

CSE138 Lecture 14

this time:

- ✓ - announcements
- ✓ - briefly: other consensus protocols besides Paxos
- ✓ - passive vs. active replication
- eventual consistency
- strong convergence and strong eventual consistency
- network partitions
- consistency / availability tradeoff (aka the CAP theorem)

- our first reading assignment: Dynamo
- Class on Zoom next week!
- Check Zulip for an updated A3 test script.

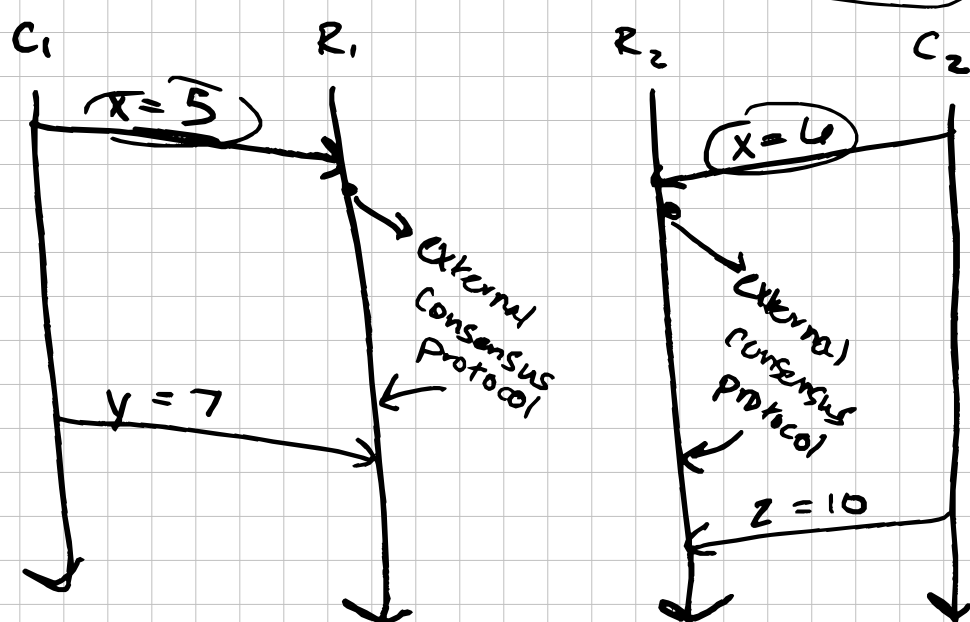
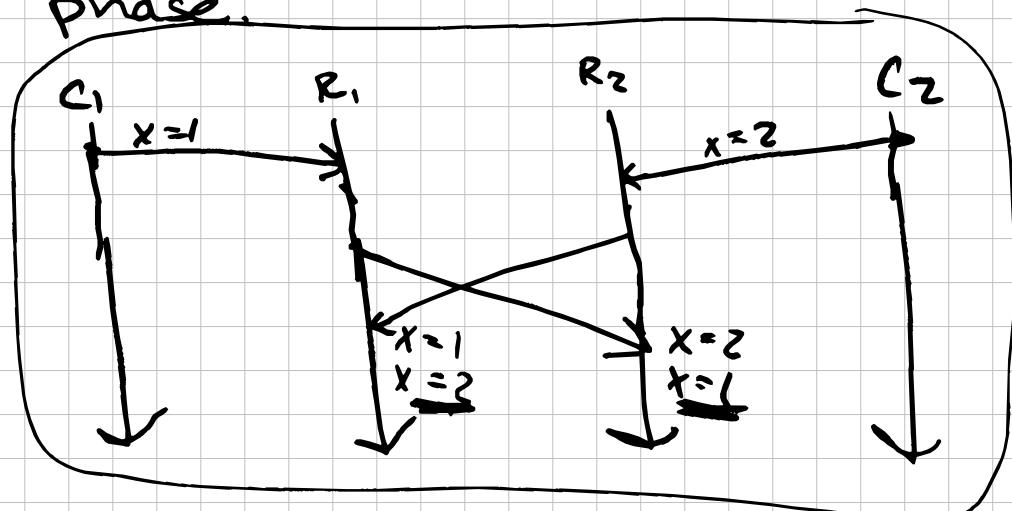
- Viewstamped Replication 1988 (Brian Oki & Barbara Liskov)

- Zab (Zookeeper Atomic Broadcast) Yahoo! Research, late 2000s

- Raft: In Search of an Understandable Consensus Protocol (Diego Ongaro & John Ousterhout, 2014)

- These are all for achieving consensus on a sequence of values. (i.e., like Multi-Paxos).

