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 came 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

<table>
<thead>
<tr>
<th>Fail stop</th>
<th>Byzantine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working perfectly or not working at all</td>
<td>Arbitrarily bad things can happen</td>
</tr>
<tr>
<td>Easier</td>
<td>Harder</td>
</tr>
</tbody>
</table>
Top-level concerns for consensus algorithms (and fault tolerance algorithms in general)

<table>
<thead>
<tr>
<th>Types of failures tolerated</th>
<th>Paxos is fail-stop, Bitcoin (next week) is byzantine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicas needed for N failures</td>
<td>2PC is N+1, Paxos is 2N+1, Byzantine is 3N+1</td>
</tr>
<tr>
<td>Failover speed</td>
<td>2PC is blocking, Paxos is not</td>
</tr>
<tr>
<td>Message complexity</td>
<td>2PC and Paxos is linear, Byzantine can be exponential</td>
</tr>
</tbody>
</table>

Trade-offs, trade-offs, trade-offs, …
Over to Darren and Logan