

0011110001010001101001111100

数据库技术与 信息系统

DATABASE AND INFORMATION SYSTEM

主编 赵凯 张淑红

天津科学技术出版社

数据库技术与信息系统

主 编 赵 凯 张淑红
副主编 杨 珂 袁立华

天津科学技术出版社

图书在版编目 (CIP) 数据

数据库技术与信息系统 / 赵凯, 张淑红主编. —天津:
天津科学技术出版社, 2012. 1
ISBN 978-7-5308-6038-0

I. ①数… II. ①赵… ②张… III. ①数据库系统
IV. ①TP311. 13

中国版本图书馆CIP数据核字 (2011) 第008053号

责任编辑:傅雪莹

责任印制:张军利

天津科学技术出版社出版

出版人:蔡 颢

天津市西康路 35 号 邮编 300051

电话 (022) 23332400 (编辑室) 23332393 (发行部)

网址: www.tjkjcs.com.cn

新华书店经销

天津午阳印刷有限公司印刷

开本 787×1092 1/32 印张 10.25 字数 253 000

2012年1月第1版第1次印刷

定价: 32.00 元

前 言

所有信息在计算机上都以数据的形式存在，计算机完成各种任务的过程其实就是对数据进行加工和处理的过程。随着人们对客观世界认识的深入，数据的容量越来越大，结构越来越复杂，如何对海量数据进行有效的分类、组织、编码、存储、检索和维护已经成为计算机技术领域研究的重要课题，数据库技术正是为了解决这一课题而提出的。

随着 Internet 的飞速发展，数据库技术也越来越广泛地应用在企事业的各个方面。无论是 Web 应用，企业级的 MIS 乃至 ERP，越是大的应用，数据库的作用就越明显。可以说离开了数据库就不可能有一个信息化的社会。

数据库技术是信息系统的一个核心技术。是一种计算机辅助管理数据的方法，它研究如何组织和存储数据，如何高效地获取和处理数据。数据库技术是研究、管理和应用数据库的一门软件科学。我们将从数据库理论、技术以及与信息系统的关系入手，针对当今主流的数据库技术，如 Oracle 等进行阐述，以及数据库技术与信息系统的关系等内容。

由于我们才疏学浅，加之时间仓促，书中错误之处在所难免，恳请专家、读者批评指正。

编 者

目 录

第一章 数据库系统基础	1
1.1 数据库技术概述.....	1
1.2 数据库的发展简史.....	2
1.3 基本概念.....	4
1.4 数据库系统模型.....	6
1.5 数据库管理系统.....	12
1.6 Codd 十二条法则.....	15
1.7 关系数据库的范式.....	17
1.8 EAR 方法.....	27
第二章 Oracle 基础知识	32
2.1 Oracle 的历史.....	32
2.2 Oracle 系统的特点.....	44
2.3 Oracle9i 组件.....	45
2.4 Oracle9i 的主要特点.....	50
2.5 Oracle 的今天与未来.....	53
2.6 Oracle 认证.....	55
2.7 网上电子公告板 (BBS)	56
第三章 Oracle 体系结构	59
3.1 Oracle 体系结构概述.....	59
3.2 Oracle 文件.....	61
3.3 进程.....	66
3.4 内存结构.....	73
3.5 数据字典.....	77

3.6 事务管理	78
3.7 数据库和实例的启动与关闭	82
3.8 并行服务器数据库	84
3.9 分布式数据库	87
3.10 Oracle9i 系统的高可用性	93
第四章 Oracle 网络数据库	99
4.1 概念和术语	99
4.2 三层结构	105
4.3 电子商务给网络体系结构带来的改变	108
4.4 Oracle 网络计算体系结构	110
4.5 认识 Oracle 应用服务器	111
4.6 因特网文件系统	114
4.7 Oracle Net	116
4.8 使用网络配置助手	120
4.9 Oracle 与防火墙	133
4.10 理解 RAID	135
第五章 SQL * Plus 工具	139
5.1 SQL*Plus 启动与关闭	139
5.2 SQL*Plus 环境设置	147
5.3 编辑和执行命令	158
5.4 输出格式控制	164
5.5 SQL*Plus Worksheet	176
5.6 iSQL*Plus 的应用	179
第六章 数据库对象	185
6.1 表	185
6.2 索引	189
6.3 视图	193
6.4 数据库触发器	197

