

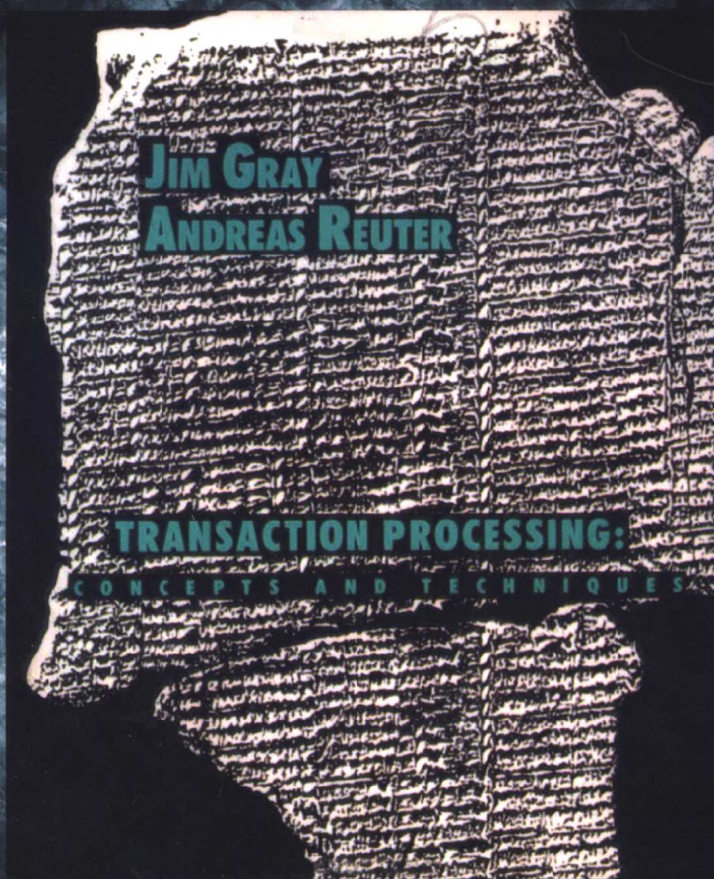


计 算 机 科 学 丛 书

事务处理

概念与技术

(美) Jim Gray Andreas Reuter 著 孟小峰 于戈 等译



Transaction Processing
Concepts and Techniques

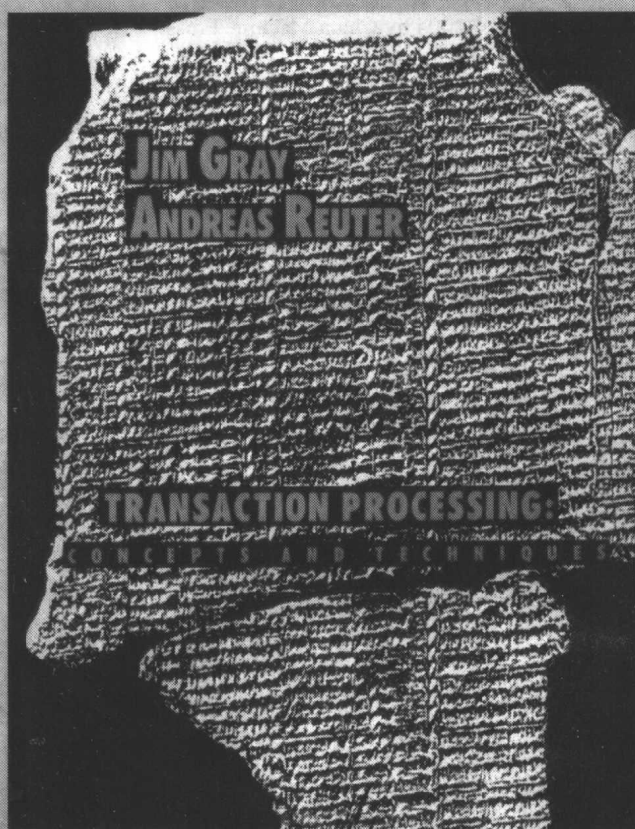


机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

事务处理 概念与技术

(美) Jim Gray Andreas Reuter 著 孟小峰 于戈 等译



Transaction Processing
Concepts and Techniques



机械工业出版社
China Machine Press

事务处理广泛地应用于数据库和操作系统领域，并在现代计算机系统中监控、控制和更新信息。本书向读者展示了大型的、分布的、异构的计算机系统是如何进行可靠工作的。作者使用事务作为基本概念，说明了在有限的资金和风险下如何构建高性能的高可用性应用。书中还详细阐述了各种可能发生的问题，以及解决这些问题的实际可用的技术。本书列举了大量成功的商业和研究系统的实例，此外，列出了许多事务处理算法的可编译的C代码片段。本书对于那些对实现分布式系统或客户-服务器结构感兴趣的人来说，是值得一读的。

本书适合作为高校的教科书，也可作为计算机专业人员的参考书。

Jim Gray and Andreas Reuter: Transaction Processing: Concepts and Techniques (ISBN: 1-55860-190-2).

