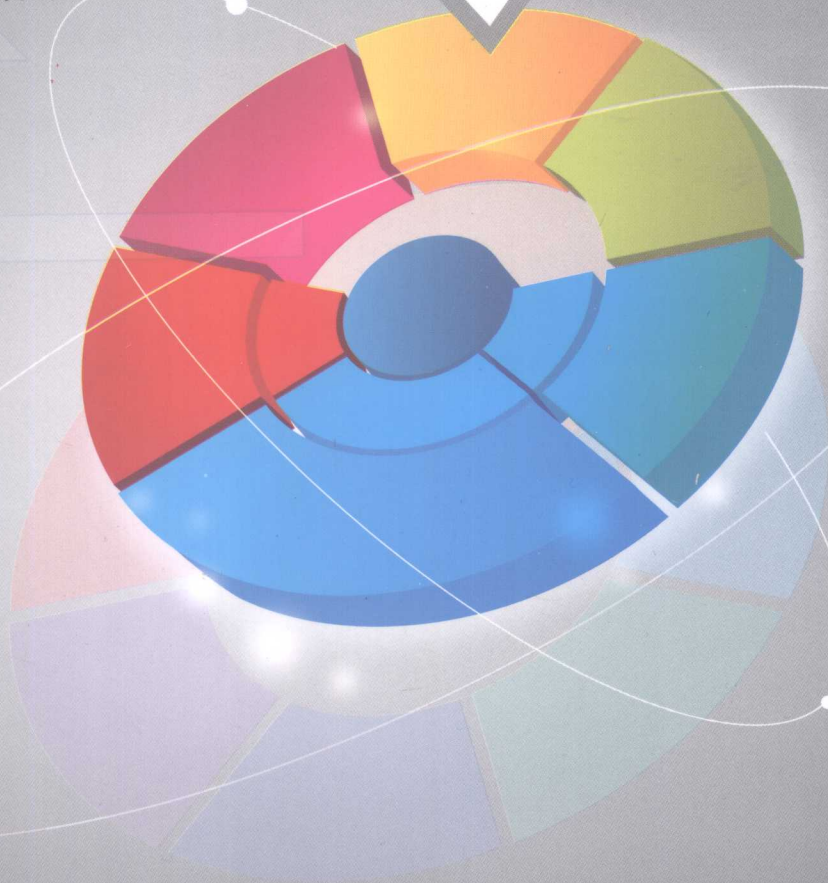


数据库 技术及应用

主编 陶程仁 李 巍 郝宁波



中国商务出版社
CHINA COMMERCE AND TRADE PRESS

数据库 技术及应用

主 编 陶程仁 李 巍 郝宁波
副主编 谌章义 朱献文 张利新 张永霞 徐雪霞



中国商务出版社
CHINA COMMERCE AND TRADE PRESS

图书在版编目(CIP)数据

数据库技术及应用/陶程仁,李巍,郝宁波主编. —北京:中国商务出版社,2009.7

ISBN 978-7-5103-0108-7

I. 数… II. ①陶…②李…③郝… III. 数据库系统
IV. TP311.13

中国版本图书馆 CIP 数据核字(2009)第 126700 号

数据库技术及应用

SHUJUKU JISHU JI YINGYONG

陶程仁 李 巍 郝宁波 主编

湛章义 朱献文 张利新 张永霞 徐雪霞 副主编

中国商务出版社出版

(北京市东城区安定门外大街东后巷 28 号)

邮政编码:100710

电话:010-64269744(编辑室)

010-64283818(发行部)

网址:www.cctpress.com

E-mail:cctp@cctpress.com

北京中商图出版物发行有限责任公司发行

三河市铭浩彩色印装有限公司印刷

787 毫米×1092 毫米 16 开本

24.75 印张 633 千字

2009 年 7 月第 1 版

2009 年 7 月第 1 次印刷

印数:1~1000

ISBN 978-7-5103-0108-7

定价:38.00 元

前言

当今是一个信息爆炸的时代,信息已经成为社会和经济发展的主要支柱之一。大量信息的产生、处理、存储、传播和使用推动了社会的进步和经济的发展。信息系统是一种以加工处理信息为主的计算机系统。数据库技术作为一种存储和使用信息的信息系统核心技术发挥着越来越重要的作用。数据库系统已在当代的社会中获得了广泛的应用,并渗透到工农业生产、商业、行政管理、科学研究、教育、工程技术和国防军事等各行各业,而且围绕数据库技术已形成了一个巨大的软件产业,即数据库管理系统和各类工具软件的开发与经营。

关系型数据库系统是所有数据库系统中发展得最为成熟的一种,其中 Microsoft 公司推出的 SQL Server 数据库管理系统作为大型关系数据库系统中的佼佼者,它建立在成熟而强大的关系模型基础上,具有操作简单、功能齐全、安全可靠等特点,能够满足各种类型的企事业单位对构建数据库的需求。

