BigTable读后感

首先，我阅读了摘要的部分，大致地了解了BigTable究竟是什么，正如摘要中所描述的，Bigtable 是一个分布式的结构化数据存储系统，它被设计用来处理海量数据，而这些数据通常是分布在数千台普通服务器上的 PB 级的数据。在许多项目中Google都运用到了BigTable来进行存储，尽管项目需求的差异度很大，但是BigTable仍然能提供灵活的解决方案。

在第一节介绍中，大致能够更深入了解什么是BigTable，知道BigTable的特点，那就是对 BigTable 而言，数据是没有格式的，用数据库领域的术语说，就是数据没有 Schema，用户自己去定义 Schema等等。同时还能知道整篇论文各个章节的内容。

在第二节中，作者主要描述了关于数据模型更多细节方面的内容，主要是数据模型。作者以存储海量的网页及相关信息为例，指出这些数据可以用于很多不同的项目，这个特殊的表可被称为为 Webtable。在 Webtable 里，使用URL作为行 关键字，使用网页的某些属性作为列名，网页的内容存在“contents:”列中， 并用获取该网页的时间戳作为标识。这里就提到了BigTable中的行列以及时间戳的定义，通过例子的描述能使得数据结构的概念更易于理解。

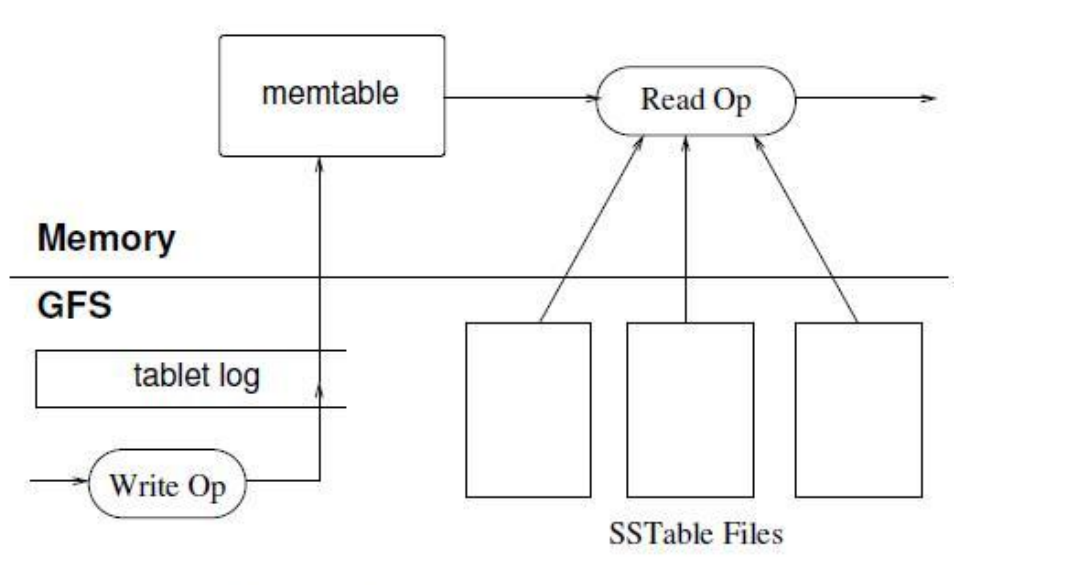
在第三节中，作者概要地介绍了客户端所能使用的API。

在第四节中，作者简要介绍了BigTable底层使用的Google的基本框架，也就是BigTable的构建。。BigTable 使用 Google 的

分布式文件系统(GFS)存储日志 文件和数据文件。BigTable 内部存储数据的文件是 Google SSTable 格式的。BigTable 还依赖一个高可用的、序列化的分布式锁服务组件，叫做Chubby。

通过Chubby的使用，可以确保在任何给定的时间内最多只有一个活动的 Master 副本；存储 BigTable 数据 的自引导指令的位置以及查找 Tablet 服务器。

在第五节中，作者描述了BigTable的关键实现部分。一个可以连接到每个用户的库文件，一个主要服务器控制其他，许多目录服务器来分配文件的储存。主服务器负责把目录分配到目录服务器上，检测目录服务器的改变，平衡目录服务器负载，对谷歌文件系统进行垃圾收集，还有控制不同列族等内容的改变。首先是介绍了Tablet的位置，使用一个三层的、类似Ｂ+树的结构存储 Tablet 的位置信息。然后是Tablet的分配，在任何一个时刻，一个 Tablet 只能分配给一个 Tablet 服务器。Master 服务器 记录了当前有哪些活跃的 Tablet 服务器、哪些 Tablet 分配给了哪些 Tablet 服 务器、哪些 Tablet 还没有被分配。然后是通过该图片形象地指出了Tablet的服务。BigTable具有以下特点全球大规模海量数据；各地的几千台计算机同时进行都可以，效率极高；没有太大限制，扩展空间高；微机都可以使用；适合存取，不适合编辑；以及不适用于传统关系型数据库；



在第六节中作者提到了Google团队为了提高BigTable的性能采用的一些精细化方案，包含了局部性群组、压缩、通过缓存提高读操作性能、Bloom过滤器、Commit日志实现、Tablet恢复、利用不变性等解决方案。

第七节第八节分别展现了使用BigTable的性能数据以及Google内部使用BigTable的例子，如论文中所提到的，随着系统中的 Tablet 服务器从 1 台增加到 500 台，系统的整体吞吐量激增，增长的倍率超过了 100倍，足见BigTable系统的功能之强大。在实际使用中Google Analytics，Google Earth等都使用了BigTable系统，可见其实用性同样很强。

第九节中作者记录下了在设计和后期开发过程中所积累得到的一些经验与教训，在第十节中列出了相关的研究性工作，以及十一节中罗列出了一些结论。