Workshop: Implementing Distributed Consensus

Dan Lüdtke

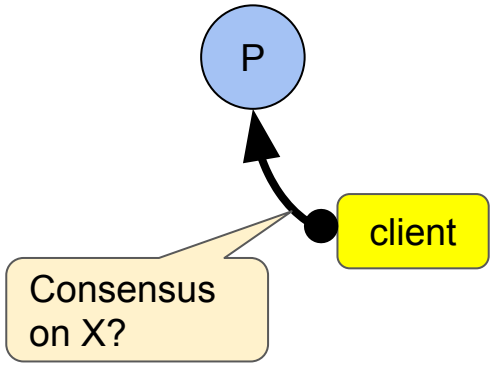
danrl@google.com

Disclaimer This work is not affiliated with any company (including Google). This talk is the result of a personal education project!

The More You Know:
Paxos Roles

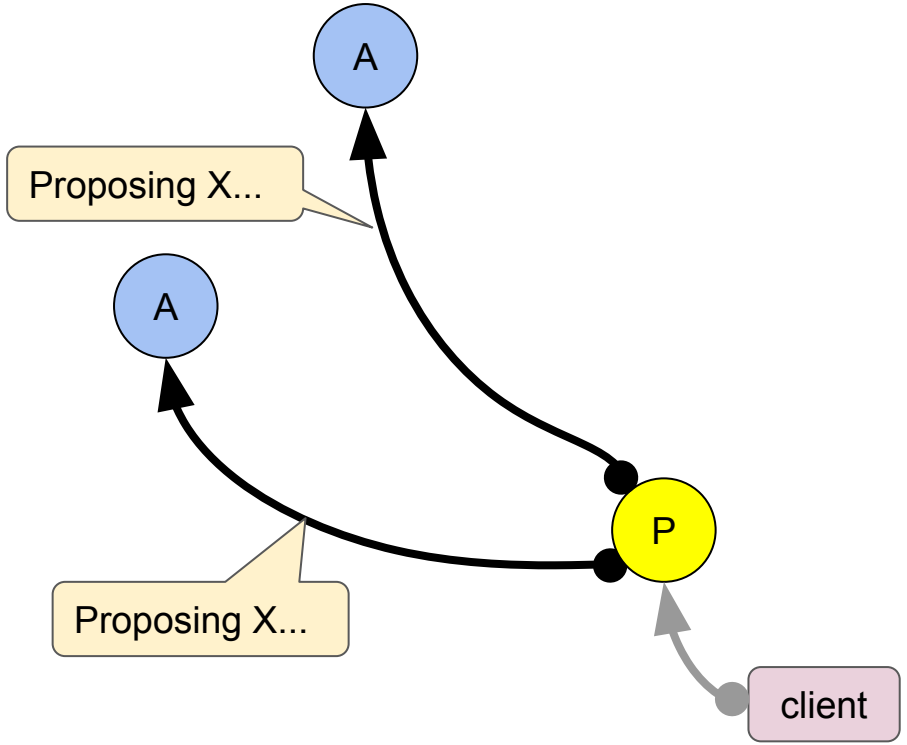
Paxos Roles

- Client
  - Issues request to a *proposer*
  - Waits for response from a *learner*
    - Consensus on value X
    - No consensus on value X
- Proposer
- Acceptor
- Learner
- Leader



