

高等院校信息技术规划教材

数据库原理及应用

——SQL Server 2014

夏保芹 刘春林 徐小平 编著

课外借



清华大学出版社

高等院校信息技术规划教材

数据库原理及应用

——SQL Server 2014

夏保芹 刘春林 徐小平 编著

清华大学出版社
北京

内 容 简 介

本书全面介绍数据库的原理及应用,内容涵盖数据库应用开发的相关技术,既注重数据库基本概念、基本理论的讲解,又注重基本技术的讲解和相关应用。全书共分12章:第1~3章是数据库原理部分,主要介绍数据库的基本概念和基本理论,包括数据库系统的组成、数据库系统的三级模式结构和两级映像、数据模型、关系代数、数据库设计步骤等;第4~7章是数据库应用基础部分,主要介绍SQL Server数据库的基本应用,包括数据库和表对象的基本操作、表中数据的操作、T-SQL基本语法等;第8~11章是数据库应用高级部分,主要包括视图、索引、游标、存储过程、触发器、数据库备份与还原、数据导入与导出、事务处理、数据库安全管理等;第12章是数据库应用案例部分,通过两个综合案例介绍基于Java的数据库应用系统开发过程。

本书实例丰富,图文并茂,既重视理论知识的讲解,又重视实际的操作应用,“实践”和“案例”理念贯穿始终。书中提供了大量应用实例,每章后面都附有习题;同时为了突出实践应用性,书中最后给出了两个综合应用案例。

本书适合作为高等院校计算机、信息管理、软件工程等相关专业数据库类课程的教材,也适合作为从事数据库系统研究、数据库管理和数据库系统开发者的参考用书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据库原理及应用:SQL Server 2014/夏保芹,刘春林,徐小平编著. —北京:清华大学出版社,2018

(高等院校信息技术规划教材)

ISBN 978-7-302-48999-3

I. ①数… II. ①夏… ②刘… ③徐… III. ①关系数据库系统—高等学校—教材 IV. ①TP311.138

中国版本图书馆CIP数据核字(2017)第293573号

责任编辑:白立军 张爱华

封面设计:常雪影

责任校对:白 蕾

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座

邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印装者:北京嘉实印刷有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:21.25

字 数:493千字

版 次:2018年1月第1版

印 次:2018年1月第1次印刷

印 数:1~1500

定 价:49.00元

产品编号:070257-01

前言

foreword

数据库技术已成为计算机科学技术发展最快的领域之一，它的应用范围遍及各行各业，成为计算机信息系统与应用系统的核心技术和重要基础，也是高等院校计算机相关专业的核心课程。本书编者长期从事本科数据库类课程教学工作，不仅具有丰富的教学经验，同时还具有多年的数据库开发经验。根据长期的教学经验，深知数据库课程的重点和难点，了解学生学习该课程时遇到的主要困难及问题，总结出如何组织教材内容使之更有利于教师教学和学生自学，从而形成本书的结构体系。

本书全面介绍数据库的原理及应用，内容涵盖数据库应用开发的相关技术，既注重数据库基本概念、基本理论的讲解，又注重基本技术的讲解和应用。全书共分 12 章，各章主要内容如下：

第 1 章为数据库系统概述，主要介绍数据管理技术的发展史、数据库系统的组成、数据库系统的三级模式和两级映像、数据模型等。

第 2 章为关系数据库，主要介绍关系的概念、关系的完整性、关系代数和关系规范化理论。

第 3 章为数据库设计，主要介绍数据库设计各个阶段的任务、方法和步骤，通过实例详细介绍数据库设计的具体方法和步骤。

第 4 章为 SQL Server 2014 概述，主要介绍 SQL Server 的发展史、SQL Server 2014 安装过程，并简要介绍常用的管理工具。

第 5 章为数据库和表，主要介绍数据库和表的概念，数据库的基本操作和表的基本操作，包括数据库的创建、修改和删除等操作，表的约束的创建和删除，以及表中数据的基本操作，包括表中数据的增加、删除和修改，最后通过实例介绍数据库文件的分离和附加方法。

