

高等应用型院校计算机教育规划教材

贾铁军 主编

数据库技术 及应用

T
ECHNIQUE &
APPLICATION OF DATABASE



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

高等应用型院校计算机教育规划教材

数据库技术及应用

主 编 贾铁军

副主编 沈学东

编 著 刘泓漫 熊玉梅 闫俊英

数据库技术及应用 \ 贾铁军主编. —北京: 中国铁道出版社, 2009. 9

(高等应用型院校计算机教育规划教材)

ISBN 978-7-113-09954-1

I. ① 贾... II. ② 刘... III. ③ 数据库系统—高等学校—教材

IV. TP311.13

中国版本图书馆CIP数据库(2009)第100894号

主 编 贾铁军

副主编 沈学东

责任编辑 李 佳

封面设计 曹 佳

印刷制作 曹 佳

封面设计 曹 佳

责任编辑 李 佳

出版发行: 中国铁道出版社(北京西城区广安门内大街8号) 邮政编码: 100045

印 刷: 三河市华丰印刷厂

版 次: 2009年7月第1版

开 本: 787mm×1092mm 1/16 印张: 18 字数: 451千

印 数: 5 000册

书 号: ISBN 978-7-113-09954-1/TP·3240

定 价: 28.00元

京铁出[2009]第100894号

中国铁道出版社

CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

本书结合流行的 SQL Server 2005, 重点介绍了数据库的基本原理和技术方法。全书共分为 9 章, 包括数据库基础知识、SQL Server 的基本结构和功能特点、数据库管理、表和视图管理、索引技术、查询更新方法、T-SQL 语句及其操作、完整性与安全性管理、备份和恢复技术及数据库应用系统设计、PowerBuilder 工具与数据库应用软件开发案例、数据库新技术及实验等。

本书配有实验指导、练习与实践习题和部分参考答案, 并通过出版社网站提供多媒体课件, 便于师生的教学和学生课外延伸学习。

本书适合作为应用型本科院校计算机类、信息类、电子商务类和管理类各专业的数据库相关课程的教材, 也可作为数据库培训班参考用书。

图书在版编目 (CIP) 数据

数据库技术及应用/贾铁军主编. —北京: 中国铁道出版社, 2009. 6

(高等应用型院校计算机教育规划教材)

ISBN 978-7-113-09954-1

I. 数... II. 贾... III. 数据库系统—高等学校—教材
IV. TP311.13

中国版本图书馆 CIP 数据核字 (2009) 第 100689 号

书 名: 数据库技术及应用

作 者: 贾铁军 主编

策划编辑: 秦绪好 杨 勇

责任编辑: 秦绪好

编辑部电话: (010) 63583215

编辑助理: 张爱华

封面设计: 唐韵设计

封面制作: 白 雪

责任印制: 李 佳

出版发行: 中国铁道出版社 (北京市宣武区右安门西街 8 号 邮政编码: 100054)

印 刷: 三河市华丰印刷厂

版 次: 2009 年 7 月第 1 版 2009 年 7 月第 1 次印刷

开 本: 787mm×1092mm 1/16 印张: 18 字数: 421 千

印 数: 5 000 册

书 号: ISBN 978-7-113-09954-1/TP·3240

定 价: 28.00 元

版权所有 侵权必究

本书封面贴有中国铁道出版社激光防伪标签, 无标签者不得销售

凡购买铁道版的图书, 如有缺页、倒页、脱页者, 请与本社计算机图书批销部调换。

信息、物资和能源已经成为人类赖以生存和发展的重要保障,信息技术的快速发展为人类社会带来了深刻的变革。数据库技术是数据管理的高新技术,是计算机科学的重要分支,与计算机网络、人工智能一起被称为计算机技术界三大热门技术,是现代化管理的有力工具。

数据库技术是计算机技术中发展最快、应用最广泛的一项技术,已经成为各类计算机信息系统的核心技术和重要基础。现代信息社会,数据已经成为重要信息资源,数据库技术已经应用到各行各业的各个层面,管理信息系统、企业资源计划、供应链管理系统、客户关系管理系统、电子商务系统、决策支持系统、数据挖掘信息系统等,都离不开数据库技术强有力的支持,数据库技术具有广阔的发展和应用前景。

SQL Server 2005 是微软公司具有里程碑意义的企业级数据库产品,在企业级支持、商业智能应用、管理开发效率等方面具有显著功能,是集数据管理与商业智能分析于一体的新式数据管理与分析平台,并具有完整的关系数据库创建、管理、设计和开发功能。

