

数据库技术

大数据业务

云计算环境

NoSQL数据库

Q

数据库

✕

数据库云平台

理论与实践

马献章 著

- ✓ **针对需求** 分析了关系型数据库与NoSQL数据库的适应场景
- ✓ **针对安全** 介绍了国产主流自主可控数据库
- ✓ **针对管理** 介绍了数据库中的事务、数据恢复和一致性
- ✓ **针对业务** 介绍了数据库应用系统的设计、优化和集成技术
- ✓ **针对开发** 介绍了主流关系数据库和NoSQL数据库的编程技术

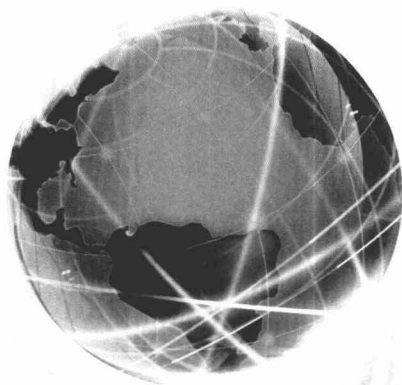
大学出版社



数据库云平台

理论与实践

马献章 著



清华大学出版社
北京

内 容 简 介

数据库云平台是当今乃至今后一段时期信息化领域普遍关注的一个热点领域。云技术的本质是分布式计算，而数据库云平台揭示的正是分布式计算在数据处理领域的本质问题。本书介绍了数据库的起源与发展，分析了关系型数据库与 NoSQL 数据库的适应场景，介绍了国产自主可控数据库云平台发展现状；针对云计算环境大数据时代对结构化和非结构化数据的管理需要，介绍了关系型数据库中的事务、数据恢复、SQL、分布式数据库和 NoSQL 数据库的一致性与事务等理论；针对越来越多的大数据业务，介绍了数据库应用系统的设计、优化和集成技术。结合发展趋势，重点介绍了当前流行的关系数据库和 NoSQL 数据库的编程技术；针对信息化建设演进式发展，介绍了数据库重构技术。

本书作为计算机、网络工程、信息管理院校研究生的高端教材，适合具有一定计算机基础知识的读者学习，也可作为数据分析师、系统架构师的数据库、云技术培训教材，以及各企事业单位组织实施信息化建设、流程再造、大数据的生态系统构建和信息化基础知识训练的参考书。

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

数据库云平台理论与实践 / 马献章著. —北京：清华大学出版社，2016
ISBN 978-7-302-42150-4

I. ①数… II. ①马… III. ①关系数据库系统-研究 IV. ①TP311.138

中国版本图书馆 CIP 数据核字（2015）第 271770 号

责任编辑：冯志强
封面设计：吕单单
责任校对：徐俊伟
责任印制：李红英

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

印装者：清华大学印刷厂

经 销：全国新华书店

开 本：185mm×260mm 印 张：32.5 插 页：1 字 数：818 千字

版 次：2016 年 1 月第 1 版 印 次：2016 年 1 月第 1 次印刷

印 数：1~3000

定 价：79.00 元

前 言

人类活动的空间延伸到哪里，数据便从哪里产生。数据是人类活动的重要资源。数据管理技术的优劣，直接影响数据处理的效率，直接影响决策的时效。数据库技术正是瞄准这一目标发展起来的专业技术，它主要研究如何存储、使用和管理数据。数据库系统就是研究如何高效地存取和科学地管理数据的计算机系统。

近年来，数据库技术是计算机科学技术中发展最快的领域之一，它的出现使得计算机应用渗透到工农业生产、商业管理、科学研究、工程技术以及国防军事等各个领域，大到国民经济、国家安全，小到个人网上购物、通信录管理，与我们每一个人息息相关。现在，数据库系统的建设规模、处理能力以及应用程度，已经成为衡量一个企事业单位甚至政府机关、军队信息化程度的重要标志。数据库技术是计算机信息系统与应用系统的核心技术，信息安全的新战场，企事业单位甚至国家间核心竞争能力的新抓手。