第 6 章为数据库查询，主要介绍数据库查询语句的基本语法，通过大量例题详细介绍 SELECT 查询语句的用法，包括单表查询、连接查询、子查询和集合运算查询，以及在数据操作中使用查询。

第 7 章为 T-SQL 编程，主要介绍 T-SQL 语言基本语法，主要内容包括标识符、数据类型、注释、常量和变量、运算符和表达式、批处理、流程控制语句和函数等。

数据库技术已成为计算机科学技术发展最快的领域之一，它的应用范围遍及各行各业，成为计算机信息系统与应用系统的核心技术和重要基础，也是高等院校计算机相关专业的核心课程。

本书编者长期从事本科数据库类课程教学工作，不仅具有丰富的教学经验，同时还具有多年的数据库开发经验。根据长期的教学经验，深知数据库课程的重点和难点，了解学生学习该课程时遇到的主要困难及问题，总结出如何组织教材内容使之更有利于教师教学和学生自学，从而形成本书的结构体系。

本书全面介绍数据库的原理及应用，内容涵盖数据库应用开发的相关技术，既注重数据库基本概念、基本理论的讲解，又注重基本技术的讲解和应用。全书共分 12 章，各章主要内容如下：

第 1 章为数据库系统概述，主要介绍数据管理技术的发展史、数据库系统的组成、数据库系统的三级模式和两级映像、数据模型等。

第 2 章为关系数据库，主要介绍关系的概念、关系的完整性、关系代数和关系规范化理论。

第 3 章为数据库设计，主要介绍数据库设计各个阶段的任务、方法和步骤，通过实例详细介绍数据库设计的具体方法和步骤。

第 4 章为 SQL Server 2014 概述，主要介绍 SQL Server 的发展史、SQL Server 2014 安装过程，并简要介绍常用的管理工具。

第 5 章为数据库和表，主要介绍数据库和表的概念，数据库的基本操作和表的基本操作，包括数据库的创建、修改和删除等操作，表的约束的创建和删除，以及表中数据的基本操作，包括表中数据的增加、删除和修改，最后通过实例介绍数据库文件的分离和附加方法。

第 6 章为数据库查询，主要介绍数据库查询语句的基本语法，通过大量例题详细介绍 SELECT 查询语句的用法，包括单表查询、连接查询、子查询和集合运算查询，以及在数据操作中使用查询。

第 7 章为 T-SQL 编程，主要介绍 T-SQL 语言基本语法，主要内容包括标识符、数据类型、注释、常量和变量、运算符和表达式、批处理、流程控制语句和函数等。

第8章为视图、索引和游标,主要介绍视图、索引和游标的基本概念,以及相应的创建、使用方法。

第9章为存储过程和触发器,主要介绍存储过程和触发器的基本概念、使用它们的优点以及创建和使用方法。

第10章为数据库管理,通过实例介绍数据库的备份与还原、数据库中数据的导入与导出,最后介绍事务以及事务处理的概念。

第11章为数据库的安全管理,主要介绍数据库身份验证模式、登录账号和用户账号的作用和使用、角色的概念和使用、用户权限管理等。

第12章为基于Java的数据库应用系统开发,主要介绍JDBC技术常用的类和接口、JDBC数据库访问的步骤,通过两个案例详细介绍基于Java的数据库应用系统开发步骤。

学习本书,可以为后续课程,如Web开发技术、SSH框架技术、信息系统开发等打下良好的基础。本书重视对学生工程实践能力的培养和训练,“实践”和“案例”理念贯穿始终,将理论知识的学习融于小项目实训中,将知识的学习渗透于每个小项目中。为方便任课教师备课,本书还提供丰富的教学资源,如教学课件、课后习题答案、实验和案例的源代码等。

