高等学校教材

数据库系统及应用

（第二版）

崔　巍　编著

高等教育出版社

需要全本请访问 www.ertongbook.com

内 容 提 要

本书是“北京市高等教育精品教材”立项项目。

本书是在第一版的基础上修订而成，与第一版比较，在整体结构上作了适当调整，根据数据库技术的发展增加了一些新的内容。本书在科学性、完备性、实用性和教材的可用性方面与第一版相比均有一定的提高。

全书包括数据库基础、数据库系统、数据库设计、数据库应用和数据库新技术等内容，共分 16 章，具体内容主要包括 数据库的基本概念、数据模型、关系数据库、关系数据库系统、关系数据库标准语言 SQL、数据完整性、数据库安全、事务管理、并发控制、数据库恢复、关系数据理论、数据库设计、面向对象数据库、数据库应用的体系结构与开发、分布式数据库、数据仓库以及数据库技术新的研究领域和应用领域等。

本书既可以作为高等院校计算机、信息管理与信息系统等相关专业数据库课程的教材，也可供从事计算机软件以及数据库应用、管理和开发的科技人员、工程技术人员及其他有关人员阅读参考。

初 版 前 言

数据库技术始于 20 世纪 60 年代，经历了最初的基于文件的初级系统、20 世纪 60 ~ 70 年代流行的层次系统和网状系统，而现在广泛使用的是关系数据库系统。数据库应用也从简单的事务管理扩展到各个应用领域，如用于工程设计的工程数据库、用于因特网的 Web 数据库、用于决策支持的数据仓库技术、用于多媒体技术的多媒体数据库等，但应用最广泛的还是在基于事务管理的各类信息系统领域。数据库的体系结构也从最初的集中式数据库变化为基于客户/服务器机制的分布式数据库。随着面向对象技术的发展，关系对象数据库系统正在逐步完善和投入使用。而随着时代的进步和发展，数据库的应用领域会越来越广泛，数据库技术也将是所有信息技术和信息产业的基础。

本书分为四个部分，第一部分是数据库基础；第二部分利用 SQL Server 介绍基于客户/服务器机制的数据库管理系统、数据库保护技术和分布式数据库；第三部分介绍数据库系统的开发、数据库设计等；第四部分则概要介绍一下数据库技术的最新研究领域和应用领域及其发展。

第一部分分为四章(第一章至第四章)。第一章主要介绍什么是数据库及与之相关的一些基本概念；第二章介绍数据模型和三层模式数据库结构，包括实体 - 联系方法、概念数据模型、传统的三大数据模型、数据库的三层模式和数据库管理系统等内容；第三章介绍关系数据库、关系代数以及关系数据库的标准数据语言 SQL，并用大量的实例介绍了 SQL 的使用方法；第四章介绍关系数据库理论，包括函数依赖、公理系统、规范化与模式分解等内容。

第二部分分为五章(第五章至第九章)，在这一部分我们选择 Microsoft SQL Server 作为数据库管理系统的实例来介绍数据库管理系统的功能、管理和使用方法，其目的不是为了教会读者如何使用 SQL Server，也不是为了单纯介绍 SQL Server，而是通过 SQL Server 了解客户/服务器机制的数据库体系结构，了解一个完备的数据库管理系统应该具备的功能，了解应该如何管理和使用数据库管理系统。通过这一部分的学习，可以为日后管理和使用这类数据库管理系统打下一个良好的基础。第五章介绍客户/服务器数据库系统的基础知识，包括客户/服务器数据库的基本概念、SQL Server 的基本概念以及 SQL Server 的数据语言 Transact - SQL 等；第六章介绍如何充分利用数据库服务器的功能，包括 SQL Server 数据库的管理、数据完整性、存储过程及触发器的应用等；第七章介绍数据库保护技术，包括并发控制、安全控制、转储和恢复及镜像等；第八章介绍客户端的开发方法，包括 ODBC 的概念以及 Visual FoxPro、PowerBuilder 和 Visual Basic 等一些常用的客户端开发工具，并介绍如何利用这些工具进行开发等；第九章介绍分布式数据库和 SQL Server 的数据复制等内容。

第三部分是第十章,介绍一些常用数据库系统的开发方法、数据库设计以及规划客户/服务器数据库应用系统的方法和要考虑的一些问题。

第四部分是第十一章,介绍数据库的一些最新研究领域和应用领域及其发展,包括数据仓库技术、因特网上的 Web 数据库、面向对象数据库、主动数据库、并行数据库、工程数据库和知识库等。

在教材的编写过程中,笔者尽可能引入新的技术和方法,力求能反映当前的技术水平和未来的发展方向。但由于学识浅陋、水平有限,定有许多不足之处,还望同行和专家批评指正。

编 者

1999 年 2 月

再 版 前 言

本书第一版于 1999 年出版，通过几年的教学实践活动，对原有的教学内容和教材作了很多必要的补充和修订。另外数据库的技术也发展很快，无论是产品的版本还是应用的体系结构都有了一些新的变化。这些都促使作者对教材作一次幅度比较大的修订，以使教材可以更好地为教学服务。

