

高等学校规划教材
国家精品课程系列教材



数据库原理 与实践 (Access版)

董卫军 邢为民 索琦 编著
耿国华 主审

计算机公共基础课程群



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

高等学校规划教材
国家精品课程系列教材

数据库原理与实践

(Access 版)

董卫军 邢为民 索琦 编著
耿国华 主审

电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书是国家精品课程“计算机基础”系列课程“数据库原理与技术”的主教材，全书由数据库基础理论篇和 Access 应用实践篇两大部分组成，第一部分为数据库基础理论，介绍数据库组织、管理和使用的一般知识，包括数据库基础、数据库设计和数据库安全 3 个章节；第二部分为 Access 应用实践，主要介绍数据库原理在 Access 上的实现，内容包括 Access 简介、创建数据库、查询数据、窗体的使用、数据报表、数据访问页、宏、VBA 和模块、综合实例 9 个章节。

本书基础理论部分与 Access 应用实践部分相辅相成，既照顾到理论基础的坚实，又强调技术实践的应用。同时，在编写时兼顾了计算机等级考试的要求。为方便教学，本书还配有电子课件，任课教师可以登录华信教育资源网（www.hxedu.com.cn）免费注册下载。

本书可作为高等院校计算机基础课程，以及相关专业的数据库技术课程的教材，也可作为全国计算机等级考试二级 Access 的培训或自学教材。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目(CIP)数据

数据库原理与实践：Access 版 / 董卫军，邢为民，索琦编著. —北京：电子工业出版社，2011.6

高等学校规划教材

ISBN 978-7-121-13385-5

I. ①数… II. ①董…②邢…③索… III. ①关系数据库—数据库管理系统，Access—高等学校—教材
IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2011)第 074940 号

策划编辑：索蓉霞

责任编辑：索蓉霞 特约编辑：李 涛

印 刷：三河市鑫金马印装有限公司

装 订：

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×1092 1/16 印张：15 字数：384 千字

印 次：2011 年 6 月第 1 次印刷

定 价：29.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：(010) 88258888。

前 言

数据库技术是计算机科学技术发展的重要内容,是构成信息系统的重要基础。数据库技术已广泛应用于各类管理信息系统(MIS)、决策支持系统(DDS)、计算机辅助软件工程(CASE)等领域。同时,数据技术与网络通信技术、多媒体技术等新技术的结合又促进了数据库技术的进一步发展。在信息已作为战略资源的今天,建设以数据库为核心的各类信息系统和应用系统,对提高企业效益、改善部门管理具有重要意义。因此,学习和掌握数据库技术的基本知识和基本技能已成为大学生必备的素质要求。

本书是国家精品课程“计算机基础”系列课程“数据库原理与技术”的主教材,教材从培养学生分析问题和解决问题的能力入手,立足于“以理论为基础,以实例为引导,以应用为目的”,采用理论结合实例的模式努力把知识点融入到案例中,做到通俗易懂,循序渐进,满足应用型人才培养的特点和需求。在体系结构安排上,尽可能地将理论、知识点、实例相结合,力求在内容选定、教学方式和方法、对学生的学习方法的指导等方面全面地贯彻素质教育理念。

本教材共12章,分为数据库基础理论篇和Access应用实践篇两部分。

数据库基础理论篇主要介绍数据库组织、管理和使用的一般知识,包括数据库基础、数据库设计和数据库安全3个章节,主要包含数据库系统、数据模型、关系运算、数据库设计、关系规范化、数据库安全等方面的知识,重点介绍了从基本原理到系统设计再到系统安全三环节的技术和方法,使读者能够对数据库有一个从外到里,由浅入深的理解。

Access应用实践篇从实用性出发,以Access为基础介绍了Access数据库程序设计所涉及的基本概念、数据表设计方法和程序设计方法。主要内容包括:Access简介、创建数据库、查询数据、窗体的使用、数据报表、数据访问页、宏、VBA和模块、综合案例9个章节。通过引例,循序渐进地介绍了数据库的设计、建立与使用方法,能够让读者在很短的时间内掌握Access数据库设计,实现信息的有效管理。

本书基础理论部分与Access应用实践部分相辅相成,既照顾到理论基础的坚实,又强调技术实践的应用。同时,在编写时兼顾了计算机等级考试的要求。为方便教学,本书还配有电子课件,任课教师可以登录华信教育资源网(www.hxedu.com.cn)免费注册下载。

本书由多年从事计算机教学的一线教师编写,董卫军编写第1~3章和第10~11章,邢为民编写第4~7章;索琦编写第8~9章和第12章;全书由董卫军统稿,西北大学耿国华教授主审。感谢教学团队成员的帮助,感谢西北大学教务处多年来的支持。由于水平有限,书中难免有不妥之处,恳请指正。

