

CSE 550: *Systems for all*

Au 2021

Ratul Mahajan

What is consensus?

Everyone (who matters) agreeing on something

- Occurrence: Did it rain today?
- Ordering: Did the chicken come before the egg?
- Key-value: Is the dress blue/black or white/gold?

Not concerned with correctness
of what is agreed upon



Why do we want consensus?

Replication is a fundamental to fault tolerance

What if replicas disagree ... ?

Why might replicas disagree?

Faults

- Messages can get lost / delayed
- Messages can get corrupted
- Storage can get corrupted
- Nodes can fail and then come up
- Network may get partitioned
- ...

The nature of faults

Fail stop	Byzantine
Working perfectly or not working at all	Arbitrarily bad things can happen
Easier	Harder

Top-level concerns for consensus algorithms (and fault tolerance algorithms in general)

Types of failures tolerated	Paxos is fail-stop, Bitcoin (next week) is byzantine
Replicas needed for N failures	2PC is $N+1$, Paxos is $2N+1$, Byzantine is $3N+1$
Failover speed	2PC is blocking, Paxos is not
Message complexity	2PC and Paxos is linear, Byzantine can be exponential

Trade-offs, trade-offs, trade-offs, ...

Over to Darren and Logan