本书由三部分组成。第一部分数据库云平台导论，由第 1~4 章组成。这部分是背景知识，专为 IT 主管部门、企事业单位的 CEO、CIO 们以及院校本科生、研究生的学习数据库而准备，介绍了数据库的起源，分析了关系型数据库与 NoSQL 数据库的优缺点，讲述了两者的如何从不同角度来解决数据存储问题。在混合持久化的新环境下，二者互为补充，相辅相成；针对当前云计算的大潮，介绍了数据库云平台的概念，以及国产自主可控数据库云平台的代表产品——虚谷云数据库和南大通用列存数据库，提出了在国家安全前提下，企事业单位选择使用数据库云平台时如何权衡国外产品与国产自主可控产品。

第二部分数据库云平台理论基础，由第 5~9 章组成。这部分既包含了经典的数据库理论，又包含了前沿的数据库发展理论，一定篇幅涉及了构建数据库管理系统，特别是将关系型数据库的时态数据支持、窗口和窗口函数、管线化数据操作、查询取回数量的控制等 SQL 标准的最新内容，以及 NoSQL 数据库的一致性与事务等新内容首次纳入教学中，使读者可以掌握最前沿的知识，适合具有一定数据库理论基础的读者学习。

第三部分数据库云平台的应用，由第 10~13 章组成。主要内容是关系型数据库与 NoSQL 数据库的编程、调优、应用设计和重构等知识，其主要考虑是大部分学生在未来要实现或重构数据库及其应用程序，只有很少一部分学生会去构建数据库管理系统。因此，其篇幅很大，分量很重，是教学的重点。数据库重构技术，也是数据库领域专家必备的知识。此外，本书包含丰富的材料来描述事务用来访问数据库的语言和 API，比如嵌入式 SQL、动态 SQL、ODBC、JDBC 和 ADO.NET 接口等，100 多个精彩示例代码可以帮助读者快速掌握编程，避免作者探索中曾经付出的代价，适合具有一定数据库理论基础的读者学习。

本书在编写过程中，许多人从该书的最初策划到框架结构的确定和具体内容的撰写，倾注了大量心血，并提出了非常宝贵的意见，在此谨表示衷心地感谢。特别是戴浩院士、张景中院士、李德毅院士、凌永顺院士、张锡祥院士、黄先祥院士、于全院士、李乐民院士、朱中梁院士、陈鲸院士、尹浩院士、陈志杰院士、赵晓哲院士、吕跃光院士、赵捷教

授、袁文先教授、宋自林教授、王振国教授、秦志光教授、裘杭萍教授等资深专家对书稿进行了帮助指导，戴浩院士亲自撰写了序言；孔辉博士、滕明贵博士、柳虔林博士、钟军博士、王汉瑛高工对本书的内容给出了广泛和宝贵的反馈意见；王汉瑛高工帮助我编写了虚谷云数据库一节，江一民工程师、段光明工程师、孔令梅工程师对书中例子进行详细的验证。他们为本书的编写、审定和出版付出了辛勤的劳动，贡献了卓越的智慧，在本书付梓之际，谨表示最诚挚的感谢和崇高的敬意。感谢我的妻子王丽平，在我创作这本书的整个过程中对我一如既往的支持。

本书在撰写过程中，汲取、借鉴了国内外一些学者和同行的最新研究成果，在此向他们表示衷心的感谢！正是有了他们的劳动成果才使得我能够站在“巨人肩上”看得更远，也才能使这部专著得以问世。

由于数据库云平台理论尚在研究与发展之中，许多学术问题有待进一步深入研究，尽管对此做了很大努力，但由于能力水平有限，仍有不尽人意之处，恳请读者批评指正。

编者

2015年5月

序

近年来，大数据热潮使得非关系型数据库（NoSQL）受到了人们更多的关注，甚至有人提出：非关系型数据库将主导未来大数据分析领域，关系型数据库（SQL）将走向死亡。事实上，关系型数据库和非关系型数据库在大数据时代都有很大的发展空间。大数据的特点之一是价值密度低，数据挖掘可以说是沙里淘金、海里捞针，有可能如获至宝，也有可能一无所获。例如北京市“公交一卡通”每天刷卡高达四千万次，地铁流量达一千万人次，运用 NoSQL 技术分析这些大数据，其结果对于优化北京公交线路具有重要的参考价值。类似的成功案例还有许多。但 NoSQL 并不能全面解决数据库技术面临的挑战，传统的关系型数据库技术仍有用武之地，它具有数据结构易于理解、操纵语言使用方便、数据完型性有充分保证等特点，并在事务处理等领域取得了相当大的成功。40 多年来，人们在数据查询和应用开发等方面积累了丰富的经验和研究成果，关系型数据库可作为开发大数据管理系统研究的框架或起点。