编 者
于西安·西北大学

目 录

数据库基础理论篇

第 1 章 数据库基础	1	1.6.1 三级模式	22
1.1 数据库基本概念	1	1.6.2 两种映射	23
1.1.1 数据与数据处理	1	1.7 访问数据库的主要过程	24
1.1.2 数据库	2	1.8 关系数据库	24
1.1.3 数据库管理系统	4	1.8.1 基本概念	24
1.1.4 数据库系统的构成	4	1.8.2 关系数据库的体系结构	25
1.2 数据管理的发展	6	1.8.3 关系模型的完整性规则	27
1.2.1 人工管理阶段	6	1.8.4 关系代数	29
1.2.2 文件系统阶段	6	习题 1	31
1.2.3 数据库系统阶段	6	第 2 章 数据库设计	36
1.2.4 分布式数据库系统阶段	7	2.1 工程化设计思想	36
1.2.5 数据管理新技术	7	2.1.1 软件工程	36
1.3 常用数据库管理系统	10	2.1.2 软件生命周期	36
1.3.1 IBM 的 DB2	10	2.2 数据库设计概述	37
1.3.2 Oracle	11	2.3 数据库设计的基本步骤	37
1.3.3 Informix	12	2.3.1 规划	38
1.3.4 Sybase	12	2.3.2 需求分析	39
1.3.5 SQL Server	14	2.3.3 概念结构设计	40
1.3.6 PostgreSQL	14	2.3.4 逻辑结构设计	41
1.3.7 MySQL	15	2.3.5 物理设计	44
1.3.8 Access 数据库	15	2.3.6 系统实施	45
1.3.9 FoxPro 数据库	16	2.3.7 运行和维护	45
1.4 客观世界的表示	16	2.4 关系模式的规范化	46
1.4.1 现实世界	17	2.4.1 函数依赖和键	47
1.4.2 概念世界	17	2.4.2 关系模式的范式	48
1.4.3 数据世界	19	2.4.3 模式分解	49
1.5 数据模型	19	习题 2	52
1.5.1 数据模型的概念	19	第 3 章 数据库安全	56
1.5.2 数据模型组成要素	20	3.1 数据库安全	56
1.5.3 逻辑模型的基本类型	20	3.1.1 系统安全的主要风险	56
1.6 数据库的体系结构	22	3.1.2 数据库安全的特征	57

3.2 数据库的安全层次	58	3.3.1 数据库加密技术	60
3.2.1 网络系统层次安全	59	3.3.2 存取管理技术	61
3.2.2 操作系统层次安全	59	3.3.3 安全审计技术	62
3.2.3 数据库管理系统层次安全	59	3.3.4 备份与恢复	65
3.3 数据库安全技术	60	习题 3	69

Access 应用实践篇

第 4 章 Access 简介	71	删除	92
4.1 Access 简介	71	5.5.2 图片、声音和影像的输入	93
4.1.1 Access 产生与发展	71	5.5.3 建立超链接	93
4.1.2 Access 安全方式	72	5.5.4 设置数据表格形式	94
4.2 Access 基本概念	73	5.5.5 查找和替换数据	94
4.2.1 数据库和数据库对象	73	5.5.6 排序与筛选	94
4.2.2 Access 基本功能	75	5.6 数据的导入和导出	96
4.2.3 Access 操作模式	76	5.6.1 导入和链接	96
4.3 Access 的启动和退出	76	5.6.2 链接外部数据库	96
4.3.1 启动 Access	76	5.6.3 导入外部数据库	97
4.3.2 退出 Access	77	5.6.4 导出数据	98
习题 4	78	习题 5	98
第 5 章 创建数据库	80	第 6 章 查询数据	102
5.1 创建数据库	80	6.1 查询与表	102
5.1.1 创建一个空数据库	80	6.2 常见的查询	102
5.1.2 通过模板建立数据库	81	6.2.1 选择查询	102
5.2 创建表	82	6.2.2 参数查询	102
5.2.1 用表向导建立表	82	6.2.3 交叉表查询	103
5.2.2 通过设计器创建表	83	6.2.4 操作查询	103
5.3 表间关系的创建	87	6.2.5 SQL 查询	103
5.3.1 表的索引	87	6.3 创建选择查询	104
5.3.2 表间关系	88	6.3.1 利用查询设计视图建立	104
5.3.3 创建关系	89	6.3.2 利用查询向导建立	107
5.3.4 编辑和删除关系	90	6.4 创建参数查询	107
5.4 修改数据库结构	90	6.5 创建交叉表查询	108
5.4.1 对表的操作	90	6.6 操作查询	109
5.4.2 对字段的操作	91	6.6.1 删除查询	109
5.4.3 对表的行与列操作	92	6.6.2 更新查询	110
5.5 表中数据的操作	92	6.6.3 生成表查询	110
5.5.1 普通数据的添加、修改和		6.6.4 追加查询	111
		6.7 Access SQL 查询	112

