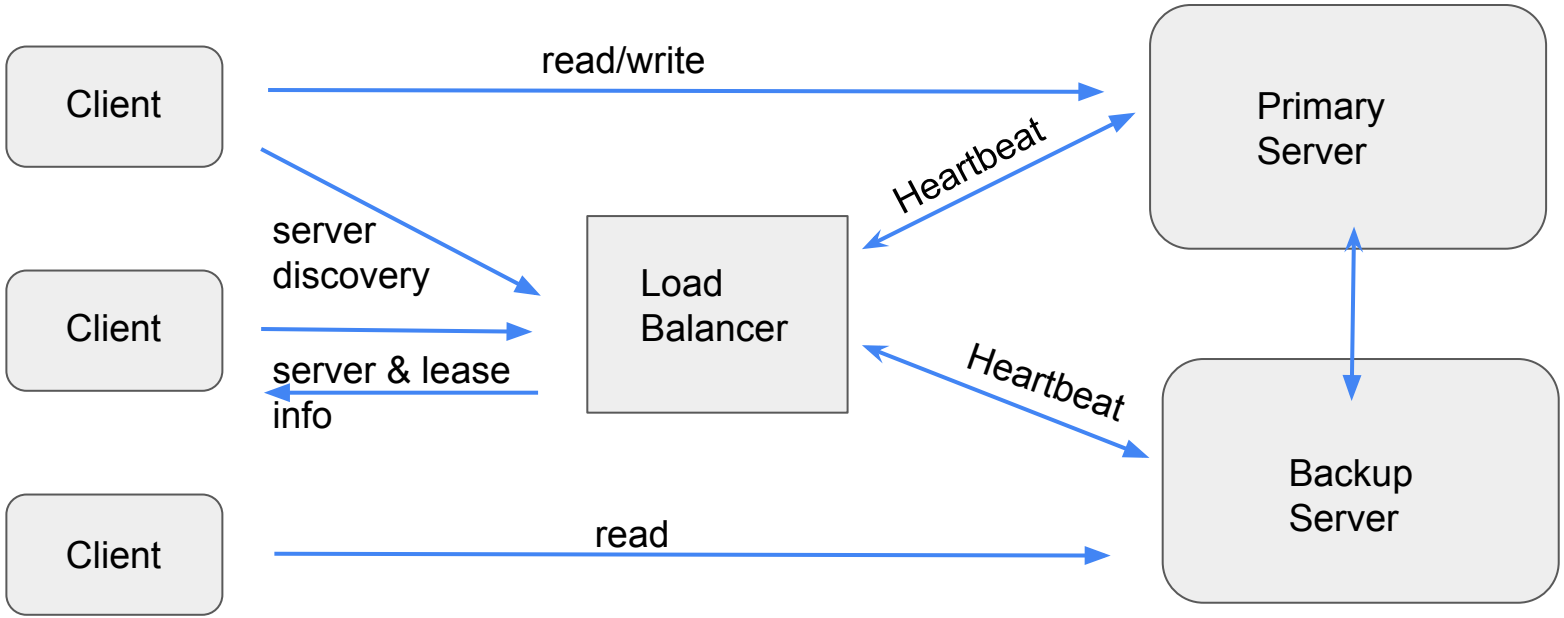
CS 739 Distributed Systems

P3 : Replicated Block Store

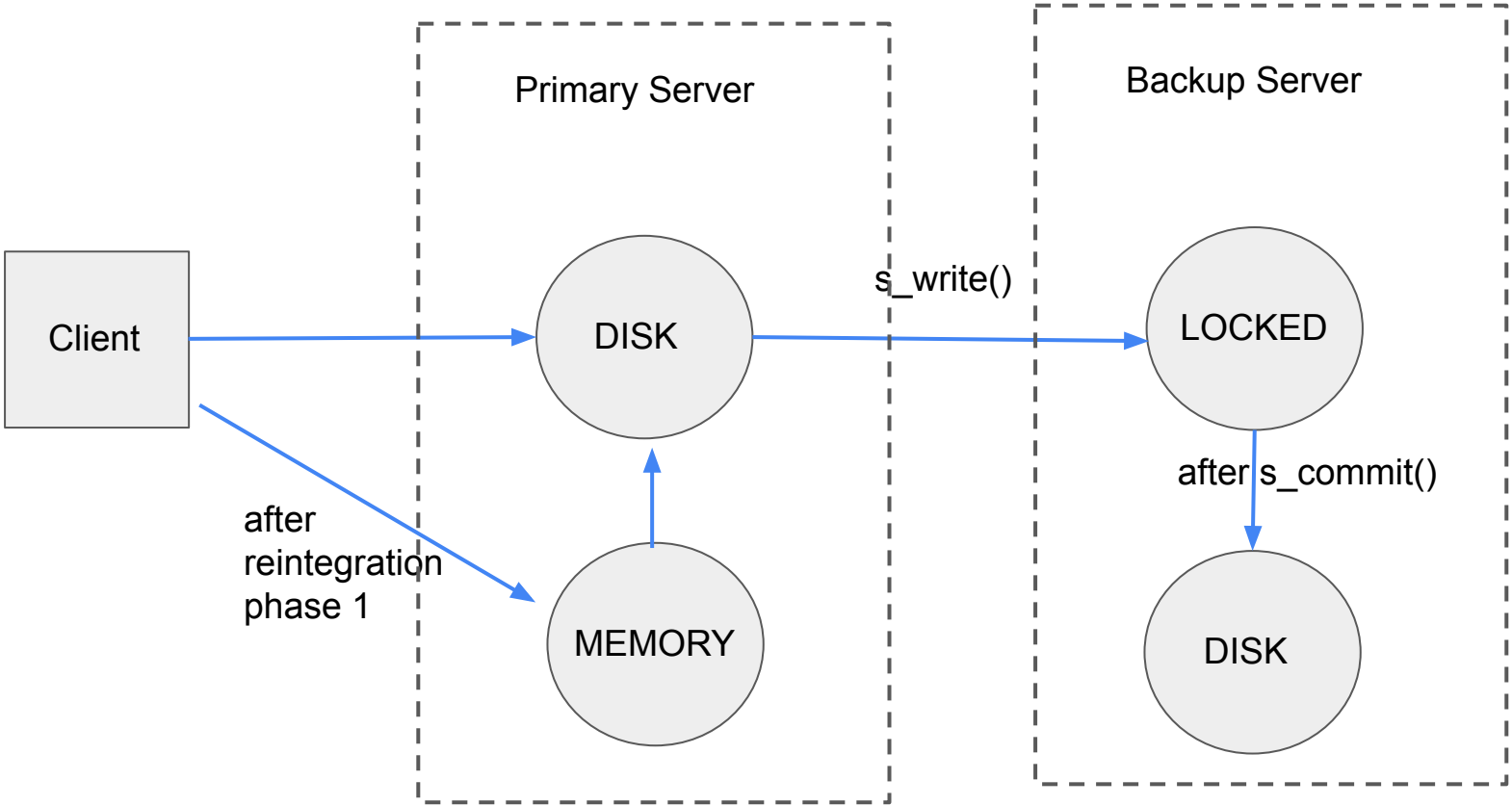
Group 4 :
Abhinav Agarwal,
Kaustubh Khare,
Prabhav Adhikari,
Chahak Tharani

