

SHUJUKU YUANLI
SHUJUKU YUANLI
SHUJUKU YUANLI
SHUJUKU YUANLI
SHUJUKU YUANLI

数据库 原理

郑若忠 宁 洪 等编著

 YUANLI

TP311.13

Z40

数 据 库 原 理

郑若忠 宁洪 阳国贵 尹俊文 田兴彦 编著

7

国防科技大学出版社

· 长沙 ·

图书在版编目 (CIP) 数据

数据库原理/郑若忠; 宁洪; 阳国贵; 尹俊文; 田兴彦编著. —长沙: 国防科技大学出版社, 1998. 8

ISBN 7-81024-490-6

I. 数据库原理

II. 郑若忠 宁洪 阳国贵 尹俊文 田兴彦

III. ①数据库②数据库原理

IV. TP311.13

国防科技大学出版社出版发行

电话: (0731) 4555681 邮政编码: 410073

E-mail: gfkdcbs @ public.cs.hn.cn

责任编辑: 黄 煌 责任校对: 潘 生

新华书店总店北京发行所经销

国防科技大学印刷厂印装

*

787×1092 1/16 印张: 21.25 字数: 491 千

1998 年 8 月第 1 版 1998 年 11 月第 2 次印刷 印数: 5001-15100 册

*

定价: 25.00 元

内 容 提 要

本书系统地介绍了数据库原理，概述了数据库前沿技术。全书共分十章：第一章，数据库系统基本概念；第二章，关系数据模型及语言；第三章，其它数据模型；第四章，数据的物理组织；第五章，查询处理与优化；第六章，事务管理；第七章，完整约束与安全保护；第八章，关系模型设计理论；第九章，数据库应用系统设计；第十章，数据库前沿技术。

本书内容丰富，全面，系统，深度广度兼顾。可作为高等院校和科研院所的计算机专业和相关专业的数据库课程的教材，也可作为有关人员学习、研究数据库理论与技术的参考书。

前 言

数据库技术经过 30 多年的迅速发展,取得了辉煌的成就,众多品牌的数据库系统已深入计算机应用的各个领域,可以毫不夸张地说,哪里有计算机,哪里就有数据库。数据库技术在人类社会信息化、信息服务社会化、社会经济数字化的历史进程中推波助澜,它的重要性是不言而喻的。

1983 年,我们出版了国内第一本数据库教材《数据库原理与方法》,受到国内外广大读者的欢迎,许多读者鼓励我们再版。遗憾的是,我们未能及时满足大家的希望。时过境迁,数据库系统的面貌日新月异,数据库课程也广为普及。我们原来的教材,有的内容已经陈旧,有的观点已经过时,而新的技术和新的发展方向层出不穷,再版已不能解决问题,需要重写一本涵盖这些变化和发展的新教材,以替代原来的教材。重写工作相当艰巨,要求作者不但有多年的数据库教学和科研的经验,还要有相关领域的知识和扎实的理论基础;既要熟练掌握数据库领域相对稳定的理论与技术,又要了解数据库的发展前沿技术。我们经过两年多的努力,几易其稿,终于将面貌一新的《数据库原理》一书奉献给读者。

我们认为本书有两大特点:(1)为了给读者提供一个合理的知识结构,并适应不同读者的需要,选材经过仔细推敲:既重点考虑当前有价值特别是有实用价值的内容,又考虑一些与数据库发展历史有关材料;既考虑数据库管理系统的设计、研究,又适当考虑数据库应用系统的开发;既考虑严格、系统的理论基础,又注意拓宽知识面;既考虑成熟的实用技术,又顾及数据库发展的前沿技术。(2)本书结构清晰、合理,采用自顶向下的叙述方法,概念清楚,层次分明,系统性好。全书共分十章,第一章为数据库系统概论;第二、三章介绍各种数据模型;第四至七章介绍 DBMS 设计的有关内容;第八、九章介绍数据库系统的设计方法;第十章简介数据库领域的发展前沿技术。本书适用于计算机专业本科教学,去掉打*号的章节,也可供非计算机专业读者使用。

