

- 21世纪应用型本科教材
- 上海市教育委员会高校重点教材建设项目

# 数据库原理及应用

- 上海市教育委员会 组编
- 张小全 柏海芸 刘梅 荣祺 编著

上海交通大学出版社

21 世纪应用型本科教材

上海市教育委员会高校重点教材建设项目

# 数据库原理及应用

张小全 柏海芸 编著  
刘 梅 荣 祺

上海市教育委员会 组编

上海 交 通 大 学 出 版 社

## 内 容 提 要

本书主要围绕着数据库系统的基本概念、原理、方法和发展趋势等进行了介绍,同时又从应用设计的角度介绍了两个使用最广泛的数据库管理系统:MS Access 和 MS SQL Server,以使读者能理论与实际相结合,初步具备数据库应用的开发能力。

本书是上海市教委“十五”规划教材,在编写中力求体现出“讲求基础、侧重应用”的特色,以达到培养计算机应用型人才的目的。

书中的习题能有效地帮助读者加深对教材内容的理解和掌握。

本书可作为高等院校计算机及相关专业的教学用书,也可作为从事相关领域工作人员的自学参考用书。

### 图书在版编目(CIP)数据

数据库原理及应用/张小全等编著. —上海:上海交通大学出版社,2004

21 世纪应用型本科教材

ISBN7-313-03770-8

I. 数... II. 张... 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(2004)第 057768 号

### 数据库原理及应用

张小全 柏海芸 编著  
刘 梅 荣 棋

上海市教育委员会 组编

上海交通大学出版社出版发行

(上海市番禺路 877 号 邮政编码 200030)

电话:64071208 出版人:张天蔚

太仓市印刷厂有限公司 印刷 全国新华书店经销

开本:787mm×960mm 1/16 印张:18.5 字数:348 千字

2004 年 7 月第 1 版 2004 年 7 月第 1 次印刷

印数:1-2 050

ISBN7-313-03770-8/TP·597 定价:26.00 元

---

版权所有 侵权必究

# 前 言

数据库技术是 20 世纪 60 年代后期产生和发展起来的一项计算机数据管理新技术, 它的产生和发展极大地推动了计算机应用的普及与深入。各行各业的信息系统, 包括在互联网上, 数据库技术的应用比比皆是, 数据库已成为信息社会的重要基础。同时, 也正由于数据库技术的广泛应用, 反过来又极大地促进了数据库技术自身不断的发展与完善。目前, 由于数据库技术与其他信息技术互相交叉与渗透, 一大批新型的数据库系统相继涌现, 数据库技术的新思想、新方法和新手段等也日渐成熟。因此, 数据库技术是计算机科学技术中发展最快、应用最广泛的技术之一, 并且仍在快速地发展之中。

本书作为上海市教委的重点建设教材, 在编写过程中进行了深入的分析与思考。在充分把握本课程内容和特点的基础上, 结合编者多年的教学经验, 并汲取了其他同类教材的诸多精华, 力求使本教材体现出“强化基础、侧重方法、跟踪发展、突出应用”的方针, 充分满足培养计算机应用型人才的目的。

本书共分 10 章, 第 1~8 章围绕着数据库系统的基本概念、原理、方法和发展趋势等展开介绍, 第 9~10 章则重点从应用设计的角度对 MS Access 和 MS SQL Server 分别进行了介绍。其中: 第 1 章让读者对数据库技术的应用情况及其作用等有一个初步的了解, 并掌握一些数据库的基本概念和术语; 第 2 章则对数据模型、数据库系统的体系结构以及数据库管理系统等方面的知识进行进一步的介绍; 第 3 章重点对关系数据库系统的概念、理论、方法以及关系代数等内容作重点描述; 在此基础上, 在第 4 章中对依据关系运算理论提出的关系数据库的标准语言 SQL 作了较为详细的介绍; 第 5 章从判别关系的好坏引出关系的规范化理论, 并可进一步来规范关系; 第 6 章的内容是读者必须重点掌握的知识, 本章系统地介绍了数据库设计中普遍采用的规范设计法, 并对其中的主要部分展开详细讨论; 第 7 章则较为全面、扼要地介绍了数据库安全保护方面的知识, 这部分内容也是目前计算机安全技术方面讨论的热点; 第 8 章向读者介绍目前在数据库技术研究与开发方面的热点及趋势, 让读者感受并有所了解数据库技术的发展情况; 第 9、第 10 章从数据库程序设计的角度, 分别将桌面型的关系数据库系统 MS Access 和分布式的关系数据库系统 MS SQL Server 介绍给读者, 让其对具体的关系数据库管理系统有一定的熟悉和了解, 并初步具备数据库应用的开发能力。书中带“\*”