本书适合作为高等院校计算机、信息管理、软件工程等相关专业数据库类课程的教材,也适合作为从事数据库系统研究、数据库管理和数据库系统开发者的参考用书。

本书在编写过程中,得到很多老师的帮助,除本书列出的参考文献外,还参考了一些博客、论坛等,在此一并表示感谢。

本书由夏保芹、刘春林、徐小平编著。其中,第1、4、6、8、9、11章由夏保芹编写,第2、5、7、10章由刘春林编写,第3、12章由徐小平编写。全书由夏保芹统稿。

由于作者水平有限,书中难免有遗漏之处,恳请业界同仁及读者朋友提出宝贵意见,编者将不胜感激,并在修订时进一步完善。

编者

2017年10月

目录

Contents

第 1 章 数据库系统概述	1
1.1 数据库管理技术发展史	1
1.1.1 数据处理技术	1
1.1.2 数据管理技术的发展	2
1.2 数据库系统的介绍	5
1.2.1 数据库系统的组成	5
1.2.2 数据库系统的三级模式结构	7
1.2.3 数据库系统的两级映像	8
1.3 数据模型	9
1.3.1 信息的三种世界	9
1.3.2 概念模型	10
1.3.3 常见的数据模型	15
1.4 本章小结	22
习题 1	22
第 2 章 关系数据库	25
2.1 关系的概念	25
2.1.1 笛卡儿积	25
2.1.2 关系的数学定义	27
2.1.3 关系模式	27
2.1.4 关系的性质	28
2.2 关系的完整性	28
2.2.1 实体完整性	28
2.2.2 参照完整性	29
2.2.3 用户自定义完整性	30
2.3 关系代数	30
2.3.1 传统的集合运算	30

2.3.2 专门的关系运算	32
2.4 关系的规范化	36
2.4.1 函数依赖理论	36
2.4.2 范式及关系规范化	38
2.5 本章小结	41
习题 2	42
第 3 章 数据库设计	46
3.1 数据库设计概述	46
3.2 需求分析	46
3.3 概念结构设计	47
3.4 逻辑结构设计	47
3.5 物理结构设计	49
3.6 数据库实施、运行和维护	50
3.7 数据库设计实例	50
3.8 本章小结	54
习题 3	54
第 4 章 SQL Server 2014 概述	55
4.1 SQL Server 2014 简介	55
4.1.1 SQL Server 发展史	55
4.1.2 SQL Server 2014 版本	56
4.1.3 SQL Server 2014 新增性能	57
4.2 SQL Server 2014 安装	61
4.2.1 SQL Server 2014 运行环境	61
4.2.2 SQL Server 2014 安装过程	62
4.3 SQL Server 2014 管理工具介绍	72
4.3.1 SQL Server Management Studio	72
4.3.2 SQL Server 配置管理器	74
4.4 本章小结	76
习题 4	76
第 5 章 数据库和表	77
5.1 数据库概述	77
5.1.1 系统数据库	77
5.1.2 数据文件和文件组	78
5.2 数据库的操作	79

5.3 表的操作	86
5.3.1 表操作	87
5.3.2 约束	91
5.3.3 表中数据操作	99
5.3.4 数据库文件的复制	103
5.4 本章小结	105
习题 5	105
第 6 章 数据库查询	110
6.1 SELECT 语法格式	110
6.2 单表查询	111
6.2.1 投影查询	111
6.2.2 选择查询	117
6.2.3 聚合函数	122
6.2.4 对查询结果排序	123
6.2.5 对查询结果分组	125
6.3 连接查询	127
6.3.1 内连接查询	127
6.3.2 自连接查询	130
6.3.3 外连接查询	131
6.3.4 交叉连接查询	133
6.4 子查询	134
6.4.1 无关子查询	134
6.4.2 相关子查询	141
6.5 集合运算查询	144
6.5.1 集合的并	144
6.5.2 集合的交	145
6.5.3 集合的差	146
6.6 数据操作中使用查询语句	147
6.6.1 在 INSERT 语句中使用查询语句	147
6.6.2 在 UPDATE 语句中使用查询语句	148
6.6.3 在 DELETE 语句中使用查询语句	148
6.7 本章小结	149
习题 6	150
第 7 章 T-SQL 编程	154
7.1 T-SQL 编程基础	154