本书作者长期从事计算机相关专业的教学与科研工作,不仅积累了丰富的教学经验,而且还具有多年数据库应用系统的开发设计经历。本书是 2009 年上海市重点课程“数据库原理及应用”建设项目的特色教材。

本书主要突出“实用、特色、新颖、操作性”,力求技术先进、实用性强。本书共分 9 章,结合流行的 SQL Server 2005,重点介绍了数据库的基本原理和技术方法。主要内容包括:数据库基础知识、SQL Server 2005 的基本结构和功能特点、数据库管理、表和视图管理、索引技术、查询更新方法、T-SQL 语句及其操作、完整性与安全性管理、备份和恢复技术及数据库应用系统设计、PowerBuilder 工具与数据库应用软件开发案例、数据库新技术等,还提供了各章的实验指导。书中带“*”部分为选学内容。

编著本书旨在重点介绍数据库的最新成果、基本原理和技术方法,数据管理技术、方法和实际应用。本教材主要是专门针对应用型人才培养编写的,其特点是:

1. 内容丰富,结构新颖。本书吸收了国内外大量的新知识、新技术和新方法,注重科学性、先进性、操作性,图文并茂、学以致用。配有“学习要点”和“课堂讨论”。
2. 注重实用性和特色。坚持“实用、特色、规范”原则,突出实用并注重素质能力培养,增加大量案例和实验,使本书在内容安排上将理论知识与实际应用有机结合起来。
3. 资源配套,便于教学。为了方便师生教学,配有同步实验指导、练习与实践习题及部分参考答案,还通过出版社网站 <http://edu.tqbooks.net> 提供电子教案。

本书由贾铁军教授任主编、统稿并编著第 1 章、第 6 章和第 8 章,沈学东任任副主编并编著第 5 章和第 9 章,刘泓漫编著第 2 章和第 3 章,熊玉梅编著第 7 章,闫俊英编著第 4 章,王坚完成了部分习题解答和课件制作,邹飞和于森参加了本书大纲的讨论、编著审校等工作。邹佳芹女士查阅资料并对全书的文字、图表进行了校对编排等。

前言

非常感谢中国铁道出版社为本书的编著、出版提供了重要帮助并提出指导意见。同时，感谢对本书编著给予大力支持和帮助的上海市教委和上海电机学院。在本书编著过程中参阅了大量的重要文献资料，难以完全准确注明，在此深表诚挚谢意！

由于数据库技术更新迅速，编者时间仓促及水平有限，书中难免存在不妥之处，敬请海涵！欢迎提出宝贵意见和建议。

编者

2009年6月

目录

CONTENTS

第 1 章 数据库系统概论	1
1.1 数据库的概念	1
1.1.1 数据与信息	1
1.1.2 数据管理与数据库系统	2
1.1.3 数据库技术的特点及应用	2
1.2 数据库技术的发展	4
1.2.1 人工管理阶段	5
1.2.2 文件管理阶段	5
1.2.3 数据库管理阶段	5
1.2.4 高级数据库管理阶段	6
*1.2.5 数据库技术的发展趋势	7
1.3 数据库系统的组成	8
1.3.1 数据库系统的组成概述	8
1.3.2 数据库系统结构类型	9
1.4 数据库的体系结构	11
1.4.1 数据库的三级体系结构	11
1.4.2 数据库的二级映像与数据独立性	13
1.5 数据库管理系统	13
1.5.1 DBMS 的工作模式	13
1.5.2 DBMS 的主要功能	15
1.5.3 DBMS 的模块组成	15
1.6 数据模型	16
1.6.1 数据模型的概念	16
1.6.2 实体联系模型	16
1.6.3 数据模型的组成要素	21
1.6.4 层次模型	22
1.6.5 网状模型	23
1.6.6 关系模型	23
1.6.7 面向对象模型	26
本章小结	30
练习与实践一	31
第 2 章 关系数据库	33
2.1 关系模型的基本概念	33

