

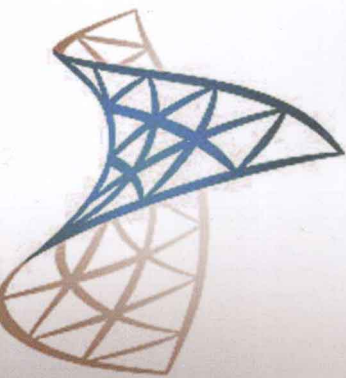


SQL Server 2012是微软公司21世纪初具有重要意义的数据库新产品。作为新一代的数据平台，数据管理能力强大，全面支持云技术与多种系统，可快速构建相应的解决方案实现私有云与公有云之间数据的扩展与应用的迁移。

高等院校规划教材

上海市精品课程特色教材

上海市教育高地建设项目



数据库原理应用 与实践

SQL Server 2012

贾铁军 甘泉 主编



科学出版社

高等院校规划教材
上海市精品课程特色教材
上海市教育高地建设项目

数据库原理应用与实践

SQL Server 2012

贾铁军 甘 泉 主编

科学出版社

北京



内 容 简 介

本书为上海市精品课程“数据库原理及应用”特色教材。主要突出“实用、新颖、操作性强的特色”，重点结合最新的 SQL Server 2012 技术及应用，介绍数据库的基本原理和技术方法。全书共 11 章，包括数据库概述、关系数据库基础、SQL Server 2012 概述及有关操作、索引及视图、T-SQL 应用编程、关系数据库的规范化、存储过程与触发器、数据库设计、数据库安全与保护、数据库新技术、数据库技术应用实践。

本书由科学出版社提供多媒体课件，并配有实验指导、辅助教材——《数据库原理及应用学习与与实践指导》，内容包括学习要点、实验及课程设计指导、习题与实践练习、自测试卷和答案等。通过课程网站提供动画视频、程序、教案等资源。

本书可作为高等院校计算机类、信息类、理工及电子商务类和管理类各专业本科生相关课程的教材，高职院校可选用，也可作为培训及其他参考用书。

图书在版编目(CIP)数据

数据库原理应用与实践 SQL Server 2012/贾铁军,甘泉主编. - 北京:科学出版社,2013. 1

(高等院校规划教材·上海市精品课程特色教材·上海市教育高地建设项目)

ISBN 978-7-03-036421-0

I. ①数… II. ①贾…②甘… 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(2012)第 319100 号

责任编辑:谭宏宇 王艳丽 / 责任校对:宣 慧

责任印制:刘 学 / 封面设计:殷 靓

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

上海叶大印务发展有限公司印刷

科学出版社发行 各地新华书店经销

*

2013 年 1 月第 一 版 开本:787×1092 1/16

2013 年 1 月第一次印刷 印张:20 3/4

字数:499 000

定价:48.00 元

(如有印装质量问题,我社负责调换)

**《数据库原理应用与实践 SQL Server 2012》
编辑委员会**

主 编 贾铁军 甘 泉
副主编 沈学东 常 艳 连志刚 胡 静
编 委 李向臣 曾 刚 邹佳芹

前 言

数据库技术是计算机技术中发展最快、应用最广泛的一项技术,已经成为各类计算机信息系统的核心技术和重要基础。数据库技术是数据管理与处理的高新技术,是计算机科学的重要分支,与计算机网络、人工智能一起被称为计算机技术界三大热门技术,是现代化管理的有力工具。

进入 21 世纪,信息技术的快速发展为现代信息化社会带来了深刻的变革。信息、物资和能源已经成为人类赖以生存和发展的重要保障,数据已经成为重要信息资源和新拓展能源,数据管理与处理已经广泛应用于各种业务,数据库技术及应用已经遍布各行各业的各个层面,电子商务系统、网络银行、管理信息系统、企业资源计划、供应链管理系统、客户关系管理系统、决策支持系统、数据挖掘信息系统等,都离不开数据库技术强有力的支持,数据库技术具有广阔的发展和应用前景。

SQL Server 2012 是微软公司 21 世纪初具有重要意义的数据库新产品。作为新一代的数据平台,数据管理能力强大,全面支持云技术与多种系统,可快速构建相应的解决方案实现私有云与公有云之间数据的扩展与应用的迁移。提供对企业基础架构最高级别的支持——专门针对关键业务应用的多种功能与解决方案,可以提供高级别的可用性及性能。在业界领先的商业智能领域,提供了更多更全面的功能以满足不同人群对数据信息的需求,包括支持来自于不同网络环境的数据的交互,全面自助分析等创新功能。在企业级支持、商业智能应用、管理开发效率等方面具有显著功能,是集数据管理与商业智能分析于一体的新式数据管理与分析平台,并具有完整的关系数据库创建、管理、设计和开发功能。

