

■ 微软新技术教材

SQL Server 数据库原理

—— 设计与实现

周立柱 冯建华 孟小峰 陈立军
杨冬青 王 珊 唐朝晖 [美] Don Vilen 编著 罗晓沛 审校



清华大学出版社

微软新技术教材系列

SQL Server 数据库原理

——设计与实现

周立柱 冯建华 孟小峰
陈立军 杨冬青 王 珊 编著
唐朝晖 [美] Don Vilen
 罗晓沛 审校

清华大学出版社
北 京

内 容 简 介

本书从如何设计数据库管理系统的角度出发,深入介绍了 SQL Server 2000 的相关内容,目的是希望读者在阅读此书之后,能够学习到像 SQL Server 2000 这样优秀的数据库管理系统软件的设计与实现技术,进而掌握、提高设计这种大型复杂的系统软件的能力。全书围绕 SQL Server 2000 的核心部分——体系结构、存储系统、查询处理、并发控制,以及事务管理,分别叙述了各个部分的设计思想、基本算法,以及具体实现。此外,本书还对 SQL Server 2000 新的扩充部分数据仓库联机分析处理、XML 等也作了介绍。

本书可作为高年级计算机及相关专业数据库教材,适于高校师生选用。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

SQL Server 数据库原理:设计与实现/周立柱等编著. —北京:清华大学出版社,2004.3
ISBN 7-302-08089-5

I. S… II. 周… III. 关系数据库—数据库管理系统, SQL Server 2000 IV. TP311.138

中国版本图书馆 CIP 数据核字(2004)第 009926 号

出 版 者:清华大学出版社
<http://www.tup.com.cn>
社总机:010-62770175

地 址:北京清华大学学研大厦
邮 编:100084
客户服务:010-62776969

组稿编辑:丁 岭

文稿编辑:丁 岭 闫红梅 魏江江 许振伍

封面设计:付剑飞

印 装 者:清华大学印刷厂

发 行 者:新华书店总店北京发行所

开 本:185×230 印张:26 字数:575 千字

版 次:2004 年 3 月第 1 版 2004 年 3 月第 1 次印刷

书 号:ISBN 7-302-08089-5/TP·5848

印 数:1~6000

定 价:42.00 元

Forward by Jim Gray, Microsoft

to

Introduction to Microsoft SQLServer2000

—**From the Perspective of its Design and Implementation.**

By Prof. Lizhu Zhou, Prof. Jianhua Feng, Prof. Xiao Feng Meng, Prof. Lijun Chen, Prof. Shan Wang, Prof. Dongqing Yang, and Don Vilen

This textbook by six of China's leading database experts and an expert from SQL Server Team gives a logical introduction to database concepts. It explains how to use a database system and how database systems work inside. The authors begin by explaining the system's architecture, then step you through the system's layers, starting with low-level files, then tables (including text and image datatypes), and then they explain indexing structures. Having laid this groundwork, they explain relational algebra and how the query processor compiles non-procedural queries into efficient query plans. Finally the book covers advanced topics such as transaction recovery and concurrency, analytic processing, and XML databases.

In writing this book, the professors used Microsoft SQLServer as their example database. They spent several weeks in Redmond Washington living with the SQLServer developers to get an intimate understanding of the system's internals, and they have corresponded with the developers to clarify questions since that visit. With this knowledge and their own extensive background, they worked for one year to produce a very informative book intended to be used as a text at university level. They debugged the book by using it in their classes.

I do not know anyone else who has taken such a systematic and scholarly approach to writing a database textbook. This book will certainly be a classic text in China, and will likely be of interest worldwide. I confidently recommend this book to you if you are interested in learning database concepts or learning the implementation technologies that underlie database systems.



Jim Gray
Distinguished Engineer
Microsoft Corporation.