本书的此次修订列为“北京市高等教育精品教材”立项项目。

本书按顺序可以分为 4 个部分：

第 1 部分为数据库基础部分（第 1 章至第 3 章）。第 1 章介绍数据库的基本概念以及数据管理技术的发展过程。第 2 章介绍 E－R 方法、概念数据模型、传统的三大数据模型、数据库的三层模式以及数据库管理系统的结构等。第 3 章重点、全面介绍了关系数据库和关系数据模型，包括关系的形式化定义、关系数据模型、关系完整性约束、关系代数和关系数据库的三层模式结构等。

第 2 部分为数据库系统部分（第 4 章至第 8 章）。这一部分强调数据库课程的实践性，建议以 SQL Server 2000 为教学和实践环境，书中涉及到的实践环节和例子均可以在 SQL Server 2000 下运行。第 4 章首先介绍了客户/服务器体系结构的概念，特别是数据库服务器的作用，然后介绍了 SQL Server 数据库系统的结构以及如何建立 SQL Server 数据库。第 5 章详细介绍了关系数据库标准语言 SQL，在表定义功能中介绍了实体完整性、参照完整性和用户定义完整性的实现，还通过大量的实例介绍了 SQL 查询功能的实现，最后介绍了 SQL 的宿主使用方式和动态使用方式等。第 6 章介绍了存储过程、触发器和规则等数据库技术。第 7 章介绍了数据库的安全问题，包括用户管理和权限管理等内容。第 8 章介绍数据库的事务管理，包括事务的性质、并发控制和数据库恢复等内容。通过这一部分内容的学习和实践，希望学生不仅掌握数据库的基本概念和技术，同时还具备使用和管理数据库的基本实践技能。

第 3 部分为数据库设计部分（第 9 章和第 10 章）。其中第 9 章介绍关系数据理论，包括函数依赖、公理系统、规范化和模式分解等内容。第 10 章介绍数据库设计，包括数据库设计的一般过程和内容，同时也希望学生对数据库建模工具能够有所了解和掌握，为此介绍了数据库建模工具 PowerDesigner。

第 4 部分为数据库应用和数据库新技术部分（第 11 章至第 16 章）。第 11 章介绍面向对象数据库，包括面向对象数据模型和面向对象数据库的特点等。第 12 章介绍了数据库应用的一般体系结构和开发方法，包括开放的客户端开发（客户/服务器应用）、分布式组件开

需要上本请访问 www.ertongbook.com

发(多层客户/服务器应用)和 Web 应用开发(浏览器/服务器应用)。通过这一章的学习,希望学生能够了解和掌握各种数据库应用的体系结构,并具备基本的开发思路和能力。这一章中的实践环节介绍了 PowerBuilder,也希望学生能够初步掌握一种现代的数据库应用开发工具。第 13 章介绍数据库应用的开发方法,包括生命周期方法、快速原型方法和面向对象方法,另外还简单介绍了面向对象的建模标准 UML。第 14 章介绍了分布式数据库和分布数据管理,包括分布式数据库的概念和特点以及 SQL Server 的分布数据管理技术。第 15 章介绍数据仓库,包括数据仓库技术以及 SQL Server 的数据仓库解决方案。第 16 章介绍了其他一些数据库研究和应用的新领域,包括并行数据库、主动数据库、工程数据库和知识库等内容。

本教材可按 60 ~ 70 学时使用,学时可以大致分配如下:第 1 部分数据库基础约 12 ~ 14 学时,第 2 部分数据库系统约 20 ~ 24 学时,第 3 部分数据库设计约 12 ~ 14 学时,第 4 部分数据库应用和数据库新技术约 16 ~ 18 学时。

为了加强教学的实践环节,这次修订增加了上机实验的题目和说明(见附录),通过这些实验,学生可以更好地理解和掌握数据库的相关知识,并具备一定的实践技能。

另外,为了方便教师选用本教材,本书还配备了电子教案、习题解答和部分应用程序代码及样板,如有需要可与作者联系(Email:cuiwei@biti.edu.cn)或到高等教育出版社的网站 http://www.hep.edu.cn/booksintro/jsj/index.html 下载。

在教材修订的过程中,笔者广泛听取了其他院校一些老师的意见以及一些学生的建议,力争克服第一版教材中的不足,在科学性、完备性、实用性和教材的可用性方面与第一版相比能有一定幅度的提高。但由于水平有限,定有许多不足之处,还望同行和专家批评指正。同时也感谢对第一版教材提出修改建议的老师和学生们。