6.7.1	SQL 的特点	112	8.3	使用报表向导建立报表	146
6.7.2	SQL 数据库的体系结构	113	8.4	自动报表与图表向导	147
6.7.3	Access SQL 的特点	113	8.4.1	创建自动报表	147
6.7.4	Access SQL 的数据定义	115	8.4.2	创建图表	148
6.7.5	Access SQL 的数据查询	117	8.5	通过设计视图创建报表	149
6.7.6	Access SQL 的数据更新	120	8.5.1	创建过程	149
6.7.7	建立 SQL 查询	121	8.5.2	报表控件	150
6.8	查询的打开与修改	122	8.5.3	在报表中添加分组	151
6.8.1	打开查询	122	8.5.4	添加计算字段	152
6.8.2	修改查询	122	8.6	修饰报表	153
习题 6	123	8.6.1	添加文字	153
第 7 章	窗体的使用	128	8.6.2	设置内容的显示效果	153
7.1	窗体的功能与构成	128	8.6.3	调整显示对齐方式	153
7.1.1	窗体的功能	128	8.6.4	调整行间距	154
7.1.2	窗体的构成	128	8.7	打印报表	154
7.2	创建窗体	129	8.7.1	页面设置	154
7.2.1	自动窗体创建	129	8.7.2	预览与打印报表	154
7.2.2	向导创建	130	习题 8	155
7.3	窗体设计视图与控件	130	第 9 章	数据访问页	158
7.3.1	窗体的设计视图	130	9.1	数据访问页的基本概念	158
7.3.2	工具箱的使用	131	9.1.1	数据访问页	158
7.3.3	窗体中的常见控件	131	9.1.2	数据访问页视图	158
7.3.4	给对象添加控件	132	9.1.3	数据访问页的组成	159
7.4	创建子窗体	136	9.2	创建数据访问页	159
7.5	通过窗体处理数据	137	9.2.1	自动创建	159
7.5.1	窗体视图工具栏	137	9.2.2	使用向导创建	160
7.5.2	记录导航按钮集	137	9.2.3	使用设计器创建	160
7.5.3	处理数据	138	9.3	美化数据访问页	161
7.6	创建切换面板	138	9.3.1	添加标签	161
7.6.1	切换面板的作用	138	9.3.2	添加命令按钮	162
7.6.2	切换面板的创建	139	9.3.3	添加滚动文字	162
习题 7	140	9.3.4	添加主题和背景	163
第 8 章	数据报表	143	9.3.5	添加 Office 电子表格	164
8.1	报表的功能与类型	143	9.3.6	建立超链接	164
8.1.1	报表的功能	143	习题 9	164
8.1.2	报表类型	143	第 10 章	宏	166
8.2	报表设计的组成	144	10.1	理解宏	166
8.2.1	报表的节	144	10.1.1	宏的分类	166
8.2.2	报表的常见节	145	10.1.2	宏的介绍	166

10.2 宏创建与运行	170	11.6.3 选择结构	197
10.2.1 创建宏	170	11.6.4 循环结构	200
10.2.2 运行宏	175	11.6.5 其他辅助控制语句	202
10.3 宏向 Visual Basic 代码转换	176	11.7 VB 的过程	204
10.3.1 转换窗体或报表中的宏	176	11.7.1 Sub 过程	204
10.3.2 转换全局宏	176	11.7.2 Function 过程	206
习题 10	176	11.7.3 参数的传递	208
第 11 章 VBA 和模块	179	11.7.4 过程和变量的作用域	208
11.1 VBA 简介	179	11.8 面向对象程序设计基础	
11.1.1 宏与 VBA 的差异	179	知识	210
11.1.2 VBA 与 VB 的区别	179	11.8.1 几个基本概念	210
11.2 VBA 开发环境	180	11.8.2 VBA 程序设计的基本	
11.2.1 开发环境界面	180	过程	211
11.2.2 编写代码	180	11.9 应用举例	214
11.2.3 模块的调试	181	习题 11	219
11.3 VBA 基础	181	第 12 章 综合实例	224
11.3.1 VB 语言元素	181	12.1 需求分析	224
11.3.2 书写规则	183	12.2 系统设计	224
11.4 Visual Basic 语言基础	184	12.3 系统实现	226
11.4.1 数据类型	184	12.3.1 创建数据库	226
11.4.2 常量与变量	187	12.3.2 查询设计	226
11.4.3 表达式与运算符	190	12.3.3 报表设计	227
11.5 数组	193	12.3.4 窗体设计	227
11.5.1 基本概念	193	12.4 系统测试及运行	230
11.5.2 一维数组的使用	193	习题 12	231
11.5.3 多维数组的使用	195	参考文献	232
11.6 基本程序设计	196		
11.6.1 程序的基本结构	196		
11.6.2 顺序结构	196		

数据库基础理论篇

第 1 章 数据库基础

数据是计算机处理的对象，因而如何有效地管理数据就是一个重要的问题。20 世纪 50 年代中期以前没有专门用于数据管理的软件。操作系统出现以后，可以通过操作系统管理数据。但由于文件之间缺乏联系，很难解决重复存储和不一致的问题。20 世纪 60 年代末提出了数据库的概念，它能为多个用户、多种应用所共享，具有最小的冗余度，较高的数据独立性。

1.1 数据库基本概念

1.1.1 数据与数据处理

在实际的工作和学习中，每个人都有很多亲戚和朋友，为了保持与他们的联系，我们常常通过一定的方法将他们的姓名、地址、电话等信息都记录在通讯录中（存于一个笔记本中，或存于手机中），这样在查找时就很方便了。这个“通讯录”就是一个最简单的“数据库”，而每个人的姓名、地址、电话等信息就是这个数据库中的“数据”。我们可以根据需要随时在这个“数据库”中添加新朋友的个人信息，也可以修改某个人的电话号码“数据”。

1. 数据

数据是事物特性的反映和描述，是符号的集合。数据不仅包括狭义的数值数据，而且包括文字、声音、图形、图像等一切能被计算机接收并处理的符号。数据在空间上的传递称为通信（以信号方式传输），数据在时间上传递称为存储（以文件形式存取）。信息是和数据关系密切的另外一个概念，数据是信息的符号表示（或称为载体），信息则是数据的内涵，是对数据语义的解释。