1. Design & Implementation

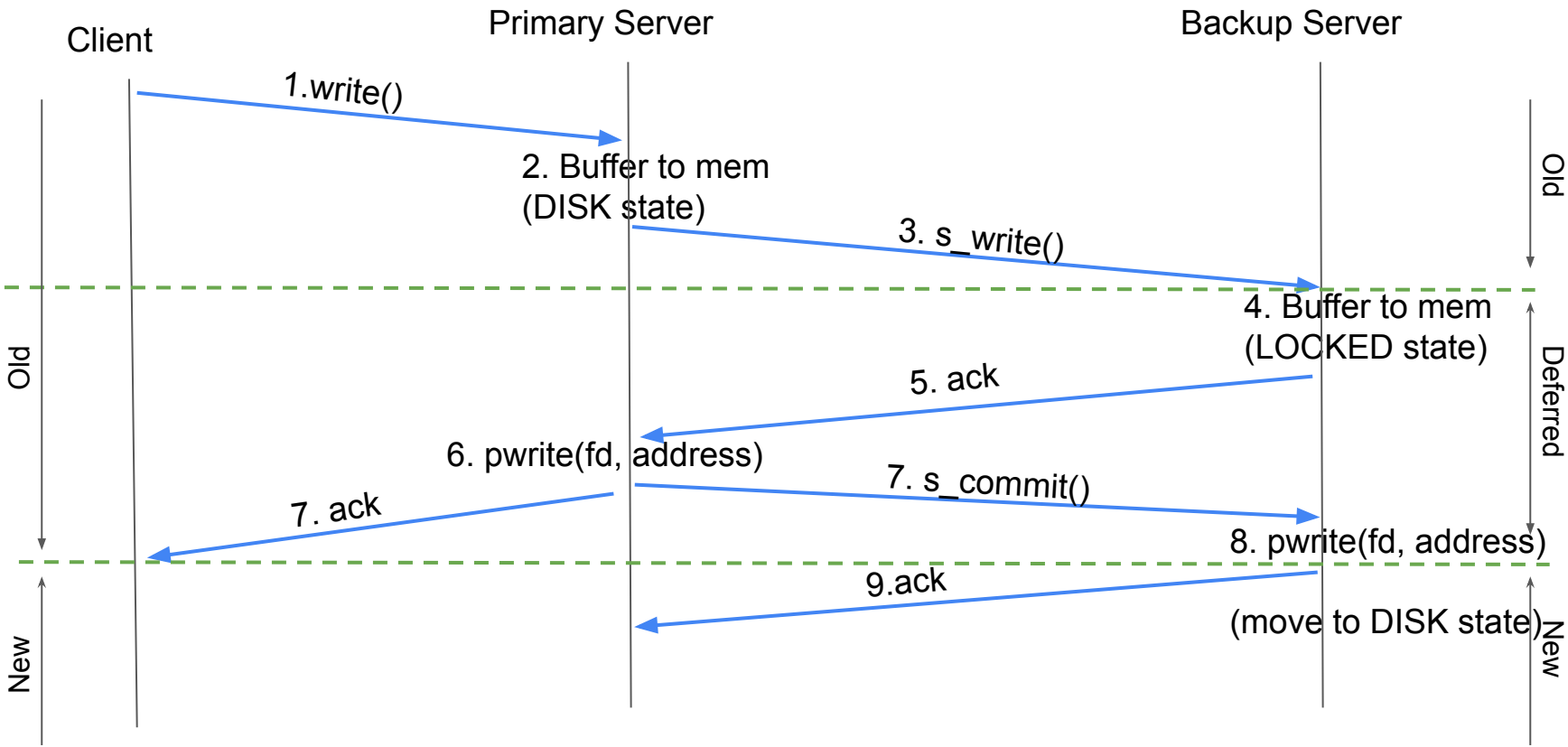
High Level System Design



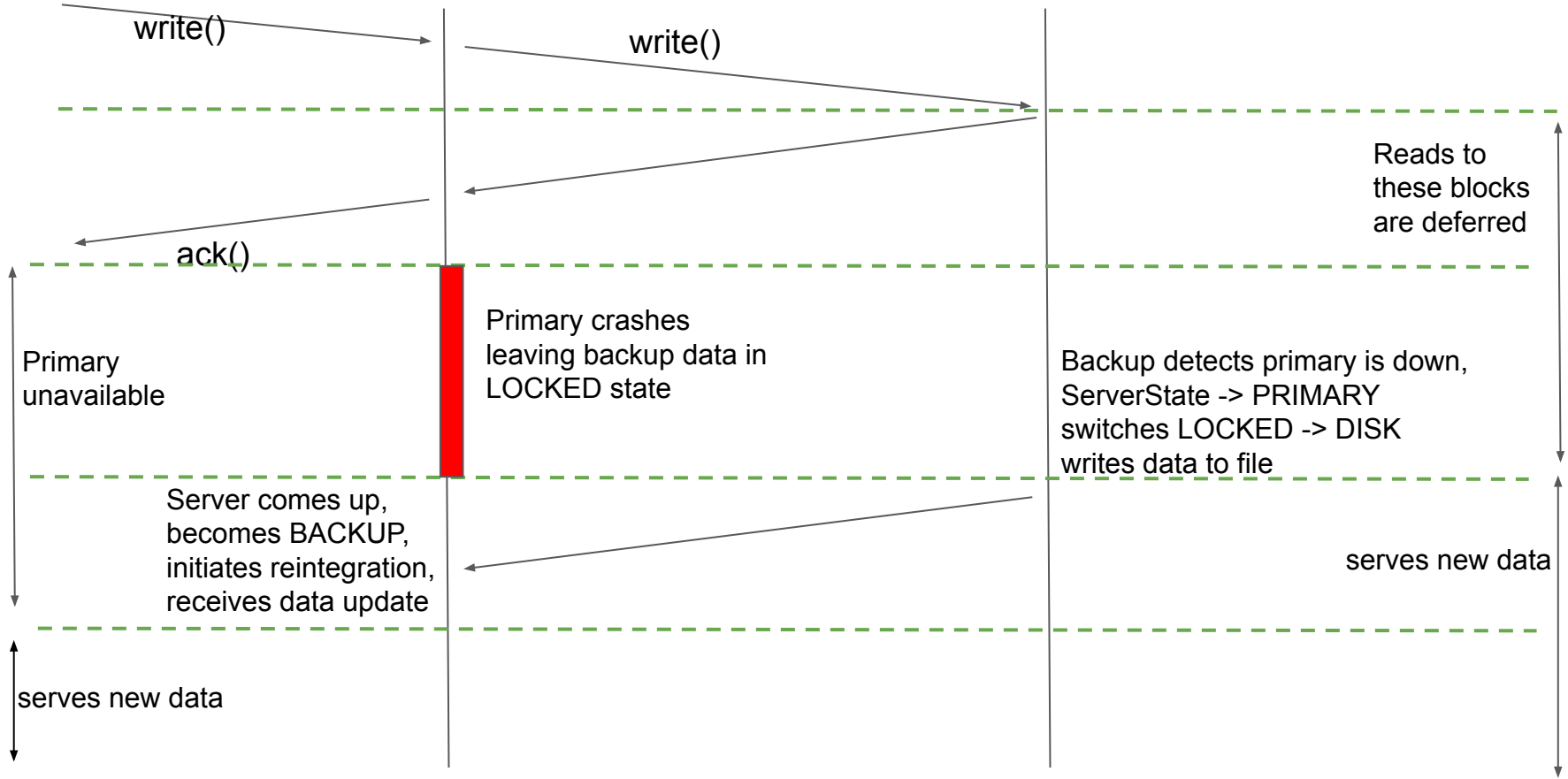
Data design


Primary Server
Backup Server
Client
DISK
s_write()
LOCKED
after s_commit()
after
reintegration
phase 1
MEMORY
DISK