本书共分 12 章,第 1 章介绍了数据库系统基础知识;第 2 章主要介绍了关系型数据库的理论基础;第 3 章详细论述了数据库中各种对象的建立和维护,结合实例,讲解了对数据库和表的创建、修改、删除等操作;第 4 章 Transact-SQL 编程,详细讲解了常量和变量、标识符、运算符、数据类型、函数以及流程控制语句;第 5 章数据查询,主要介绍了使用 SQL 语句对数据库进行查询和更新操作的基本方法;第 6 章主要介绍了存储过程和触发器技术;第 7 章数据库保护技术,针对数据库的完整性、安全性、账号管理、角色管理、架构管理以及事务处理和并发控制做了详细论述;第 8 章介绍数据库的备份与恢复;第 9 章讲述数据库技术的应用与开发;第 10 章和第 11 章分别介绍了 C/S 体系结构和 B/S 体系结构的应用技术与应用实例;第 12 章则对数据库新技术的未来发展趋势做了一番探讨。

本书的特点是理论与实践相结合,将数据库基本原理、SQL Server 数据库管理操作以及数据库系统应用开发三者有机地结合起来进行讲述,同时本书深入浅出、通俗易懂,因此它不仅适合数据库系统开发的软件人员参考,同时对高校计算机专业的老师进行教学和研究也有很重要的参考价值。

全书由陶程仁、李巍和郝宁波担任主编,由谌章义、朱献文、张利新、张永霞和徐雪霞担任副主编,并由陶程仁、李巍和郝宁波负责统稿。其具体分工如下:

第 10 章,第 11 章,第 12 章:陶程仁(贺州学院);

第 1 章,第 2 章:李巍(兰州职业技术学院);

第 7 章,第 9 章第 1 节与第 2 节:郝宁波(黄淮学院);

第 5 章,第 4 章第 7 节与第 8 节:谌章义(洛阳师范学院);

第 6 章,第 3 章第 1 节:朱献文(黄淮学院);

第 3 章第 2 节~第 5 节:张利新(内蒙古科技大学);

第4章第1节~第6节,第9章第4节~第7节:张永霞(宁夏师范学院数学与计算机科学学院);

第8章,第9章第3节:徐雪霞(武汉体育学院)。

在本书的编写过程中,得到了多方人士的大力支持,并且参考了大量的文献资料,在此谨致谢意。另外,由于时间仓促、编者水平有限,虽经过认真的整理和修改,书中难免有疏漏和不足之处,希望广大读者批评指正,以便我们进一步完善。

编 者
2009.6

目 录

第 1 章 数据库系统概论	(1)
1.1 数据与数据管理	(1)
1.2 数据库系统的组成	(6)
1.3 数据模型.....	(10)
1.4 数据库系统结构.....	(18)
1.5 数据库管理系统.....	(24)
1.6 数据库技术的演变和发展.....	(29)
第 2 章 关系数据库理论基础	(35)
2.1 关系数据库概述.....	(35)
2.2 关系代数.....	(38)
2.3 关系模型的完整性规则.....	(44)
2.4 关系的规范化理论.....	(45)
2.5 数据库的设计方法.....	(51)
第 3 章 数据库对象建立与维护	(55)
3.1 SQL Server 数据库的基础知识	(55)
3.2 数据库的创建.....	(59)
3.3 数据库的修改与删除.....	(73)
3.4 SQL Server 表概述	(76)
3.5 表的创建与维护.....	(86)
第 4 章 Transact-SQL 编程	(96)
4.1 Transact-SQL 概述	(96)
4.2 批处理、脚本和注释	(96)
4.3 常量和变量.....	(98)
4.4 标识符	(101)
4.5 运算符	(104)
4.6 数据类型	(109)
4.7 函数	(121)
4.8 流程控制语句	(130)
第 5 章 数据查询	(138)
5.1 SQL 概述	(138)
5.2 数据查询	(140)
5.3 数据更新	(158)
第 6 章 存储过程和触发器	(162)
6.1 存储过程概述	(162)