2. 数据处理

数据处理可定义为对数据的收集、存储、加工、分类、检索、传播等一系列活动。这些活动是数据处理业务的基本环节，是任何数据处理业务中必不可少的共有部分。对于数据处理而言，其基本目的是从大量的、杂乱无章的、难以理解的数据中抽取并推导出对某些特定用户有价值、有意义的数据。数据管理和数据处理有一定的联系，数据管理技术的优劣，将直接影响数据处理的效率。

1.1.2 数据库

1. 数据库的概念

数据库是存放数据的仓库,是对现实世界有用信息的抽取、加工和处理,并按一定格式长期存储在计算机内的、有组织的、可共享的数据集合。数据库的概念包括两层意思:

(1) 数据库是能够合理保管数据的“仓库”,用户在该“仓库”中存放要管理的事务数据,“数据”和“库”两个概念相结合构成数据库;

(2) 数据库包含数据管理的新方法和新技术,通过数据库可以方便地组织数据、维护数据、控制数据和利用数据。

2. 数据库的特点

相对于传统的人工管理和文件技术,数据库具有如下特点。

(1) 数据结构化

数据库中的数据是从全局观点出发建立的,按一定的数据模型进行组织、描述和存储。其结构基于数据间的自然联系,数据面向全组织,不针对特定应用,具有整体的结构化特征。

(2) 数据独立性

数据独立性是指数据的逻辑组织方式和物理存储方式与用户的应用程序相对独立。数据库的数据独立性包括物理独立性和逻辑独立性两个方面。

数据的物理独立性是指当数据的物理存储改变时,应用程序不用改变。换言之,用户的应用程序与数据库中的数据是相互独立的。数据在数据库中的存储形式是由 DBMS 管理的,应用程序要处理的只是数据的逻辑结构。

数据的逻辑独立性是指当数据的逻辑结构改变时,用户应用程序不用改变。也就是说,用户的应用程序与数据库的逻辑结构是相互独立的。

通过数据独立性可以将数据的定义和描述从应用程序中分离出来。数据的存取由 DBMS 管理,用户不必考虑存取路径等细节,从而简化了应用程序的编制,大大减少了应用程序的维护和修改工作量。

(3) 数据冗余低

数据的冗余度是指数据重复的程度。数据库系统从整体角度描述数据,数据可以被多个应用共享。这样可以节约存储空间、减少存取时间,而且可以避免数据之间的不相容和不一致。但还需要必要的冗余,必要的冗余可保持数据间的联系。

(4) 统一的数据管理和控制

数据库对系统中的用户来说是共享资源。计算机的共享一般是并发的,即多个用户可以同时存取数据库中的数据,甚至可以同时存取数据库中同一个数据。因此,数据库管理系统必须提供以下几方面的数据控制保护功能。

① 数据的安全性保护

数据的安全性保护可以防止不合法的使用所造成的数据泄密和破坏,使每个用户只能按规定对某种数据以某些方式进行使用和处理。例如,可以使用身份鉴别、检查口令或其他手段来检查用户的合法性,只有合法用户才能进入数据库系统。

② 数据的完整性控制

数据的完整性指数据的正确性、有效性和相容性。完整性检查可以保证数据库中的数据在输入和修改过程中始终符合原来的定义和规定,在有效的范围内保证数据之间满足一定的

关系。例如，月份是 1~12 之间的正整数，性别是“男”或“女”，成绩是大于等于 0 且小于等于 100 的整数，学生的学号是唯一的。

③ 数据库恢复

计算机系统的硬件故障、软件故障、操作员的失误以及人为的攻击和破坏，都会影响数据库中数据的正确性，甚至会造成数据库部分或全部数据丢失。数据库恢复机制能够及时发现和修复故障，从而防止数据被破坏。

④ 并发控制

当多个用户的并发进程同时存取、修改数据库时，可能会发生由于相互干扰而导致结果错误的情况，并使数据库完整性遭到破坏。因此，必须对多用户的并发操作加以控制和协调。

3. 数据的组织级别

数据库中数据的组织一般可以分为四级：数据项、记录、文件和数据库，它们之间的关系如图 1.1 所示。

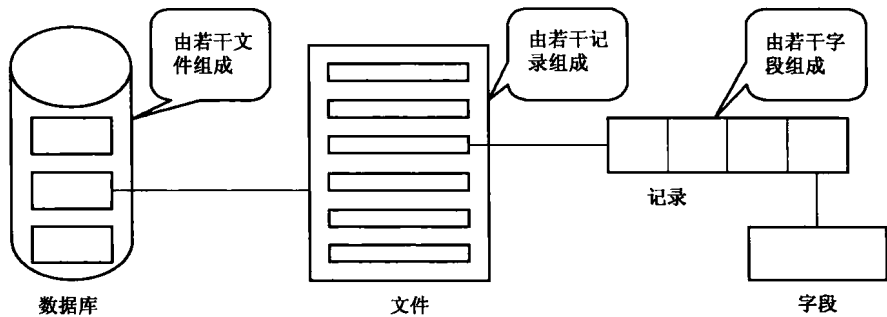


图 1.1 数据组织级别

(1) 数据项

数据项是定义数据的最小单位，也叫元素、基本项、字段等。数据项与现实世界实体的属性相对应，每个数据项都有一个名称，称为数据项名。数据项的值可以是数值、字母、字母数字、汉字等形式。数据项的取值有一定的范围，称为域，域以外的任何值对该数据项都是无意义的。

