Chubby (OSDI '06)

strong consistency, by Google

Problem

- everyone needs consistency
- Paxos isn't the right abstraction (?)

Solution

distributed lock manager & filesystem

Lock API

- `Acquire, Release`
- *sequencer* operations

File API

- Open, Close
- Create, Delete
- GetContentsAndStat
- SetContents, SetACL
- event notifications

Sessions

- **problem**: clients can crash while holding locks

- **solution**: locks get leases, too!
- `KeepAlive`

Internal Architecture

- quorum of **five** replicas
- **master** elected with lease
- database log replicated via consensus

Safety

- crashes
- communication delays
- **limited** clock drift

Failover

- master fails to be reelected, lease expires

- new master gathers state from clients before servicing requests

- grace period freezes clients

Handles

Sequencers

API

Use Cases