6.2	创建存储过程	(164)
6.3	存储过程的管理	(171)
6.4	触发器概述	(173)
6.5	触发器的创建与使用	(180)
6.6	触发器的管理	(188)
第 7 章	数据库保护技术	(191)
7.1	数据库的完整性	(191)
7.2	数据库的安全性	(196)
7.3	账号管理	(202)
7.4	角色管理	(207)
7.5	架构管理	(214)
7.6	事务处理	(216)
7.7	并发控制	(224)
第 8 章	数据库的备份与恢复技术	(234)
8.1	备份与恢复的基本知识	(234)
8.2	SQL Server 2005 备份类型	(241)
8.3	备份策略	(247)
8.4	备份数据	(251)
8.5	数据库系统故障的种类	(255)
8.6	数据恢复技术	(257)
第 9 章	数据库技术应用与开发	(262)
9.1	数据库设计概述	(262)
9.2	需求分析	(267)
9.3	概念结构设计	(273)
9.4	逻辑结构设计	(282)
9.5	物理结构设计	(285)
9.6	数据库的实施	(288)
9.7	数据库的运行和维护	(289)
第 10 章	C/S 体系结构的应用技术与应用实例	(291)
10.1	C/S 体系结构概述	(291)
10.2	C/S 体系结构的数据库应用系统开发	(308)
第 11 章	B/S 体系结构的应用技术与应用实例	(323)
11.1	B/S 体系结构概述	(323)
11.2	B/S 体系结构的数据库应用系统开发	(326)
第 12 章	数据库新技术的未来发展趋势	(355)
12.1	现代数据库概述	(355)
12.2	数据库技术与其他技术相结合产生的新领域	(358)
12.3	面向应用领域的数据库新技术	(385)
参考文献		(390)

第 1 章 数据库系统概论

数据库技术是计算机领域中最活跃的技术领域之一,是计算机科学的重要分支,它的出现极大地推动了各行各业对计算机的应用,尤其是进入 20 世纪 90 年代以来,随着计算机网络技术的深入发展和广泛应用,数据库技术进入了一个崭新阶段。数据库技术的应用对许多大中型企业、事业单位提高科学管理水平,起到了举足轻重的作用。

数据库是许多数据的集合,DBMS 是管理这些数据集合的计算机软件系统,简称为数据库系统。数据库系统主要用于管理大量数据、控制多用户访问、定义数据库构架以及执行数据库操作等。数据库系统随着信息技术的发展而发展。信息技术的发展基础是计算机硬件性能和软件工具的共同发展以及信息需求的不断增长,同样,数据库技术也是随着硬件技术和软件技术的发展以及应用范围的拓展而演变。从当前市场上流行的数据库产品来看,有些趋势已经非常明显,例如市场上已经出现了作为关系数据库系统向面向对象数据库过渡的对象关系数据库、具有网络计算功能的分布式数据库和数据仓库等。学习完本章,读者应该对数据库系统有一个初步了解,为后面的学习打下一个良好的基础。

1.1 数据与数据管理

数据库是计算机信息管理的基础,其研究对象是数据。因此,在介绍数据库技术之前,有必要了解数据与信息的基本概念和数据管理技术的发展历程。

1.1.1 数据与数据处理

1. 数据与信息

数据和信息是数据处理中的两个基本概念,有时可以混用,如平时讲数据处理就是信息处理,但有时必须分清。一般认为,数据是人们用于记录事物情况的物理符号。为了描述客观事物而用到的数字、字符以及所有能输入到计算机中并能被计算机处理的符号都可以看做是数据。例如,谭浩强的基本工资为 1350 元,职称为教授,这里的“谭浩强”、“1350”、“教授”就是数据。在实际应用中,有两种基本形式的数据,一种是可以参与数值运算的数值型数据,如表示成绩、工资的数据;另一种是由字符组成、不能参与数值运算的字符型数据,如表示姓名、职称的数据。此外,还有图形、图像、声音等多媒体数据,如人的照片、商品的商标等。

信息是数据中所包含的意义。通俗地讲,信息是经过加工处理并对人类社会实践和生产活动产生决策影响的数据。不经过加工处理的数据只是一种原始材料,对人类活动产生不了决策作用,它的价值只是在于记录了客观世界的事实。只有经过提炼和加工,原始数据才发生了质的变化,给人们以新的知识和智慧。

数据与信息既有区别,又有联系。数据是表示信息的,但并非任何数据都能表示信息,信息

只是加工处理后的数据,是数据所表达的内容。另一方面信息不随表示它的数据形式而改变,它是反映客观现实世界的知识,而数据则具有任意性,用不同的数据形式可以表示同样的信息。例如一个城市的天气预报情况是一条信息,而描述该信息的数据形式可以是文字、图像或声音等。

2. 数据处理

人类社会的一切活动都离不开对信息的处理,由于信息世界中的绝大部分信息都可以用数据来表示,因此信息处理实际上也可以说是数据处理。从数据处理技术发展的历史来看,大致分为手工处理、机械处理和电子处理三个阶段。

从原始社会到 19 世纪末,由于社会生产力和科学技术的限制,数据处理处于低级的手工处理阶段。这一阶段主要使用的计算工具有算盘、基于齿轮结构的六位加法器、计算尺和微分机等。该阶段的特点是计算工具低级、精确度差、处理能力低,并且离不开手工操作。

1890 年,美国中央统计局的 H. Hollorith 为了编制人口统计表的需要,发明了卡片制表机,使数据处理跨入了机械处理阶段。此卡片制表机能以半自动方式进行卡片的穿孔、校验、分类、整理和制表等工作。与手工阶段相比,一部分手工操作由机械所代替,因此在数据处理的能力和效率上都有很大提高,但其效率受到机械设备性能的限制。

