**Chubby**

Given an implementation of Paxos, what is a useful higher-level abstraction for programmers wanting to do consensus-like-things in a distributed system?

Common things a programmer might want to do:

- elect a leader
- store the location of root of a hierarchy of nodes (or the root of some other data structure)

Both problems require coming to consensus on a value, so clearly a protocol like Paxos needs to be involved. But, Paxos is a fairly low-level primitive – it would be convenient to have a higher-level abstraction that subsumes consensus.

Chubby:

- a small file system built on top of Paxos
- files have names, values, and locks associated with them
- chubby provides upcalls on common events

- makes it very simple to solve the above common problems
    - leader election
        - name a file "service_name.leader"
        - whoever owns the lock is the leader
    - root of data structure
        - name a file "service_name.root"
        - store the name of the root node in the file
        - grab a lock to update the root node

Engineering lessons

- read traffic massively outweighs write traffic
    - read caches on clients
    - kept coherent with invalidation mechanism
- locks and distributed systems are complex
    - what happens if a lock holder fails?
        - Chubby: timeout the lock holder's session after a minute, and as a side-effect, release the lock
    - what happens if an action post-lock-grab is delayed until after the lock is released?
        - introduce sequencers – basically sequence numbers tied to lock state sequence number
- even Chubby can be up but service unavailable
    - maintenance, network partitions