



普通高等教育“十一五”国家级规划教材

数据库技术 及应用

陆桂明 主编



机械工业出版社
CHINA MACHINE PRESS



普通高等教育“十一五”国家级规划教材

数据库技术及应用

主 编 陆桂明

副主编 王 峰

参 编 杨 彬 海 燕 张 帆

主 审 甘 勇 刘建华



机械工业出版社

本书为普通高等教育“十一五”国家级规划教材。全书共分十一章，前五章介绍数据库系统的基础知识和数据模型；第6章介绍数据库系统的分析与设计；第7章介绍客户/服务器数据库；第8~10章介绍SQL Server基本知识、数据库系统开发工具和开发实例；第11章介绍目前数据库发展的新技术。

本书可作为高等学校计算机专业本、专科学生学习数据库课程的教材，也可供其他专业师生和软件开发技术人员参考。

本书配有电子教案，欢迎选用本书的老师索取，索取邮箱：wxld2677@163.com。

图书在版编目（CIP）数据

数据库技术及应用/陆桂明主编. —北京：机械工业出版社，2008.9
普通高等教育“十一五”国家级规划教材
ISBN 978-7-111-24749-4

I. 数… II. 陆… 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（2008）第 118400 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

责任编辑：王小东 版式设计：霍永明 责任校对：刘志文

封面设计：姚毅 责任印制：邓博

北京四季青印刷厂印刷（三河市杨庄镇环伟装订厂装订）

2008 年 9 月第 1 版第 1 次印刷

184mm×260mm·19.5 印张·458 千字

标准书号：ISBN 978-7-111-24749-4

定价：33.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
销售服务热线电话：（010）68326294

购书热线电话：（010）88379639 88379641 88379643

编辑热线电话：（010）88379728

封面无防伪标均为盗版

前言

计算机处理信息使用最多的手段就是数据库技术，而且数据库技术的应用已经渗透到各个行业和各个领域，学习和掌握数据库的基本原理、基本的开发技术，以及利用数据库系统进行数据处理成为计算机专业学生必须具备的能力之一。

本书作者从事多年“数据库技术及应用”教学和相关软件开发工作，对教学和实践有较丰富的经验，本书的初稿作为讲义已在教学中使用多年，取得了较好的教学效果。现在正式出版，希望对相关专业的学生和工程技术人员有所帮助。

《数据库技术及应用》一书共分十一章。第1章介绍了数据库系统发展和现状；第2章介绍了关系数据库的概念；第3章介绍了结构化查询语言SQL；第4章介绍了关系数据库中的模式设计问题；第5章介绍了数据库的安全与保护及技术，第6章介绍了数据库系统的分析与设计；第7章介绍了客户机/服务器模式和分布式数据库系统；第8章介绍了SQL Server 2005的特性和常用工具；第9章介绍了数据库应用开发工具的应用；第10章通过一个具体的实例，演示了数据库管理系统的建立步骤；第11章介绍了当前数据库技术发展的新技术。

本书由陆桂明、王峰、杨彬、海燕、张帆编写，陆桂明编写第1章，张帆编写第2、8、11章，杨彬编写第3、4章，海燕编写第5、6章，王峰编写第7、9、10章，全书由陆桂明负责统稿，甘勇、刘建华为本书主审。

由于编者水平有限，疏漏和错误之处在所难免，恳请各位专家、老师和广大读者不吝指正。

编者

目 录

前言	3.1.1 SQL 的产生与发展	24
第 1 章 数据库系统概述	3.1.2 SQL 的基本概念	25
1.1 信息与数据	3.1.3 SQL 的主要特点	26
1.1.1 信息	3.2 SQL 数据定义	27
1.1.2 数据	3.2.1 基本数据类型	28
1.1.3 信息与数据的关系	3.2.2 创建、修改和删除基本表	28
1.1.4 信息管理系统	3.2.3 创建和维护索引	31
1.2 数据库的产生与发展	3.3 SQL 数据查询	32
1.2.1 人工管理阶段	3.3.1 SELECT 查询命令	32
1.2.2 文件系统管理阶段	3.3.2 简单查询	34
1.2.3 数据库管理阶段	3.3.3 连接查询	48
1.3 数据库系统的组成	3.3.4 嵌套查询	51
1.4 数据库系统的体系结构	3.3.5 集合查询	57
1.4.1 单用户数据库系统	3.4 SQL 数据更新	58
1.4.2 主从式结构数据库系统	3.4.1 插入数据	58
1.4.3 客户机/服务器结构的数据库系统	3.4.2 修改数据	60
1.4.4 分布式结构的数据库系统	3.4.3 删除数据	60
1.5 数据库系统的安全	3.5 视图	61
1.5.1 软件方面的安全设置	3.5.1 视图的定义和删除	62
1.5.2 硬件方面的安全设置	3.5.2 查询视图	64
小结	3.5.3 更新视图	65
第 2 章 关系数据库	3.5.4 视图的优点	67
2.1 关系数据结构及概念	3.6 SQL 数据控制	68
2.1.1 关系数据结构	3.6.1 权限与角色	68
2.1.2 基本概念	3.6.2 权限与角色的授予与收回	68
2.2 关系的完整性约束	3.7 嵌入式 SQL	71
2.3 关系模型的数据操作	3.7.1 嵌入式 SQL 的一般形式	71
2.4 关系代数	3.7.2 嵌入式 SQL 语句与主语言之间的通信	72
小结	3.7.3 不用游标的 SQL 语句	74
习题	3.7.4 使用游标的 SQL 语句	77
第 3 章 结构化查询语言——SQL	3.7.5 动态 SQL 简介	80
3.1 SQL 概述	小结	81