从理论上讲，关系型数据库和非关系型数据库两者呈互补关系。数据库的 CAP 理论指出：一致性（Consistency）、可用性（Availability）、分区容错性（tolerance to network Partitions）三者无法兼顾，任何分布式系统只能同时满足其中两点。该理论已经实践证明成为业界的共识。

众所周知，各种关系型数据库的理论基础是 ACID，即原子性（Atomicity）、一致性（Consistency）、隔离性（Isolation）和持久性（Durability）。ACID 模型体现了传统数据库的设计思想，但也限制了系统的某些性能，如横向扩展能力、高效存储与访问能力、非结构化查询与分析能力等，尤其在应对社会网络上大容量、高并发数据时显得力不从心。非关系型数据库的理论基础是 BASE 模型，即基本可用（Basically Available）、软状态（Soft state）、最终一致性（Eventually consistent）。ACID 和 BASE 在化学中分别指代“酸”和“碱”，在分布式信息系统中，BASE 和 ACID 正好代表了两种不同的数据管理方法。

NoSQL 的正确解读是“不仅仅是 SQL”（Not only SQL），它泛指非关系型、分布式的数据存储系统，包括列存储数据库、键/值存储数据库、文档型数据库、图形数据库、对象数据库等。BASE 模型的核心思想是牺牲强一致性要求，可以基于日志实现异步复制，可以处理超大量的数据；数据分区存放，具有天然的冗余备份特征；根据应用的特点采用比较简单的数据模型，不需要预定义表结构。虽然非关系型数据库的结构化查询统计能力偏弱，在数据一致性方面还需应用层保障，但因 NoSQL 除具有较高的可用性和柔性的可靠个性，还有很强的横向、纵向扩展能力，在特定的应用领域它仍有很大的发展空间。

随着 SNS（社会性网络服务）类型的 Web 2.0 网站的兴起，用户每天生产的数据量剧增，要求数据库有高扩展性，用户高并发的读写请求，要求网站能及时响应；与此同时，数据的一致性显得不是那么重要了，于是非关系型数据库应运而生。严格地说，非关系型数据库是一种数据结构化存储方法的集合，而云计算时代海量数据管理系统就是为可扩展

性、弹性、容错性、自管理性等目标设计的，因此典型 NoSQL 数据库系统离不开云平台的支持。

针对云计算环境下结构化和非结构化数据的管理需要，马献章高工撰写了《数据库云平台理论与实践》一书。该书由三部分组成。导论介绍的是背景知识，重点分析 SQL 与 NoSQL 数据库的优缺点，并在云计算的架构下引出平台服务的概念——数据库云平台。第二部分是理论基础，包含了经典的和前沿的数据库理论，特别是介绍了 SQL 标准的最新进展，如时态数据支持、窗口和窗口函数、管线化数据操纵语言、查询结果的数量控制等知识。第三部分是数据库云平台应用，包括应用程序的优化、NoSQL 数据库性能调优、数据库重构等知识。作者结合近 30 年的开发实践，介绍了若干精彩的应用案例，读者从中可以领略其开发思想和编程技巧。

《数据库云平台理论与实践》内容丰富、贴近实践，是云计算时代决策者、咨询者和技术研发者不可多得的一部工具书和普及读物。它紧跟信息时代的发展潮流，深入浅出、循序渐进地解读了云计算时代基于数据管理的各类应用，对于提高企事业单位的信息化水平和核心竞争力具有较强的参考价值。我与作者相识多年，马高工长期在基层一线工作，曾主持过“珠峰”可信数据库的开发和多项信息系统建设，勤学敬业，思维敏捷，笔耕不辍，论著颇丰。《数据库云平台理论与实践》就是他多年心血的结晶，我向这本专著的出版表示祝贺。同时感谢清华大学出版社的编辑独具慧眼，为我们提供了一部好教材。希望本书能成为广大读者的良师益友。