6.5 同义词	197
6.6 回滚段	199
6.7 数据库链接	200
6.8 序列	201
6.9 包、过程和函数	202
6.10 快照	203
第七章 Oracle 对象—关系数据库	206
7.1 什么是面向对象数据库系统	206
7.2 面向对象数据库系统的对象理论	208
7.3 对象—关系数据库	213
7.4 定义对象类型	215
7.5 对象的使用	219
第八章 数据库设计	224
8.1 数据库设计的基本任务	224
8.2 数据库设计的基本过程	225
8.3 需求分析阶段	226
8.4 概念结构设计	231
8.5 逻辑结构设计	232
8.6 建立数据库	238
第九章 系统开发	246
9.1 数据库的生存周期	246
9.2 应用程序开发方法	250
9.3 对象的发现和标志	256
9.4 数据分析	263
9.5 管理技术	265
9.6 Oracle Developer	267
9.7 Oracle9i JDeveloper	275
9.8 Oracle WebDB	276

第十章 Oracle 优化和调试	282
10.1 选择平台	282
10.2 内存/CPU 建议	284
10.3 优化程序	287
10.4 SQL 的调整	291
10.5 调试一个 Oracle 数据库	294
10.6 调试 SGA	297
10.7 Oracle 网络计算环境的调整与优化	299
10.8 调整 Windows NT 以优化 Oracle9i	303

第一章 数据库系统基础

本章主要介绍数据库系统的发展简史、关系型数据库系统的概念、关系数据库系统的模型和结构、Codd 十二条法则、关系型数据库的规范化、EAR 方法等内容。

1.1 数据库技术概述

人类社会正在进入信息化的社会，人们在政治、经济、科学研究、文化教育等各个领域都将产生大量的信息，这些信息需要加工和处理，需要交流和应用。例如，一项商务活动显然必须跟踪客户、订单、产品库存以及员工信息。此外，相关业务信息的分析有助于一个公司更具有竞争力。例如，销售分析可利用当前的销售数据来预测未来的销售，并可能有助于确定提高所有商务收益率的发展方向。

在当今这个高技术的时代，多数信息由计算机管理，因为计算机使它们容易组织、存储，而且计算机还能保护有价值的信息。功能强大的个人计算机和网络的激增已经使所有商务（或大或小）能快速、安全地将信息传送到需要访问它的人员之处。

客观世界所有信息在计算机上都以数据的形式存在，计算机完成各种任务的过程其实就是对数据进行加工和处理的过程。随着人们对客观世界认识的深入，数据的容量越来越大，结构越来越复杂，如何对海量数据进行有效的分类、组织、编码、存储、检索和维护已经成为计算机技术领域研究的重要课题，数据库技术正是为了解决这一课题而提出的。

随着 Internet 的飞速发展，数据库技术也越来越广泛地应用在企

业的各个方面。无论是 Web 应用，企业级的 MIS 乃至 ERP，越大的应用，数据库的作用就越明显。可以说离开了数据库就不可能有一个信息化的社会。

1.2 数据库的发展简史

20 世纪 60 年代以前，计算机主要用于科学计算。进入 20 世纪 60 年代以后，在社会大系统中出现了巨大的信息流和相伴随的宏大的数据流，为了适应和满足社会发展的需要，必须采用新的技术和手段，对这些数据进行收集、存储、加工、检索、分类、统计和传输等，数据库技术是应数据管理任务的需要而产生的。这就是数据库技术产生的历史背景。

数据库技术产生于 20 世纪 60 年代末。首先使用“Data Base”一词的是美国系统发展公司在为美国海军基地研制数据中引用。

1968 年，IBM 公司在数据库管理系统方面率先研制成功了集成数据存储系统，它可以让多个程序共享数据库。

1969 年 10 月，CODASYL 数据库研制者提出了网状模型数据库系统规范报告 DBTG，使数据库系统开始走向规范化和标准化。

1970 年 IBM 公司 San Jose 研究所的 E·F·Codd 在美国计算机学会会刊《Communication of the ACM》上发表的题为“A Relation Model of Data for Shared Data Bank”的论文，开创了数据库系统的新纪元。这三件事情奠定了数据库技术的基础。