序 言

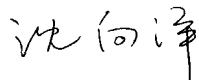
数据库技术作为信息产业发展的关键技术，近年来发展十分迅速，目前已成为现代计算领域不可或缺的重要组成部分。其中，SQL Server 数据库系列作为当今应用最广泛的数据库之一更是受到了企业界和学术界普遍关注，但市场上相关的技术书籍多集中在应用和操作的层面。在 IT 院长高峰会上，许多院长都提出希望能够写一本权威的 SQL Server 数据库原理方面的教材，这使我们萌发了组织编写本书的想法。

在这个提议下，微软亚洲研究院同周立柱、冯建华、孟小峰、陈立军、杨冬青、王珊等几位(以编写章节顺序排名)中国数据库领域的专家以设计和实现为重点，合作编写了这本 SQL Server 数据库原理教科书。几位教师在深入研究 SQL Server 数据库构架和特性的基础上进行了分工编写和统筹组稿。为进一步加深对数据库开发的理论和过程的理解，微软亚洲研究院于 2002 年邀请六位教师访问了微软公司西雅图总部的软件开发部门，同 SQL Server 开发组的工程师进行了开放、热烈的讨论，其中资深项目经理唐朝晖博士和 Don Vilen 更志愿为本书的编写提供咨询和相关支持。在编写的关键阶段，他们还专程从总部飞至北京同各位老师一起讨论交流。

微软亚洲研究院从创立之初就致力于同中国教育界的各项交流与合作，在支持高校精品课程建设方面也做出了不懈的努力，本书的出版是我们同教育界合作的又一次成功探索。

在此，我代表微软公司感谢为此书编写付出心血的各位老师，同时感谢罗晓沛教授的认真审校和清华大学出版社的大力支持，也感谢微软亚洲研究院高校关系部各位同事在组织编写中付出的辛勤工作。

希望本书能够给致力于数据库研究和应用开发的朋友一些有益的启示!



微软亚洲研究院 院长
2004 年 2 月 20 日

目 录

第 1 章 绪论	1
1.1 数据库管理系统的通用功能与系统组成	1
1.2 SQL Server 2000	3
1.3 SQL Server 2000 技术特色	4
1.4 本书的各章安排与内容	5
参考文献	5
第 2 章 SQL Server 的体系结构	6
2.1 SQL Server 的客户/服务器方式	6
2.1.1 网络程序库(Net-Library)	7
2.1.2 用户态调度器(User Mode Scheduler, UMS)	9
2.1.3 开放数据服务(Open Data Services, ODS)	9
2.2 关系引擎与存储引擎	10
2.3 存取方法管理器	12
2.3.1 行操作管理器与索引管理器	12
2.3.2 页面管理器与文本管理器	14
2.4 事务管理器	15
2.5 锁管理器	16
2.6 内存管理器	17
2.7 事务日志与故障恢复	18
2.7.1 日志管理	18
2.7.2 检查点操作	18
2.7.3 恢复过程	19
2.7.4 使用页面 LSN 的恢复技术	20
2.8 SQL 内核与操作系统间的交互	21
2.8.1 线程处理与对称多进程处理	22
2.8.2 工作者线程池	23
2.8.3 磁盘 I/O 操作	24
2.9 本章小结	24

参考文献	25
第3章 数据库与数据库文件	26
3.1 SQL Server 系统数据库	27
3.1.1 master 数据库	27
3.1.2 model 数据库	27
3.1.3 tempdb 数据库	28
3.1.4 pubs 数据库	28
3.1.5 Northwind 数据库	29
3.1.6 msdb 数据库	29
3.2 数据库文件	29
3.2.1 基本知识	30
3.2.2 SQL Server 数据库文件	30
3.3 SQL Server 数据库文件组	31
3.3.1 默认的文件组	32
3.3.2 使用多个文件	32
3.3.3 利用文件组创建数据库的实例	33
3.4 SQL Server 数据库文件结构	34
3.4.1 基本知识	34
3.4.2 SQL Server 用户数据库	35
3.4.3 SQL Server 的空间分配	37
3.5 数据库的备份与恢复	40
3.5.1 备份类型	40
3.5.2 恢复模式	41
3.6 本章小结	43
参考文献	43
第4章 表的物理存储	44
4.1 内部存储概述	45
4.2 数据页	47
4.2.1 基本知识	47
4.2.2 数据页的格式	47
4.2.3 SQL Server 的数据页	50
4.2.4 检查数据页	52
4.3 数据行的结构	56

