



工业和信息化人才培养规划教材
Industry And Information Technology Training Planning Materials

Technical And Vocational Education

高职高专计算机系列

数据库技术与应用 ——SQL Server 2008 (第2版)

Technology and Applications of
Database-SQL Server 2008 (version 2)

张建伟 梁树军 金松河 等 © 编著

系统讲述数据库系统的原理

详细介绍SQL Server 2008的基础及操作

以综合实例串起知识内容, 实际体验数据库系统的开发过程



人民邮电出版社
POSTS & TELECOM PRESS



工业和信息化人才培养规划教材
Industry And Information Technology Training Planning Materials

Technical And Vocational Education
高职高专计算机系列

数据库技术与应用 ——SQL Server 2008 (第2版)

Technology and Applications of
Database-SQL Server 2008 (version 2)

张建伟 梁树军 金松河 等 © 编著

系统讲述数据库系统的原理

详细介绍SQL Server 2008的基础及操作

以综合实例串起知识内容, 实际体验数据库系统的开发过程



人民邮电出版社
POSTS & TELECOM PRESS

图书在版编目 (CIP) 数据

数据库技术与应用 : SQL Server 2008 / 张建伟等
编著. — 2版. — 北京 : 人民邮电出版社, 2012.2
工业和信息化人才培养规划教材. 高职高专计算机系
列

ISBN 978-7-115-27068-9

I. ①数… II. ①张… III. ①关系数据库—数据库管
理系统, SQL Server 2008—高等教育—教材 IV.
①TP311.138

中国版本图书馆CIP数据核字(2012)第003962号

内 容 提 要

本书系统地讲述了数据库的原理与 SQL Server 2008 的功能及应用。

全书分 3 篇, 共 16 章。第 1 篇数据库系统原理, 主要介绍关系型数据库的基本原理和数据库设计的步骤; 第 2 篇 SQL Server 2008 基础及操作, 深入研究了 Microsoft SQL Server 2008 系统的基本结构和功能特点、安装规划和配置技术、数据库管理、Transact-SQL 语言、表、数据操纵技术、索引技术、数据安全性与完整性技术、视图技术、存储过程技术、触发器技术、备份和还原技术等; 第 3 篇 SQL Server 2008 应用, 以一个学生成绩管理系统为例, 详细介绍了利用 C# 和 SQL Server 2008 完成学生成绩管理系统的开发过程。

本书内容翔实、结构合理、示例丰富、语言简洁流畅。适合作为高等院校本/专科计算机软件、信息系统、电子商务等相关专业的数据库课程教材, 同时也适合作为各种数据库技术培训班的教材以及数据库开发人员的参考资料。

工业和信息化人才培养规划教材·高职高专计算机系列 数据库技术与应用——SQL Server 2008 (第 2 版)

- ◆ 编 著 张建伟 梁树军 金松河 等
责任编辑 桑 珊
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京隆昌伟业印刷有限公司印刷
- ◆ 开本: 787×1092 1/16
印张: 19.75 2012 年 2 月第 2 版
字数: 491 千字 2012 年 2 月北京第 1 次印刷

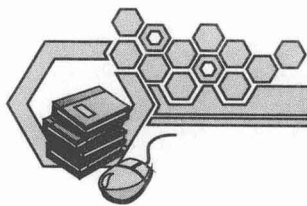
ISBN 978-7-115-27068-9

定价: 39.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223

反盗版热线: (010)67171154

广告经营许可证: 京崇工商广字第 0021 号



第2版前言

数据库技术是计算机科学技术中发展最快的领域之一,也是应用最广的技术之一。数据库管理系统是国家信息基础设施的重要组成部分,也是国家信息安全的核心技术之一。信息技术的飞速发展大大推动了社会的进步,也逐渐改变了人们的生活、工作和学习方式。数据库技术和网络技术是信息技术中的重要支柱。当今热门的信息系统,如管理信息系统、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、智能信息系统等,都离不开数据库技术强有力的支持。

Microsoft SQL Server 是一个典型的关系型数据库管理系统,从 SQL Server 7.0、SQL Server 2000、SQL Server 2005 发展到现在的 SQL Server 2008,随着版本的不断升级,功能越来越强大。SQL Server 2008 可以为各类用户提供完整的数据库解决方案,可以帮助用户建立自己的电子商务体系,增强用户对外界变化的反应能力,提高用户的市场竞争力。

本书初版自 2008 年 4 月出版以来,受到了老师和同学们的欢迎和青睐,这让编者由衷地感到欣慰和感激。第 2 版在保留初版基本框架的基础上,做了以下修改。

1. 软件版本升级,使用 SQL Server 2008 版。
2. 把原来的章节结构分为 3 个篇章:数据库系统原理、SQL Server 2008 基础及操作、SQL Server 2008 应用,使教材结构更加清晰。
3. 丰富了数据库设计、运行和维护的内容。
4. 增加了 SQL Server 2008 组件的介绍、标示符列的使用等内容。
5. 更新了部分例题和实训内容。