中国工程院院士

戴润

2015 年 7 月

专家推荐语

大数据、云计算，近年来受到广泛关注，非专业人员恐怕也能说出一些自己的理解。但是，数据的增长到底带来哪些技术上的挑战，如何应对这些挑战，大数据需要云计算吗，云计算又能够提供哪些支撑，等等，回答这些问题，有赖于严谨深入的专业性工作。

《数据库云平台理论与实践》以作者长期一线工作的深厚工程经验为基础，介绍了数据库管理系统实现的核心关键技术，也介绍了数据库应用系统的设计、优化和集成技术，内容组织循序渐进、深入浅出，贴近实践。尤其可贵的是，作者敏锐把握信息技术和产业发展趋势，吸收了云计算、大数据等方面新的进展，是云计算时代数据库研发者与应用者不可多得的一部工具书和普及读物。

中国科学院深圳先进技术研究院研究员 冯圣中

《数据库云平台理论与实践》以作者长期一线工作的深厚工程经验为基础，紧跟信息时代的发展潮流，既介绍了数据库管理系统实现的核心关键技术，也介绍了数据库应用系统的设计、优化和集成技术，内容组织循序渐进、深入浅出，贴近实践，是云计算时代数据库研发者与应用者不可多得的一部工具书和普及读物。

国防科技大学计算机学院院长（教授） 廖湘科

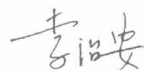
《数据库云平台理论与实践》是一部好书，既介绍了数据库云平台的经典理论，又介绍了前沿知识，还结合大型企事业单位数据管理和决策需要，介绍了数据库应用系统的设计、编程、优化和集成，是云计算时代决策者、咨询者和技术开发者不可多得的一部通向数据管理和利用的知识普及书和导游图，内容丰富，贴近实践，符合目前企事业单位的现状和发展潮流，对于提高单位在本领域的信息化水平和核心竞争力具有较强的实用性和参考价值。

军事科学院战略研究部博士生导师（研究员、少将） 姜春光

数据库技术发展近年来发展很快，随着计算机技术的广泛运用，以及云计算的生成和实践，数据库技术与云结合的越来越紧密，数据库的强大离不开云平台的支撑，云计算应用的普及与拓展离不开数据库的支撑。该书正是在这种发展的需求中应运而生，对于如何充分依托云平台，进行存储、使用和管理数据库进行了详细阐述，对于解决提高各行各业基于云环境下的数据科学管理和使用，提高信息化能力

和水平具有极其重要的参考价值。该书第一部分从数据库的起源与发展现状入手,详细介绍了数据库的分类,关系模型与数据库的优缺点,分析了 NoSQL 数据库解决的核心难题,阐述了数据库云平台的概念、特点和分类,可控云平台,以及应用和存在问题等;第二部分的事务原子性、并发性和锁机制,数据备份与恢复时涉及的技术,关系数据库的查询语言,分布式数据库的构建原理和方法,以及 NoSQL 数据库解决一致性问题的机制,使读者掌握了数据库云平台的理论基础;通过第三部分介绍的关系数据库和 NoSQL 数据库的编程、调优、应用设计和重构等知识,使读者掌握开发云环境下应用的核心关键方法。从总体看,该书对于读者系统学习数据库应用编程知识,掌握开发准则和编程技巧,在云计算环境下结构化和非结构化数据的管理具有很强的帮助,可作为高等院校信息类专业研究生教学使用教材。

空军工程大学信息与导航学院教授(大校)



《数据库云平台理论与实践》一书紧贴大数据时代信息处理技术的发展前沿,系统介绍了具有云架构的数据库技术的最新进展和应用。全书内容精练、结构清晰、逻辑严谨、深入浅出,既有理论阐述,又有实例介绍。适合高等院校信息类专业高年级学生和研究生阅读,也可供具有一定实战经验的 IT 工程师参考。

解放军理工大学指挥信息系统学院教授(少将)



《数据库云平台理论与实践》以作者长期理论探索和工程实践为基础,介绍了数据库管理系统实现的核心关键技术,以及数据库应用系统的设计、优化和集成技术。内容组织结构清晰、系统性强,知识介绍循序渐进、深入浅出,并紧密地结合了实践。作者很好地把握了信息技术和产业的发展趋势,吸收了云计算、大数据等方面的新兴技术,是目前数据库研发者与应用者不可多得的一部工具书和普及读物,也可作为研究生和博士生学习使用。