本书由郑若忠、宁洪牵头,集体讨论大纲,分头编写,集体修改定稿。宁洪编写第一、二、三、五、七、八章;阳国贵编写第十章;尹俊文编写第四、六章;田兴彦编写第九章;翟桂英负责全书文字和图表的录入工作。

本书的编写出版工作得到国防科技大学各级领导、机关的支持,特别是国防科技大学出版社和计算机学院的支持,在此表示感谢。由于数据库技术发展极为迅速,涉及的知识面宽、内容广,虽经我们艰苦努力,但错漏之处在所难免,欢迎专家、读者批评、指正。

编者

1998 年 3 月

目 录

第一章 数据库系统基本概念

1.1 数据库	(1)
1.1.1 数据管理技术的发展	(1)
1.1.2 数据库系统	(4)
1.2 现实世界的描述	(6)
1.2.1 信息的三个领域	(6)
1.2.2 实体模型	(7)
1.2.3 数据模型	(11)
1.3 数据库的体系结构	(18)
1.3.1 数据库的分级结构	(18)
1.3.2 模式及其映射	(19)
1.3.3 数据独立性	(22)
1.4 数据库管理系统	(22)
1.4.1 DBMS 的功能	(22)
1.4.2 DBMS 的程序组成	(23)
1.4.3 数据语言	(24)
1.4.4 数据字典	(26)
1.4.5 用户访问数据的过程	(27)
1.4.6 DBMS 的分类	(27)
习题	(28)

第二章 关系数据模型及语言

2.1 关系模型	(30)
2.1.1 关系及其基本术语	(30)
2.1.2 关系数据库模式及完整约束	(32)
2.2 关系运算	(40)
2.2.1 关系代数	(40)
2.2.2 元组关系演算	(48)
2.2.3 域关系演算	(54)

* 2.2.4	关系运算的安全围	(55)
* 2.2.5	三类关系运算表达能力的等价性	(58)
2.3	关系数据库语言	(62)
2.3.1	三种典型的关系查询语言	(62)
2.3.2	SQL 语言	(74)
2.4	几种常用的关系数据库	(83)
2.4.1	Oracle	(83)
2.4.2	IBM DB2	(84)
2.4.3	Sybase	(84)
2.4.4	Informix	(86)
2.4.5	Ingres	(86)
2.4.6	Microsoft SQL Server	(87)
习题		(87)

第三章 其它数据模型

3.1	层次数据模型	(90)
3.1.1	层次数据库的基本概念	(90)
3.1.2	层次模型的完整约束	(95)
* 3.1.3	IMS 层次模型的数据定义及操作	(96)
3.2	网络数据模型	(105)
3.2.1	网络数据库的基本概念与数据定义	(105)
3.2.2	网络模型的完整约束	(116)
* 3.2.3	网络数据库的操作	(118)
3.3	面向对象数据模型	(120)
3.3.1	面向对象数据库	(120)
3.3.2	面向对象数据模型	(121)
* 3.4	基于逻辑的数据模型	(124)
3.4.1	Datalog	(125)
3.4.2	逻辑数据库	(126)
习题		(127)

第四章 数据的物理组织

4.1	文件基本的组织方法	(130)
4.1.1	流水文件	(130)
4.1.2	顺序文件	(132)
4.1.3	Hash 文件	(136)
4.1.4	文件空间动态变化的 Hash 技术	(144)
4.2	文件的索引结构	(148)