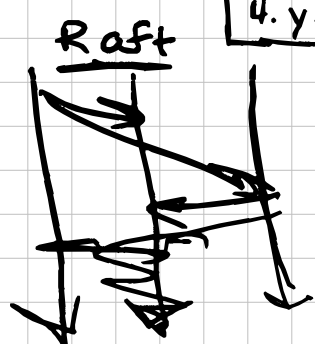
- All have a dedicated leader election phase.



log
1. x=6
2. x=5
3. z=10
4. y=7

log
1. x=6
2. x=5
3. z=10
4. y=7

log
1. ~
2. ~
3. no op
4. no op
5. no op
6. ~

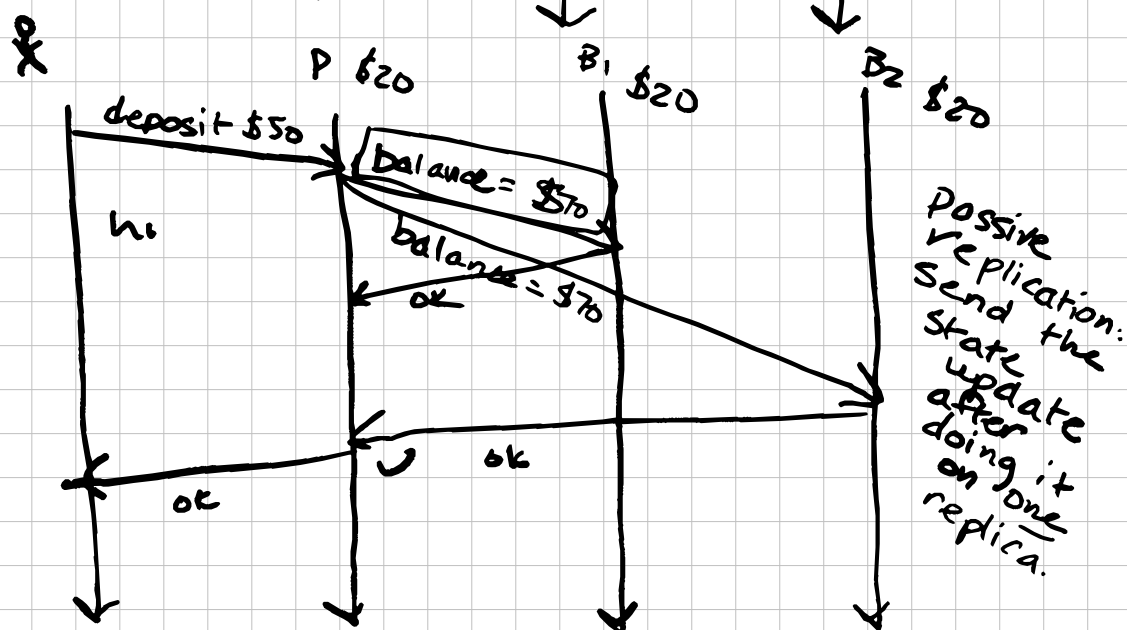
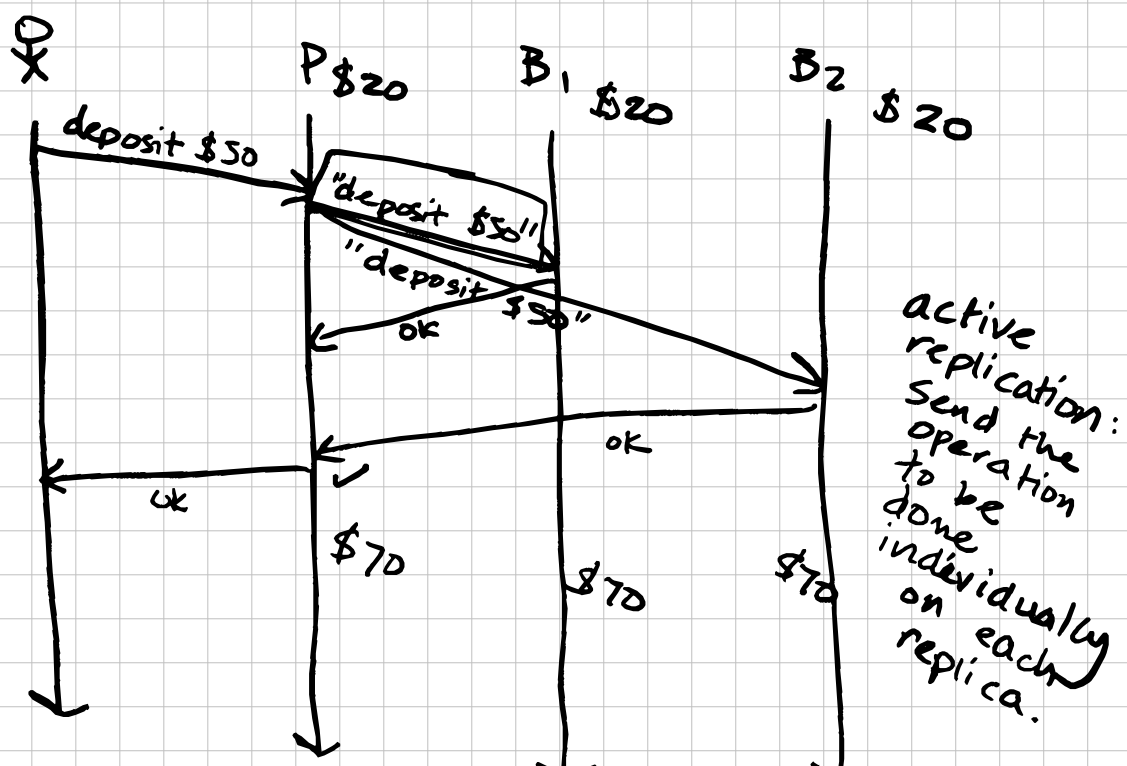