之后，E·F·Codd 连续发表了多篇论文，成功地奠定了关系数据理论的基石。由于他的杰出贡献，1981 年 E·F·Codd 获得了计算机科学的最高奖——ACM 图林奖。

1971 年美国数据系统语言协会在正式发表的 DBRG 报告中，提出了三级抽象模式，即对应用程序所需的那部分数据结构描述的外模式，对整个客体系统数据结构描述的概念模式，对数据存储结构描述的内模式，解决了数据独立性的问题。

20 世纪 70 年代末，关系方法的理论研究和软件的研制均取得了很大成果，IBM 公司的 San Jose 实验室在 IBM370 系列机上研制的关系数据库实验系统 System R 获得成功。1981 年 IBM 公司又宣布了具有 System R 全部特征的新的数据库软件产品 SQL/DS 问世。

与 System R 同期，美国加州大学伯克利分校也研制了 Ingress 关系数据库实验系统，并由 Ingress 公司发展成为 Ingress 数据库产品。

1976 年美籍华人陈平山提出了数据库逻辑设计的实际联系方法。1978 年新奥尔良发表了 DBDWD 报告，他把数据库系统的设计过程划分为四个阶段：需求分析、信息分析与定义、逻辑设计和物理设计。1980 年 J·D·Unman 所著的《数据库系统原理》一书正式出版。

1984 年 David Marer 所著的《关系数据库理论》一书，标志着关系数据库在理论上的成熟。

在关系数据库管理系统 (RDBMS) 产生之前，层次 (IMS) 和网状 (IDMS) 模式是常见的。在这些模式之前，使用平面文件 (操作系统文件不一定平面) 来建立数据库，并且使用第三代语言 (3GL) 访问例程。实际上一些专用系统仍然是按这种方式建立，只是进行了修改或没有变化。在大型机和微机中依然存在着许多这样的遗留数据库。CODASYL (数据系统语言协会) 是数据库任务组 (Database Task Group, DBTG) 创建的一种数据库标准，这是一种基于 COBOL 的网络数据库标准，并且 IDMS 是一个厂商的实现。但是，从 20 世纪 70 年代起，RDBMS 已经逐渐地控制了市场，如 Oracle、Sybase、Informix 和 Ingress。

最近，面向对象 (Object-Oriented OO) 的数据库管理系统 (DBMS) 已经成为最为突出的数据库管理系统，并找到了许多适当的应用环境，如 CAD/CAM、工程、多媒体等。

面向对象 DBMS 适于在这些领域中应用，因为在一个几乎非事

务性的环境中，它们具有控制复型数据类型的实力。由于竞争，RDBMS 厂商为了提供包括文本、音频、图像和视频数据类型的面向对象/多媒体性能，已经制造了商业可用的通用服务器。Oracle 的 Universal Server 就是一个例子。另外，用户定义的数据类型或可扩展类型，已经被扩大或增加到核心数据库服务器中，Oracle9i 就提供了这样的性能。类似这样的 RDBMS 产品被认为是混合的，然而它们明显比以前的 RDBMS 更具有主流性。

1.3 基本概念

数据、数据库、数据库管理系统和数据库系统是与数据库密切相关的四个基本概念。

1.3.1 数据

为了了解世界、研究世界和交流情况，人们需要描述各种事物。用自然语言来描述虽然很直接，但事物用它来描述未免太繁琐，也不便于形式化。为此，人们常常只描述那些感兴趣的事物的特征或属性，作为事物的描述。

1.3.2 数据库

数据库 (Data Base, 简称 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存储设备上, 而且数据是按一定的格式存放的。

人们收集并抽取一个应用所需要的大量数据之后, 应将其保存起来, 以供进一步加工处理, 进一步抽取。在科学技术飞速发展的今天, 人们的视野越来越广, 数据量急剧增加。过去人们把数据存放在文件柜里, 现在人们借助计算机和数据库技术科学地存放和管理大量的数据, 以便能方便而充分地利用这些宝贵的信息资源。