2.1.1	基本术语	33
2.1.2	关系的定义和性质	34
2.1.3	关系模式	35
2.1.4	E-R 模型向关系模型的转换	35
2.2	关系的完整性	37
2.3	关系代数	38
2.3.1	关系代数的运算符及分类	38
2.3.2	传统的集合运算	38
2.3.3	专门的关系运算	40
*2.4	关系演算	44
2.4.1	元组关系演算	44
2.4.2	域关系演算	45
2.5	查询优化	46
2.5.1	关系代数等价变换规则	46
2.5.2	关系表达式的优化算法	48
	本章小结	48
	练习与实践二	49
第3章	关系数据库语言	51
3.1	SQL 的基本概念与特点	51
3.1.1	SQL 的基本概念	51
3.1.2	SQL 的主要特点	52
3.2	SQL 的数据定义	53
3.2.1	创建和删除 SQL 模式	53
3.2.2	数据类型	54
3.2.3	创建、修改和删除基本表	54
3.3	索引	56
3.3.1	索引的作用	56
3.3.2	索引的分类	57
3.3.3	创建索引	57
3.3.4	删除索引	57
3.4	SQL 的数据查询	58
3.4.1	SELECT 语句命令的格式	58
3.4.2	SELECT 语句中的限定	58
3.4.3	SELECT 查询语句	59
3.5	SQL 的数据更新	62
3.5.1	插入数据	62
3.5.2	修改数据	62

3.5.3 删除数据	63
3.6 视图	63
3.6.1 创建和删除视图	63
3.6.2 更新视图	64
3.6.3 查询视图	64
*3.7 嵌入式 SQL	65
本章小结	66
练习与实践三	66
第4章 关系数据库的规范化设计	68
4.1 规范化问题的提出	68
4.1.1 规范化理论的主要内容	68
4.1.2 不合理的关系模式存在的存储异常问题	69
4.2 函数依赖	70
4.2.1 函数依赖的定义	70
4.2.2 函数依赖的逻辑蕴涵	71
4.2.3 函数依赖的推理规则	71
4.2.4 属性集的闭包	72
4.2.5 候选键的求解理论和算法	72
4.2.6 函数依赖推理规则的完备性	73
4.2.7 函数依赖集	74
*4.3 关系模式的分解	75
4.3.1 模式分解问题	75
4.3.2 无损分解	76
4.3.3 无损分解的测试方法	76
4.3.4 保持函数依赖的分解	78
4.4 关系模式的范式	78
4.4.1 第一范式 (1NF)	79
4.4.2 第二范式 (2NF)	79
4.4.3 第三范式 (3NF)	80
4.4.4 BCNF (Boyce-Codd NF)	81
4.4.5 第四范式 (4NF)	83
4.5 关系模式的规范化	83
4.5.1 关系模式规范化的目的和原则	83
4.5.2 关系模式规范化的步骤	83
4.5.3 关系模式规范化的要求	84
本章小结	84
练习与实践四	85

第5章 数据库设计	88
5.1 数据库设计概述	88
5.1.1 数据库设计的任务、内容和特点	88
5.1.2 数据库设计方法概述	89
5.1.3 数据库设计的步骤	90
5.2 数据库应用系统设计	92
5.2.1 系统需求分析	92
5.2.2 概念结构设计	95
5.2.3 逻辑结构设计	102
5.2.4 数据库物理设计	106
5.2.5 数据库实施	107
5.2.6 数据库运行和维护	108
5.3 数据库设计文档	109
5.4 数据库应用系统设计案例	111
5.4.1 引言	111
5.4.2 用户需求	111
5.4.3 概念结构设计	116
5.4.4 逻辑结构设计	120
5.4.5 物理结构设计	121
5.4.6 数据库的实施和维护	126
本章小结	130
练习与实践五	130
第6章 数据库安全技术	133
6.1 数据库的安全性	133
6.1.1 数据库安全性问题	133
6.1.2 数据库权限管理	135
6.1.3 安全监控方法	141
6.1.4 用户与角色管理	143
6.2 完整性控制	145
6.2.1 数据的完整性	145
6.2.2 完整性规则构成	145
6.2.3 完整性约束条件的分类	146
6.2.4 数据完整性的实施	147
6.2.5 完整性规则实现	151
6.2.6 默认值	153
6.3 并发控制与封锁	154
6.3.1 并发操作带来的问题	154

