

高等院校计算机教育系列教材

0 10 01 101 10 0
010 010 010 011 01
0101100 01 01 010
011 01 01 010 0 0
01 010 00 010 0 0
010 0 010
100 0 00 101 100 0 00 1
10011010
010 0101 0
0101100



SQL Server

数据库原理及应用

(第2版)

张莉 主编

- 知识点新，突出实践教学，强化能力培养
- 理论知识+感性认识+动手实践，完美结合
- 内容简明扼要，突出知识要点
- 以实用为宗旨，实例丰富，用实例引导读者模仿学习

0101100 01
0101101000
0101101010 1 01 10
011 0 100 1 10100 10100010
0110 01 101 10 0 10
011 01 01
01 010 00
010 0 010

赠送
电子课件

清华大学出版社



高等院校计算机教育系列教材

SQL Server 数据库原理及应用 (第2版)

张 莉 主编

清华大学出版社

北 京

内 容 简 介

本书选用以网络环境为背景,目前广泛使用的 SQL Server 关系型数据库管理系统为数据库系统平台,全面系统地介绍数据库技术的基础理论、数据库系统的设计方法、现代数据库系统的开发应用等内容。包括了数据库设计理论、设计方法、Transact-SQL 编程、网络数据库系统安全、面向对象数据库技术、并行数据库系统和分布式数据库系统、数据仓库、数据挖掘等内容。本书第 2 版仍以广泛应用的 SQL Server 列举基本案例,通过 ODBC 和 ADO 数据库访问技术,把 ASP 技术或 JSP 技术制作的 Web 网页与数据库之间连接起来,技术方法结构清晰,基本覆盖了现代数据库技术应用知识要点。

本书在内容编排上系统全面,力求重点突出、新颖实用、可读性强,是系统学习和掌握现代数据库技术理论与实践的理想教材或参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

SQL Server 数据库原理及应用/张莉主编. —2 版. —北京:清华大学出版社,2009.5
(高等院校计算机教育系列教材)
ISBN 978-7-302-19735-5

I. S… II. 张… III. 关系数据库—数据库管理系统, SQL Server—高等学校—教材 IV. TP311.138
中国版本图书馆 CIP 数据核字(2009)第 038855 号

责任编辑:彭欣

封面设计:山鹰工作室

版式设计:杨玉兰

责任校对:王晖

责任印制:何芊

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京季蜂印刷有限公司

装 订 者:三河市李旗庄少明装订厂

经 销:全国新华书店

开 本:185×260 印 张:26.25 字 数:635 千字

版 次:2003 年 1 月第 1 版 2009 年 5 月第 2 版

印 次:2009 年 5 月第 1 次印刷

印 数:46501~50500

定 价:36.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:032017-01

前 言

数据库技术是现代信息技术的重要组成部分。随着计算机技术的发展与广泛应用，无论是在数据库技术的基础理论、数据库技术应用、数据库系统开发，还是在数据库商品软件的推出方面，都有了长足的、迅速的进步与发展。

随着计算机网络技术的日益发展与广泛应用，支持网络技术的数据库系统也不断推陈出新，推动着数据库技术的发展和應用。新的数据库系统，在继承了传统数据库技术的同时，不仅在数据库系统性能和功能的完善方面，而且在数据库系统管理的易操作性、安全性、易维护性方面以及系统的开发技术方面，都较以前的数据库系统产品有了很大的发展与进步，并且在现代信息技术系统开发和新技术应用的理念方面，都更加符合当今社会对信息技术发展的需要与应用。尽管数据库系统软件不断推陈出新，但数据库技术的核心仍然万变不离其宗，保持着传统的数据库基础理论和设计方法。

SQL Server 是在 Microsoft 公司与 Sybase 公司合作的基础上推出的关系型数据库系统，它一经推出便迅速成为 Windows NT 网络方案的首选数据库系统。而新一代的 SQL Server 2000 着眼于 Internet 背景下的网络数据库的应用与开发，是 Microsoft 公司继 Windows 2000 之后的一部很有影响力的系统软件。它除了具有现代数据库管理系统所具备的各种强大的功能外，在支持 Internet 技术方面，还提供了 HTTP 的支持，并为 Web 用户提供了功能完善的网络数据管理与数据分析解决方案以及开发应用数据仓库和电子商务的便利环境。

本书选用网络环境广泛使用的 SQL Server 作为数据库系统平台，系统地介绍网络数据库技术的基础理论、实现方法、设计过程与开发应用等。在内容编排上由理论到实践、从技术基础到综合开发与新技术应用，使学习过程能够由浅入深、由点到面、由具体实践到系统提高逐步过渡。本书在技术基础内容方面包括传统数据模型的建立、数据库系统结构组成、关系模型的数据结构、数据运算、数据库设计理论与设计方法、数据库系统分析与设计、标准通用查询语言 SQL 等；在数据库系统开发和应用方面，介绍了 SQL Server 2000 关系数据库系统结构组成、管理工具、语言与编程接口，以及 SQL Server 2000 数据管理、SQL Server 数据库操作等应用开发功能。

