CS 582: Distributed Systems

Consistency and Consensus

Dr. Zafar Ayyub Qazi

Fall 2024

Today's Agenda

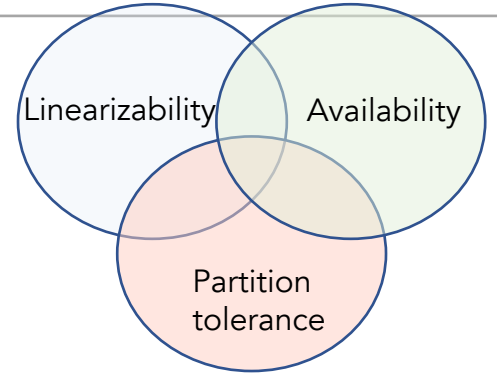
- Eventual Consistency

Specific learning outcomes

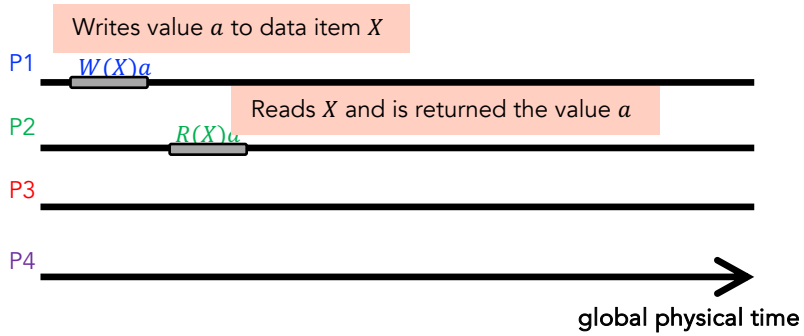
By the end of today's lecture, you should be able to:

- ❑ Compare and contrast sequential and causal consistency
- ❑ Explain and analyze the eventual consistency model

Recap: Consistency Models



- CAP Theorem
- Consistency models
  - Linearizability
  - Sequential consistency
  - Casual consistency



- Linearizability and sequential consistency:
  - Operations take place in some total order
  - But sequential consistency may not preserve real-time ordering
- Causal consistency
  - Only (potentially) causally-related operations seen in the same order

Discussion

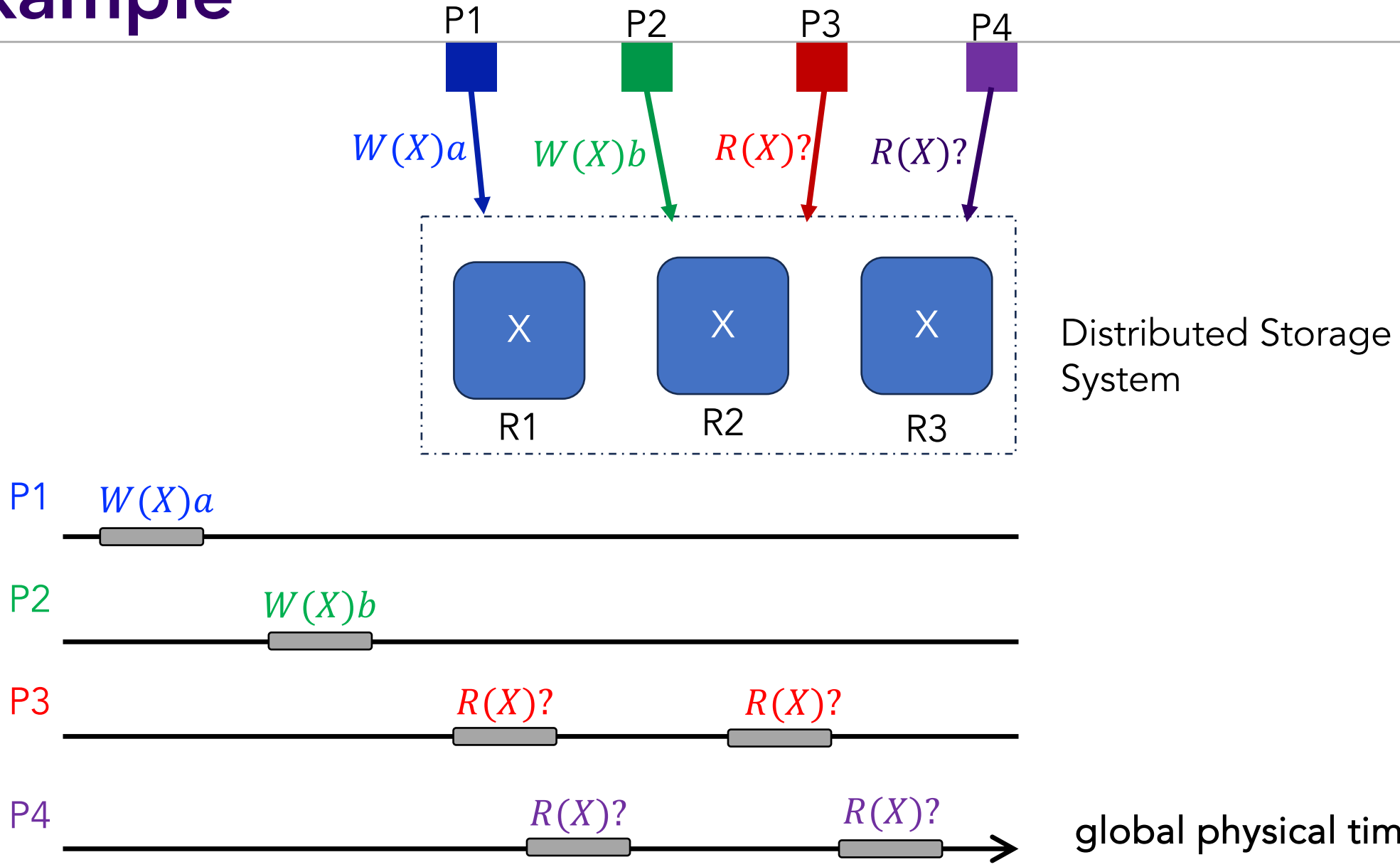
- Q. Is causal consistency strictly weaker than sequential consistency?

