

<<数据库系统概念>>

图书基本信息

书名：<<数据库系统概念>>

13位ISBN编号：9787111375296

10位ISBN编号：7111375297

出版时间：2012-3

出版时间：机械工业出版社

作者：（美）Abraham Silberschatz,（美）Henry F.Korth,（美）S.Sudarshan

页数：805

译者：杨冬青,李红燕,唐世渭

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## &lt;&lt;数据库系统概念&gt;&gt;

## 前言

前言：Database System Concepts, 6E数据库管理已经从一种专门的计算机应用发展为现代计算环境中的一个重要成分，因此，有关数据库系统的知识已成为计算机科学教育中的一个核心的部分。

在本书中，我们讲述数据库管理的基本概念。

这些概念包括数据库设计、数据库语言、数据库系统实现等多个方面。

本书可作为本科生三年级或四年级数据库入门课程的教科书，也可作为研究生一年级的教科书。

除了作为入门课程的基本内容外，本书还包括了可作为课程补充或作为高级课程介绍性材料的高级内容。

我们仅要求读者熟悉基本的数据结构、计算机组织结构和一种高级程序设计语言，例如Java、C或Pascal。

书中的概念都以直观的方式加以描述，其中的许多概念都基于我们大学运行的例子加以阐释。

本书中包括重要的理论结果，但省略了形式化证明，取而代之的是用图表和例子来说明为什么结论是正确的。

对于形式化描述和研究结果的证明，读者可以参考文献注解中列出的研究论文和高级教材。

本书中所包括的基本概念和算法通常是基于当今的商品化或试验性的数据库系统中采用的概念和算法。

我们的目标是在一个通常环境下描述这些概念和算法，而没有与某个特定的数据库系统绑定。

特定的数据库系统的细节将在第九部分“实例研究”中讨论。

在本书第6版中，我们保持了前面版本的总体风格，同时对内容和结构进行了扩展来反映数据库设计、管理和使用的方式所发生的变化。

我们还考虑了数据库概念的教学方面的趋势，并在适当的地方做出了推动这种趋势的修改。

本书的组织本书组织成十个主要部分：综述（第1章）。

第1章对数据库系统的性质和目标进行了一般性综述。

我们解释了数据库系统的概念是如何发展的，各数据库系统的共同特性是什么，数据库系统能为用户做什么，以及数据库系统如何与操作系统交互。

我们还引入了一个数据库应用的例子：包括多个系、教员、学生和课程的一个大学机构。

这个应用作为贯穿全书的运行实例。

这一章本质上是诱导性、历史性和解释性的。

第一部分：关系数据库（第2章至第6章）。

第2章介绍了数据的关系模型，包括基本概念，诸如关系数据库的结构、数据库模式、码、模式图、关系查询语言和关系操作等。

第3~5章主要介绍最具影响力的面向用户的关系语言——SQL。

第6章介绍形式化的关系查询语言，包括关系代数、元组关系演算和域关系演算。

这部分描述了数据操纵，包括查询、修改、插入和删除（假设已有一个模式设计）。

关于模式设计的问题延迟到第二部分讲述。

第二部分：数据库设计（第7章至第9章）。

第7章给出了数据库设计过程的概要介绍，主要侧重于用实体-联系数据模型来进行数据库设计。

实体-联系模型为数据库设计问题，以及我们在数据模型的约束下捕获现实应用的语义时所遇到的问题提供了一个高层视图。

UML类图表示也在这一章中讲述。

第8章介绍关系数据库设计理论。

这一章讲述函数依赖和规范化，重点强调提出各种范式的动机，以及它们的直观含义。

这一章以关系设计的概览开始，依赖于对函数依赖的逻辑蕴涵的直观理解。

这使得规范化的概念可以在函数依赖理论的完整内容之前先作介绍。

函数依赖理论将在本章中稍后部分讨论。

教师可以只选用8.1节至8.3节这些较前面的章节，而不会丢失连贯性。

## &lt;&lt;数据库系统概念&gt;&gt;

不过，完整地讲授这一章将有利于学生对规范化概念形成较好的理解，从而诱导出函数依赖理论中一些较艰深的概念。

第9章讲述应用设计和开发。

这一章侧重于用基于Web的界面构建数据库应用。

另外，这一章还讲述了应用安全性。

第三部分：数据存储和查询（第10章至第13章）。

第10章讨论存储设备、文件和数据存储结构。

在第11章中介绍多种数据存取技术，包括B+树索引和散列。

第12章和第13章阐述查询执行算法和查询优化。

这两章使用户能更好地理解数据库的存储和检索的内部机制。

第四部分：事务管理（第14章至第16章）。

第14章着重介绍事务处理系统的基本概念，包括原子性、一致性、隔离性和持久性。

它还提供了用于保证这些特性的方法的一个概述，包括封锁和快照隔离性。

第15章重点讲述并发控制，并介绍保证可串行化的几种技术，包括封锁、时间戳和乐观（有效性检查）技术。

在这一章中还讨论死锁问题，并介绍保证可串行化的其他方法，特别是详细讨论广泛使用的快照隔离方法。

第16章讨论在系统崩溃和存储器故障情况下保证事务正确执行的主要技术。

这些技术包括日志、检查点和数据库转储。

被广泛使用的ARIES算法也在这里做了介绍。

第五部分：系统体系结构（第17章至第19章）。

第17章介绍计算机系统体系结构，并描述了作为基础的计算机系统对于数据库系统的影响。

在这一章中讨论了集中式系统、客户-服务器系统、并行和分布式体系结构。

在第18章关于并行数据库的讨论中，我们探讨了各种并行技术，包括I/O并行、查询间并行和查询内并行，以及操作间并行和操作内并行。

这一章中还讨论了并行系统设计。

第19章讨论分布式数据库系统，在分布式数据库系统的环境下重新讨论数据库设计、事务管理、查询执行和优化问题。

这一章还包括了故障时的系统可用性问题，并介绍了异构分布式数据库、基于云的数据库和分布式目录系统。

第六部分：数据仓库、数据挖掘与信息检索（第20章和第21章）。

第20章介绍数据仓库和数据挖掘的概念。

第21章描述用于查询文本数据的信息检索技术，包括在Web搜索引擎中使用的基于超链接的技术。

第六部分使用了第一部分和第二部分的模型和语言概念，但并不依赖于第三部分、第四部分或第五部分。

因此它可以很容易地结合到侧重于SQL和数据库设计的课程中。

第七部分：特种数据库（第22章和第23章）。

第22章介绍基于对象的数据库。

该章讲述了对象-关系数据模型，该模型扩展了关系数据模型以支持复杂数据类型、类型继承和对象标识。

该章还描述了用面向对象的编程语言来访问数据库。

第23章介绍数据表示的XML标准，它正日益广泛地应用于复杂数据交换和存储。

这一章还描述了XML的查询语言。

第八部分：高级主题（第24章至第26章）。

第24章讨论应用开发中的高级话题，包括性能调整、性能评测标准、数据库应用测试和标准化。

第25章介绍空间和地理数据、时间数据、多媒体数据以及移动和个人数据库管理中的问题。

最后，第26章讨论高级事务处理。

## <<数据库系统概念>>

这一章的内容包括事务处理监控器、事务 workflow、电子商务、高性能事务系统、实时事务系统和持续长时间的事务。

第九部分：实例研究（第27章至第30章）。

在这一部分我们对四个领先的数据库系统进行实例研究，包括 PostgreSQL、Oracle、IBM DB2 和 Microsoft SQL Server。

这几章中列举了上述每一种系统的独有特性，描述了它们的内部结构，提供了关于各个产品的丰富的有用信息，帮助读者了解前面各部分描述的各种实现技术是如何使用到实际系统中的。

这几章中还包括实际系统设计中的几个有趣的方面。

第十部分：附录（附录A ~ 附录E）。

我们提供5个附录，包括一些历史性的和高级的内容；这些附录只在本书的Web站点（<http://www.db-book.com>）中联机提供。

只有附录A（详细的大学模式）例外，它给出了我们的大学模式的细节，包括完整的模式、DDL和所有的表。

这个附录出现在纸质版本中。

附录B（高级关系数据库设计）描述了高级关系数据库设计，包括多值依赖理论、连接依赖、投影连接和域-码范式。

这个附录是为希望更详细地研究关系数据库设计理论的读者，以及希望在课程中这样做的教师准备的。

。

这个附录同样只是联机提供，就在本书的网站上。

附录C（其他关系查询语言）描述其他的关系查询语言，包括 QBE Microsoft Access 和 Datalog。

虽然大多数新的数据库应用系统使用关系模型或对象-关系模型，但网状的和层次的数据模型在一些遗留应用中也仍然在使用。

为了满足希望了解这些数据模型的读者的需要，我们给出了描述网状和层次的数据模型的附录，分别为附录D（网状模型）和附录E（层次模型）。

## <<数据库系统概念>>

### 内容概要

本书是经典的数据库系统教科书《Database System Concepts》的最新修订版，全面介绍数据库系统的各种知识，透彻阐释数据库管理的基本概念。本书内容丰富，不仅讨论了关系数据模型和关系语言、数据库设计过程、关系数据库理论、数据库应用设计和开发、数据存储结构、数据存取技术、查询优化方法、事务处理系统和并发控制、故障恢复技术、数据仓库和数据挖掘，而且对性能调整、性能评测标准、数据库应用测试和标准化、空间和地理数据、时间数据、多媒体数据、移动和个人数据库管理以及事务处理监控器、事务 workflow、电子商务、高性能事务系统、实时事务系统和持续长时间的事务等高级应用主题进行了广泛讨论。