崔 巍

2003 年 2 月

目　录

第 1 章　绪论 …… (1)
1.1　什么是数据库 …… (1)
1.1.1　相互关联的数据的集合 …… (1)
1.1.2　用综合的方法组织数据 …… (2)
1.1.3　低冗余与数据共享 …… (2)
1.1.4　数据具有较高的独立性 …… (2)
1.1.5　保证数据的安全、可靠 …… (2)
1.1.6　最大限度地保证数据的正确性 …… (3)
1.1.7　数据可以并发使用并能同时保证数据的一致性 …… (3)
1.2　数据库管理系统 …… (3)
1.3　数据库管理和数据库管理员 …… (4)
1.4　数据库系统 …… (5)
1.5　数据库的过去、现在和未来 …… (5)
1.5.1　人工管理阶段 …… (5)
1.5.2　文件系统阶段 …… (6)
1.5.3　数据库系统阶段及其发展过程 …… (7)
1.5.4　数据库技术的广泛应用领域和发展方向 …… (9)
本章小结 …… (10)
习题一 …… (10)
第 2 章　数据模型和三层模式数据库 …… (11)
2.1　信息结构与 E－R 方法 …… (11)
2.1.1　数据的三种范畴 …… (11)
2.1.2　实体－联系方法 …… (12)
2.2　概念数据模型 …… (15)
2.2.1　概念数据模型中的基本联系 …… (15)
2.2.2　理解和转换多对多联系 …… (19)
2.2.3　连接陷阱 …… (19)
2.2.4　概念数据模型——实例研究 …… (21)
2.2.5　扩展 Bachman 图 …… (24)
2.3　传统的三大数据模型 …… (25)
2.3.1　层次数据模型 …… (25)
2.3.2　网络数据模型 …… (27)
2.3.3　关系数据模型 …… (29)
2.4　数据独立性与三层结构 …… (31)
2.4.1　存储数据独立性 …… (32)
2.4.2　概念数据独立性 …… (33)
2.5　数据库管理系统的结构 …… (34)
2.5.1　数据库的三层模式结构 …… (35)
2.5.2　模式说明实例 …… (36)
2.5.3　数据库管理系统的总体结构 …… (40)
本章小结 …… (41)
习题二 …… (41)
第 3 章　关系数据库 …… (43)
3.1　关系数据库系统概述 …… (43)
3.1.1　关系数据库的发展 …… (43)
3.1.2　关系数据模型概述 …… (43)
3.2　关系数据模型 …… (45)
3.2.1　关系模型的数据结构和基本术语 …… (45)
3.2.2　关系的形式定义 …… (47)
3.2.3　对关系的限定（关系的性质） …… (49)
3.3　关系模型的完整性约束 …… (49)
3.3.1　实体完整性规则 …… (49)
3.3.2　参照完整性规则 …… (50)

3.3.3 用户定义完整性 …… (52)
3.3.4 完整性约束的作用 …… (52)
3.4 关系代数 …… (53)
3.4.1 基本概念和几个符号 …… (53)
3.4.2 传统的集合运算 …… (55)
3.4.3 专门的关系运算 …… (56)
3.4.4 基本运算及变换 …… (60)
3.5 关系数据库系统的三层模式结构 …… (61)
3.5.1 关系概念模式 …… (61)
3.5.2 关系存储模式 …… (62)
3.5.3 关系外部模式 …… (62)
本章小结 …… (62)
习题三 …… (63)
第 4 章 Microsoft SQL Server 数据库基础 …… (64)
4.1 客户/服务器体系结构 …… (64)
4.1.1 客户/服务器和其他应用结构 …… (64)
4.1.2 文件服务器与客户/服务器的数据库操作 …… (67)
4.1.3 客户/服务器结构的特点 …… (68)
4.1.4 三层客户/服务器模型 …… (69)
4.2 Microsoft SQL Server 基础 …… (71)
4.2.1 SQL Server 的数据库结构 …… (71)
4.2.2 系统数据库的作用 …… (72)
4.2.3 SQL Server 的企业管理器和查询分析器简介 …… (73)
4.2.4 建立数据库 …… (74)
4.3 Transact - SQL 简介 …… (78)
4.3.1 变量、数据类型与表达式 …… (78)
4.3.2 函数 …… (84)
4.3.3 流程控制和输入/输出命令 …… (87)
本章小结 …… (88)
习题四 …… (88)
第 5 章 关系数据库标准语言——SQL …… (89)
5.1 SQL 语言概述 …… (89)
5.2 SQL 的数据定义功能 …… (90)
5.2.1 基本表的定义 …… (90)
5.2.2 修改表结构 …… (94)
5.2.3 删除表 …… (95)
5.2.4 索引的定义 …… (95)
5.2.5 索引的删除 …… (97)
5.3 SQL 的数据查询功能 …… (97)
5.3.1 简单查询 …… (98)
5.3.2 排序 …… (102)
5.3.3 连接查询 …… (103)
5.3.4 外(OUTER)连接查询 …… (105)
5.3.5 嵌套查询 …… (108)
5.3.6 分组及计算查询 …… (115)
5.4 视图(View) …… (119)
5.4.1 从单个表派生出的视图 …… (119)
5.4.2 从多个表派生出的视图 …… (120)
5.4.3 视图中的虚列 …… (121)
5.4.4 视图的删除 …… (122)
5.5 SQL 的数据操作功能 …… (122)
5.5.1 插入操作 …… (122)
5.5.2 更新操作 …… (123)
5.5.3 删除操作 …… (124)
5.6 SQL 的数据控制功能 …… (124)
5.7 SQL 的宿主使用 …… (125)
5.7.1 嵌入识别与预编译 …… (125)
5.7.2 数据通信区与主变量 …… (125)
5.7.3 游标 …… (127)
5.7.4 利用游标进行删除和更新操作 …… (129)
5.8 动态 SQL …… (129)
5.8.1 动态 SQL 语句的划分 …… (129)
5.8.2 动态定义功能 …… (130)
5.8.3 动态操作功能 …… (130)
5.8.4 动态查询功能 …… (131)
本章小结 …… (132)
习题五 …… (132)
第 6 章 存储过程、触发器和数据完整性 …… (134)
6.1 存储过程 …… (134)
6.1.1 基本概念 …… (134)