本书作者长期从事计算机相关专业的教学与科研工作,不仅积累了丰富的教学经验,而且还有多年数据库应用系统的研发设计经历。本书是 2012 年上海市精品课程“数据库原理及应用”的特色教材和“校企-校校合作的新成果”,特奉献给广大师生进行教学和交流。

本书共分 11 章,重点结合最新的 SQL Server 2012,介绍数据库的基本原理、新技术、新应用和新方法。主要内容包括数据库概述、关系数据库基础、SQL Server 2012 概述及有关操作、索引及视图、T-SQL 应用编程、关系数据库的规范化、存储过程与触发器、数据库设计、数据库安全与保护、数据库新技术、数据库技术应用实践。书中带“*”部分为选学内容。

本书主要突出“实用、特色、新颖、操作性”。旨在重点介绍数据库的最新成果、基本原理、新技术、新方法和实际应用。其特点是:

(1) 内容先进,结构新颖。吸收了国内外大量的新知识、新技术和新方法。注重科学性、先进性、操作性。图文并茂、学以致用。每章配有“教学目标”、“案例”和“讨论交流”等。

(2) 注重实用性和特色。坚持“实用、特色、规范”原则,突出实用及素质能力培养,增加大量案例和同步实验,在内容安排上将理论知识与实际应用有机结合。

(3) 资源配套,便于教学。为了方便师生教学,通过科学出版社提供多媒体课件等。配有实验指导、辅助教材——《数据库原理及应用学习与实践指导》,内容包括学习要点、详尽实验及课程设计指导、习题与实践练习、自测试卷和答案等。通过上海市精品课程网站 www.jiatj.com。

sdju.edu.cn 提供动画视频、程序代码、教学大纲及教案等资源。

本书由上海市精品课程“数据库原理及应用”负责人贾铁军教授任主编、统稿并编著第 3 章、第 5 章和第 9 章,甘泉总经理(大连信源网络科技有限公司)任主编并编著第 1 章,常艳教授(辽宁警官学院)任副主编并编著第 6 章,上海电机学院沈学东副教授任副主编并编著第 8 章和第 11 章,胡静副教授任副主编并编著第 7 章,连志刚副教授任副主编并编著第 10 章,李向臣(上海海洋大学)编著第 2 章,曾刚(辽宁警官学院)编著第 4 章,邹佳芹女士完成了部分习题解答和课件制作,并对全书的文字、图表进行了校对编排及查阅资料等。邹飞和于森参加本书编著大纲的讨论、编著审校等工作。

非常感谢科学出版社为本书的编著与出版提供了重要的帮助和指导意义。同时,感谢对本书编著给予大力支持和帮助的院校及企业领导和同仁。对编著过程中参阅大量的重要文献资料难以完全准确注明,在此深表诚挚谢意!

由于内容庞杂、技术更新迅速、时间仓促及水平有限,书中难免存在不妥之处,敬请海涵见谅! 欢迎提出宝贵意见和建议。

编 者

2012 年 9 月于上海

目 录

前言

第 1 章 数据库概述	1
1.1 数据库的概念	1
1.2 数据库技术的发展	5
1.3 数据库系统的构成	9
1.4 数据库的模式结构.....	13
1.5 数据库管理系统.....	15
1.6 数据模型.....	17
1.7 本章小结.....	32
第 2 章 关系数据库基础	33
2.1 关系模型概述.....	33
2.2 关系模型的完整性.....	38
2.3 关系运算基础.....	40
2.4 关系演算	48
2.5 查询优化	50
2.6 本章小结.....	52
第 3 章 SQL Server 2012 概述及有关操作	53
3.1 SQL 的概念和新特点	53
3.2 SQL Server 2012 的特点和功能	55
3.3 SQL Server 2012 结构及数据库种类	59
3.4 常用的数据类型.....	66
3.5 SQL Server 2012 安装配置和登录	69
3.6 常用的数据库和表操作.....	76
3.7 数据查询.....	83
3.8 数据更新方法.....	88
3.9 本章小结.....	92
第 4 章 索引及视图	93
4.1 索引概述.....	93
4.2 索引的基本操作.....	99
4.3 视图及其应用	105
4.4 视图的常用操作	107
4.5 特殊类型视图的应用	115
4.6 本章小结	120