7.1.1	标识符	154
7.1.2	数据类型	155
7.1.3	注释	160
7.1.4	常量和变量	160
7.1.5	运算符和表达式	164
7.1.6	批处理	165
7.2	T-SQL 控制流语句	166
7.2.1	BEGIN...END 和 PRINT	166
7.2.2	IF...ELSE	167
7.2.3	CASE	168
7.2.4	WHILE、BREAK 和 CONTINUE	170
7.2.5	GOTO、RETURN 和 WAITFOR	171
7.3	函数	173
7.3.1	系统内置函数	173
7.3.2	用户定义函数	177
7.4	本章小结	181
	习题 7	182
第 8 章	视图、索引和游标	185
8.1	视图	185
8.1.1	视图概述	185
8.1.2	创建视图	186
8.1.3	修改视图	192
8.1.4	使用视图	193
8.1.5	删除视图	197
8.2	索引	197
8.2.1	索引的类型	198
8.2.2	创建索引	199
8.2.3	删除索引	204
8.3	游标	205
8.3.1	游标的类型	206
8.3.2	游标的操作	207
8.4	本章小结	213
	习题 8	214
第 9 章	存储过程和触发器	217
9.1	存储过程	217

9.1.1	使用存储过程的优点	217
9.1.2	存储过程的类型	218
9.1.3	创建和执行存储过程	220
9.1.4	修改和删除存储过程	233
9.1.5	存储过程的其他操作	235
9.2	触发器	240
9.2.1	触发器的类型	241
9.2.2	创建触发器	243
9.2.3	触发器的其他操作	255
9.3	本章小结	260
习题 9		261
第 10 章	数据库管理	263
10.1	数据库的备份与还原	263
10.1.1	数据库的备份	263
10.1.2	数据库的还原	265
10.2	数据的导入与导出	268
10.2.1	数据的导出	269
10.2.2	数据的导入	273
10.3	事务处理	275
10.3.1	事务	275
10.3.2	锁机制	276
10.4	本章小结	278
习题 10		278
第 11 章	数据库的安全管理	281
11.1	SQL Server 身份验证模式	281
11.1.1	身份验证模式概述	281
11.1.2	身份验证模式设置	283
11.2	登录账号和用户账号	285
11.2.1	服务器登录账号	285
11.2.2	数据库用户账号	288
11.3	角色	290
11.3.1	服务器角色	290
11.3.2	数据库角色	291
11.3.3	应用程序角色	295
11.4	用户权限管理	297

11.4.1	权限的类型	297
11.4.2	用户权限操作	299
11.5	本章小结	306
习题 11		307
第 12 章 基于 Java 的数据库应用系统开发		309
12.1	JDBC 相关知识介绍	309
12.1.1	JDBC 概述	309
12.1.2	JDBC 常用类和接口	309
12.1.3	JDBC 数据库访问步骤	312
12.2	基于 Java 的图书管理系统开发	313
12.2.1	系统需求分析	313
12.2.2	系统功能模块设计	314
12.2.3	系统数据库设计	314
12.2.4	系统实现	315
12.3	基于 Java 的学生成绩管理系统开发	320
12.3.1	系统需求分析	320
12.3.2	系统功能模块设计	320
12.3.3	系统数据库设计	321
12.3.4	系统实现	321
12.4	本章小结	328
习题 12		328
参考文献		329
附录 A 索引		329
附录 B 参考文献		329
附录 C 参考文献		329
附录 D 参考文献		329
附录 E 参考文献		329
附录 F 参考文献		329
附录 G 参考文献		329
附录 H 参考文献		329
附录 I 参考文献		329
附录 J 参考文献		329
附录 K 参考文献		329
附录 L 参考文献		329
附录 M 参考文献		329
附录 N 参考文献		329
附录 O 参考文献		329
附录 P 参考文献		329
附录 Q 参考文献		329
附录 R 参考文献		329
附录 S 参考文献		329
附录 T 参考文献		329
附录 U 参考文献		329
附录 V 参考文献		329
附录 W 参考文献		329
附录 X 参考文献		329
附录 Y 参考文献		329
附录 Z 参考文献		329