的章节为选学内容。

本书内容的介绍循序渐进、脉络清晰，希望使读者在掌握基础知识、基本理论的同时，强调对应用技能的掌握；同时，也让读者对新理论、新技术和新方法有所涉猎，为适应数据库技术的飞速发展打下基础。书中的习题可以有效地帮助读者加深对本教材内容的理解和掌握。

本书可作为高等院校计算机及相关专业的教学用书，也可作为从事相关领域工作人员的自学参考用书。

本书由张小全任主编，负责全书的统编、修改及总纂定稿。其中，第1、2、6、8章由张小全编写；第3、4章及第8章部分由柏海芸编写，并负责全书的校对及编排工作；第5、7章由刘梅编写；第9、10章由荣祺编写。

本书承蒙史济民、龚正良、曹渠江等教授进行了认真的审阅，并提出了宝贵的意见，编者在此深表谢意。

本书的编写参考了相关的资料，谨向有关作者表示感谢。

由于作者水平有限，书中的疏漏错误之处敬请读者批评指正。

编 者

2004年5月

# 目 录

|                            |     |
|----------------------------|-----|
| 第 1 章 数据库系统引论 .....        | 1   |
| 1.1 数据库技术的应用 .....         | 1   |
| 1.2 数据库技术的产生和发展 .....      | 5   |
| 1.3 数据库技术的术语 .....         | 12  |
| 第 2 章 数据库的系统结构 .....       | 14  |
| 2.1 信息的三个世界及其描述 .....      | 14  |
| 2.2 数据模型 .....             | 17  |
| 2.3 数据库的体系结构 .....         | 24  |
| 2.4 数据库系统的组成 .....         | 27  |
| 2.5 数据库管理系统 .....          | 29  |
| 2.6 数据库应用系统的体系结构 .....     | 32  |
| 2.7 数据库系统的效益 .....         | 35  |
| 第 3 章 关系数据库系统 .....        | 38  |
| 3.1 关系数据模型 .....           | 38  |
| 3.2 关系代数 .....             | 47  |
| 3.3 关系演算 .....             | 55  |
| 3.4 关系代数表达式的优化策略 .....     | 57  |
| 第 4 章 关系数据库标准语言——SQL ..... | 61  |
| 4.1 SQL 概述及特点 .....        | 61  |
| 4.2 SQL 的数据定义 .....        | 64  |
| 4.3 SQL 数据查询 .....         | 72  |
| 4.4 SQL 数据更新 .....         | 85  |
| 4.5 视图的定义与操作 .....         | 89  |
| 4.6 SQL 数据控制 .....         | 92  |
| 4.7 嵌入式 SQL .....          | 94  |
| 第 5 章 关系数据库的规范化理论 .....    | 104 |
| 5.1 问题的提出 .....            | 104 |
| 5.2 函数依赖 .....             | 106 |