习题	81	5.4.2 数据库的并发控制	121
第4章 关系数据理论	83	5.4.3 并发的目的	121
4.1 规范化问题的提出	83	5.4.4 并发所引起的问题	121
4.2 函数依赖	85	5.4.5 并发控制方法	123
4.2.1 函数依赖的定义	86	5.4.6 并发调节的可串行性	129
4.2.2 函数依赖的基本性质	87	5.4.7 两段锁协议	131
4.2.3 完全函数依赖与部分函数 依赖	87	5.5 数据库的恢复	132
4.2.4 传递函数依赖	88	5.5.1 故障种类	133
4.2.5 码	88	5.5.2 恢复基本策略	134
4.3 范式	89	5.5.3 数据库备份	136
4.3.1 第一范式 (1NF)	90	小结	137
4.3.2 第二范式 (2NF)	91	习题	138
4.3.3 第三范式 (3NF)	93	第6章 数据库系统的分析与设计	139
4.3.4 BC 范式 (BCNF)	94	6.1 数据库设计的概述	139
4.3.5 多值依赖	97	6.1.1 数据库设计的任务和内容	139
4.3.6 第四范式 (4NF)	100	6.1.2 数据库设计的方法	140
4.3.7 规范化小结	101	6.1.3 数据库设计的步骤	141
4.4 模式分解	103	6.1.4 数据库设计的工具	143
4.4.1 模式分解的准则	103	6.2 需求分析	144
4.4.2 分解的函数依赖保持性和 无损连接性	105	6.2.1 需求分析的任务	144
4.4.3 模式分解的算法	109	6.2.2 需求分析的方法	145
小结	110	6.3 数据库概念结构设计	149
习题	110	6.3.1 概念结构设计的方法	150
第5章 数据库保护技术	112	6.3.2 概念结构设计的步骤	153
5.1 安全与保护概述	112	6.4 数据库逻辑设计	160
5.2 数据库的完整性	112	6.4.1 逻辑结构设计的任务和步骤	160
5.2.1 数据库的完整性分类	113	6.4.2 E-R 模型向关系模型的转换	161
5.2.2 数据库的完整性约束	113	6.4.3 数据模型的优化	167
5.3 数据库的安全性	115	6.4.4 外模式设计	168
5.3.1 用户的标识与鉴别	116	6.5 数据库物理设计	169
5.3.2 存取权限控制	116	6.5.1 确定数据库的物理结构	169
5.3.3 视图机制	117	6.5.2 评价物理结构	173
5.3.4 跟踪审查	118	6.6 数据库的实施和维护	173
5.3.5 数据加密存储	118	6.6.1 数据库的实施	174
5.4 数据库的并发控制	119	6.6.2 数据库的维护	175
5.4.1 事务及特性	119	小结	176
5.4.2 数据库的并发控制	121	习题	176
5.4.3 并发的目的	121	第7章 客户机/服务器数据库	178
5.4.4 并发所引起的问题	121		
5.4.5 并发控制方法	123		
5.4.6 并发调节的可串行性	129		
5.4.7 两段锁协议	131		
5.5 数据库的恢复	132		
5.5.1 故障种类	133		
5.5.2 恢复基本策略	134		
5.5.3 数据库备份	136		
小结	137		
习题	138		