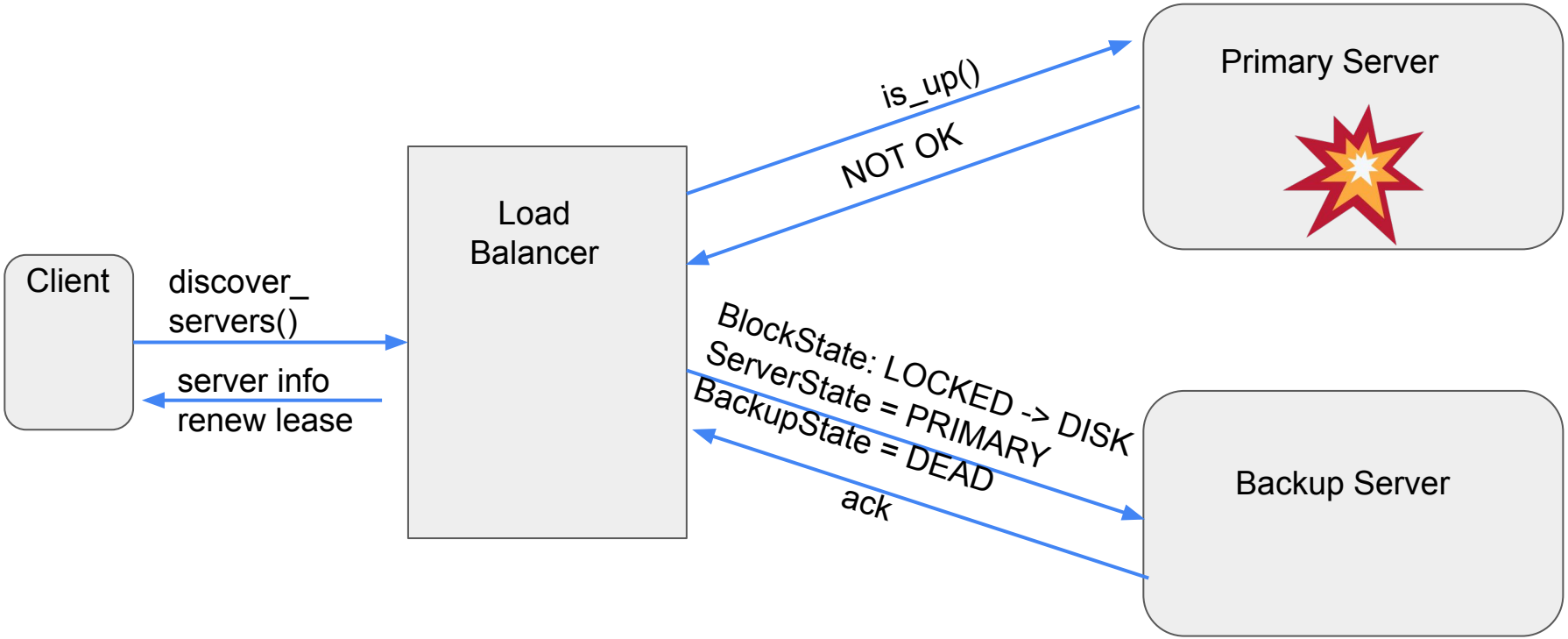
Write protocol



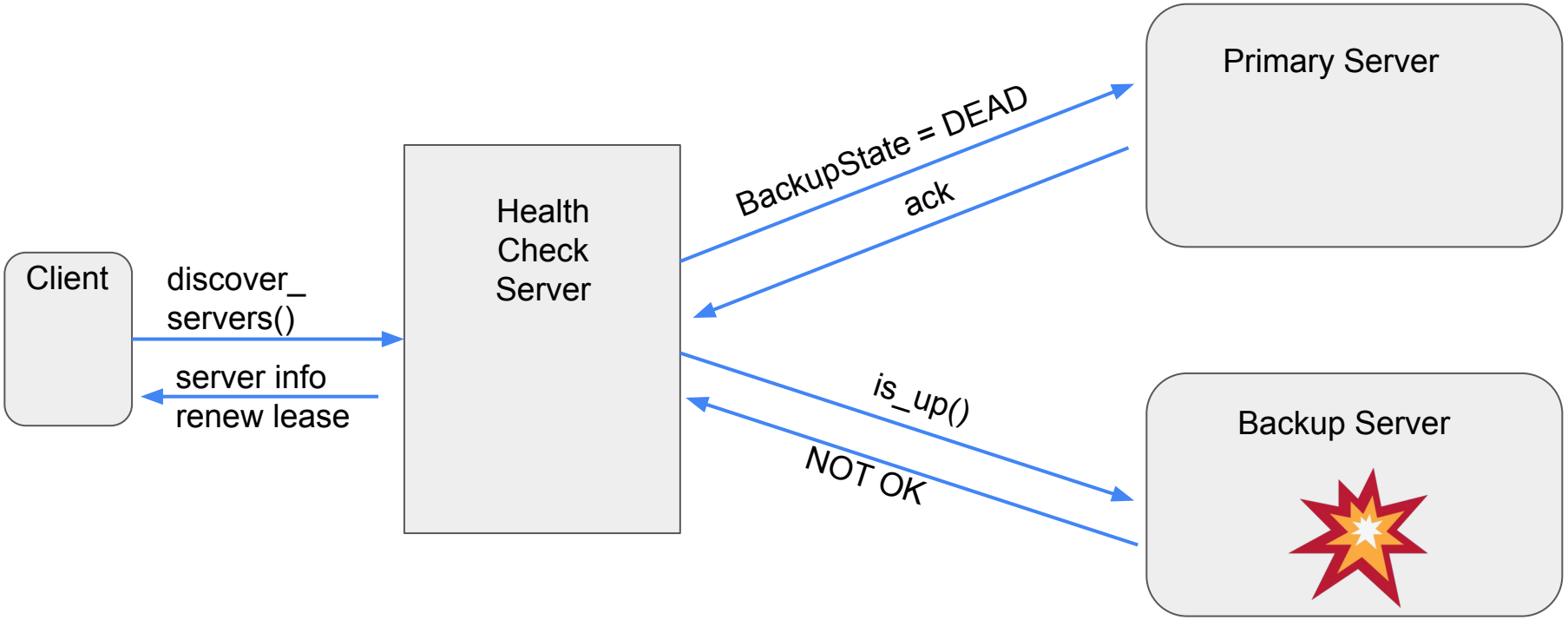
Write Failure Scenario



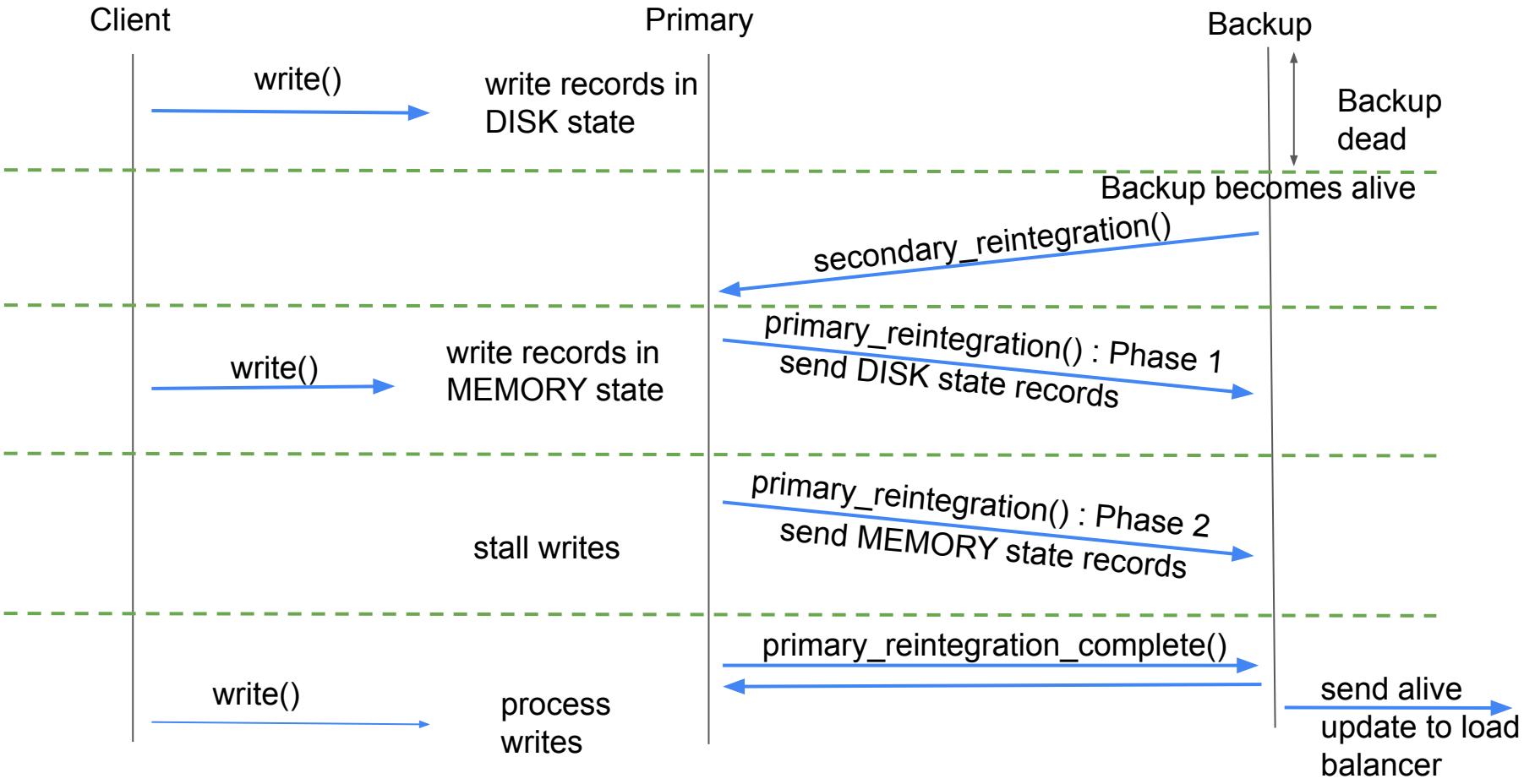
Primary Crash Scenario (transparent to client)



Backup Crash Scenario (transparent to client)



Reintegration (backup comes up)