Recap: Causal and Sequential Consistency

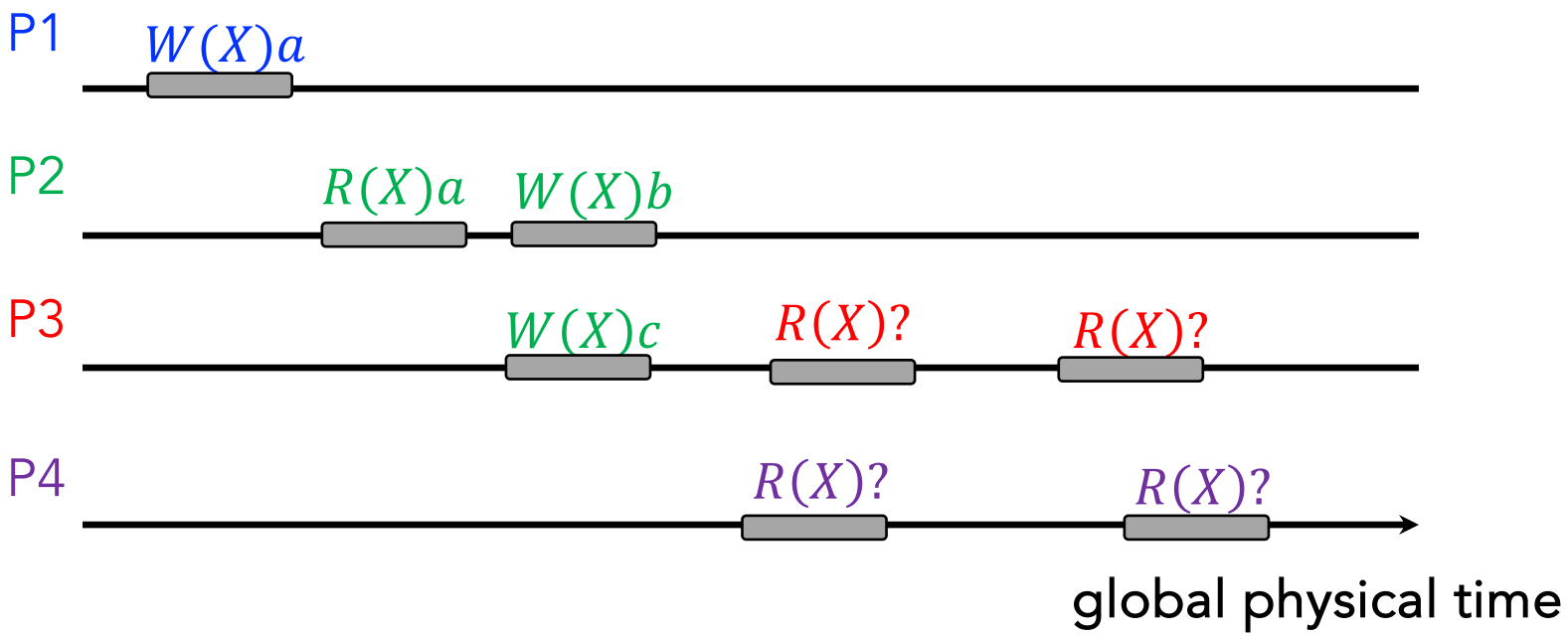
- Causal consistency: potentially causally related operations must be executed in the same order by all replicas
  - In other words, if $a \rightarrow b$, then a must execute before b on all replicas
  - All concurrent ops may be seen in different orders