(2) 记录

记录由若干相关联的数据项组成，是处理和存储信息的基本单位，是关于一个实体的数据总和。构成记录的数据项表示实体的若干属性。记录有“型”和“值”的区别，“型”是同类记录的框架，它定义记录的组成；而“值”是记录反映的实体的内容。为了唯一标识每个记录，就必须有记录标识符（也叫关键字）。能唯一标识记录的关键字称为主关键字，其他标识记录的关键字称为辅关键字。

(3) 文件

文件是一给定类型的记录的全部具体值的集合，文件用文件名称标识。由于数据库文件可以看成是具有相同性质的记录的集合，因而数据库文件具有以下特性：

- ① 文件的记录格式相同，长度相等；
- ② 不同的行是不同的记录，因而具有不同的内容；
- ③ 不同的列表示不同的字段，同一列中的数据的性质（属性）相同；
- ④ 每一行各列的内容是不能分割的，但行的顺序和列的顺序不影响文件内容的表达。

(4) 数据库

数据库是比文件更大的数据组织形式，数据库是具有特定联系的数据的集合，也可以看

成是具有特定联系的多种类型的记录的集合。数据库的内部构造是文件的集合，这些文件之间存在某种联系，不能孤立存在。

1.1.3 数据库管理系统

数据库管理系统 (Database Management System, DBMS) 是位于用户与操作系统 (OS) 之间的一层数据管理软件，是数据库系统的核心。它为用户或应用程序提供访问数据库的方法，包括数据库的建立、查询、更新以及各种数据控制。DBMS 总是基于某种数据模型 (层次型、网状型、关系型和面向对象型等)。DBMS 具有以下基本功能。

1. 对象定义功能

DBMS 通过提供数据定义语言 (Data Definition Language, DDL) 实现对数据库中数据对象的定义，如对外模式、模式和内模式加以描述和定义；数据库完整性的定义；安全保密的定义，如用户口令、级别、存取权限；存取路径的定义，如索引的定义。

2. 数据操纵功能

DBMS 提供数据操纵语言 (Data Manipulation Language, DML) 实现对数据库中数据的一些基本操作，如检索、插入、修改、删除和排序等。DML 一般有两类：一类是嵌入主语言的 DML，如嵌入到 C++ 或 PowerBuilder 等高级语言 (称为宿主语言) 中；另一类是非嵌入式语言 (包括交互式命令语言和结构化语言)，它的语法简单，可以独立使用，由单独的解释或编译系统来执行，所以一般称为自主型或自含型的 DML。

3. 数据库控制功能

DBMS 提供的数据库控制语言 (Data Control Language, DCL) 保证数据库操作都在统一的管理下协同工作以确保事务处理的正常运行，同时保证数据库的正确性、安全性、有效性和多用户对数据的并发使用以及发生故障后的系统恢复等。数据库控制功能包括并发控制和存取控制，如安全性检查、完整性约束条件检查、数据库内部 (如索引) 和数据字典的自动维护、缓冲区大小的设置等。

4. 数据组织、存储和管理

数据组织和存储的基本目标是提高存储空间利用率和方便存取，提供多种存取方法 (如索引查找、Hash 查找、顺序查找等) 提高存取效率。DBMS 要分类组织、存储和管理各种数据，包括数据字典、用户数据、存取路径等。要确定以何种文件结构和存取方式在存储级上组织这些数据，如何实现数据之间的联系。

1.1.4 数据库系统的构成

数据库系统 DBS (Database System) 主要由硬件、数据库、软件 and 用户 4 部分构成，其基本结构如图 1.2 所示。

1. 硬件

硬件是数据库依存的物理设备，运行数据库系统的计算机硬件需要足够大的内存以存放系统软件，需要足够多的联机直接存取设备 (如大容量磁盘) 存储庞大的数据，需要足够多的脱机存储介质 (如磁盘、光盘、磁带等) 以存放备份数据，需要较高的通道能力，以提高数据传送速率，并且要求系统互联，以实现数据共享。

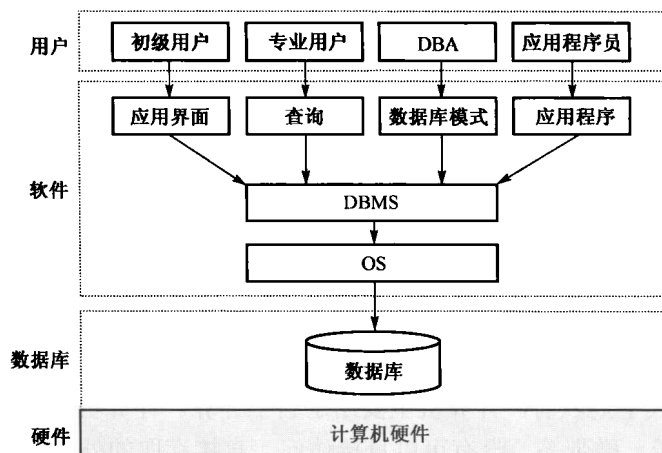


图 1.2 DBS 的基本构成

2. 数据库

数据库是数据库系统的核心和管理对象，是存储在一起相互有联系的数据集合。数据库中的数据是集成的、共享的、最小冗余的，能为多种应用服务。数据按照数据模型所提供的形式框架存放在数据库中。