4.3.1	基本知识	56
4.3.2	SQL Server 的数据行结构	58
4.3.3	列偏移数组	60
4.4	SQL Server 中定长和变长行的存储	60
4.4.1	定长行的存储	60
4.4.2	变长列的存储	63
4.5	SQL Server 的页链接	65
4.6	文本和图像数据	65
4.6.1	LOB 数据存储在表外	66
4.6.2	LOB 数据存储在数据行	67
4.7	SQL Server 的 PRIMARY KEY 和 UNIQUE 约束	71
4.7.1	列值的可空性	71
4.7.2	索引属性	72
4.7.3	选择码	72
4.8	SQL Server 的临时表	73
4.8.1	私有临时表(#)	73
4.8.2	全局临时表(##)	73
4.8.3	直接使用 tempdb	74
4.8.4	临时表上的约束	74
4.9	SQL Server 的系统表	75
4.10	本章小结	78
	参考文献	78
第5章	索引及行操作	79
5.1	预备知识	80
5.1.1	堆文件	80
5.1.2	B 树与 B ⁺ 树索引结构	80
5.2	SQL Server 的索引组织	84
5.2.1	簇集索引	85
5.2.2	非簇集索引	86
5.3	在 SQL Server 中创建索引	87
5.3.1	主要参数说明	88
5.3.2	其他选项说明	88
5.3.3	约束与索引	89
5.4	SQL Server 的索引页结构	90

5.4.1	具有惟一标识符的簇集索引	90
5.4.2	索引行格式	93
5.4.3	簇集索引的结点行	94
5.4.4	非簇集索引的叶级行	99
5.4.5	非簇集索引的结点行	104
5.5	索引的空间需求	108
5.5.1	B 树的大小	109
5.5.2	索引的实际大小与估计大小	110
5.6	索引的管理	111
5.6.1	索引碎片的类型	111
5.6.2	索引碎片的检测	111
5.6.3	删除索引碎片	113
5.7	SQL Server 的特殊索引	114
5.7.1	先决条件	114
5.7.2	计算列上的索引	116
5.7.3	索引视图	117
5.8	索引的使用	117
5.8.1	查找行	118
5.8.2	连接运算	118
5.8.3	排序	118
5.8.4	分组聚集	119
5.8.5	维护惟一性	119
5.9	插入行	119
5.10	分裂页	120
5.10.1	分裂索引的根页	120
5.10.2	分裂索引的中间页	120
5.10.3	分裂数据页	120
5.11	删除行	124
5.11.1	从堆中删除行	124
5.11.2	从 B 树中删除行	127
5.11.3	从索引的中间结点删除行	130
5.11.4	回收页	130
5.12	更新行	130
5.12.1	移动行	131
5.12.2	管理转向指针	133

5.12.3 在位更新	133
5.12.4 异位更新	133
5.13 表级与索引级的数据更新	136
5.14 本章小结	136
参考文献	136
第6章 SQL Server 查询处理	137
6.1 SQL Server 查询处理概述	137
6.1.1 SQL Server 查询的体系结构	138
6.1.2 查询处理特点	140
6.2 执行模型	142
6.2.1 基本知识	142
6.2.2 多种执行模型的介绍	143
6.2.3 SQL Server 查询执行模型	144
6.2.4 物理操作的顺序	148
6.3 基本运算	148
6.3.1 选择运算	148
6.3.2 哈希	149
6.3.3 排序	152
6.3.4 其他操作	154
6.4 连接	156
6.4.1 嵌套循环连接	156
6.4.2 归并连接	157
6.4.3 Hash 连接	158
6.5 聚集	167
6.5.1 基于嵌套循环的聚集算法	167
6.5.2 基于排序的聚集	168
6.5.3 基于哈希的聚集	169
6.6 子查询处理	169
6.6.1 假设	170
6.6.2 异步 I/O	172
6.6.3 避免 I/O	172
6.6.4 数据流	176
6.6.5 控制流	178
6.7 本章小结	181