1946 年第一台电子计算机 ENIAC 的诞生,标志着一个崭新的数据处理阶段——电子处理阶段的开始。电子计算机以其自动、快速的处理,大容量的数据存储,灵活的输入/输出,彻底改变了以往数据处理效率低、手工操作多、可靠性差、与当时社会生产力的发展不相适应的落后状态。随着计算机硬件和软件的发展,特别是高速中央处理器和大容量磁盘存储器的应用,使计算机不仅能进行各类数值计算,还能进行文字和图像等各种数据处理,使数据处理技术获得突飞猛进的发展。

1.1.2 数据管理的三个阶段

人类社会错综复杂,其中存在着各种各样的事物、事物的属性、事物之间的联系以及人们对各种事物的管理活动。所有的管理活动都离不开信息,农民从事农业生产需要土质、产量、肥料价格、气候情况等信息;工厂制订生产计划需要市场需求、材料来源、生产能力等有关信息;在日常生活中需要天气、物价、亲友的通信地址及电话号码、日程安排等信息。总之,人们的一切活动时刻都和大量的各种信息打交道,所有的这些信息构成了信息世界。

数据管理是指对数据的组织、存储、检索和维护等工作,它是数据处理的核心。随着计算机软件与硬件的发展,数据管理技术也经历了三个阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前,计算机数据管理的能力很差,这一阶段称为人工管理阶段。那时计算机没有磁盘等直接存储设备,没有操作系统,没有管理数据的软件,数据处理方式是批处理,数据管理的任务主要由应用程序员自己承担,系统所提供的数据管理功能仅仅是一些简单的 I/O 操作,如图 1-1 所示。

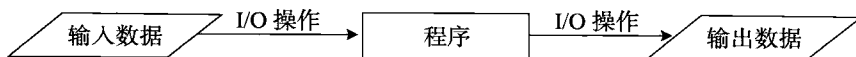


图 1-1 程序中数据的输入与输出

在人工管理阶段,不同的应用程序处理不同的数据,数据与程序之间是一一对应的关系,如图 1-2 所示。

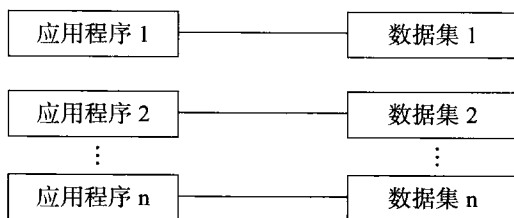


图 1-2 人工管理阶段数据与程序之间的对应关系

人工管理数据存在以下缺点:

(1) 数据不保存

当时的计算机主要用于科学计算,一般不需要将数据长期保存,只是在计算某一具体题目时将数据输入,运行完后得到输出结果,输入、输出和中间数据都不保存。这与信息系统中对数据的管理思想是完全不一样的,因为在信息系统中,数据作为一个重要资源,不仅要对它进行加工处理,还要予以保存,以便以后检索和修改。

(2) 数据不共享

在人工管理阶段,数据是面向应用的。一组数据只对应一个应用程序,当多个应用程序要用到相同数据时,必须各自定义,不能共享。

(3) 数据冗余度大

由于数据不能共享,必然会出现相同数据的多个副本,不同的副本对应不同的应用程序。这样会导致数据冗余度大。

(4) 数据缺乏独立性

数据与程序是紧密结合在一起的,数据的逻辑结构、物理结构、存取方式都由程序规定,当数据的逻辑结构、物理结构、存取方式发生变化时,必须对应用程序作相应修改,因此程序员的负担很重。

(5) 数据的不统一性

在一个组织中,要想将数据作为一种资源共享,必须对数据的命名、格式、存取方式等标准进行统一规定。但是在人工管理阶段,数据与程序紧密结合,不同应用程序会对同一数据做不同的定义,因而往往会出现“同名异物”和“同物异名”的现象。

2. 文件系统阶段

20 世纪 50 年代到 60 年代中期,计算机数据管理技术进入文件系统阶段。这时已经有了磁盘、磁鼓等直接存储的硬件设备,也有了专门的数据管理软件,一般称为文件系统;处理方式上不仅有了批处理,而且还有了能够实现联机实时处理的处理方式。

在文件系统阶段,文件系统把数据组织成相互独立的数据文件,数据可以长期保存在存储设备上,应用程序利用“按文件名访问,按记录进行存取”的管理技术,可以对文件中的数据进行修改、插入和删除等操作。文件系统实现了一定的数据独立性,它将数据的逻辑结构与物理结构分离,应用“存取方式”实现逻辑结构与物理结构之间的映射,如图 1-3 所示。