本书既可作为高年级本科生或低年级研究生的数据库课程教材，也可供数据库领域的技术人员参考。

## <<数据库系统概念>>

### 作者简介

Abraham Silberschatz

于纽约州立大学石溪分校获得博士学位，现为耶鲁大学计算机科学Sidney J. Weinberg教授，计算机科学系主任，曾任贝尔实验室信息科学研究中心副主任。

他是ACM Fellow 和 IEEE

Fellow，曾获得IEEE Taylor L. Booth 教育奖、ACM Karl V. Karlstrom

杰出教育者奖、ACM SIGMOD 贡献奖和IEEE

计算机学会杰出论文奖。

他的研究兴趣包括操作系统、数据库系统、存储系统、网络管理和分布式系统。

Henry F. Korth

于普林斯顿大学获得博士学位，现为利哈伊大学计算机科学与工程系Weiseman教授，曾任贝尔实验室数据库原理研究中心主任。

他是ACM

Fellow 和 IEEE Fellow，是VLDB

10年贡献奖的获得者。

他的研究兴趣包括为现代计算架构（多核、多线程、多级缓存）设计的数据库算法、基于Web的大型数据仓储、实时数据库系统和并行系统。

S. Sudarshan

于威斯康星大学麦迪逊分校获得博士学位，现为印度理工学院计算机科学与工程系教授，曾为贝尔实验室数据库研究组技术人员。

他的研究兴趣包括查询处理和优化、关系数据和图结构数据的关键字查询，以及构建和测试数据库应用系统的工具。

译者序：

Database System Concepts, 6E

数据库系统是对数据进行存储、管理、处理和维护的软件系统，是现代计算环境中的一个核心成分。

随着计算机硬件、软件技术的飞速发展和计算机系统在各行各业的广泛应用，数据库技术的发展尤其迅速，引人注目。

有关数据库系统的理论和技术是计算机科学技术教育中必不可少的部分。

《数据库系统概念》是一本经典的、备受赞扬的数据库系统教科书。

其内容由浅入深，既包含数据库系统的基本概念，又反映数据库技术的新进展。

本书被国际上许多著名大学所采用，并多次再版。

我们先后将本书的第3版、第4版和第5版译成中文，由机械工业出版社分别于2000年、2003年和2006年出版发行。

国内许多大学采用 《数据库系统概念》作为本科生和研究生数据库课程的教材或主要教学参考书，收到了良好的效果。

现在，我们又翻译了该书第6版。

第6版保持了前5版的总体风格，同时对内容进行了扩充，对结构进行了调整，以更好地符合数据库教学的需求，反映数据库设计、管理和使用方式的发展和变化。

第6版的内容大体上可以分为五个部分。

第1~9章讲述数据库系统的基本概念，包括对数据库系统的性质和目标的综述，对关系数据模型和关



## <<数据库系统概念>>

系语言的介绍，对数据库设计过程、关系数据库理论以及数据库应用设计和开发（包括基于Web的界面构建数据库应用和应用安全性问题等）的详细讨论。

第10～19章主要讨论数据库系统实现技术，包括数据存储结构、数据存取技术、查询优化方法、事务处理系统的基本概念和并发控制、故障恢复技术，还包括在并行数据库系统和分布式数据库系统中所采用的一些主要策略和技术。

第20～23章主要讨论数据管理与应用的深入话题，包括对数据仓库和数据挖掘概念与技术的较详细的介绍，以及对用于查询文本数据的信息检索技术（包括在Web搜索引擎中使用的基于超链接的技术）的介绍。

这一部分还介绍了新型的数据库系统，包括对象-关系数据库模型、数据表示的XML标准，以及XML的查询语言。

第24～26章是一些高级话题，内容包括应用开发中的诸如性能调整、性能评测标准、数据库应用测试和标准化等高级话题，以及空间和地理数据、时间数据、多媒体数据、移动和个人数据库管理中的问题。

这一部分还讨论了事务处理监控器、事务 workflow、电子商务、高性能事务系统、实时事务系统和持续长时间的事务等高级事务处理问题。

第27～30章对PostgreSQL、Oracle、IBM DB2和Microsoft SQL Server这四个领先的数据库系统进行实例研究，结合这几个具体系统来讨论前面各部分描述的各种实现技术是如何使用到实际系统中的。

上述五大部分中第一部分的主要内容，以及第二、第三、第四部分的部分内容可以作为本科生数据库概论课程的教材或主要参考资料，第二、第三和第四部分的其余内容可以用于研究生的数据库课程教学，第五部分可以作为帮助学生了解实际系统的补充材料。