所谓数据库是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张性, 并可为各种

用户共享。

1.3.3 数据库管理系统

了解了数据和数据库的概念，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（Data Base Management System，简称 DBMS）。数据库管理系统就是管理一个数据库的软件，它充当所有数据的知识库，并列它的存储、安全、一致性、并发操作、恢复和访问负责。

DBMS 有一个数据词典（有时被称为系统目录），其中贮存着它拥有的每个事物的数据，例如名字、结构、位置和类型，这种关于数据的数据也被称为元数据（metadata）。在一条数据的生存周期里（从它的创建到删除），这条数据的逻辑和物理信息都被记录在数据词典中。数据库系统管理员（Database Administrator，DBA）应该熟悉 DBMS 的数据词典；在数据库的整个生命周期内，数据词典为他（她）服务。

数据库管理系统的主要功能包括以下几个方面：

1.数据库的定义功能

DBMS 提供了数据的定义语言（Data Definition Language，简称 DDL），用户通过它可以方便地对数据库中的数据对象进行定义。

2.数据操纵功能

DBMS 还提供了数据操纵语言（Data Manipulation Language，简称 DML），用户可以使用 DML 操纵数据，实现对数据库的基本操作，如查询、插入、删除和修改等。

3.数据库的运行管理

数据库在建立、运用和维护时由数据库管理系统统一管理、统一控制，以保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。

4.数据库的建立和维护功能

它包括数据库初始数据的输入、转换功能，数据库的转储、恢复功能，数据库的重组功能和性能监视、分析功能等。这些功能通常是由一些实用程序完成的。

数据库管理系统是数据库系统的一个重要组成部分。

1.3.4 数据库系统

数据库系统 (DataBase System, 简称 DBS) 的简单结构如图 1-1 所示。图中的数据是数据的汇集，它们以一定的组织形式存于存储介质上，一般是磁盘。数据库管理系统 (DBMS) 是管理数据库的软件，它实现数据库系统的各种功能。图中的应用是指以数据库为基础的各种应用程序，应用程序必须通过 DBMS 访问数据库。数据库既然是共享的，就需要有人进行数据库的规划、设计、协调、维护和管理等工作，负责这些工作的人员或集体成为数据库管理员 (DataBase Administrator, 简称 DBA)。应用程序、数据库管理系统 (及其开发工具)、数据库和数据库管理员构成数据库系统。数据库系统有时也简称数据库。

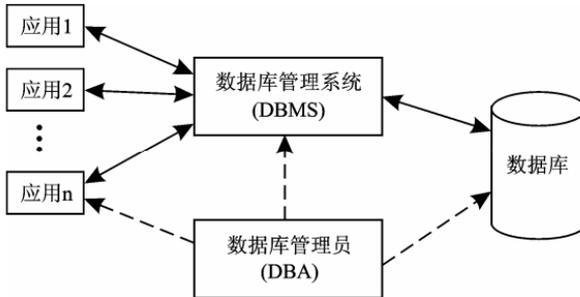


图 1-1 数据库系统

1.4 数据库系统模型

模型是现实世界的模拟和抽象。数据模型 (Data Model) 也是一种模型，它是现实世界数据特征的抽象。数据模型是数据库系统

中的一个关键概念，是实体与实体之间的一个轮廓视图，是数据库系统用以提供信息表示和操作手段的形式框架。

数据库中的数据是高度结构化的，即数据库不仅要考虑记录内的数据项间的联系，还要考虑记录之间的联系。

下面对数据库模型进行介绍：

在数据库的发展史上，最有影响的数据库模型通常有以下 4 种：

(1) 层次模型 (Hierarchical Model)，用树型结构描述实体及其联系。

(2) 网状模型 (Network Model)，用网状结构描述实体及其联系。

(3) 关系模型 (Relational Model)，用二维表结构描述实体及其联系。

(4) 数据独立存取模型，将数据表现为四级模型。即：实体集模型、串模型、编码模型、物理设备模型。

层次模型和网状模型又称为格式化模型，这类模型的数据结构可用图来描述。实体用记录型来表示，一个记录型对应于一个结点，记录型之间的联系对应于结点之间的连接弧。对应于树型图的数据模型为层次模型，对应于网状图的数据模型为网状模型。关系模型为非格式化模型。在关系模型中，使用单一的二维表结构表示实体与实体之间的关系。下面详细介绍三种常用模型。