Copyright © 1993 by Morgan Kaufmann Publishers, Inc.

Translation Copyright © 2004 by China Machine Press.

All rights reserved.

本书中文简体字版由美国Morgan Kaufmann公司授权机械工业出版社独家出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

版权所有，侵权必究。

本书版权登记号：图字：01-2001-2210

图书在版编目（CIP）数据

事务处理：概念与技术 /（美）格雷（Gray, J）等著；孟小峰等译. -北京：机械工业出版社，2004.1

（计算机科学丛书）

书名原文：Transaction Processing: Concepts and Techniques

ISBN 7-111-12641-6

I. 事… II. ①格… ②孟… III. 事务计算程序 IV. TP319

中国版本图书馆CIP数据核字（2003）第062116号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：温莉芳 温丹丹

北京昌平奔腾印刷厂印刷·新华书店北京发行所发行

2004年1月第1版第1次印刷

787mm×1092mm 1/16·53.75印张

印数：0 001-4 000册

定价：96.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换
本社购书热线：（010）68326294

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅筹划了研究的范畴，还揭开了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall, Addison-Wesley, McGraw-Hill, Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum, Stroustrup, Kernighan, Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及收藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业

的教学度身订造的。其中许多教材均已为M. I. T., Stanford, U.C. Berkeley, C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程,而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下,读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑,这些因素使我们的图书有了质量的保证,但我们的目标是尽善尽美,而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正,我们的联系方法如下:

电子邮件: hzedu@hzbook.com

联系电话: (010) 68995264

联系地址: 北京市西城区百万庄南街1号

邮政编码: 100037

专家指导委员会

(按姓氏笔画顺序)

尤晋元	王 珊	冯博琴	史忠植	史美林
石教英	吕 建	孙玉芳	吴世忠	吴时霖
张立昂	李伟琴	李师贤	李建中	杨冬青
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
周立柱	周克定	周傲英	孟小峰	岳丽华
范 明	郑国梁	施伯乐	钟玉琢	唐世渭
袁崇义	高传善	梅 宏	程 旭	程时端
谢希仁	裘宗燕	戴 葵		

秘 书 组

武卫东 温莉芳 刘 江 杨海玲

Preface to the Chinese Edition

The fundamental premise of this book is that the transaction concept is a key abstraction for distributed computation. One implements transactions at a one layer, and then all higher layers have a much simpler failure semantics (all or nothing) and much simpler error handling.

This was a radical view when the book first appeared ten years ago. Since then the Internet has blossomed, distributed computing has become more tangible with CORBA, DCOM, and now Web Services. TP monitors have been supplanted by web servers and object request brokers. Database systems have matured enormously.

Today we can see that many have been converted to this view. Transaction techniques are being used within most operating systems, and transactional file systems and transaction managers are being added to most operating systems. Transactional resource managers like queue managers, workflow systems, persistent object systems, and main memory database systems are proliferating now that this enabling technology is in place.

Some things about this book are dated-even obsolete-but the core concepts and techniques have moved from the specialty of database systems to the broader context of distributed computing.

Andreas Reuter and I are grateful to you for undertaking the arduous task of translating his book from English to Chinese. Chinese, being a more compact language, will make the book shorter (and lighter), and having it in Chinese will make it accessible to a more than a billion people. I hope you find this book informative and helpful.

Jim Gray
Microsoft Research
18 October, 2002

中文版序

本书的基本前提在于，事务概念是分布式计算的一个重要抽象。在某一层次上实现了事务，那么所有更高的层次上会有一个简化的失败语义（要么全做，要么全不做），并且错误处理也会简单的多。

十年前当本书初次面世时，这是一个基本的观点。自此以后，因特网蓬勃发展，分布式计算在CORBA、DCOM和现在的Web Service的支持下，变得更为切实可行。TP监控器已经被Web服务器和对象请求代理所取代。数据库系统也已经成熟多了。

今天我们看到，许多技术都已接受了这一观点。事务技术被广泛用于操作系统、事务型文件系统，并且事务管理器被加入到了大多数的操作系统中。由于事务技术已经成熟，所以事务型资源管理器（如队列管理器、工作流系统、持久对象系统和主存数据库系统等）日益推广流行。

