Sami Emre Erdogan
Student ID: 2019280513

Homework 5: GFS

1. The master stores three major types of metadata: the file and chunk namespaces, the mapping from files to chunks, and the locations of each chunk's replicas. While the first two type of data are persisted by master, the locations of each chunk are not persisted in the master side.
Q1: How does the master node get the locations of each chunks at startup?

The master does not store chunk location information persistently. Instead, it asks each chunkserver about its chunks at master startup and whenever a chunkserver joins the cluster.

Q2: What is the benefit of this approach comparing with the approach that the master persists this information?

It allows us to update the master state simply, reliably, and without risking inconsistencies in the event of a master crash.

2. Assume in a cluster of GFS of 1000 servers. Each server has 10 disks with 10TB storage capacity and 100MB/s I/O bandwidth for each disk. The ethernet that connects servers has bandwidth of 1Gbps.
Q1: What is the minimum time required to recovery a node failure (i.e. distribute its replica to other survived server nodes)?

For a failed node we need to look at 1 server node and 1 server node has:
1 server node = 10 x 10TB = 100TB
The node can receive data at the speed of 1Gbps, so to restore the node:
100TB / 1Gb ~= 800 000 seconds ~= 222 hours is roughly needed.

Q2: For quality of service, usually the recovery traffic is throttled. If the bandwidth used for recovery is 100Mbps per machine, what is the roughly time required to recover a failure node?

100TB / 100 mb ~= 2222 hours is roughly needed.

Sami Emre Erdogan
Student ID: 2019280513

Q3: Assume the server node has 10000 hours MTBF. How many server failures are likely to have in a year in this cluster? What is the mean time between node failure in this cluster?

We have 1000 servers with 10000 hours MTBF. One year has 8760 hours. So, in this cluster the mean time between node failure will be roughly 10 hours.
8760/10 =867 failures.
In 1 year, we will get around 876 failures.

Q4: Comparing the time you got from Q2 and Q3, what is the implication of these calculated time to the number of replicas that used in GFS? Read the GFS paper and answer the following questions:

If roughly 2222 hours is needed for recovery then, three replicas are not enough it will take too long but if it takes around 222 hours, three replicas should be good enough to handle.