1.4.1 层次模型

层次模型是以记录型为结点构成的树，是一个以记录型为结点的有根的定向树或者森林。

层次模型把客观问题抽象为一个严格的自上而下的层次关系。树的结构由结点和连线组成，结点用来表示实体，连线用来表示实体和实体之间的关系。客观世界存在的这种关系只能是 1 : N 关系，树的最高位置只有一个称为根结点。每一个结点的上方结点是该结点的父结点，该结点的下方是该结点的子结点。没有子结点的结点

称为叶。

在数据库中定义满足下面两个条件的基本层次联系的集合为层次模型：

- (1) 有且仅有一个根结点，无父结点。
- (2) 其他结点有且仅有一个父结点。

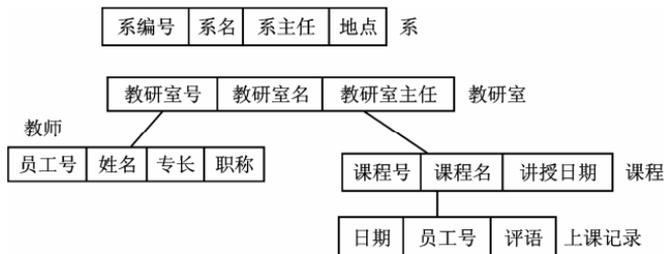


图 1-2 层次模型示意图

1) 层次模型的优点主要有：

- (1) 层次数据模型本身比较简单。
- (2) 对于实体间联系是固定的，而预定定义好的应用系统，采用层次模型来实现，其性能优于关系模型，不低于网状模型。
- (3) 层次数据模型提供了良好的完整性支持。

2) 层次模型的缺点主要有：

- (1) 现实世界中很多联系是非层次性的，如多对多联系、一个结点具有多个双亲等，层次模型表示这类联系的方法很笨拙，只能通过引入冗余数据（易产生不一致性）或创建非自然的数据组织（引入虚拟结点）来解决。
- (2) 对插入和删除操作的限制比较多。
- (3) 查询子女结点必须通过双亲结点。
- (4) 由于结构严密，层次命令趋于程序化。

可见用层次模型对具有一对多的层次关系的部门描述非常自然、直观，容易理解。这是层次数据库的突出优点。

1.4.2 网状模型

网状模型能很好地反映客观世界中较为复杂的事物之间的联系，网状模型是以记录型为结点的网络。网状模型的基本特征是一个父结点允许有多个子结点，一个子结点也可以有多个父结点。

在数据库中，把满足以下两个条件的基本层次联系集合称为网状模型：

- (1) 有一个以上结点，无父结点。
- (2) 至少有一个结点有多于一个的父结点。

网状模型反映实体间复杂关系的表达能力比较强，也就是说，它既能表达实体间的纵向联系，又能表达实体间的横向联系。但是，网状模型在概念、结构和使用方面都比较复杂，对计算机的软件和硬件环境要求比较高。在如图 1-3 所示的网状模型示意图中，4 种零件由 3 个不同的厂家供应，并且由这些零件组装成 3 种部件。

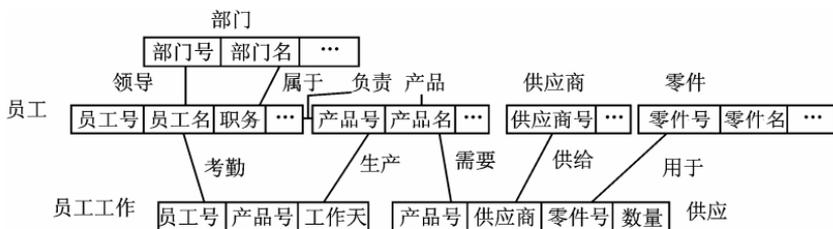


图 1-3 网状模型示意图

用 $M:N$ 关系可以直接描述网状模型。从逻辑上看，层次模型和网状模型都是使用连线表示实体和实体之间的关系，使用结点表示实体。从物理上看，层次模型和网状模型都是使用指针表示实体之间的关系。

1) 网状数据模型的优点主要有：

(1) 能够更为直接地描述现实世界，如一个结点可以有多个双亲。