

Day 5

CSD 1380



Last Glass

RPC Semantics & Mechanism

at-least-once

at-most-once

exactly-once

request retry
duplicate suppress

response replay

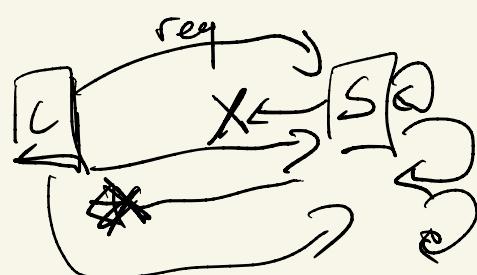
different
combinations

provide

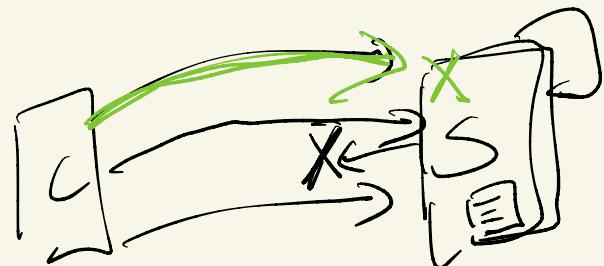
request-retry



"at-least-once"

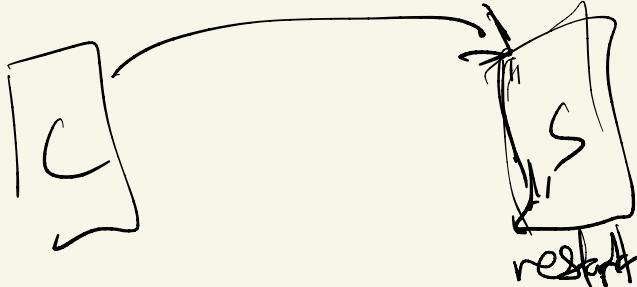


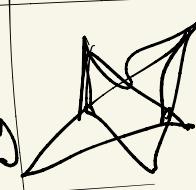
request retry + suppress
duplicates



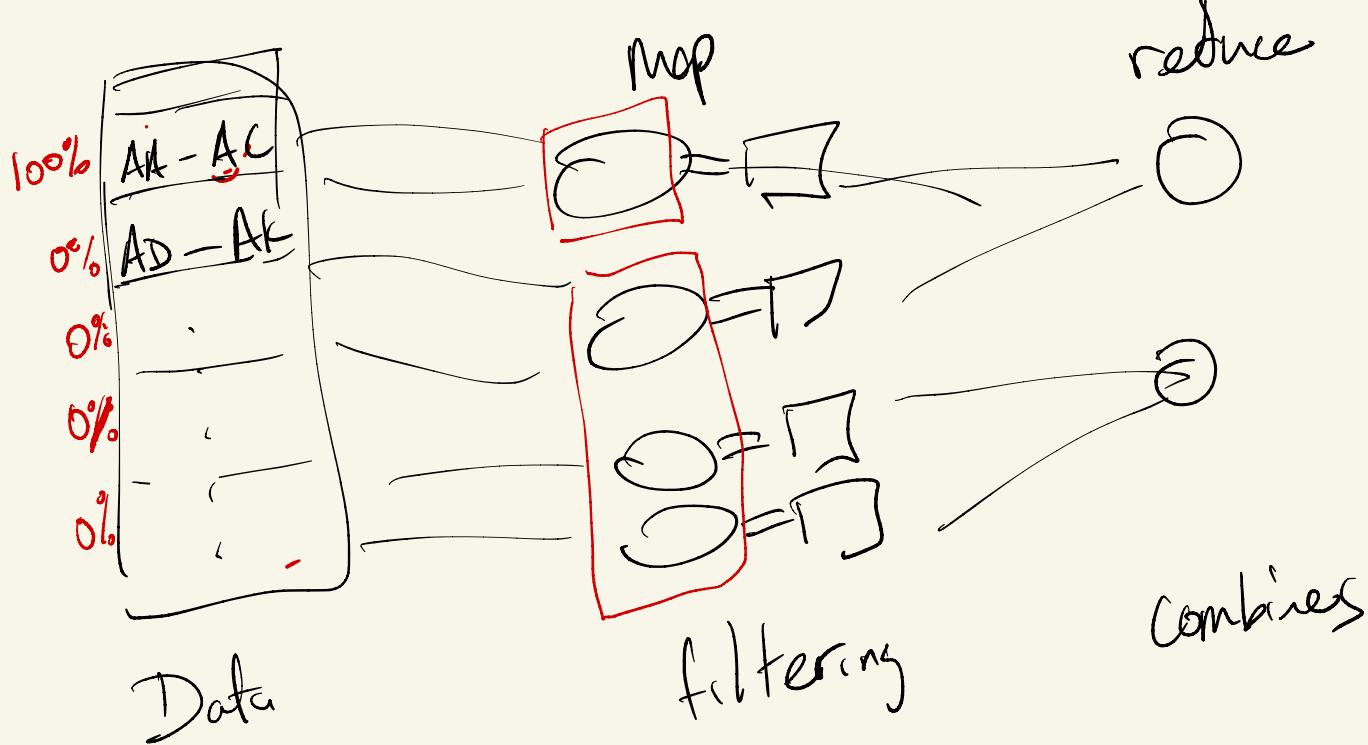
+
suppress
duplicates

+
response
replay



at least once	at most once	exactly once
retry	retry + suppression + response replay	

Map / reduce



select count(*) from Brown where name = "theo"

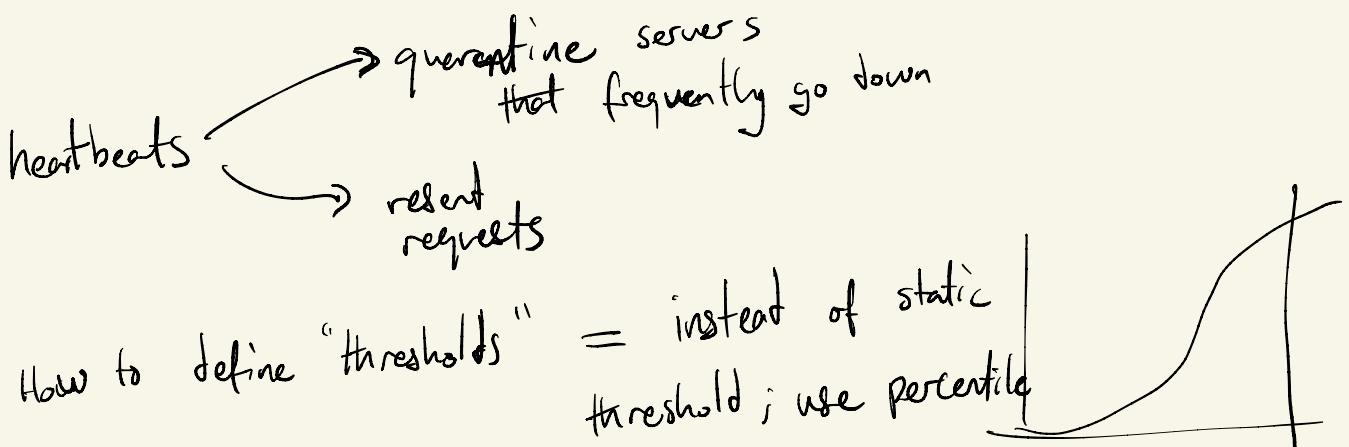
filter [map]

reduce

- ① distance (network distance)
 - ② uneven data partition
 - ③ Background mappers speeds (Server speeds)
 - ④ servers failures / lost jobs
- CPU monitoring
- I/O throughput
- H/W specification
- load balancing
- Heartbeats
- n/w topology

Problem detection

- ① heartbeats (servers that keep going down) → Solution (Designs)
- ② progress reports (helps track fast V slow)
- ③ monitoring (CPU / IO / queue length)



productivity rating = send job that are proportional in size to ratings

↑
increase / decrease the # of process or that server

cloning request — where two servers do the same thing & you pick fastest

④ reduce total time

- ① original server might be slow now but they can change
 ② just care about who finishes first

monitoring \Rightarrow weights for either load balancing
or placement of map/reduce workers/tasks

Summary of today	Map reduce (structure)	(heartbeats progress report monitoring / fw-spec + topology)
	* signals for detecting issues * ways to address fix issues (retry, cloning, quarantining)	

Next week: Load Balancing