|               |                                       |            |
|---------------|---------------------------------------|------------|
| 5.3           | 关系范式.....                             | 110        |
| 5.4           | 函数依赖理论.....                           | 114        |
| 5.5           | 关系分解原则.....                           | 119        |
| <b>第 6 章</b>  | <b>数据库设计.....</b>                     | <b>125</b> |
| 6.1           | 数据库设计概述.....                          | 125        |
| 6.2           | 需求分析阶段.....                           | 130        |
| 6.3           | 概念结构设计阶段.....                         | 134        |
| 6.4           | 逻辑结构设计阶段.....                         | 142        |
| 6.5           | 物理设计阶段.....                           | 146        |
| 6.6           | 数据库实施阶段.....                          | 148        |
| 6.7           | 数据库运行和维护阶段.....                       | 150        |
| <b>第 7 章</b>  | <b>数据库保护.....</b>                     | <b>153</b> |
| 7.1           | 数据库的安全性.....                          | 153        |
| 7.2           | 数据库的完整性.....                          | 158        |
| 7.3           | 数据库的并发控制.....                         | 167        |
| 7.4           | 数据库恢复.....                            | 176        |
| <b>第 8 章</b>  | <b>数据库技术的研究与发展热点.....</b>             | <b>182</b> |
| 8.1           | 面向对象数据库系统.....                        | 182        |
| 8.2           | 分布式数据库系统.....                         | 188        |
| 8.3           | 并行数据库系统.....                          | 192        |
| 8.4           | 数据仓库.....                             | 197        |
| 8.5           | 新一代数据库技术的特征.....                      | 201        |
| <b>第 9 章</b>  | <b>Microsoft Access 数据库程序设计.....</b>  | <b>203</b> |
| 9.1           | Access 数据库技术.....                     | 203        |
| 9.2           | Access 中的基本操作.....                    | 210        |
| 9.3           | Access 中的编程示例.....                    | 220        |
| <b>第 10 章</b> | <b>Microsoft SQL Server 程序设计.....</b> | <b>240</b> |
| 10.1          | MS SQL Server 数据库.....                | 240        |
| 10.2          | Transact-SQL 概述.....                  | 250        |
| 10.3          | SQL Server 程序设计示例.....                | 266        |
| 10.4          | 在 Access 中访问 SQL Server.....          | 281        |
|               | <b>参考文献.....</b>                      | <b>289</b> |

# 第1章 数据库系统引论

数据库技术是当今计算机领域中应用最广泛、发展最迅速的一个重要分支。它诞生于 20 世纪 60 年代后期，是一项计算机数据管理的新技术。发展至今，已形成了一整套较为完整的理论体系和一大批实用系统。计算机在许多领域的广泛应用与数据库技术的不断发展、完善和推广密切相关。

本章首先通过对数据库技术应用示例的简单介绍，使读者对该技术及其作用有一个初步的、感性的认识；同时也让读者了解市场上流行的数据库产品；然后对数据管理技术的发展作一个简要的回顾；最后介绍了数据管理的新技术，数据库技术的形成、特点及发展趋势。

本章学习要求：

- 初步了解数据库系统的应用情况及数据库技术的作用；
- 初步了解市场上常见的数据库产品；
- 了解数据管理技术的发展过程；
- 了解数据库技术的产生、发展过程及趋势；
- 掌握数据库技术的特点；
- 掌握数据库系统的基本术语、定义。

## 1.1 数据库技术的应用

很多计算机应用系统，如办公信息系统（OIS）、计算机辅助设计与制造（CAD/CAM）、计算机集成制造系统（CIMS）、地理信息系统（GIS）以及其他大量的管理信息系统（MIS）等，都是以数据库技术为基础的。在科学技术迅猛发展的今天，通过与网络、通信、多媒体以及其他各种技术相结合，数据库技术的研究、开发和应用仍然方兴未艾。数据库技术已越来越成为现代信息技术的重要组成部分，是现代计算机信息系统和诸多其他计算机应用系统的基础和核心。

有这样一种说法，一个国家的数据库建设规模（指数据库的个数和种类）、数据库信息量的大小和使用频度已成为衡量这个国家信息化程度的重要标志之一。

### 1.1.1 数据库系统应用示例

在现代社会生活中，只要稍加观察，就可以罗列出许多数据库系统应用的具体实例。如学校的教学管理系统、医疗部门的医疗信息管理系统、图书馆的图书管理系统等。很难想像，离开了这些计算机应用系统的支持，结果将会怎样。

下面，以图书管理系统为例，介绍数据库技术的应用。

#### 1. 存储数据信息

一个图书管理系统所涉及的数据信息有很多，为简单起见，下面只罗列一些主要的信息，如：

- 书籍信息：书号、书名、类别、作者、出版日期、出版社、定价、复本数。
- 出版社信息：名称、地址、电话、法人。
- 借阅信息：借书证、姓名、书号、借书日期、还书日期、备注。