7.1 客户机/服务器简介	178	8.5.1 创建表	213
7.1.1 客户机/服务器计算模式		8.5.2 添加、删除、修改字段	213
7.1.1.1 概念	178	8.5.3 创建、删除、修改约束	217
7.1.2 客户机/服务器环境下应用		8.5.4 删除表	218
7.1.2.1 成分的分布	179	8.6 数据更新	218
7.2 客户机/服务器模式的体系结构	180	8.6.1 用 Insert 增添数据	218
7.2.1 体系结构	180	8.6.2 用 Update 更新数据	221
7.2.2 客户机/服务器模式的定位	181	8.6.3 用 DELETE 删除数据	222
7.3 客户机/服务器模式的特性和优点	186	8.7 数据查询	223
7.3.1 客户机/服务器模式的特性	186	8.8 视图的创建与使用	224
7.3.2 客户机/服务器模式的优点		8.8.1 创建视图	224
7.3.2.1 与不足	186	8.8.2 查看视图	228
7.4 构建客户机/服务器系统	187	8.8.3 修改视图	229
7.4.1 客户机/服务器系统开发		8.8.4 重命名视图	229
7.4.1.1 工具	187	8.8.5 使用视图	229
7.4.2 客户机/服务器模式应用系统的		8.8.6 删除视图	231
7.4.2.1 开发特点	190	8.9 索引的创建与使用	232
小结	191	8.9.1 创建索引	232
习题	191	8.9.2 查看、修改、维护和删除	
第8章 SQL Server 数据库管理		8.9.2.1 索引	234
8.1 SQL Server 2005 概述	192	8.10 触发器的创建与使用	236
8.2 Transact-SQL 程序设计	193	8.10.1 概述	236
8.2.1 数据类型	193	8.10.2 创建触发器	237
8.2.2 变量	196	8.10.3 查看触发器	238
8.2.3 运算符	196	8.10.4 修改触发器	239
8.2.4 函数	197	8.10.5 删除触发器	239
8.2.5 流程控制语句	200	8.11 SQL Server 安全性管理	239
8.3 存储过程的创建与使用	202	8.11.1 SQL Server 2005 的安全	
8.3.1 创建存储过程	202	8.11.1.1 机制	240
8.3.2 执行存储过程	204	8.11.2 服务器安全性管理	241
8.3.3 查看、修改存储过程	205	8.11.3 数据库用户管理	245
8.3.4 重命名、删除存储过程	207	8.11.4 数据库角色管理	246
8.4 数据库的创建与管理	207	8.11.5 权限管理	247
8.4.1 数据库的存储结构	207	8.12 在 VB、Delphi 和 ASP 中使用	
8.4.2 创建与删除数据库	208	8.12.1 SQL 数据库	248
8.4.3 修改数据库	210	8.12.1.1 在 VB 中使用 SQL 数据库	248
8.5 表的创建与管理	213	8.12.1.2 在 Delphi 中使用 SQL	
		8.12.1.2.1 数据库	250

8.12.3 在 ASP 中使用 SQL 数据库·····	251	10.3.4 奖惩管理子系统数据流图与 数据字典·····	277
小结·····	252	10.3.5 毕业管理子系统数据流图与 数据字典·····	278
习题·····	252	10.3.6 注册管理子系统数据流图与 数据字典·····	279
第 9 章 数据库应用开发工具 ·····	253	10.4 数据库分析·····	279
9.1 数据库开发工具的分类·····	253	10.4.1 学籍管理系统 E-R 图·····	279
9.1.1 桌面型数据库·····	253	10.4.2 利用 Microsoft Visio 反向获取 E-R 图·····	279
9.1.2 网络型数据库·····	253	10.5 数据库设计·····	283
9.2 桌面型数据库开发工具·····	254	10.5.1 创建数据库·····	283
9.2.1 xBase·····	254	10.5.2 创建各数据表·····	285
9.2.2 Visual FoxPro·····	256	小结·····	293
9.2.3 Access·····	256	习题·····	294
9.2.4 Paradox·····	258	第 11 章 数据库新技术 ·····	295
9.3 网络型数据库开发工具·····	259	11.1 与多学科技术相结合形成的数据库 新技术·····	295
9.3.1 Oracle·····	259	11.1.1 面向对象数据库系统·····	295
9.3.2 Sybase·····	262	11.1.2 分布式数据库系统·····	296
9.3.3 Informix·····	263	11.1.3 多媒体数据库系统·····	297
9.3.4 SQL Server·····	264	11.1.4 知识数据库系统·····	297
9.3.5 IBM DB2·····	265	11.1.5 并行数据库系统·····	298
9.3.6 前台开发工具·····	266	11.1.6 模糊数据库系统·····	298
9.4 数据库建模工具·····	268	11.2 与特定的应用领域相结合形成的 数据库新技术·····	299
小结·····	270	11.2.1 数据仓库·····	299
习题·····	270	11.2.2 工程数据库·····	301
第 10 章 数据库管理系统开发 实例 ·····	271	11.2.3 统计数据库·····	301
10.1 引言·····	271	11.2.4 空间数据库·····	302
10.2 任务概述·····	271	11.3 数据库建设中应注意的几个问题·····	302
10.2.1 目标·····	271	小结·····	303
10.2.2 具体要求·····	271	习题·····	303
10.3 数据描述·····	272	参考文献 ·····	304
10.3.1 顶层数据流图与数据字典·····	273		
10.3.2 第二层数据流图与数据 字典·····	274		
10.3.3 第三层数据流图与数据 字典·····	275		

第1章 数据库系统概述

随着人类向信息化社会迈进,信息成为人们日常生活、学习和工作等各个方面都不能缺少的资源。在我们日常生活和工作的各个方面都存在着大量信息,利用计算机有效地组织和管理信息,可以大大提高信息的获取和使用效率。例如图书馆所提供的图书信息中,对图书编号、图书类别、书名、作者、关键词等信息进行分类,为读者提供了快速查阅的手段,如果这些信息按一定方式存入计算机,当输入要查询的信息后,计算机系统 will 自动索引出满足查询条件的图书,同时显示图书的内容简介、借阅及库存情况等相关信息。