4.2.1	索引顺序文件	(149)
4.2.2	索引无序文件	(149)
4.2.3	索引文件的查找	(150)
4.2.4	索引的组织	(151)
4.2.5	聚集索引	(155)
4.3	B 树	(157)
4.3.1	B 树	(157)
4.3.2	B ⁺ 树	(159)
* 4.3.3	B 树的变种	(163)
4.4	多关键字检索	(163)
4.4.1	倒排文件	(164)
4.4.2	多重表	(166)
* 4.4.3	分部 Hash 函数	(167)
* 4.4.4	多维树	(169)
* 4.4.5	联想处理	(171)
* 4.5	数据压缩技术	(172)
4.5.1	统计编码	(173)
4.5.2	预测编码	(176)
4.5.3	变换编码	(177)
4.5.4	基于字典的压缩方法	(180)
* 4.6	Microsoft SQL Server 的存储结构	(181)
4.6.1	基本结构	(181)
4.6.2	页的内容	(181)
4.6.3	索引结构	(182)
4.6.4	数据存取	(182)
	习题	(183)

第五章 查询处理与优化

5.1	查询处理	(185)
5.2	查询处理的访问例程	(186)
5.2.1	选择运算的实现	(186)
5.2.2	连接运算的实现	(187)
5.2.3	投影运算的实现	(189)
5.2.4	集合运算的实现	(189)
5.2.5	组合操作的例程	(191)
5.3	查询优化	(192)
5.3.1	规则优化	(192)
5.3.2	物理优化	(195)

* 5.4 代价估算优化	(196)
5.4.1 询问执行代价的组成	(196)
5.4.2 代价估算信息	(196)
5.4.3 代价估算	(197)
习题	(200)

第六章 事务管理

6.1 事务管理的基本概念	(201)
6.1.1 事务的状态	(201)
6.1.2 事务的特性	(202)
6.2 并发控制	(203)
6.2.1 串行化调度	(203)
6.2.2 加锁技术	(205)
6.2.3 基于时标的并发控制技术	(210)
* 6.2.4 其它并发控制技术	(211)
6.3 恢复	(213)
6.3.1 恢复的一般原则	(213)
6.3.2 恢复的基本机制	(214)
6.3.3 基于延迟修改的恢复技术	(216)
6.3.4 基于立即修改的恢复技术	(216)
* 6.3.5 其它的恢复技术	(218)
* 6.4 Microsoft SQL Server 中的事务管理	(219)
6.4.1 Microsoft SQL Server 中的事务	(219)
6.4.2 Microsoft SQL Server 中的日志和检查点	(220)
6.4.3 SQL Server 的加锁机制	(220)
习题	(222)

第七章 完整约束与安全保护

7.1 数据库保护的基本内容	(224)
7.2 完整约束	(226)
7.2.1 完整约束说明	(226)
7.2.2 完整约束检验	(227)
7.3 安全保护	(228)
7.3.1 系统安全分级	(228)
7.3.2 安全保护的一般策略	(230)
7.3.3 用户标识与鉴别机制	(231)
* 7.3.4 信息流动控制	(232)
* 7.3.5 推断控制	(233)

7.3.6 访问控制	(236)
* 7.3.7 数据加密	(239)
7.3.8 审计	(241)
7.3.9 数据库安全检查纵览	(242)
习题	(243)

第八章 关系模型设计理论

8.1 函数相关性	(245)
8.1.1 函数相关性的定义	(245)
8.1.2 FD 的蕴涵与公理	(246)
8.1.3 FD 集合的等价性	(250)
8.1.4 FD 集合的正则覆盖	(252)
* 8.2 多值相关性	(253)
* 8.3 连接相关性	(256)
8.4 关系模型的规范理论	(256)
8.4.1 关系模型的评价	(256)
8.4.2 关系框架的分解	(258)
8.4.3 关系的 1NF 和 2NF	(262)
8.4.4 关系的 3NF 和 BCNF	(264)
* 8.4.5 关系的 4NF 和 5NF	(267)
习题	(269)

