Homework 4: GFS

Introduction to Big Data Systems course

Due: October 16, 2022 23:59 China time. Late submission results in lower (or even no) scores.

For questions or concerns, contact TA (Huanqi Cao, Mingzhe Zhang) by WeChat. Or send an email to caohq18@mails.tsinghua.edu.cn or zmz21@mails.tsinghua.edu.cn if you could not use WeChat.

Overview

Read the GFS paper and answer the following questions.

Submit a PDF report to Tsinghua web learning, or by email if you can't access web learning.

Part 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?

Q2

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

Part 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 me required to recovery a node failure (i.e. distribute its replica to other survived server nodes)?

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 me required to recover a failure node?

Q3

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

Q4

Comparing the me you got from Q2 and Q3, what is the implication number of replicas that used in GFS?