本书中的一些内容已经过时了——甚至是作废了——但其核心概念和技术已从数据库系统扩展到了分布式计算的范畴。

我和Andreas Reuter感谢译者的辛勤工作。中文，作为更为简洁的语言，会使本书更薄些（也更轻了），中文版还会使十多亿人有机会阅读本书。我希望本书对各位是有价值和有帮助的。

Jim Gray
微软研究院
2002.10

译者序

值此译稿完成之时，从DBWORLD惊悉关系模型的发明人E·F Codd博士于2003年4月18日不幸辞世（1923年~2003年）。人虽逝去，但关系模型这一知识财富却在造福着人类。在整个计算机软件领域，恐怕难以找到第二个像关系模型这样简单，但所带来的市场价值却如此巨大的概念。其实，传承这一发明并使之真正实用化的关键人物有两人：一是C·J Date，他使关系模型概念普及化，为大众所接受；另一个当属本书的作者之一Jim Gray博士，他使关系模型的技术实用化，为大众服务。在IBM工作期间，Jim Gray参与和主持过IMS、System R、SQL/DS、DB2等项目的开发。其中，除了System R仅作为研究原型，没有成为产品外，其他几个项目都发展成数据库市场上有影响力的产品。这其中Jim Gray博士在事务处理技术方面的贡献功不可没。

在关系数据库理论基本成熟后，各大公司在关系数据库管理系统（RDBMS）的实现和产品开发中都遇到一系列技术问题。主要是在数据库的规模愈来愈大，数据库的结构愈来愈复杂，又有愈来愈多的用户共享数据库的情况下，如何保障数据的完整性、安全性、并发性以及故障恢复的能力，这些问题成为数据库产品是否能实用化并最终为用户接受的关键因素。Jim Gray在解决这些重大技术问题，使RDBMS成熟并顺利进入市场的过程中，起到了关键性作用。概括地说，解决上述问题的主要技术手段和方法是：把对数据库的操作划分为称为“事务”的基本原子单位，一个事务要么全做，要么全不做（即all-or-nothing原则）；用户在对数据库发出操作请求时，需要对有关的数据“加锁”，防止不同用户的操作之间互相干扰；在事务运行过程中，采用“日志”记录事务的运行状态，以便发生故障时进行恢复；对数据库的任何更新都采用“两阶段提交”策略。以上方法及其他各种方法总称为“事务处理技术”。Jim Gray在事务处理技术上的创造性思维和开拓性工作，使他成为这一领域公认的权威人士，并于1998年获得图灵奖，成为自图灵奖诞生32年来第三位因在数据库技术的发展中做出重大贡献而获此殊荣的学者（见《ACM图灵奖（1966~1999）》，高等教育出版社，2000年）。

本书即是Jim Gray博士和德国斯图加特大学Andreas Reuter教授多年来在事务处理领域所取得成就的结晶。本书有如下几个显著特征：

- **内容的组织和叙述的方法独树一帜** 计算机科学在其60多年的发展中已形成了一个综合性的学科体系，可分为硬件、软件、计算机原理、操作系统……等诸多领域。现有的书通常是某个学科领域的知识总结。本书却是一个例外。它好像是一位大师挥舞手中的巨擘，打破时空的界限，勾画出的印象派杰作。作者用一个系统实现者的眼光阐述有关问题，力图使读者“所学即能为我所用”，因此数据结构和程序片段在书中随处可见。它更注重实效，更加独树一帜。
- **内容的广泛和技术的深入独一无二** 本书的目的是解决如何建造大型的、分布的、异构的计算机系统的问题。事务处理技术只是作者阐述这一复杂问题的“一种集成的概念框架”，书中内容分为七个部分，即事务处理基础、容错技术、面向事务的计算、并发控制、恢复、资源管理器实例、系统综述；分别涉及计算机基本知识、数据库系统、分布式系统、操作系统等。该书的内容涉猎广泛，而在事务处理技术上讲述得深入透彻而且完整全面。本书更像是一本关于事务处理的小型百科全书，这在计算机界是独一无二的。

• **技术的实践和学术的研究完美结合** 其实本书在整体上主张技术实践优先论。事务处理是一个技术实践引导理论研究的领域。通常商业系统实现了某些思想之后很久,该思想才会在学术界出现。Jim Gray常年在工业界的研究机构工作,大部分时间从事实际系统的开发。他实际是在通过本书向进行学术研究的人们提个醒儿,计算机的技术创新应该从研究系统开始,两者的贡献是不分高下的。