2014: Vive la Différence: Paxos. vs. VSR vs. Zab.

van Renesse, Schiper, Schneider

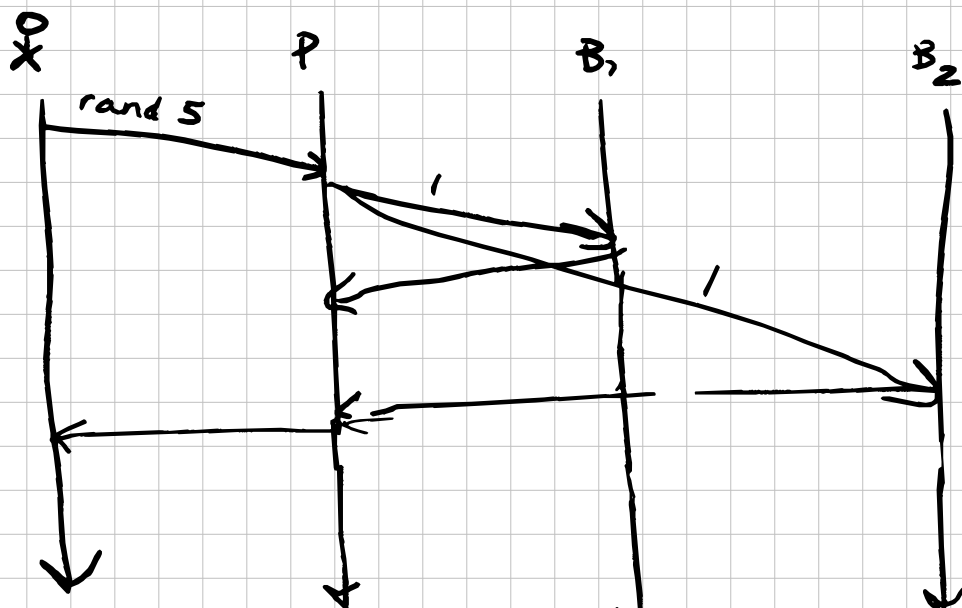
passive vs active replication

↑ aka: state machine replication



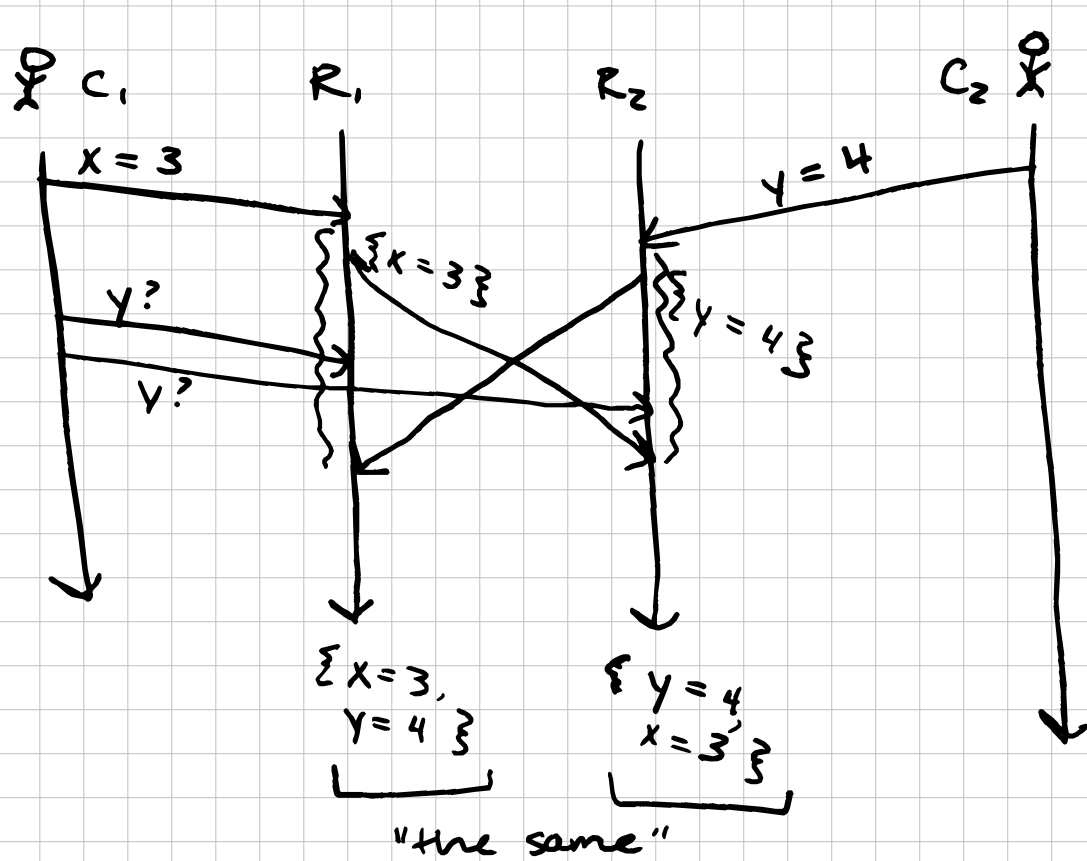
Why pick active or passive?

- Active might be a good choice if the updated state is large.
- Passive might be a good choice if the computation is expensive to run.



One reason to do passive replication is if the computation is nondeterministic!

ideas relevant to the Dynamo paper!



Strong consistency (informally):

Clients can't tell that the data is replicated.

The above execution is not strongly consistent!

(Clients can make intermediate observations that reveal that the data is replicated.)

But it is eventually consistent:

Replicas eventually agree, as long as clients eventually stop submitting updates.
informal definition of eventual consistency.

