



论文阅读报告

Architecture of a Database System

2021 年 6 月 10 日

学院: 智能与计算学部

班级: 网络安全一班

姓名: 谢远峰

学号: 3019244283

1 综述

- 综述

数据库作为早期多用户服务系统之一，其研究催生了诸多为保证系统可扩展性及稳定性的系统开发技术，被应用于诸多领域。数据库管理系统广泛存在于当前的计算机系统领域并占据重要地位，尤其在当今大数据时代，DBMS 的地位更加突出。

- 动机

全文聚焦于数据库工作系统的系统架构设计问题，以当前流行的商业开源软件作为参考，从体系架构角度探讨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、事务处理系统、处理查询及优化结构及其具有代表性的共享组件和应用。以宏观的视角讨论数据库管理系统的结构组成及各个部分粗略介绍。系统地将如何构建数据库系统的知识进行梳理，从应用角度对数据库架构进行全局讲解。对于常见教科书中省略的部分进行详细描述。本文旨在记录在实际环境中实施数据库的艺术，并有助于记录学术研究与实际实施之间的差距。

- 解决方案

本文没有为数据库架构提供任何规定的解决方案。相反，它是一篇描述数据库系统研究和实施状态的调查论文。本文的重点放在关系数据库的体系结构上。尽管具有局限性，但架构思想仍然在很大程度上适用于现代系统。

- 论文贡献

为数据库关系架构提供理论性概述和整理，可依据本文完成 DBMS 的改进。NoSQL 系统的兴起导致许多替代架构的出现，但这些较新的系统仍然可以从使用本文中描述的通用架构中受益。例如，并发控制和事务管理是现代数据库中持续关注的问题。

- 获得的启发

任何数据库实现都可以遵循本文中的一些指导原则——无论它们是否完全按照介绍的方式实现。通过阅读该论文，读者能够了解成功实现数据必须考虑的设计问题，并能够使用提供的资源深入研究感兴趣的课题。

2 第一章

介绍当前主流的数据库类型及架构

数据库的架构设计，包括个人计算机客户端控制、进程管理器、查询处理管理、存储与缓冲管理、共享组件及工具共 5 个部分。从 SQL 查询的整个生命周期中查看数据库系统的主要部分的体系结构概述。这包括如何接收、解析和优化查询，以及如何将结果数据作为事务的一部分从存储中返回。

3 第二章

介绍 DBMS 进程，线程模型及其优化

从简单的支持线程单片机模式陈述多种进程模型的实现及优缺点对比（进程模型、线程模型、进程池）。基于进程模型的优化提出共享空间和缓冲区的优化两个概念。缓冲区优化分为磁盘 I/O 缓冲（缓冲池和日志尾部的处理）和客户端通信缓冲（数据预提取）。提高硬件性能，允许多进程的实现，采用工作者拥有一个进程、一个线程、进程/线程池三种方式进行实现，并进行对应的优缺点分析。提供系统过载导致性能衰退的处理方案：控制客户端连接数，准入控制器状态信息查询并进行动态调整。

4 第三章

介绍集群技术、多进程系统和多核系统

- 共享内存

所有处理机可使用相同的内容和硬件设施，具有大致相同的性能。遵循单处理机方式，系统通常支持作业（进程或线程）被透明地分配到每个处理器上，并且共享的数据结构继续可以被所有的作业所访问。

- 无共享内存

由多个独立计算机集群组成，可以通过网络进行高速互连通信，或在商业网络组件上增强通信频率。由

DBMS 完成不同机器任务的协调。每个处理器可以独立于其他处理器执行。

- 共享磁盘

所有的处理器可以访问具备大致相同性能的磁盘，但不能访问彼此的 RAM。优点为管理成本低，单个 DBMS 处理节点发生故障不会影响其他节点访问整个数据库。需在多台计算机间进行数据共享协调。

- 非均衡内存访问

拥有独立内存的集群系统，提供共享内存编程模型，群中的每个系统都能快速访问本地内存，DBMS 可通过忽略内存访问的非均衡性在共享内存系统上运行。优化时，应采用两个基本准则：（1）尽量使用本地为处理器分配内存使用（2）尽可能保证一个 DBMS 使用者被安排到与之前相同的硬件处理器。

- 线程和多处理器

轻量级 DBMS 线程包：所有线程运行在一个单一的操作系统进程中，单一的进程一次只能在一个处理器上执行。解决方法：每个物理处理器产生一个进程，可最大化硬件固有的并行度，同时最小化每个进程的内存开销。

