

Problem 3 (8 points)

Google opted to build a distributed storage system that can leverage the storage resources of compute nodes in a data center, unfortunately, this approach may increase network overhead due to replication. In this question you will estimate the performance of a distributed file system.

To write a file in our distributed file system:

- A client sends the data to one node (primary node). The file is stored in a buffer in memory.
- Once the entire file is received. The primary node writes the data from memory to disk and concurrently replicates the file to two secondary nodes.
- The secondary replicas write the file to disk then acknowledge the operation to the primary.

Only when the primary gets the acknowledgment from the two secondary nodes, the primary will acknowledge the operation to the client. For reliability reasons, Google selects one secondary to be in the same rack as the primary node, and one secondary in a different rack. For reads, the client can read from any of the three nodes. The client is located on a rack that is different from the primary and secondary nodes.

Assume the following hardware characteristics. (Simplification: assume GB and MB are base 2 numbers.)

	Latency	Bandwidth
RAM	100ns	20 GBps
SSD	60 μ s	200 MBps
Network within Rack	70 μ s	1,280 MBps
Network in DC	500 μ s	320 MBps

For the following questions, assume there is only one client in the system. Consider each sub-question separately, i.e., an optimization in one part does not affect the other parts.

- a) What is the response time when a client writes a 1GB file? Show your calculation.
- b) What will be the throughput and response time of writing a 1GB file if the client only waits until the data is stored at the primary and one of the secondaries (not two of the secondaries as in part a)? Show your calculation.