6.1.2 创建和执行存储过程 …… (135)
6.1.3 存储过程的返回值和状态信息 …… (137)
6.1.4 存储过程的修改和删除 …… (137)
6.1.5 妥善利用存储过程 …… (138)
6.2 触发器及其用途 …… (138)
6.2.1 基本概念 …… (138)
6.2.2 建立触发器 …… (139)
6.2.3 插入视图和删除视图 …… (140)
6.2.4 删除类触发器 …… (140)
6.2.5 插入类触发器 …… (140)
6.2.6 更新类触发器 …… (141)
6.3 数据完整性 …… (142)
6.3.1 规则 …… (142)
6.3.2 默认值 …… (143)
6.3.3 用户定义数据完整性 …… (144)
本章小结 …… (145)
习题六 …… (145)
第7章 安全性 …… (147)
7.1 安全性概述 …… (147)
7.1.1 安全性措施的层次 …… (147)
7.1.2 数据库管理系统的安全功能 … (148)
7.1.3 自主存取控制与强制存取控制 …… (149)
7.1.4 SQL Server 的身份验证模式…… (149)
7.2 用户管理和角色管理 …… (151)
7.2.1 用户的分类 …… (151)
7.2.2 登录用户和数据库用户 …… (152)
7.2.3 用户管理 …… (152)
7.2.4 角色管理 …… (154)
7.2.5 SQL Server 的预定义角色…… (156)
7.3 权限管理 …… (157)
7.3.1 概述 …… (158)
7.3.2 授予权限 …… (158)
7.3.3 收回权限 …… (161)
7.3.4 禁止权限 …… (162)
7.3.5 角色与存取控制 …… (163)
7.4 其他安全问题 …… (163)
7.4.1 数据加密 …… (163)
7.4.2 审计 …… (164)
7.4.3 统计数据库 …… (164)
7.4.4 用户定义的安全性措施 …… (165)
本章小结 …… (166)
习题七 …… (166)
第8章 事务管理 …… (167)
8.1 事务 …… (167)
8.1.1 事务的概念 …… (167)
8.1.2 事务的性质 …… (168)
8.1.3 SQL 对事务的支持 …… (169)
8.2 并发控制 …… (171)
8.2.1 干扰问题 …… (171)
8.2.2 可串行性 …… (174)
8.2.3 封锁 …… (174)
8.2.4 死锁 …… (178)
8.2.5 隔离级别 …… (180)
8.2.6 封锁与隔离级别 …… (181)
8.3 恢复 …… (181)
8.3.1 数据库恢复概述 …… (181)
8.3.2 故障类型 …… (182)
8.3.3 备份类型 …… (183)
8.3.4 日志的概念 …… (183)
8.3.5 恢复模型 …… (184)
8.3.6 备份或转储 …… (184)
8.3.7 恢复或还原 …… (189)
本章小结 …… (192)
习题八 …… (192)
第9章 关系数据理论 …… (194)
9.1 基本概念 …… (194)
9.1.1 函数依赖 …… (194)
9.1.2 术语和符号 …… (195)
9.1.3 为什么要讨论函数依赖 …… (195)
9.1.4 模式分解 …… (196)
9.2 函数依赖的公理系统 …… (196)
9.2.1 Amstrong 公理的内容及正确性 …… (197)
9.2.2 Amstrong 公理的推论 …… (197)