第九章 数据库应用系统设计

9.1 数据库应用系统设计概述	(271)
9.2 系统需求分析	(272)
9.2.1 确定设计范围	(273)
9.2.2 数据收集与分析	(273)
9.2.3 需求说明书	(274)
9.3 概念模型设计	(275)
9.3.1 局部视图设计	(275)
9.3.2 视图集成	(276)
* 9.4 逻辑模型设计	(277)
9.4.1 逻辑模型设计概述	(277)
9.4.2 DBMS 的选择	(278)
9.4.3 逻辑转换	(279)
9.4.4 模型评价	(280)
* 9.5 物理模型设计	(281)
* 9.6 数据库应用程序的开发	(282)

9.6.1 开发过程·····	(282)
9.6.2 Oracle 应用开发工具·····	(283)
9.6.3 PowerBuilder·····	(283)
习题·····	(285)

第十章 数据库前沿技术

10.1 概述·····	(286)
10.2 分布式数据库·····	(287)
10.2.1 目标与分类·····	(287)
10.2.2 主要研究内容·····	(288)
10.2.3 产品与原型系统·····	(291)
* 10.3 并行数据库·····	(292)
10.3.1 并行性与体系结构·····	(293)
10.3.2 并行算法与查询优化·····	(294)
10.3.3 物理组织与数据划分·····	(295)
* 10.4 知识库与演绎数据库·····	(296)
10.4.1 概述·····	(296)
10.4.2 相关问题的研究·····	(297)
10.4.3 原型系统·····	(298)
10.5 对象—关系数据库·····	(300)
10.5.1 背景·····	(300)
10.5.2 对象—关系数据库技术·····	(302)
10.5.3 ORDB 系统与实现途径·····	(304)
10.6 多媒体数据库·····	(306)
10.6.1 多媒体数据管理的基本要求·····	(306)
10.6.2 多媒体数据库的功能与结构·····	(308)
10.6.3 多媒体数据库系统的实现途径·····	(309)
10.7 数据仓库·····	(310)
10.7.1 产生背景·····	(310)
10.7.2 数据仓库的特点、组成及工具·····	(311)
10.7.3 数据仓库的建立过程·····	(313)
10.8 ODBC 及数据库与 Web 的连接·····	(314)
10.8.1 ODBC·····	(314)
10.8.2 JDBC·····	(316)
10.8.3 数据库与 Web 的连接·····	(316)
* 10.9 主动数据库·····	(318)
10.9.1 主动数据库及研究内容·····	(319)
10.9.2 原型系统·····	(321)

* 10.10	其它数据库	(321)
10.10.1	模糊数据库.....	(321)
10.10.2	模糊演绎数据库.....	(322)
10.10.3	工程数据库.....	(322)
10.10.4	空间数据库.....	(323)
10.10.5	统计与科学数据库.....	(323)
10.10.6	实时数据库.....	(323)
10.10.7	时态和历史数据库.....	(324)
10.10.8	主存数据库.....	(324)
参考文献.....		(325)

第一章 数据库系统基本概念

20 世纪 60 年代末,数据库技术初露头角,随即得到迅速发展,成为数据处理的公用支撑技术。时至今日,数据库系统已遍布政府机关、社会团体和各个行业,存储着它们赖以正常运转的数据资源,显著地提高了工作效率和质量,产生出极大的社会、经济效益。数据库技术的发展一方面受到应用需求的牵引和驱动,另一方面又促进和拓展了计算机的应用领域,同时也改变了人们运用信息的方式,有力地推动了社会信息化的进程。本章简要介绍数据库技术的发展历史及有关概念。

1.1 数据库

1.1.1 数据管理技术的发展

人类活动的整个历史,离不开对信息和数据的收集、保存、传输、处理和利用。自 20 世纪 40 年代电子计算机问世以后,人们将十进制数变为二进制数以便于计算机的存储和处理,随后又将文字编码成位串形式,进而将图形、图像、声音等多媒体信息数字化。时至今日,除暗示外,几乎所有信息均可表示成计算机能识别的字符串或位串,为信息的迅速传播和处理提供了可能,人们对信息、数据的利用和处理逐步实现自动化、网络化和社会化,整个人类社会已成为信息化的社会。