参考文献	181
第7章 SQL Server 查询优化技术	183
7.1 概述	183
7.1.1 基于代价的方法和启发式方法	184
7.1.2 语法优化和语义优化	185
7.1.3 新的优化方法	185
7.2 SQL Server 的查询优化概述	186
7.2.1 Cascades 查询优化器框架	187
7.2.2 MEMO 结构	189
7.3 SQL Server 优化策略	193
7.3.1 优化模型	193
7.3.2 优化时间	195
7.3.3 转换及转换规则	195
7.3.4 多阶段优化	199
7.4 连接顺序的选择方法	201
7.5 子查询的优化	210
7.6 SQL Server 的物理优化	216
7.6.1 查询分析	216
7.6.2 如何选择索引	221
7.6.3 连接算法的选择	229
7.7 代价计算	230
7.7.1 一般的代价计算方法	230
7.7.2 选择运算	231
7.7.3 排序和连接	235
7.7.4 SQL Server 中统计信息的管理	235
7.8 其他的辅助功能	237
7.8.1 如何正确选择计划	237
7.8.2 提示技术	237
7.8.3 更新计划	238
7.9 本章小结	238
参考文献	239
第8章 事务	240
8.1 事务概念	240

8.1.1	事务基本性质	241
8.1.2	事务生命周期	243
8.2	SQL Server 中的事务定义	244
8.2.1	事务定义模式	244
8.2.2	事务执行模式	245
8.2.3	批处理、触发器中的事务	247
8.3	SQL Server 中的事务模型	248
8.3.1	平面事务	248
8.3.2	嵌套事务	249
8.3.3	保存点	251
8.3.4	分布式事务	253
8.4	事务调度	254
8.5	事务隔离性级别	258
8.5.1	并发调度中的不一致性现象	258
8.5.2	SQL Server 中事务隔离性级别的定义	260
8.6	本章小结	264
第9章	并发控制	265
9.1	基于锁的并发控制	266
9.2	锁模式	266
9.2.1	基本锁模式	266
9.2.2	意向锁	269
9.2.3	其他锁模式	271
9.2.4	码区间锁	272
9.3	锁资源	276
9.4	锁持有度	278
9.5	锁拥有者	279
9.6	封锁协议	280
9.6.1	两段锁协议	280
9.6.2	两阶段提交协议	281
9.7	显示锁信息	281
9.8	锁管理器	290
9.8.1	封锁体系结构	290
9.8.2	锁块	292
9.8.3	锁资源块	292

9.8.4	锁拥有者块	294
9.9	锁升级	294
9.10	封锁产生的问题	295
9.10.1	阻塞	295
9.10.2	活锁	298
9.10.3	死锁	298
9.11	绑定连接	303
9.12	强制封锁类型	305
9.13	乐观并发控制原理	306
9.13.1	有效性检查协议	307
9.13.2	游标并发	307
9.14	本章小结	309
第10章	恢复	310
10.1	故障类型	310
10.1.1	事务故障	310
10.1.2	系统故障	311
10.1.3	介质故障	311
10.2	日志	312
10.2.1	基本日志结构	312
10.2.2	基于日志的恢复	313
10.2.3	检查点	316
10.3	ARIES 算法介绍	318
10.3.1	ARIES 概述	318
10.3.2	ARIES 中数据结构	320
10.3.3	ARIES 的常规处理	322
10.3.4	重启处理	323
10.4	SQL Server 中的备份与恢复	328
10.4.1	备份类型	328
10.4.2	将数据库还原到前一个状态	332
10.4.3	恢复模型	333
10.5	SQL Server 事务日志物理构架	335
10.5.1	创建日志文件	336
10.5.2	虚拟日志文件	336
10.5.3	截断事务日志	338

