In order to meet the needs of Google's rapidly growing data processing, Google designed and implemented the Google file system. GFS has many same goals with distributed file systems of the past, such as performance, scalability, reliability, and availability. However, its design is also influenced by Google's application load and technical environment.

The nodes of the GFS system can be divided into three roles: master, chunkserver and client. A GFS cluster consists of a single master and multiple chunkservers and is accessed by multiple clients. The master server is the heart of the GFS file system architecture, and it ensures the mapping between chunk and files, the namespace of files, and so on. The files stored in the GFS are divided into fixed-size chunks, and the GFS Master allocates a world-only chunk handle and stores three copies. The client is the interface that GFS provides to the application, which is a set of specialized interfaces that are provided as library files. When the client accesses GFS, it first accesses the master server node, obtains the ChunkServer information that interacts with it, and then directly accesses the ChunkServer to complete the data access work.

The size of chunk is 64MB. This can reduce the network overhead by getting a large chunk at a time for network request. This can also reduce the master overhead because the client can store more information of chunk which reduce the disk fragmentation. This design can lead to hotspot problems when thousands of requests all visit the same chunk at the same time. This can lead to local overload, which can be solved by increasing the number of replicas or get it through other client.

Operation Log is almost the lifeblood of the entire system, because it records the changes to metadata. The system can recover through log, and can also achieve concurrent operations through log. At the same time, in order to ensure the validity of the log, it will also be stored on many different nodes. When the master needs failure recovery, it needs to rely on the log. Firstly, the master will automatically create a checkpoint and then reproduce the log after the checkpoint to ensure that the recovery can be faster.

The paper introduces how the master and chunkserver communicate with each other. The lease mechanism is primarily designed to ease the burden on the master by delegating rights to the master chunk, which serialize the operation while being able to synchronize to other replicas. The master only needs to maintain communication with the master chunk and does not need to communicate with other replicas. In addition this lease is based on the time period and the master chunk can apply for an extension, and the master can control the cancellation and grant.

The Google file system is built on cheap servers. It regards server failure as a normal phenomenon, and automatically tolerates fault through software, which greatly reduces the cost of the system while ensuring the reliability and availability of the system. GFS has successfully realized storage needs, both as a storage platform for research and development and as a data processing platform for production systems.