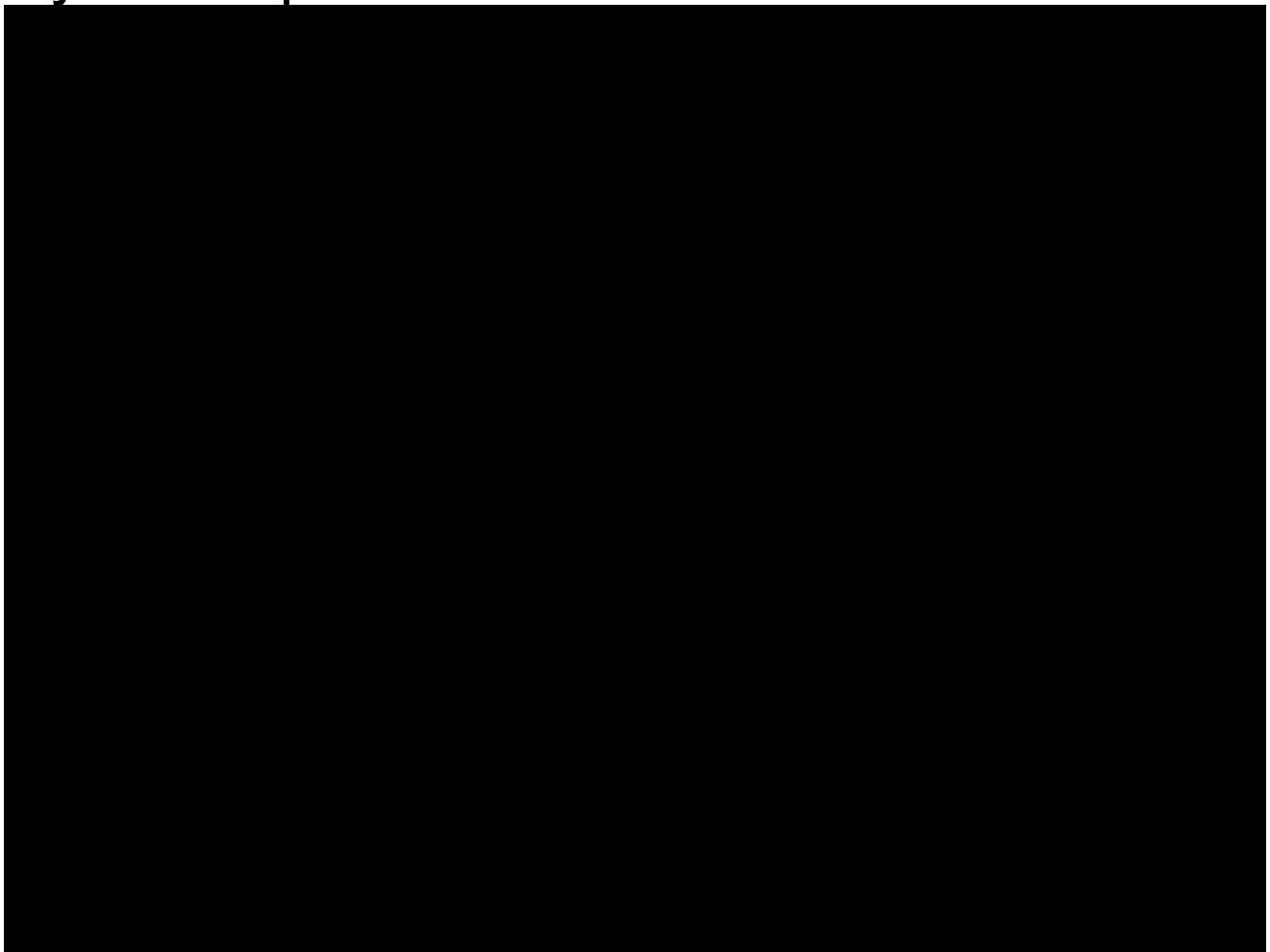


2. Testing & Measurement

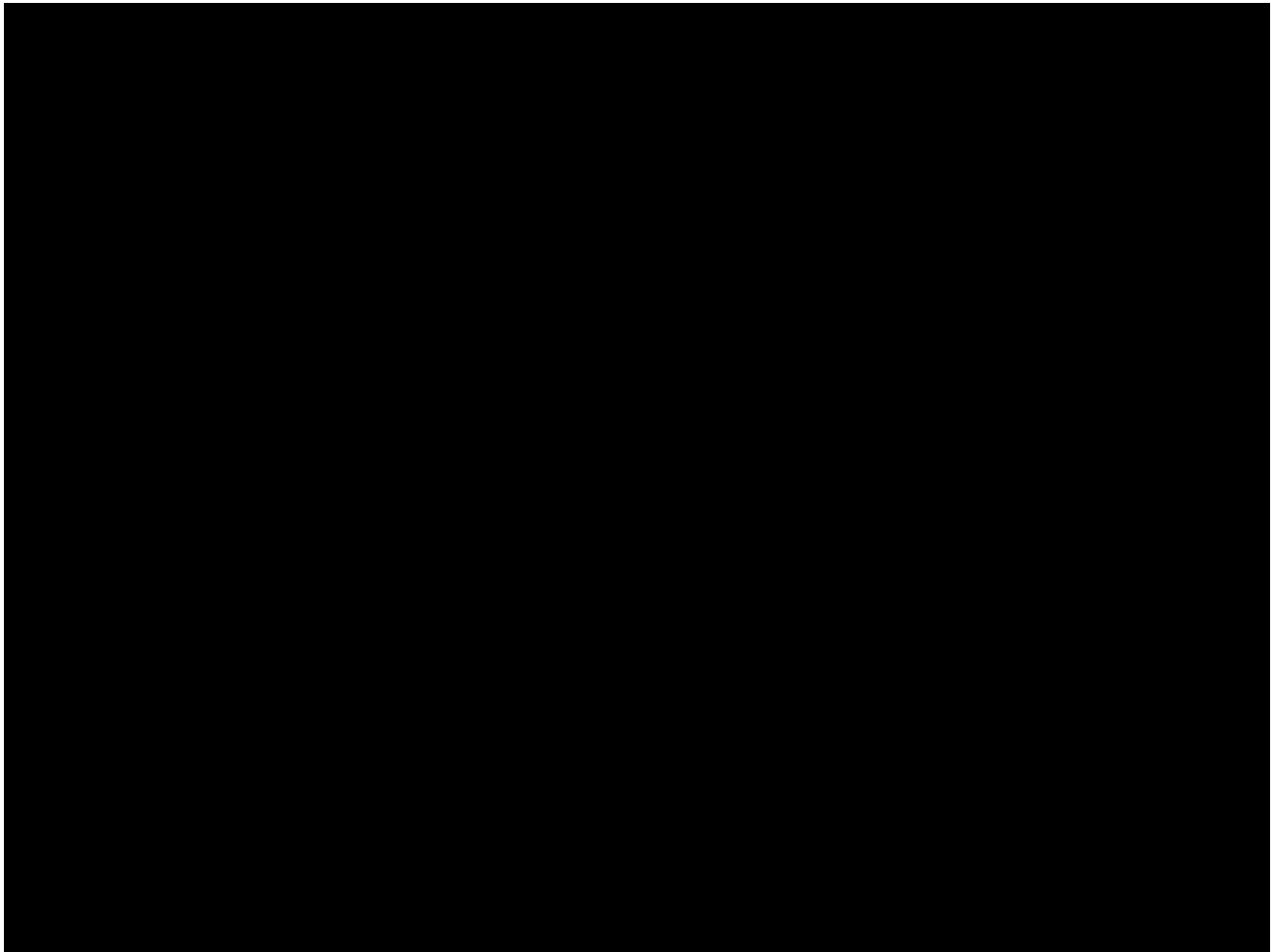
2.1 Correctness

- All reads in clients are sent to all the active servers (to demo if all reads are the same). This feature is driven by configuration.
- The Tests fail if reads are inconsistent
  - Reads are consistent if the test doesn’t halt