当这些数据信息送入到计算机中时，数据库系统就能有效地把它们管理起来。比如说做下面一些工作：

- (1) 把这些数据信息合理、有效地保存起来。如定义框架、表格等，对数据的类型、格式等作出相应规定，用数据库文件的形式保存在计算机中。
- (2) 对这些数据信息做有关的整理工作。如建立索引等。
- (3) 为防止信息出错或丢失信息等，对数据库信息提供各种保护措施。

#### 2. 实现相关操作

在有效保存数据信息的基础上，系统还能完成如下操作：

- (1) 查询功能：管理人员或读者能多视角查询馆藏书籍的相关信息。
- (2) 借/还书籍：读者借阅、归还书籍的整个过程由系统操作处理，并将相应信息记录保存。
- (3) 书库管理：书籍的采购入库、分类索引、登记造册、淘汰处理等。
- (4) 各种统计工作：为统计之需，系统要做各种数据信息的分析、计算、汇总等工作，并生成各种报表。

当然，一个实际使用的图书管理系统的功能应该还要更多、更完善。但通过上面的粗略介绍，读者对数据库技术就有了一个初步的认识。运用数据库技术开发的系统，不但能有效地管理大量数据信息，而且能简化实际工作中的许多繁琐

操作，并且高效、可靠。

### 1.1.2 数据库系统的作用

直观地来说，数据库就是存放大量数据的仓库。这里有两层意思：一是量大，既能满足现代社会信息量大的需求，又能充分体现出计算机能高效处理大量数据的能力；二是能对仓库中大量数据进行井井有条的管理。这就是数据库管理系统（DataBase Management System, DBMS）的作用。一般而言，数据库管理系统主要用来管理大量数据、控制多用户访问、定义数据库架构、执行数据库各种操作等。数据库系统就是由数据库、数据库管理系统和用户应用系统等组成的。

#### 1. 管理大量数据

大量的数据到底有多大呢？有人做了这样一项说明，目前数据库市场上较先进的数据库产品能支持 10TB 量的数据，这样大的数据究竟是一个怎样的概念呢？下面的一些例子有助于对 10TB 数据量概念的理解。

● 10TB 的数据是 10 995 116 277 760 个字节大小，相当于存储 5 497 558 138 880 个汉字。

- 在航空售票系统中，大约可以存储 183 亿个航班信息。
- 在银行业务系统中，大约可以存储 38 亿个流水账户信息。
- 在超市的销售业务系统中，大约可以存储 26 亿个商品信息。
- 在工厂的 MIS 系统中，大约可以存储 32 亿个客户信息。
- 在学校的教学管理系统中，大约可以存储 42 亿个学生信息。
- 在图书管理系统中，大约可以存储 45 亿本书的简要信息，或者说可以存储 2800 万本 400 页厚的图书的所有内容。

#### 2. 数据库的定义功能

为了有效地管理大量的数据，就需要对数据库本身作一系列的说明和规定。设想一下，即使是一个普通的货物仓库，也需对各种物品按规定整理到位；也就是对仓库中物品进行登记、整理，包括物品的品名、规格、价格、数量、日期、供货商等各种信息，以及物品具体存放的位置等，以便满足收发货物及做各种统计之需，进而达到高效率地管好仓库。而对存放大量数据信息的数据库来说，这样的工作就更必需了。

数据库系统中提供了定义数据库和数据库架构的相关功能。这部分操作由一种被称为数据描述语言（Data Description Language, DDL）的专门工具来实现。

诸如建立数据库, 指定数据库的架构, 如表、视图、索引等对象, 并对其中数据的性质、规格等多方面进行定义、说明。同时还可以对数据库和数据库架构进行修改、删除等操作。

例如, 在最常见的 Microsoft SQL Server 系统中, 可以用其中相应的 Transact-SQL 语言中的下列语句来实现上述操作:

- CREATE DATABASE: 创建数据库。
- ALTER DATABASE: 修改数据库。
- DROP DATABASE: 删除数据库。
- CREATE TABLE: 创建表。
- ALTER TABLE: 修改表。
- DROP TABLE: 删除表。
- CREATE VIEW: 创建视图。
- ALTER VIEW: 修改视图。
- DROP VIEW: 删除视图。

### 3. 数据库的信息查询功能

当数据存储到数据库中之后, 用户就可以使用其中的数据。最常用到的功能就是数据查询操作。即使是修改变化的数据还是删除无用的数据等其他操作, 一般也须先查找到该数据后才能实施。数据库中存放有大量的数据, 因此查询操作是否迅速、高效、方便就显得尤其重要。数据库系统提供了数据库的查询 (Query) 语言, 这种语言也可称为数据操纵语言 (Data Manipulation Language, DML)。使用该语言可以在数据库中执行诸如: 检索指定的数据、插入需要的数据、修改变化或错误的数据库、删除无用的数据等。

同样, 在 Microsoft SQL Server 系统中, 可以实现上述功能的 Transact-SQL 语言中相应的语句有:

- SELECT: 检索数据。
- INSERT: 插入数据。
- UPDATE: 修改数据。
- DELETE: 删除数据。

### 4. 数据库中数据的共享及控制

因为数据库中存放有大量的数据且卓有成效地进行着管理, 因此数据库还有一个重要的作用就是可以最大限度地实现数据共享, 也即各种各样的用户原则上都可以按需使用数据库中的数据。由此也带来了对数据库本身及库中数据的安全

保护问题。事实上,数据库可以让各种各样用户实现数据共享,但也会设置一些限制。一般来说,除了可以公开访问的数据,不同的用户只能使用适合自己需要的数据和实施允许自己执行的操作。这样系统就得设置一些用户的级别、权限等,以确保信息的安全,防止越权和失密。此外,不同的用户在同时使用库中数据时,也得尽可能控制和避免相互间发生干扰而影响正常的操作。再就是万一出现人为的或意想不到的故障而造成库中数据被损坏或丢失时,数据库系统也会采取一系列措施尽可能地恢复数据,减少损失。

上述这些就是以后要讨论的在多用户共享环境下数据库管理要实现的四个方面:数据库完整性、安全性、并发控制和数据库的恢复。

### 1.1.3 常见的数据库系统

面对数据库应用这样一个巨大的市场,各数据库产品供应商竞相推出自己的产品。从早期的 dBASE、FoxBASE 等,到当前被用户使用最多的 Visual FoxPro、Microsoft Access、Sybase、Informix、Oracle 和 Microsoft SQL Server 等,数据库产品始终在不断地推陈出新、升级换代。这些系统各有特点,各具所长,均在数据库市场上占有一席之地,并拥有自己的用户群。本书将在稍后章节中介绍 Microsoft Access 和 Microsoft SQL Server。

## 1.2 数据库技术的产生和发展

计算机诞生初期主要用于做科学计算工作,但从 20 世纪 50 年代起,计算机的应用已由科学研究部门逐步扩展到企事业、行政部门等。至 60 年代,数据处理已成为计算机的主要应用方向。数据处理也称为信息处理,是指从某些已知的数据出发,推导加工出一些新的数据。随着数据处理量的不断增长,要求不断提高,数据管理技术也应运而生。数据管理是指对数据的收集、整理、组织、存储、维护、检索、传送等操作。在数据处理中,通常计算工作比较简单,数据管理工作却比较复杂。但这数据管理工作却是数据处理业务的基本环节,也是任何数据处理业务中必不可少的共有部分。因此可以说,数据管理是数据处理的基础和核心。

## 1.2.1 数据管理技术的发展

数据管理技术的发展,实际上是伴随着计算机软、硬件技术的发展而发展起来的。计算机的数据处理工作,首先需要把大量的数据存放在存储器中,然后才谈得上对其实施有效的数据管理工作。因此,存储器的类型、容量、存取速度等直接影响着数据管理技术的发展。同样,计算机自身数据处理技术能力的加强,计算机应用领域的不断扩展、渗透,以及对数据处理工作不断增长的需求等,都推动了数据管理技术的发展。数据管理技术经历了如下几个不同的发展阶段,并仍在不断地发展之中。