就数据库编程来说，Transact-SQL 是标准 SQL 语言的增强版本，用 Transact-SQL 语言可以直接和数据库引擎进行交互，书中系统地介绍了 Transact-SQL 程序设计应用与开发。数据库系统安全是数据库技术的重要组成部分，面向对象技术广泛应用于现代数据库技术，并行数据库系统和分布式数据库系统是现代数据库的重要内容，这些内容书中均有完整的介绍。最后介绍了 SQL Server 2000 作为网络编程的后台数据库，可以通过 ODBC 和 ADO 这两种数据库访问技术，用于 ASP 技术或 JSP 技术制作的 Web 网页与数据库之间的连接。

本书符合现代信息技术的教育理念，旨在引导和培养读者的综合实践与创新能力。通过本书，读者不仅可以系统地掌握数据库技术的基础理论、设计方法，还可以系统地掌握现代数据库系统的开发技术。

本书由张莉教授根据多年从事相关教学及科研成果主持编写。参加本书编写的其他教师在数据库技术原理与应用方面，也都拥有着多年科研开发与教学实践经验。特别是具有丰富理论与扎实经验的研究生和本科生对案例的设计与调试，更为本书增加了新颖性、实用



性和可读性。本书在编写过程中参考了大量的文献资料，是教学、科研与实践协作的结果。

本书由张莉教授担任主编，具体参加修订编写的人员还有王永进、杨丽丽、陈雷、王强、董莉、赵文坊、段清玲、张筠、李琼飞、马钦、王德宇、马志远、张一飞、张媛媛、朱杨等。本书涉及数据库技术应用原理的新内容较多，修订版经过多年教学实践检验，其内容编排符合现代教育认知理念，取得了良好的教学效果和社会反响。第 2 版仍然以应用研发为目的，理论结合实践，注重案例分析。

本书整个编写过程中力求系统简练，但难免有遗漏、不足之处，恳请学界同仁和广大读者及时批评指正，并提出宝贵的意见和建议。

编者



目 录

第 1 章 数据库系统概述1	
1.1 概述.....1	
1.1.1 数据库技术与信息技术.....1	
1.1.2 数据库技术的应用及特点.....2	
1.1.3 SQL Server 数据库系统与 网络分布式操作.....2	
1.1.4 网络数据库系统编程.....3	
1.2 数据库技术的发展.....3	
1.2.1 人工管理阶段.....4	
1.2.2 文件系统阶段.....5	
1.2.3 数据库系统阶段.....6	
1.2.4 高级数据库系统阶段的 应用与发展.....8	
1.2.5 数据库系统发展阶段的 划分.....10	
1.3 信息描述与数据模型.....11	
1.3.1 数据模型及其三要素.....12	
1.3.2 概念模型及其表示方法.....13	
1.3.3 实体联系模型.....16	
1.4 三种常见的数据模型.....18	
1.4.1 层次模型.....18	
1.4.2 网状模型.....19	
1.4.3 关系模型.....20	
1.5 数据库系统.....21	
1.5.1 数据库系统的 三级模式结构.....21	
1.5.2 数据库系统的组成.....23	
1.5.3 数据库管理系统的功能.....24	
1.5.4 数据库系统的视图.....25	
1.5.5 数据库系统的访问.....26	
思考练习题.....27	
第 2 章 关系数据库系统模型29	
2.1 关系模型的基本概念.....29	
2.1.1 关系模型.....29	
2.1.2 数据结构.....32	
2.1.3 关系操作.....32	
2.1.4 关系的完整性.....33	
2.2 关系运算.....35	
2.2.1 传统的集合运算.....35	
2.2.2 专门的关系运算.....37	
2.2.3 关系数据检索实例.....42	
2.3 关系演算.....44	
2.3.1 元组关系演算语言.....45	
2.3.2 域关系演算语言 QBE.....50	
思考练习题.....53	
第 3 章 关系数据库的设计理论55	
3.1 实体类型的属性关系.....55	
3.2 数据的函数依赖.....56	
3.2.1 函数依赖.....56	
3.2.2 依赖的逻辑蕴涵.....57	
3.2.3 关键字.....58	
3.3 关系数据库模式的规范化理论.....59	
3.3.1 关系模式规范化设计.....60	
3.3.2 第一范式(1NF).....61	
3.3.3 第二范式(2NF).....62	
3.3.4 第三范式(3NF).....63	
3.3.5 BCNF 范式.....64	
3.3.6 多值函数依赖关系的 第四范式(4NF).....65	
3.3.7 第五范式(5NF).....67	
3.4 关系模式的分解算法.....69	
3.4.1 关系模式的分解.....69	
3.4.2 无损连接.....69	
3.4.3 无损连接的检验.....70	
3.4.4 保持函数依赖的分解.....72	
思考练习题.....73	
第 4 章 数据库设计方法75	
4.1 数据库系统的设计过程.....75	

