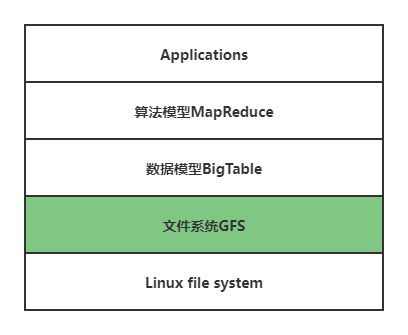
Google File System读后感

1. **GFS介绍**

Google搜索引擎的支撑架构图如下，本文将思考并分析其中的GFS部分：



GFS是google整个分布系统的基础，其他存储系统将直接或间接地构建在其之上。它使用系统来监测错误并自动恢复，是一个可扩展的分布式文件系统，能够用于大型分布式数据相关的应用，且能够运行在较普遍的商用硬件上。

* 1. **背景介绍**

GFS的诞生是基于以下的四个问题：组件的失效比异常更常见；传统标准的文件量巨大，管理困难；多数文件的变化是增加而非重写；统一设计应用和文件系统API对提升灵活度有好处。

1. **设计原理**
   1. **如何保存文件**

基于操作系统中文件系统保存文件的原理，当我们需要保存文件时，需要往硬盘的相应位置存储文件的名称、创建时间、大小、索引等信息，通过索引记录文件在硬盘上的位置来寻找文件。此时，我们使用的结构为Block，传统意义上一个Block为1024Byte。

* 1. **如何保存大文件**

对于大文件来说，Block结构不太适合使用。因为存储一个文件需要的Block过多，GFS中使用新的数据结构Chunk，一个Chunk相当于64MB，这样大大减少了Metadata的数据量。但这样存储小文件时就非常浪费空间。

* 1. **如何保存超大文件**

那么对于大于超大文件来说，仅使用Chunk也无法满足需求。因此，GFS中使用主从结构。通过Master保存所有原始信息，并在其中建立索引，索引信息指向各个Chunk Server，并记录相关的偏移量。而这个结构的缺点则是，每次Chunk Server进行变化都需要传输到Master中进行记录，出现浪费。

* 1. **如何减少Master存储的数据和流量**

首先，GFS对Master + Chunk Server的结构进行优化，它将index从Master中转移到了Chunk Server中，仅在Master中记录文件所处的Chunk但不记录偏移量的索引值。这样，只用数据出现Chunk的变化时才需要修改Master中的值。

* 1. **数据损坏与修复**

GFS通过校验和备份的方式来防止并检测数据是否损坏。首先，Chunk由许多的block组成，GFS通过在每个block保存一个checksum值，每次进行文件读取时都要对block中的checksum进行你校验。其次，同一个Chunk将被保存在多个Chunk Server中，当其中一个Server挂掉，可以通过其他的进行数据恢复。

而当某个Chunk出现问题时，Master会通过定位其他Server中的副本来进行恢复。

* 1. **如何恢复数据**

那么，如果是Chunk Server本身出现了掉线等问题，将如何恢复数据呢？在GFS中，通过定时“握手”的方式来判断Server是否出现掉线等问题。如果长时间Server都不进行回复，其他Server会首先呼叫此Server，若失败则启动修复进程。

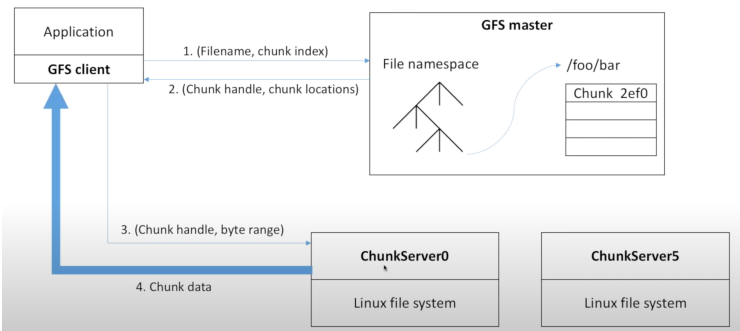
在GFS中，Chunk的修改记录将会被记录在操作日志中。这个日志是元数据的唯一持久化及记录，也作为定义同步操作顺序的逻辑时间基线。修复进程会根据此记录按照Chunk存活数据量递增的形式进行Chunk修复，因此存活最少的将首先修复。

* 1. **如何对应热点数据**

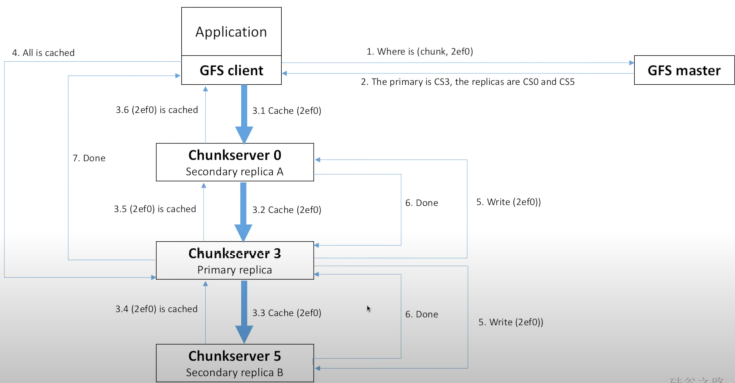
在Master中有一个热点平衡线程，它会记录每个Chunk的访问热度以及Server剩余空间和带宽。因此当某Chunk达到繁忙时，可以将其中的数据进行多次复制，通过多个Server的方式减缓繁忙问题。

* 1. **如何读写文件**

基于论文中的图进行理解，读文件的流程如下：

****

基于论文中的图进行理解，写文件的流程如下：



1. **总结**

GFS是Google搜索引擎的基础，它实现了在普通硬件上进行大规模数据吞吐、分布式的文件存储系统。但它不具有完善的错误处理机制，而将错误解决留给了客户端以及更高层的架构，主要实现基础功能，保持系统的简易、快捷的程度。