

# 读后感

---

19301086 高时玉

## 1. Google MapReduce

---

MapReduce是一种编程模型，用于大规模数据集的并行运算。概念"Map"和"Reduce"，是它们的主要思想，都是从函数式编程语言里借来的，还有从矢量编程语言里借来的特性。它极大地方便了编程人员在不会分布式并行编程的情况下，将自己的程序运行在分布式系统上。MapReduce最早是由Google公司研究提出的一种面向大规模数据处理的并行计算模型和方法。MapReduce的推出给大数据并行处理带来了巨大的革命性影响，使其已经成为事实上的大数据处理的工业标准。总的来讲，Google MapReduce 所执行的分布式计算会以一组键值对作为输入，输出另一组键值对，用户则通过编写 Map 函数和 Reduce 函数来指定所要进行的计算。MapReduce 编程模型的原理是：利用一个输入 key/value pair 集合来产生一个输出的 key/value pair 集合。MapReduce 库的用户用两个函数表达这个计算：Map 和 Reduce。用户自定义的 Map 函数接受一个输入的 key/value pair 值，然后产生一个中间 key/value pair 值的集合。MapReduce 库把所有具有相同中间 key 值 I 的中间 value 值集合在一起后传递给 reduce 函数。用户自定义的 Reduce 函数接受一个中间 key 的值 I 和相关的一个 value 值的集合。Reduce 函数合并这些 value 值，形成一个较小的 value 值的集合。一般的，每次 Reduce 函数调用只产生 0 或 1 个输出 value 值。通常我们通过一个迭代器把中间 value 值提供给 Reduce 函数，这样我们就可以处理无法全部放入内存中的大量的 value 值的集合。MapReduce 模型可以有多种不同的实现方式。如何正确选择取决于具体的环境。例如，一种实现方式适用于小型的共享内存方式的机器，另外一种实现方式则适用于大型 NUMA 架构的多处理器的主机，而有的实现方式更适合大型的网络连接集群。

## 2. Google Bigtable

---

Bigtable 是一种压缩的、高性能的、高可扩展性的，基于 Google 文件系统（Google File System，GFS）的数据存储系统，用于存储大规模结构化数据，适用于云端计算。Bigtable 的设计是为了能可靠地处理 PB 级的海量数据，使其能够部署在千台机器上。Bigtable 借鉴了 parallel databases（并行数据库）和 main-memory databases(内存数据库)的一些特性，但是提供了一个完全不同接口。Google 的很多项目使用 Bigtable 存储数据，包括 Web 索引、Google Earth、Google Finance。这些应用对 Bigtable 提出的要求差异非常大，无论是在数据量上（从 URL 到网页到卫星图像）还是在响应速度上（从后端的批量处理到实时数据服务）。尽管应用需求差异很大，但是，针对 Google 的这些产品，Bigtable 还是成功的提供了一个灵活的、高性能的解决方案。Bigtable系统实现了数据的分布式管理和结构的组织。Bigtable包含了3个主要的组件：链接到每个客户的库，一个Master服务器和多个Tablet服务器。根据负载情况的变化，Bigtable可以动态的向集群中添加或者删除Tablet服务器。

Bigtable重要的贡献是证明了在分布式的系统中，针对超大规模的数据量，使用排序大表的来设计数据库是可行的。这直接带动了LSM Tree的流行，在后来的HBase，LevelDB中都使用了这种方式处理数据。另外，Bigtable系统中Chubby的使用，还告诉工业界分布式协调组件的重要性，这也引导了Zookeeper的设计实现，而其仍然是今天的分布式系统中重要的组件。

## 3. Google File System

---

GFS是一个可扩展的分布式文件系统，用于大型的、分布式的、对大量数据进行访问的应用。它运行于廉价的普通硬件上，并提供容错功能。它可以给大量的用户提供总体性能较高的服务。GFS系统主要存储文件系统两类数据：文件元数据：包括file system namespace（目录结构、文件元数据等）、文件存储位置信息等；文件数据：文件在磁盘上的存储格式。

GFS具备以下特点：高度扩展性，单个GFS集群文件系统可扩展至几千个节点，数PB的容量；高可靠性，GFS运行在普通x86服务器上，但GFS通过多副本等机制保证可靠性；满足POSIX语义，应用程序无需任何修改即可使用GFS。Google 文件系统展示了一个使用普通硬件支持大规模数据处理的系统的特质。虽然一些设计要点都是针对我们的特殊的需要定制的，但是还是有很多特性适用于类似规模的和成本的数据处理任务。

## 总结

---

Map Reduce和Big Table，它们则是基于GFS研发的。这三种基本的核心技术构建了一个完整的分布式计算体系结构。与map reduce一样，它是一种大规模数据集并行操作的编程模型。Map和reduce (reduction)是其主要思想和概念，借鉴了函数式编程语言，具有矢量编程语言的特点。它极大地便利了程序员在分布式系统上运行他们的程序，而无需进行分布式并行编程。目前，软件实现是指定一个映射函数将一组键值对映射为一组新的键值对，指定一个并发reduce函数，以确保每个映射的键值对共享相同的键组。大表是一种分布式数据存储系统，是一种处理海量数据的非关系型数据库。Big able是一个非关系数据库。它是一个稀疏、分布式和持久的多维排序映射。它适用于廉价的设备、大规模的海量数据以及分布式和并发的数据处理。它易于扩展，效率高，支持动态扩展。