6.3.2	事务	154
6.3.3	封锁技术	156
6.3.4	并发操作的调度	157
6.4	数据备份与恢复	157
6.4.1	数据备份	158
6.4.2	数据恢复	160
6.5	数据库安全解决方案	165
6.5.1	数据库安全策略	165
6.5.2	数据加密	168
6.5.3	数据库审计	169
	本章小结	170
	练习与实践六	170
第 7 章	数据库新技术	172
7.1	分布式数据库	172
7.2	主动数据库	173
7.3	多媒体数据库	176
7.4	数据仓库与联机分析处理技术	179
7.4.1	数据仓库技术	179
7.4.2	联机分析处理技术	180
7.4.3	数据挖掘技术	182
7.5	数据库新技术应用发展趋势	184
	本章小结	185
	练习与实践七	186
第 8 章	PowerBuilder 与数据库应用开发	187
8.1	PowerBuilder 简介	187
8.1.1	PowerBuilder 概述	187
8.1.2	PB 特点及访问数据库方式	190
8.1.3	PB 常用术语及主窗口菜单	193
8.2	PowerScript 语言概述	194
8.3	利用 PB 开发数据库应用程序	197
8.3.1	创建 PB 应用程序的过程	197
8.3.2	基于 Web 的库存管理及订购系统开发案例	199
	本章小结	213
	练习与实践八	213
第 9 章	数据库应用实践	215
9.1	SQL Server 界面及功能应用实验	215

9.1.1	实验目的	215
9.1.2	实验内容	215
9.1.3	实验步骤	215
9.2	关系模式及模型应用实验	223
9.2.1	实验目的	223
9.2.2	实验应用背景及步骤	223
9.2.3	练习	230
9.3	SQL 基本操作实验	232
9.3.1	数据定义实验	232
9.3.2	索引应用操作	238
9.3.3	数据查询实验	240
9.3.4	视图操作实验	244
9.4	关系数据库的规范化设计解决方案	246
9.4.1	实验目的	246
9.4.2	实验内容及步骤	247
9.5	数据库安全实验	251
9.5.1	实验目的	251
9.5.2	实验内容及步骤	251
9.5.3	练习	261
9.6	综合应用实践	261
9.6.1	实验目的	261
9.6.2	实验内容	261
	本章小结	261
	课程设计	262
附录 A	练习与实践部分参考答案	264
参考文献		275

第 1 章

数据库系统概论

数据库技术是数据管理的最新技术，是计算机科学的重要分支，与计算机网络、人工智能一起被称为计算机技术界三大热门技术，是现代化管理的重要工具。随着 IT 的发展，数据库技术应用已从事务处理扩大到计算机辅助设计、人工智能、决策支持系统和网络应用等新领域。各行业大量的重要数据需要经过数据库才能进行有效组织、存储、处理和共享。通过学习数据库系统有关基本知识，可为以后的学习奠定重要基础。

学习要点

- 数据库的基本概念
- 数据库技术特点、应用及发展趋势
- 数据库系统的组成及数据库的体系结构
- DBMS 的工作模式、主要功能和组成
- 概念模型与数据模型

教学目标

- 理解数据、数据管理和数据库系统的基本概念
- 掌握数据库技术特点、应用及发展趋势
- 了解数据库系统的组成及数据库的体系结构
- 掌握 DBMS 的工作模式、主要功能和组成
- 理解概念模型与数据模型

1.1 数据库的概念

1.1.1 数据与信息

信息 (information) 是现实世界在人们头脑中的反映，是对客观事物状态和特征的描述。信息通过载体才能进行传播与处理。

数据 (data) 是信息的表达方式，是人们描述客观事物及其活动的抽象表示。数据的概念包括描述事物特性的数据内容和存储在某一种媒体上的数据形式。

数据分为数值数据和非数值数据，可以是数字、文字、符号、图形、表格、图像、声音、视

频等形式。数据是数据库中存储的基本对象,人们收集并抽取所需要的大量数据之后,将其保存并进行进一步加工处理,从而得到有用信息。

数据与信息互相依存又有区别。数据是信息的具体表示形式和载体,信息反映数据的含义。信息和数据是数据库管理的基本内容和对象。信息源于数据,信息以数据的形式存储和处理。信息可用不同的数据形式来表示,信息不随数据表现形式而改变,而数据的表示方式却可以选择。

1.1.2 数据管理与数据库系统

数据处理(data processing)是对数据进行加工的过程。对数据进行的查找、统计、分类、修改、变换、运算等都属于加工。其目的是根据需要,从大量的数据中抽取有意义、有价值的数

据(信息),作为决策和行动的依据。其实质是信息处理。

数据管理(data management)以数据的管理为目的,在数据处理的一系列活动中,以数据收集、存储、传播、检索、分类等操作为基本环节,这些基本环节统称为数据管理。