第1章

chapter 1

数据库系统概述

本章学习目标

- 掌握数据管理技术的发展历程。
- 熟练掌握数据库系统的组成。
- 掌握数据库系统的三级模式和两级映像。
- 了解信息的三种世界。
- 熟练掌握概念模型的基本概念和表示。
- 了解数据模型的基本概念和常用的数据模型。

当今信息社会中数据库技术的发展已经成为信息技术重要组成部分,绝大多数计算机应用系统都离不开数据库的支撑。本章首先介绍数据管理技术的发展历程,主要介绍每个阶段的特点;其次介绍数据库系统的组成、数据库系统的三级模式和两级映像;最后介绍数据模型,包括信息的三种世界、概念模型的基本概念和表示方法以及数据模型的分类等内容。

1.1 数据库管理技术发展史

1.1.1 数据处理技术

1. 数据

信息社会中我们每时每刻都在和信息打交道,那么我们经常用到的“信息”这个词,是什么意思呢?简单地说,信息能够反映物质和能量的形态、结构、状态等特征,可以用人的感官感觉到,也可以用仪器、仪表和各种传感器进行探测。信息具有实效性、有用性和知识性,它是客观世界的反映。

数据(Data)是承载信息的符号记录,是信息的载体,是信息的具体表现形式。在日常生活中,人们直接用自然语言描述、传递信息;在计算机中,为了存储和处理这些信息,就要抽取出信息中包含的事物及其特征,以及和其他事物的联系,用数据来表示它们。数据通常由一组“数字”组成,用以表示某一事物或事件的特征、数量等。例如,描述一个学生信息,我们关心的是该学生的学号、姓名、年龄、性别、班级信息,则可以表示为(1401010201,

张三,19,男,信息管理与信息系 2014 级 2 班)。

必须指出的是,在有些不严格的情况下,会把“数据”和“信息”两个概念混为一谈,即对它们不加区分。其实,数据不等于信息,数据只是信息的一种表达形式。

数据有“型”和“值”之分。数据的型是指数据的结构,数据的值是指数据的具体取值。例如,学生数据由学号、姓名、年龄、性别、班级属性构成,这是学生数据的型,而一个具体的学生数据(1401010201,张三,19,男,信息管理与信息系 2014 级 2 班)就是数据的值。

数据本身和其语义是不可分的。计算机处理的数据是经过抽象的,它必须联系其语义解释才有意义。例如,数据 19,这可能是一个学生的年龄,可能是一个团队的人数,也可能是一棵树的高度。所以,离开语义解释,数据就失去了意义。又如,(1401010201,张三,19,男,信息管理与信息系 2014 级 2 班),了解语义解释的人知道,这个数据表示信息管理与信息系 2014 级 2 班的一个学生信息,学号是 1401010201,姓名是张三,年龄是 19 岁,性别是男。

2. 数据处理

数据处理(Data Process)是指对数据进行收集、分类、组织、整理、存储、查询、维护、加工、计算、传播、打印等一系列的活动。数据处理的基本目的是从大量的、可能是杂乱无章的、难以理解的数据中抽取并推导出某些特定的对人们来说是有价值、有意义的数据。数据处理贯穿于社会生产和社会生活的各个领域,数据处理技术的发展及其应用的广度和深度,极大地影响着人类社会发展的进程。随着计算机的日益普及,通过计算机数据处理进行信息管理已成为主要的应用,如仓库管理、财会管理、交通运输管理、技术情报管理、办公室自动化等。