- Sequential consistency: implies the following:
  1. Operations take place in some total order
  2. The sequence is consistent with the order of operations on each individual client process

Example


P1
P2
P3
P4
$W(X)a$
$W(X)b$
$R(X)?$
$R(X)?$
X
X
X
R1
R2
R3
Distributed Storage System
P1
$W(X)a$
P2
$W(X)b$
P3
$R(X)?$
$R(X)?$
P4
$R(X)?$
$R(X)?$
global physical time

Example



Why Causal Consistency?

- Causal consistency is strictly weaker than sequential consistency although can give strange results, as you have seen
  - If system is sequentially consistent → it is also causally consistent
- BUT: it also offers more possibilities for concurrency
  - Concurrent operations (which are not causally-dependent) can be executed in different orders by different servers
  - In contrast to sequential consistency, we do not need to enforce a global ordering of all operations
  - Hence, one can get better performance than sequential consistency

Eventual Consistency

- Allow divergent replicas
- Allow reads to see stale or conflicting data
- Resolve multiple versions when failures go away
- **Eventually** the replicas in the system reach a convergent state

Eventual Consistency (Cont'd)

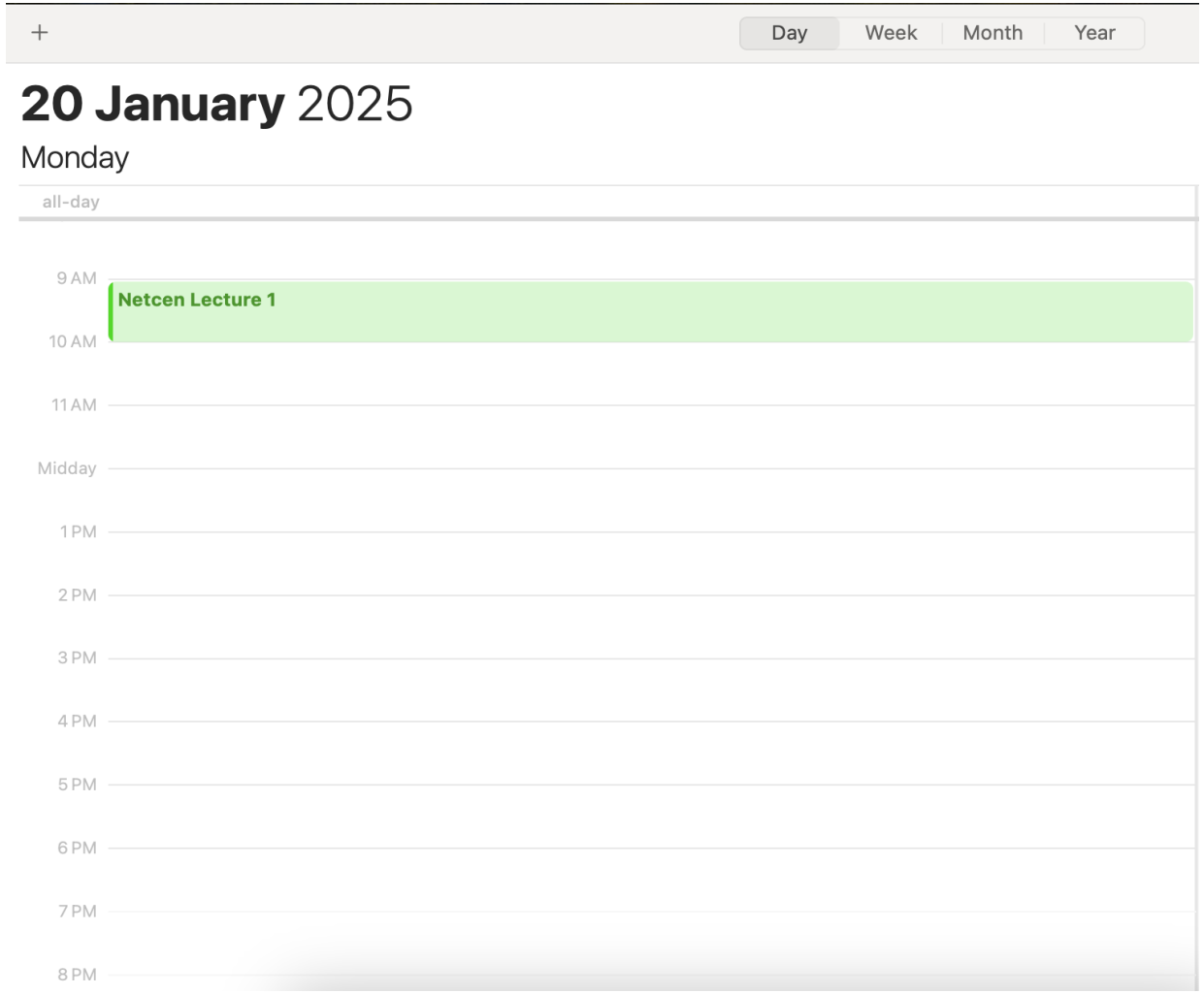
- Note the "eventual" bit in the definition
  - Doesn't tell us when the system will reach a consistent state
  - It is a very weak constraint

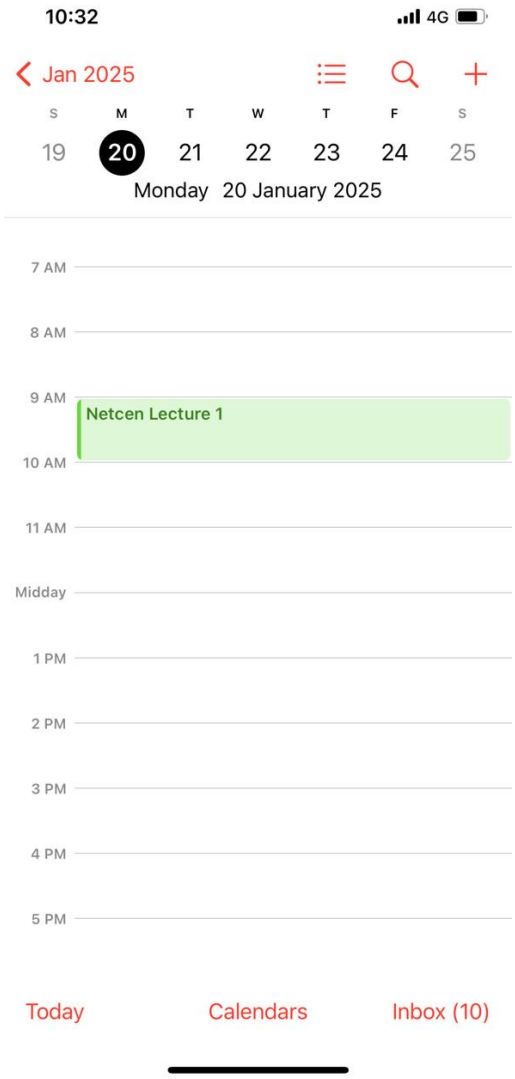
- Since it allows nodes to disagree temporarily
  - Need a conflict resolution mechanism to resolve conflicts and allow the nodes to agree on a common state

Why Eventual Consistency?

