之前下载了厦门大学 林子雨老师翻译的中文版pdf，包含

1 概述；

2 进程模型；

3 并行体系结构；进程和内存协调；

4 关系查询处理器；

5存储管理；

6事物；并发控制和恢复；

7 共享组件；

8 结束语

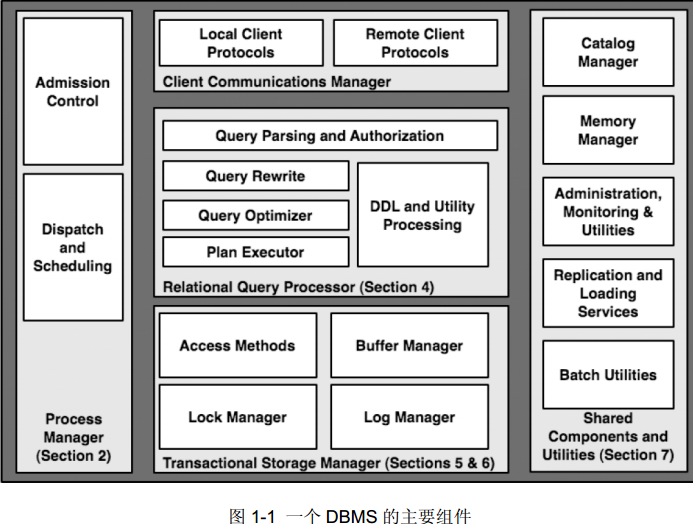
教科书中主要讲的是相关算法和数据库相关概念，但是关于如何让一个数据库工作的系统设计问题却鲜有资料。本文从数据库设计的一些准则包括处理模型，并行架构，存储系统设计，事物处理系统，查询处理及优化机构以及具有代表性的共享组件和应用。

#######################################

一：概述

应用数据库群体较小，市场只能支撑几个高水平的竞争者，因此只有一小部分成功的数据库产品存在。从事数据库设计和应用的人们彼此联系紧密，他们往往来自于同一所学校，研究同样的知名项目，然后合作开发几个相同的产品。另一方面，数据库管理系统的教学领域往往忽视对体系架构问题的讲解。数据库教材一直关注那些易于教学、研究和考试的算法和理论知识点，没有从应用角度对数据库架构有一个全局的讲解。总而言之，关于如何构建一个数据库方面的知识，并不是保密的，可是，它并没有被系统地写下来并供人们讨论交流

下一部分开始具体讲，一个查询语句的具体处理过程：



这个在《MySQL性能调优与架构设计》中也有几个具体的流程图，步骤也差不多；具体到哪一步要干嘛，经历的过程有什么等等

这在开篇就讲到了数据库的整体架构，然后介绍各种切实可行的方案，首先是 单处理机，然后是多种并行处理架构。关于核心服务器系统架构的讨论，适用于其他系统，但是很多是在DBMS中得到应用的。我们从一个简单的查询的角度开始，重点关注查询处理器。之后研究存储架构和事物管理设计。

具体查询过程参考另外一篇文章吧！另开一篇《MySQL Server 系统架构：逻辑模块组成和各模块工作的配合》

###################################

第二章：进程模型

对比了几种进程模型：进程，线程，轻量级线程。具体每一种的区别操作系统原理里都有讲的。或者再参照我写的另外专门的对比**《进程 线程 轻量级线程》**

然后根据每个DBMS工作者采用的不同进程模型分为从简单到复杂的三种：每个DBMS工作者采用一个进程；每个DBMS工作者采用一个线程；进程池；尽管被简化了但是商业DBMS一般采用这三种

然后就是这三种方式的原理对比和几种商业数据库各自采用情况看**《DBMS进程线程采用对比》：**

每个DBMS工作者采用一个进程： DB2在不支持线程的系统采用这种方式，支持线程的系统采用每一个工作者采用一个线程，这也是Oracle的进程模型默认设计。Oracle也支持上文的进程池。PostgreSQL在所有的系统都只运行每个DBMS工作者拥有一个进程模型。

每个DBMS工作者采用一个线程：这种还分为两种：1**每个DBMS工作者拥有一个系统线程**：IBM DB2运行在良好的线程支持系统上采用这种设计，MySQL也是采用这种设计；2**每个DBMS工作者采用一个DBMS线程**：这种的DBMS工作者调度通过系统进程或者线程去调度轻量级线程实现的。规避了一些系统调度的一些问题但是开发成本高，没有良好的开发工具，需要运营商长期维护。