• **扎实的功底和渊博的知识力透纸背** 书中叙述的有关“Sphere of Control”的概念闻所未闻(第4章,译为控制区域),事务理论详尽透彻(第7章),文件系统的实现滴水不漏(第13~15章)。在本书的翻译过程中,我们常常为作者深厚的功底和渊博的知识所折服,同时又为自己的知识欠缺而惭愧,自叹不能精确地传达作者精妙的论述。

当然本书也不是完美无缺的。由于本书的内容过多,在叙述上有些内容不够简练,有些内容的叙述甚至太过冗余。如事务的ACID特性,几乎每章都用到,但每次都要重新解释。这与本书的写作风格有关,力图使每章内容都相对保持独立。

全书共8个部分19章。本书的翻译、统稿和审校由孟小峰和于戈共同组织完成。孟小峰负责第1~8章,于戈负责第9~16章,其余部分共同负责。具体翻译分工如下:序、前言由孟小峰翻译;第1章由李凌伟、孟小峰翻译;第2章由王博、张荣翻译;第3章由白芸、王建华翻译;第4章由王海燕、丁锐翻译;第5章由豆昱、李明晴翻译;第6章由张新宇、肖震翻译;第7章由孟小峰翻译;第8章由张在建、罗学君、孟小峰翻译;第9~11章由宋宝燕翻译;第12章由王国仁、夏秀峰翻译;第13章由鲍玉斌翻译;第14章由王丹翻译;第15章由于亚新翻译;第16章由于戈翻译;第18、19章由于戈、孟小峰翻译。

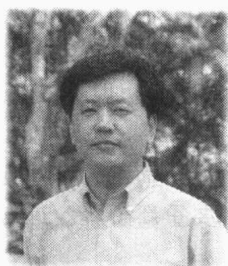
译者曾在2001年在罗马VLDB大会上有幸见到Jim Gray,谈到翻译本书的打算,得到他的极大鼓励,并在随后的版权购买过程中给予很大的帮助。2002年在北京再次见到Jim Gray博士,谈起我们的翻译工作已接近尾声,Jim显得极为高兴,并欣然在回美国的旅途中给本书中译本作了序。

译者感谢机械出版社华章公司的编辑们,没有他们的努力,本书是不可能完成的。

本书涉及面广,内容丰富,术语量大,这在一般的书籍中是不常见的,翻译难度可想而知。本书译词主要遵从教科书中的习惯用法,并参考《英汉计算机词典》(修订本,1998)等。为了搞清其他领域(如容错、操作系统等)中的词汇,我们专门请教了一些这方面的专家。但在翻译中仍感力不从心,译文中不当之处在所难免。诚恳读者批评指正并不吝赐教。如果你有任何建议或意见,欢迎发Email给xfmeng@public.bta.net.cn, yuge@mail.neu.edu.cn。

译者
2003年4月

译者简介



孟小峰，1964年生，博士，教授，博士生导师，中国人民大学信息学院计算机系副主任。现任中国计算机学会理事、中国计算机学会数据库专委会委员秘书长、中国计算机学会青年计算机科技论坛（YOCSEF）学术委员会主席（2003~2004年）、《计算机研究与发展》编委、IEEE CS会员、ACM SIGMOD会员。曾先后在香港中文大学、香港城市大学、新加坡国立大学访问研究。主持或参加过十多项国家科技攻关项目、国家自然科学基金以及国家863项目。获国家科技进步二等奖1项、电子部科技进步特等奖1项、北京市科技进步二等奖2项、第七届“中创软件人才奖”等奖励。研制开发的主要软件产品有国产数据库系统COBASE、嵌入式移动数据库系统“小金灵”、中文自然语言查询系统NChiq1、并行数据库系统PBASE/1等。近十年在国内外杂志及国际会议上发表论文60余篇，有关数据库方面的著译作7部。多次应邀担任国际会议ICDE'03、DASFAA'03-04、MDM'03-04、WISE'00、01、WAIM'01-03等的程序员委员会委员，并担任WAIM2002 程序委员会主席。近期主要研究领域为Web数据集成、XML数据库、移动数据管理等。