3. 软件

数据库软件主要是指数据库管理系统 DBMS，DBMS 是为数据库存取、维护和管理而配置的软件，它是数据库系统的核心组成部分，DBMS 在操作系统支持下工作。数据库软件除了 DBMS 之外，还要有数据库应用系统，数据库应用系统是为特定应用而开发的数据库应用软件。数据库管理系统为数据的定义、存储、查询和修改提供支持，而数据库应用系统是对数据库中的数据进行处理和加工的软件，它面向特定应用。如管理信息系统、决策支持系统和办公自动化等都属于数据库应用系统。

4. 用户

数据库系统中存在一组参与分析、设计、管理、维护和使用数据库的人员，他们在数据库系统的开发、维护和应用中起着重要的作用。分析、设计、管理和使用数据库系统的人员主要是：数据库管理员、系统分析员、应用程序员和最终用户。

(1) 数据库管理员

数据库管理员负责全面管理和控制数据库系统，主要职责包括：设计与定义数据库系统，帮助最终用户使用数据库系统，监督与控制数据库系统的使用和运行，改进和重组数据库系统，调整数据库系统的性能，转储与恢复数据库，重构数据库。

(2) 系统分析员

系统分析员是数据库系统建设期的主要参与人员，负责应用系统的需求分析和规范说明，确定系统的基本功能和数据库结构，设计应用程序和硬软件配置并组织整个系统的开发。

(3) 应用程序员

应用程序员根据系统的功能需求负责设计和编写应用系统的程序模块，并参与对程序模块的测试。

(4) 最终用户

数据库系统的最终用户是有不同层次的，不同层次的用户其需求的信息以及获得信息的

方式也是不同的。一般可将最终用户分为操作层、管理层和决策层。他们通过应用系统的用户接口使用数据库。

1.2 数据管理的发展

数据管理包括数据组织、分类、编码、存储、检索和维护。随着硬件、软件技术的发展及计算机应用范围的扩大，数据管理经历了四个阶段。

1.2.1 人工管理阶段

20 世纪 50 年代中期以前，计算机主要用于科学计算。计算机的软硬件均不完善，硬件方面只有卡片、纸带、磁带等，没有可以直接访问、直接存取的外部存取设备。软件方面还没有操作系统，也没有专门管理数据的软件，数据由程序自行携带，数据与程序不能独立，数据不能长期保存，如图 1.3 所示。

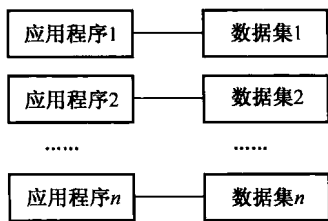


图 1.3 人工管理数据

在人工管理阶段，程序员在程序中不仅要规定数据的逻辑结构，还要设计其物理结构，包括存储结构、存取方法、输入输出方式等。当数据的物理组织或存储设备改变时，用户程序就必须重新编制。由于数据的组织是面向应用的，不同的计算程序之间不能共享数据，使得不同的应用之间存在大量重复数据。

1.2.2 文件系统阶段

20 世纪 50 年代中期至 60 年代中期，由于计算机大容量存储设备（如硬盘）的出现，推动了软件技术的发展，而操作系统的出现标志着数据管理步入一个新的阶段。在文件系统阶段，数据以文件为单位存储在外存，且由操作系统统一管理。操作系统为用户使用文件提供了友好界面。在这个阶段，程序与数据有了一定的独立性，程序与数据分开，有了程序文件与数据文件的区别。数据文件可以长期保存在外存上多次存取，进行诸如查询、修改、插入、删除等操作。文件的逻辑结构与物理结构脱钩，程序和数据分离，使数据与程序有了一定的独立性。用户的程序与数据可分别存放在外存储器上，各个应用程序可以共享一组数据，实现了以文件为单位的数据共享，如图 1.4 所示。

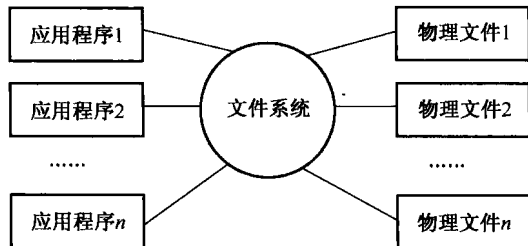


图 1.4 文件系统

但在文件系统阶段，由于数据的组织仍然是面向程序的，所以存在大量的数据冗余。而且数据的逻辑结构不能方便地修改和扩充，数据逻辑结构的微小改变都会影响到应用程序。由于文件之间互相独立，因而它们不能反映现实世界中事物之间的联系，操作系统不负责维护文件之间的联系信息。如果文件之间有内容上的联系，只能由应用程序去处理。

1.2.3 数据库系统阶段

20 世纪 60 年代后，随着计算机在数据管理领域的普遍应用，人们对数据管理技术提出了更高的要求。要求以数据为中心组织数据，减少数据冗余，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，以降低应用程序研制与维护的费用。数据库技术正是在这样一个应用需求的基础上发展起来的。