杨冬青、李红燕、唐世渭组织并参加了本书的翻译和审校工作；参加翻译的还有范红杰、程序、苗高杉、邹森、陈巍、王婧、王林青、孟必平。

限于译者水平，译文中疏漏和错误难免，欢迎批评指正。

译者  
于北京大学

## <<数据库系统概念>>

### 书籍目录

出版者的话

译者序

前言

#### 第1章 引言

1.1数据库系统的应用

1.2数据库系统的目标

1.3数据视图

1.3.1数据抽象

1.3.2实例和模式

1.3.3数据模型

1.4数据库语言

1.4.1数据操纵语言

1.4.2数据定义语言

1.5关系数据库

1.5.1表

1.5.2数据操纵语言

1.5.3数据定义语言

1.5.4来自应用程序的数据库访问

1.6数据库设计

1.6.1设计过程

1.6.2大学机构的数据库设计

1.6.3实体-联系模型

1.6.4规范化

1.7数据存储和查询

1.7.1存储管理器

1.7.2查询处理器

1.8事务管理

1.9数据库体系结构

1.10数据挖掘与信息检索

1.11特种数据库

1.11.1基于对象的数据模型

1.11.2半结构化数据模型

1.12数据库用户和管理员

1.12.1数据库用户和用户界面

1.12.2数据库管理员

1.13数据库系统的历史

1.14总结

术语回顾

实践习题

习题

工具

文献注解

#### 第一部分 关系数据库

#### 第2章 关系模型介绍

2.1关系数据库的结构



## <<数据库系统概念>>

2.2数据库模式

2.3码

2.4模式图

2.5关系查询语言

2.6关系运算

2.7总结

术语回顾

实践习题

习题

文献注解

### 第3章 SQL

3.1SQL查询语言概览

3.2SQL数据定义

3.2.1基本类型

3.2.2基本模式定义

3.3SQL查询的基本结构

3.3.1单关系查询

3.3.2多关系查询

3.3.3自然连接

3.4附加的基本运算

3.4.1更名运算

3.4.2字符串运算

3.4.3select子句中的属性说明

3.4.4排列元组的显示次序

3.4.5where子句谓词

3.5集合运算

3.5.1并运算

3.5.2交运算

3.5.3差运算

3.6空值

3.7聚集函数

3.7.1基本聚集

3.7.2分组聚集

3.7.3having子句

3.7.4对空值和布尔值的聚集

3.8嵌套子查询

3.8.1集合成员资格

3.8.2集合的比较

3.8.3空关系测试

3.8.4重复元组存在性测试

3.8.5from子句中的子查询

3.8.6with子句

3.8.7标量子查询

3.9数据库的修改

3.9.1删除

3.9.2插入

3.9.3更新

## &lt;&lt;数据库系统概念&gt;&gt;

## 3.10总结

术语回顾

实践习题

习题

工具

文献注解

## 第4章 中级SQL

## 4.1连接表达式

## 4.1.1连接条件

## 4.1.2外连接

## 4.1.3连接类型和条件

## 4.2视图

## 4.2.1视图定义

## 4.2.2SQL查询中使用视图

## 4.2.3物化视图

## 4.2.4视图更新

## 4.3事务

## 4.4完整性约束

## 4.4.1单个关系上的约束

## 4.4.2not null约束

## 4.4.3unique约束

## 4.4.4check子句

## 4.4.5参照完整性

## 4.4.6事务中对完整性约束的违反

## 4.4.7复杂check条件与断言

## 4.5SQL的数据类型与模式

## 4.5.1SQL中的日期和时间类型

## 4.5.2默认值

## 4.5.3创建索引

## 4.5.4大对象类型

## 4.5.5用户定义的类型

## 4.5.6create table的扩展

## 4.5.7模式、目录与环境

## 4.6授权

## 4.6.1权限的授予与收回

## 4.6.2角色

## 4.6.3视图的授权

## 4.6.4模式的授权

## 4.6.5权限的转移

## 4.6.6权限的收回

## 4.7总结

术语回顾

实践习题

习题

文献注解

## 第5章 高级SQL

## 5.1使用程序设计语言访问数据库

## &lt;&lt;数据库系统概念&gt;&gt;

5.1.1JDBC

5.1.2ODBC

5.1.3嵌入式SQL

5.2函数和过程

5.2.1声明和调用SQL函数和过程

5.2.2支持过程和函数的语言构造

5.2.3外部语言过程

5.3触发器

5.3.1对触发器的需求

5.3.2SQL中的触发器

5.3.3何时不用触发器

5.4递归查询\*\*

5.4.1用迭代来计算传递闭包

5.4.2SQL中的递归

5.5高级聚集特性\*\*

5.5.1排名

5.5.2分窗

5.6OLAP\*\*

5.6.1联机分析处理

5.6.2交叉表与关系表

5.6.3 SQL中的OLAP

5.7总结

术语回顾

实践习题

习题

工具

文献注解

## 第6章 形式化关系查询语言

6.1关系代数