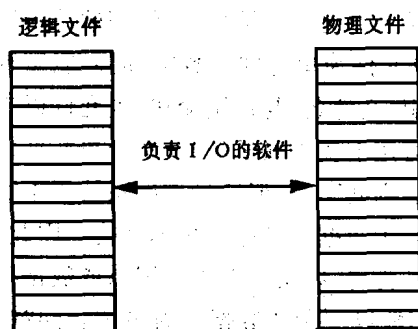


图 1.1 初期文件系统特征

数据库技术正是在满足这日益增长的数据处理需要的过程中,由文件系统发展起来的一种理想的数据管理技术。60 年代早期的文件系统是数据管理方法的雏形,其文件在外存的物理结构与用户观点的逻辑结构完全一致,用户的数据文件主要存储在磁带上,其组织方式是顺序式的,这时的数据管理软件属于操作系统的一部分,其主要功能是完成 I/O 设备的输入/输出操作,如图 1.1 所示。显然,这种组织形式只能应付批处理,不适用于实时访问。由于让用户各自建立文件,用户不但要编制应用程序,还要花费很大精力为数据文件设计物理细节,而且文件不宜共享,数据冗余度高,当文件的物理结构发生变化或更换外存设备时,需修改或重编应用程序,用户负担沉重,人们不得不探索新的数据管理方法。

大约到 60 年代中后期,直接访问设备——磁鼓、磁盘的性能改善使数据组织发生了如下变化:除磁带外,磁鼓、尤其是磁盘成为联机的主要外存设备;文件的物理结构与逻辑

结构之间已有所区别(其关系如图 1.2 所示),在文件的物理结构中增加了链接和索引等形式,因而对文件中的记录可顺序和随机地访问;数据管理软件(仍属操作系统的一部分)提供从逻辑文件到物理文件的“访问方法”是这一时期数据管理的主要特征。对系统而言,