4.2 需求分析.....77	思考练习题.....119
4.2.1 收集用户需求.....77	
4.2.2 分析用户需求.....78	
4.2.3 撰写需求说明书.....82	
4.3 概念设计.....83	
4.3.1 概念模型的特点及设计方法...83	
4.3.2 局部概念模型的设计.....84	
4.3.3 合成全局概念模型.....86	
4.4 逻辑结构的设计.....89	
4.4.1 概念模型向关系模型的转换...90	
4.4.2 应用规范化理论优化 逻辑模型.....92	
4.4.3 设计用户子模式.....93	
4.5 物理设计.....93	
4.5.1 影响物理设计的因素.....94	
4.5.2 确定关系模式的存取方法.....94	
4.5.3 确定数据库的存储结构.....95	
4.6 数据库的实施、运行和维护.....96	
4.6.1 数据库的实施.....97	
4.6.2 数据库的试运行.....97	
4.6.3 数据库的运行和维护.....98	
思考练习题.....99	
第5章 结构化查询语言 SQL.....101	
5.1 SQL 结构化查询语言.....101	
5.1.1 SQL 语言的特点.....101	
5.1.2 SQL 数据库的 三级模式结构.....102	
5.2 SQL 的数据描述与操作.....103	
5.2.1 定义基表.....103	
5.2.2 定义索引.....105	
5.3 SQL 语言的数据检索功能.....106	
5.4 SQL 语言的数据更新功能.....111	
5.4.1 修改语句.....111	
5.4.2 插入语句.....112	
5.4.3 删除语句.....112	
5.4.4 SQL 语言对视图的操纵.....113	
5.5 SQL 的数据控制功能.....116	
5.5.1 SQL 的数据控制方法.....117	
5.5.2 SQL Server 数据操作.....118	
第6章 SQL Server 关系 数据库系统.....121	
6.1 SQL Server 系统概述.....121	
6.1.1 SQL Server 2000 的 功能与特点.....121	
6.1.2 SQL Server 2000 产品系列.....125	
6.2 SQL Server 2000 的数据库结构.....126	
6.2.1 数据库逻辑体系结构.....127	
6.2.2 数据库物理体系结构.....128	
6.2.3 客户机和服务器的 交互方式.....133	
6.2.4 SQL Server 2000 分布式查询.....134	
6.3 SQL Server 客户机/服务器体系 结构及软件组成.....134	
6.3.1 SQL Server 2000 客户端组件.....135	
6.3.2 SQL Server 2000 服务器端组件.....137	
6.3.3 SQL Server 2000 对 XML 的支持.....139	
6.3.4 SQL Server 2000 传输组件.....140	
6.4 SQL Server 2000 的 系统管理与开发.....142	
6.4.1 分布式管理.....142	
6.4.2 管理工具.....142	
6.4.3 语言与编程接口.....144	
思考练习题.....145	
第7章 SQL Server 2000 数据库 应用与管理.....147	
7.1 SQL Server 2000 的安装.....147	
7.1.1 环境要求.....147	
7.1.2 SQL Server 2000 的安装.....149	
7.1.3 从 SQL Server 7.0 升级.....153	
7.2 启动和登录数据库服务器.....154	
7.2.1 数据库服务器的启动.....154	
7.2.2 登录数据库服务器.....156	

7.3 SQL Server 2000 数据管理	157	8.5 对数据库中数据表的操作.....	205
7.3.1 SQL Server 2000 的 数据定义语言	157	8.5.1 按条件查询数据	205
7.3.2 SQL Server 2000 的 数据操作语言	160	8.5.2 使用通配符和子查询.....	206
7.4 SQL Server 数据库操作	162	8.5.3 数据连接	207
7.4.1 用 Enterprise Manager 创建数据库	162	8.5.4 数据排序	208
7.4.2 用 Enterprise Manager 定义和 管理数据表	163	8.5.5 数据统计	209
7.4.3 用视图创建向导创建视图	166	8.6 数据库索引	210
7.4.4 用工具和向导创建 数据库图表	169	8.6.1 索引的概念	211
7.4.5 用 Enterprise Manager 创建存储过程和触发器	170	8.6.2 索引类型	211
7.4.6 数据库的安全管理	174	8.6.3 索引的创建与删除.....	212
思考练习题	178	8.6.4 索引的设计	213
第 8 章 Transact-SQL 语言结构	179	8.7 数据库视图	214
8.1 T-SQL 语言概述	179	8.7.1 视图的创建与删除.....	215
8.1.1 T-SQL 语言特点	179	8.7.2 视图与数据修改	216
8.1.2 T-SQL 附加语言	180	8.7.3 查询和视图的比较.....	217
8.1.3 SQL Server Query Analyzer 工具	185	8.8 函数	217
8.1.4 osql 实用程序.....	186	8.8.1 聚合函数	217
8.2 Transact-SQL 数据类型.....	187	8.8.2 数学函数	222
8.3 Transact-SQL 语言运算符.....	193	8.8.3 字符串函数	223
8.3.1 算术运算符	193	8.8.4 系统函数	228
8.3.2 赋值运算符	195	8.8.5 日期和时间函数	231
8.3.3 位运算符	195	8.8.6 文本和图像函数	233
8.3.4 比较运算符	196	8.9 SQL Server 2000 数据库复制.....	235
8.3.5 逻辑运算符	197	8.9.1 数据库复制的技术体系.....	235
8.3.6 一元运算符	197	8.9.2 数据库复制原理	238
8.3.7 运算符的优先顺序	198	8.9.3 复制的拓扑结构	240
8.4 数据表的创建与维护	198	8.9.4 事务复制的实现	242
8.4.1 创建数据表	198	8.9.5 SQL Server 数据备份与 恢复	248
8.4.2 删除数据表	201	思考练习题	254
8.4.3 向数据表中插入记录	202	第 9 章 Transact-SQL 程序设计	255
8.4.4 从数据表中删除记录	203	9.1 T-SQL 的变量定义和 流程控制语句	255
8.4.5 修改和更新记录	204	9.1.1 变量的定义和使用.....	255
		9.1.2 流程控制语句	258
		9.2 存储过程	265
		9.2.1 存储过程概述	266
		9.2.2 创建存储过程	267
		9.2.3 调用存储过程	269