6.1.1基本运算

6.1.2关系代数的形式化定义

6.1.3附加的关系代数运算

6.1.4扩展的关系代数运算

6.2元组关系演算

6.2.1查询示例

6.2.2形式化定义

6.2.3表达式的安全性

6.2.4语言的表达能力

6.3域关系演算

6.3.1形式化定义

6.3.2查询的例子

6.3.3表达式的安全性

6.3.4语言的表达能力

6.4总结

术语回顾

实践习题

习题

## &lt;&lt;数据库系统概念&gt;&gt;

文献注解

第二部分 数据库设计

第7章 数据库设计和E-R模型

7.1设计过程概览

7.1.1设计阶段

7.1.2设计选择

7.2实体-联系模型

7.2.1实体集

7.2.2联系集

7.2.3属性

7.3约束

7.3.1映射基数

7.3.2参与约束

7.3.3码

7.4从实体集中删除冗余属性

7.5实体 - 联系图

7.5.1基本结构

7.5.2映射基数

7.5.3复杂的属性

7.5.4角色

7.5.5非二元的联系集

7.5.6弱实体集

7.5.7大学的E-R图

7.6转换为关系模式

7.6.1具有简单属性的强实体集表示

7.6.2具有复杂属性的强实体集表示

7.6.3弱实体集表示

7.6.4联系集表示

7.7实体-联系设计问题

7.7.1用实体集还是用属性

7.7.2用实体集还是用联系集

7.7.3二元还是n元联系集

7.7.4联系属性的布局

7.8扩展的E-R特性

7.8.1特化

7.8.2概化

7.8.3属性继承

7.8.4概化上的约束

7.8.5聚集

7.8.6转换为关系模式

7.9数据建模的其他表示法

7.9.1E-R图的其他表示法

7.9.2统一建模语言UML

7.10数据库设计的其他方面

7.10.1数据约束和关系数据库设计

7.10.2使用需求：查询、性能

7.10.3授权需求

## &lt;&lt;数据库系统概念&gt;&gt;

7.10.4数据流、工作流

7.10.5数据库设计的其他问题

7.11总结

术语回顾

实践习题

习题

工具

文献注解

## 第8章 关系数据库设计

8.1好的关系设计的特点

8.1.1设计选择：更大的模式

8.1.2设计选择：更小的模式

8.2原子域和第一范式

8.3使用函数依赖进行分解

8.3.1码和函数依赖

8.3.2Boyce-Codd范式

8.3.3BCNF和保持依赖

8.3.4第三范式

8.3.5更高的范式

8.4函数依赖理论

8.4.1函数依赖集的闭包

8.4.2属性集的闭包

8.4.3正则覆盖

8.4.4无损分解

8.4.5保持依赖

8.5分解算法

8.5.1BCNF分解

8.5.23NF分解

8.5.33NF算法的正确性

8.5.4BCNF和3NF的比较

8.6使用多值依赖的分解

8.6.1多值依赖

8.6.2第四范式

8.6.34NF分解

8.7更多的范式

8.8数据库设计过程

8.8.1E-R模型和规范化

8.8.2属性和联系的命名

8.8.3为了性能去规范化

8.8.4其他设计问题

8.9时态数据建模

8.10总结

术语回顾

实践习题

习题

文献注解

## 第9章 应用设计和开发

## <<数据库系统概念>>

- 9.1应用程序和用户界面
  - 9.2Web基础
    - 9.2.1统一资源定位符
    - 9.2.2超文本标记语言
    - 9.2.3Web服务器和会话
  - 9.3servlet和JSP
    - 9.3.1一个servlet的例子
    - 9.3.2servlet会话
    - 9.3.3servlet的生命周期
    - 9.3.4servlet支持
    - 9.3.5服务器端脚本
    - 9.3.6客户端脚本
  - 9.4应用架构
    - 9.4.1业务逻辑层
    - 9.4.2数据访问层和对象-关系映射
    - 9.4.3Web服务
    - 9.4.4断连操作
  - 9.5快速应用开发
    - 9.5.1构建用户界面的工具
    - 9.5.2Web应用框架
    - 9.5.3报表生成器
  - 9.6应用程序性能
    - 9.6.1利用缓存减少开销
    - 9.6.2并行处理
  - 9.7应用程序安全性
    - 9.7.1SQL注入
    - 9.7.2跨站点脚本和请求伪造
    - 9.7.3密码泄露
    - 9.7.4应用程序认证
    - 9.7.5应用级授权
    - 9.7.6审计追踪
    - 9.7.7隐私
  - 9.8加密及其应用
    - 9.8.1加密技术
    - 9.8.2数据库中的加密支持
    - 9.8.3加密和认证
  - 9.9总结
  - 术语回顾
  - 实践习题
  - 习题
  - 项目建议
  - 工具
  - 文献注解
- 第三部分 数据存储和查询
- 第10章 存储和文件结构
    - 10.1物理存储介质概述
    - 10.2磁盘和快闪存储器