本书分为 3 篇,共 16 章。

第 1 篇数据库系统原理包括第 1 章~第 3 章,介绍数据库的基本理论知识,内容包括数据库概述、数据模型、关系数据库、数据库设计。

第 2 篇 SQL Server 2008 基础及操作包括第 4 章~第 15 章,介绍 SQL Server 2008 的安装、配置和管理;SQL Server 2008 的组件和管理工具;数据库管理系统 SQL Server 2008 的使用,内容包括 T-SQL 语言,数据库和表,数据库查询,视图、索引和游标,存储过程和触发器,数据库的备份还原与导入导出,系统安全管理和完整性控制。

第 3 篇 SQL Server 2008 应用包括第 16 章,介绍数据库的应用实例,通过实例介绍了用 C# 和 SQL Server 2008 开发数据库系统的方法。

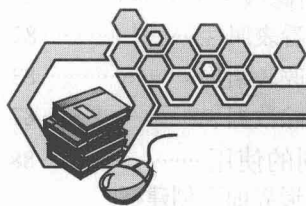
本书由张建伟、梁树军、金松河、毛艳芳、王治国、张保威等编著。全书由梁树军统稿并定稿,由张建伟教授主审。在本书的编写和出版过程中得到了郑州轻工业学院教务处的大力支持和帮助,在此由衷地向他们表示感谢!

本书除了可用作高等院校、高职高专学生的教材外,还可供计算机应用开发人员在学习数据库技术时参考。

由于编写时间仓促,加之水平有限,书中难免有错误之处,恳请广大读者不吝赐教。

编 者

2011 年 9 月



目 录

第 1 篇 数据库系统原理

第 1 章 数据库基础知识	2
1.1 数据库、数据库管理系统与数据库系统	2
1.1.1 数据库	2
1.1.2 数据库管理系统 (DBMS)	3
1.1.3 数据库系统 (DBS)	3
1.2 数据库技术的产生与发展	4
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 高级数据库阶段	8
1.3 数据库系统的组成与结构	9
1.3.1 数据库系统的组成	9
1.3.2 数据库系统的结构	10
1.4 数据库系统的作用与特点	12
1.4.1 数据库系统的作用	12
1.4.2 数据库系统的特点	12
1.5 数据模型	13
1.6 概念模型	14
本章小结	17
习题	17
第 2 章 关系数据库	18

2.1 关系数据库及其特征	18
2.1.1 关系数据库的特点	18
2.1.2 关系模型的基本术语	19
2.1.3 关系的性质	20
2.2 关系模式	21
2.3 关系的完整性	22
2.4 关系数据库语言 SQL	23
本章小结	25
习题	26

第 3 章 数据库的设计

3.1 需求分析	28
3.1.1 需求分析任务	28
3.1.2 需求分析的基本步骤	28
3.2 概念结构设计	29
3.3 逻辑结构设计	30
3.3.1 逻辑结构设计的步骤	30
3.3.2 概念模型转换为一般的关 系模型	31
3.4 物理结构设计	32
3.5 数据库的实施	32
3.6 数据库的运行和维护	33
本章小结	34
习题	34

第 2 篇 SQL Server 2008 基础及操作

第 4 章 SQL Server 2008 概述	36
4.1 SQL Server 2008 简介	36
4.1.1 SQL Server 2008 概述	36
4.1.2 SQL Server 2008 新增	

功能特性	37
4.2 SQL Server 2008 的安装	38
4.2.1 SQL Server 2008 的版本	38
4.2.2 安装 SQL Server 2008 的 软硬件要求	39