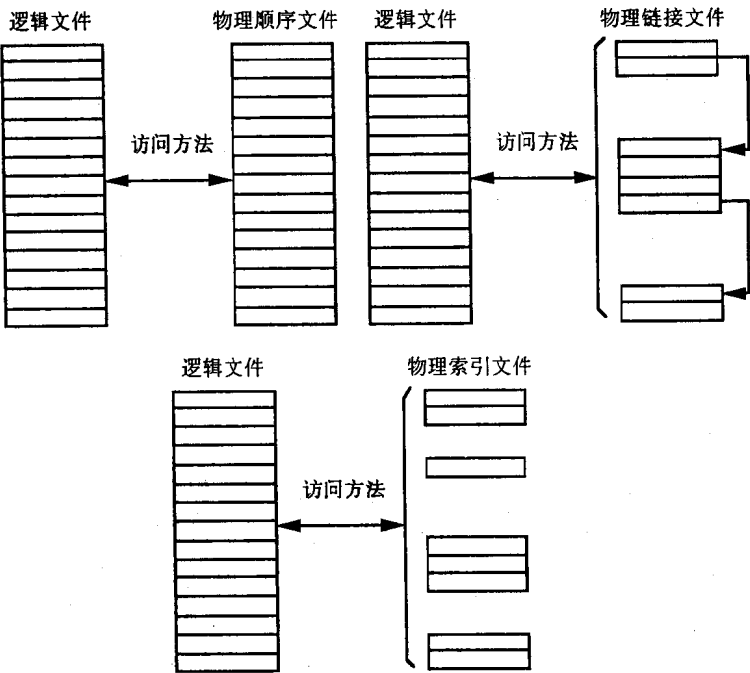


图 1.2 带有访问方法的数据组织

组成文件的最小单位是记录,记录之间仅存在顺序关系而无其它结构;在管理软件中还增加了安全、保密检查机构;部分系统允许用户之间以文件为单位共享数据,但未能实现以记录和数据项为单位的数据共享;用户仍以文件标识(文件名)与系统交往,也允许以文件中的记录标识访问数据。它不但适用于批处理,也可用于实时联机任务;系统更换外设也无需用户修改应用程序;可以实现以文件为单位的数据共享等。但这种文件系统仍不够理想,未能体现用户观点下的数据逻辑结构较大地独立于数据在外存的物理结构。因此,数据物理存储的改变,仍然需要修改用户的应用程序。再者,以文件而不以记录或数据项为单位共享数据,必然导致数据的大量冗余,用户也不能以记录或数据项为单位访问数据。同时,文件系统亦难于增删数据以适应新的应用要求。这些亟待解决的问题,促使人们研究一种新的数据管理技术。

60 年代后期出现的三件大事,标志着数据管理技术进入数据库时代。第一是美国 IBM 公司 1968 年研制成功、1969 年正式推出的商品化的层次数据库管理系统 IMS (Information Management System);第二是美国 CODASYL 组织(Conference On Data System Language)1969 年公布的 DBTG(DataBase Task Group)报告,提出网络数据库系

统;第三是自 1970 年起,IBM 公司 E. F. Codd 发表了一系列论文,奠定了关系数据库的理论基础。初期的数据库方法通常具有如下特点:(1) 对用户观点的数据进行严格细致的描述,使得文件的记录、数据项等数据单位之间的联系清晰,结构简单;(2) 允许用户以记录或数据项为单位进行访问,也允许多关键字检索和文件之间的交叉访问;(3) 数据的物理存储可以很复杂,同样的物理数据可以导出多个不同的逻辑文件,用户以简单的逻辑结构操作数据而无需考虑数据的存储情况、改动数据的物理位置和存储结构、不必修改或重写应用程序,用户逻辑数据与它们的物理存储之间的转换由数据管理软件完成(如图 1.3 所示),从而解决了数据的应用独立于数据的存储的问题。

但初期数据库方法中数据的整体逻辑结构仅是用户逻辑文件的简单并集,在用户越来越多、系统为每个用户提供的逻辑文件日渐庞杂的情况下,数据库的组织也就越来越乱。为了提高效率,减少冗余,增加新的数据,常常需要改变数据的整体逻辑结构,这就必然导致用户逻辑文件的修改,进而导致用户应用程序的修改。特别是对某些系统来说,改变整体逻辑结构已成为系统活动方式,这样就提出了用户的数据逻辑结构尽量不受整体逻辑结构变化的影响问题,促使人们把用户观点的逻辑结构从整体逻辑结构中独立出来,形成今天的数据库方法。

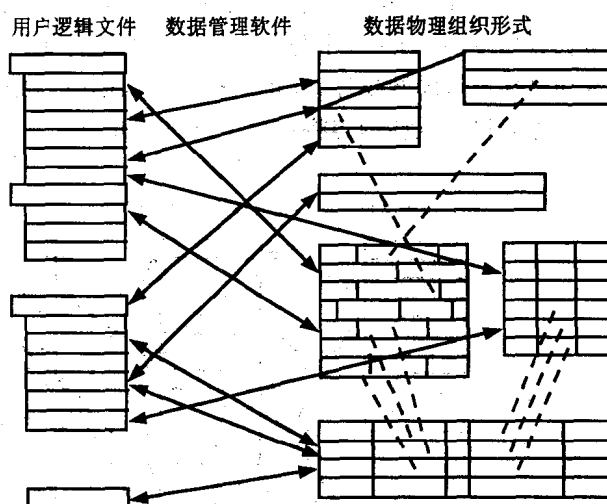


图 1.3 初期的数据库结构

数据库方法的最大特点就是三级结构和两级数据独立性,即在用户数据逻辑结构与数据的物理存储结构之间加入数据的整体逻辑结构,使得数据的物理存储结构的变化尽量不影响数据的整体逻辑结构或用户的应用程序;数据整体逻辑结构的改变也尽量不影响用户应用程序。采用此方法构造的系统的结构如图 1.4 所示。