于戈，1962年生，博士，教授，博士生导师。1982年和1985年于东北大学获学士和硕士学位；1996年于日本九州大学获工学博士学位。现任东北大学信息学院副院长兼计算机系系主任，中国计算机学会数据库专业委员会副主任委员、中国计算机学会青年计算机论坛学术委员会委员、沈阳市电脑协会副理事长、辽宁省计算机学会副理事长、中国电子学会高级会员、ACM会员、SIGMOD会员、IEEE 会员、IEEE CS会员、日本IPSJ学会会员、WAIM国际学术会议指导委员会委员，ICDE' 2002、CODAS' 2001等多届国际学术会议程序委员会委员，《东北大学学报》（自然科学版）、《控制与决策》、《控制工程》等学术期刊的编委。承担或完成了“基于CORBA软件总线的CIMS信息集成平台的研究”等10多项国家863课题、国家自然科学基金课题和“新一代数据库基础技术的研究”等多项国际合作课题。获得国家科技进步奖二等奖2项，省部级科技进步二等奖1项、三等奖2项。获得教育部“跨世纪优秀人才基金”、“霍英东青年教师基金”、“中国高校青年教师奖”。出版《分布式数据库》等专著和论文集4部，在《J. of Computer Sci. and Tech.》、《计算机学报》、《软件学报》等权威学术期刊和ICDE'96、DASFAA'95等著名国际会议上发表学术论文50余篇。

序

企业、政府、科学、文化等领域的活动，正在变得越来越依赖于计算机信息资源。随着计算机系统获取和维护的信息量和种类日益增多，如何使用、管理和保护这些信息的技术对人类和现代工业社会变得至关重要。

事务处理技术是有效管理和使用计算机信息资源的关键。

事务处理涵盖了对所存储的信息和应用程序加以管理两方面的技术，其中应用程序负责对信息的解释和具体操作。从数据库恢复和并发控制，到事务监控器，负责初始化和控制应用的执行。这些事务处理技术为关键信息资源的管理和保护提供了必要的机制和措施，而这些关键信息资源其实已成为商业、科学、文化活动的基礎。

为使日益剧增的计算机信息有用，要求它们必须能够准确反映现实世界的情况，并能保证应用程序可获得它们。通常，使用所存储的信息是指访问和修改有关的数据，这些数据集中描述或刻画了现实世界中的某些现象或活动的状态和演化。由于许多数据要一起被访问和修改以便正确反映现实世界，因此必须保证相关数据的一致性。任何对相关数据修改的中断，或对相关数据访问或修改的交错，都会造成数据的不一致性。

保证数据一致性的关键是要明确数据访问和更新的序列。这一序列称为事务。事务处理技术就是要确保一个事务要么完整地执行，要么根本不执行，并保证并发执行的事务彼此互不干扰，如同在隔离执行一样。这一技术的意义尤其在于，在一些复杂情况下上述保证仍能得到支持，如计算机部件发生故障、数据分布在不同计算机上或不同事务的交叉或并行执行等。

过去25年中，企业和大学的实验室在事务处理技术方面付出了巨大的努力，已使得该技术可以保证事务的“all-or-nothing”（要么全做，要么全不做）执行，以及并发事务的隔离性。本书第一次全面介绍了事务处理系统的技术和方法，用来控制和保护事务处理系统管理的有价值的信息资源。作者详细描述了商品化的和实验性的事务处理系统的技术现状。重点论证了这些技术的有效性和高效性。详细解释了为什么要面对这些问题以及如何解决这些问题，所有这些使得本书不仅对求知的学生，而且对系统开发人员都是十分有益的。

作者Jim Gray博士和Andreas Reuter教授，在商品化和实验性事务处理系统的实现技术方面，加起来已有五十多年的直接经验。他们对事务处理技术做出了巨大的贡献，并因他们的学术成就而享誉世界。本书既是他们对事务处理问题深刻理解的结晶，也是对大部分行之有效的实务处理实现技术的难得的评价。作者具有的对基本概念和使用方法的辨识能力，给读者理解事务处理系统的问题和技术，提供了坚实和实用的基础。

本书覆盖了事务处理技术的所有方面。开始的几章为读者介绍了事务概念和事务执行的计算环境的基本知识，其中给出了有关计算机部件故障的重要假设，即事务处理系统必须对故障部件容错。书中阐述了事务处理监控器的作用。事务处理监控器主要控制应用程序和所提供功能的激活和执行。这为讨论并发控制和恢复技术打下了基础。事务隔离性的讨论覆盖了并发控制问题，它涉及了从硬件层到记录和索引的隔离语义的各方面问题。书中详细讨论了故障情况下重要而复杂的事务恢复技术。从记录管理到分布提交协议，对所需要的用于保证事务的“要么全做，要么全不做”执行和数据永久性的恢复技术详加阐述。其后，介绍了

