EGALITARIAN PAXOS: THERE IS MORE CONSENSUS IN EGALITARIAN PARLIAMENTS
Iulian Moraru, David Andersen (Carnegie Mellon University), Michael Kaminsky (Intel)

PAXOS OVERVIEW

- State Machine Replication:
  - Fault tolerance through redundancy
  - All replicas execute the same commands in the same order
  - Tolerates F failures with 2F+1 replicas
  - Replicas can fail by crashing (non-Byzantine)

BOTTLENECK IN (MULTI-)PAXOS

- Leader brokers all communication with clients
- Handles O(N) messages per command
- State machine unavailable until new leader is elected after a failure

EGALITARIAN PAXOS

- Clients submit commands to any replica
- No contention for instances
- Available without interruption if F+1 replicas are non-faulty (2F+1 replicas total)
- Leader brokers all communication with clients
- Handles O(N) messages per command
- State machine unavailable until new leader is elected after a failure

EXECUTION ALGORITHM

- Performed independently on each replica:
  1. Wait until command C is committed
  2. Build C's dependency graph recursively
  3. Find strongly connected components (SCCs)
  4. Execute:
     - Execute SCCs in inverse topological order
     - Execute commands within each SCC in increasing sequence number order

CONCLUSIONS

- High throughput due to load balancing
- Optimal commit latency in wide area when tolerating 1 and 2 faults
- Constantly available if majority of replicas alive
- Better handling of slow replicas than previous Paxos versions