例如,从“学生成绩表”中查找分数最高的学生、按分数从高到低排序、修改成绩或打印等都属于数据管理。而进行成绩统计汇总或制作数据图则属于数据处理。

数据库(database, DB)是结构化的数据集合,是长期存储在计算机内的、有组织的、可共享的数据集合,可视为按一定结构“存放数据的仓库”。数据库中的数据按一定的数据模型组织、描述和存储,具有较小的冗余度,较高的数据独立性和易扩展性,并可共享。数据库具有集成性、共享性、海量性和持久性的特点。

数据库系统(database system, DBS)是采用了数据库技术的计算机系统,是实现有组织地、动态地存储大量关联数据,方便多用户访问的计算机软硬件和数据资源组成的系统。其主要特性为:实现数据共享,减少数据冗余度,保持数据一致性和数据

独立性;提高系统的安全保密性,并发控制及故障恢复。

数据库管理系统(database management system, DBMS)是建立、运用和维护数据库,并进行统一管理、统一控制的数据管理软件。它便于用户定义和操作数据,并保证数据的安全性、完整性,多用户对数据的并发使用及发生故障后的数据库恢复等。DBMS 是数据库系统的核心,允许在数据库内共享对数据的访问。DBMS 由一个数据库和一组管理数据的程序构成,其地位如图 1-1 所示。常见的大型关系 DBMS 有微软的 SQL Server, IBM 的 DB2, 以及 Oracle、Sybase、Informix 等,桌面型关系数据库有 FoxPro、Access 等。



图 1-1 数据库管理系统的地位

1.1.3 数据库技术的特点及应用

1. 数据库技术的主要特点

(1) 实现数据的高度集成

数据处理应用系统中的数据来源于各项业务,且数据之间又相互关联。如在一个商品供销管理信息系统中,进货数据来源于供货管理,销售数据来源于售货管理,员工数据来源于人力资源

管理等。对这些数据进行集中管理,保持数据之间的正确关联,才能完成所需的综合数据处理。利用数据库技术和 DBMS 提供的数据库管理功能可以实现数据集成。

(2) 提供有效的数据共享

在一个数据库应用系统中,需要对集中管理的数据进行共享。如在商品销售管理系统中,供货管理需要参考近期的销售数据来确定进货种类与数量,销售单价的确定又必须参照最近的进货单价等,这就需要数据共享。可以利用数据库技术的数据共享功能来实现。

(3) 减少数据冗余

传统的数据处理应用系统中的每一项业务都需拥有各自的数据文件,其中的数据在若干不同业务的数据文件中分别保存,出现数据冗余。大量冗余数据的存在将导致应用系统存储、管理和维护上的困难。数据库技术对所有数据集中管理,并利用有效的数据共享功能,不再需要各项业务单独保存各自的数据文件,减少了大量的数据冗余。

(4) 保证数据一致性

数据一致性是指保存在数据库中不同数据集合中的相同数据项必须具有相同的值。一个数据库由许多数据文件组成,数据文件之间通过公共数据项相联系,当对一个数据文件中的这种数据项进行修改时,相关联文件中的对应数据项也必须自动修改才能始终保持数据库数据的一致性和正确性。通过 DBMS 自动实现或编程,可以实现对数据库数据进行插入和删除等操作的一致性。

(5) 实施统一的数据标准

数据标准是指数据项的名称、数据类型、数据格式、有效数据的判定准则等数据项特征值的取值规则。

(6) 控制数据的安全性和完整性

针对数据库所进行的各项操作都必须根据用户所拥有的权限进行鉴别,鉴别机制由 DBMS 提供,每个用户的权限设定则由数据库管理员(DBA)负责建立,以保障数据库应用系统数据的安全、保密和完整性。

(7) 实现数据的独立性

数据独立性是指存储在数据库中的数据独立于处理数据的应用程序。数据不会因为某项应用的需要而改变其结构。应用程序对某一项应用是通过处理数据获取信息的过程实现的,只能根据具体数据来设计所需要的数据处理方法而保持数据独立。

(8) 减少应用程序开发与维护工作量

由于实现了数据的独立性,使得在应用程序开发时,不再需要考虑所处理的数据组织问题,从而减少了应用程序的开发与维护工作量。但在应用系统开发初期,必须完善规划数据库,设计数据库中的各个数据集,规范数据库中相关数据间的关联。只有满足规范化设计要求的数据库,才能够真正实现各类不同的应用需求。