数据管理(Data Management)是指数据的收集整理、组织、存储、维护、检索、传送等操作,是数据处理业务的基本环节,是所有数据处理过程必有共同部分。数据处理中,通常计算比较简单,且数据处理业务中加工计算因业务的不同而不同,需要根据业务的需要来编写应用程序加以解决。而数据管理则比较复杂,由于可利用的数据呈爆炸性增长,且数据的种类繁多,从数据管理角度而言,不仅要使用数据,而且要有效地管理数据,因此需要一个通用的、使用方便且高效的管理软件,把数据有效地管理起来。数据管理技术的优劣将对数据处理的效率产生直接影响,而数据库技术就是针对该需求目标进行研究并发展和完善起来的计算机应用的一个分支。

1.1.2 数据管理技术的发展

数据管理技术是随着计算机硬件和软件的发展而不断发展的,也是随着社会对数据处理任务的要求不断提高而发展的。数据库是数据管理的产物,随着计算机应用的不断发展,数据处理越来越占主导地位,数据库技术得到不断的发展,数据库技术的应用也越来越广泛。从数据管理的角度,数据库技术经历了人工管理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前,计算机处于发展初期,主要用于科学计算,数据管理能力很差。此时的计算机硬件和软件发展都相对落后。从硬件来看,外存只有磁带、卡片、纸带,没有磁盘等直接存取的存储设备;从软件来看,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。

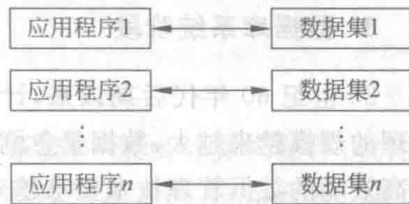
这个时期数据管理的特点主要有以下几点:

(1) 数据不能长期保存。由于当时计算机水平的限制,还没有磁盘等外存储设备,且计算机主要用于科学计算,因此一般不需要将数据长期保存。当某一计算需要数据时,将数据输入,用完后不保存原始数据也不对计算结果保存,下次需要时再进行人工输入。

(2) 没有对数据进行管理的软件系统。应用程序中需要用到数据时,程序员不仅要规定数据的逻辑结构,而且要设计物理结构,包括存储结构、存取方法、输入/输出方式等。此时的数据管理完全由应用程序完成。

(3) 数据不独立。当数据逻辑结构或物理结构改变后,必须由程序员对程序做相应修改,此时的数据不能离开应用程序单独修改,这也加重了程序员的负担。

(4) 数据不共享。数据的组织方式是由程序员编写程序时自行设计的,某一数据对应某一程序,即使两个程序用到相同的数据,也必须各自定义、各自组织,数据无法共享,这也导致应用程序之间有大量重复的数据,数据冗余。



人工管理阶段应用程序与数据之间的对应关系 图 1.1 人工管理阶段应用程序与数据之间的对应关系

2. 文件系统阶段

20 世纪 50 年代后期至 20 世纪 60 年代中期,计算机得到一定发展,应用范围越来越广泛,不仅用于科学计算,在数据管理方面也有了一定应用。此时,在计算机硬件方面,出现了磁盘、磁鼓等作为直接的外存储设备;在软件方面,出现了操作系统,有了专门用于管理数据的软件,称为文件系统,处理方式不仅有文件批处理,而且能够进行联机实时处理。

这个时期数据管理的特点主要有以下几点:

(1) 数据可以长期保存。此时的计算机有了磁盘这样的外存储设备,使得数据可以长期存储在外存上,便于对数据进行反复的查询、修改、插入和删除等各种操作。

(2) 有专门的软件进行数据管理。操作系统提供了文件管理功能和访问文件的存取方法,数据的存取以记录为基本单位,文件系统实现了记录内的结构化,即给出记录内各种数据间的关系,但是,数据文件从整体来看是无结构的。此阶段出现了多种文件组织,如顺序文件、索引文件、随机文件等。

(3) 数据独立性低。文件系统可以对数据进行管理,使数据具有了一定的独立性,但文件系统只实现了数据的物理独立性,而没有实现数据的逻辑独立性。文件系统中文

件是为某一特定应用服务的,一旦数据的逻辑结构改变,必须修改应用程序,修改文件结构的定义;而应用程序的改变,如改用不同的高级语言,也将引起文件数据结构的改变,因此数据与应用程序之间的独立性较低。

