The Google File System

The paper discusses the google file system (GFS) interface extension which is designed to support the distributed applications. Later in the paper the performance of the system is compared with the micro-benchmarks and real world usage standards. A large number of clients can access the google file system easily and at high performance on cheap commodity hardware.

Four major features of Google file system makes it resilient. One, it’s capability of handling component failures that are caused by OS bugs, human errors, networking etc. Second, the system can handle multi-GB files. Third, it handles the situation where files are mutated by appending new data by multiple clients. Lastly, co-designing the applications and the file system API. The paper dives deep into the design decision and the architecture of GFS. Following are some of the points discussed in the paper.

GFS organizes files hierarchically into directories and identifies them by path names and supports operations like snapshots and record append in addition to basic operations like create, delete, open, close, read and write files.

With GFS clusters, there is a single master that maintains all metadata, and multiple chunkservers that can be accessed by multiple clients. The master manages communication with chunkservers and collects data from them. It enables the clients to directly establish the connection with the chunkserver by routing the client to the right chunkserver.

GFS chunk size is 64GB to minimize client-master interaction, network overhead, and metadata size on the master. Having larger chunk sizes has disadvantages since the chunkservers can become hot spots for clients.

The GFS keeps track of failed chunkservers and record appends by regular handshakes between the master and chunkservers. Data transfers over TCP connections are pipelined to reduce data transfer latency.

Through the decoupling of data flow and network flow, atomic record appending, and snapshots, the GFS is designed to minimize the master's involvement in all operations which are achieved by managing namespaces, chunks, and coordinating system-wide activities.

GFS doesn't delete files when they're deleted. It renames the files to a hidden name and stores them for a configurable period of time, and then the master deletes all replicas and file metadata unless they're reused or renamed.

Last but not least, although the principles are specific to the Google environment, they can be applied to data processing tasks of similar magnitude in other contexts, such as research and data processing.