目前全世界已有成千上万个数据库系统在运行,它们几乎触及到人类社会生活的所有方面,无论是企事业内部的信息管理,还是各行业(银行、交通、邮电、税务、公安等等)的业务处理,以及一般的信息加工和情报检索无不以数据库技术为基础。近年来随着分布处理、高速网络、多媒体、数据压缩等一批高新技术的发展,数据库技术的应用更加普遍深入,分布式数据库系统也得到了迅速的发展,目前几乎所有主要的数据库管理系统都提供了分布数据管理的功能。特别是随着 Internet/Intranet 如火如荼地发展,成百上千的站点都以数据库服务器作为低层支持,使人们能更方便有效地共享各类信息。与此同时,数据库技术也遇到了新应用领域的挑战,风靡整个 80 年代的关系数据库技术由于其数据结构

平板,数据类型和操作简单、固定,只能处理短事务等弊病,不再能有效地支持像计算机辅助设计与辅助加工、计算机辅助软件工程、图像处理、超文本、多媒体等一类新的应用,人们热切呼唤新一代数据库技术。美国高级 DBMS 功能委员会在 1990 年曾提出一个“第三代数据库系统宣言”。目前较普遍的看法是新一代数据库技术应支持面向对象的数据模型,其体系结构应适应功能扩展,能处理复杂的数据类型和长寿事务,同时能与 30 多年来陆续开发的各类数据库共存。为此,自 20 世纪 80 年代中后期以来,数据库工作者在如下两方面开展了卓有成效的工作:一是改造和扩充关系数据库,以适应新的应用需求;二是改用新的数据模型(例如面向对象数据模型、基于逻辑数据模型等)构造数据库,人们期待着一个“后关系数据库”(Postrelational database)时代的到来。此外,新一代数据库技术的另一显著特征是与其它学科的发展高度结合,例如数据库技术与分布处理技术结合导出分布式数据库,数据库技术与并行处理技术结合导出并行数据库,数据库技术与人工智能技术结合导出演绎数据库、知识库和主动数据库,数据库技术与多媒体技术结合导出多媒体数据库等等。这些内容将在第十章——介绍。

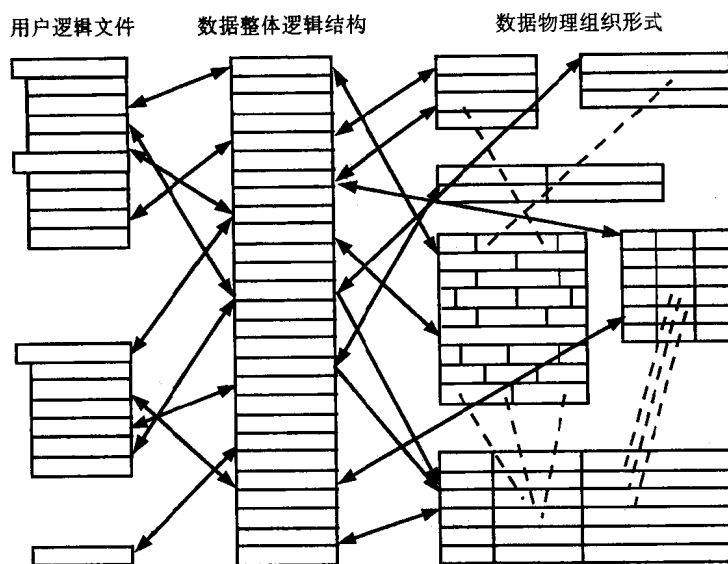


图 1.4 今天的数据库方法

值得一提的是,存储技术的发展为数据库提供了良好的硬件基础,极大推动了数据库技术的进步。1956 年第一台磁盘机(RAMAC 磁盘)问世,单片磁盘容量仅为 5M 字节,到 1978 年扩大到 6×10^8 字节,1997 年达到 2G 字节。近年迅速发展起来的光盘技术,因其容量大、不易磨损、便于保存等诸多特点,很快成为数据库的存储介质,目前许多文献资料数据库(例如大百科全书等)已存储在光盘上,在市面流通。