9.2.4 修改存储过程	271	11.3 数据仓库	318
9.2.5 删除存储过程	273	11.3.1 从数据库到数据仓库	318
9.2.6 存储过程的优化	273	11.3.2 数据仓库的定义	320
9.3 用户自定义的数据类型	274	11.3.3 数据仓库系统的结构	320
9.4 触发器的应用	275	11.3.4 数据仓库的数据组织	322
9.4.1 SQL Server 触发器的用途	276	11.3.5 数据仓库的数据模型	324
9.4.2 创建触发器	276	11.3.6 SQL Sever 2000 的 数据仓库组件	325
9.4.3 使用触发器	278	11.3.7 SQL Server 2000 数据仓库的建设流程	327
9.4.4 查看触发器的信息	283	11.4 数据挖掘	329
9.4.5 修改和删除触发器	285	11.4.1 数据挖掘概述	329
9.5 事务和锁	286	11.4.2 数据挖掘模型	330
9.5.1 事务的概念	286	11.4.3 使用决策树模型从关系 数据库中挖掘数据	332
9.5.2 事务编程	288	11.4.4 应用聚类分析模型 挖掘数据	336
9.5.3 锁	290	思考练习题	336
9.5.4 使事务更有效地工作	292	第 12 章 网络数据库编程	338
思考练习题	293	12.1 网络数据库的连接	338
第 10 章 数据库的安全及其维护	295	12.1.1 ODBC 的管理	338
10.1 SQL Server 的安全性管理	295	12.1.2 通过 ODBC 连接数据库	345
10.2 SQL Server 数据库的 安全性管理与控制	297	12.1.3 ADO 简介	350
10.3 SQL Server 的安全体系结构和 安全认证	302	12.2 ASP 编程	358
10.4 SQL Server 的权限管理	304	12.2.1 ASP 概述	358
思考练习题	310	12.2.2 ASP 的内置对象	361
第 11 章 现代数据库技术	311	12.2.3 ASP 的内置组件	374
11.1 面向对象的数据库系统	311	12.2.4 ASP 实现对数据库的 访问	378
11.1.1 面向对象的数据模型	312	12.3 JSP 编程	381
11.1.2 面向对象数据库语言	313	12.3.1 JSP 概述	381
11.1.3 面向对象数据库的 模式演进	314	12.3.2 JSP 编程概要	382
11.1.4 面向对象程序设计	315	12.3.3 JSP 数据库访问技术	388
11.2 并行数据库系统和分布式 数据库系统	315	思考练习题	392
11.2.1 并行数据库系统	315	参考答案	394
11.2.2 分布式数据库系统	317		

第 1 章 数据库系统概述

数据库技术是现代信息科学与技术的重要组成部分，是计算机数据处理与信息管理系统核心。数据库技术研究和解决在计算机信息处理过程中如何有效地组织和存储大量数据的问题，以及在数据库系统中减少数据存储冗余、实现数据共享、保障数据安全以及高效地检索数据和处理数据的问题。

随着计算机技术与网络通信技术的发展，数据库技术已经成为信息社会中对大量数据进行组织与管理的重要手段及软件技术，以及网络信息化管理系统的基础。本章主要介绍数据库技术的应用与发展、网络数据库特点，以及数据模型的建立和数据库系统结构等内容，是学习和掌握现代数据库技术的基础。

1.1 概 述

从 20 世纪 60 年代末开始到现在，数据库技术已经发展了几十年。在计算机技术发展的过程中，人们在数据库技术的理论研究和系统开发方面都取得了辉煌的成就，而且已经展开对新一代数据库系统的深入研究。数据库系统已经成为现代计算机系统的重要组成部分。

从 20 世纪 70 年代后期开始，数据库作为计算机专业的一门重要课程进入国外各个大学的课堂；我国高等院校从 20 世纪 80 年代开始，也把数据库作为计算机专业本科生和研究生的主要课程。并且随着计算机技术在国民生产中地位的逐年上升，数据库技术作为计算机技术的重要分支，其重要性越来越被人们所广泛认识。时至今日，数据库技术已经成为数据处理的公用支撑技术，数据库系统也早已遍布政府机关、社会团体和企业单位，存储着他们赖以正常运转的数据资源，从而显著地提高了工作效率和质量，产生出巨大的社会效益和经济效益。