### 1. 人工管理阶段

在这一阶段(20世纪50年代中期以前),计算机主要用于科学计算,其他工作还没有展开。外部存储器只有磁带、卡片和纸带等,还没有磁盘等直接存储设备。软件只有汇编语言,尚无数据管理方面的软件。数据处理的方式基本上是批处理。这个时期的数据管理有下列特点:

(1) 数据不保存在计算机内。计算机主要用于计算,计算完后,数据也不需要长期保存。在进行某一课题计算时,将原始数据随程序一起输入内存,计算处理结束后将结果数据输出。随着计算任务的完成,用户作业退出计算机系统,数据空间随着程序空间一起被释放。

(2) 没有专用的软件对数据进行管理。每个应用程序自身都要包括对数据的存储结构、存取方法、输入输出方式等方面内容的设计。程序中存取代码随着数据存储结构的改变而改变,因而数据与程序不具备独立性。数据存储结构改变时,应用程序必须修改,否则程序就不能正确运行。由于程序是直接面向存储结构的,因此数据的逻辑结构与物理结构没什么区别。

(3) 只有程序(Program)概念,没有文件(File)概念。数据的组织方式必须由程序员自行设计和安排。

(4) 数据面向程序。即一组数据对应一个程序。

### 2. 文件系统阶段

在这一阶段(20世纪50年代后期至60年代中期),计算机不仅用于科学计算,还用于信息处理。但随着数据量的增大,数据的存储、检索和维护等问题就凸显了出来,数据结构和数据管理技术也就迅速发展了起来。此外计算机系统的一些软、硬件环境也得到了改善。如:外部存储器已有磁盘、磁鼓等直接存储设

备；软件领域出现了高级语言和操作系统；操作系统中的文件系统是专门管理外存的数据管理软件；数据处理的方式有批处理，也有联机实时处理等。这一阶段的数据管理有以下特点：

(1) 数据以“文件”形式可长期保存在外部存储器的磁盘上。由于计算机的应用转向信息管理，因此对文件要进行大量的查询、修改和插入等操作。

(2) 数据的逻辑结构与物理结构有了区别，但比较简单。程序与数据之间具有“设备独立性”，即程序只需用文件名就可以与数据打交道，不必关心数据存放的物理位置，由操作系统中的文件系统进行读、写。

(3) 文件组织已呈多样化。有索引文件、链接文件和直接存取文件等。但文件之间相互独立、缺乏联系，数据之间的联系要通过程序去构造。

(4) 数据面向应用。数据不再属于某一特定的程序，可以重复使用。但是文件结构的设计仍然基于特定的用途，程序基于特定的物理结构和存取方法，因此程序与数据结构之间的依赖关系并未根本改变。或者说，数据与程序的独立性程度并不高。

(5) 对数据的操作以记录为单位。这是由于文件中只存储数据，不存储文件记录的结构描述信息。因此，文件系统的基本数据存取单位是记录，即只能先读取整条记录，经处理后才能获得其中的数据项信息，而不能直接对记录中的数据项进行数据存取操作。

文件系统阶段是数据管理技术发展中的一个重要阶段。在这一阶段中，数据结构和算法的充分发展丰富了计算机科学，为数据管理技术的进一步发展打下了基础，且现在仍是计算机软件科学的基础。

