DANIEL CODY
PROFESSOR LABOUSEUR
THE GOOGLE FILE SYSTEM
DECEMBER 8TH, 2014

Ghemawat, Sanjay, Howard Gobioff, and Shun-Tak Leung. "The Google File System." (n.d.): n. pag. 19 Oct. 2003. Web.

The Google File System

- Scalable, for large scale applications and many users
- Designed for data-intensive applications
- Inexpensive commodity hardware, so it can easily be replaced if there is an error
- Fault Tolerance, in the case of an error

The Google File System

- Contains thousands of inexpensive storage machines, which is good for scalability and replacing hardware
- Uses a single Master to control the file system
- Stores Data in chunks, 64MB, which is good for quick access to information
- Uses Writes & Mutations, which minimizes the computing time

The Google File System

- Scalable, large data software is way of the future, so this file system is needed
- Cheaper, easy to replace hardware is best, because hardware failure is imminent
- Master is extremely interesting as the controller of the file system

A Comparison of Approaches to Large-Scale Data Analysis

- Map & Reduce is similarly efficient, but not as scalable
- No Indexing relies too much on the programmer
- Strong fault tolerance provides better reliability

Advantages & Disadvantages

- GFS provides more scalability, which is important for a huge company
- MR has very strong fault tolerance, which insures correct data and always being available
- GFS relies too much on the master to handle situations, which could cause errors if the master is not coded properly
- MR relies much more on the programmer, which can always cause errors