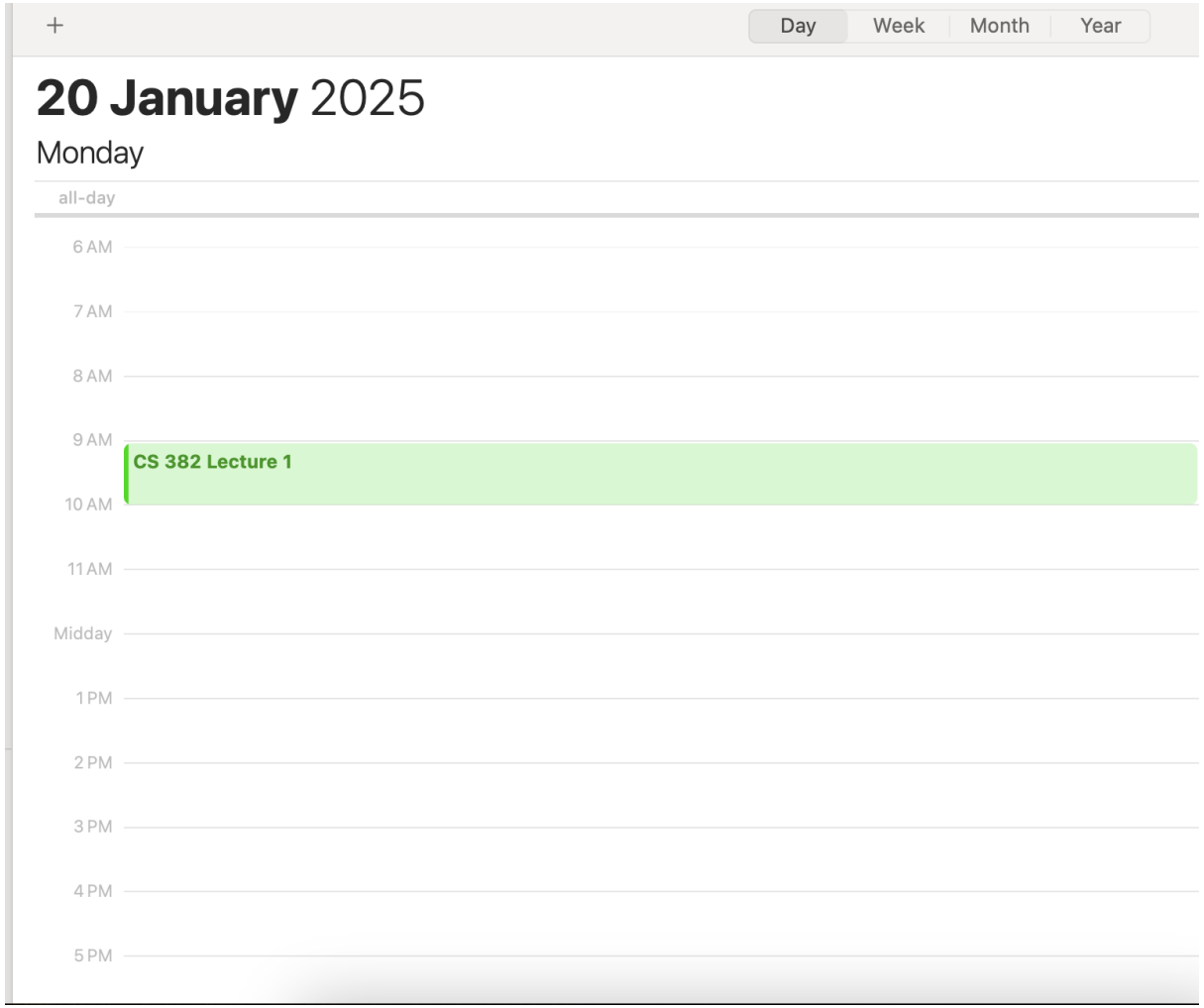
- Support disconnected operations or network partitions
  - Better to read a stale value than nothing
  - Better to save writes somewhere than nothing
- Support for increased parallelism
- Issues
  - Potentially anomalous application behavior
  - Stale reads and conflicting writes…

