**谷歌GFS读后感**

GFS作为谷歌三驾马车之首，是谷歌对计算机——大数据领域重大贡献的开端，也是谷歌自chrome之后迈向伟大公司的巨大一步。

GFS的全名是Google File System，即谷歌文件管理系统。它提供了适应于当代大数据环境的数据管理方式，又称为：分布式文件管理系统，是现在大数据蓬勃发展的基石。

不同于传统的文件管理方式，GFS成功的实现了数百TB的数据管理。其核心设计思路为：

1. 组件失效被认定为常态事件，而不是意外事件。而处理方式则采用监控、侦测、自动恢复的方式自检。
2. 增加标准衡量的基础单位，将数据基数增大到GB和TB，而不是KB，例如：在设计I/O流和Block的尺寸上。
3. 对于文件的修改 ，采用追加数据方式，而并不覆盖原有的数据，对于已经写入的数据基本只进行读操作。
4. 应用程序和文件系统API协同设计，已提高灵活性。

采用上述的设计思路，Google当时已经实现了超过300BT数据的存储与应用。在现在的情况看来，更是大数据合理、可行的管理方式。

而其中的系统交互又以最小化所有操作和节点交互为原则，使得对数据的管理容错率更高。再加上Master与chunk的存储设计使伸缩性大大提升，才能突破传统意义上数据的树状存储方式，实现数据的集群管理，为互联网对大数据的处理打下基石。也是hadoop时代的重要基础。

我想这也是IT行业发展如火如荼，一直满足摩尔定律的关键。由雅虎起初定义的发展模式——开源，使得谷歌的三驾马车将人们带进了大数据时代，使现在的5G和人工智能成为可能。