(4) 数据共享性差。在文件系统阶段,数据仍然面向某一应用程序,一个数据文件基本上对应于一个应用程序,当多个应用程序使用相同的数据时,也必须建立各自的数据文件,而不能共享一个数据文件。因此,在文件系统阶段,数据冗余度仍然很大,而且大量数据重复存储也容易导致数据的不一致性。

文件系统阶段应用程序与数据之间的对应关系如图 1.2 所示。

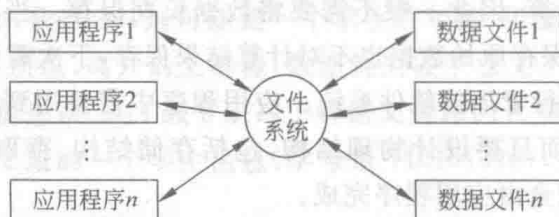


图 1.2 文件系统阶段应用程序与数据之间的对应关系

3. 数据库系统阶段

20 世纪 60 年代后期以来,计算机技术发展迅速,计算机应用越来越广泛,用于数据管理的规模越来越大,数据量急剧增长,传统的文件系统已经不能满足数据管理的需求,对高效能的数据管理技术要求越来越强烈。

当时计算机的软硬件也得到了很大发展。在硬件方面,有了大容量的磁盘存储设备,且价格逐渐降低,处理方式上,联机实时处理的要求更高,并开始提出和考虑分布处理;软件方面出现了能够统一管理和共享数据的数据库管理系统(DBMS)。

这个时期数据管理的特点主要有以下几点:

(1) 数据结构化。数据结构化是数据库系统阶段与文件系统阶段的根本区别。在文件系统阶段,文件的记录内部是有结构的,但数据文件之间无联系,数据文件是面向特定应用程序的。而在数据库系统阶段,数据库中的数据不再针对某一特定应用,而是面向整个应用系统,它是对整个系统各个应用整体考虑后建立的总的数据结构,在描述数据时不仅描述数据本身,还描述数据之间的联系,能够展现整体数据的结构化。

(2) 具有较高的数据独立性。在数据库系统阶段,数据不再和某一应用程序对应,它们之间不存在相互依赖的关系,应用程序不再随着数据存储结构的变化而变化。数据库系统中的数据既具有逻辑独立性,又具有物理独立性。逻辑独立性使得当数据的逻辑结构发生改变时,应用程序不必修改;物理独立性使得当数据的存储结构发生改变时,可以使数据的逻辑结构不变,从而应用程序不必修改。数据独立性使数据的定义从程序中分离出去,由专门的数据库管理系统管理,从而简化了应用程序的编写,减轻了程序员的负担。

(3) 具有较高的共享性、较低的冗余度。在数据库系统阶段,数据库中的数据是面向整个应用系统的,允许多个用户、多个应用程序同时存取数据而互不影响,这就提高了数

据的共享性,大大减少了重复数据,使数据具有较低的冗余度。由于减少了重复数据,因此避免了数据之间的不相容性和不一致性。

(4) 数据由数据库管理系统统一管理和控制。在数据库系统阶段,数据的管理完全由数据库管理系统完成,除了管理功能,数据库管理系统还保证数据的完整性、安全性、并发控制和数据恢复等。

数据库系统阶段应用程序与数据之间的对应关系如图 1.3 所示。



图 1.3 数据库系统阶段应用程序与数据之间的对应关系

1.2 数据库系统的介绍

1.2.1 数据库系统的组成

数据库系统是计算机应用系统中引入数据库后的系统,一般由计算机硬件、软件和人员构成,具体包括计算机硬件、操作系统、数据库、数据库管理系统、应用开发工具、应用系统、用户等。计算机硬件指存储和运行数据库系统的硬件设备,包括 CPU、内存、大容量的存储设备、输入/输出设备等。下面仅介绍其中的一部分组成。