4.2.3 SQL Server 2008 安装过程	40	6.3 数据表的修改	82
4.3 SQL Server 2008 的组件	45	6.3.1 查看表属性	82
4.4 SQL Server 2008 的管理工具	47	6.3.2 修改表结构	83
4.4.1 Management Studio	48	6.3.3 删除数据表	87
4.4.2 SQL Server 配置管理器	52	6.4 标识符列的使用	88
4.4.3 SQL Server Profiler	54	6.4.1 图形界面下创建标识符列	88
本章小结	56	6.4.2 使用 SQL 命令创建标识符列	89
习题	56	本章小结	90
本章实训	57	习题	90
第 5 章 创建和管理 SQL Server 2008 数据库	58	本章实训	91
5.1 系统数据库概述	58	第 7 章 操纵数据表中的数据	92
5.1.1 Master 数据库	59	7.1 概述	92
5.1.2 Msdb 数据库	59	7.2 数据操作	93
5.1.3 Model 数据库	59	7.2.1 用 INSERT 语句插入数据	93
5.1.4 Tempdb 数据库	59	7.2.2 用 UPDATE 语句更新数据	95
5.2 创建数据库	60	7.2.3 用 DELETE 语句删除数据	96
5.2.1 数据库文件	60	7.3 数据检索	97
5.2.2 使用 Management Studio 创建数据库	61	7.4 使用 SELECT 子句进行简单查询	97
5.2.3 用 SQL 命令创建数据库	64	7.5 使用 WHERE 子句选择数据	101
5.3 管理数据库	66	7.5.1 确定查询	101
5.3.1 查看数据库属性	66	7.5.2 模糊查询	103
5.3.2 修改数据库	69	7.5.3 带查找范围的查询	105
5.3.3 收缩数据库	71	7.6 聚合函数	106
5.3.4 删除数据库	72	7.7 分组查询	107
本章小结	73	7.8 排序	109
习题	73	7.9 连接查询	111
本章实训	74	7.9.1 交叉连接	111
第 6 章 创建和管理 SQL Server 2008 数据表	75	7.9.2 内连接	112
6.1 表的概念	75	7.9.3 外连接	114
6.2 数据表的创建	76	7.10 子查询技术	115
6.2.1 在图形界面下创建数据表	76	7.10.1 带 IN 的嵌套查询	116
6.2.2 用 SQL 命令创建数据表	79	7.10.2 带比较运算符的嵌套查询	117

7.10.3 带 ANY 或 ALL 的嵌套 查询	118	8.6 用户自定义函数	152
7.10.4 带 EXISTS 的嵌套 查询	120	8.7 游标技术	153
7.11 集合运算	122	8.7.1 游标概述	153
7.11.1 并操作	122	8.7.2 声明游标	154
7.11.2 交操作	123	8.7.3 打开游标	156
7.11.3 差操作	123	8.7.4 从游标中提取记录	156
本章小结	124	8.7.5 关闭游标	157
习题	124	8.7.6 释放游标	158
本章实训	125	8.7.7 游标的应用	158
第 8 章 Transact-SQL 语言	127	本章小结	160
8.1 数据类型	128	习题	160
8.1.1 精确数字类型	128	本章实训	161
8.1.2 近似数字类型	130	第 9 章 视图	162
8.1.3 日期和时间类型	130	9.1 视图的作用和基本类型	162
8.1.4 字符数据类型	132	9.2 视图的创建	163
8.1.5 二进制数据类型	133	9.2.1 在图形界面下创建 视图	163
8.1.6 其他类型	135	9.2.2 用 SQL 语句创建视图	165
8.1.7 用户自定义类型	135	9.3 视图的修改	167
8.2 变量	135	9.4 通过视图查询数据	167
8.2.1 局部变量	136	9.5 通过视图更新数据	168
8.2.2 全局变量	136	9.6 视图的删除	169
8.3 运算符及表达式	137	本章小结	169
8.3.1 运算符	137	习题	170
8.3.2 表达式	138	本章实训	171
8.3.3 注释符	139	第 10 章 索引	172
8.3.4 通配符	139	10.1 索引简介	172
8.4 控制语句和批处理	139	10.2 索引的类型和特点	172
8.4.1 IF...ELSE	139	10.3 创建索引	174
8.4.2 BEGIN...END	140	10.3.1 在图形界面下创建 索引	174
8.4.3 WHILE...CONTINUE... BREAK	140	10.3.2 用 SQL 语句创建索引	176
8.4.4 CASE	141	10.4 修改索引	179
8.4.5 RETURN	141	10.4.1 通过 SQL 语句修改 索引	179
8.4.6 批处理	142	10.4.2 通过图形界面 修改索引	181
8.4.7 其他命令	142	10.5 删除索引	181
8.5 常用函数	143	10.6 索引优化向导	182
8.5.1 聚合函数	144		
8.5.2 标量函数	144		