在这种方式下,应用程序只涉及数据的逻辑结构,当数据的物理结构改变时,不会导致应用程序的修改,这就是数据的物理独立性。数据的物理独立性使应用程序脱离了数据的物理结构,

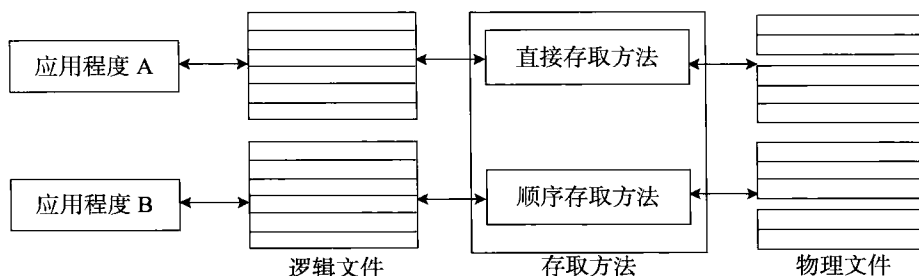


图 1-3 文件系统的数据处理

也使其适应性得以提高。同时,程序员在编写程序和对程序进行维护时不必过多考虑数据的物理存储细节,只需将精力集中在算法上,因而生产效率得以大大提高。另外,数据的物理独立性使数据共享成为可能,多个应用程序可以对同一文件进行操作。

与早期的人工管理阶段相比,文件系统已经有了很大的进步。但文件系统管理数据还存在以下不足:

(1)数据的共享性较差

文件系统提供了数据的物理独立性,实现了一定的数据共享,但它只能实现文件级共享而不能在记录或数据项级实现数据共享。

(2)数据的冗余度较大

件系统中,文件的逻辑结构是根据它的应用而设计的,数据的逻辑结构与应用程序之间相互依赖。即使不同应用程序具有部分相同的数据(记录或数据项),也必须构造各自的文件,这样就存在大量的数据冗余,浪费大量的存储空间。

(3)数据存在不一致性

数据的冗余度大与数据的不一致性是密切相关的。同一数据在多个地方同时存放,同一数据在不同存放地的值可能不相同,这会降低信息的价值,有时甚至会造成重大损失。

(4)数据的独立性较差

文件系统只实现了数据的物理独立,而没有实现数据的逻辑独立。数据的逻辑结构对应一个特定的应用,当应用发生变化时,数据的逻辑结构也要发生改变,当数据的逻辑结构发生变化时,程序也要做相应的修改。因此,文件系统的数据与程序之间缺乏逻辑独立性。

为了说明上面的问题,现在来看如下一段 C 语言程序。

程序 A:

```
#include "stdio.h"
main()
{
    FILE *fp;
    fp=fopen("Employee.c","w");
    putw(200005,fp);
    puts("tanlin",fp);
    putw(20,fp);
}
```

```

    putw(200009,fp);
    fputs("xunbin",fp);
    putw(21,fp);
    fclose(fp);
}

```

程序 B:

```

#define SIZE 2
struct Employee_type
{
    int num;
    char name[8];
    int age;
}Employee[SIZE];
main()
{
    int i;
    FILE *fp;
    fp=fopen("Employee. c", " r");
    for(i=0;i<SIZE;i++)
    {
        fread(&Employee[i],sizeof(struct Employee_type),1,fp);
        printf("%8d%8s%4d\n",
            Employee[i]. num,Employee[i]. name,Employee[i]. age);
    }
}

```

这是对文件进行操作的一段程序,文件中保存员工记录。程序 A 将两个员工的员工号、姓名、年龄输入到一个文件中,程序 B 从文件中读取记录,存入变量中。程序 B 中结构型 Employee 的设计必须与文件的逻辑结构一致,即程序设计与文件逻辑结构是紧密相关的。

可以看到,程序 B 对文件操作时必须知道文件的逻辑结构,或者说,若文件的逻辑结构发生了变化,就必须修改程序。而在关系数据库系统中,只需用一个 SQL(结构化查询语言)语句即可完成记录的插入和读取,只要不违反数据库完整性和安全性约束,写 SQL 语句完全无需考虑数据文件的逻辑结构,而数据库完整性和安全性控制是由数据库管理系统自动实现的。

3. 数据库系统阶段

20 世纪 60 年代后期以来,计算机数据管理技术开始进入数据库系统阶段。这时计算机技术发展迅速,硬件方面有了大容量磁盘,硬件价格下降;软件方面出现了包括操作系统在内的大量的系统软件;在处理方式上,联机实时处理要求增多,并开始提出和考虑分布式处理方法。数据库系统阶段开始的标志是产生了一种称为数据库管理系统的专门用于数据管理的软件。