## &lt;&lt;数据库系统概念&gt;&gt;

- 10.2.1磁盘的物理特性
- 10.2.2磁盘性能的度量
- 10.2.3磁盘块访问的优化
- 10.2.4快闪存储
- 10.3RAID
- 10.3.1通过冗余提高可靠性
- 10.3.2通过并行提高性能
- 10.3.3RAID级别
- 10.3.4RAID级别的选择
- 10.3.5硬件问题
- 10.3.6其他的RAID应用
- 10.4第三级存储
- 10.4.1光盘
- 10.4.2磁带
- 10.5文件组织
- 10.5.1定长记录
- 10.5.2变长记录
- 10.6文件中记录的组织
- 10.6.1顺序文件组织
- 10.6.2多表聚簇文件组织
- 10.7数据字典存储
- 10.8数据库缓冲区
- 10.8.1缓冲区管理器
- 10.8.2缓冲区替换策略
- 10.9总结
- 术语回顾
- 实践习题
- 习题
- 文献注解

## 第11章 索引与散列

- 11.1基本概念
- 11.2顺序索引
- 11.2.1稠密索引和稀疏索引
- 11.2.2多级索引
- 11.2.3索引的更新
- 11.2.4辅助索引
- 11.2.5多码上的索引
- 11.3B+树索引文件
- 11.3.1B+树的结构
- 11.3.2B+树的查询
- 11.3.3B+树的更新
- 11.3.4不唯一的搜索码
- 11.3.5B+树更新的复杂性
- 11.4B+树扩展
- 11.4.1B+树文件组织
- 11.4.2辅助索引和记录重定位
- 11.4.3字符串上的索引

## &lt;&lt;数据库系统概念&gt;&gt;

- 11.4.4B+树索引的批量加载
  - 11.4.5B树索引文件
  - 11.4.6闪存
  - 11.5多码访问
    - 11.5.1使用多个单码索引
    - 11.5.2多码索引
    - 11.5.3覆盖索引
  - 11.6静态散列
    - 11.6.1散列函数
    - 11.6.2桶溢出处理
    - 11.6.3散列索引
  - 11.7动态散列
    - 11.7.1数据结构
    - 11.7.2查询和更新
    - 11.7.3静态散列与动态散列比较
  - 11.8顺序索引和散列的比较
  - 11.9位图索引
    - 11.9.1位图索引结构
    - 11.9.2位图操作的高效实现
    - 11.9.3位图和B+树
  - 11.10SQL中的索引定义
  - 11.11总结
  - 术语回顾
  - 实践习题
  - 习题
  - 文献注解
- 第12章 查询处理
- 12.1概述
  - 12.2查询代价的度量
  - 12.3选择运算
    - 12.3.1使用文件扫描和索引的选择
    - 12.3.2涉及比较的选择
    - 12.3.3复杂选择的实现
  - 12.4排序
    - 12.4.1外部排序归并算法
    - 12.4.2外部排序归并的代价分析
  - 12.5连接运算
    - 12.5.1嵌套循环连接
    - 12.5.2块嵌套循环连接
    - 12.5.3索引嵌套循环连接
    - 12.5.4归并连接
    - 12.5.5散列连接
  - 12.6其他运算
    - 12.6.1去除重复
    - 12.6.2投影
    - 12.6.3集合运算
    - 12.6.4外连接

## &lt;&lt;数据库系统概念&gt;&gt;

12.6.5 聚集

12.7 表达式计算

12.7.1 物化

12.7.2 流水线

12.8 总结

术语回顾

实践习题

习题

文献注解

第13章 查询优化

13.1 概述

13.2 关系表达式的转换

13.2.1 等价规则

13.2.2 转换的例子

13.2.3 连接的次序

13.2.4 等价表达式的枚举

13.3 表达式结果集统计大小的估计

13.3.1 目录信息

13.3.2 选择运算结果大小的估计

13.3.3 连接运算结果大小的估计

13.3.4 其他运算的结果集大小的估计

13.3.5 不同取值个数的估计

13.4 执行计划选择

13.4.1 基于代价的连接顺序选择

13.4.2 采用等价规则的基于代价的优化器

13.4.3 启发式优化

13.4.4 嵌套子查询的优化\*\*

13.5 物化视图\*\*

13.5.1 视图维护

13.5.2 增量的视图维护

13.5.3 查询优化和物化视图

13.5.4 物化视图和索引选择

13.6 查询优化中的高级话题\*\*

13.6.1 top-K 优化

13.6.2 连接极小化

13.6.3 更新的优化

13.6.4 多查询优化和共享式扫描

13.6.5 参数化查询优化

13.7 总结

术语回顾

实践习题

习题

文献注解

第四部分 事务管理

第14章 事务

14.1 事务概念

14.2 一个简单的事务模型

## &lt;&lt;数据库系统概念&gt;&gt;

- 14.3存储结构
  - 14.4事务原子性和持久性
  - 14.5事务隔离性
  - 14.6可串行化
  - 14.7事务隔离性和原子性
    - 14.7.1可恢复调度
    - 14.7.2无级联调度
  - 14.8事务隔离性级别
  - 14.9隔离性级别的实现
    - 14.9.1锁
    - 14.9.2时间戳
    - 14.9.3多版本和快照隔离
  - 14.10事务的SQL语句表示
  - 14.11总结
  - 术语回顾
  - 实践习题
  - 习题
  - 文献注解
- 第15章 并发控制
- 15.1基于锁的协议
    - 15.1.1锁
    - 15.1.2锁的授予
    - 15.1.3两阶段封锁协议
    - 15.1.4封锁的实现
    - 15.1.5基于图的协议
  - 15.2死锁处理
    - 15.2.1死锁预防
    - 15.2.2死锁检测与恢复
  - 15.3多粒度
  - 15.4基于时间戳的协议
    - 15.4.1时间戳
    - 15.4.2时间戳排序协议
    - 15.4.3Thomas写规则
  - 15.5基于有效性检查的协议
  - 15.6多版本机制
    - 15.6.1多版本时间戳排序
    - 15.6.2多版本两阶段封锁
  - 15.7快照隔离
    - 15.7.1更新事务的有效性检验步骤
    - 15.7.2串行化问题
  - 15.8插入操作、删除操作与谓词读
    - 15.8.1删除
    - 15.8.2插入
    - 15.8.3谓词读和幻象现象
  - 15.9实践中的弱一致性级别
    - 15.9.1二级一致性
    - 15.9.2游标稳定性

