Summary

GFS is a large data-intensive distributed file system developed by Google for its internal use to meet their storage use and data processing. It is developed using inexpensive commodity hardware. The file system has common characteristics of other distributed file systems such as performance, scalability, reliability, and availability. The developers explained following observations they considered while developing GFS. First, machines fail and it is a common thing. We cannot expect all machines to be up and running all the time. Second, large files are increasingly common in file system and it is not efficient to manage large number of small sized files. Third, random writes are rare and most writes are appended. Fourth, more flexibility is achieved if applications and file system are co-designed.

Key idea

The GFS design consists of a single master node which holds all the metadata of all the files and various chunkservers which holds the actual file data. Files are treated as fixed size chunks. Usual operations such as create, delete, open, close, read and write files are implemented. Other mechanisms provided are snapshot and record append. Corrupt data is identified using checksumming for maintaining data integrity. Fast recovery and master replication is done to make the system highly available.

Strengths

* Using record append, multiple client request to append to files is possible without the requirement of lociking while writing a file. Merging appends from various clients appends is easier. It also ensures atomicity of each clients changes.
* Each file is replicated on different chunkservers. The usual replication factor is 3. This replication increases the reliability of the system.
* The system gets more simple and reliable with lazy garbage collection. The file which is to be deleted is first renamed to a hidden name. During the master’s regular scan, if it founds any hidden name after 3 days, it is deleted.

Weakness

* No support for random writes at all. Although the developers observed that files are mostly appended and fewer times files are written or modified in between, random write is a basic operation in any traditional file system which GFS does not provide.
* Master server introduces single point of failure. A Ddos attack on the master or any sort of failure will make the whole system fail.

Additional thoughts on paper

GFS has been build around considering Google’s workload for data processing using commodity hardware. The decisions made for designing the system are specific to Google’s use case. The paper presents research & expermiemtns of this cluster system and the system performed well on this workloads. The design of GFS has inspired another file system HDFS by hadoop, which is used to manage dig data in similar distributed fashion.