies among operations



Partition Recovery

- **Concurrent Version System (CVS)** - uses version vector

- **Merge conflicts**
  - **Manual merging** - offline wiki system
  - **Automatic merging** - Google Docs

- **Commutative Replicated Data types (CRDTs)**
  - Merge concurrent modifications, always, in any order.
  - Rearrange operations into a preferred consistent global order

Partition Recovery - AMAZON Carts

Before Partition

During Partition

After Partition Recovery

3,2

Compensation in an ATM

- **Key invariant:** Balance should be zero or higher
- **Essential operations:** Deposit, Withdraw (violate invariant) and Check balance
- Availability preferred over Consistency
  - Withdraw limit - bounds risk
  - Compensation - Extra fee and repayment of money
- **Check Kiting**

Check Kiting

$1000

$0

Check of $5000

$-4000

$5000

Withdraw $3000

$-4000

$2000

Check of $5000

$1000

$-3000

Repeat cycle

THANK YOU !!