Brian Farrell

GFS Paper Review

Category:

This paper is about the Google File System, a distributed file and storage system designed and implemented by Google. This was a new system at the time this paper was written.

Context:

The time this was written in, 2003, was a time where data usage at companies like google was increasing exponentially. Therefore, a new system needed to be designed to handle this data and potential faults. The author drew sources from multiple different fields within computer science, including operating systems, programming languages, computer systems, and distributed file systems.

Correctness:

The system successfully met its fault tolerance in both experiments with losing chunk servers. The first experiment with a single chunk server being down led to the system being able to restore the chunks without compromising the other tasks of the file system. The second experiment was also a success because two chunk servers were restored within two minutes.

Contributions:

The main contribution of this paper is the new file system implemented at Google. The system is an improvement on previous iterations as outlines in the paper. The system is able to run very efficiently and built with enough redundancy to fix the faults that are guaranteed to occur. The system was also able to achieve high performance by reducing strict consistency requirements.

Clarity:

This paper is very well written. The abstract, introduction, and conclusion give a good summary of the file system. The sections are broken down nicely to make the paper more readable and digestible.