应该说计算机是目前信息管理最有效的工具,因此研究信息在计算机中如何进行存储和管理,以什么样的数据结构来组织数据,对信息管理至关重要,我们还以图书馆的图书管理系统为例,假如数以万计的图书数据无序地堆积在一起,无序地输入到计算机中,读者可能需要花费很长的时间在计算机中按翻阅和遍历的方式查找一本图书的各种信息,显然,这种方式的效率不是我们想要的。因此数据如何管理正是需要研究的问题,数据库(DataBase, DB)技术就是一门研究数据管理和应用的技术。

数据库技术就是利用计算机技术,按某种有序的信息组织方式把数据存储在计算机中,同时为用户提供数据的输入、修改、查询、显示、打印等功能,因此数据库技术正是建立信息管理系统,实现信息管理的重要手段。数据库技术产生于20世纪60年代中期,随着计算机技术的发展以及信息量不断增加,计算机在数据处理方面的工作越来越多,这项技术得到了各国计算机厂商的重视和发展,而且极大地促进了数据库应用向各行各业的渗透。

1.1 信息与数据

数据库技术是人们为了实现信息管理而产生的一项技术。为了便于理解数据库系统的相关知识,先来了解一下人们对信息、数据和信息管理系统的认识。

1.1.1 信息

通常认为,信息是对事物的状态、运动方式和特征的描述,反映的是客观系统中某一事物的属性或表现形式。它是人们进行各种活动所需要的知识。

这里的“事物”泛指一切可能的研究对象,包括客观世界和主观世界。既包含看得见摸得着的具体实体,也包含那些看不见摸不着的抽象概念,如:人、图片、声音、计算机、磁场、质量、长度等。信息是现实世界中的事物,通过人的感官感知并经过人脑的抽象后形成的概念。因此一切事物都含有信息。信息的表现需要载体,信息的载体有很多种,包括图书、报纸、广播、电视、网络等,同一信息可以有多种的载体。现在由计算机或计算机网络向人们提供的信息与从前通过报纸、电视等方式获得的信息并无本质的

区别。

为了在计算机中存储和处理信息,必须提取事物的属性和特征,例如要了解一组学生对象时,可以从“学号、姓名、性别、民族、生日、身高”等属性来加以描述,具体形式为:

200602104 马 丽 女 回 1988-04-17 1.58

200602105 李立峰 男 汉 1988-06-05 1.73

200602106 刘蓓蓓 女 汉 1988-03-23 1.61

200602107 刘延法 男 汉 1988-11-02 1.76

这种表示信息的符号就是数据。

1.1.2 数据

数据是指用符号记录下来的可以鉴别的信息,用于记录事物的情况,数据的内容通常是事物特征的反映或描述。所采用的形式可以是数字、文字、图形、图像或其他特殊符号。因此数据有数据类型和数据值之分。不同的数据类型记录事物的性质是不一样的。例如数值型数据可以用来表示身高、价格等;字符型数据可表示姓名、家庭地址等;还有特殊类型的数据,如声音、图像等。

1.1.3 信息与数据的关系

数据和信息两个概念,既有区别,又有联系。数据是信息的载体,是承载信息的符号,而信息是数据有意义的表现。数据不能独立出现,必须伴随其属性,例如,“体重:65kg”,如果独立出现65,则无法知道其含义。

但在许多场合,数据和信息又很难区分,因为信息本身就是数据化的,是用数据来描述和记载的,而数据本身包含了各种信息。所以有时我们也将数据和信息作为同一个概念。在一些不需要严格分辨的场合,可以把信息处理说成数据处理,把信息管理系统说成是数据管理系统。

1.1.4 信息管理系统

信息管理系统就是管理数据的系统。信息的产生随处可见,而且往往需要把信息记录下来,作为资料保存。如图书馆的图书目录,这必然导致信息的整理、分类、保存、索引、检索和处理等一系列管理问题。

随着现代科学技术的发展,信息的开发与利用日益受到广泛重视。目前信息的特点是数量大、种类多、来源广。特别是计算机网络的逐步普及,使得信息资源数量急剧增加。而无序的信息降低了这些信息资源的使用效率,因此,必须采取有效的方法对信息加以组织、处理和控制。

信息管理系统的主要功能就是收集、组织、存储、查询、统计和更新信息,以及安全、可靠地控制和共享信息等。其实,信息管理系统在计算机技术出现之前就有,如图书馆的档案管理,只是计算机技术运用于信息管理中可以大幅度地提高效率,现在,人们提到信息管理系统时,通常就是指计算机信息管理系统。

1.2 数据库的产生与发展

数据库技术是随着计算机应用由科学计算发展到数据处理而产生的一种技术,它的发展是与计算机对数据处理的速度、规模、要求以及计算机软硬件技术发展是密不可分的。计算机实现数据管理经历了三个发展阶段,即人工管理阶段(20世纪50年代中期以前)、文件系统阶段(20世纪50年代后期至60年代中期)和数据库系统阶段(20世纪60年代中后期开始)。

1.2.1 人工管理阶段

早期的计算机大多采用穿孔纸带作为外存存储设备,进行数据处理时,要将原始数据和程序(记录在穿孔纸带中)一起输入主存,运算处理后将结果直接打印输出,数据处理的方式基本上是批处理。

