The Google File System

Ghemawat, Gobioff, and Leung. *The Google File System* November 24,2013

Charlie Ropes
November 24, 2013

Main Idea

The Main Idea of this paper is to explain how the Google File System is designed, set up, and how it works. The Google File Systems is suppose to distribute large data-intensive applications efficiently and effectively.

Implemented

- Using Cheap Hardware
  - Breaks easily but is not costly to replace.
- Using Chunk Servers
  - Chunks of data are stored on different servers.
  - Chunks are passed between servers to get where they need to go.
  - Chunks are replicated on different servers for higher availability.

Opinion

The idea and implementation of this system help solve a lot of problems. A system like this will allow for easier distribution of large files. The way they implement it is also cost effective and prevents many types of system failures. If multiple servers go down they can easily be replaced and rebooted.

Advantages and Disadvantages

Advantages

- High Availability
  - Restarting is quick so downtime is minimal.
- Chunk Replication
  - Chunks have replicas stored on different servers.

Disadvantages

- Cheap Hardware
  - Equipment will break more frequently.

Real World Use

A system like this could be used in large companies that have to store mass amounts of data and retrieve it quickly. A system like this would be cheap to implement and maintain. This system also has many precautions to make sure the system stays up in running with little to no downtime which is a must I most investment companies.