5 第四章

介绍关系查询处理器，聚焦于常见的 SQL 命令

- 查询解析、授权、重写

检查查询定义的正确性、解决名字和引用、将查询优化为优化器使用的内部形式、核实用户的执行权限。查询处理器调用目录管理器、检查表注册情况，通过后进行检查授权并调到重写模块实现查询的简化。

- 查询优化

将内部的查询表示转化为高效的查询执行计划。优化领域包括但不限于计划空间、选择性估算、自动调优、搜索算法、并行处理。

- 查询执行

主流的查询执行器采用迭代器模型，操作原子为数据流图中的结点，为迭代器类中的一个子类，迭代器连接控制流和数据流并预分配固定数量的元组描述符，

- 访问方法

对系统支持的基于磁盘的数据结构访问进行管理，包含无须的文件和多种索引。

- 数据仓库

数据仓库服务支持于决策支持、包含大量历史数据的数据库，OLTP 系统中的数据更新周期性加载进数据仓库中，工作者构建工作流系统，不断从系统挂取数据装载到仓库中。利用批量加载机技术可以提高热门查询的速度及性能，且没有 SQL 层的开销，充分利用面向 B-树的特殊批量加载方法的优点。

- 数据库扩展性

主流数据库系统支持主流编程语言描述的多种数据类型，如抽象数据类型，结构化类型，XML。关系数据库查询引擎的底层架构设计区别主要体现于优化器搜索策略和查询执行器控制流模型。

6 第五章

介绍 DBMS 的存储管理管理器，分析 DBMS 控制存储能力，讨论存储层次

- 存储管理器分类

直接与底层的面向磁盘的块模式设备驱动程序进行交互 (原始模式访问); 使用标准的 OS 文件系统设施

- 空间控制

将数据直接存储到“原始”磁盘设备，导致无法提供给需要使用文件系统接口的工具，且由于接口的独特性，DBMS 的可移植性变差。采用地址偏移量定位数据的方法，在较少包含密集 I/O 工作负载状态下，系统开销可以有效降低。

- 时间控制

软件及硬件运行失败的情况下，若无对磁盘写操作时间和顺序进行有效的显式控制，则无法实现 DBMS 原子恢复。采用钩子，使程序规避对于文件的缓冲，确保写操作请求写入磁盘时，无双冲突，且 DBMS 可实现页面置换策略。

- 缓冲管理

DBMS 根据系统需要和可用资源动态调整缓冲池大小，缓冲池被拆分组织成帧数组并执行哈希映射，块从磁盘整体无格式改变地复制至缓冲池中。操作系统采用传统的 LRU 策略进行表的全局扫描。

7 第六章

介绍基础架构，分析数据库组件内部的依赖关系

- 事务特性及可串行化

原子性、一致性、隔离性、持久性。可串行化有三种并发控制技术：严格两段锁 (2PL)、多版本并发控制 (MVCC)、乐观并发控制 (OCC)。2PL 应用最为广泛。

- 锁和锁存器

锁关联每一个事务，事务有对应 ID；锁存器作为对数据库锁的补充，被用于处理互斥问题。

- 日志管理器

数据库标准恢复机制采用写前日志协议。采用“DIRECT,STEAL/NON-FORCE”的模式运行来将快速通道速度最大化。

- 索引的锁及日志

B+-树的 latch 作为轻量级锁采用三种方案 (保守方案、联结锁方案、右链接方案) 实现树物理结构改变。索引通过日志逻辑保证日志记录和恢复效果。

- 事物存储的相互依赖

恢复逻辑依赖于并发策略；存取算法并发协议依赖于回复逻辑，保证缓冲区管理器高自由度和高独立性。

8 第七章

介绍数据库系统中共享组件和工具

- 目录管理器

将系统中数据的相关信息以元数据格式保存；实体名称及之间关系以表格式存储。

- 内存分配器

使用基于内存池的内存分配器，利用一个连续的虚拟内存区域列表。

- 磁盘管理子系统

将表映射到设备或文件上，维护存储的时间和空间。

- 备份服务

在网络上备份数据库且周期性更新。备份服务主要分为物理备份、基于触发器的备份、基于日志备份。

- 管理、监控和工具

DBMS 提供一系列工具管理系统，可实现优化统计信息收集、物理重组、索引建立等