##########################################

第三章：**并行架构：进程和内存协调**

**这一章都很重要：**

**专门写到了另一个word中：《并行架构：进程和内存协调》**

##########################################

第四章：关系查询处理器

这章主要讲SQL语句的执行：先是解析和授权然后查询重写，就是视图重写、表达式重写、语义优化之类的，查询优化器这部分不会，这几天刚上了一本《数据库查询优化器的艺术》，想看下这个再看看。

##########################################

第五章：存储管理

参考写的《数据库中的存储管理》

###########################################

第六章：事物：并发控制和恢复

这一章也是很基础但是很重要的部分，书上都学了原理，这里对照实际商业产品的实现讲了些例子。

也是单独列一篇吧！

《事物：并发控制和恢复》

###########################################

第七章： 共享组件

这章主要是一些商业DBMS中都存在，却在研究文献中很少涉及到的共享组件和工具。

一 目录管理器

数据库的相关信息以元数据的形式在数据库的目录管理器中保存。目录管理器记录了系统中基本实体（用户，模式，表，列，索引等）的名称及他们的关系，并以一系列表的形式存储在数据库中，元数据与数据保持一样的格式，可以使系统在使用上更加紧凑。

出于效率的原因我们常采用不同的防水处理基本的数据目录和通常的表。目录中经常被访问的部分常根据需求物化在内存中，所用的数据结构一般是对目录的扁平关系结构进行非规范化处理。。。。。。

二 内存分配器

教材介绍的DBMS的内存管理往往全部集中在缓冲池上，实际上数据库系统也分配了大量内存用于其他任务。

为查询操作分配内存的注意事项：

数据库厂商们为空间密集型的操作符（如哈希连接和排序）所采用的内存分配方案各有不同。一些系统（比如DB2 for zSeries）允许DBA（数据库管理员）控制这些操作能使用的RAM数量，并且保证每个查询在执行时都能获得该数量的RAM。准入控制策略确保了这一保证。在这样的系统中，操作符通过内存分配器从堆中分配它们的内存。这些系统提供了很好的性能稳定性，但是，会强制DBA决定如何去平衡各个子系统（如缓冲池和查询操作符）之间的物理内存使用。

其他系统（比如Microsoft SQL Server），从DBA手中接管了内存分配任务，实现了内存分配的自动化管理。这些系统尽量智能地在查询执行的多个组件中分配内存，包括缓冲池中的页面缓存和查询操作符的内存使用。用于所有这些任务的内存池即缓冲池本身。因此，在这些系统中的查询操作符通过DBMS实现的内存分配器从缓冲池中取得内存，并且，只在连续地请求超过缓冲池页面大小的内存时才使用操作系统的分配器。

这个区别呼应了第6.3.1节中关于查询准备的讨论。前一类系统假设由DBA来完成复杂的调优工作，DBA对系统内存各种参数进行仔细配置后，系统的工作负载将会服从于这些配置好的参数。在这些条件下，这样的系统应该总是能够像预期的那样很好地执行。后一类系统则假设DBA不能正确地设置这些参数，并努力以软件逻辑来代替DBA的手工调整。

系统保留了自适应改变其相关分配的权力，这为在可变的工作负载上获得更好的性能提供了可能性。如第6.3.1节所讨论的那样，**从这个区别中不仅可以看出这些数据库厂商希望其产品如何被使用**，也可以看出他们顾客的管理经验（和财务资源）。

7.3 磁盘管理子系统

我感觉这部分应该和第五章的内容结合起来看，主要讲的还是磁盘部分的内容。。

7.4 备份服务

7.5 管理，监控和工具

###########################################

第8章 结束语

从本文中应该清楚的是，现代商业数据库系统是构建在两个基础之上的，一个是学术研究，另一个是**为高端客户开发工业级别的产品所积累的大量经验**。编写和维护一个高性能、全功能的关系型DBMS的任务，需要投入巨大的时间和精力。然而，许多关系型D**BMS的经验已转化到新的领域。Web服务、网络附加存储、文本和电子邮件库、通知服务和网络监控等，这些领域都可以从DBMS的研究和经验中受益**。**数据密集型的服务是当今计算的核心，数据库系统设计的知识是可以被广泛应用到各个领域的技能**，既包括数据库领域也包括其他领域。这些新的应用方向，也带来了一些数据库管理方面的研究问题，这为数据库社区和其他计算领域之间的交互开辟了新的道路。