数据库系统的产生是企业海量数据处理需求的必然结果。随着计算机用于管理的规模越来

越大,应用越来越广泛,数据量急剧增大,企业对数据管理技术的要求也越来越高。首先,企业要求数据作为企业组织的公共资源而集中管理控制,为企业的各种用户共享,因此,应大量地消去冗余数据,节省存储空间。其次,当数据变更时,能减少对多个数据副本的多次变更操作,从而可大大节省计算时间,更为重要的是,不会因遗漏某些副本的变更而使系统给出一些不一致的数据。再次,要求数据具有更高的独立性,不但要具有物理独立性,而且还要具有逻辑独立性,即当数据逻辑结构改变时,不影响用户的应用程序,从而降低应用程序开发和维护的成本。所有这些,是文件系统所不能满足的,而数据库管理系统却可以做到。

与文件系统不同,数据库系统是面向数据的而不是面向程序的,数据是系统的中心,各处理功能处于外围,它们通过数据管理软件从数据库中获取所需数据和存储处理结果。按数据库的方法,数据处理的过程如图 1-4 所示。

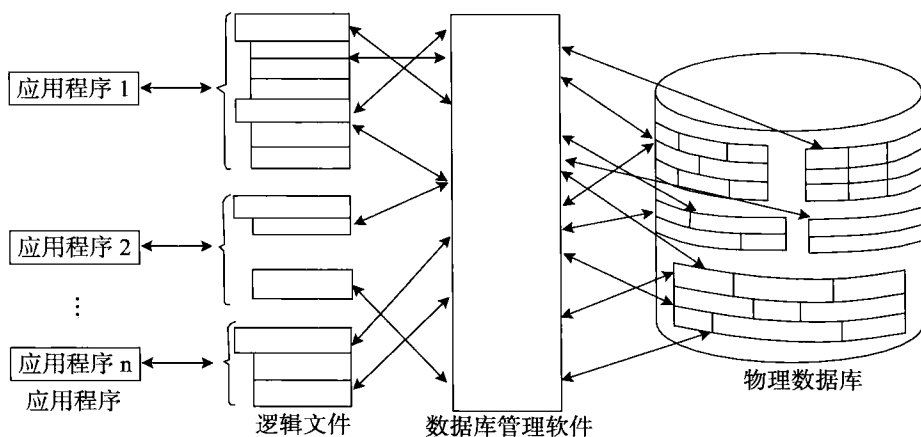


图 1-4 数据库系统的数据处理

1.2 数据库系统的组成

数据库系统(Database System,DBS)由 4 个组成部分,即数据库(Database,DB)、数据库管理系统(Database Management System,DBMS)、数据库应用和人员,如图 1-5 所示。其中数据

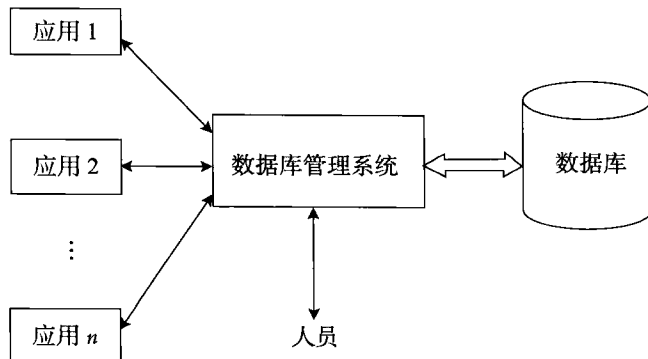


图 1-5 数据库系统的组成

库用于存储数据;数据库管理系统是用于操纵数据库的系统软件;数据库应用是为满足用户各种需求而设计的程序,如报表、查询等;数据库管理员是专门从事数据库管理的人员。下面详细介绍数据库系统的每一个组成部分。

1.2.1 数据库

数据库是存贮在一起的相关数据的集合,这些数据无有害的或不必要的冗余,为多种应用服务;数据的存贮独立于使用它的应用程序,对数据库插入新数据,修改和检索原有数据均能按一种公用的和可控制的方法进行,数据被结构化。

存储在系统中的数据被分成一或多个数据库,但在后面的讨论中我们假定只一个数据库,这并不影响讨论。

一个数据库可以定义为满足一个或多个组织、多个用户的信息要求的集成数据存储库。此定义包含了一个数据库的要素:集成、共享、存储、信息。

数据库的存储指的是它存储于计算机系统中。一个数据库包含的数据量很大,一般用数以兆字节计,所以只能存于外存设备(主要是磁盘)上。数据库在运行过程中还将不断扩大,故要求的外存储器容量较大。另一方面,数据库少包含的数据类型多,数据间的联系复杂,因而有一个数据在存储设备上如何组织及如何有效存取的问题。同时,还要根据用户的地理分布要求确定数据应是存储于单个计算机系统中还是分散于多个计算机系统中。

数据库通常由两大部分组成:

(1)有关应用所需要的工作数据的集合,称为物理数据库,它是数据库的主体。

(2)关于各级数据结构的描述数据,称做描述数据库(数据字典)。

例如,有销货合同库,它包含下列信息:

合同号	订货日期	品名	规格	型号	数量	单价	金额	交货日期	订货单位	到站	联系人	电话
-----	------	----	----	----	----	----	----	------	------	----	-----	----

作为生产部门,要关心的内容为 3~6 项和 9 项。作为仓库部门,要关心的内容 3~6 项。作为财会部门,要关心的内容为 3~8 项。而销售部门,则关心全部内容。

手工管理通常是一式多份,冗余度大,用计算机管理则只有一份就可共享,这种数据是工作数据。类似于下面的信息库:

库 名	用户标记	所属模式	口 令	对应物理设备
销货合同库	A,B,C	...	...	...

它主要用于描述数据,供 DBMS 使用,或供数据库管理员用。

1.2.2 数据库管理系统

数据库管理系统是一个很复杂的系统软件。它很像一个操作命令语言解释器,把用户程序的数据操作语句转换为对系统存贮文件的操作;它又像一个向导,把用户对数据库的一次访问,从用户级带到概念级,再导向物理级,有效地实现数据库三级之间的转换。概括而言,数据库管理系统担负着对数据库中的数据资源进行统一管理的任务,并且负责执行用户给数据库系统发布的各种请求命令。

数据库管理系统的功能随系统的大小而异。大型系统功能较强较全,小型系统功能较弱。

一般说来,它包括以下几个方面:

1. 数据库的描述

它们数据描述语言所描述的全局和局部的逻辑数据结构、存贮结构、保密定义以及信息格式等各项内容从源形式转换成目标形式存放在数据库中供系统查阅(也称为数据库定义功能或数据字典)。

2. 数据库的管理

它控制整个数据库系统的运行;控制用户的并发性访问,执行对数据的安全、保密和完整性检验;实施对数据检索、插入、删除、修改等操作(也称为对数据库的控制功能)。

3. 数据库的建立和维护

其功能包括:初始时装入数据库;运行时记录工作日志、监视数据库性能;在性能变坏时重装组织数据库;在用户要求或系统设备变化时修改和更新数据库;在系统软、硬件发生故障时恢复数据库等。

4. 数据通信

通常与操作系统协同处理数据的流向。这些数据可能来自应用程序、计算机终端或其他系统,也可能有系统内运行的进程所产生。它们可能被送到调用队列缓冲区、终端或正在执行的某个进程中。

1.2.3 人员

开发、管理和使用数据库系统的人员主要是:数据库管理员、系统分析员和数据库设计人员、应用程序员和最终用户。不同的人员涉及不同的数据抽象级别,具有不同的数据视图,如图 1-6 所示。

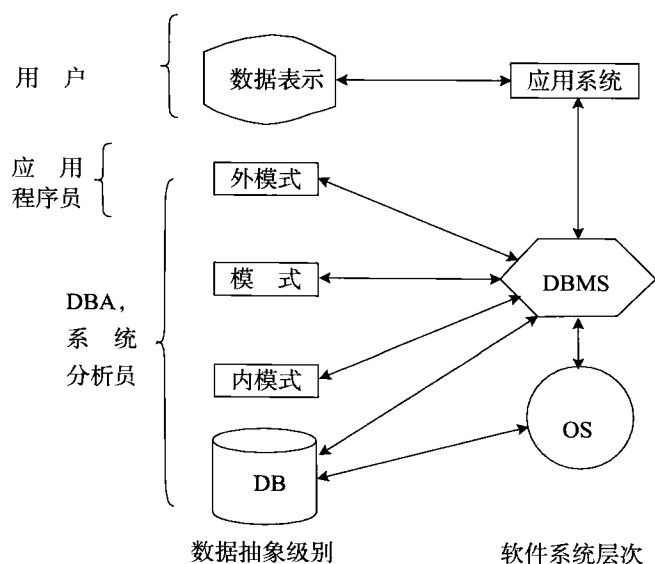


图 1-6 各种人员的数据视图

其各种人员的职责分别如下：

1. 数据库管理员 (DataBase Administrator, 简称 DBA)

在数据库系统环境下,有两类共享资源。一类是数据库,另一类是数据库管理系统软件。因此需要有专门的管理机构来监督和管理数据库系统。DBA 则是这个机构的一个(组)人员,负责全面管理和控制数据库系统。具体职责包括：