将事务恢复技术和隔离技术用于记录级存储和关联索引的设计与实现。从这些章中，数据库系统方面的学生和开发人员都可以获得有用的信息。本书最后，从商业和学术两个方面对事务处理系统加以综述。

贯穿全书，作者对基本问题给出了深入细致的讨论，对已证明行之有效的技术给出了详细具体的描述。对概念方面的内容，辅以一些仔细设计的图表加以说明。对技术方面的内容，则给出代码片段，有效提高了读者对实现问题的认识。

本书既综合覆盖了事务处理技术，又详细描述了相关问题和算法，对学生是很好的教材，对专业人士则是资源宝库。事务处理技术对工业社会的信息管理需求十分重要，这需要人们对该技术深入理解和广泛应用。对于即将应用和拓展事务处理概念和技术的人士，本书可以充当这方面的指南和参考书。

Bruce Lindsay
IBM Almaden Research Center
圣荷塞，加利福尼亚

前言

买书的同时若能买到读书的时间，将是再好不过了。

亚瑟·叔本华：附录和补遗

写作此书的目的

本书的目的就是帮助读者理解大型、分布、异构的计算机系统可靠工作的原理。与常见的分布计算的复杂方法不同，本书力图展示一种易于使用的分布系统应用开发方法。为什么书名不用诸如分布系统、高可靠性、互操作性或客户/服务器等词汇，而使用了一个比较一般的术语——事务处理呢？对很多人来说这是一个表示老式数据处理的词汇，指面向批处理和大型主机的工作方式。

问题的关键是，一个大型的应用系统可能具有上千个终端，用到上百台计算机，需提供绝对没有停机时间的服务，对这种系统的设计、实现和操作，只从一个角度来阐述清楚是不可能的，这也是本书如此浩繁的原因。因此需要用一种综合的观察角度和方法来解决分布系统的问题。我们的宗旨是说明，事务提供了这样一种集成的概念框架，而面向分布事务的操作系统可以作为具体的实现技术。客户/服务器范型提供了一种结构化系统和开发应用的好方式，但仍然需要用事务来控制客户/服务器的交互操作。简而言之：没有事务，分布式系统不能满足典型实际应用的需要。

不过这并不是最关键的理由；更重要的是，这是许多人（包括系统实现者、系统管理者和应用开发人员）从实践中得到的收获。当然，开发大型系统的概念已经由来已久。事实上，在批处理处于全盛的时期，一些主要思想老早就被人们提出了，而且至今它们也不显得过时。事务处理概念是为解决单处理器联机应用的复杂性而被提出的。其实，这些概念对具有更高复杂性的海量分布系统的有效实现，显得更为至关重要。本书主要阐述事务概念是如何用于解决分布系统问题的，以及这些概念如何使我们在有限的资金和风险范围内，能够建立高性能、高可用性的应用。我们试图通过一些我们所熟悉的“历史教训”来阐述这一开发理念。归结一点，事务有助于掩盖，甚至避免分布系统的复杂性。

目前有很多的书是关于数据库系统（传统的数据库和分布数据库）、操作系统、计算机通信、应用开发的——还可以列很多。围绕这些术语的学科划分已深深地根植于世界各地的计算机科学的教学体系中。相应的教育和专家也按此分门别类。有关的书籍通常是以一种枚举的风格，将过去十多年在这一领域的技术文献中的主要思想罗列给读者。这种写作风格试图提供很多观点和可供选择的方案，但很少告诉读者什么观点是好的，什么是不怎么好的，以及为什么好与不好。更确切地说，你若要真的设计和实现一个实际系统，这些有关算法的一般综述几乎不能说明什么，例如，如何开始以及从哪里开始你的系统设计。

我们的宗旨是帮助读者解决实际问题。本书的重点在于，对问题只提供一两个解决方法和有关的说明，而对许多其他的方法并不一一涉及，也就是说本书不采用百科全书式的叙述风格。但本书的内容是广泛的，它将从一个系统的角度全面阐述事务处理的概念和技术。要使一个大型的系统正常运行，必须采用一种真正的“一条龙”的观点，即从一个请求开始，通过系统的各个层次和构件，直到结果被安全地输出为止。这其中要涉及终端上的表示管理、

通信子系统、操作系统、数据库、程序设计语言的执行系统以及应用开发环境等。设计一个集成度如此之高的系统，所需要的一套设计方案将完全有别于功能有限的算法设计。这种全面整体性的方法是在其他有关分布系统和数据库的著作中所未曾见到的。自1986年我们开始写作本书以来，我们一直深信这种方法是必需的。

