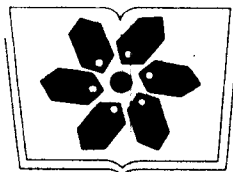


李建中 孙文隽 著

并行关系 数据库 管理系统引论

数据库丛书

科学出版社



中国科学院科学出版基金资助出版 国家自然科学基金委员会资助出版

数据库丛书

并行关系数据库管理系统引论

李建中 孙文雋 著

科学出版社

1998

内 容 简 介

本书全面地总结了 80 年代以来有关并行关系数据库管理系统的研究成果,系统地论述了有关并行关系数据库管理系统的理论、技术和方法,为并行关系数据库管理系统的基础研究、设计实现和应用开发提供了有价值的参考资料。本书首先讨论了并行计算结构、并行数据库机器和并行数据库的基本概念,然后着重论述了并行关系数据库管理系统的实现技术,包括并行数据库的物理组织、并行数据操作算法、并行查询优化方法等。

本书既可作为计算机专业研究生或高年级本科生数据库课程的教材或参考书,也可以供从事数据库教学与科研工作的科技工作者参考。

图书在版编目(CIP)数据

并行关系数据库管理系统引论/李建中,孙文隽著.北京:
科学出版社,1998.6
(数据库丛书)
ISBN 7-03-006522-0

I. 并… II. ①李… ②孙… III. 关系数据库-数据库管
理系统-基本知识 IV. TP311.13

中国版本图书馆 CIP 数据核字(98)第 05735 号

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1998 年 7 月第 一 版 开本:787×1092 1/16

1998 年 7 月第一次印刷 印张:16 1/2

印数:1—2 800 字数:365 000

定价:25.00 元

《数据库丛书》是我国数据库专家学者团结协作、合力撰写的。一套系列著作，它比较全面地反映了国际数据库技术的丰富内容与最新发展，和我国数据库科技工作者多年来的主要研究成果，具有较高的理论水平和学术价值。

数据库是计算机科学技术中发展最快的领域之一，也是应用最广泛的技术之一。是计算机信息系统与应用系统的构成基础。相信《数据库丛书》的编辑出版，必将有助于推动我国数据库技术的研究与发展，促进我国数据库技术的普及与提高，加快数据库应用的推广与深入，为我国社会经济、信息化作出贡献。

張致祥

一九八二年六月

《数据库丛书》编委会

主 编	萨师煊			
副主编	罗晓沛	王 珊		
编 委	王能斌	施伯乐	郑怀远	童 颖
	唐世渭	周立柱	徐秋元	周龙骧
	徐洁磐	郑振楣	何新贵	马应章
	李建中	张大洋	董继润	瞿兆荣
	张作民	何守才	姚卿达	唐常杰
	冯玉才	尹良琨	杨冬青	邵佩英
	李昭原	周傲英	于 戈	

序

数据库是计算机领域发展最快的学科之一,因为它既是一门非常实用的技术,也是一门涉及面广、研究范围宽的学科。因此,它吸引了理论研究、系统研制和应用开发等不同方面众多的学者、专家和技术人才致力于其研究和实践。

数据库系统所管理、存储的数据是各个部门宝贵的信息资源。在信息化时代来临、Internet高速发展的今天,信息资源的经济价值和社会价值越来越明显。建设以数据库为核心的信息系统和应用系统,对于提高企业的效益、改善部门的管理、改进人们的生活均具有实实在在的意义。正因为数据库技术与经济、社会的发展和信息化建设有着密切的关系,这门学科才获得了巨大的源动力和深厚的应用基础。

数据库系统已从第一代网状、层次数据库系统发展到第二代关系数据库系统和第三代以面向对象为主要特征的数据库系统。数据库技术与网络通信技术、面向对象技术、并行计算技术、多媒体技术、人工智能技术等互相渗透,互相结合,成为当前数据库技术发展的主要特征。它使数据库领域中新的技术内容层出不穷,新的学科分支不断涌现,形成了新一代数据库系统的大家族。与传统的数据库相比,当今数据库的整体概念、技术内容、应用领域,甚至某些原理都有了重大的发展和变化。

面对如此丰富的学术内容和技术方法,如此广阔的研究方向和应用领域,从事数据库研究、开发和应用的科技人员,攻读数据库方向的研究生都迫切希望有一套丛书能系统而全面地介绍数据库学科的多个分支和相关领域。

《数据库丛书》的编写宗旨是把当前数据库学科各个分支的最新学术成果介绍给读者,以促进国内的学术研究;同时,又介绍数据库技术的发展过程,各分支之间的内在联系及在数据库大家族中的位置,以促进数据库和计算机科学的其他领域技术的结合。

本丛书由各分册组成,包括《数据库进展》、《分布式数据库》、《分布式数据库管理系统实现技术》、《并行关系数据库管理系统引论》、《数据仓库技术与联机分析处理》等。本丛书的每一分册涉及数据库学科的一个或几个分支。其中《数据库进展》则与其他分册有所不同,是本丛书的总纲、指南和补充,是给本丛书穿针引线、铺垫基础,从而使丛书成为一个各部分既相互独立又相互联系的整体。