本章小结	185	12.3.1 创建 DDL 触发器	213
习题	185	12.3.2 测试 DDL 触发器的 功能	214
本章实训	186	12.3.3 查看和修改 DDL 触发器	214
第 11 章 存储过程	187	12.4 登录触发器	215
11.1 存储过程简介	187	本章小结	216
11.1.1 存储过程的类型	187	习题	216
11.1.2 存储过程的优点	188	本章实训	217
11.2 存储过程的创建与执行	189	第 13 章 数据库的备份与还原	218
11.2.1 在图形界面下创建存 储过程	189	13.1 概述	218
11.2.2 用 SQL 语句创建存储 过程	192	13.2 数据库备份方式	219
11.2.3 存储过程的执行	196	13.3 备份设备	220
11.3 修改存储过程	198	13.3.1 物理设备与逻辑设备	220
11.4 重命名存储过程	198	13.3.2 创建与管理备份设备	221
11.5 删除存储过程	199	13.4 数据库备份	223
本章小结	199	13.4.1 完整备份	223
习题	199	13.4.2 差异备份	226
本章实训	200	13.4.3 事务日志备份	227
第 12 章 触发器	201	13.4.4 文件/文件组备份	228
12.1 触发器简介	201	13.5 数据库还原	230
12.1.1 触发器的概念	201	13.6 数据库的分离和附加	237
12.1.2 触发器的功能	201	13.6.1 分离数据库	237
12.1.3 触发器的类型	202	13.6.2 附加数据库	238
12.2 DML 触发器	203	本章小结	241
12.2.1 DML 触发器的类型	203	习题	241
12.2.2 DML 触发器的工作 原理	203	本章实训	241
12.2.3 创建 DML 触发器的注意 事项	203	第 14 章 数据传输	243
12.2.4 创建 AFTER 触发器	204	14.1 DTS 概述	243
12.2.5 创建 INSTEAD OF 触发器	208	14.2 数据导出	244
12.2.6 查看 DML 触发器	209	14.3 数据导入	248
12.2.7 修改 DML 触发器	210	本章小结	251
12.2.8 删除 DML 触发器	211	习题	251
12.2.9 禁用与启用 DML 触发器	211	本章实训	251
12.3 DDL 触发器	212	第 15 章 SQL Server 2008 数据库 的安全性和完整性管理	253
		15.1 数据库的安全性	253
		15.1.1 身份验证简介	253

15.1.2 验证模式的修改	254
15.2 管理服务器登录	254
15.2.1 使用 Management Studio 管理登录账户	254
15.2.2 使用 Transact-SQL 管理 登录账户	257
15.3 角色和用户管理	258
15.3.1 角色管理简介	258
15.3.2 角色的管理	260
15.3.3 用户管理简介	264
15.3.4 用户的管理	264
15.4 SQL Server 2008 权限	266
15.5 权限管理	268

15.6 数据库的完整性	271
15.7 约束的概念和类型	272
15.7.1 PRIMARY KEY 约束	272
15.7.2 FOREIGN KEY 约束	273
15.7.3 UNIQUE 约束	274
15.7.4 CHECK 约束	274
15.7.5 DEFAULT 定义	274
15.7.6 允许空值	274
15.8 管理约束	275
本章小结	275
习题	276
本章实训	276

第 3 篇 SQL Server 2008 应用篇

第 16 章 学生成绩管理系统的 设计与实现	280
16.1 ADO.NET 数据库访问对象 模型	280
16.1.1 ADO.NET 结构	281
16.1.2 数据集介绍	282
16.2 系统功能设计	283
16.3 数据库和表设计	284
16.4 程序开发	288
16.4.1 创建项目	288
16.4.2 登录窗口	288

16.4.3 主窗口	292
16.4.4 基础资料	294
16.4.5 教学管理	296
16.4.6 用户管理	301
16.4.7 “关于”窗口	302
本章小结	303

附录 ——创建学生成绩管理 数据库	304
----------------------------	-----

参考文献	306
------------	-----

第 1 篇

数据库系统原理

第1章

数据库基础知识

数据库技术已成为计算机科学的一个重要分支，是数据管理的最新技术，也是计算机技术中发展最快的领域之一。许多信息系统都是以数据库为基础建立的。数据库技术已经成为人们存储数据、管理信息、共享资源的最先进、最常用的技术。

本章介绍数据库系统的基本概念，包括数据管理技术的发展过程、数据库系统的基本概念、数据模型及数据库系统的体系结构等。读者从中可以学习到使用数据库的原因及其重要性。本章是学习后面各章节的预备和基础。

1.1 数据库、数据库管理系统与数据库系统

数据库技术是计算机技术中发展最为迅速的领域之一，已经在科学、技术、经济、文化和军事等领域发挥着重要作用。

1.1.1 数据库

数据库 (Database, DB)，顾名思义，是存放数据的仓库。只不过这个仓库是在计算机的存储设备上，而且数据是按照一定的数据模型组织并存放在外存上的一组相关数据的集合。通常这些数据是面向一个组织、企业或部门的。例如，学生成绩管理系统中，学生的基本信息、课程信息、成绩信息等都是从学生成绩管理数据库的。

除了用户可以直接使用的数据，还有另外一种数据。它们是有关数据库的定义信息的，如数据库的名称，表的定义，数据库用户名及密码、权限等。这些数据用户不会经常使用，但是对数据库非常重要。这些数据通常存放在“数据字典 (Data Dictionary)”中。数据字典是数据库管理系统中非常重要的组成部分，它是由数据库管理系统自动生成并维护的一组表和视图。数据字典是数据库管理系统工作的依据。