2. 数据库技术应用

随着 IT 技术的发展,数据库技术的应用从数据处理与管理扩展到计算机辅助设计、人工智能、决策支持系统和网络应用等新领域。现在已经渗透到各行各业和业务部门的各个层面。因此,数据库系统已构成计算机应用系统中的重要支撑性软件。

【案例 1.1】数据库技术各行业应用示例。

- ① 销售业：用于存储供应商、商品、客户信息以及商品的销售信息。
- ② 航空业：用于存储航班和票务信息。航空业是最先以地理上分布的方式使用数据库的行业之一，分布于世界各地的终端，通过通信网络或其他数字网络来访问中央数据库系统。
- ③ 金融业：用于存储股票、债券等金融票据的持有、出售和买入等交易信息，还可以用于存储银行客户的信息、账户、贷款以及银行的交易记录。
- ④ 电信业：用于存储通信网络的信息，存储通话记录，存储用户付费业务记录及产生每月通信账单等。
- ⑤ 制造业：用于存储产品的订单，产品原料的供应情况，跟踪产品的产量以及仓库产品的详细清单。
- ⑥ 教育系统：用于存储教职员的信息，存储工资、津贴和纳税的信息，产生工资单；存储学生信息、课程注册和成绩信息等。

数据库技术是数据管理的最新技术，给人们带来了前所未有的便利。如通过网络查询信息、预订机票、付费等。数据库的应用广泛而深入。

随着信息技术的快速发展，数据库技术也产生了一些新的应用领域，主要有五个方面：

- ① 多媒体数据库：主要存储与多媒体相关的数据，如声音、图像和视频等数据。多媒体数据最大的特点是数据量大、数据类型多且数据类型间差距较大，因此需要较大的存储空间和较复杂的数据管理。
- ② 空间数据库：将数据库应用于地理信息系统 (GIS) 和计算机辅助设计与制造 (CAD/CAM)，与地球上的空间位置相关的空间数据是 GIS 的重要组成部分；而设计数据（如机械及集成电路设计等）则是 CAD/CAM 的核心。
- ③ 移动数据库：是在移动计算机系统上发展起来的，如笔记本式计算机、掌上计算机等。其最大特点是通过无线数字通信网络进行传输，用户可以随时随地访问和获取数据，为一些商务应用和应急情况带来了极大的便利。
- ④ 信息检索系统：为一种典型的联机文档管理系统，一直与数据库技术同步发展。
- ⑤ 决策支持系统：联机分析处理 (OLAP) 是数据库系统的主要应用，它支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。它可使分析管理人员或执行人员从多角度对信息进行快速、一致、交互地存取，获得对数据更深入的了解。

课堂讨论：

1. 什么是数据管理？与数据处理有何区别？
2. 数据库系统与数据库管理系统的区别有哪些？
3. 数据库技术的主要特点有哪些？

1.2 数据库技术的发展

利用计算机进行数据处理，极大地促进了数据处理技术不断丰富和发展，到目前基本经历了人工管理、文件管理、数据库管理以及高级数据库管理四个发展阶段。

1.2.1 人工管理阶段

在计算机诞生初期,只有硬件系统,主要是运算器、控制器和存储器,输入/输出设备非常简单,只有纸带或卡片机,工作效率极低,只能输入/输出极少数据。

早期的计算机主要用于科学计算,数据处理都是通过手工方式进行的。当时外存没有磁盘等直接存取的存储设备;软件没有操作系统,数据的处理是批处理。

人工管理数据具有如下特点:

① 数据不存入计算机。由于计算机软硬件技术有限,数据随程序一起输入计算机,处理结束后输出结果,不能长期保存,数据空间随着程序一起被释放。

② 只有程序的概念,没有文件的概念。数据的组织方式由程序员设计与安排,数据需要通过应用程序进行管理,没有相应的软件来处理数据。

③ 数据面向程序。一组数据对应一个程序。数据冗余大且不共享,程序之间互不干扰。④ 数据无独立性。当应用程序改变时,数据的逻辑结构和物理结构也相应发生变化。

1.2.2 文件管理阶段

从20世纪50年代中期到60年代中期,晶体管取代了运算器和控制器中的电子管,可将成批数据单独组成文件存储到外部存储设备上,出现了操作系统、汇编语言和一些高级语言。计算机不仅用于科学计算,还大量用于管理,在操作系统中有专门的数据管理软件,称为文件系统。它是数据库系统发展的初级阶段,并非真正的数据库系统。