内容的选择和组织

由于采用“一条龙”的观点，这需要我们覆盖很多基础性的东西，我们将重点放在事务处理的基本概念上：如简单的TP系统的结构化问题、简单的事务模型、简单的加锁、简单的写日志、简单的恢复等等。随便翻翻目前的教材和参考书，就会发现它们在基本概念上是模糊不清的。例如，在现有的教材中，我们还没有发现有关B树的真正实现的描述。而B树目前是用于数据库、文件系统、信息检索系统等之中的基本存取路径结构，是真正基本的东西。本书更像一本编译课程的教材或Tanenbaum的操作系统书。其中提供了很多代码片段用于说明基本算法和数据结构。

本书注重实效，详尽阐述并覆盖基本的事务处理问题。本书的写作使我们确信这是一种好的方法，不过这种写作方法和风格可能看起来比较陌生。我们的动机就是要以注重实效的观点阐述问题。目前没有结构化复杂系统的理论；因此重大决策不是依赖理论，更合适的做法是依赖良好的基础修养和依从好的工程原则——即注重实效的标准。我们相信，这些从面向事务处理的基本概念中导出的原则，会在未来若干年内显得至关重要。

本书主体部分有7个主题，共分16章，阅读上可或多或少地相互独立，并按不同的次序阅读。

第一个内容是事务处理的总体概述（第1章）。它讲述了总体的系统概貌。介绍了基本的事务特性：原子性（要么全做，要么全不做）、一致性（状态的正确转换）、隔离性（没有并发冲突）、永久性（已提交的更新不受任何系统故障的影响）——简称为ACID。没有技术背景的读者阅读到此，就能够对本领域有大致地了解。

第2章是为那些对计算机科学基本术语不甚熟悉的读者准备的。本章介绍了有关硬件、软件、协议标准等方面最为重要的术语和概念。所有这些术语是后续讨论所必需的。

第3章阐述为什么系统会失败，并给出如何避免此类故障的建议。本章回顾了硬件和软件的容错概念和技术（如彻底失败、冗余、模块化、修复）。若你想使用不完善的现成组件开发一个具有1万年平均无故障时间的模块，本章可以提供这方面的内容。第3章解释了在构建高可用软件时事务的重要意义。

第4~6章讲述了事务的有关理论和使用。第4章详细讨论了把应用系统构造为事务的具体含义。重点讨论了目前不能被扁平事务很好支持的计算类型。进行这些计算的应用需要事务概念的扩展和广义性。第5章通过描述TP监控器的作用，解释了面向事务的计算究竟对操作系统和其他低级别部件意味着什么。同时还解释了事务程序如何与这些系统服务交互作用。第6章是面向程序设计人员的。它提供了很多程序控制结构和代码片段来解释事务型的远程过程调用的工作原理、请求调度的处理，以及其他精妙之处。对位级（bit-level）事件不感兴趣的读者可略过本章的前半部分，直接从6.4节事务队列的内容开始阅读。

第7、8章介绍了并发的理论和实践。事务处理系统将应用中的所有并行执行部分掩盖起来，因此支持了ACID特性中的隔离性。第7章叙述了这些技术背后的理论，第8章则展示了如何用封锁技术实现这一理论。

第9~12章介绍了事务管理和恢复技术，即与事务原子性和永久性有关的内容。第9章讲述写日志和归档的处理。第10章讨论如何编写事务型资源管理器，如数据库系统或队列管理

器。它阐述了资源管理器如何与事务相结合，如何写日志记录、获得锁和参与事务的提交和回滚。书中给出了一个简单的资源管理器（一位资源管理器）来具体展示这方面的技术。第11章介绍事务管理器，它具体负责确定事务的状态和了解参与此事务的资源管理器的情况。本章给出了一个简单事务管理器的具体实现——当然对具体细节不感兴趣的读者同样可以略过这些内容。第12章是事务管理器所用到的高级概念和技术的汇总。

第13~15章讨论一个相对独立的内容：一个非常重要的资源管理器——事务型文件系统的实现。从裸金属（磁盘）开始，讨论空间管理问题，并详细论述了系统中缓冲管理器的角色。接下来，介绍了在定长页面中变长元组的组织，以及支持面向元组存取的所有文件组织方式。最后，第15章讨论了关联存取，主要介绍B树及其在高并行环境下的实现问题。所有这些都以一种支持ACID特性的方式实现，即产生的文件、元组和存取路径均是事务型对象。