这个时期数据管理的特点是:

- 1) 数据不保存。由于外设没有可以擦写的介质,过程和结果无法保存。
- 2) 数据的独立性差。这个时期的数据包含在程序之中,数据与程序不具有独立性,程序和数据相互依赖。程序员设计程序时,不仅要设计处理数据的步骤,还要设计数据的组织方式。而且各应用程序处理的数据之间毫无联系,处理过程人工干预比较多。
- 3) 只有程序(Program)的概念,没有文件(File)的概念。
- 4) 数据无共享。数据不独立,因而无法共享。

1.2.2 文件系统管理阶段

随着计算机软硬件的发展,外存已有磁盘、磁鼓等可以擦写的存储设备。软件方面有了高级语言和操作系统。操作系统中的文件管理系统提供了管理数据文件的功能。文件管理系统的方式就是把相关的数据组织成数据文件,程序和数据都以文件的方式存储在磁盘上。在程序运行中可以打开数据文件进行存取,数据文件的具体存储位置由操作系统来管理。

这一阶段数据管理的特点是:

- 1) 数据以文件形式长期保存在磁盘上,用户可以通过程序随时对数据文件进行查询、修改、添加和删除等处理。并且处理的中间或最后结果都可以以文件的形式保存在磁盘上。
- 2) 程序员只需通过文件名存取数据,不必过多关心数据的物理位置,而把精力放在程序设计上即可。
- 3) 文件的形式已多样化。有索引文件、顺序文件等,对文件中的数据既可顺序访问也可随机访问,便于存储和查询。
- 4) 数据与程序具有一定的独立性。有了存储文件以后,数据不再属于某个特定的程序,一定程度上可以共享。

与前一阶段相比,文件系统在数据管理方面有了很大进步,但根本性问题仍然没有得

到解决,随着数据处理规模的急剧增大,逐渐显露出很多缺陷,主要表现在以下三个方面:

- 1) 数据一致性差:各程序之间不能共享相同的数据,同样的数据在不同的文件中有可能不一样。
- 2) 数据冗余度大:由于文件之间缺乏联系,造成每个应用程序都有对应的数据文件,可能出现同样的数据在多个文件中重复存储。
- 3) 数据联系弱:数据文件相互独立,缺乏联系。

由于这些原因,促使人们研究一种新的数据管理技术,克服以上的缺陷,这就是 20 世纪 60 年中后期产生的数据库技术。

1.2.3 数据库管理阶段

从 20 世纪 60 年代中后期开始,信息处理的规模越来越大,对数据管理的技术提出了更高的要求,文件系统已不能胜任数据共享和处理。

这个时期磁盘技术取得了重大进展,大容量和快速存取的磁盘陆续进入市场,为数据库技术的实现提供了硬件基础。各种数据库系统也相继问世,其主要标志是 20 世纪 60 年代后期的三件大事:

- 1) 1968 年美国 IBM 公司推出的层次模型的 IMS 数据库系统。
- 2) 1969 年美国数据系统语言协会 (CODASYL) 的数据库任务组 (DBTG) 发表关于网状模型的 DBTG 报告,它们为统一管理与共享数据提供了有力的支持。

这两种数据库系统由于都是由文件系统发展而来的,数据结构比较繁琐,程序受数据库文件中的结构影响较大,用户在使用数据库时需要对数据的结构有详细的了解,这对数据库的使用造成很多困难。同时,数据结构过于繁琐,影响了复杂数据结构的实现。

- 3) 1970 年美国 IBM 公司 E. F. Codd 博士连续发表一系列论文,奠定了关系数据库的理论基础。

关系数据库在 20 世纪 70 年代中期至 20 世纪 80 年代得到了充分发展,它具有简单的结构与较少的物理表示,使用和操作符合人们日常的处理方式,因此在 20 世纪 80 年代初逐步取代层次与网状数据库系统成为数据库系统的主导,如 20 世纪 80 年代初出现的一批商品化的关系数据库系统,如 Oracle、SQL/DS、DS、DB2、IMGRES、INFORMIX、UNIFY 和 DBASE 等。

与文件系统相比,数据库系统克服了文件系统的缺陷,提供了对数据更高级更有效的管理。概括起来,数据库技术的管理方式具有以下特点:

1. 数据共享性好、冗余减少

数据库的数据面向整个应用系统。系统要考虑所有用户对数据的要求,数据库中提供了所有用户的数据成分,不同用户可以使用库中全部数据或部分数据。这样数据不再面向特定的某个或多个应用,而是面向整个应用系统。数据冗余明显减少,实现了数据共享。

2. 数据的结构化

数据库是存放在磁盘上的数据集合,是按一定的数据结构组织起来的。数据库系统中

的数据文件既相对独立又是相互联系的,从总体上遵循一定的结构形式。数据库正是通过数据文件之间的联系反映现实事物之间的某种联系的。

3. 减少了数据的不一致性

由于数据库包含了整个系统所有用户的所有数据,用户操作的数据是通过数据库管理系统从数据库中映射出来的某个子集,不是独立的文件,实际上所有用户使用的是物理存储的一个文件,这就避免了数据的不一致性。