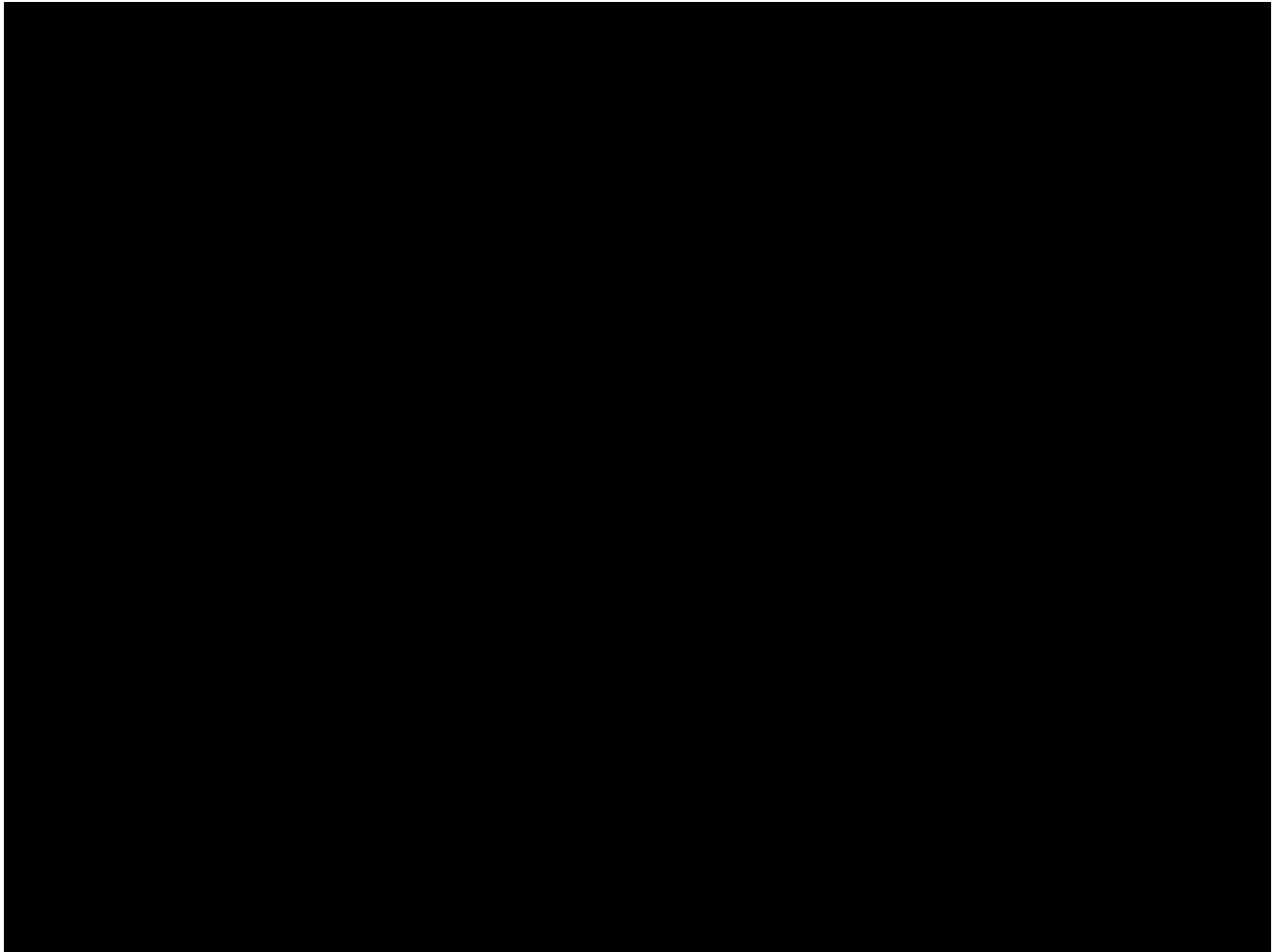
2.2 Primary/Backup Crashes Scenario



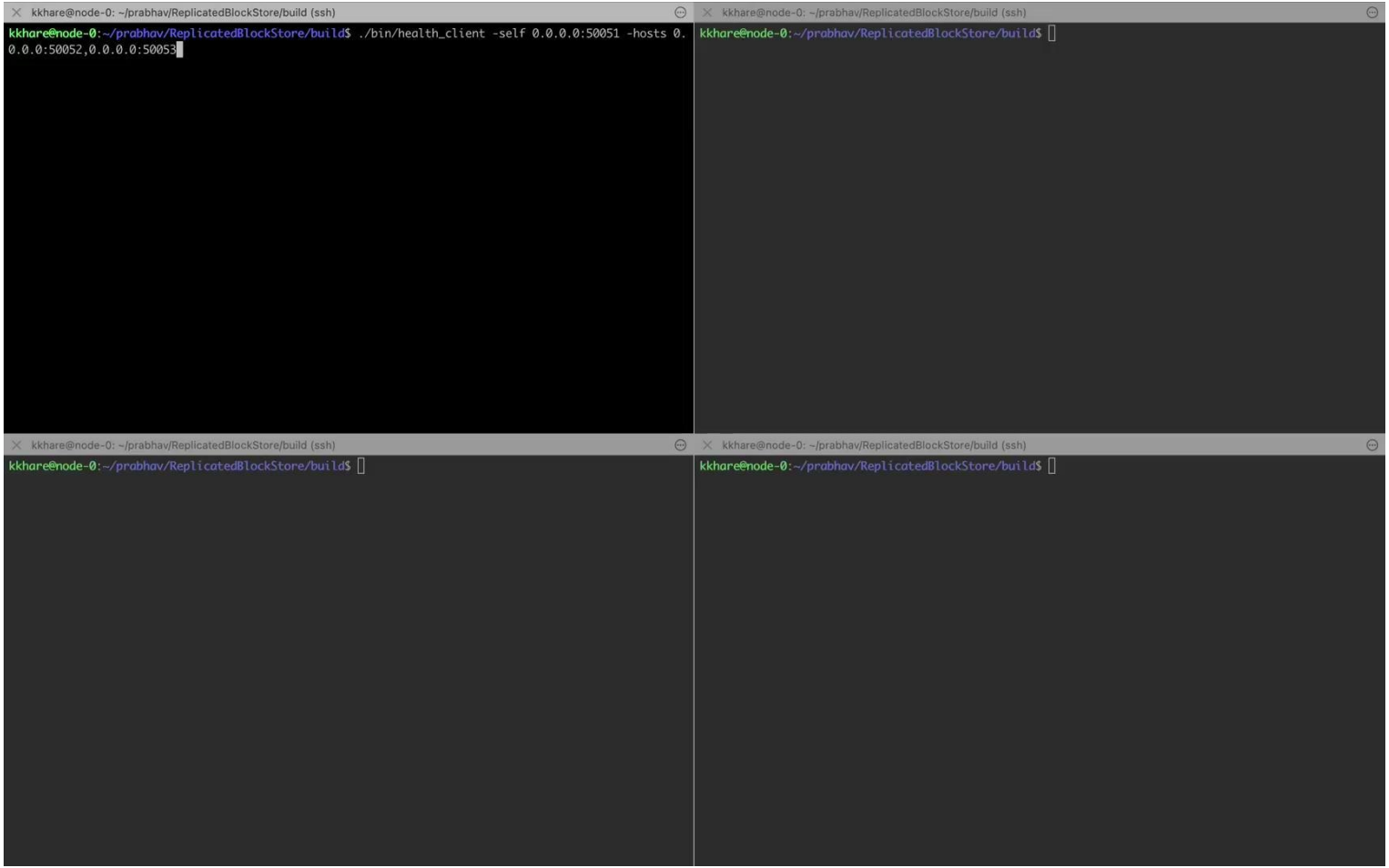
2.2 Primary/Backup Crashes Scenario - unaligned addresses