## &lt;&lt;数据库系统概念&gt;&gt;

15.9.3跨越用户交互的并发控制

15.10索引结构中的并发\*\*

15.11总结

术语回顾

实践习题

习题

文献注解

## 第16章 恢复系统

16.1故障分类

16.2存储器

16.2.1稳定存储器的实现

16.2.2数据访问

16.3恢复与原子性

16.3.1日志记录

16.3.2数据库修改

16.3.3并发控制和恢复

16.3.4事务提交

16.3.5使用日志来重做和撤销事务

16.3.6检查点

16.4恢复算法

16.4.1事务回滚

16.4.2系统崩溃后的恢复

16.5缓冲区管理

16.5.1日志记录缓冲

16.5.2数据库缓冲

16.5.3操作系统在缓冲区管理中的作用

16.5.4模糊检查点

16.6非易失性存储器数据丢失的故障

16.7锁的提前释放和逻辑undo操作

16.7.1逻辑操作

16.7.2逻辑undo日志记录

16.7.3有逻辑undo的事务回滚

16.7.4逻辑undo中的并发问题

16.8ARIES\*\*

16.8.1数据结构

16.8.2恢复算法

16.8.3其他特性

16.9远程备份系统

16.10总结

术语回顾

实践习题

习题

文献注解

## 第五部分 系统体系结构

### 第17章 数据库系统体系结构

17.1集中式与客户-服务器体系结构

17.1.1集中式系统

## &lt;&lt;数据库系统概念&gt;&gt;

- 17.1.2客户-服务器系统
  - 17.2服务器系统体系结构
    - 17.2.1事务服务器
    - 17.2.2数据服务器
    - 17.2.3基于云的服务器
  - 17.3并行系统
    - 17.3.1加速比和扩展比
    - 17.3.2互连网络
    - 17.3.3并行数据库体系结构
  - 17.4分布式系统
    - 17.4.1分布式数据库示例
    - 17.4.2实现问题
  - 17.5网络类型
    - 17.5.1局域网
    - 17.5.2广域网
  - 17.6总结
  - 术语回顾
  - 实践习题
  - 习题
  - 文献注解
- 第18章 并行数据库
- 18.1引言
  - 18.2I/O并行
    - 18.2.1划分技术
    - 18.2.2划分技术比较
    - 18.2.3偏斜处理
  - 18.3查询间并行
  - 18.4查询内并行
  - 18.5操作内并行
    - 18.5.1并行排序
    - 18.5.2并行连接
    - 18.5.3其他关系运算
    - 18.5.4运算的并行计算代价
  - 18.6操作间并行
    - 18.6.1流水线并行
    - 18.6.2独立并行
  - 18.7查询优化
  - 18.8并行系统设计
  - 18.9多核处理器的并行性
    - 18.9.1并行性与原始速度
    - 18.9.2高速缓冲存储器和多线程
    - 18.9.3适应现代体系架构的数据库系统设计
  - 18.10总结
  - 术语回顾
  - 实践习题
  - 习题
  - 文献注解