尽管如此，随着数据管理规模的扩大，数据量的剧增，文件系统管理阶段的缺陷仍被显现出来。具体表现在：

(1) 数据冗余 (Redundancy)。由于文件间缺乏联系，造成每个应用程序都有对应的文件，有可能同样的数据在多个文件中重复存储。

(2) 数据不一致性 (Inconsistency)。这往往是数据冗余造成的。在进行数据更新操作时稍有不慎，就可能使同样的数据在不同的文件中变得不一样。

(3) 数据联系弱 (Poor data relationship)。这是由于文件之间相互独立，缺乏联系造成的。

上述问题是文件系统本身所难以解决的。由此造成了数据处理效率低、成本高，并有许多潜在的问题需要解决。这也正是数据库系统产生的背景。

### 3. 数据库系统阶段

20 世纪 60 年代后期以来，为了克服文件系统的弊病，适应迅速增长的数据

处理的需求,人们开始探索新的数据管理方法和工具。这一时期,磁盘存储技术取得重要进展,大容量和快速存取的磁盘相继投入市场,给新型数据管理技术的研制提供了良好的物质基础。为了解决多用户、多应用共享数据的需求,使数据为尽可能多的应用服务,数据库技术由此应运而生,出现了统一管理数据的专门软件系统:数据库管理系统(DataBase Management System, DBMS)。

## 1.2.2 数据库技术的发展及特点

数据库(DataBase, DB)一词的提出可以追溯到20世纪50年代后期。据日本的木道夫说:“那时,美国军队为了作战和装备,想把各种数据集中在一个地方作为数据基地。于是就把各个部队分散的情报集中起来管理,根据统一取得的最新数据考虑战略部署。据说,这就是数据库术语的起源”。然而,具有实用价值的真正数据库系统是20世纪60年代后期出现的。这一时期数据管理技术有了突破性的进展,有三个重要事件标志着数据库管理技术进入了数据库系统阶段,这三个事件被称为数据库系统发展的三个里程碑。

### 1. 数据库系统发展的三个里程碑

(1) IMS 系统(1968年)。为满足阿波罗计划的需求,美国的IBM公司与Rockwell公司合作,研制成世界上第一个实用的数据库系统:IMS(Information Management System),为阿波罗飞船1969年顺利登月提供了重要保证。IMS采用了以层次数据结构为基础的数据模型,是层次数据库系统的典型代表。1968年研制出第一个版本:IMS/1,1969年9月投入市场。以后几年相继推出了IMS/2、IMS/VS等版本。虽然这是一个庞大的、耗费资源的、不太灵巧的系统,但它是数据库系统中第一个商用产品,20世纪70年代获得广泛应用。IMS系统对数据库技术的发展产生了重要的影响。

(2) DBTG 报告(1969年)。美国数据系统语言协会(Conference On Data System Language, CODASYL)是由用户和厂商自发组织的团体,成立于1959年。该组织有两大主要贡献,一是在1960年提出的COBOL语言;另一是在1969年提出的DBTG报告。该组织下属一个数据库任务小组(Data Base Task Group, DBTG),专门研究数据库语言。1969年10月该小组提出了一份“数据库建议书”,并在1971年4月正式得到通过,这就是著名的“DBTG报告”。该报告给出了网状数据库系统的方案,为建立网状数据库系统提供了完整的系统设计和语言规范,是网状数据库系统的标准,在数据库系统的发展过程中做出了重要的贡献。以后,根据DBTG报告实现的系统一般都称为DBTG系统(或CODASYL系统)。DBTG系

统在 20 世纪 70 年代至 80 年代中期得到了广泛的应用。

(3) E.F.Codd 的文章 (1970 年)。1970 年 6 月, IBM 公司 San Jose 研究所的 E.F.Codd 在美国计算机协会会刊《Communication of the ACM》发表了题为“大型共享数据库的数据关系模型”(A Relational Model of Data for Large Shared Data Banks)的著名论文,首次全面论述了关系数据库的概念,提出了关系模型,引进了关系代数,推导了关系演算,阐述了数据间存在的函数相关性,概括了关系规范,从而在计算机科学中开创了研究数据库理论与方法的新领域。连同 E.F.Codd 以后发表的多篇论文,为关系数据库的全面开发奠定了坚实的理论基础。作为关系数据库的创始人和奠基人,1981 年 11 月在 ACM (美国计算机协会) 洛杉矶年会上, E.F.Codd 荣获了计算机科学的最高荣誉: 图灵奖。

由于关系模型极其简单, 又有严谨的关系理论为依托, 以此开发的关系数据库产品一经推出即大受欢迎, 并逐步取代层次、网状数据库系统, 成为主流产品。当前用户使用的绝大多数都是关系数据库系统。