第16章概述事务处理范畴内许多商业系统的情况。我们试图揭示每一系统的特征。这不是一个竞争性的比较，而是对每一系统优点的正面描述。

最后，附有一个关于事务处理术语的词汇表。拥有如此大量的词汇，这在一般的教材中是不常见的。然而，正如第5章引言所指出的，事务处理领域的术语是全面而定义完好的。因此词汇表可以起到两个目的：首先，若对某个术语不明确，可查词汇表找到有关的解释。其次，通过这种方式，词汇表有助于促进本领域术语的统一。

事务处理基础	1. 全书概述 2. 基本术语
容错基础	3. 故障和容错
面向事务的计算	4. 事务模型 5. 事务型的程序设计风格 6. TP监控器的工作原理
并发控制	7. 并发理论（可串行性） 8. 并发控制（加锁）
恢复	9. 写日志 10. 事务型资源管理器 11. 事务管理器 12. 高级TP技术
资源管理器实例	13. 文件和页面的实现 14. 记录和字段的实现 15. 存取路径的实现
系统综述	16. 流行TP系统的概述

大多数章包含了一节历史评注，试图说明事情的由来、一些思想的首次出处等。与计算机科学的许多其他领域不同，事务处理技术主要是由工业界研究机构和开发实验室提出的。

一些在商业产品中提出并实现的思想，是在若干年后被重新发现并作为科学成果加以发表的。在历史评注中，我们尽我们所知，尽量包含来自两方面的贡献。

本书大部分后章还配有习题，小到简单的内容复习题，大到一个学期项目。其中大部分题目还在章尾给出了答案。仿照唐纳德·克努特的做法，对每一道题目有一个形如[节，级别]的修饰。“节”用于说明习题涉及哪一节的内容，“级别”用于指明习题的难度系数：

[10] 指可以在1分钟时间完成的题目（用于检查是否阅读了书中的有关内容）。

[20] 指可以在15~20分钟时间完成的题目。

[25] 指可以在1小时时间完成的题目。

[30] 指短的（程序设计）项目：可以在一天内完成。

[40] 指较大型的（程序设计）项目：可以在两周时间完成。

另外，对难度超过[40]的题目，用[项目]标识。而对标有[讨论]的题目，则需要读者走出去，好好地调查研究一番。

本书原本的计划是要包含SQL的实现和有关事项的应用设计。随着写作的推进，我们意识到已没有空间和时间来讨论这些内容。于是，正如现在所看到的，本书就是一本关于事务如何实现的书。Ceri和Pelagatti有很出色的书[1984]讨论了高级别数据库的问题（如SQL、规范化、优化以及一些事务管理的问题）。

有关本书的学习

本书主要面向高年级本科生、研究生和一些可能需要了解所用事务处理系统情况的专业程序员（如CICS/DB2用户）或需要一本基本参考书的专业程序人员。

本书的内容在目前大学计算机科学的教学体系中没有对应的科目。分门别类是现有课程结构的本质特征，而且至今还没有标准的事务处理课程。然而，在本书的写作过程中，各个阶段的初稿已经被用于许多本科和研究生的课程中。我们认为我们所采用的写作方法在计算机科学现存结构中是合适的。但我们希望，这一方法的出现将有助于消除现有分门别类的思想。以下是将本书用于各种课程的建议：

基本了解：第1章，2.7、4.1~4.2、5.1~5.3、5.5~5.7节，第16章。

事务处理导论：第1、2章，3.1~3.6节，第4、5章，6.4~6.5、7.1~7.6节，第9、10、16章。

数据库系统：第1、4、7、10、13、14、15章。

分布式系统：第1、2、3、4、7、8、10、12、16章。

操作系统：第1、2、3、5、6、7、8、10、11、13、16章。

高级内容：在高级课程中，第1、3章可以略过，或者简要地过一遍。另外，可以采用Tanenbaum的有关操作系统和计算机网络的书，Ceri和Pelagatti的分布数据库的书，Osuzu和Valduries的分布数据库原理的书，和Date的任何一本数据库的书。

当然还可以想到许多其他的组合方式。对已经熟悉本领域内容的读者，可以略过第1章、第2章；不过我们建议大家浏览一下这部分内容，以便了解本书用到的术语。

第3章可以跳过，但我们建议大家阅读一下。事务乍看起来太显而易见。然而第3章仔细地阐述了为什么事务是正确的异常处理模型和开发高可用系统的关键。这里讨论的技术有助于更好地理解容错在系统级的含义，并且它们可以随时用于应用系统开发之中。

我们在叙述中尽量保持各章内容互不依赖。如果你阅读了全书，就会发现各章间有一些重复内容。这保证了各章使用上的独立性。

为使本书更便于教学上的使用，教师可以考虑通过Transarc公司的Encina大学项目获取有关