电子科技大学软件工程学院院长(教授)



目 录

第一部分 数据库云平台导论

第 1 章	数据库的起源与发展现状	2
1.1	数据管理	2
1.1.1	数据管理的 3 个阶段	2
1.1.2	数据库系统发展的 3 个里程碑	4
1.1.3	数据库管理系统的 3 个发展阶段	5
1.2	数据库的分类	6
1.2.1	层次型数据库	7
1.2.2	网状型数据库	7
1.2.3	关系型数据库	8
第 2 章	关系模型与数据库的优缺点	10
2.1	关系模型	10
2.1.1	关系模型概述	10
2.1.2	关系代数	14
2.1.3	关系演算	18
2.2	关系模式与关系数据库	20
2.3	关系型数据库的优势	20
2.3.1	持久存储大量数据	21
2.3.2	通过事务保证数据的强一致性	21
2.3.3	通用性好和高性能	21
2.3.4	以标准化为前提	21
2.4	关系型数据库的不足	22
2.4.1	大量数据的写入操作	22
2.4.2	对海量数据的高效存储和访问	24
2.4.3	为有数据更新的表做索引或表结构变更	24
2.4.4	对简单查询需要快速返回结果的处理	25
2.4.5	字段不固定时的应用	25
第 3 章	关系型数据库的补充——NoSQL	26
3.1	NoSQL 的概念	26
3.2	NoSQL 的起源	26
3.3	NoSQL 数据库解决的核心难题	28
3.3.1	实现高速查询	28

3.3.2	满足多写需要	28
3.4	NoSQL 数据库的分类	29
3.4.1	面向列的有序存储数据库	29
3.4.2	键/值存储数据库	30
3.4.3	面向文档数据库	31
3.4.4	图形数据库	32
3.4.5	对象数据库	33
第 4 章	数据库云平台	35
4.1	数据库云平台的概念	35
4.2	数据库云平台的特点	35
4.3	数据库云平台的分类	37
4.3.1	数据库即服务	37
4.3.2	分布式数据库	37
4.4	国产自主可控数据库云平台产品介绍	37
4.4.1	虚谷云数据库	38
4.4.2	南大通用列存数据库	41
4.5	为何推荐使用自主可控产品	43
4.5.1	当前国际主流数据库的安全隐患	44
4.5.2	信息安全最终要落实到数据库安全	44
4.5.3	满足需求就好	45
4.5.4	一站式服务方式是优势	46

第二部分 数据库云平台理论基础

第 5 章	关系型数据库中的事务	48
5.1	事务处理的重要性	48
5.2	事务的特性要求	49
5.2.1	一致性	50
5.2.2	原子性	50
5.2.3	持久性	51
5.2.4	隔离性	51
5.3	事务的状态	51
5.4	系统日志	53
5.5	并发控制	53
5.5.1	基于锁的协议	54
5.5.2	两阶段加锁	58
5.5.3	死锁处理	59
5.5.4	基于时间戳排序的并发控制	61
5.5.5	多版本并发控制	63
5.5.6	饥饿处理	64

5.5.7	索引中使用加锁进行并发控制	65
5.5.8	其他并发控制问题	66
5.6	锁的粒度	67
5.6.1	加锁的粒度级别	68
5.6.2	多粒度级别加锁	68
5.7	集中式系统中的事务处理	71
5.7.1	单用户系统的组织	71
5.7.2	集中式多用户系统的组织	72
5.8	分布式系统上的事务处理	73
5.8.1	数据库服务器分布的关键因素	73
5.8.2	分布式系统的组织	73
5.8.3	会话与上下文信息	77
5.8.4	队列事务处理	79
5.8.5	分布式事务基本两阶段提交协议	79
5.8.6	分布式事务基本两阶段提交协议的改进	80
5.9	结构化查询语言中的事务支持	82
第 6 章	关系型数据库的数据恢复	84
6.1	数据库数据恢复的概念	84
6.1.1	数据库故障的种类	84
6.1.2	恢复算法概要及分类	85
6.1.3	故障恢复技术	86
6.1.4	潜入/非潜入与强制/非强制	88
6.1.5	日志中的检查点	88
6.1.6	事务故障的恢复	90
6.2	延迟更新恢复技术	90
6.2.1	单用户环境下使用延迟更新的恢复	91
6.2.2	多用户环境下采用并发执行方案的延迟更新	92
6.2.3	不影响数据库的事务动作	93
6.3	即时更新恢复技术	93
6.3.1	单用户环境下即时更新的恢复	94
6.3.2	多用户环境下采用并发执行方案即时更新的恢复	94
6.4	镜像分页	94
6.5	多数据库系统中数据的恢复	96
6.6	系统与介质故障的恢复	96
第 7 章	关系型数据库结构化查询语言 SQL	98
7.1	SQL 基础知识	98
7.1.1	语句	98
7.1.2	名称	101
7.1.3	数据类型	102
7.1.4	常量	107
7.1.5	表达式	108

