## Recitation 19: Raft II

MIT - 6.033 Spring 2021 Henry Corrigen-Gibbs

Plan \* Recitation Qs Logistics \* Recap: Big picture \* DB hands - on due tomorrow Mrz 5 \* Normal operation \* Operation under failure \* Design project dre May 21 \* Scenarios \* Rost of class is about security...

Recitation Questions

1. The anthors of Raft were boling for an "understandable" consensus alg. what does this mean?

## 2. How understandable is Rast?

3. Why is understandability in ortant? Is 7?

The Big Picture Admin DENY ( ..., ...) Senser that Allow (..., ...) never Sails Door Allovéo? (~) Goal: Implement a never-fail server using a cluster of some-times-fail server. Raft Cluster 7 Þ )~ Raft dient Code

in Raft Normal operation - 12 2 3 - 12 2 3 - 12 2 3 - 12 2 3 - 12 2 3 - 12 2 3 Allow (omir, 329) Allow (omir, 329) Client IP2 \* On alps 1 2 2 3 1. Client talks to leader. Leader pushes update out to all servers. 2 3. When a majority of serves reply to leader \* leader applies change to "state mehine" \* leader replics to dient.

Why a majority of servers?

Imagine a partitioned network







A big mess? Imagine is this happenel with Corio Pass.... prony bad.

The only tricky detail: What happens when the leader Sails? (Problem with all benevilent dictatorships.) Raft des NOT protect against "Byzantine" failures ALLOLIED (mir, 32, 97? 大丁 Client Raft DES prevent against "Sail Stop" Failurs, as long as not "too mary" nots fail. How many? Q D ALLOWED (amily 32,9) ? 

How does cluster recover from failure?

La Leader election.

Vote for Note for Ne? Vote for Ne? Vote for Ne? Vote for Ne? \* When follower doesn't hear from leader after a while -> run election \* Node with most up-to-date log wins Later term or longert by (if bet term equal) Important: Follovers may delete log entries... The only log entries that are committed are ones that LEADER has decided are committed Safter hearing back from My of notes.

Example of how Can't tell commitment From logs alone. orperation of the second secon append 1 2 2 3 N CRASH Servers keep extra state to trach

which entries the cluster has committed.

L> See Sig 2 "commit Index" "Inst Applied"

When a new leader takes over ...

-> It's log is authoritative.

-> Committed cutries stay committed. -> Non-leader logs may change.

Folloners' logs may be in consistent with leaver's log. Loader takes precédence

Rules to remember \* In election, server with most yoto-date log \* Terms increase with each election \* Need a majority of votes to win?





Back to the big picture

The illusion of taking to a rever-fail server, but constantial from many some-times-fail servers.

SVery powerful iten 17