第 5 章 T-SQL 应用编程	121
5.1 T-SQL 基础概述	121
5.2 批处理、脚本及事务	127
5.3 常量变量函数和表达式	129
5.4 流程控制语句	150
* 5.5 知识扩展: SQL Server 2012 对 T-SQL 的增强	157
* 5.6 嵌入式 SQL 概述	158
5.7 本章小结	159
第 6 章 关系数据库的规范化	161
6.1 规范化的主要问题	161
6.2 函数依赖概述	163
* 6.3 关系模式的分解	168
6.4 关系模式的范式	172
6.5 关系模式的规范化	177
6.6 本章小结	178
第 7 章 存储过程与触发器	179
7.1 存储过程概述	179
7.2 存储过程的实现	180
* 7.3 触发器应用	185
7.4 本章小结	190
第 8 章 数据库设计	191
8.1 数据库设计概述	191
8.2 数据库应用系统设计	195
8.3 数据库设计文档	212
8.4 数据库应用系统设计案例	214
8.5 本章小结	235
第 9 章 数据库安全与保护	237
9.1 数据库安全性概述	237
9.2 数据库安全技术及机制	240
9.3 数据库的访问权限及控制	243
9.4 数据的完整性	253
9.5 并发控制与封锁	262
9.6 数据备份与恢复	265
9.7 本章小结	272
第 10 章 数据库新技术	273
10.1 数据库新技术概述	273
10.2 面向对象数据库	274
10.3 分布式数据库	278
10.4 数据仓库与数据挖掘	281

10.5	其他新型数据库	286
10.6	数据库新技术发展趋势	295
10.7	本章小结	296
第 11 章	数据库技术应用实践	297
11.1	实验一 SQL Server 2012 界面及功能	297
11.2	实验二 关系模式及模型应用	298
11.3	实验三 常用数据操作	300
11.4	实验四 数据库基本操作	304
11.5	实验五 T-SQL 程序结构	305
11.6	实验六 存储过程及触发器	307
11.7	实验七 数据库应用系统设计	308
11.8	实验八 数据库安全	311
11.9	实验九 分布式数据库应用	314
11.10	综合应用实践(课程设计指导)	316
11.11	本章小结	318
参考文献		319

第 1 章 数据库概述

进入 21 世纪的现代信息化社会,数据库技术已经成为各领域的各种业务数据管理的重要工具和最新技术,与计算机网络、人工智能一起被称为计算机技术界三大热门技术。随着 IT 技术的快速发展和广泛应用,数据库技术的应用已经从事务处理扩展到计算机网络服务、商务智能、计算机辅助设计和决策支持系统等新领域,各行业大量的重要数据,需要经过数据库才能进行有效组织、存储、处理和共享。通过学习数据库有关技术知识,可以为以后的业务学习或就业奠定重要基础。



教学目标

- 熟悉数据、数据管理和数据库的基本概念
- 掌握数据库技术特点、应用及发展趋势
- 了解数据库系统的组成及数据库的体系结构
- 掌握 DBMS 的工作模式、主要功能和组成
- 理解概念模型与数据模型

1.1 数据库的概念

熟练掌握数据、数据库、数据管理与处理和数据库系统等基本概念,对数据库技术和后续知识及内容的学习极为重要。

1.1.1 信息和数据的概念

物资、信息和能源是人类赖以生存和发展的三大支柱,各种信息只有经过数据表示才能进行采集、传输、存储与处理。

1. 信息的概念

在客观世界,信息无处不在无处不用,各种业务及生活的衣食住行都离不开信息。

信息(information)是人们对现实事物的状态和特征的描述,是进行决策的重要依据,如知道下雨的气象信息,就会带雨伞。

信息是各种客观事物的存在方式、运动形态和具体特征,及其之间相互联系等要素在头脑中的反映,通过人脑的抽象后形成的概念及描述。

客户资料、供应商情况、进货的采购订单、销售的销售报表、企业的库存资料、产品资料等都是企业的营销信息。

2. 数据的概念

各种信息必须经过数据载体的描述和表示,才能进行采集、传输、存储与处理。

数据(data)是信息的表达方式和载体,是人们描述客观事物及其活动的抽象表示,是描述事物的符号记录,是利用信息技术进行采集、处理、存储和传输的基本对象。数据的概念包括描述事物特性的数据内容和存储在某一种媒体上的数据形式。