数据库管理系统借助数据字典来理解数据库中数据的组织,并完成对数据库中数据的管理与维护。数据库用户可通过数据字典获取有用的信息,如用户创建了哪些数据库对象,这些对象是如何定义的,这些对象允许哪些用户使用等。但是,数据库用户是不能随便改动数据字典中的内容的。

在收集并抽取出一个应用所需要的大量数据之后,应将其保存起来供进一步查询和加工处理,以获得更多有用的信息。过去人们把数据存放在文件柜里,数据越来越多,从大量的文件中查找数据就会十分困难。现在人们借助数据库,科学地保存和管理大量复杂的数据,从而能方便而又充分地利用这些宝贵的信息资源。

严格地讲,数据库是长期存储在计算机内,有组织的、大量的、可共享的数据集合。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,并可为用户共享。

简而言之,数据库中的数据具有永久存储、有组织和可共享 3 个基本特点。

1.1.2 数据库管理系统 (DBMS)

在建立了数据库之后,下一个问题就是如何科学地组织和存储数据,如何高效地获取和维护数据。完成这个任务的是一个系统软件——数据库管理系统 (Database Management System, DBMS)。

DBMS 是指数据库系统对数据进行管理的软件系统,它是数据库系统的核心组成部分,数据库系统的一切操作,包括查询、更新及各种控制,都是通过 DBMS 进行的。DBMS 是基于数据模型的,因此可以把它看成是某种数据模型在计算机系统上的具体实现。根据所采用数据模型的不同,DBMS 可以分成网状型、层次型、关系型、面向对象型等。但在不同的计算机系统中,由于缺乏统一的标准,即使是同种数据模型的 DBMS,它们在用户接口、系统功能等方面也常常是不同的。

DBMS 把用户对数据库的操作从应用程序带到外部级、概念级,再导向内部级,进而操纵存储器中的数据。一个 DBMS 的主要目标是使数据成为一种可管理的资源。DBMS 应使数据易于为各种用户所共享,应该增进数据的安全性、完整性及可用性,并提供高度的数据独立性。

1.1.3 数据库系统 (DBS)

数据库系统 (Database System, DBS) 是指在计算机系统中引入数据库后的系统,一般由数据库、数据库管理系统 (及其开发工具)、应用系统和数据库管理员构成。应当指出的是,数据库的建立、使用和维护等工作只靠一个 DBMS 是远远不够的,还要有专门的人员来完成,这些人被称为数据库管理员 (Database Administrator, DBA)。

在不引起混淆的情况下,人们常常把数据库系统简称为数据库。数据库系统组成如图 1.1 所示。数据库系统在计算机系统的位置如图 1.2 所示。

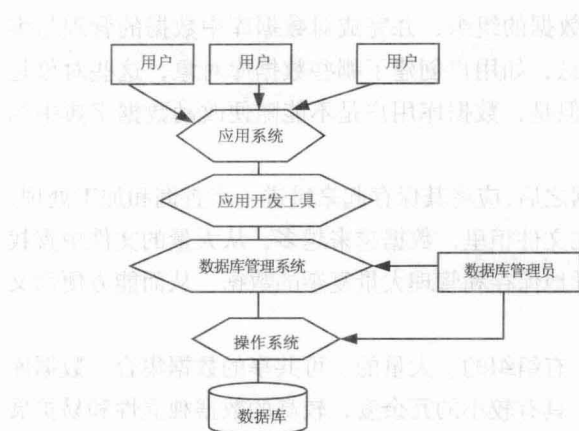


图 1.1 数据库系统



图 1.2 数据库系统在计算机系统中的地位

1.2 数据库技术的产生与发展

在使用计算机之后，数据处理的速度及规模都是过去人工或机械方式无法比拟的。随着数据处理量的不断增加，数据管理技术应运而生，其演变过程随着计算机硬件和软件的发展不断变化，以一个学校的教务处对学生、课程和成绩的管理为例，在没有使用计算机的时候，教务处的工作人员将学生的信息抄写在一张张的卡片上。为了方便查找，将同一个系、同一个年级、同一个班级的学生卡片存放在相邻的地方，并对不同的系、年级和班级做上标签。每门课程的信息也抄写在卡片上，将同一个专业的卡片放在一起并做上标签。每个学期的期末将同一个班的各门课的成绩单收集起来存放在档案中。

当查找一个学生的信息时，如果知道他所在的系和班级，按照标签可以很快找到该学生的卡片。如果只知道他的姓名，那么只有在所有的学生卡片中一个一个查找，需要花费很多时间。

当计算一个学生某个学期的平均成绩时，首先在档案中找到该学生所在班级这个学期的所有成绩单，从中找出该学生各门课程的成绩，再算平均成绩。统计某一门课程的成绩分布时也要进行类似的处理。

