Egalitarian Paxos
Iulian Moraru, David Andersen, Michael Kaminsky* (CMU, *Intel)

EPAXOS OVERVIEW
- Optimal latency in the wide-area for 1 and 2 failures
- Higher throughput in a cluster (due to load balancing)
- Constant availability if majority of replicas alive
- Better handling of slow replicas

BOTTLENECK IN (MULTI-)PAXOS
Clients send commands to a stable leader
Longer delay if leader crashes
Distant clients too (worse latency)

EGALITARIAN PAXOS
Clients choose any replica

INTUITION
Paxos: Pre-ordered instance space
EPaxos: Instances ordered at commit time (ordering constraints)

EVALUATION

CONSTANT AVAILABILITY

HIGHER THROUGHPUT
- Order only commands that interfere
- Ordering constraints: dependency list, approx. seq. #
- Fast path (1RTT) quorum: $\frac{F}{2} + 1$
  - Better than Generalized Paxos by 1
- Slow path (2RTT) quorum: $F + 1$

LOWER LATENCY

OPTIMAL WIDE-AREA LATENCY