1. 文件系统管理数据的特点

① 数据可长期保存。

② 简单的数据管理功能。由文件系统进行数据管理,程序和数据之间有了一定的独立性,可减少程序员的工作量。

③ 数据共享性差。在文件系统中,文件仍然面向应用,当不同文件具有相同数据时,须建立各自的文件,而不能共享数据,致使数据的冗余度大,浪费存储空间。

④ 数据的独立性差。数据结构发生改变时,需要修改应用程序和文件结构的定义;应用程序的改变也会改变数据结构,文件系统仍然是一个无结构的数据集合。

2. 文件系统的缺陷

随着数据管理规模的扩大,数据量急剧增加,文件系统的缺陷主要表现在:

① 数据联系弱。文件之间相互独立,缺乏联系。

② 数据冗余。由于文件之间缺乏联系,造成每个应用程序都有对应的文件,这样相同的数据会在多个文件中重复存储。

③ 数据不一致。由于数据冗余,在进行更新操作时,稍不注意,就可能使相同的数据在多个文件中不一致。例如修改时遗漏,修改时造成错改。

1.2.3 数据库管理阶段

从20世纪60年代中期以来,计算机软硬件技术得到快速发展。含有运算器、控制器和内存储器逐渐向超大规模集成电路发展,为存储和处理大数据量的数据库提供了有力的支持和保证。同时,操

作系统得到了发展,而且各种 DBMS 软件不断涌现,使得数据库管理技术越来越成熟和完善,成为计算机领域中最具影响力和发展潜力,应用范围最广、成果最显著的技术之一,形成了“数据库时代”。数据库系统建立了数据之间的有机联系,实现了统一、集中、独立地管理数据,实现了数据共享。

1. 数据的集成性

数据库系统中采用统一的数据结构方式,使数据结构化;全局的数据结构由多个应用程序共用,各程序调用局部结构的数据,全局与局部的结构模式构成数据集成。

2. 数据高度共享,低冗余

数据库系统从整体角度看待和描述数据,数据面向整个系统而不再面向某个应用,因此数据可以被多用户、多应用所共享。数据库技术与网络技术的结合扩大应用。数据共享程度极大地减少数据的冗余度,节约存储空间,又能避免数据之间的不相容性和不一致性。

3. 数据独立性高

数据的独立性是指用户的应用程序与数据库中数据是相互独立的,即当数据的物理结构和逻辑结构发生变化时,不影响应用程序对数据的使用。

4. 数据的管理和控制能力

数据库系统对数据统一管理和控制,保证了数据的安全性和完整性。数据库系统对访问用户身份及其操作的合法性进行检查;自动检查数据的一致性、相容性,保证数据符合完整性约束条件;以并发控制手段有效控制多用户程序同时对数据操作,保证共享及并发操作;恢复功能保障当数据库遭到破坏时能自动恢复到正确状态。

1.2.4 高级数据库管理阶段

从 20 世纪 80 年代以来,数据库技术在商业的巨大成功刺激了其他领域对其需求的快速增长。不仅开辟了新的应用领域,而且推动了数据库技术的研究和发展,特别是面向对象数据库系统。同时不断与其他技术结合,向高级数据库技术发展。

1. 分布式数据库技术

随着地理上分散的用户对数据共享的需求和网络技术的发展,产生了分布式数据库系统。其具有如下五个主要特点:

① 大部分数据在本地进行分布处理,提高了系统处理效率和可靠性。数据复制技术是分布式数据库的重要技术。

② 克服了中心数据库的弱点,降低了数据传输代价。数据库的数据物理上分布在各地,逻辑上是个相互联系的整体,从而实现了数据的物理分布性和逻辑整体性。

③ 提高系统的可靠性。局部系统发生故障,其他部分仍可继续工作。

④ 各地终端由数据通信网络相连。本地终端单独不能胜任的处理任务,可以通过通信网络取得其他数据库和终端的支持。

⑤ 数据库位置透明,方便系统扩充。从用户看整个数据库仍是个集中的数据库,用户不用关心数据的分片、物理位置分布和副本的一致性,由分布式 DBMS 实现分布。

分布式数据库系统兼顾集中管理和分布处理两项任务,因而具有良好的性能,其具体结构如图 1-2 所示。详细内容将在第 7 章进行介绍。