计算机及数据库技术的出现，使数据管理人员从繁重的数据管理中解放出来，由计算机和数据库完成了大量的数据管理工作，如自动检索、数据的自动保存、数据分类等。

总的来说，数据库技术的发展经历了以下几个阶段。

1.2.1 人工管理阶段

计算机没有应用到数据管理领域之前，数据管理的工作是由人工完成的。这种处理方式经历了很长时间。

20 世纪 50 年代中期以前，计算机主要用于科学计算。当时的硬件状况是，外存只有纸带、卡片、磁带，没有磁盘等直接存取的存储设备；软件状况是，没有操作系统，没有管理数据的软件，因此称这一阶段的数据管理方式为人工管理数据。人工管理数据具有如下特点。

(1) 数据不保存。由于当时计算机主要用于科学计算，一般不需要将数据长期保存，只是在计算某一课题时将数据输入，用完不保存。

如果要用计算机统计分析全校每一门课的成绩，就要编写统计分析程序，在运行该程序时读

入相应的学生选修课程成绩单等数据,计算完成后数据和程序都不在计算机中保存。

(2) 应用程序管理数据。数据需要由应用程序自行管理,没有相应的软件系统负责数据的管理工作。应用程序中不仅要规定数据的逻辑结构,而且要设计物理结构,包括存储结构、存取方法、输入方式等,因此程序员负担很重。

(3) 数据不共享。数据是面向应用的,一组数据只能对应一个程序。当多个应用程序涉及某些相同的数据时,由于必须各自定义,无法互相利用、互相参照,因此程序与程序之间有大量的冗余数据。例如,教务处既要统计某一门课的成绩又要分析某一个学生的成绩时,就要编写两个程序,尽管都要使用学生选修课成绩单,但是每个程序要分别定义两个成绩单数据,分别输入,分别使用,如图 1.3 所示。

(4) 数据不具有独立性。数据的逻辑结构或物理结构改变后,必须对应用程序做相应的修改,这就进一步加重了程序员的负担。例如,学生成绩由 5 级记分制改为百分制时,上面两个统计分析程序都要修改。

在人工管理阶段,程序与数据之间的对应关系如图 1.4 所示。

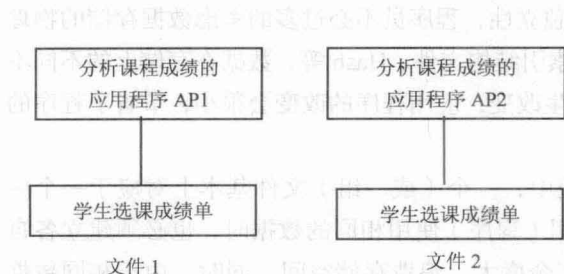


图 1.3 两个应用程序使用同一数据



图 1.4 人工管理阶段应用程序与数据之间的对应关系

1.2.2 文件系统阶段

20 世纪 50 年代后期到 60 年代中期,硬件方面已有了磁盘、磁鼓等直接存储设备;软件方面,操作系统中已经有了专门的数据管理软件——文件系统。可以把相关的数据组织成一个文件存放在计算机中,需要时只要提供文件名,计算机就能从文件系统中找出所要的文件,把文件中存储的数据提供给用户进行处理。

例如,在学校教务处对学生学籍的管理中,为了改变查找、计算工作量大及花费时间长的被动局面,将学生卡片、课程卡片和学生学习成绩单中的内容存放到文件“Student”、“Course”、“Study”中,并对每个文件编写一组程序用于数据维护,包括增加、删除、修改和查询一条记录。在此基础上根据需要编写一些查询和报表打印程序,如根据学生的姓名、学生编号查找学生的信息,统计某学期某个学生的平均成绩,统计某门课的平均成绩等。

管理系统投入应用后,教务处的工作效率大大提高。例如,每学期末将各门课的考试成绩输入计算机以后,可以很快计算出学生的平均成绩,并打印出需要补考的学生的名单。

但是,由于数据的组织仍然是面向程序,所以仍然存在大量的数据冗余,经过一段时间的使用后,教务人员发现有时必须修改程序和文件结构才能适应工作的需要。例如,学校领导要求统计不同生源地的学生成绩,因为原来的学籍管理软件中没有实现这个功能,必须编写一段程序来实现。又如,当需要在“Student”文件中增加“个人网址”属性时,这涉及改变文件的结构,需要若干步骤才能完成。