7.1.6	内嵌函数	108
7.1.7	NULL 值	109
7.1.8	别名和匿名	110
7.1.9	时间段	110
7.2	SQL 的数据定义	111
7.2.1	在 SQL 中创建和删除数据库	112
7.2.2	SQL 中基本表的创建、修改、删除	113
7.2.3	SQL 中索引的建立与删除	117
7.3	SQL 的数据查询	119
7.3.1	查询语句的基本结构	119
7.3.2	单表查询	120
7.3.3	连接查询	128
7.3.4	嵌套查询	129
7.3.5	集合查询	133
7.3.6	窗口和窗口函数	134
7.3.7	查询取回数量的控制	138
7.4	SQL 的数据操纵	139
7.4.1	插入数据	139
7.4.2	删除数据	140
7.4.3	更新数据	142
7.4.4	管线化数据操纵语言	143
7.5	SQL 的视图	144
7.5.1	DBMS 如何处理视图	145
7.5.2	创建视图	145
7.5.3	删除视图	147
7.5.4	查询视图	147
7.5.5	更新视图	147
7.5.6	物化视图	148
7.6	时态 SQL	149
7.6.1	时态的概念	149
7.6.2	应用时间段表	152
7.6.3	系统版本表	156
7.6.4	双时态表	160
7.7	数据库持久化存储模块	161
7.7.1	存储过程与函数	161
7.7.2	时态 SQL 与 PSM	164
7.8	SQL 的触发器	166
7.8.1	创建触发器	167
7.8.2	触发器使用的特殊表	168
7.8.3	修改触发器	174
7.8.4	删除触发器	174

7.8.5	触发器的使用限制	175
7.9	SQL 的访问控制	175
7.9.1	授予权限	175
7.9.2	取消权限	176
7.9.3	视图机制与 SQL 安全	177
第 8 章	分布式数据库	178
8.1	分布式数据库系统的概念	178
8.1.1	分布式数据库系统的定义	178
8.1.2	分布式数据库系统的特点	180
8.1.3	分布式数据库系统的分类	181
8.1.4	分布式数据库系统的特色功能	182
8.2	分布式数据库系统的体系结构	183
8.2.1	分布式数据库系统的模式结构	183
8.2.2	分布式数据库管理系统的组成	185
8.2.3	网络应用程序	186
8.3	分布式数据库的数据分片、复制和分配技术	189
8.3.1	数据分片和分配	189
8.3.2	数据复制	192
8.3.3	数据分配	196
第 9 章	NoSQL 型数据库的一致性与事务	197
9.1	一致性问题	197
9.1.1	更新一致性	197
9.1.2	读取一致性	198
9.2	放宽一致性约束	201
9.2.1	CAP 定理	202
9.2.2	BASE 理论	207
9.2.3	NWR 理论	207
9.3	放宽持久性约束	208
9.4	版本戳	209
9.4.1	“商业事务”与“系统事务”	209
9.4.2	在多结点环境中生成版本戳	210
9.5	键/值数据库的一致性与事务	212
9.5.1	键/值数据库的一致性	212
9.5.2	键/值数据库的事务	212
9.6	文档数据库的一致性与事务	213
9.6.1	文档数据库的一致性	213
9.6.2	文档数据库的事务	214
9.7	列族数据库的一致性与事务	214
9.7.1	列族数据库的一致性	214
9.7.2	列族数据库的事务	216
9.8	图形数据库的一致性与事务	216

