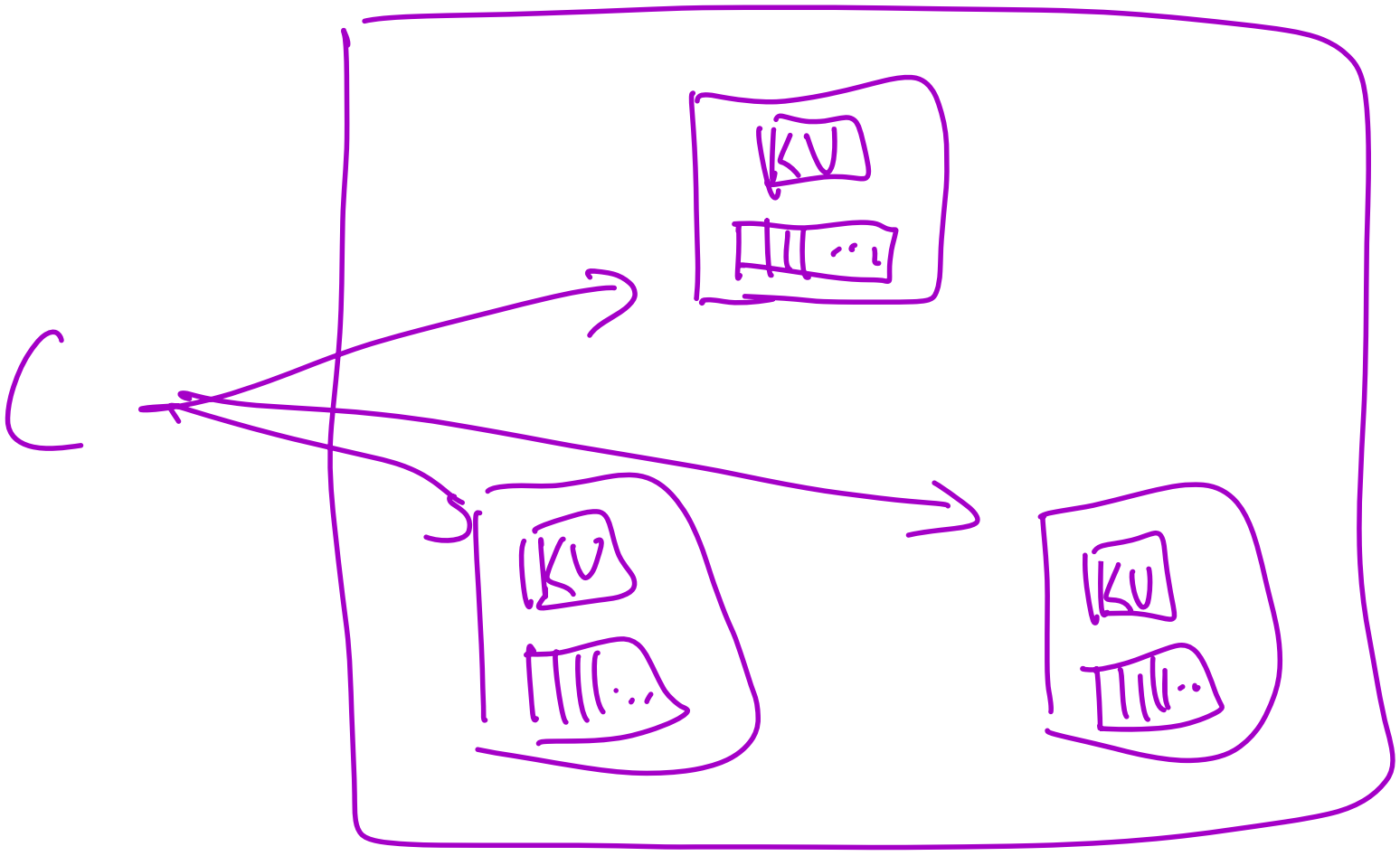
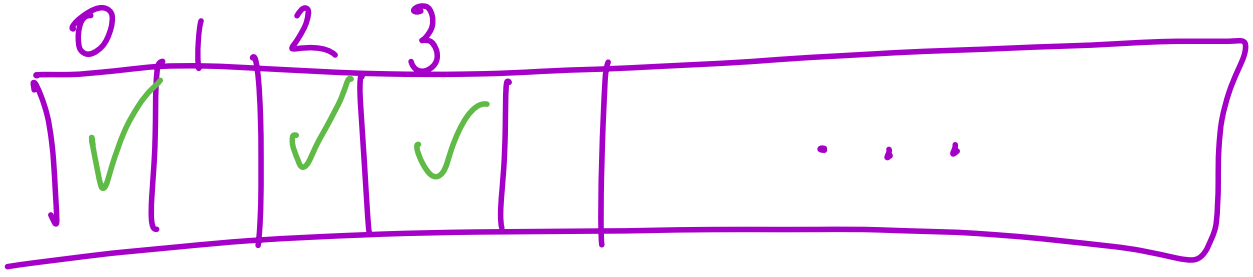
CSE 452