数据的概念包括两方面含义:一是数据的内容是信息;二是数据的表现形式是符号。

通常,数据分为数值数据和非数值数据两大类,可以是数字、文字、符号、图形、表格、图像、声音、录像、视频等形式。数据是数据库中存储与管理的基本对象,人们收集并抽取所需要的大量数据之后,将其保存经过进一步加工处理,从而得到有用信息。

数据库中的数据具有两个特性:

(1) 整体性。数据库中的数据都是从全局观点出发建立的,按一定的数据模型(结构)进行组织、描述、存储、管理和控制。

(2) 共享性。数据库中的数据是为多用户共享建立的,已经摆脱了具体程序的限制和制约;不同的用户可以按各自的用法使用数据库中的数据。

3. 信息与数据的区别

数据与信息既有区别又互相依存。数据是信息的具体表示形式和载体,信息反映数据的含义。数据是数据库管理的基本内容和基本对象,是信息的一种符号化表示方法,采用一定的符号表示信息,而具体用哪种形式的符号及表示方式方法,则是人为规定。信息来源于数据,数据是信息的具体表现形式,信息以数据的形式存储、管理、传输和处理,数据经过处理后可得到更多有价值的信息。信息是观念性的,数据是物理性的。信息可用数据的不同形式来表示,数据的表示方式可以选择,而信息不随数据表现形式而改变。

1.1.2 数据库与数据库管理系统

1. 数据处理与管理

数据处理(data processing)是对数据进行加工的过程。对数据进行的查询、分类、修改、变换、运算、统计、汇总等都属于加工。其目的是根据需要,从大量的数据中抽取有意义、有价值的信息,作为决策和行动的根据,其实质是信息处理。

数据管理(data management)是对原有基本数据进行管理为目的,在数据处理过程中,数据收集、存储、检索、分类、传输等基本环节统称为数据管理。

△**注意:**数据处理与数据管理的区别:狭义上一般使数据发生较大根本性变化的数据加工称为数据处理,如汇总;而广义上时常不加区别地统称为数据管理。

【案例 1-1】从货物信息管理系统的“货物价格数据表”中,查找价格最高的货物、按价格从高到低排序、修改价格或打印等操作都属于数据管理,而进行价格统计汇总或制作货物数据图则属于数据处理。

2. 数据库与数据库系统

数据库(database, DB)是存储在计算机上的结构化的相关数据集合。可理解为“按一定结构存管数据的仓库”,是在计算机内的、有组织(结构)的、可共享、长期存储的数据集合。数据库中的数据可按一定的数据模型(结构)进行组织、描述和存储,具有较高的数据独立性和易扩展性,具有较小的冗余度,并可共享。数据库还具有集成性、共享性、海量性和持久性等特点。数据库技术的主要目标是根据用户需求自动处理、管理和控制大量业务数据。

数据库系统(database system, DBS)是具有数据库功能特点的计算机系统,是实现有组织

地、动态地存储大量关联数据、方便多用户访问的计算机软硬件和数据资源组成的系统。其主要特性为:实现数据共享,减少数据冗余度;保持数据一致性和数据独立性;提高系统的安全保密性,并发控制及故障恢复。

3. 数据库管理系统

数据库管理系统 (database management system, DBMS) 是建立、运用、管理和维护数据库,并对数据进行统一管理和控制的软件。便于用户定义和操纵数据,并保证数据的安全性、完整性、多用户对数据同时并发使用及发生意外时的数据库恢复等。DBMS 是整个数据库系统的核心,对数据库中的各种数据进行统一管理、控制和共享。DBMS 的功能和结构将在 1.5 节中介绍,其地位如图 1-1 所示。常见的大型关系型 DBMS 如微软的 SQL Server、IBM 的 DB2,以及 Oracle、Sybase、Informix 等,桌面单机型 FoxPro、Access 等。

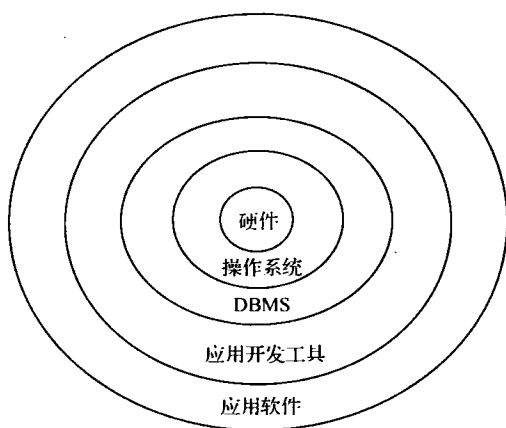


图 1-1 DBMS 的地位

常见的大型关系型 DBMS 如微软的 SQL Server、IBM 的 DB2,以及 Oracle、Sybase、Informix 等,桌面单机型 FoxPro、Access 等。