第一步,建立一个新文件“Student-new”,其结构是在“Student”的结构中加入“个人网址”这一项。

第二步,编写一个程序将文件“Student”中的数据转存到“Student-new”中。

第三步,删除文件“Student”。

第四步,将文件“Student-new”重命名为“Student”。

这项工作到此并没有结束,因为文件中保存的是数据,不保存数据的结构,数据结构是在程序中定义的,旧“Student”文件的结构写到了所有使用它的程序中,必须一一修改这些程序以适应新的文件结构,否则程序运行就会出错。

从这个例子可以看到用文件系统管理数据的优点和不足。一般地讲,用文件系统管理数据具有如下特点。

(1) 数据可以长期保存。数据可以组织成文件长期保存在计算机中反复使用。

(2) 由文件系统管理数据。文件系统把数据组织成内部有结构的记录,实现“按文件名访问,按记录进行存取”的管理技术。

文件系统使应用程序与数据之间有了初步的独立性,程序员不必过多的考虑数据存储的物理细节。例如,文件系统中可以有顺序结构文件、索引结构文件、Hash 等,数据在存储上的不同不会影响程序的处理逻辑。如果数据的存储结构发生改变,应用程序的改变会很小,节省了程序的维护工作量。但是,文件系统仍存在以下缺点。

(1) 数据共享性差,冗余度大。在文件系统中,一个(或一组)文件基本上对应于一个应用(程序),即文件是面向应用的。当不同的应用(程序)使用相同的数据时,也必须建立各自的文件,而不能共享相同的数据。因此数据的冗余度大,浪费存储空间。同时,由于相同数据的重复存储、各自管理,容易造成数据的不一致性,给数据的修改和维护带来了困难。

(2) 数据独立性差。文件系统中的文件是为某一特定应用服务的,文件的逻辑结构对该应用来说是优化的,因此要对现有的数据再增加一些新的应用会很困难,系统不容易扩充。一旦数据的逻辑结构发生改变,就必须修改应用程序,修改文件结构的定义。因此数据与程序之间仍缺乏独立性。文件系统阶段程序与数据之间的关系如图 1.5 所示。

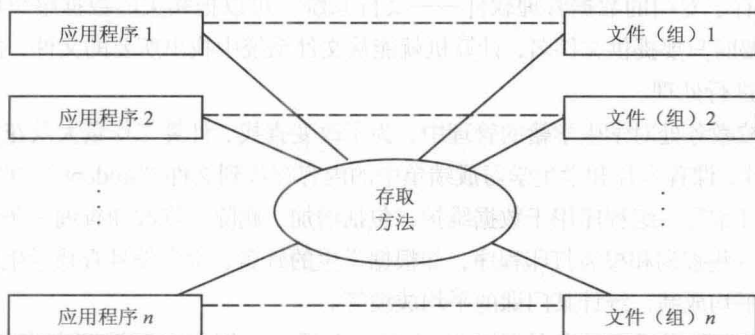


图 1.5 文件系统阶段应用程序与数据之间的对应关系

1.2.3 数据库系统阶段

20 世纪 60 年代后期,计算机用于管理的规模越来越大,应用越来越广泛,数据量急剧增长,

同时对多种应用、多种语言互相覆盖的共享数据集合的需求也越来越强烈。

这时已有大容量磁盘,硬件价格下降;软件价格则上升,为编制和维护系统软件及应用程序所需的成本相对增加。在这种背景下,以文件系统作为数据管理手段已经不能满足应用的需求。于是,为解决多用户、多应用共享数据的需求,使数据为尽可能多的应用服务,数据库技术便应运而生,出现了统一管理数据的专用软件系统——数据库管理系统。

用数据库系统来管理数据和使用文件系统相比具有明显的优点,从文件系统到数据库系统,标志着数据管理技术的飞跃。

例如,对教务管理系统,学校决定采用数据库技术,购买了一个关系数据库管理系统(RDBMS),在这个RDBMS之上建立一个应用系统,将教务处和学生工作处保存的学生数据进行综合设计,供全校各院系的教师 and 教务人员共享访问和使用。

在系统中要建立3个关系:学生基本表“Student”、课程基本表“Course”和学习成绩基本表“SC”。在数据库系统中只要用DDL语言向RDBMS提交CREATE TABLE语句就可以了,例如建立关系“Student”:

```
CREATE TABLE Student
(
    Sno      CHAR(10) NOT NULL UNIQUE,
    Sname    CHAR(10),
    Sex      CHAR(2),
    Age      INT,
    Classno  CHAR(6),
    Dept     CHAR(12) );
```

这条语句在数据库中建立了一个关系“Student”,用来保存学生的信息。更重要的是,RDBMS将“Student”的结构也保存到数据库的数据字典中。

向关系中增加、删除、修改一个元组(记录)用RDBMS提供的语句INSERT、DELETE、UPDATE来完成。这些语句如何操作磁盘上的数据是由RDBMS来完成的,程序员不用编写专门的程序,从而节省了程序员大量的时间和精力。