10.5.4 收缩日志文件·····	339
10.6 SQL Server 事务日志逻辑构架·····	341
10.6.1 事务日志内容·····	341
10.6.2 日志管理器·····	342
10.6.3 SQL Server 的检查点·····	343
10.7 本章小结·····	345
参考文献·····	345
第 11 章 Analysis Services 2000 概述 ·····	346
11.1 多维数据模型及相关概念·····	346
11.1.1 数据仓库、OLAP 与 Data Mining 基本框架·····	346
11.1.2 基本概念·····	348
11.1.3 星形模式(Star Schema)·····	351
11.1.4 雪花模式(Snow Flake Schema)·····	351
11.1.5 聚集(aggregation)·····	352
11.2 Analysis Services 的体系结构和功能·····	354
11.3 Analysis Server 的多维查询语言 MDX 及其处理·····	356
11.3.1 MDX 的坐标系统·····	356
11.3.2 MDX 的语法格式·····	357
11.3.3 轴维和切片维·····	357
11.3.4 MDX 查询示例·····	358
11.3.5 MDX 语句的查询处理·····	360
11.4 Analysis Server 的 Cube 处理·····	362
11.5 Analysis Server 的数据挖掘·····	363
11.5.1 OLE DB for DM·····	363
11.5.2 数据挖掘模型·····	364
11.5.3 Analysis Services 数据挖掘三层体系结构·····	366
11.6 本章小结·····	369
参考文献·····	369
第 12 章 SQLXML——SQL Server 对 XML 的支持 ·····	371
12.1 SQL Server 对 XML 的支持·····	372
12.1.1 连接两个世界的桥梁·····	372
12.1.2 SQL Server 中支持 XML 访问的体系结构·····	373
12.1.3 SQLXML·····	373

12.2	SQL FOR XML 查询	375
12.2.1	关系视图上的 XML 查询	375
12.2.2	FOR XML 查询的几种模式	375
12.2.3	生成 XML 形式查询结果的机制	378
12.3	建立在关系数据上的 XML 视图	379
12.3.1	XML 视图概述	379
12.3.2	利用带注释的映象模式在关系数据库上建立 XML 视图	380
12.3.3	基于 XML 视图的 XPath 查询	384
12.3.4	基于 XML 视图的更新	385
12.4	XML 数据转换为关系数据	386
12.4.1	OpenXML	387
12.4.2	XML Bulk Load	389
12.5	本章小结	394
	参考文献	394

第 1 章 绪 论

数据库系统是现代信息管理不可缺少的重要基础设施。它被广泛地用于各种行业、各种组织、各种部门的数据保存、业务处理、决策分析等日常工作。在今天的信息化社会，数据库已经成为一个单位、组织赖以发展的命脉所在，数据库管理系统也发展成为继操作系统之后最复杂的系统软件。

数据库是计算机科学与技术学科的一个重要研究领域，它的应用实践与原理的学习是计算机科学与技术专业教学课程设置中的重要内容。这些课程大体上可以分为两个层次：一是以应用为主，配以基本原理教学的本科的专业基础课；二是以原理为主，讨论数据库管理系统设计与实现的本科高年级或研究生的专业课。本书的主要目的是为第二个层次的教学服务，为学习、研究数据库管理系统的软件设计与实现提供一本教材。与其他同类的教材相比，本书没有采用常用的系统、全面介绍通用原理与技术的写法，而是紧紧结合微软公司的 SQL Server 2000 产品，介绍它的设计思想、原理、算法、具体实现等细节。

本章是全书的开篇，其主要内容如下：首先简要介绍数据库管理系统的一般性功能与系统组成，接着讨论为什么选择 SQL Server 2000 来学习数据库管理系统的设计与实现，然后是关于 SQL Server 2000 系统特点的简单介绍，最后是本书的有关内容以及各章安排。