(1) 决定数据库中的信息内容和结构

数据库中要存放哪些信息,DBA 要参与决策。因此,DBA 必须参加数据库设计的全过程,并与用户、应用程序员、系统分析员密切合作共同协商,搞好数据库设计。

(2) 决定数据库的存储结构和存取策略

DBA 要综合各用户的应用要求,和数据库设计人员共同决定数据的存储结构和存取策略以求获得较高的存取效率和存储空间利用率。

(3) 定义数据的安全性要求和完整性约束条件

DBA 的重要职责是保证数据库的安全性和完整性。因此 DBA 负责确定各个用户对数据库的存取权限、数据的保密级别和完整性约束条件。

(4) 监控数据库的使用和运行

DBA 还有一个重要职责就是监视数据库系统的运行情况,及时处理运行过程中出现的问题。比如系统发生各种故障时,数据库会因此遭到不同程度的破坏,DBA 必须在最短时间内将数据库恢复到正确状态,并尽可能不影响或少影响计算机系统其他部分的正常运行。为此,DBA 要定义和实施适当的后备和恢复策略。如周期性的转储数据、维护日志文件等。有关这方面的内容将在下面做进步讨论。

(5) 数据库的改进和重组重构

DBA 还负责在系统运行期间监视系统的空间利用率、处理效率等性能指标,对运行情况进行记录、统计分析,依靠工作实践并根据实际应用环境,不断改进数据库设计。不少数据库产品都提供了对数据库运行状况进行监视和分析的实用程序,DBA 可以使用这些实用程序完成这项工作。

另外,在数据运行过程中,大量数据不断插入、删除、修改,时间一长,会影响系统的性能。因此,DBA 要定期对数据库进行重组,以提高系统的性能。

当用户的需求增加和改变时,DBA 还要对数据库进行较大的改造,包括修改部分设计,即数据库的重构造。

2. 系统分析员和数据库设计人员

系统分析员负责应用系统的需求分析和规范说明,要和用户及 DBA 相结合,确定系统的硬件软件配置,并参与数据库系统的概要设计。

数据库设计人员负责数据库中数据的确定、数据库各级模式的设计。数据库设计人员必须参加用户需求调查和系统分析,然后进行数据库设计。在很多情况下,数据库设计人员就由数据库管理员担任。

3. 应用程序员

应用程序员负责设计和编写应用系统的程序模块,并进行调试和安装。

4. 用户

这里用户是指最终用户(End user)。最终用户通过应用系统的用户接口使用数据库。常用的接口方式有浏览器、菜单驱动、表格操作、图形显示、报表书写等,给用户提供了简明直观的数据表示。

最终用户可以分为如下三类:

(1)偶然用户。这类用户不经常访问数据库,但每次访问数据库时往往需要不同的数据库信息,这类用户一般是企业或组织机构的高中级管理人员。

(2)简单用户。数据库的多数最终用户都是简单用户。其主要工作是查询和修改数据库,一般都是通过应用程序员精心设计并具有友好界面的应用程序存取数据库。银行的职员、航空公司的机票预定工作人员、旅馆总台服务员等都属于这类用户。

(3)复杂用户。复杂用户包括工程师、科学家、经济学家、科学技术工作者等具有较高科学技术背景的人员。这类用户一般都比较熟悉数据库管理系统的各种功能,能够直接使用数据库语言访问数据库,甚至能够基于数据库管理系统的 API 编制自己的应用程序。

1.3 数据模型

模型是现实世界特征的模拟和抽象。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。在实际数据库系统中,人们(主要是数据库管理员 DBA)用数据库管理系统 DBMS 提供的数据库描述语言 DDL,按照该 DBMS 支持的某种数据模型的要求来描述系统的概念模式。

研究构造一个数据模型是不容易的,这是因为一个数据模型必须要有足够丰富的构件,以便能描述现实世界中各种有意义的方面,而且还应使实现概念模式的有效方法成为可能。这样,在数据库技术中,数据模型也历经了一个从低级到高级、从对简单对象的建模到复杂对象的建模,以及从具体到抽象的发展过程。

20 世纪 60 年代末 70 年代初在文件系统后相继出现了层次数据模型、网状数据模型和关系数据模型,即现在常称的三大经典数据模型。层次数据模型采用简单的树(或森林)结构来表达实体及实体间的联系;网状数据模型用网状结构来表达实体及实体间的联系;关系数据模型则以关系的集合论为基础用统一的关系(二维表格数据)来表达实体及实体间的联系,具有较高的数据独立性和较严密的数学理论基础。这三大经典数据模型是目前应用得最广泛的模型,现行大部分实用 DBMS 均是基于它们的。三大经典模型的主要特点是它们都基本上面向计算机世界,即它们易于有效地实现数据的存储和处理。但其追求的目标不是建模对象的语义,它们有限的表达能力也远不能自然地按直接方式来描述和模拟现实世界中的复杂应用,因而它们基本上不是面向现实世界的。下面简要介绍层次模型、网状模型和关系模型。

1.3.1 层次模型

层次模型是数据库系统中最早出现的数据模型,层次数据库系统采用层次模型作为数据的组织方式。层次数据库系统的典型代表是 IBM 公司的 IMS(Information Management System)