4. 有较高的数据独立性

数据独立性包括物理独立性和逻辑独立性。物理独立性是指数据库的物理结构(即数据的组织、存储、存取方式和外部存储设备等)发生变化时,不会影响到数据库的整体逻辑结构和用户的逻辑结构,由于应用程序是根据用户的逻辑结构编写的,所以应用程序不必改动,这样数据库就达到了物理独立性。

逻辑独立性是指数据库的整体逻辑结构改变时,由数据库管理系统改变整体逻辑结构与用户逻辑结构之间的映像,使用户逻辑结构不变,从而应用程序不变,这样就实现了数据库的逻辑独立性。

5. 统一的数据控制功能

为确保数据库的正确性和有效性以及数据库系统的正常运行,数据库管理系统提供了四个方面的数据控制功能:

(1) 数据完整性 是指保证数据库始终存储正确的数据。用户可设计一些完整性规则以确保数据值的正确性。例如数值型数据只可含有0~9的数字,并对数据值之间的联系进行各种检验。

(2) 数据安全性 系统提供口令或其他手段来验证用户的合法性,保证数据的安全和机密,防止数据丢失或被窃取。

(3) 数据库的并发控制 避免多用户或并发程序之间对数据的干扰,防止数据库中的数据被破坏。

(4) 数据的恢复 在数据库被破坏时或数据不可靠时,系统有能力把数据库恢复到最近某个时刻的正确状态。

在数据库系统阶段,程序和数据的联系可用图1-1表示。

从文件系统发展到数据库技术是信息处理领域的一个重大变化。在文件系统阶段,程序设计处于主导地位,数据服从程序设计的需要;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最关心的问题,而应用程序设计则围绕既定的数据结构和人机界面的处理上做工作。

数据库技术从产生发展到今天,技术的更新发展很快。目前,被广泛应用的关系数据库系统已经是数据库技术的第二代产品,而20世纪80年代以后出现的基于对象模型的数

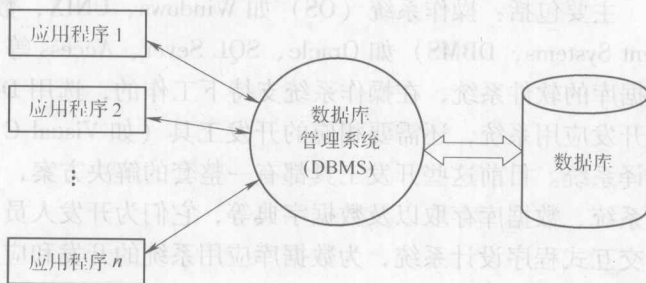


图 1-1 程序和数据的联系图

数据库系统又在向关系数据库系统发起挑战。现在,数据库技术与网络通信技术、人工智能技术、面向对象技术、并行计算技术及分布式技术等相互渗透、相互结合,成为数据库技术发展的主要特征。

1.3 数据库系统的组成

数据库系统(DataBase System, DBS)是指包含数据库应用的计算机系统,表明不仅是一组对数据进行管理的软件(即通常所说的数据库管理系统),也不仅仅是一个数据库,而是一个可运行的,按照数据库方式组织、存储、维护和向应用系统提供数据支持的系统。它是存储介质、处理对象和管理系统的集合体,一般由数据库、硬件、软件、数据库管理员和用户五部分构成。

1. 数据库

数据库(DataBase, DB)是指与各项应用相关的全部数据的集合。主要由两大部分组成:一部分是应用所需要的数据集合,称为物理数据库;另一部分是关于各级数据结构的描述,称为描述数据库。

数据库是存储在计算机系统的数据集合,具有较小的冗余度和较高的数据独立性,具备完善的自我保护能力和数据恢复能力,能把各种相关数据集中管理防止意外破坏,同时对用户程序提供共享。

2. 硬件支持系统

硬件是数据库赖以保存的物理设备,包括CPU、内存、外存和数据通信等各种存储、处理和传输数据的设备。为了尽可能提高整个系统对用户的响应速度,数据库系统对硬件的要求是:需要足够大的内存和cache,用来存放操作系统、应用程序以及数据库,提供系统和用户工作缓冲区;外部存储要配备高速的、大容量的存储设备,如果需要存储异地数据,还需要考虑系统支持联网能力和配备必要的后备存储设备等因素,而网络还需要有较高的带宽和数据传输速度。

3. 软件支持系统

主要包括:操作系统(OS)如Windows、UNIX,数据库管理系统(DataBase Management Systems, DBMS)如Oracle、SQL Server、Access等,各种应用程序等。DBMS是管理数据库的软件系统,在操作系统支持下工作的,选用DBMS时还要考虑操作系统的支持;为开发应用系统,还需要相应的开发工具(如Visual C++、Delphi、ASP、Java等)及其编译系统。目前这些开发工具都有一整套的解决方案,如人机界面生成、报表生成器、表格系统、数据库存取以及数据字典等,它们为开发人员和最终用户提供的高效率、多功能的交互式程序设计系统,为数据库应用系统的开发和应用提供了良好的环境,极大地提高了程序开发效率。