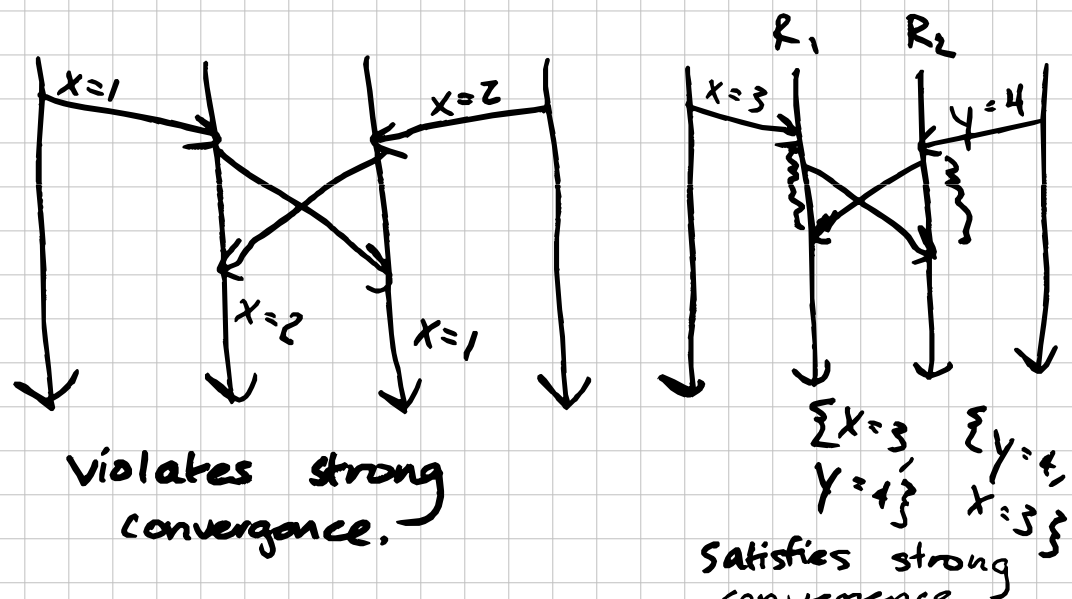
Strong
causal
FIFO
these consistency policies are all safety properties.

Eventual consistency is a liveness property, not a safety property.

But! Here's a safety property:

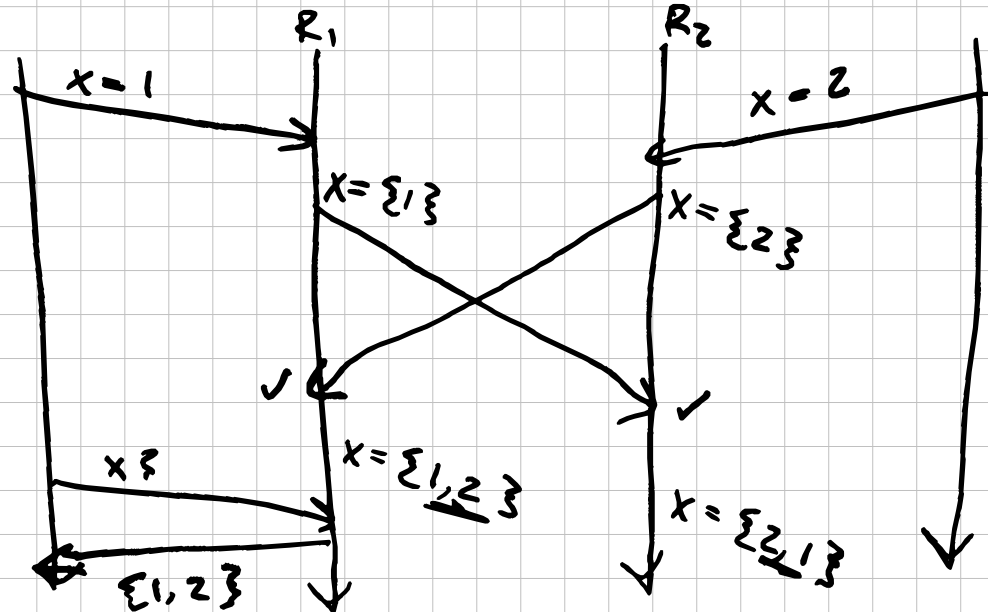
Strong convergence:

Replicas that have delivered the same set (in any order)! of updates have equivalent state.