1. 数据库

数据库(DataBase,DB)是长期存储在计算机内有组织的、可共享的数据集合。这些数据按一定的数据模型组织,无有害的或不必要的冗余,并为多种应用服务,数据的存储独立于使用它的程序。可以直观地认为,数据库就是存储数据的仓库,只是这个仓库是计算机的大容量存储设备,仓库中的数据是要按一定的数据结构存储,且具有较高的共享性、独立性,较低的冗余度,为多种应用服务。

2. 数据库管理系统

数据库管理系统(DataBase Management System,DBMS)是一种操纵和管理数据库的软件,用于建立、使用和维护数据库,它对数据库进行统一的管理和控制,以保证数据库的安全性和完整性。DBMS 是建立在操作系统之上的、位于用户和操作系统之间的数据管理软件,为用户或应用程序提供访问数据库的方法。数据库管理员也通过 DBMS 进行数据库的维护工作,包括数据库的建立、数据操作、安全性、完整性等操作。

DBMS 是数据库系统的核心,是管理数据库的软件。DBMS 的功能主要有以下几点:

(1) 数据定义功能。DBMS 提供数据定义语言(Data Definition Language, DDL), 供用户定义数据库的三级模式结构、两级映像以及完整性约束和保密限制等。例如, DBMS 提供的结构化查询语言 SQL (Structured Query Language) 使用 CREATE、DROP、ALTER 等语句分别用来建立、删除和修改数据库。DDL 所描述的结构仅仅给出了数据库的框架, 定义的内容保存在数据字典(Data Dictionary, DD)中, 数据字典是 DBMS 存取数据的基本依据。

(2) 数据操纵功能。DBMS 提供数据操纵语言(Data Manipulation Language, DML), 供用户对数据库中的数据进行操作, 主要操作包括对数据进行查询、插入、删除、修改等。DML 主要有两类: 一类是自主型或自含型, 这一类是交互式命令语言, 语法简单, 可独立使用; 另一类是宿主型, 把对数据库的存取语句嵌入到高级语言中, 例如嵌入到 C、Java 等。

(3) 数据库运行管理功能。数据库的运行管理功能是 DBMS 的核心部分, 主要包括对数据库进行并发控制、安全性检查、完整性约束条件的检查和执行、数据库的内部维护等。所有访问数据库的操作都要在这些控制程序的统一管理下进行, 以保证数据的安全性、完整性、一致性和多用户对数据库的并发控制。

(4) 数据库的建立和维护功能。数据库的建立包括数据库的初始数据的装入与数据转换等。数据库的维护包括数据库的转储、恢复、重组与重构造、系统性能监视与分析等。

(5) 数据组织、存储和管理。数据库中存放多种数据, 包括数据字典、用户数据、存储路径等, 数据库管理系统负责对这些数据进行分门别类的组织、存储和管理, 以确定用何种文件结构和存取方式物理地组织这些数据, 以提高存储空间利用率, 提高对数据进行查询、增加、删除、修改等操作的效率。

(6) 数据通信功能。DBMS 提供与其他软件系统进行通信的功能。DBMS 提供了与其他 DBMS 或其他软件的接口, 从而实现不同 DBMS 间数据的转换、异构数据库之间的互操作等, 通常这些功能要与操作系统协调完成。

3. 用户

用户是开发、管理和使用数据库系统的人员, 主要包括数据库管理员、应用程序员和终端用户。

1) 数据库管理员

数据库管理员(DataBase Administrator, DBA)负责设计、建立、管理和维护数据库以及协调用户对数据库的各种要求。DBA 应熟悉计算机的软硬件系统, 具有较全面的数据处理知识。DBA 的主要职责包括设计数据库的逻辑结构; 决定数据库的存储结构和存取策略; 定义数据的安全性和完整性; 监督和控制数据库系统的使用和运行; 数据库的改进和重组重构。

2) 应用程序员

应用程序员(Application Programmer)负责分析、设计、开发、维护数据库系统的程序模块。应用程序员使用软件工程的方法对系统进行需求分析, 与终端用户及数据库管