1.1.1 数据库技术与信息技术

信息技术(Information Technology, IT)是当今使用频率最高的名词之一，它随着计算机技术在工业、农业以及日常生活中的广泛应用，已经被越来越多的个人和企业作为自己赶超世界潮流的标志之一。而数据库技术则是信息技术中一个重要的支撑。没有数据库技术，人们在浩瀚的信息世界中将会手足无措。

数据库技术是计算机科学技术的一个重要分支。从 20 世纪 50 年代中期开始，计算机应用从科学研究部门扩展到企业管理及政府行政部门，人们对数据处理的要求也越来越高。1968 年，世界上诞生了第一个商品化的数据管理系统 IMS(Information Management System)，从此，数据库技术得到了迅猛发展。在互联网日益被人们接受的今天，Internet 又使数据库



技术、知识、技能的重要性得到了充分的放大。现在,数据库已经成为信息管理、办公自动化、计算机辅助设计等应用的主要软件工具之一,帮助人们处理各种各样的信息数据。

1.1.2 数据库技术的应用及特点

数据库最初在大公司或大机构中被用作大规模事务处理的基础。后来随着个人计算机的普及,数据库技术被移植到 PC(Personal Computer, 个人计算机)上,最初用于数据库单用户模式。接着,由于 PC 机在工作组内连成网,数据库技术就移植到工作组级。现在,数据库已经应用于 Internet 和内联网。

20 世纪 60 年代中期,数据库技术是用来解决文件处理系统问题的。当时的数据库处理还很脆弱,常常出现应用不能提交的情况。20 世纪 70 年代关系模型的诞生为数据库专家提供了构造和处理数据库的标准方法,推动了关系数据库的发展和应用。1979 年,Ashton-Tate 公司引入了微机产品 dBase II,并称之为关系数据库管理系统,从此数据库技术移植到了个人计算机上。20 世纪 80 年代中期到后期,终端用户开始使用局域网技术将独立的计算机连接成网络,终端之间可以共享数据库,形成了一种新型的多用户数据处理,称为客户机/服务器数据库结构。现在,数据库技术正在被用来同 Internet 技术相结合,以便在机构内联网、部门局域网甚至 WWW 上发布数据库数据。

1.1.3 SQL Server 数据库系统与网络分布式操作

SQL Server 是由 Sybase、Microsoft 和 Ashton-Tate 联合开发的应用于 OS/2 系统上的数据库系统,1988 年正式投入使用。1990 年,Ashton-Tate 公司退出了 SQL Server 的开发,1994 年,Sybase 公司也将重点投入到 UNIX 版本的 SQL Server 开发上,而 Microsoft 公司则致力于将 SQL Server 移植到 NT 平台上。1996 年,Microsoft 公司推出了 SQL Server 6.5;1998 年,升级到 7.0 版本;到了 2000 年,Microsoft SQL Server 2000 面世,性能稳定,广为使用。目前升级到 Microsoft SQL Server 2005。

SQL Server 是基于 SQL 客户/服务器(C/S)模式的数据库系统,建立在 Microsoft Windows NT 平台上,可以提供强大的企业数据库管理功能。

Microsoft SQL Server 2000 数据库系统是在 Windows NT 环境下开发的一种全新的关系型数据库系统,是发展最快的关系数据库,占世界市场份额的 38%。SQL Server 2000 具有大型数据库的一些基本功能,支持事务处理功能、支持数据库加密、设置用户组或用户的密码和权限等。它为用户提供了大规模联机事务处理(OLTP)、数据仓库和电子商务应用程序所需的最新的优秀数据库平台。

SQL Server 2000 是一个具备完全 Web 支持的数据库产品,提供了以 Web 标准为基础的扩展数据库编程功能,提供了对可扩展标记语言(XML)的核心支持以及在 Internet 上和防火墙外进行查询的能力。使用 SQL Server 2000 可以获得非凡的可伸缩性和可靠性。通过向上伸缩和向外扩展的能力,SQL Server 2000 可以满足苛刻的电子商务和企业应用程序的要求。它还是 Microsoft .NET Enterprise Server 的数据管理与分析中枢,并包括加速从概念到最后交付开发过程的工具。

1.1.4 网络数据库系统编程

有了数据库的支持,扩展网页的功能,设计交互式页面,构造功能强大的后台管理系统以及更新、维护网站都将变得轻而易举。随着网络应用的深入,Web 数据库技术将日益显现出其重要地位。下面简单介绍 Web 数据库开发的相关技术。

1. 通用网关接口(CGI)编程

通用网关接口(Common Gateway Interface, CGI)是一种通信标准,它的任务是接受客户端的请求,经过辨认和处理,生成 HTML 文档并重新传回到客户端。这种交流过程的编程就叫做通用网关接口编程。CGI 可以运行在多种平台上,具有强大的功能,可以使用多种语言编程,如 Visual Basic、Visual C++、Perl、AppleScript 等,比较常见的是用 Perl 语言编写的 CGI 程序。但是 CGI 有其致命的弱点,即速度慢和安全性差等。