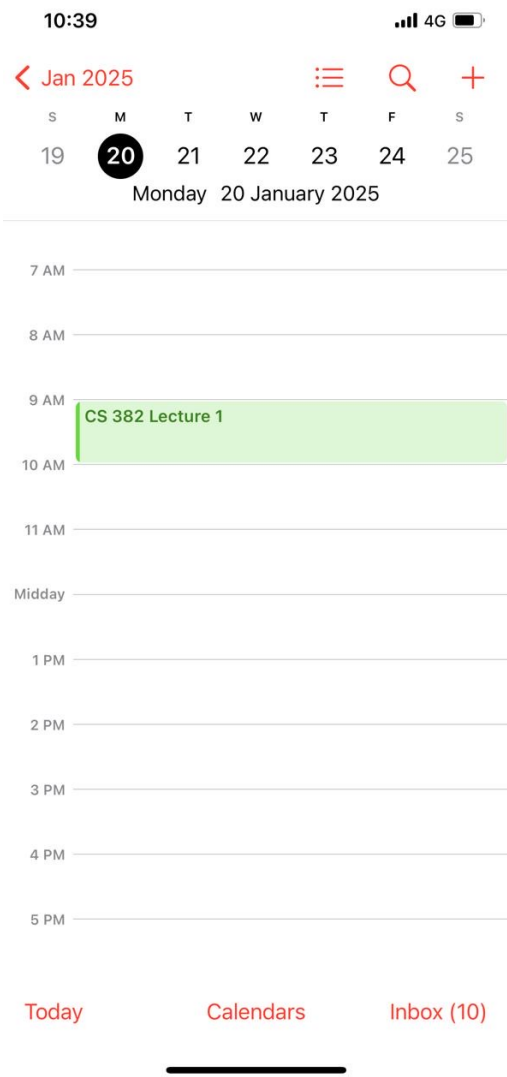
Example: Calendar Application



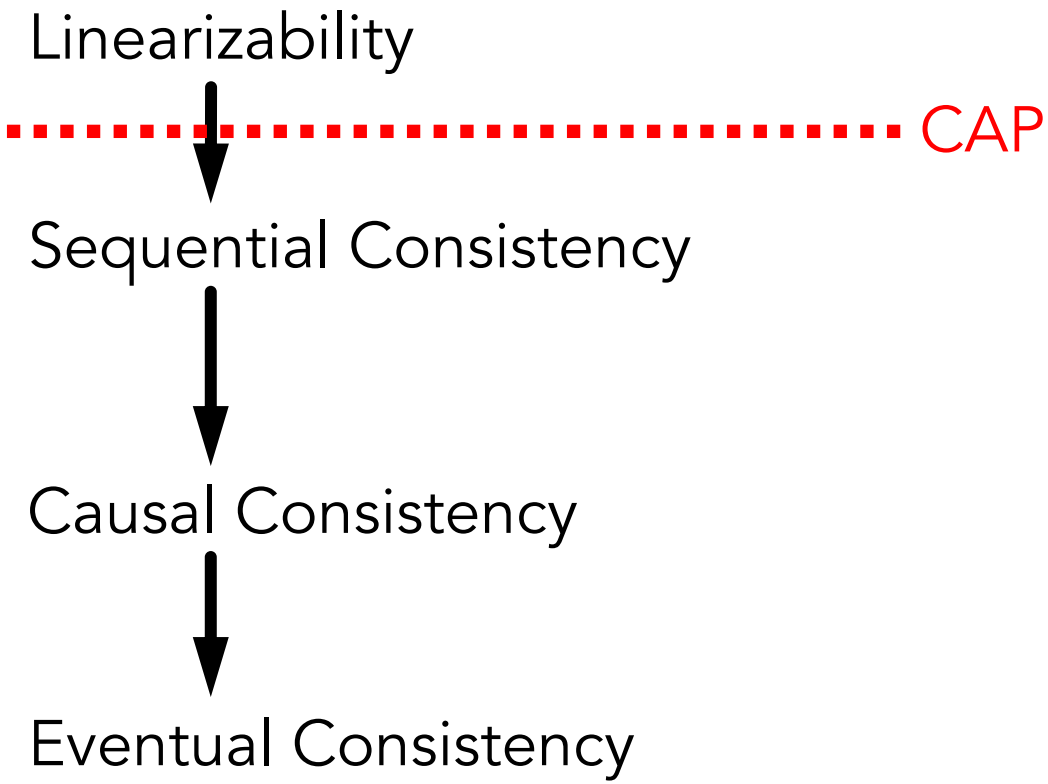


Example: Calendar Application

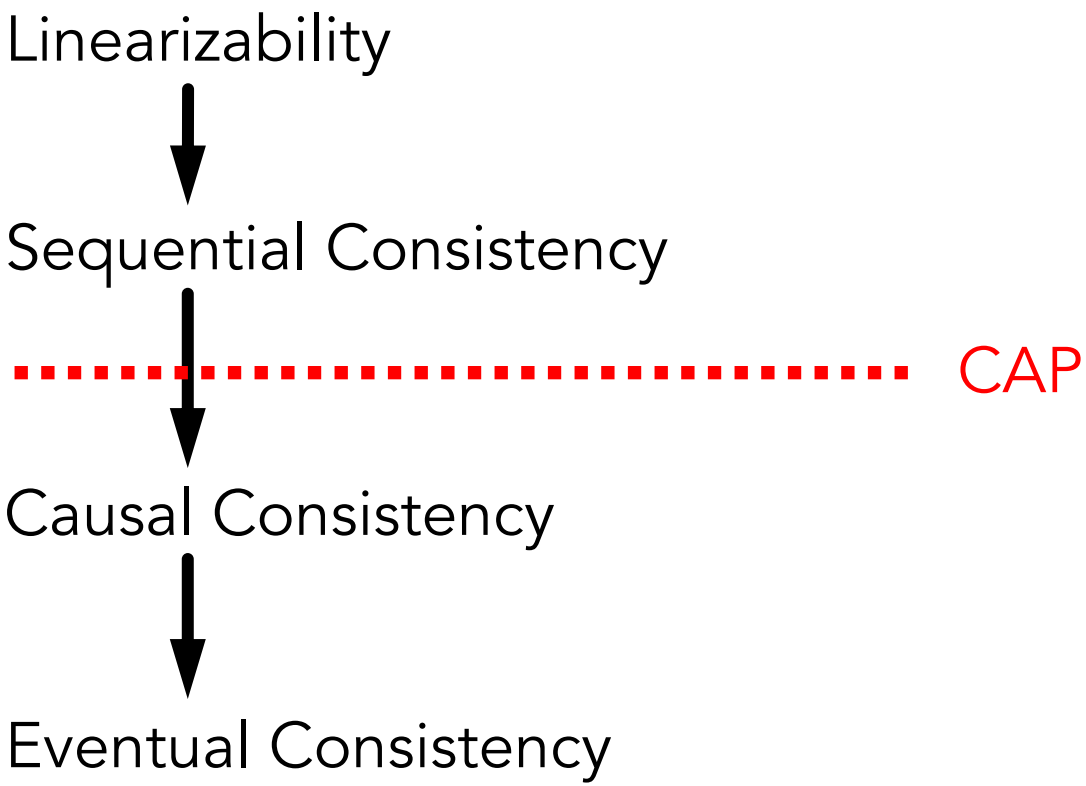




Consistency Hierarchy



Consistency Hierarchy



Consistency Models

Linearizable

Sequential

Causal

Eventual

Unavailable

Not available during network partitions. Replicas will have to pause operations to ensure safety.

Sticky Available

Available on every non-faulty node, as long as clients only talk to the same replicas, instead of switching to new ones

Totally Available

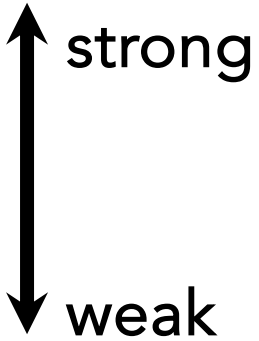
Available on every non-faulty node, even when there is a network partition

Other Consistency Models

- Serializability
- Monotonic reads
- Monotonic writes
- External consistency
- Read your writes
- …
- Read the book from Tanenbaum, Chapter 7.3, if interested

Summary: Consistency Models

- Consistency models
  - Linearizability
  - Sequential consistency
  - Casual consistency
  - Eventual consistency



Variations in:
- Ordering of writes
- Staleness of reads

- Linearizability and sequential consistency:
  - Same sequence of updates at all replicas
  - All replicas agree on the order of the updates