Google designed and implemented a scalable distributed file system for large distributed data-intensive applications called the Google File System (GFS) in that demands of Google’s data processing needs are increasing day by day. Though there are already quite a few previous distributed file systems like Microsoft’s Windows Distributed File System (DFS) and HDFS by Apache Software Foundation, GFS has a majority of different and innovative points compared with them. First, GFS features constant monitoring, error detection, fault tolerance and automatic recovery, since problems could emerge because of human or machine errors. Second, design assumptions and parameters such as I/O operation and block size are included since nowadays people are working with large data sets of many TBS, even PBS which are still growing where it is unwise to manage billions of KB-sized files. Third, appending operation on files is optimized and atomicity guaranteed as most files now are modified by appending instead of overwriting existing information. Fourth, the applications and the file system API are co-designed so that the system’s more flexible. At last, the system is now extensively deployed within Google where the largest ones have over 1000 storage nodes which are accessed by thousands of clients continuously.