2. 动态服务器页面(ASP)

动态服务器页面(Active Server Pages, ASP)是微软公司推出的一种用以取代 CGI 的技术,是一种真正的简便易学、功能强大的服务器编程技术。ASP 实际上是微软开发的一套服务器端脚本运行环境,通过 ASP 我们可以建立动态的、交互的、高效的 Web 服务器应用程序。用 ASP 编写的程序都在服务器端执行,程序执行完毕后,再将执行的结果返回给客户端浏览器,这样不仅减轻了客户端浏览器的负担,大大提高了交互速度,而且避免了 ASP 程序源代码的外泄,提高了程序的安全性。有关 ASP 技术的详细介绍和应用将在第 12 章介绍。

3. Java 服务器页面(JSP)

Java 服务器页面(Java Server Pages, JSP)是 Sun 公司发布的 Web 应用开发技术,一经推出,就得到了人们的广泛关注。JSP 技术为创建高度动态的 Web 应用提供了一个独特的开发环境,它能够适应市场上 85% 的服务器产品。

JSP 使用 Java 语言编写服务器端程序,当客户端向服务器发出请求时,JSP 源程序被编译成 Servlet 并由 Java 虚拟机执行。这种编译操作仅在对 JSP 页面进行第一次请求时发生。因此,JSP 程序能够提供更快的交互速度,其安全性和跨平台性也很优秀。

目前,ASP 技术与 JSP 技术是市场上并驾齐驱的两种 Web 应用开发技术,各自都占有一定的市场份额。我们将在第 12 章对这两种技术做进一步的学习和讨论。

1.2 数据库技术的发展

数据库技术是计算机科学技术中发展最快的重要分支之一,它已经成为计算机信息系统和应用系统的重要技术支撑。数据库技术所研究的问题就是如何科学地组织和存储数据,如何高效地获取和处理数据,而数据处理的中心问题是数据管理,因此,数据管理的发展是数据库技术发展的一个重要标志。

数据管理指的是对数据的分类、组织、编码、存储、检索和维护。自 20 世纪 60 年代末 70 年代初以来,随着数据库技术的不断发展和完善,数据管理技术在近 40 年中主要经历了 4 个阶段:人工管理阶段、文件系统阶段、数据库系统阶段和高级数据库系统阶段。

1.2.1 人工管理阶段

早期的数据处理,限于计算机技术的发展主要用于科学计算,数据本身不能独立存储和提供应用,数据只能是附属于计算机程序的一部分,随程序一起运行与消失。

1. 人工管理阶段的背景

20 世纪 50 年代中期以前,计算机主要用于科学计算。可使用的外部存储设备只有磁带、卡片、纸带,没有磁盘等直接存取的设备。而且此时的计算机没有操作系统,没有管理数据的软件,数据处理方式是批处理,所有的数据完全由人工进行管理,因此这个阶段被称为人工管理阶段。

2. 人工数据管理的特点

- 数据不保存。一组数据对应于一个应用程序,应用程序与其处理的数据结合成一个整体。在进行计算时,系统将应用程序与数据一起装入,用完后就将它们撤销,释放被占用的数据空间与程序空间。不仅对用户数据如此处置,对系统软件有时也是这样。
- 没有软件对数据进行管理。程序员不仅要规定数据的逻辑结构,还要在程序中设计物理结构,如存储结构、存取方法、输入/输出方式等。因此,数据与程序不具有独立性,如果数据在存储上改变了,程序员就必须修改程序。
- 没有文件概念。数据的组织方法由应用程序开发人员自行设计和安排。
- 数据面向应用。即如果两个应用程序使用相同的数据,也必须各自定义自己的数据存储和存取方式,不能共享相同的数据定义,因此造成了程序和程序之间大量的相同数据的冗余。

3. 人工数据管理模型

人工管理阶段的数据库管理模型如图 1.1 所示。

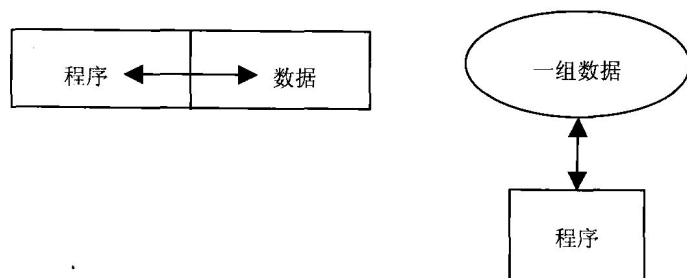


图 1.1 人工数据管理模型

1.2.2 文件系统阶段

随着计算机硬件技术的发展，特别是有了可以直接存储的外存储设备，软件上也有了数据管理软件以后，使得计算机管理的数据可以以文件的形式保留在外存上，可以通过对数据文件的存取进行数据的查询、插入、修改、删除等操作，但是对文件数据的访问操作只能以数据记录为单位，不能以数据项为单位进行。