Distributed Systems

MultiPaxos





slot = index into log

Differences from single decree

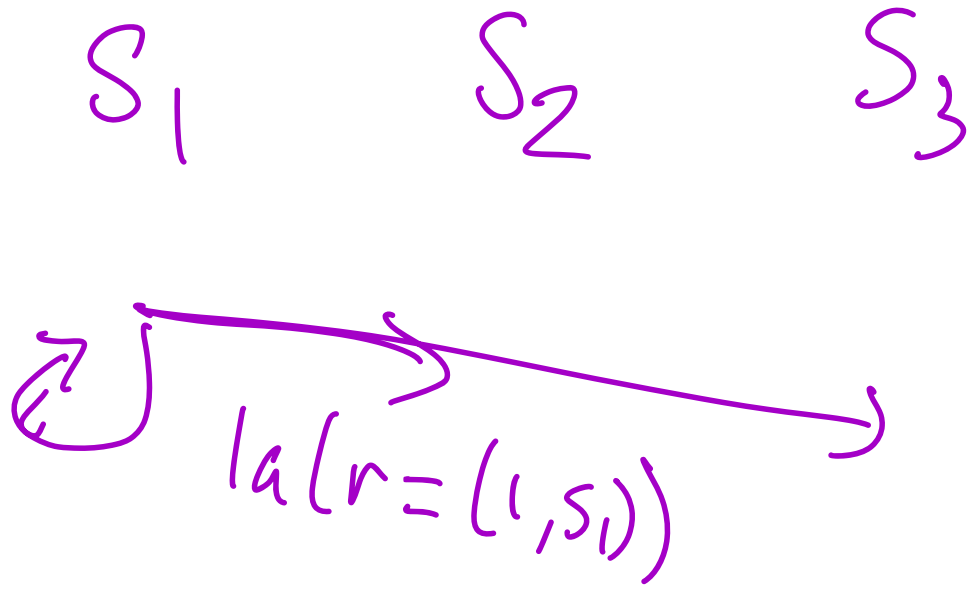
- all nodes play all roles
- ballot "numbers" are pairs

  (seq num, server-id)