## &lt;&lt;数据库系统概念&gt;&gt;

## 第19章 分布式数据库

## 19.1同构和异构数据库

## 19.2分布式数据存储

## 19.2.1数据复制

## 19.2.2数据分片

## 19.2.3透明性

## 19.3分布式事务

## 19.3.1系统结构

## 19.3.2系统故障模式

## 19.4提交协议

## 19.4.1两阶段提交

## 19.4.2三阶段提交

## 19.4.3事务处理的可选择性模型

## 19.5分布式数据库中的并发控制

## 19.5.1封锁协议

## 19.5.2时间戳

## 19.5.3弱一致性级别的复制

## 19.5.4死锁处理

## 19.6可用性

## 19.6.1基于多数的方法

## 19.6.2读一个、写所有可用的方法

## 19.6.3站点重建

## 19.6.4与远程备份的比较

## 19.6.5协调器的选择

## 19.6.6为可用性而牺牲一致性

## 19.7分布式查询处理

## 19.7.1查询转换

## 19.7.2简单的连接处理

## 19.7.3半连接策略

## 19.7.4利用并行性的连接策略

## 19.8异构分布式数据库

## 19.8.1数据统一视图

## 19.8.2查询处理

## 19.8.3多数据库中的事务管理

## 19.9基于云的数据库

## 19.9.1云上的数据存储系统

## 19.9.2云上的传统数据库

## 19.9.3基于云的数据库的挑战

## 19.10目录系统

## 19.10.1目录访问协议

## 19.10.2LDAP：轻量级目录访问协议

## 19.11总结

## 术语回顾

## 实践习题

## 习题

## 文献注解

## 第六部分 数据仓库、数据挖掘与信息检索

## &lt;&lt;数据库系统概念&gt;&gt;

## 第20章 数据仓库与数据挖掘

20.1 决策支持系统

20.2 数据仓库

20.2.1 数据仓库成分

20.2.2 数据仓库模式

20.2.3 面向列的存储

20.3 数据挖掘

20.4 分类

20.4.1 决策树分类器

20.4.2 其他类型的分类器

20.4.3 回归

20.4.4 分类器验证

20.5 关联规则

20.6 其他类型的关联

20.7 聚类

20.8 其他类型的数据挖掘

20.9 总结

术语回顾

实践习题

习题

工具

文献注解

## 第21章 信息检索

21.1 概述

21.2 使用术语的相关性排名

21.2.1 使用TF-IDF的排名方法

21.2.2 基于相似性的检索

21.3 使用超链接的相关性

21.3.1 流行度排名

21.3.2 PageRank

21.3.3 其他的流行度度量

21.3.4 搜索引擎作弊

21.3.5 将TF-IDF和流行度排名度量方法结合

21.4 同义词、多义词和本体

21.5 文档的索引

21.6 检索的有效性度量

21.7 Web的抓取和索引

21.8 信息检索：网页排名之外

21.8.1 查询结果的多样化

21.8.2 信息抽取

21.8.3 问答系统

21.8.4 查询结构化数据

21.9 目录与分类

21.10 总结

术语回顾

实践习题

习题

## &lt;&lt;数据库系统概念&gt;&gt;

工具

文献注解

## 第七部分 特种数据库

## 第22章 基于对象的数据库

22.1概述

22.2复杂数据类型

22.3SQL中的结构类型和继承

22.3.1结构类型

22.3.2类型继承

22.4表继承

22.5SQL中的数组和多重集合类型

22.5.1创建和访问集合体值

22.5.2查询以集合体为值的属性

22.5.3嵌套和解除嵌套

22.6SQL中的对象标识和引用类型

22.7 O-R特性的实现

22.8持久化程序设计语言

22.8.1对象的持久化

22.8.2对象标识和指针

22.8.3持久对象的存储和访问

22.8.4持久化C++系统

22.8.5持久化Java系统

22.9对象-关系映射

22.10面向对象与对象-关系

22.11总结

术语回顾

实践习题

习题

工具

文献注解

## 第23章 XML

23.1动机

23.2XML数据结构

23.3XML文档模式

23.3.1文档类型定义

23.3.2XML Schema

23.4查询和转换

23.4.1XML树模型

23.4.2XPath

23.4.3XQuery

23.5XML应用程序接口

23.6XML数据存储

23.6.1非关系的数据存储

23.6.2关系数据库

23.6.3SQL/XML

23.7XML应用

23.7.1存储复杂结构数据

## <<数据库系统概念>>

23.7.2 标准化数据交换格式

23.7.3 Web服务

23.7.4 数据中介

23.8 总结

术语回顾

实践习题

习题

工具

文献注解

### 第八部分 高级主题

#### 第24章 高级应用开发

24.1 性能调整

24.1.1 提高面向集合的特性

24.1.2 批量加载和更新的调整

24.1.3 瓶颈位置

24.1.4 可调参数

24.1.5 硬件调整

24.1.6 模式调整

24.1.7 索引调整

24.1.8 使用物化视图

24.1.9 物理设计的自动调整

24.1.10 并发事务调整

24.1.11 性能模拟

24.2 性能基准程序

24.2.1 任务集

24.2.2 数据库应用类型

24.2.3 TPC基准程序

24.3 应用系统开发的其他问题

24.3.1 应用系统测试

24.3.2 应用系统移植

24.4 标准化

24.4.1 SQL标准

24.4.2 数据库连接标准

24.4.3 对象数据库标准

24.4.4 基于XML的标准

24.5 总结

术语回顾

实践习题

习题

文献注解

#### 第25章 时空数据和移动性

25.1 动机

25.2 数据库中的时间

25.2.1 SQL中的时间规范

25.2.2 时态查询语言

25.3 空间与地理数据

25.3.1 几何信息表示

## &lt;&lt;数据库系统概念&gt;&gt;

- 25.3.2设计数据库
- 25.3.3地理数据
- 25.3.4空间查询
- 25.3.5空间数据的索引
- 25.4多媒体数据库
  - 25.4.1多媒体数据格式
  - 25.4.2连续媒体数据
  - 25.4.3基于相似性的检索
- 25.5移动性和个人数据库
  - 25.5.1移动计算模型
  - 25.5.2路由和查询处理
  - 25.5.3广播数据
  - 25.5.4连接断开与一致性
- 25.6总结
- 术语回顾
- 实践习题
- 习题
- 文献注解

## 26章 高级事务处理

- 26.1事务处理监控器
  - 26.1.1TP监控器体系结构
  - 26.1.2使用TP监控器进行应用协调
- 26.2事务 workflow
  - 26.2.1 workflow 说明
  - 26.2.2 workflow 的故障原子性需求
  - 26.2.3 workflow 执行
  - 26.2.4 workflow 恢复
  - 26.2.5 workflow 管理系统
- 26.3电子商务
  - 26.3.1电子目录
  - 26.3.2市场
  - 26.3.3订单结算
- 26.4主存数据库
- 26.5实时事务系统
- 26.6长事务
  - 26.6.1不可串行化的执行
  - 26.6.2并发控制
  - 26.6.3嵌套事务和多级事务
  - 26.6.4补偿事务
  - 26.6.5实现问题
- 26.7总结
- 术语回顾
- 实践习题
- 习题
- 文献注解