Primary fails after write

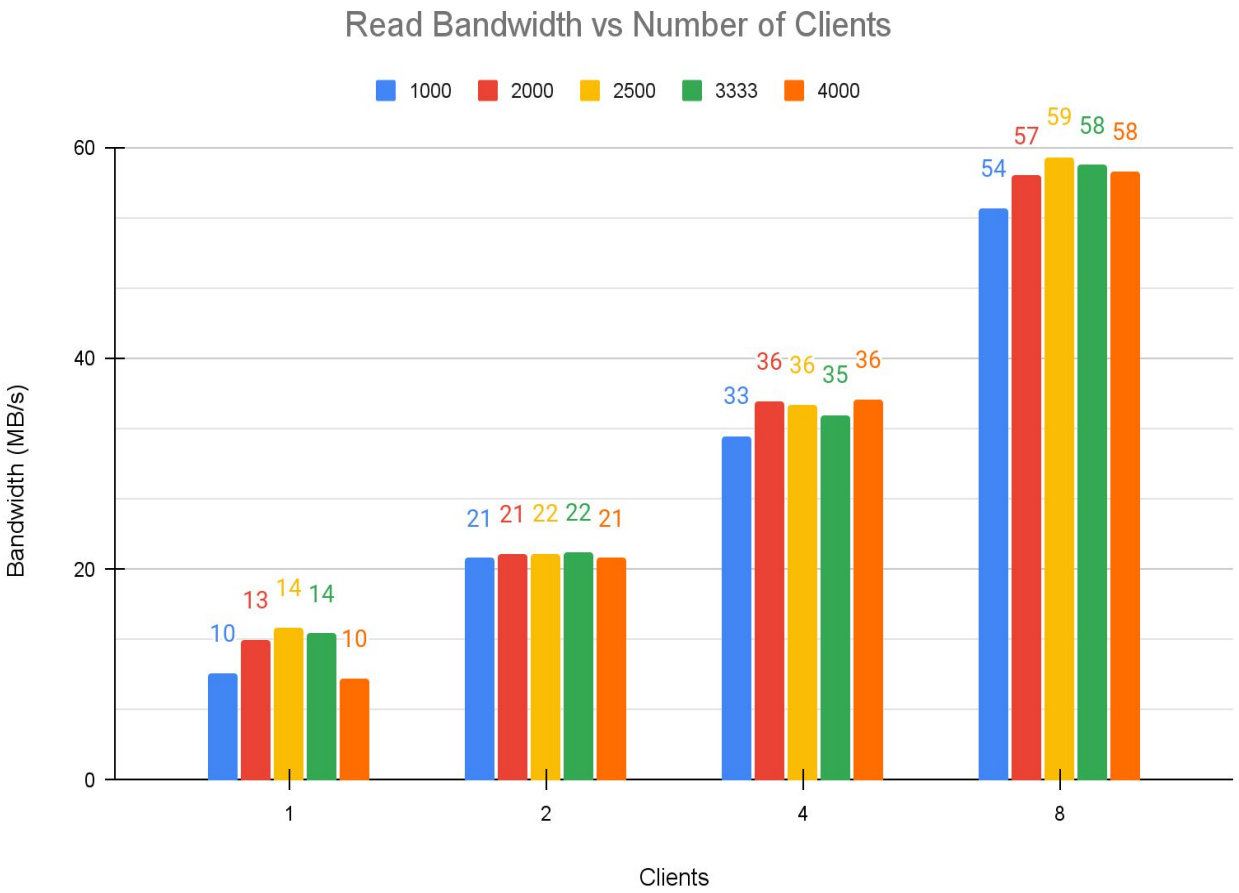


Back up crash during reintegration



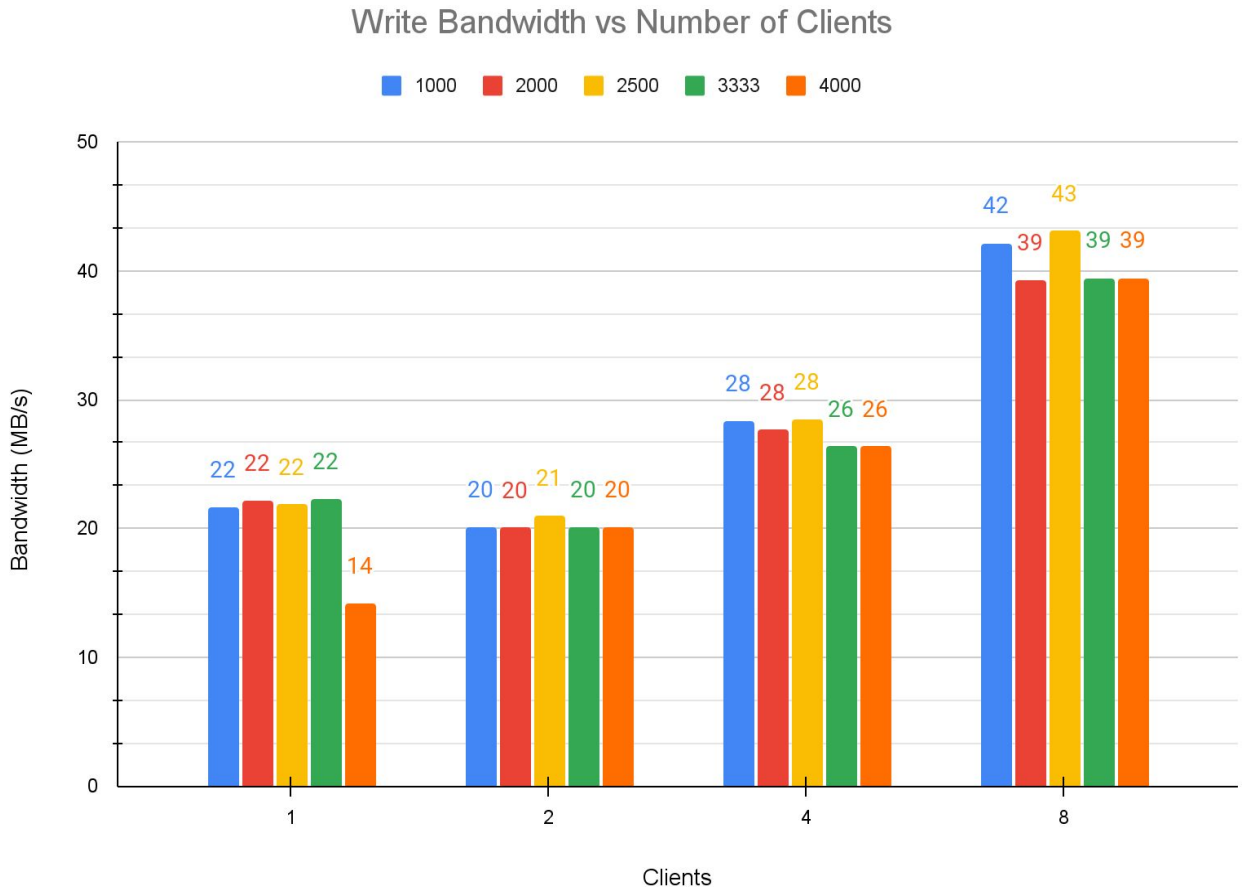
3. Performance

Read Performance



- Total bandwidth increases with increasing clients
- Increasing the total number of reads reduces the bandwidth slightly

Write Performance

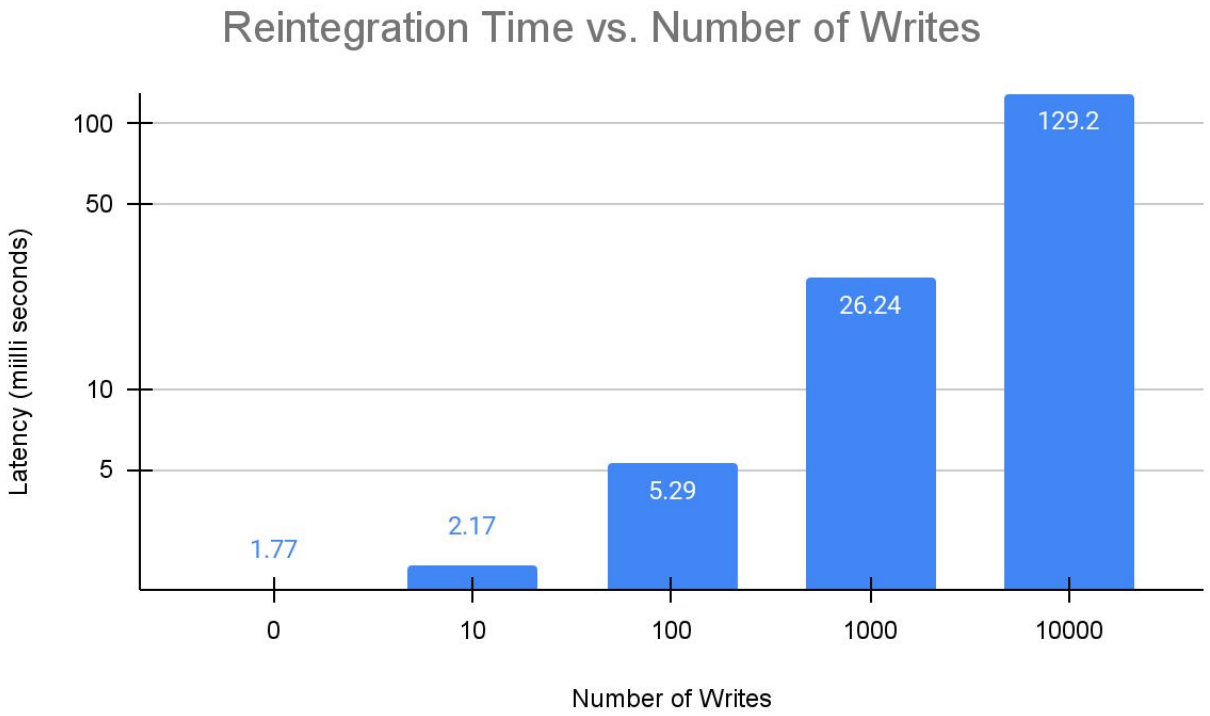


- Total bandwidth increases with increasing clients
- Increasing the total number of writes reduces the total bandwidth slightly

Concurrent Read/Write Performance

- Ready heavy workload has more bandwidth

Reintegration Performance



- Reintegration time increases as number of writes while backup is dead increases

Questions