1. 文件系统阶段的背景

从 20 世纪的 50 年代后期到 60 年代中期，计算机不仅用于科学计算，还大量用于管理。磁盘、磁鼓等直接存取设备也相继被使用，软件中也有了高级语言和操作系统，在操作系统中已经有了专门的管理数据软件，一般称为文件系统，数据处理方式不再是单一的批处理，产生了联机实时处理的方式。

2. 文件数据管理的特点

- 数据能够长期保存在磁盘上。用户可以经常对在外存上保存的数据进行查询、修改、插入和删除等操作。
- 有了提供文件与程序之间进行数据存取方法转换的软件。文件的逻辑结构与存储结构由系统进行转换，使程序与数据有了一定的独立性。程序员不必关心数据的物理位置，可以集中精力考虑算法，并且节省了维护程序的工作量。
- 文件形式的多样化。由于有了直接存取存储设备，也就有了索引文件、链接文件、直接存取文件等。对文件的记录可以顺序访问、随机访问。文件之间相互独立，必须用程序来实现文件与文件之间的联系。
- 数据存取以记录为单位。

3. 文件数据管理模型

文件系统阶段的数据管理模型如图 1.2 所示。

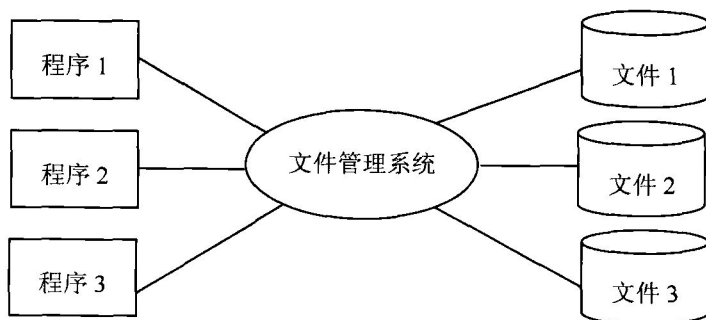


图 1.2 文件系统阶段的模型



4. 文件系统的缺陷

文件系统比第一阶段有很大的改进。但是文件管理下的数据仍然是无结构的信息集合，它可以反映现实世界中客观存在的事物，却不能反映各事物之间的联系。文件系统的主要缺陷如下。

- 数据冗余度大。文件系统中的数据还是面向应用的，因此可能有相同的数据在多个应用程序的文件中重复存储。
- 不一致性。由于数据有冗余，在进行修改时，同样的数据可能会在不同的文件中产生不一致性。
- 数据联系弱。这是文件之间缺乏联系造成的。
- 数据和程序缺乏独立性。数据的逻辑结构对某一特定应用服务来说是优化的，因此很难扩充，一旦数据的逻辑结构改变，则必须修改应用程序；而应用程序的改变又会影响数据结构的变化。因此文件系统仍然是一个不具有弹性的无结构的数据集合。

1.2.3 数据库系统阶段

随着计算机技术的快速发展，出现了大容量硬盘，硬件的价格也趋于下降。计算机信息处理对大量数据进行管理的需求越来越多，要求也越来越高，并随之发展完善了坚实的数据库理论基础，出现了数据库管理系统商品软件，使数据库技术应用日益普及，理论发展趋于成熟。

数据库技术的产生，使数据与程序可以相互独立，能最大限度地减少存储数据的重复冗余，最大限度地为多个程序或用户所共享。

1. 数据库系统阶段的背景

从 20 世纪 60 年代后期开始，计算机用于管理的规模越来越庞大，存储技术也取得了很大发展，有了大容量的磁盘存储数据；系统对联机实时处理的要求越来越多，并开始提出和考虑分布处理。在这样的背景下，为了解决多用户、多应用共享数据的需求，人们着手开发和研制更加完美的数据管理模式，出现了数据库这样的数据管理技术，数据库的概念从此产生。进入 20 世纪 70 年代后，数据库技术得到迅猛发展，并成为计算机科学的一个重要分支。

2. 数据库系统的特点

数据库系统相比文件系统来说，克服了文件系统的缺陷，在数据管理方面有了一次重大的飞跃。这一阶段数据管理的主要特点如下。

- 数据库中的数据是结构化的。由于文件系统中的数据是无结构的或半结构化的，数据库系统不仅考虑数据项之间的联系，而且还要考虑记录之间的联系，这种联系是通过存储路径来实现的。
- 数据冗余度小，易扩充。文件系统中的数据均是与某个应用相联系的，而数据库