## 第九部分 实例研究

## 第27章 PostgreSQL

## &lt;&lt;数据库系统概念&gt;&gt;

## 27.1概述

## 27.2用户界面

## 27.2.1交互式终端界面

## 27.2.2图形界面

## 27.2.3编程语言接口

## 27.3SQL变化和扩展

## 27.3.1PostgreSQL类型

## 27.3.2规则和其他主动数据库特征

## 27.3.3可扩展性

## 27.4PostgreSQL中的事务管理

## 27.4.1PostgreSQL的并发控制

## 27.4.2恢复

## 27.5存储和索引

## 27.5.1表

## 27.5.2索引

## 27.6查询处理和优化

## 27.6.1查询重写

## 27.6.2查询规划和优化

## 27.6.3查询执行器

## 27.6.4触发器和约束

## 27.7系统结构

## 文献注解

## 第28章 Oracle

## 28.1数据库设计和查询工具

## 28.1.1数据库和应用设计工具

## 28.1.2查询工具

## 28.2SQL的变化和扩展

## 28.2.1对象-关系特性

## 28.2.2Oracle XML DB

## 28.2.3过程化语言

## 28.2.4维度

## 28.2.5联机分析处理

## 28.2.6触发器

## 28.3存储和索引

## 28.3.1表空间

## 28.3.2段

## 28.3.3表

## 28.3.4索引

## 28.3.5位图索引

## 28.3.6基于函数的索引

## 28.3.7连接索引

## 28.3.8域索引

## 28.3.9划分

## 28.3.10物化视图

## 28.4查询处理和优化

## 28.4.1执行方法

## 28.4.2优化



## &lt;&lt;数据库系统概念&gt;&gt;

- 28.4.3并行执行
- 28.4.4结果高速缓存
- 28.5并发控制与恢复
  - 28.5.1并发控制
  - 28.5.2恢复的基本结构
  - 28.5.3Oracle数据卫士
- 28.6系统体系结构
  - 28.6.1专用服务器：内存结构
  - 28.6.2专用服务器：进程结构
  - 28.6.3共享服务器
  - 28.6.4Oracle Real Application Clusters
  - 28.6.5自动存储管理器
  - 28.6.6Oracle Exadata
- 28.7复制、分布以及外部数据
  - 28.7.1复制
  - 28.7.2分布式数据库
  - 28.7.3外部数据源
- 28.8数据库管理工具
  - 28.8.1Oracle企业管理器
  - 28.8.2自动工作负载存储
  - 28.8.3数据库资源管理
- 28.9数据挖掘
- 文献注解

## 第29章 IBM DB2 Universal Database

- 29.1概述
- 29.2数据库设计工具
- 29.3SQL的变化和扩展
  - 29.3.1XML特性
  - 29.3.2数据类型的支持
  - 29.3.3用户自定义函数和方法
  - 29.3.4大对象
  - 29.3.5索引扩展和约束
  - 29.3.6Web服务
  - 29.3.7其他特性
- 29.4存储和索引
  - 29.4.1存储体系结构
  - 29.4.2缓冲池
  - 29.4.3表、记录和索引
- 29.5多维聚簇
  - 29.5.1块索引
  - 29.5.2块映射
  - 29.5.3设计考虑
  - 29.5.4对现有技术的影响
- 29.6查询处理和优化
  - 29.6.1存取方法
  - 29.6.2连接、聚集和集合运算
  - 29.6.3对复杂SQL处理的支持

## &lt;&lt;数据库系统概念&gt;&gt;

29.6.4多处理器查询处理特性

29.6.5查询优化

29.7物化的查询表

29.7.1查询路由到MQT

29.7.2MQT的维护

29.8DB2中的自治特性

29.8.1配置

29.8.2优化

29.9工具和实用程序

29.10并发控制和恢复

29.10.1并发与隔离

29.10.2提交与回滚

29.10.3日志与恢复

29.11系统体系结构

29.12复制、分布和外部数据

29.13商务智能特性

文献注解

### 第30章 Microsoft SQL Server

30.1管理、设计和查询工具

30.1.1数据库开发和可视化数据库工具

30.1.2数据库查询和调优工具

30.1.3SQL Server Management Studio

30.2SQL变化和扩展

30.2.1数据类型

30.2.2查询语言增强

30.2.3例程

30.2.4带过滤的索引

30.3存储和索引

30.3.1文件组

30.3.2文件组内的空间管理

30.3.3表

30.3.4索引

30.3.5分区

30.3.6在线创建索引

30.3.7扫描和预读

30.3.8压缩

30.4查询处理和优化

30.4.1编译处理概述

30.4.2查询简化

30.4.3重排序和基于代价的优化

30.4.4更新计划

30.4.5优化时的数据分析

30.4.6部分搜索和启发式搜索

30.4.7查询执行

30.5并发与恢复

30.5.1事务

30.5.2封锁

<<数据库系统概念>>

30.5.3恢复和可用性

30.6系统体系结构

30.6.1服务器上的线程池

30.6.2内存管理

30.6.3安全性

30.7数据访问

30.8分布式异构查询处理

30.9复制

30.9.1复制模型

30.9.2复制选项

30.10.NET中的服务器编程

30.10.1.NET基本概念

30.10.2SQL CLR宿主

30.10.3可扩展性协定

30.11XML支持

30.11.1本地存储和组织XML

30.11.2查询和更新XML数据类型

30.11.3XQuery表达式的执行

30.12SQL Server服务代理

30.13商务智能

30.13.1SQL Server集成服务

30.13.2SQL Server分析服务

30.13.3SQL Server报表服务

文献注解

第十部分 附录

附录A详细的大学模式

参考文献

索引

## <<数据库系统概念>>

### 章节摘录

版权页:第1章引言数据库管理系统 ( DataBase-Management System , DBMS ) 由一个互相关联的数据的集合和一组用以访问这些数据的程序组成。

这个数据集合通常称作数据库 ( database ) , 其中包含了关于某个企业的信息。

DBMS的主要目标是要提供一种可以方便、高效地存取数据库信息的途径。

设计数据库系统的目的是为了管理大量信息。

对数据的管理既涉及信息存储结构的定义, 又涉及信息操作机制的提供。

此外, 数据库系统还必须提供所存储信息的安全性保证, 即使在系统崩溃或有人企图越权访问时也应保障信息的安全性。

如果数据将被多用户共享, 那么系统还必须设法避免可能产生的异常结果。

在大多数组织中信息是非常重要的, 因而计算机科学家开发了大量的用于有效管理数据的概念和技术

。

这些概念和技术正是本书所关注的。

在这一章里, 我们将简要介绍数据库系统的基本原理。

1.1 数据库系统的应用数据库的应用非常广泛, 以下是一些具有代表性的应用: 企业信息 销售:用于存储客户、产品和购买信息。

<<数据库系统概念>>

编辑推荐

<<数据库系统概念>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>