9.8.1 图形数据库的一致性 216

9.8.2 图形数据库的事务 216

第三部分 数据库云平台的应用

第 10 章 开发数据库应用的编程 220

10.1 关系型数据库的嵌入式 SQL 220

10.1.1 程序化 SQL 技术 220

10.1.2 简单的嵌入式 SQL 语句 227

10.1.3 嵌入式 SQL 中的数据检索 240

10.1.4 基于游标的删除和更新 251

10.1.5 游标和事务处理 252

10.2 关系型数据库的动态 SQL 253

10.2.1 动态 SQL 的概念 253

10.2.2 动态语句的执行 254

10.2.3 动态 SQL 的两步动态执行 256

10.2.4 动态查询 259

10.3 关系型数据库应用编程接口 268

10.3.1 应用编程接口的概念 268

10.3.2 ODBC 和 SQL/CLI 标准 270

10.3.3 ODBC API 292

10.3.4 Java 数据库连接 298

10.3.5 ADO.NET 接口 310

10.4 NoSQL 数据库编程 314

10.4.1 MongoDB 文档型数据库的编程接口 314

10.4.2 Cassandra 列族数据库的编程接口 328

10.4.3 Redis 键/值存储数据库的编程接口 352

第 11 章 数据库性能调优 361

11.1 调优问题概述 361

11.1.1 调优的目标 361

11.1.2 识别性能问题 362

11.1.3 剖析性能问题 363

11.1.4 优化解决问题 364

11.2 关系型数据库的查询优化 366

11.2.1 查询处理的架构 367

11.2.2 基于关系代数等价性的启发式优化 368

11.2.3 查询执行计划的开销估计 371

11.2.4 选择一个计划 377

11.3 应用程序的优化 380

11.3.1 SQL 语句的优化 381

11.3.2	索引	391
11.3.3	反向规范化	393
11.3.4	实现惰性读取	394
11.3.5	引入缓存	395
11.3.6	充分利用工具	395
11.4	物理资源的管理	397
11.5	NoSQL 数据库的调优	398
11.5.1	NoSQL 数据库调优的原则	398
11.5.2	文档型数据库 MongoDB 的常用优化方案	398
11.5.3	列族数据库 Cassandra 的优化	402
第 12 章	数据库应用系统的设计	405
12.1	数据库应用系统设计的目标	405
12.2	数据库应用系统的设计方法学	405
12.2.1	数据库应用系统涉及的角色	406
12.2.2	数据库应用的设计与实现过程	408
12.2.3	使用 UML 图作为数据库设计规范说明的辅助工具	417
12.3	面向数据的关键设计	427
12.3.1	在关系数据库内查找对象	427
12.3.2	实现引用完整性	433
12.3.3	实现安全访问控制	436
12.4	支持数据库渐进式开发的潜在工具	445
12.4.1	工具	445
12.4.2	沙箱	446
12.4.3	脚本	447
第 13 章	数据库重构	449
13.1	数据库重构的重要性	449
13.2	数据库重构的概念	451
13.2.1	数据库重构的定义	451
13.2.2	数据库重构的内涵是保持语义	453
13.2.3	数据库重构的类别	453
13.2.4	重构工具	454
13.3	数据库重构的过程	454
13.3.1	确认数据库重构是必要的	454
13.3.2	选择最合适的数据库重构	456
13.3.3	确定数据清洗的需求	456
13.3.4	使原数据库模式过时	458
13.3.5	编写单元测试进行前测试、中测试和后测试	458
13.3.6	实现预期的数据库模式变化	461
13.3.7	迁移源数据	462
13.3.8	更新数据库管理脚本	462
13.3.9	重构外部访问程序	463

- 13.3.10 进行回归测试 463
 - 13.3.11 为重构编写文档..... 463
 - 13.3.12 对工作进行版本控制 464
- 13.4 数据库重构的策略..... 464
 - 13.4.1 通过小变更降低变更风险 464
 - 13.4.2 唯一地标识每一次重构 464
 - 13.4.3 转换期触发器优于视图或批量同步 465
 - 13.4.4 确定一个足够长的转换期 465
 - 13.4.5 封装对数据库的访问 466
 - 13.4.6 使建立数据库环境简单 466
 - 13.4.7 将数据库资产置于变更控制之下 467
- 13.5 数据库重构的方法..... 467
 - 13.5.1 结构重构 467
 - 13.5.2 参照完整性重构 481
 - 13.5.3 数据质量重构 493