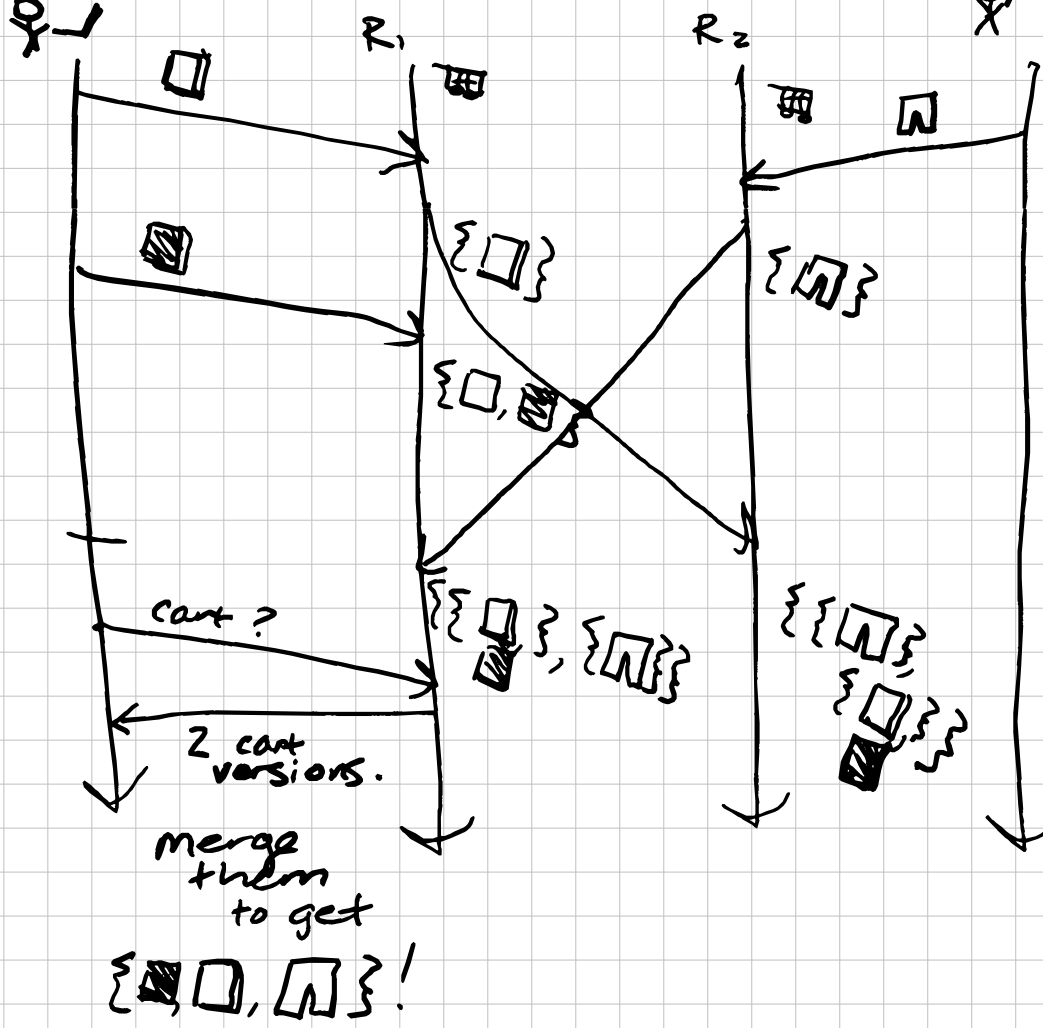


both of the above violate strong consistency.

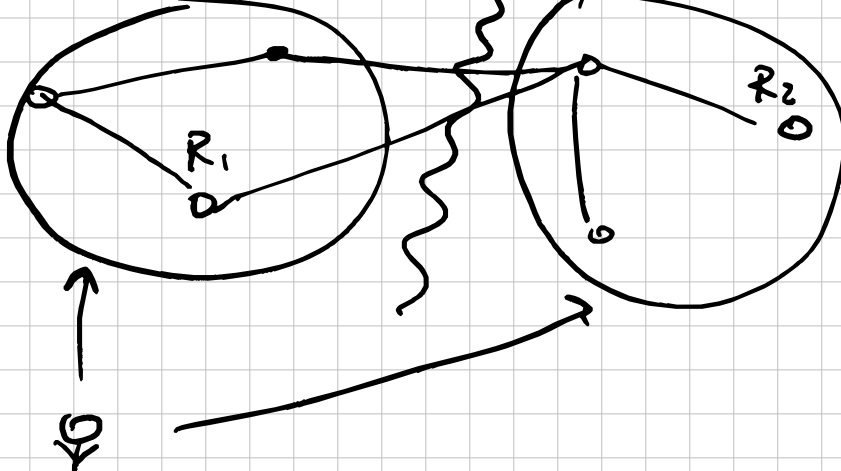
What about this?



Client may need to resolve conflicts on their end, though!



network partitions



Omission model is useful here!

During a network partition, any messages that attempt to cross the partition are lost.

Next time: the consistency/availability tradeoff.