  w/ lexicographic order

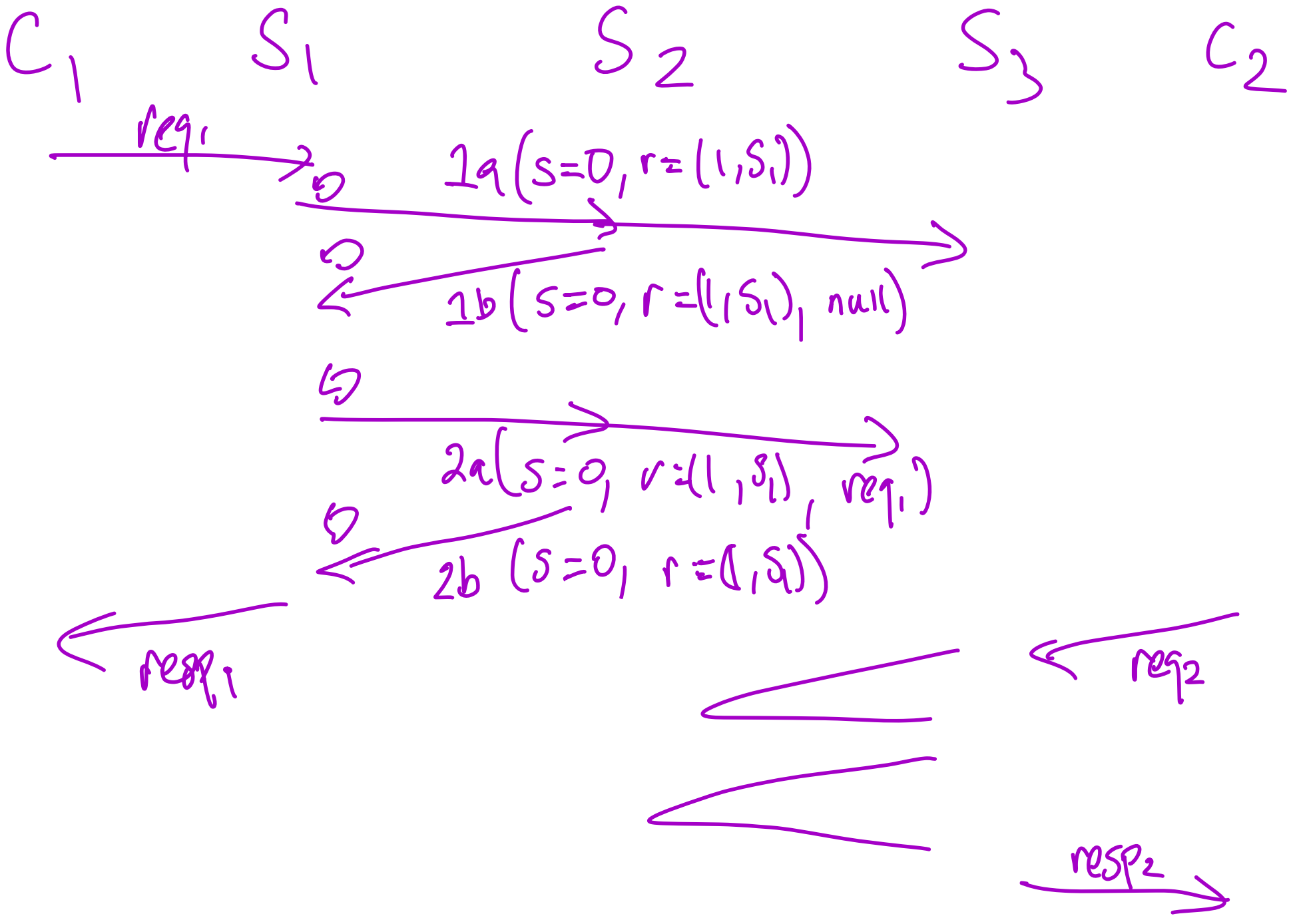
  $(1, S_2) < (4, S_1)$

  $(1, S_2) > (1, S_1)$


S_1
S_2
S_3
$la(r=(1,s_1))$

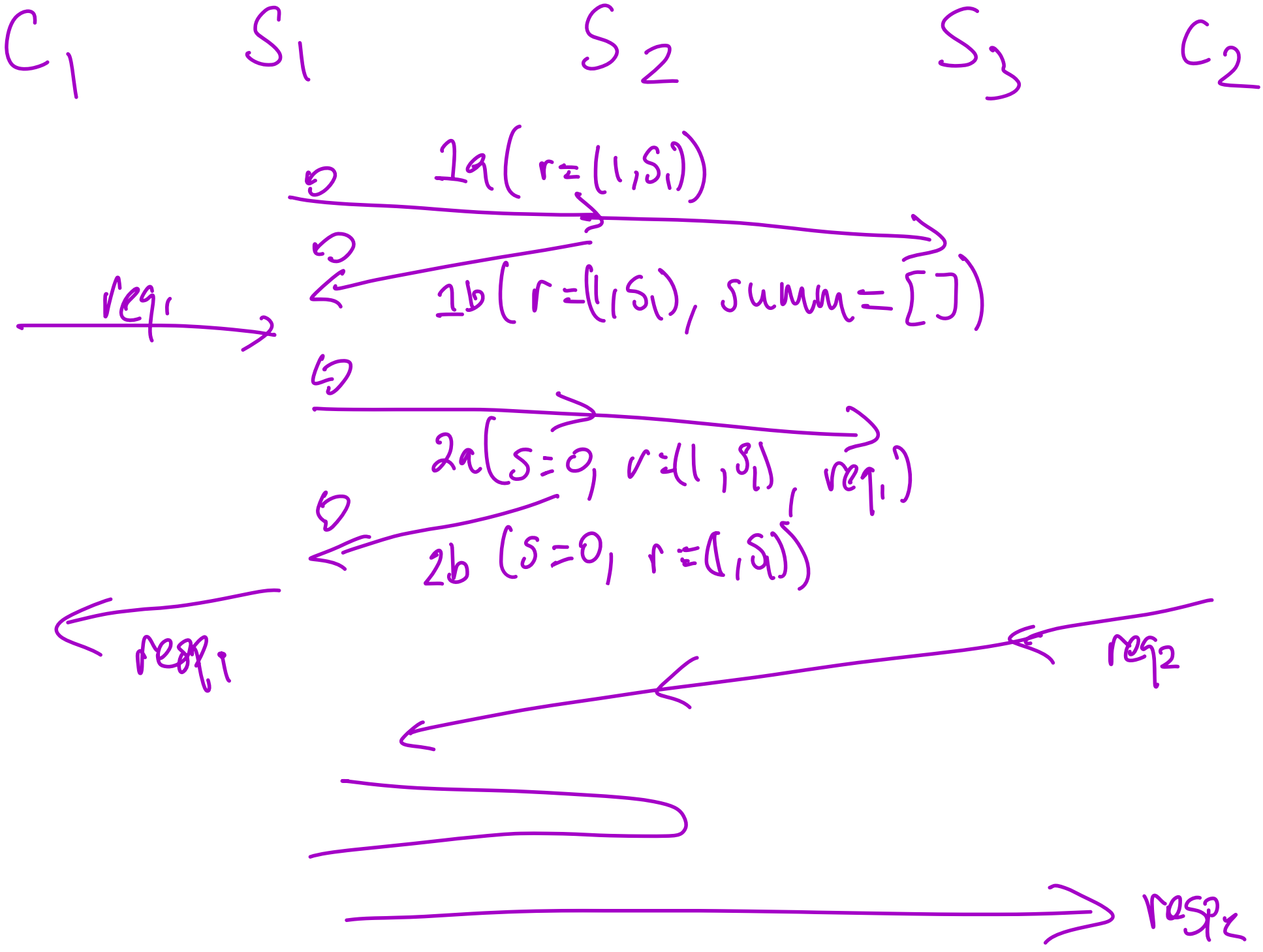
Unoptimized MultiPaxos

- SDP in each slot
- clients broadcast reqs to all servers
  - find empty slot, propose this request
- wait for prefix of log chosen
  - execute request
    - respond to client


C1
S1
S2
S3
C2
req1
1a (s=0, r=(1,S1))
1b (s=0, r=(1,S1), null)
2a (s=0, r=(1,S1), req1)
2b (s=0, r=(1,S))
resp1
req2
resp2

Distinguished Proposer optimization (Leader Election)

- use phase 1 to elect the leader
- combine phase 1 across all slots
  - delete slot # from 1a/1b
  - 1b summarize votes in all slots (list)
- phase 2 unchanged


C1
S1
S2
S3
C2
1a(r=(1,S1))
1b(r=(1,S1), summ=[])
req1
2a(s=0, r=(1,S1), req1)
2b (s=0, r=(1,S1))
resp1
req2
resp2

Heartbeats

- leader tell followers that they are still up
- followers set a timer to check if they've heard from the leader recently
  - if not, try to become leader