## 2. 数据库系统的特点

作为在文件系统基础上发展起来的一种数据管理的新技术, 数据库系统克服了以往数据管理中的诸多缺陷, 提供了对数据更高级、更有效的管理。具体表现出如下特点:

(1) 数据结构化。数据库设计的基础是数据模型。数据库中的数据是按照一定的数据模型来组织、描述和存储的, 称为数据集成化或数据结构化。在进行数据库设计时, 要站在全局需要的角度抽象和组织数据, 完整、准确地描述数据自身和数据之间的联系。因此, 在数据库系统中, 数据不再针对某一应用, 而是面向全局, 面向系统, 具有整体的结构化。

(2) 实现了数据的集中化管理。既然是面向全局、面向整体, 数据库系统通常把一个单位所有部门、所有应用所涉及到的相关数据集中到一个数据库中统一进行管理和维护, 各部门可以从中各取所需。由此实现了数据的集中化管理。

(3) 数据的冗余度小、共享度高。冗余就是不必要的重复。数据集中化管理可以带来的好处是数据在数据库中一般只存储一次, 并为各个不同的用户所共享。不像在文件系统中, 由于数据是面向某一特定用户的, 故不同的用户须自行组织数据, 哪怕其中的有些数据是重复的, 也不能共享。由此, 数据库系统最大限度地减小了数据的冗余度。

同样, 数据的结构化处理、集中化管理所带来的又一好处是实现了数据的多用户共享。数据共享节约了资源, 提高了数据的价值和利用率, 是数据库技术先进性的一个重要体现。

(4) 避免了数据的不一致性。数据的不一致性是由数据冗余所引起的。冗余造成同一数据的重复出现,由于修改等操作的不完备等原因,使其具有不同的值。没有冗余,就不可能出现不一致性。因此,减小了冗余就最大限度地避免了数据的不一致性。

(5) 数据的独立性高。数据库中的数据独立性包括数据的物理独立性和逻辑独立性。数据独立性是指数据库中的数据独立于应用程序,即数据的逻辑结构、存储结构与存取方式的改变不影响应用程序。由于数据库的系统结构分成局部逻辑结构、全局逻辑结构和物理结构三级,因此当数据的物理存储改变了,通过 DBMS 的映射变换最终使应用程序不用改变,此即物理独立性;当数据的逻辑结构改变了,也通过 DBMS 的映射变换使应用程序不用改变,此即逻辑独立性。数据独立性高可以简化应用程序的编制,大大减少应用程序的维护和修改工作。总之,数据库系统的三级结构、二级映像保证了较高的数据独立性。

(6) 数据由 DBMS 统一管理和控制。对数据库的大量管理、维护工作实际上是由数据库管理系统(DBMS)来统一管理和控制的。它提供了数据的安全性保护、完整性保护、并发控制和数据库恢复等各种功能,以确保用户能正常、正确使用数据库系统。

综上所述,数据库技术极大地克服了文件系统等的缺陷,充分展现了其自身的特点,很好地适应了当前数据处理的需求,使数据管理技术得到了飞跃。当然,作为一种计算机应用领域的新兴技术,其自身仍在不断地快速发展之中。

### 1.2.3 数据库技术的热点和发展趋势

尽管 20 世纪的 70 年代,层次、网状、关系三大数据库系统奠定了数据库技术的概念、原理和方法,数据库技术也得到了飞速发展和广泛应用,然而,在许多新的数据库应用领域中,传统的数据库技术和系统已显得力不从心。因此从 80 年代起,数据库技术不断与其他信息技术相结合,新的研究领域被拓展,新的数据库技术被推出。当前及今后一段时期数据库技术发展的重点和热点主要表现为如下几个方面:

#### 1. 分布式数据库技术

传统的数据库技术实现了数据的集中化管理,其优点明显。但随着数据量的剧增,系统越发庞大,操作也更复杂;另外,随着通信技术和网络技术的发展,远程服务应用日益普遍,如果数据集中存储,大量的通信都要通过主机,造成拥挤,效率降低。分布式数据库兼顾了集中管理和分布处理两个方面,因而具有良