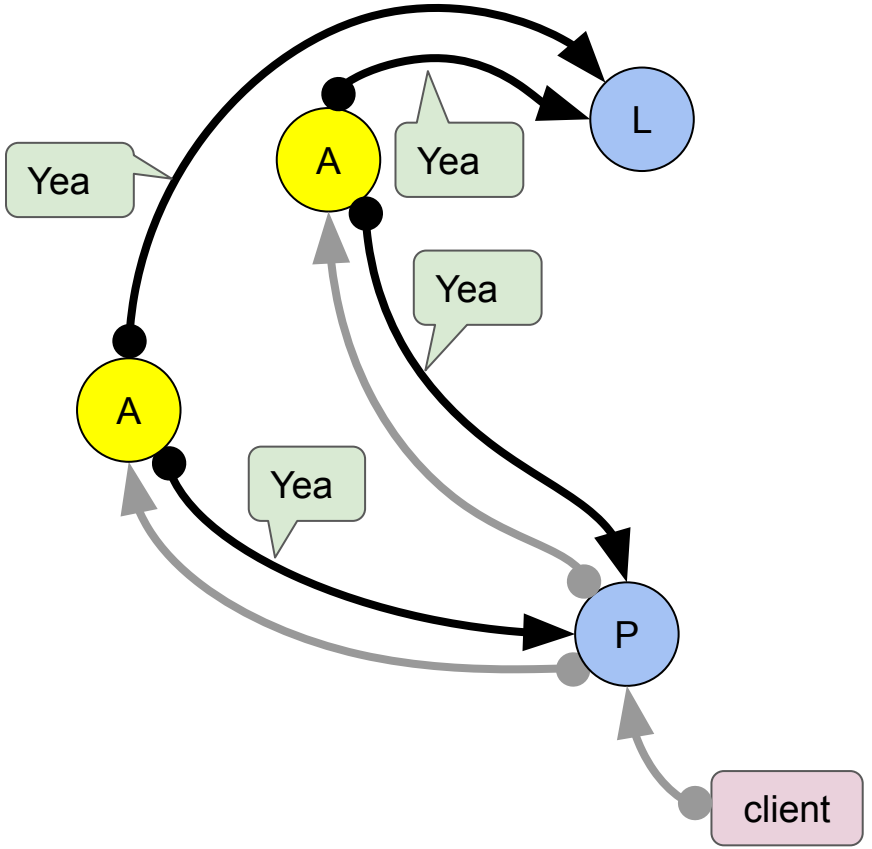
Paxos Roles

- Client
- Proposer (P)
  - Advocates a *client* request
  - Asks acceptors to agree on the proposed value
  - Move the protocol forward when there is conflict
- Acceptor
- Learner
- Leader



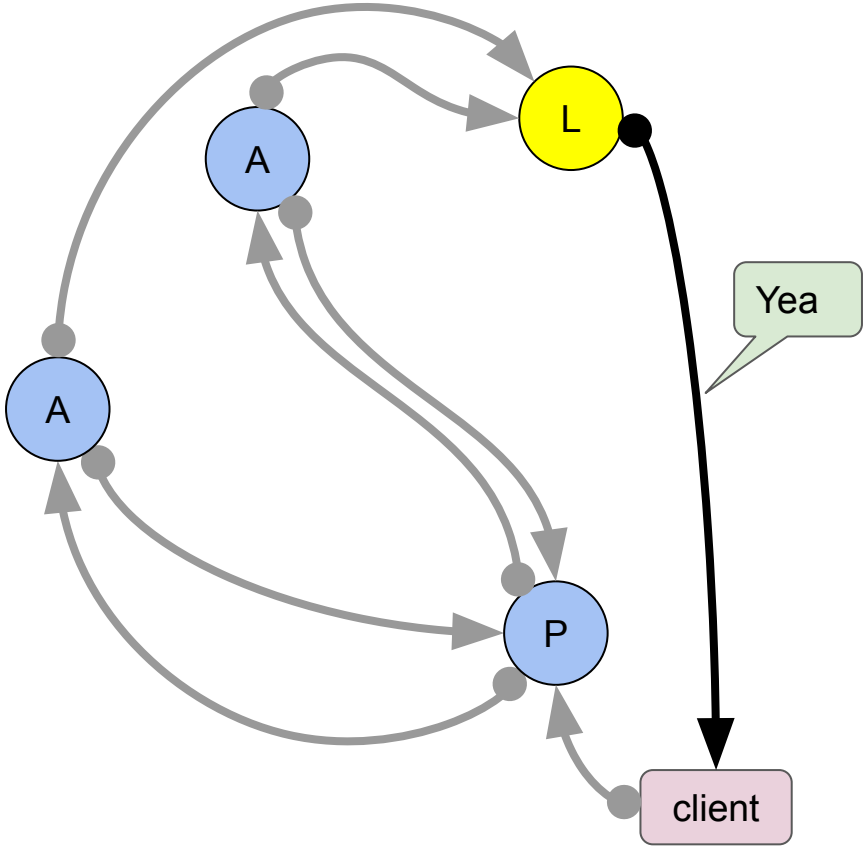
Paxos Roles

- Client
- Proposer (P)
- Acceptor (A)
  - Also called "voter"
  - The fault-tolerant "memory" of the system
  - Groups of acceptors form a *quorum*
- Learner
- Leader