1.1 数据库管理系统的通用功能与系统组成

数据库管理系统(Database Management System, 简称 DBMS)是一种软件工具，通过它的支持，用户可以系统、有序、高效地生成用于日常业务与决策的数据，这些数据能够长期保存，它是称之为“数据库”的永久性存储，例如保存在磁盘上，它们不受计算机系统软硬件的故障的影响，可以方便地进行查询。数据库管理系统的通用功能可以概括如下。

- **支持数据库的创建与管理：**这主要体现在系统提供的数据库定义语言 DDL(Data Definition Language)与管理命令上。以关系数据库管理系统为例，它的 DDL 就包括生成数据库、生成关系表、建立索引等命令。在数据库管理方面则包括建立用户、用户组、用户授权、查看数据库状态等命令。
- **支持永久性存储：**这种支持应当独立于使用数据库的应用，可供许多用户共享，数据量可以达到千兆字节(Giga Bytes)，千千兆字节(Tera Bytes)，甚至更高。不仅如此，系统还必须具有良好的数据结构，为访问这些数据提供高效的手段。

- **支持高级查询：**高级查询使用户在查询数据库时十分方便，用户只需要说明查询的条件，而如何从数据库里找到符合条件的过程则由数据库管理系统自动完成。关系数据库管理系统为用户提供 SQL 语言实现这一要求，通过 SQL，普通用户可以交互地查询到想要的数据库，程序开发者则可以在 C、C++ 等高级语言里访问数据库，完成应用程序的开发。
- **支持事务管理：**事务是数据库里具备原子性、一致性、隔离性、以及持久性等特性的特殊处理进程，数据库管理系统必须保证事务的这些特性的实现，支持多个事务同时访问数据库时对资源的并发要求，支持在系统出现故障时的事务恢复，保证数据库的一致性。
- **支持新的高级应用：**技术的进步与应用的深入，对数据库管理系统提出了许多新的需求，例如数据仓库与联机分析处理、XML 数据管理、空间数据管理、移动数据管理等等。提供上述支持，显然是数据库管理系统必须要考虑的。

大型复杂软件系统的设计通常都采用“分层”的办法，这样可以使得系统的实现变得相对简单，数据库管理系统就是如此。按照这种“分层”的思想，一个数据库管理系统可以划分为数据库管理维护、用户界面处理、查询处理、事务管理、调度、缓冲区管理，以及存储管理几大部分。图 1.1 是它的一个参考设计方案，以下是图中各个组成部分以及它们之间的交互关系的概要描述。

图 1.1 右上角的“创建与维护”子系统接收数据库维护人员，如数据库管理员生成、维护数据库的 DDL 等命令，完成前面介绍的数据库管理系统的第 1 项功能。它将命令处理的结果如数据库与关系表的结构、索引情况、用户权限分配等信息(元数据)记录在数据字典里，供给系统的其他部分使用。

“用户界面处理”模块从用户那里接收 SQL 命令，利用数据字典对命令进行语法语义检查，将命令转送给查询处理子系统。在查询处理结束时，“用户界面处理”模块将 SQL 命令的最终结果返还给用户。

“查询处理”子系统负责将 SQL 命令转换成关系代数操作，同时进行各种优化，形成由关系代数操作构成的查询执行计划，将计划输出给“事务处理”子系统。

“事务处理”子系统为查询生成事务标识，记录事务处理信息，将查询执行计划的操作送交“调度”模块，等待事务的结束。

“调度”模块对来自多个用户的请求通过加锁或时间戳等方法对访问数据库的操作排队，并依次将读写日志、读写数据库的操作输出给“恢复”子系统。

“故障恢复”子系统调用“缓冲管理”模块完成对于日志、数据库的读写操作，在这一过程中，记录数据变化的日志记录必须在数据更新之前先行写入日志，以保证当出现故障时能够对数据进行恢复。

“缓冲管理”模块查看内存里的缓冲区，根据它与“调度”模块之间对于数据与日志管理的约定，必要时调用“存储管理”子系统完成数据的读写。