4. 数据库管理员

对于较大规模的数据库系统,需要有专人来负责建立、维护和管理数据库系统,这样的人员称为数据库管理员(DataBase Administrator, DBA)。数据库管理员的任务主要包括:定义并存储数据库的内容,监督和控制数据库的使用,负责数据

库的日常维护以及必要时重新组织和改进数据库, 如果需要的话, 除了数据库管理员外, 还有系统分析员、程序开发员和测试员等。这些工作可以由一个人或几个人来完成。

5. 用户

数据库系统的使用者就是系统的最终用户, 这些用户通过开发人员设计好的界面来使用数据库, 包括对数据库的添加、修改、删除、统计及生成报表等操作。这些操作由于都是通过数据库管理人员和程序开发人员编写的程序和规则来进行的, 所以可以避免用户由于误操作而对数据库的破坏, 使系统安全、可靠和稳定。

1.4 数据库系统的体系结构

从数据库管理系统结构来看, 数据库系统是一个包含数据库 (DB)、数据库管理系统 (DBMS)、应用程序接口 (API) 三级模式结构。数据库的这种模式结构可以使整个系统安全、稳定和可靠, 虽然数据库管理员 (DBA) 既可以通过应用程序接口操作数据库, 也可以绕过数据库管理系统直接操作数据库中的数据。DBA 的这种权限是一把双刃剑, 它一方面可以直接和灵活地访问数据库中的数据, 而另一方面却可能会破坏数据的完整性和稳定性。但对最终用户来说, 他们见到的仅是数据库的应用程序, 通过预先编好的程序访问数据库, 因而用户对整个系统的安全不构成威胁。

从数据库系统的三级模式结构角度来看, 数据库系统可以分为单用户结构、主从式结构、客户机/服务器结构和分布式结构。

1.4.1 单用户数据库系统

单用户数据库系统是一种早期的最简单的数据库系统。在单用户系统中, 整个数据库系统 (包括应用程序、数据库管理系统和数据库) 都装在一台计算机上, 由一个用户独占使用, 如图 1-2 所示。不同计算机之间的数据库无法同时使用, 数据的共享只能在脱机状况下用磁盘等媒介相互复制。20 世纪 80 年代后期相继推出的 dBase III、FoxPro 等就属于这种数据库系统。

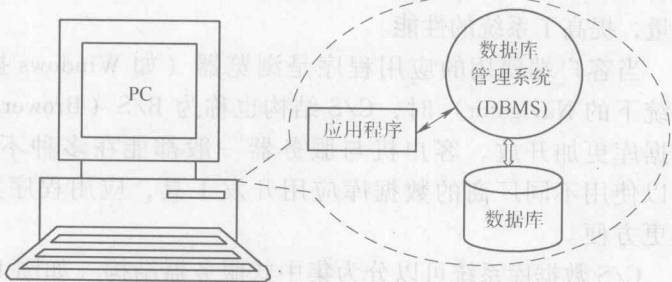


图 1-2 单用户数据库系统

在单用户系统中, 一个单位中各个部门的计算机是相互独立的, 各个部门用各自的计算机管理本部门的数据, 因此单位内部存在大量的冗余数据。例如, 学校中的学生处、教务处、各院系等部门必须重复存放学生的一些基本信息 (如学号、姓名、年龄、班级等)。在单用户系统中, 各个信息集无法实时联系, 形成信息孤岛。

1.4.2 主从式结构数据库系统

主从式结构是指一个主机（大多是运算速度和性能较高的小型机以上）带多个终端的多用户结构，如图 1-3 所示。在这种结构中，数据库系统（包括应用程序、数据库管理系统和数据库）都集中存放在主机上，所有处理任务都由主机来完成，各个用户通过终端分时地（由于计算机性能较高，也可以看作是同时）操作数据库，共享数据资源。

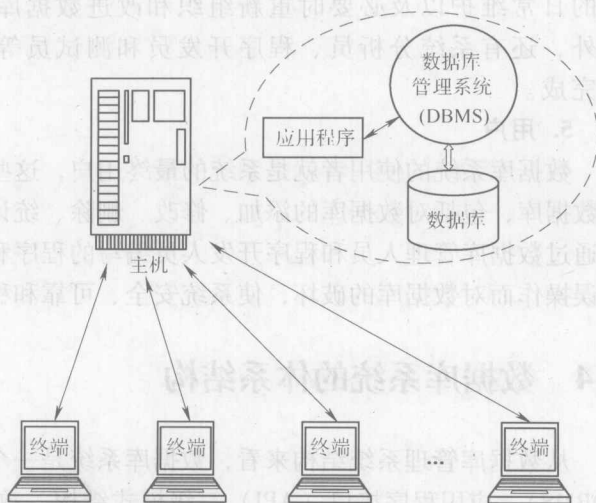


图 1-3 主从式结构数据库系统

主从式结构的优点是数据集中，易于管理与维护。缺点是当主机出现故障时，整个系统都不能使用，因此系统的可靠性不高。

1.4.3 客户机/服务器结构的数据库系统

随着计算机网络的发展和微型计算机性能的提高，人们开始把 DBMS 功能和应用分开。网络中某个节点上的计算机（性能较高）专门用于执行 DBMS 功能，称为数据库服务器，简称服务器，其他节点上的计算机（一般是个人计算机 PC）安装 DBMS 的外围应用程序，支持用户的应用，称为客户机，这就是客户机（Client）/服务器（Server）结构的数据库系统，简称 C/S 结构。