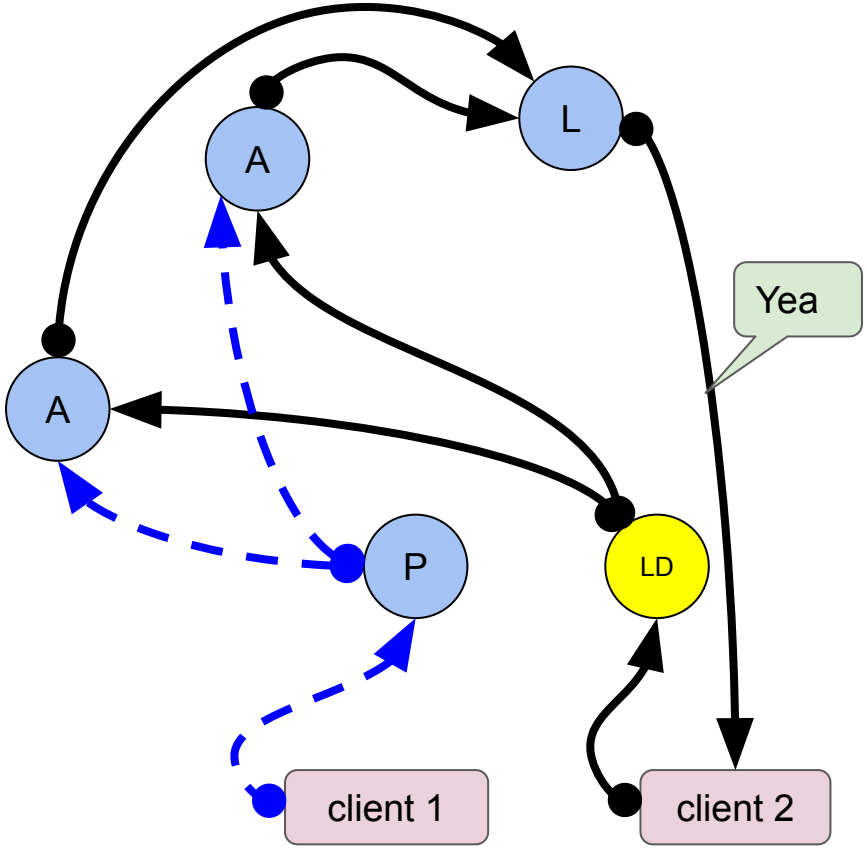
Paxos Roles

- Client
- Proposer (P)
- Acceptor (A)
- Learner (L)
  - Adds replication to the protocol
  - Takes action on learned (agreed on) values
  - E.g. respond to *client*
- Leader



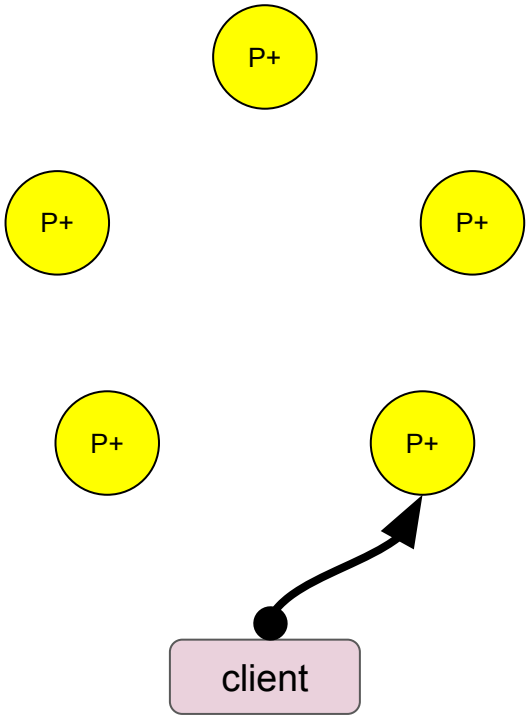
Paxos Roles

- Client
- Proposer (P)
- Acceptor (A)
- Learner (L)
- Leader (LD)
  - Distinguished *proposer*
  - The only *proposer* that can make progress
  - Multiple *proposers* may believe to be leader
  - *Acceptors* decide which one gets a majority



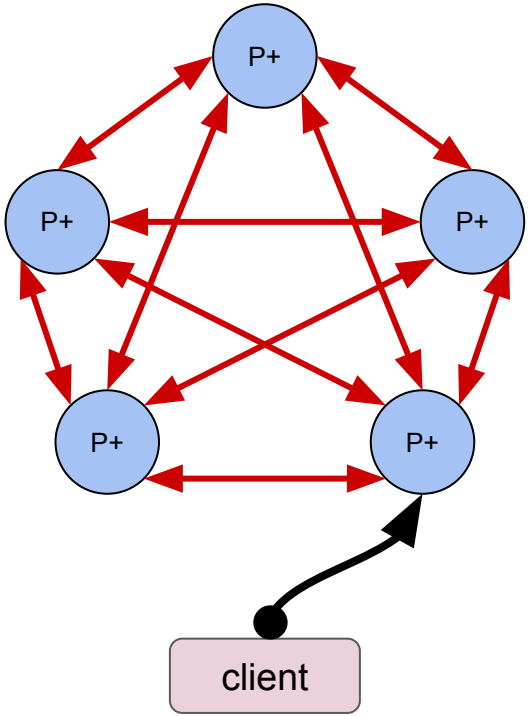
Coalesced Roles

- A single processor can have multiple roles
- P+
  - Proposer
  - Acceptor
  - Learner
- Client talks to any processor
  - Nearest one?
  - Leader?