1.1.2 数据库系统

为了更快地理解数据库方法,不妨将其与图书馆做一对比。

图书馆是负责存储和借阅图书的部门；而数据库系统则是存储数据并负责用户访问数据的机构。正像图书馆不能简单地与书库等同起来一样，也不能把数据库系统仅仅理解为存储数据的汇集，即数据库。就图书馆而言，如果把书籍胡乱地堆放在书库中，几乎无法从数以百万册计的浩瀚书海中查找出读者要借的一本书来。因此，没有一套完整的书卡作为图书馆藏书的模型，查借书籍就很困难，管理员也很难掌握藏书全貌。如果不知道书卡与书架的对应关系，管理员也难以按借书单找到该书的存放位置。一个图书馆要想提高服务质量，必须完成以下工作：

(1) 建立完善的书卡。书卡的内容和格式常包括：书号(分类号)、书名、作者名、出版社名、出版时间、内容摘要和其它细节内容。有时为了方便读者，也按不同分类编排书卡，如以书名、作者名或其它为索引进行编排。

(2) 图书应有组织地存放在书库中。图书馆藏书数量很大，书库中房间、书架很多，需要按照一定的顺序和规则(物理结构)分放图书，并列出各类书籍存放的对应关系表，使管理人员能按此表快速查找。

(3) 规定借阅权限，并建立周密的借阅管理制度。不同类型读者的借阅权限不同。如善本书只供特定的人借阅；机密图书只供有特权的人借阅；某些书只供读者在馆内翻阅等。

(4) 提供图书及各类文献的借阅服务。

读者借书要先出示借书证，图书管理员验明读者身份和借阅权限后，根据读者填写的借书单(访问请求)，按照书籍与书架的对应关系表，到书库中查找所借图书交与读者(响应)，并作某些登记(日志)；还书时管理员要按书号把交还的图书送回原来存放的书架位置上；如果还书较多，管理员又要立即“响应”其他读者借书，那么，他也可以把还书按归还次序放到一个“当日暂存书架”上，待有空闲时，再清理“当日暂存书架”上的还书，并送回书库中存放它们的原来位置上。

对数据库系统来说，也要完成类似于上述图书馆的工作。

(1) 要建立数据模型(见 1.2.3 节)，使用户可以根据数据模型访问数据库中的数据(如检索、插入、删除和修改)，而不必关心数据的物理存储位置，就像读者可以按书卡填写借书单，而不用顾及书籍存放在书库的什么位置一样。当然数据模型也要像书卡那样能反映各种数据之间的内在联系。

(2) 数据也应有组织地存放在存储设备上，并建立数据模型到物理存储位置的对应表(这种对应称为映射)，它使系统能够按照用户的访问请求，找到被访问数据的存储位置。建立数据模型和设计数据的物理存储(组织)方法，其目的是使用户对数据的应用与数据的存放位置和存储结构无关，后者的变动不影响前者(正像改变图书的存放位置不影响读者按书卡借书一样)，这称为数据独立性(见 1.3.3 节)。

(3) 数据库系统要为不同用户确定不同的访问权限并进行访问控制。

(4) 提供以数据库为基础的各种应用服务。

上述各种功能都是在数据库管理系统(DBMS)的统一管理和控制下实现的，正像一个大型公共图书馆需要有专门的工作人员负责规划、设计、协调、维护和管理一样。数据库系统一般设有数据库管理员(DBA)，负责协调和监视数据库的使用，一旦发现违反安全