1.1.3 数据库技术的特点及应用

1. 数据库技术的主要特点

(1) 数据高度集成。数据处理应用系统中的数据源于多项业务,且数据之间相互关联。如在一个商品供销管理信息系统中,进货数据来源于供货管理、销售数据来源于售货管理、员工数据来源于人力资源管理等。对这些数据进行集中管理,保持之间的正确关联,才能完成所需的综合数据处理。利用数据库技术和 DBMS 提供的数据库管理功能便可实现多种数据的集成。

(2) 数据广泛共享。在一个数据库应用系统中,可对集中管理的数据进行共享。如供货管理需要参考商品销售管理系统中近期的销售数据,确定进货种类与数量,确定销售单价又需要参照最近的进货单价等,利用数据库技术通过计算机网络可实现数据广泛共享。

(3) 数据独立冗余低。数据独立性是指数据库存储的数据与处理数据的应用程序互相独立。避免了在传统的数据库应用系统中,应用程序与相关业务数据关联,致使各种业务数据在多种不同的数据文件中分别存储,数据大量冗余且无法统一更新。数据库技术可对所有数据集中管理,并利用有效的数据库共享功能,不再需要各项业务单独保存各自的数据文件,极大地减少了数据冗余。

(4) 实施统一的数据标准。数据标准是指数据库中数据项的名称、数据类型、数据格式、有效数据的判定准则及要求等数据项特征值的取值规则。

(5) 控制数据的安全性和完整性。针对数据库所进行的各种操作都必须根据操作者所拥有的权限进行鉴别,鉴别机制由 DBMS 提供,每个用户的操作权限设定则由数据库管理员 DBA 负责。以保障数据库应用系统及数据的安全性、机密性和完整性。

(6) 保证数据一致性。数据一致性是指存储在数据库中不同数据集合(表)的相同数据项必须具有相同的值。一个数据库由多种数据文件组成,数据文件之间通过公共数据项相联系,

当对一个数据文件中的数据项更新时,相关联文件中的对应数据项也必须自动更新,才能始终保持数据库数据的一致性和正确性。通过 DBMS 可以自动实现对数据库中数据,进行追加、插入和删除等操作的一致性问题的。

(7) 应用程序开发与维护效率高。在应用程序开发时,数据的独立性可不必考虑软件和数据关联问题,以及所处理的数据组织等问题,减少了应用程序的开发与维护的工作量。只在应用系统开发初期,需要规划数据库、设计数据库中的各个数据集、规范数据库中相关数据间的关联。只有一个满足规范化设计要求的数据库,才能够真正实现各类业务不同的应用需求。

2. 数据库技术应用

随着 IT 技术的快速发展,数据库技术的应用从数据处理与管理,扩展到计算机辅助设计、人工智能、决策支持系统和计算机网络应用等新领域。在 21 世纪现代信息化社会,由于信息(数据)无处不在,所以,数据库技术的应用非常广泛深入,遍布各个领域、行业、业务部门和各个层面。网络数据库系统及数据库应用软件已成为信息化建设和应用中的重要支撑性软件产业,得到广泛应用。

【案例 1-2】 数据库技术应用行业实例。

(1) 销售业:存储、查询供应商、商品、客户及销售信息和商品的网上订购等。

(2) 金融业:包括网银,用于银行客户的信息、账户、贷款以及银行的交易记录。还可以用于存储股票、债券、金融票据的持有、出售和买卖等交易信息。

(3) 制造业:用于产供销存储产品的订单、产品原料的供应情况,跟踪产品的产量以及仓库产品的详细清单,如零部件生产加工及组装各种业务数据的管理和应用。

(4) 电信业:用于存储通信网络的信息,存储通话记录,存储用户付费业务记录及产生每月通信账单,以及交费情况等。

(5) 航空业:用于存储、查询、网络订购国内外各种航班和票务信息。航空业是最先以地理上分布的方式使用数据库的行业之一,分布于世界各地的终端通过通信网络或其他数字网络来访问中央数据库系统。

(6) 教育系统:用于存储教职员的信息,存储工资、津贴和纳税的信息,产生工资单;存储学生信息、课程及实验信息、成绩信息和大学生科创信息等。

数据库技术是数据管理的最新技术,给各种广大用户的业务发展和生活带来了极大便利。例如,通过网络查询信息、预订机票、网上购物和付费等,数据库的应用更加广泛深入。

随着信息技术的快速发展,数据库技术也产生了一些新的应用领域,主要包括:

(1) 多媒体数据库。多媒体数据库主要存储与多媒体相关的数据,如声音、图像和视频等数据。多媒体数据最大的特点是数据量大、数据类型多且数据类型间差距较大,因此,需要较大的存储空间和较复杂的数据管理。

(2) 空间(云)数据库。将数据库应用于地理信息系统 GIS 和计算机辅助设计与制造(CAD/CAM),与地球上的空间位置相关的空间数据是 GIS 的重要组成部分,而设计数据(如机械及集成电路设计等)则是 CAD/CAM 的核心。

(3) 移动数据库。移动数据库是在移动计算机系统上发展起来的,如笔记本电脑、掌上电脑等。其最大特点是通过无线数字通信网络进行传输,用户可以随时随地访问和获取数据,为一些商务应用和应急情况带来了极大的便利。

(4) 信息检索系统。信息检索系统为一种典型的联机文档管理系统,一直与数据库技术

同步发展。

(5) 决策支持系统。联机分析处理(OLAP)是数据库系统的主要应用,支持复杂的分析操作,侧重决策支持,并且提供直观易懂的查询结果。可使分析管理人员或执行人员从多角度对信息进行快速、一致、交互地存取,获得对数据更深入的了解。

讨论思考:

- (1) 什么是数据管理?与数据处理有何区别?
- (2) 数据库系统与数据库管理系统的区别有哪些?
- (3) 数据库技术的主要特点有哪些?

1.2 数据库技术的发展

计算机和网络技术的快速发展,极大地促进了数据库技术的快速发展和广泛应用。数据库技术的发展主要经历了人工管理、文件管理、数据库管理以及高级数据库管理四个发展阶段。对“数据库新技术”将在第10章进行详细介绍。

1.2.1 人工管理阶段

1946年世界上第一台计算机“ENIAC”的诞生初期,应用计算机面临的一个重要问题就是数据的存储。当时的计算机是将数据以打孔的方式存储在纸带上,既不容易检索也不容易修改。以电子管为主要元器件,主要依靠硬件系统,包括运算器、控制器、存储器和简单的输入输出设备,工作效率极低,只能计算并输入输出很少的数据。

当时的计算机主要应用于科学计算,绝大部分的数据管理基本还是以手工方式,用纸卡及报表等进行记载、储存、查询和修改。当时外存没有磁盘等直接存取的存储设备,还没有操作系统和数据库。

人工管理数据的特点是:

(1) 计算机不存储数据。由于当时计算机软硬件技术所限,数据随程序一起输入计算机,处理结束后输出结果,无法长期保存,数据空间与程序一起被释放。

(2) 数据面向应用。一组数据对应一个应用程序。若多个程序需要使用相同的数据,则需在这些程序中重复存储相同的数据,程序之间的数据不能共享,造成数据的大量冗余,容易导致数据不一致。

(3) 数据不独立。当应用程序改变时,数据的逻辑结构和物理结构也相应发生变化。

(4) 无数据文件处理软件。数据的组织方式由程序员设计与安排,数据需要由应用程序进行管理,没有相应的数据文件处理软件。

1.2.2 文件管理阶段

从20世纪50年代中期到60年代中期,计算机以晶体管取代了运算器和控制器中的电子管,由于存储介质的更新,数据以文本文件或二进制文件的形式存储。可将成批数据单独组成文件存储到外部存储设备,出现了操作系统、汇编语言和一些高级语言。计算机不仅限于科学计算,还大量用于管理等,在操作系统中有专门的数据管理软件,称为文件系统,是数据库系统发展的初级阶段,并非真正的数据库系统。

1. 文件系统管理数据的特点

(1) 数据可长期保存。主要以文件形式在计算机中保存各种数据,如电子表格数据。

(2) 数据共享性差。在文件系统中,文件仍然面向应用,当不同文件具有相同数据时,须建立各自的文件,而不能共享数据,致使数据的冗余度大,浪费储存空间。

(3) 数据的独立性弱。软件带“数据包”,数据结构发生改变时,需修改应用程序和文件结构,应用程序的改变也会改变数据结构,文件系统仍是一无结构的数据集合。

(4) 具有简单的数据管理功能。由文件系统
进行数据管理,程序和数据之间有了一定的独立性,可减少程序员的工作量。在此阶段中应用和数据文件之间的关系如图 1-2 所示。