9.2.3 逻辑蕴涵和闭包 …… (198)
9.2.4 公理的完备性 …… (199)
9.2.5 闭包的计算 …… (200)
9.2.6 函数依赖集的等价和最小化 … (201)
9.3 规范化 …… (203)
9.3.1 第一范式(1NF) …… (203)
9.3.2 第二范式(2NF) …… (204)
9.3.3 第三范式(3NF) …… (205)
9.3.4 BC 范式(BCNF) …… (206)
9.3.5 多值依赖与第四范式(4NF) … (207)
9.3.6 规范化小结 …… (210)
9.4 模式分解 …… (211)
9.4.1 模式分解的准则 …… (211)
9.4.2 3NF 无损连接和保持函数依赖算法 …… (213)
9.4.3 使分解后的关系模式数最少 … (214)
本章小结 …… (215)
习题九 …… (215)
第 10 章 数据库设计 …… (217)
10.1 完善 E – R 模型中的概念 …… (217)
10.1.1 弱实体 …… (217)
10.1.2 依赖联系 …… (217)
10.1.3 强制联系 …… (218)
10.2 数据库设计的过程 …… (218)
10.2.1 概念模型设计 …… (219)
10.2.2 逻辑数据库设计与规范化理论的应用 …… (220)
10.2.3 物理数据库设计 …… (221)
10.3 数据库设计工具 PowerDesigner …… (222)
10.3.1 PowerDesigner 简介 …… (222)
10.3.2 概念数据库设计 …… (223)
10.3.3 物理数据库设计 …… (225)
10.3.4 建立数据库 …… (227)
本章小结 …… (227)
习题十 …… (227)
第 11 章 面向对象数据库 …… (228)
11.1 新的数据库应用和新的数据类型 …… (228)
11.2 面向对象的数据模型 …… (229)
11.2.1 对象与类 …… (229)
11.2.2 对象的属性、方法和状态 …… (230)
11.2.3 对象的交互与消息 …… (230)
11.2.4 类的确定与划分 …… (231)
11.2.5 封装(encapsulation) …… (232)
11.2.6 继承性(inheritance) …… (232)
11.2.7 多态性(Polymorphism) …… (234)
11.2.8 对象标识符 …… (235)
11.3 对象关系数据库与对象数据库 … (235)
11.3.1 面向对象数据库的研究方法 …… (235)
11.3.2 ODMG 数据模型和 OQL …… (236)
11.4 面向对象数据库研究中的几个特点 …… (236)
11.4.1 面向对象模型的优势 …… (236)
11.4.2 面向对象数据库发展过程的特点 …… (237)
11.4.3 面向对象数据库系统的基本特征 …… (237)
11.4.4 面向对象数据库研究中存在的问题 …… (237)
11.5 RDBMS、ORDBMS 和 OODBMS 的比较 …… (238)
11.5.1 RDBMS 和 ORDBMS …… (238)
11.5.2 ORDBMS 和 OODBMS 的相似之处 …… (238)
11.5.3 ORDBMS 和 OODBMS 的不同之处 …… (238)
本章小结 …… (239)
习题十一 …… (239)
第 12 章 数据库应用的结构与开发环境 …… (240)
12.1 数据库应用开发概述 …… (240)
12.2 开放的客户端开发 …… (241)
12.2.1 开放式客户体系结构与 ODBC …… (241)
12.2.2 Visual FoxPro 作为客户端