在文件系统阶段，人们在信息处理中关注的中心问题是系统功能的设计，因此程序设计占主导地位；而在数据库方式下，数据开始占据了中心位置，数据的结构设计成为信息系统的首要问题。数据库不仅包括数据本身，而且包括数据之间的联系。为了让多种应用程序并发地使用数据库中的共享数据，必须使数据与程序具有较高的独立性，这就需要有一个软件系统对数据实行专门管理，提供安全性和完整性等统一控制，方便用户以交互命令或程序方式对数据库进行操作。为数据库的建立、使用和维护而配置的软件就是数据库管理系统 DBMS，如图 1.5 所示。

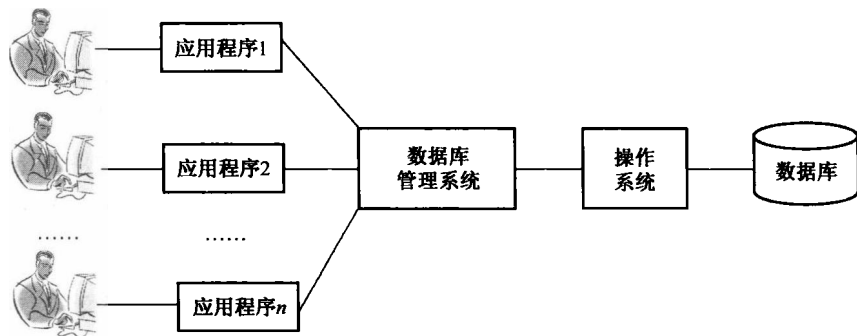


图 1.5 数据库系统阶段

1.2.4 分布式数据库系统阶段

随着数据库应用的不断发展，规模的不断扩大，集中式系统也显现了它的不足之处。如大型 DBS 的设计和操作都比较复杂，系统显得不灵活并且安全性也较差。因此，采用将数据分散的方法，把数据库分成多个，建立在多台计算机上，这种系统称为分布式系统。

分布式系统是物理上分散逻辑上集中的数据库系统，系统中的数据分布存放在计算机网络的各个场地中，每个场地都有自治处理（即独立处理）能力并能完成局部应用，而每个场地也参与（至少一种）全局应用，程序通过网络通信子系统执行全局应用，如图 1.6 所示。

分布式系统是用通信网络连接起来的结点（也称为场地）的集合，每个结点都是拥有集中式数据库的计算机系统。

分布式数据库系统主要有三个特点。

- ① 数据库的数据物理上分布在各个场地，但逻辑上是一个整体。
- ② 每个场地既可以执行局部应用（访问本地 DB），也可以执行全局应用（访问异地 DB）。
- ③ 各地的计算机由数据通信网络相联系。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 DB 和计算机的支持。

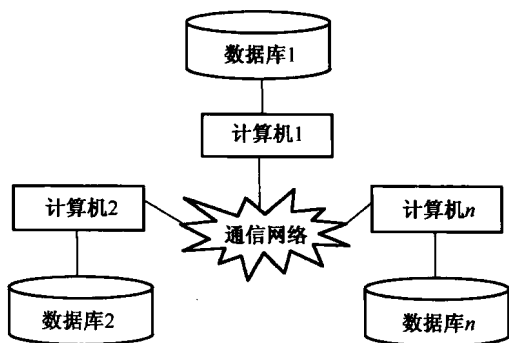


图 1.6 分布式数据库

1.2.5 数据管理新技术

从 20 世纪 60 年代末开始，数据库系统已从第一代的层次数据库、网状数据库，第二代的关系数据库系统，发展到第三代以面向对象模型为主要特征的数据库系统。随着用户应用

需求的提高、硬件技术和互联网技术的发展,促进了数据库技术与网络通信技术、人工智能技术、面向对象程序设计技术、并行计算技术等相互渗透,互相结合,形成了数据库新技术。

1. 面向对象数据库

关系型数据库系统虽然技术很成熟,但其局限性也显而易见:它能很好地处理所谓的“表格型数据”,却对越来越多的复杂类型的数据无能为力。20世纪90年代以后,人们把大量的精力花在研究“面向对象的数据库系统(Object Oriented Database)”或简称“OO数据库系统”上。

面向对象数据库系统主要有以下两个特点:

- ① 对象数据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、递归联系;
- ② 具有面向对象技术的封装性和继承性的特点,提高了软件的可重用性。

面向对象数据库的主要设计思想是用新型数据库系统来取代现有的数据库系统。这对许多已经运用传统数据库系统多年并积累了大量工作数据的客户来说,无法承受新旧数据转换带来的巨大工作量及巨额开支。另外,面向对象的关系型数据库系统所使用的查询语言极其复杂。所以,虽经数年发展,面向对象的关系型数据库系统的市场发展并不理想。

2. 多媒体数据库

多媒体是多种媒体形式的有机集成,如数字、字符、文本、图形、图像、声音、视频等。其中数字、字符等称为格式化数据,文本、图形、图像、声音、视频等称为非格式化数据,非格式化数据具有数据量大、处理复杂等特点。多媒体数据库系统(简称MDBS)是结合数据库技术和多媒体技术,能够有效实现对格式化和非格式化的多媒体数据进行存储、管理和操纵等功能的数据库系统。较于传统的数据库,多媒体数据库具有以下特征。