中的数据面向系统，减少了数据冗余，实现了不同应用间的数据共享，如图 1.3 所示。

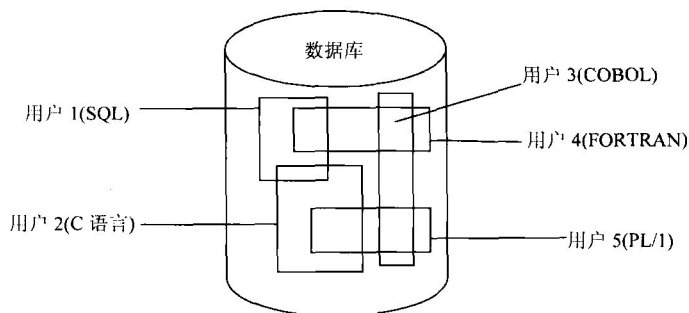


图 1.3 数据共享

- 较高的数据和程序独立性。数据库系统提供了两方面的映像功能，一方面是数据的存储结构与逻辑结构之间的映像或转换功能；另一方面是数据的总体逻辑结构与某类应用所涉及的局部逻辑结构之间的映像或转换功能。前者使得当数据的存储结构改变时，数据的逻辑结构可以不变，从而应用程序可以不必修改，称为数据和程序的物理独立性，简称数据的物理独立性。后者使得当总体逻辑结构改变时，通过对映像的相应改变而保持局部逻辑结构不变。程序员根据局部逻辑结构编写应用程序，因此应用程序可以不必改变。这就是数据和程序的逻辑独立性，简称数据的逻辑独立性。
- 数据库为用户提供了方便的接口。用户可以用数据库系统提供的查询语言和交互式命令操纵数据库，也可以用高级语言编写的程序来操纵数据库，这样就拓宽了数据库的应用范围。
- 数据的最小存取单位是数据项。在数据库中用户既可以存取数据库中某一个数据项或一组数据项，也可以存取一个记录或一组记录。

3. 数据库系统的控制功能

数据库是系统中各用户的共享资源，计算机的共享一般是并发的(Concurrency)，即许多用户同时使用数据库，因此，系统必须提供统一的数据控制功能。主要表现在以下几个方面。

- 数据的安全性(Security)控制。数据的安全性是指保护数据以防止不合法的使用造成数据的泄密和破坏。例如，在学生成绩的数据库中，只有有修改权限的用户才能对此数据库的内容进行修改，其他用户只能读取全部或部分数据。
- 数据的完整性约束。数据的完整性是指数据的正确性、有效性与相容性。在输入、修改数据库的过程中始终符合原来数据的定义和规定。例如，人的性别只有男和女；学生的学号是唯一的，不能有重复；一年只有 12 个月等。
- 并发控制。当多个用户同时存取、修改数据库中的数据时，可能会发生相互干扰，从而使数据完整性遭到破坏。数据库的并发控制避免了这种现象的发生，提高了数据库的利用率。



- 数据库的恢复。数据库系统虽然在数据管理上有较大改进，但是它不可能永远正确地工作。因此，当数据库在运行过程中发生硬件或软件故障时，数据库系统提供了将数据库恢复到最近某个时刻的正确状态的功能。

4. 数据库系统数据管理模型

数据库阶段的数据管理模型如图 1.4 所示。

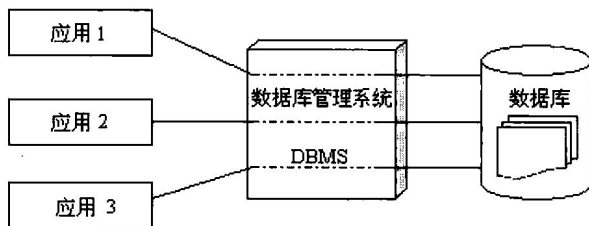


图 1.4 应用与数据库的联系

5. 数据库、数据库管理系统以及数据库系统的定义

综上所述，我们可以为数据库下一个定义：数据库是与应用彼此独立的、以一定的组织方式存储在一起的、彼此相互关联的、具有较少冗余的、能被多个用户共享的数据集合。数据库可以人工地建立、维护和使用，也可以通过计算机建立、维护和使用。当然，我们关心的是后者。计算机化的数据库一般都是通过应用程序或数据库管理系统来建立、维护 and 使用的。

数据库管理系统(DBMS)是一个通用的软件系统，它由一组计算机程序构成。数据库管理系统能够对数据库进行有效的管理，包括存储管理、安全性管理、完整性管理等。数据库管理系统提供了一个软件环境，使用户能方便快速地建立、维护、检索、存取和处理数据库中的信息。

狭义地讲，数据库和数据库管理系统加在一起就构成了数据库系统；广义地讲，数据库系统是由数据库、硬件、软件和数据库管理员 4 个部分组成的。有关这部分的定义我们在 1.5 节还会有详尽的介绍。

1.2.4 高级数据库系统阶段的应用与发展

20 世纪 70 年代中期以来，随着计算机技术的不断发展，出现了分布式数据库、面向对象数据库和智能型知识数据库等，通常被称为高级数据库技术。特别是进入 20 世纪 80 年代以后，在不断出现的数据库新产品中，关系型数据库系统居多，而且随着数据库技术应用的普及，数据库管理系统的功能也越来越强。

1. 高级数据库系统阶段的背景

进入 20 世纪 80 年代后，数据库技术在不同需求的驱动下得到了很大的发展。其特征是数据库技术与应用领域技术的结合，形成了很多新鲜的技术内容。如工程数据库是数据