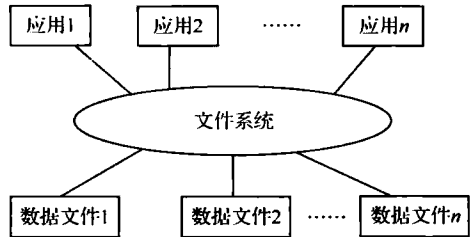


图 1-2 应用和数据文件间的关系

2. 文件系统的不足

随着数据管理规模的扩大,数据量急剧增加,文件系统的缺陷显现出来,主要表现为:

(1) 数据冗余大。由于数据文件之间缺乏联系,造成每个应用程序都有对应的数据文件,这样相同的数据会在多个数据文件中重复存储。

(2) 数据不一致。由于数据冗余,在进行更新操作时,若不注意,就容易使相同的数据在多个数据文件中不一致。如修改时遗漏或造成错改数据。

(3) 数据联系弱。文件之间相互独立却缺乏必要的关联,影响了数据管理。

1.2.3 数据库管理阶段

1969 年 IBM 公司研制的层次模型的数据库管理系统 IMS,以及 20 世纪 70 年代美国数据库系统语言协商 CODASYL 下属数据库任务组 DBTG 构建的网状模型,是第一代数据库的代表。数据库技术的发展先后经历了第一代的网状和层次数据库系统,第二代的关系数据库系统,第三代的以面向对象模型为主要特征的数据库系统。

从 20 世纪 60 年代中期以来,随着计算机硬软件技术的快速发展,CPU 向超大规模集成电路发展,为存储和处理大数据量的数据库给予极大的技术支持。同时操作系统得到了发展,而且各种 DBMS 软件不断涌现,使得数据库管理技术不断发展和完善,成为计算机领域中最具影响力和发展潜力、应用范围最广、成果最显著的技术之一,形成了“数据库时代”。数据库系统建立了数据之间的有机联系,实现了统一、集中、独立地管理数据和数据共享。数据库管理阶段的主要特点包括:

(1) 数据的集成性强。数据库系统中采用统一的数据结构方式,使数据结构化;全局的数据结构由多个应用程序共用,各程序调用局部结构的数据,全局与局部的结构模式构成数据集成。

(2) 数据高度共享冗余低。数据库系统从整体角度看待和描述数据,数据面向整个系统不再面向某个应用,因此,数据可以被多用户、多应用所共享。数据库技术与网络技术的结合扩大应用。数据共享程度极大地减少数据的冗余度,节约存储空间,又能避免数据之间的不相容性和不一致性。

(3) 数据独立性高。数据的独立性是指用户的应用程序与数据库中数据是相互独立的,即当数据的物理结构和逻辑结构发生变化时,不影响应用程序对数据的使用。

(4) 数据统一进行管理和控制。数据库管理系统对所有数据统一进行管理和控制,保证了数据的安全性和完整性。数据库系统对访问用户身份及其操作的合法性进行检查,自动检查数据的一致性、相容性,保证数据符合完整性约束条件,以并发控制手段有效控制多用户程序同时对数据操作,保证共享及并发操作,恢复功能保障当数据库遭到破坏时能自动恢复到正确状态。

1.2.4 高级数据库管理阶段

从20世纪80年代以后,数据库技术在商业领域取得巨大成功,激发了其他领域对其需求的快速增长。数据库技术新的应用领域的研究,极大地推动了数据库技术,特别是面向对象数据库系统的研究和发展。同时不断与其他技术结合,向高级数据库技术发展。

1. 分布式数据库技术

随着地理上分散的用户对数据共享的需求和网络技术的发展,产生了分布式数据库系统。它具有如下5个主要特点:

(1) 大部分数据在本地进行分布处理,提高了系统处理效率和可靠性。数据复制技术是分布式数据库的重要技术。

(2) 解决了中心数据库的不足,减少了数据传输代价。数据库的数据物理上分布在各地,逻辑上是个相互联系的整体,从而实现了数据的物理分布性和逻辑整体性。

(3) 提高系统的可靠性,局部系统发生故障,其他部分仍可继续工作。

(4) 各地终端由数据通信网络相连。本地终端单独不能胜任的处理任务,可以通过通信网络取得其他数据库和终端的支持。

(5) 数据库位置透明,方便系统扩充。从用户看整个数据库仍是个集中的数据库,用户不用关心数据的分片、物理位置分布和副本的一致性,由分布式DBMS实现分布。

分布式数据库系统兼顾集中管理和分布处理两项任务,因而具有良好的性能,其具体结构如图1-3所示。详细内容将在第10章进行介绍。

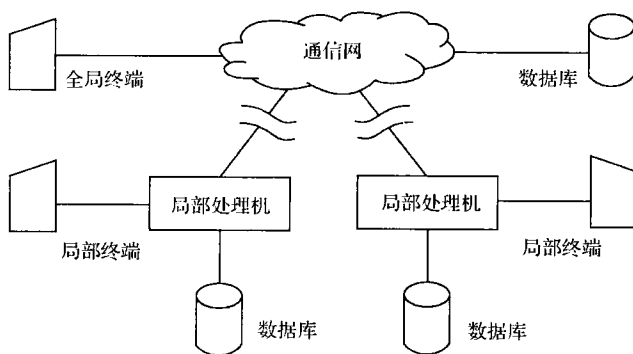


图 1-3 分布式数据库

2. 面向对象数据库技术

对于许多复杂数据结构的实际应用领域,层次、网状和关系三种模型难以满足需要。例如,多媒体数据、多维表格数据、CAD数据等应用问题,都需要更高级的数据库技术来表达,以便于管理、构造与维护大容量的持久数据,并与大型复杂程序紧密结合,将面向对象的程序设

计技术与数据库技术相结合产生了面向对象数据库技术。它主要有两个特点：

(1) 对象数据模型能完整地描述现实世界的数据结构,表达数据间嵌套、递归的联系。

(2) 具有面向对象技术的封装性(数据与操作定义一起)和继承性(继承数据结构和操作)的特点,提高了软件的可重用性。

3. 面向应用领域的数据库技术

数据库技术经过几十年的发展,形成完善的理论体系和实用技术。为了适应应用多元化的需求,结合各应用领域的特点,将数据库技术应用到特定领域,产生了工程数据库、地理数据库、统计数据库、科学数据库、空间数据库等多种数据库,同时也出现了数据仓库和数据挖掘等技术,使数据库领域中的新技术不断涌现。

最新的 SQL Server 2012 实现了一个为云做好准备的信息平台,可以有效解决日益增加的数据量带来的挑战,以帮助用户管理任何大小、任何地方(本地或云端中)的任何数据。通过提供的数据平台和工具,用户可以提取更有价值的数 据,从而作出有效决策。

SQL Server 2012 与以前版本相比,更具扩展性、更可靠,并提供了更高的性能。另外,还包含了 Power View 功能,微软称“这是行业领先的商业智能功能”,该功能通过强大的交互能力,可以将用户对于任何地方、任何数据的探索,转变成一种更加自然、轻松 的体验。此外,通过 Windows Azure 中的一项基于 Apache Hadoop 的服务,可以连接 SQL Server 并整合商业智能工具。

* 1.2.5 数据库技术的发展趋势

根据数据库应用及多家分析机构的评估,数据库技术发展将会以应用为导向,面向业务服务,并且为新型应用提供多种支持。

1. 混合数据快速发展

在 SQL Server 2012、DB2 Viper 和 Oracle 11g 都很重视软件产品的可扩展标记语言 XML (extensible markup language)特性,数据应用的主要开发平台也将转换到 XML 化的操作语 义。随着服务组件体系结构 SCA(service component architecture)和多种新型 Web 应用的普 及,XML 数据库将完成一个从文档到数据的转变。同时,“XML 数据/对象实体”的映射技术 也将得到广泛应用。

2. 数据集成与数据仓库倾向内容管理

数据仓库(data warehouse)是在企业管理和决策中一个面向主题的、集成的、相对稳定的、 反映历史变化的数据集 合,是决策支持系统和联机分析应用数据源的结构化数据环境。是一 门新兴的、快速发展的信息技术,主要侧重于对企业已有历史数据的综合分析利用,从中找出 对企业发展有用的信息,为企业 提供决策支持,帮助企业创造收入、提高利润。研究和解决从 数据库中获取信息的问题。其特征 在于面向主题、集成性、稳定性和时变性。用于支持管理决 策。新一代数据库的出现,使数据集成和数据仓库的实施更简单,连续处理、准实时处理和小 范围数据处理都将成为数据集成和分析人员所面临的新课题。随着数据应用逐步过渡到数 据服务,还应着重处理 3 个问题:关系型与非关系型数据的融合、数据分类、国际化多语言数 据。数据仓库和数据挖掘技术将在第 10 章具体介绍。

3. 主数据管理

在企业内部的应用整合和系统互联中,许多企业具有相同业务语义的数据被反复定义和