开发工具……（243）
12.3　数据库应用开发工具 PowerBuilder ……（248）
12.3.1　PowerBuilder 的数据库接口 …（249）
12.3.2　PowerBuilder 应用的构成与开发方法……（251）
12.3.3　对象、属性与事件 ……（252）
12.3.4　PowerBuilder 的开发环境 ……（253）
12.3.5　数据库操作与事务对象……（255）
12.3.6　初步体验 PowerBuilder 的开发方法……（256）
12.4　基于组件的数据库开发 ……（265）
12.4.1　组件与组件模型……（266）
12.4.2　基于组件的开发……（267）
12.4.3　有关标准……（268）
12.4.4　使用 PowerBuilder 和 EAServer 开发组件应用……（269）
12.4.5　从 EAServer 组件访问数据库……（273）
12.5　Web 数据库应用开发……（274）
12.5.1　Web 数据库应用的一般结构和方法……（274）
12.5.2　使用 PowerBuilder 和 EAServer 建立 Web 应用 ……（275）
本章小结 ……（280）
习题十二 ……（280）
第 13 章　数据库应用系统的开发方法 ……（281）
13.1　概述 ……（281）
13.2　结构化生命周期方法 ……（281）
13.3　快速原型方法 ……（284）
13.4　面向对象方法 ……（285）
13.4.1　面向对象的分析和设计方法……（285）
13.4.2　UML 简介 ……（288）
13.5　客户/服务器应用规划综述 ……（290）
13.5.1　可以量化的需求分析……（291）
13.5.2　性能需求……（293）
13.5.3　并发需求……（294）
13.5.4　数据分布需求……（294）
13.5.5　恢复需求……（294）
13.5.6　安全问题……（295）
13.5.7　系统需求……（295）
本章小结 ……（296）
习题十三 ……（296）
第 14 章　分布式数据库与分布数据管理 ……（297）
14.1　什么是分布式数据库 ……（297）
14.2　分布式数据库的数据分布方式 …（299）
14.2.1　数据的分布方式……（299）
14.2.2　数据分片……（300）
14.3　分布式数据库的特点和目标 ……（300）
14.3.1　分布式数据库系统的基本目标……（301）
14.3.2　更新传播……（302）
14.3.3　分布式查询处理……（303）
14.3.4　目录表管理……（303）
14.3.5　分布式事务管理……（304）
14.3.6　其他……（304）
14.4　SQL Server 的分布数据管理 ……（305）
14.4.1　概述……（305）
14.4.2　SQL Server 的复制及其术语……（306）
14.4.3　复制模型……（307）
14.4.4　复制的类型及复制的实现技术……（310）
本章小结 ……（315）
习题十四 ……（315）
第 15 章　数据仓库 ……（316）
15.1　什么是数据仓库 ……（316）
15.2　数据仓库的结构 ……（317）
15.3　数据仓库系统 ……（318）
15.4　建立数据仓库 ……（319）
15.5　实现数据仓库的数据库环境 ……（319）
15.6　数据仓库与决策支持 ……（320）
15.7　SQL Server 的数据仓库解决方案 ……（321）

需要全本请访问 www.ertongbook.com

15.8 将 SQL Server 的 OLTP 数据转换到 OLAP 数据仓库 …… (322)
15.8.1 合并数据 …… (323)
15.8.2 清理数据 …… (323)
15.8.3 聚合数据 …… (323)
15.8.4 将数据组织到多维数据集中 …… (323)
15.8.5 转换阶段和数据仓库组件 …… (323)
15.9 软件开发商的数据仓库解决方案 …… (324)
本章小结 …… (325)
习题十五 …… (326)
第 16 章 数据库研究和应用的新领域 …… (327)
16.1 概述 …… (327)
16.2 数据库技术与其他技术相结合产生的新领域 …… (327)
16.2.1 并行数据库 …… (328)
16.2.2 主动数据库 …… (328)
16.2.3 知识库 …… (330)
16.2.4 多媒体数据库 …… (331)
16.2.5 模糊数据库 …… (331)
16.3 面向应用领域的数据库新技术 … (332)
16.3.1 工程数据库 …… (332)
16.3.2 空间数据库 …… (333)
本章小结 …… (334)
习题十六 …… (334)
附录 上机实验 …… (335)
实验 1 建立数据库 …… (335)
实验 2 建立表和数据完整性 …… (337)
实验 3 SQL 数据操作 …… (339)
实验 4 SQL 数据查询 …… (340)
实验 5 视图的定义和操作 …… (342)
实验 6 SQL 的宿主使用方式 …… (342)
实验 7 存储过程 …… (343)
实验 8 触发器 …… (344)
实验 9 用户管理和权限管理 …… (344)
实验 10 并发控制 …… (345)
实验 11 数据库建模工具 …… (346)
实验 12 使用 Visual FoxPro 做客户端开发 …… (346)
实验 13 使用 PowerBuilder 做客户端开发 …… (347)
主要参考书目 …… (348)

第 1 章　绪　　论

数据库是数据管理的工具。数据管理经历了从手工管理阶段、文件管理阶段到数据库管理阶段的变迁。数据库技术自产生以来便为广大用户所接受,并获得了广泛的应用。无论是对学生,还是对广大计算机用户,数据库都已不再陌生。

1.1　什么是数据库

数据库,顾名思义就是存放数据的仓库,只不过这个仓库不是一般意义上的库房,它是存储在计算机存储设备上的相互关联的数据集合。

从中文字面上来看,数据库和数据仓库完全是一个概念,这种想当然的理解是不准确的。数据库对应的英文单词是 DataBase,如果直译是数据基地的意思;而数据仓库则另有其词——DataWarehouse。所以数据库和数据仓库不是同义词,数据仓库是在数据库技术的基础上发展起来的一个新的应用领域(有关数据仓库的内容将在第 15 章介绍)。

数据库技术发展到今天已经是一门非常成熟的技术。它始于 20 世纪 60 ~ 70 年代,然而无论是从数据库的技术水平,还是从数据库的应用水平,今天和过去都不可同日而语,但是数据库的最基本特征没有变,概括起来应该包括如下几个特征:

- 数据库是相互关联的数据的集合;
- 数据库用综合的方法组织数据,保证尽可能高的访问效率;
- 数据库具有较小的数据冗余,可供多个用户共享;
- 数据库具有较高的数据独立性;
- 数据库具有安全控制机制,能够保证数据的安全、可靠;
- 数据库允许并发地使用,能有效、及时地处理数据,并能保证数据的一致性和完整性。

1.1.1　相互关联的数据的集合

数据库中的数据不是孤立的,数据与数据之间是相互关联的。也就是说,在数据库中不仅要能够表示数据本身,还要能够表示数据与数据之间的联系。

比如在学籍管理中,有学生和课程两类数据,在数据库中除了要存放这两类数据之外,还要存放哪些学生选修了哪些课程或哪些课程由哪些学生选修这样的信息,这就反映了学生数据和课程数据之间的联系。

1.1.2 用综合的方法组织数据

数据库能够根据不同的需要按不同的方法组织数据,比如可以有顺序组织方法、索引组织方法、倒排数据库组织方法等。这样做的目的就是要最大限度地提高用户或应用程序访问数据的效率。

1.1.3 低冗余与数据共享

由于在数据库技术之前,数据文件都是独立的,所以任何数据文件都必须含有满足某一应用的全部数据。比如,人事部门有一个职工文件,教育部门也有一个职工文件,人事部门的职工文件的记录格式是:

职工基本情况	有关人事管理的数据

教育部门的职工文件的记录格式是:

职工基本情况	有关教育培训的数据

这样在两个部门的职工文件中都有‘职工基本情况’的数据,也就是说,这一部分数据是重复存储的,如果还有第三、第四个部门也有类似的职工文件,那么重复存储所造成的空间浪费是很大的。在数据库中,可以共享类似‘职工基本情况’这样的共用数据,从而降低数据的冗余度。降低数据冗余不仅可以节省存储空间,更重要的是可以保证数据的一致性。

1.1.4 数据具有较高的独立性

数据独立性是指数据的组织和存储方法与应用程序互不依赖、彼此独立的特性。在数据库技术之前,数据文件的组织方式和应用程序是密切相关的,当改变数据结构时相应的应用程序也必须随之修改,这样就大大增加了应用程序的开发代价和维护代价。而数据库技术却可以使数据的组织和存储方法与应用程序互不依赖,从而大大降低应用程序的开发代价和维护代价。

1.1.5 保证数据的安全、可靠

数据库技术要能够保证数据库中的数据是安全、可靠的。数据库要有一套安全机制,以便可以有效地防止数据库中的数据被非法使用或非法修改 数据库还要有一套完整的备份和恢复机制,以便保证当数据遭到破坏时(软件或硬件故障引起的),能立刻将数据完全恢复,从而保证系统能够连续、可靠地运行。

1.1.6 最大限度地保证数据的正确性

保证数据正确的特性在数据库中称之为数据完整性。在数据库中可以通过建立一些约束条件保证数据库中的数据是正确的。比如,某学生的年龄是 20 岁,当误输入为 2 岁或 200 岁时,数据库能够主动拒绝这类错误。

1.1.7 数据可以并发使用并能同时保证数据的一致性

数据库中的数据是共享的,并且允许多个用户同时使用相同的数据,这就要求数据库能够协调一致,保证各个用户之间对数据的操作不发生矛盾和冲突,即在多个用户同时使用数据库时,能够保证数据的一致性和正确性。

以上概括介绍了数据库的主要特性或特征,在后续的章节中对它们会作出更详细的解释,将会回答一些为什么……?什么是……?如何做……?等问题。

1.2 数据库管理系统

上一节提到的数据库的各种功能和特性,并不是数据库中的数据固有的,它是靠管理或支持数据库的系统软件——数据库管理系统(DataBase Management System,简称 DBMS)提供的。一个完备的数据库管理系统应该具备上一节提到的各种功能,它的任务就是对数据资源进行管理,并且使之能为多个用户共享,同时还能保证数据的安全性、可靠性、完整性、一致性,还要保证数据的高度独立性。

具体来说,一个数据库管理系统应该具备如下功能:

1) 数据库定义功能:可以定义数据库的结构和数据库的存储结构,可以定义数据库中数据之间的联系,可以定义数据的完整性约束条件和保证完整性的触发机制等;

2) 数据库操纵功能:可以完成对数据库中数据的操纵,可以装入、删除、修改数据,可以重新组织数据库的存储结构,可以完成数据库的备份和恢复等操作;

3) 数据库查询功能:可以以各种方式提供灵活的查询功能,使用户可以方便地使用数据库中的数据;

4) 数据库控制功能:可以完成对数据库的安全性控制、完整性控制、多用户环境下的并发控制等各方面的控制;

5) 数据库通信功能:在分布式数据库或提供网络操作功能的数据库中还必须提供数据库的通信功能。

1.3 数据库管理和数据库管理员

数据库的使用会改变企事业单位的管理方式,最简单的原因就是很多部门或用户把他们的数据集中放在了数据库中,这自然会带来很多好处,如:

1) 由于消灭了数据重复和数据不一致性,数据将更加可靠、实用;

2) 由于数据独立性,减少了程序的维护代价;

3) 为数据的特定查询请求提供了快速响应;

等等。

然而,如果认为仅仅靠对称之为数据库或数据库管理系统的特定技术"黑箱"进行投资就可以得到这些好处,那将是错误的。因为要把众多部门或用户的数据放在同一个数据库中,那就必须要考虑这些数据会不会产生冲突?会不会有越权使用数据的情况发生?重要的数据会不会丢失?诸如此类的问题是用户非常关心的。为了解决这些问题就要有一个数据库管理的部门,这个部门将负责和数据库管理有关的所有工作,也就是说,负责数据库的管理。

从事数据库管理工作的人员称之为数据库管理员(DataBase Administrator,简称为 DBA),DBA 有大量的工作要做,既有技术方面的工作,又有管理方面的工作。DBA 要和各种人员打交道,包括普通应用人员、计算机专业人员、企事业管理人员等,DBA 要参加数据库开发和使用的全部工作,总的来说 DBA 的工作可以概括如下:

1) 首先在数据库规划阶段要参与选择和评价与数据库有关的计算机软件和硬件,要与数据库用户共同确定数据库系统的目标和数据库应用需求,要确定数据库的开发计划;

2) 在数据库设计阶段要负责数据库标准的制定和共用数据字典的研制,要负责各级数据库模式的设计,负责数据库安全、可靠方面的设计;

3) 在数据库运行阶段首先要负责对用户进行数据库方面的培训;负责数据库的转储和恢复;负责对数据库中的数据进行维护;负责监视数据库的性能,并调整、改善数据库的性能,提高系统的效率;继续负责数据库安全系统的管理;在运行过程中发现问题、解决问题。

由此可见,数据库管理员负责数据库的全面管理工作。数据库管理员的工作是十分繁重和十分重要的,要求作为 DBA 的人员除了要掌握一定的数据处理、数据库技术之外,还应有处理好人际关系的优良素质和能力。在一个企事业中,特别是对一个规模较大的数据库,不能指望一两个人能够完成数据库管理工作,所以 DBA 通常是指数据库管理部门。

数据库管理员这一职位是非常重要、非常关键的,任何一个数据库系统如果没有数据库管理员或相应的部门执行 DBA 的职责,数据处理自动化就难以成功,数据库就会失去统一的管理和控制,从而造成数据库的混乱。所以,在开发数据库系统时,一开始就应设置数据库管理员的职位或相应的机构,要明确 DBA 的责任,也要保证 DBA 的权限。DBA 一定是

那些懂得和掌握数据库全局工作、并作为设计和管理数据库的核心人员,在数据库系统的开发过程中,DBA 应该发挥极其重要的作用。

1.4 数据库系统

前面介绍了数据库的各种特征和管理数据库的系统软件 DBMS,那么什么是数据库系统呢?

简单来说,数据库系统就是基于数据库的计算机应用系统。这样一个系统除了包括:

- 以数据为主体的数据库;
- 管理数据库的系统软件 DBMS;

此外还要包括:

- 支持数据库系统的计算机硬件环境和操作系统环境;
- 管理和使用数据库系统的人,特别是负责设计、维护数据库的技术人员——数据库管理员;
- 方便使用和管理系统的各种技术说明书和使用说明书。

至此可以看出,数据库、数据库管理系统和数据库系统是三个不同的概念,数据库强调的是数据,数据库管理系统是系统软件,而数据库系统强调的是系统。

1.5 数据库的过去、现在和未来

数据库的核心任务是数据管理,它包括数据的分类、组织、编码、存储、检索和维护等,但并不是一开始就有数据库技术,在计算机诞生的初期,计算机主要用于科学计算,虽然此时同样有数据管理的问题,但这时的数据管理是以人工的方式进行的,后来发展到文件系统,再后来才是数据库。也就是说,数据管理经历了人工管理阶段、文件系统阶段和数据库系统阶段。

1.5.1 人工管理阶段

人工管理阶段是指计算机诞生的初期(20 世纪 50 年代中期以前)。这个时期的计算机技术,从硬件看还没有磁盘这样的可直接存取的存储设备,从软件看没有操作系统,更没有管理数据的软件。所以这个时期数据管理的特点是:

1) 数据不保存。因为计算机主要用于科学计算,一般也不需要长期保存数据,只是在要完成某一个计算或课题时才将数据输入,然后不仅原始数据不保存,计算结果也不保存。

2) 还没有文件的概念。这个时期的数据组织必须由每个程序的程序员自行组织和安

需要全本请访问: www.ertongbook.com