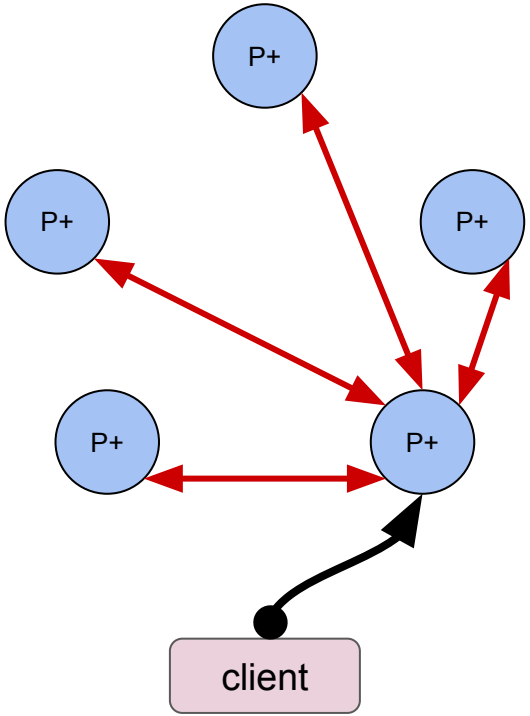
Coalesced Roles at Scale

- P+ system is a complete digraph
  - a directed graph in which every pair of distinct vertices is connected by a pair of unique edges
  - *Everyone talks to everyone*
- Let **n** be the number of processors
  - n = Quorum Size
- **Connections** = **n** * (**n** - 1)
  - Potential network (TCP) connections

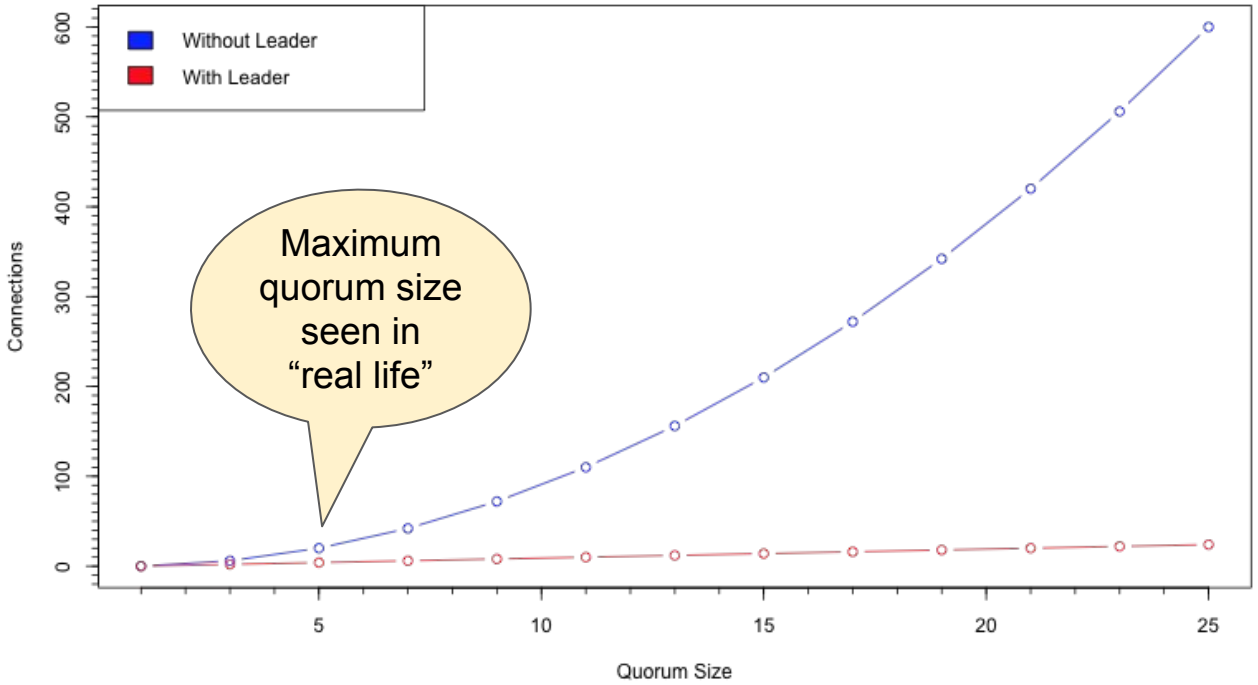


Coalesced Roles with Leader

- P+ system with a leader is a directed graph
  - *Leader talks to everyone else*
- Let **n** be the number of processors
  - a.k.a. Quorum Size
- **Connections** = **n** - 1
  - Network (TCP) connections



Coalesced Roles at Scale



Limitations

- Single consensus
- Multiple proposers may believe to be the leader
- Standard Paxos not resilient against Byzantine failures