因为关系“Student”的结构和数据都由RDBMS管理,在教务管理系统中,向关系“Student”添加“E-mail”属性时,既不需要编写转储数据的程序,也不用修改那些使用了关系“Student”的程序。

从这个例子可以看出,数据库系统阶段应用程序与数据之间的对应关系如图1.6所示。

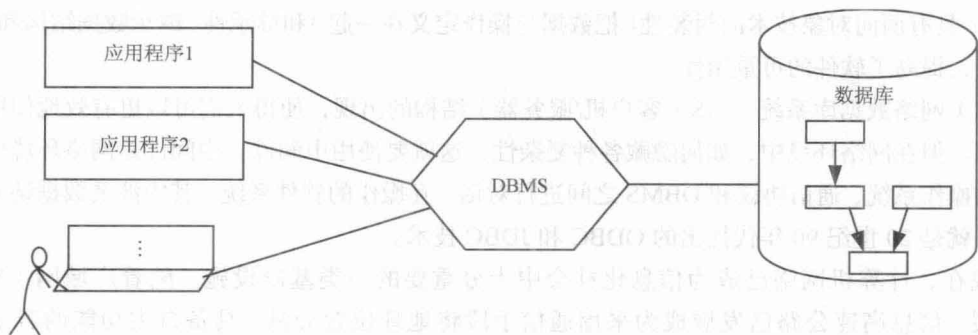


图 1.6 数据库系统阶段应用程序与数据之间的对应关系

由于数据库是以数据为中心组织数据,减少数据的冗余,提供更高的数据共享能力,同时要

求程序和数据具有较高的独立性,因此当数据的逻辑结构改变时,不涉及数据的物理结构,也不影响应用程序,这样就降低了应用程序研制与维护的费用。

1.2.4 高级数据库阶段

这一阶段的主要标志是 20 世纪 80 年代的分布式数据库系统、90 年代的对象数据库系统和 21 世纪初的网络数据库系统的出现。

(1) 分布式数据库系统。在这一阶段以前的数据库系统是集中式的。在文件系统阶段,数据分散在各个文件中,文件之间缺乏联系。集中式数据库把数据库集中在一个数据库中进行管理,减少了数据冗余和不一致性,而且数据联系比文件系统更强。但集中式系统也有弱点:一是随着数据量增加,系统会变得相当庞大,操作复杂,开销大;二是数据集中存储,大量的通信都要通过主机,造成拥挤现象。随着小型计算机和微型计算机的普及,计算机网络软件和远程通信的发展,分布式数据库系统崛起了。

分布式数据库系统主要有以下 3 个特点。

① 数据库的数据物理上分布在不同地方,但逻辑上是一个整体。

② 各个分散的数据库既可以执行局部应用(访问本地数据库),又可以执行全局应用(访问异地数据库)。

③ 各分散的计算机由数据通信网络相连。本地计算机不能单独胜任的处理任务,可以通过通信网络取得其他数据库和计算机的支持。

分布式数据库系统兼顾了集中管理和分布处理两个方面,因而有良好的性能。

(2) 对象数据库系统。在数据处理领域,关系数据库的使用已相当普遍、相当出色。但是现实世界存在着许多具有更复杂数据结构的实际应用领域,已有的层次、网状、关系三种数据模型对这些应用领域都显得力不从心。如多媒体数据、多维表格数据、CAD 数据等应用问题,需要更高级的数据库技术来支持,以便管理、构造与维护大容量的持久数据,并使它们能与大型复杂程序紧密结合。对象数据库正是由于适应这种形势而发展起来的,它是面向对象的程序设计技术与数据库技术结合的产物。

对象数据库系统主要有以下 2 个特点。

① 对象数据库模型能完整地描述现实世界的数据结构,能表达数据间嵌套、递归等关系。

② 具有面向对象技术的封装性(把数据与操作定义在一起)和继承性(继承数据结构和操作)的特点,提高了软件的可重用性。

(3) 网络数据库系统。C/S(客户机/服务器)结构的出现,使得人们可以更有效地使用计算机资源。但在网络环境中,如何隐藏各种复杂性,这就要使用中间件。中间件是网络环境中保证不同的操作系统、通信协议和 DBMS 之间进行对话、互操作的软件系统。其中涉及数据访问的中间件,就是 20 世纪 90 年代提出的 ODBC 和 JDBC 技术。

现在,计算机网络已成为信息化社会中十分重要的一类基础设施。随着广域网(WAN)的发展,信息高速公路已发展成为采用通信手段将地理位置分散,具备自主功能的若干台计算机和数据库系统有机地连接起来组成的因特网(Internet),用于实现通信交往、资源共享或协调工作等目标。这个目标在 20 世纪末已经实现,正在对社会的发展起着巨大的推进作用。