(1) 与传统数据库的差异大

多媒体数据库虽然在理论和技术上对传统数据库有很多继承,但其处理的数据对象、数据类型、数据结构、应用对象、处理方式都与传统数据库有较大差异,因此不能简单认为多媒体数据库只是对传统数据库的一种简单扩充或者试图用传统技术来处理多媒体数据。

(2) 处理对象的复杂性

多媒体数据库存储和处理的是现实世界中的复杂对象,不仅要处理包括数字、字符等格式化数据,还要处理图像、音频、视频等非格式化数据。

(3) 媒体间的独立性

多媒体数据库面临的数据有单一媒体数据和复合媒体数据(多种单一媒体数据的结合)。多媒体数据库从实用性的要求出发,强调多媒体数据库的用户可最大限度地忽略各媒体间差异,从而实现对多媒体数据的管理和操作。

3. 主动数据库

主动数据库(Active Database, ADB)是相对于传统数据库的被动性而言的。传统的数据库系统只能根据用户或应用程序的服务请求对数据库进行存储、检索等操作,而不能根据发生的事件或数据库的状态主动作出反应。主动数据库系统(ADBMS)具有主动提供服务功能,并且能以一种统一的机制实现各种主动服务的数据库系统。

一个主动数据库系统在某一事件发生时,引发数据库管理系统去检测数据库当前状态,若满足指定条件,则触发规定执行的动作,我们称之为ECA规则。一个主动数据库系统可表示为: $ADBS = DBS + EB + EM$ 。

其中,DBS代表传统数据库系统,用来存储、操作、维护和管理数据;EB代表ECA规

则库，用来存储 ECA 规则，每条规则指明在何种事件发生时，根据给定条件应主动执行什么动作；EM 代表事件监测器，一旦检测到某事件发生就主动触发系统，按照 EB 中指定的规则执行相应的动作。

一个主动数据库系统应该具有以下功能：

① 主动数据库系统应该提供传统数据库系统的所有功能，且不能因为增加了主动性功能而使数据库的性能受到明显影响。

② 主动数据库系统必须给用户和应用提供关于主动特性的说明，且该说明应该为数据库的永久性部分。

③ 主动数据库系统必须能有效地实现主动特性，且能与系统的其他部分有效地集成在一起，包括查询、事务处理、并发控制和权限管理等。

④ 主动数据库系统应能够提供与传统数据库系统类似的数据库设计和调试工具。

4. 数据仓库

数据仓库 (Data Warehouse, DW) 是指为了满足中高层管理人员预测和决策分析的需要，在传统数据库的基础上产生能够满足预测和决策分析需要的数据环境。数据仓库是面向主题的、整合的、稳定的，并且时变地收集数据以支持管理决策的一种数据结构形式。数据仓库的概念包含以下四方面的含义。

① 数据仓库是面向主题的：与传统数据库面向应用进行数据组织的特点相对应，数据仓库中的数据是面向主题进行组织的。主题是一个抽象的概念，对企业信息系统中的数据在较高层次上进行抽象综合归类并进行分析利用。在逻辑意义上，它是相应企业中某一宏观分析领域所涉及的分析对象。

② 数据仓库是集成的：数据仓库的数据主要用于分析，其数据的最大特点在于它不局限于某个具体的操作数据，而是对细节数据的归纳和整理。数据仓库中的综合数据不能从原有数据库系统中直接得到，而需从其中抽取。因此，数据在进入数据仓库之前，必须进行加工与集成，处理原始数据中的所有歧义和矛盾，如单位不统一、字段的同名异义、异名同义等，将原始数据结构进行从面向应用到面向主题的转变。这是数据仓库建设中最关键、最复杂的一步。

③ 数据仓库是稳定的：数据仓库数据反映的是一段相当长的时间内历史数据的内容，是不同时间内数据快照（来自数据库的一个表或表的子集的最新拷贝）的集合，以及基于这些快照进行统计、综合和重组的导出数据，而不是联机处理的数据，不进行实时更新。

④ 数据仓库随时间变化：数据仓库的数据稳定性是针对应用来说的，即用户进行分析处理时不能进行数据更新操作。但并不是说，在数据从集成输入到数据仓库中开始到最终被删除的整个数据生存周期之中所有数据都永久不变。

当前的数据仓库系统中，直接面向用户的部分前端工具主要有两类：联机分析处理 (OLAP) 的分析查询型工具和数据挖掘 (DM) 的挖掘型工具。

(1) 联机分析处理 (OLAP)

OLAP 的显著特征是能提供数据的多维概念视图，使最终用户从多角度、多侧面、多层次地考察数据库中的数据，从而深入地理解包含在数据库中的信息和内涵，多维数据分析是决策的主要内容。目前，OLAP 工具可分为两大类：基于多维数据库的 MOLAP 和基于关系数据库的 ROLAP。MOLAP 利用专有的多维数据库存储 OLAP 分析所需的数据，数据以多维方式存储并以多维视图方式显示。ROLAP 则利用关系表模拟多维数据，将分析的结果经多维处理转化为多维视图展示给用户。