在 C/S 结构中，客户端的用户请求通过应用程序传送到数据库服务器，数据库服务器进行处理后，将结果返回给客户端（而不是所有数据），从而减少了数据在网络上的传输量，提高了系统的性能。

当客户端使用的应用程序是浏览器（如 Windows 操作系统中的 IE 和 Linux 操作系统下的 Navigator）时，C/S 结构也称为 B/S（Browser/Server）结构，而这种结构的数据库更加开放。客户机与服务器一般都能在多种不同的硬件和软件平台上运行，可以使用不同厂商的数据库应用开发工具，应用程序具有更高的复用性，发布和安装更方便。

C/S 数据库系统可以分为集中式服务器结构（如图 1-4 所示）和分布式服务器结构（如图 1-5 所示）。前者在网络中仅有一台数据库服务器，而客户机是多台。后者在网络中有多台数据库服务器。数据库服务器既可集中组成集群，也可以分布在网络各节点。分布式服务器结构是 C/S 与分布式数据库的结合。与主从式结构相似，在集中式服务器结构中，一台数据库服务器要为众多的客户服务，往往容易成为瓶颈，制约系统的性能。在分布式服务器结构中，瓶颈问题得到了解决，但由于数据分布在不同的服务器上，给数据的处理、管理与维护带来一定的难度。B/S 数据库系统（如图 1-6 所示）在客户端与数

据库服务器之间设置 WWW 服务器，数据库服务器既可以与 WWW 服务器安装在同一台服务器中，也可以单独设置。

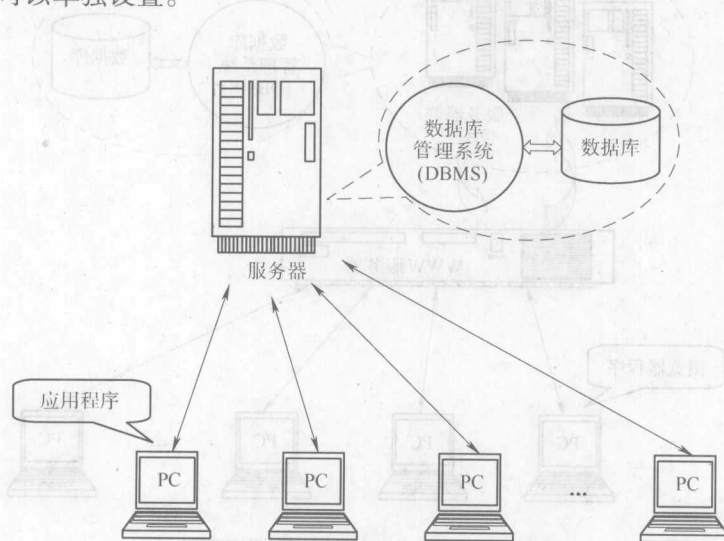


图 1-4 集中式 C/S 结构数据库系统

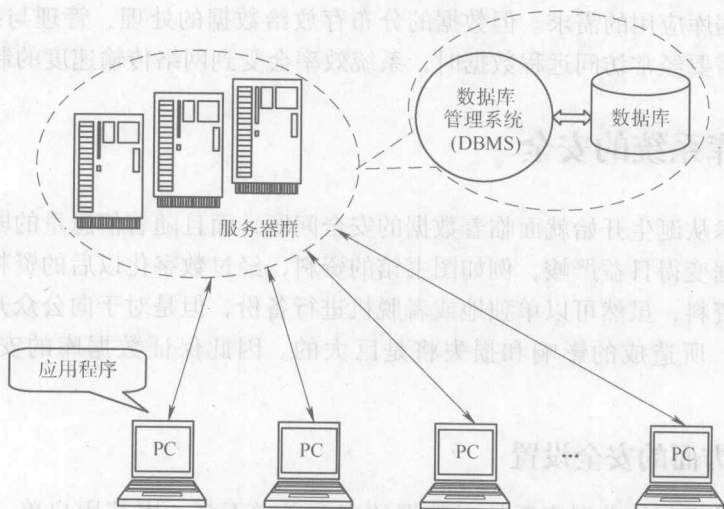


图 1-5 分布式 C/S 结构数据库系统

1.4.4 分布式结构的数据库系统

分布式结构的数据库系统是指数据库中的数据在逻辑上是一个整体，但物理上分布在计算机网络的不同节点上，结构上与分布式 C/S 结构数据库系统相似。网络中的每个节点都可以处理本地数据库中的数据，执行局部应用，也可以同时存储和处理多个异地数据库中的数据，执行全局应用。