《数据库丛书》是开放的,故丛书的分册将随着数据库学科的发展而不断补充。

本丛书各分册的主编和作者,多是长期从事数据库各分支领域研究工作的专家、学者。他们学术造诣高深,实践经验丰富,书中许多内容是他们长期研究的成果。本丛书不仅反映了国际数据库技术的最新成果和发展方向,也展示了我国数据库工作者的学术成果和研究深度,具有较高的理论水平和学术价值。它的出版是我国数据库学术界的一件大喜事。我向本丛书的所有作者和编委的辛勤工作表示崇高的敬意。

萨师煊

1998年1月

前 言

80年代以来,并行计算机系统的发展十分迅速,很多商品化的高性能并行计算机系统已经投入市场。但是,并行计算机软件系统的发展却远远落后于硬件系统的发展,至今尚未出现令人满意的商品化的并行操作系统、并行编译系统、并行数据库系统、并行智能系统等并行软件系统。最近几年,并行软件系统的基础理论和实现技术已经成为计算机科学技术领域的重要研究课题,引起了学术界和工业界两方面的重视,越来越多的计算机科学技术工作者开始投入到并行软件系统的研究行列。由于在各种应用领域中数据库管理系统占有重要的地位,并行数据库系统的研究引起了人们特别的关注。值得注意的是,随着计算机应用领域的不断扩大,数据库规模越来越大,数据库查询也越来越复杂,对数据库系统性能的要求越来越高。传统数据库系统已经难以适应迅速增长的性能要求。数据库系统性能的提高已经成为目前急需解决的问题。随着并行计算机系统的出现,许多数据库研究者认为,在并行计算机上实现数据库管理系统(即并行数据库管理系统)能够极大地提高数据库管理系统的性能,并行数据库管理系统有希望成为未来的高性能数据库管理系统。并行数据库管理系统已经成为一个新的数据库研究领域。

并行数据库系统的研究可以追溯到数据库机器的研究。数据库机器的研究主要集中在数据库管理系统专用硬件的研究。由于磁盘和处理机之间速度差别的迅速增长和数据库机器硬件设计的高开销,这方面的研究受到了很大的限制,专用数据库机器的研究收效甚微。但是,随着通用并行计算机系统的出现,并行数据库管理系统的研究取得了引人注目的进展。几个并行数据库原型系统已经成功地运行在通用并行计算机上。并行数据库系统已经成为并行计算机的主要应用系统之一。很多研究者认为,在具有数百个处理机的并行计算机上建立高性能数据库系统是可行的。

80年代以来,人们在并行数据库管理系统方面开展了大量研究,取得了很多研究成果。最近几年,随着并行计算机系统的迅速发展,人们对并行数据库系统的研究兴趣有增无减,大有方兴未艾之势。遗憾的是,至今国内外尚无一部有关并行数据库系统的学术著作问世。为了满足广大数据库理论、工程和应用工作者的要求,我们集中了国内外有关并行数据库管理系统研究的学术论文和我们多年来的研究成果,经过一年多的系统研究整理,写成了本书。

本书旨在为已经具备一定数据库基础知识的读者提供一本系统、深入介绍并行关系数据库管理系统基础理论和实现技术的参考书。本书总结了80年代以来并行数据库的主要研究成果,系统地论述了并行数据库系统的理论、技术和方法,包括并行数据库的物理组织,并行数据操作算法的设计、分析与实现,并行数据库查询优化等。

本书共六章。第一章介绍数据库系统的进展、新一代数据库的特点和传统计算机结构的不适应性,并综述了数据库机器的主要研究成果,概述了并行数据库系统的研究进展,论述了支持并行数据库系统的并行计算结构和并行数据库管理系统的软件结构。第二章介绍并行计算机系统结构和并行算法设计与分析的基本概念,包括向量计算机系统结构、

多处理机并行计算机系统结构、基于高速通信网络的计算机机群并行计算环境等。第三章论述并行数据库的物理存储结构,包括传统数据库物理存储结构的并行化方法和各种新的并行数据库物理存储结构。第四章论述实现各种关系数据库数据操作的并行算法的设计和析技术。第五章论述实现并行关系数据库管理系统的并行数据流技术。第六章论述并行数据库查询优化的理论、技术和方法。

本书不仅凝集着作者的心血,也凝集着所有对本书撰写给予鼓励、支持、关心和帮助过的科技人员的心血。

值此本书问世之际,谨向给予我们鼓励、支持和帮助的国内外数据库专家、学者和同行表示诚挚的谢意。我们特别感谢罗晓沛教授在本书的写作过程中给予的多方面指导和帮助。我们也特别感谢王珊教授,她对本书进行了认真审阅并提出了十分宝贵的意见。

我们还由衷地感谢国家自然科学基金委员会、国家高科技“863”计划有关专家组、国家教育委员会和黑龙江省科学技术委员会的支持。从1993年以来,我们的研究工作多次得到这些部门的资助。

我们还要感谢我们的94和95级研究生,他们对我们的研究工作给予了强有力的帮助,在本书的写作过程中,他们也给予了很大的支持。

最后向李建中的妻子石敏同志和孙文隽的丈夫闫跃同志表示深深的谢意。如果没有他们的理解和支持,我们就不可能专心致志地从事研究工作,这本书也就不会诞生。

本书的第一、二、三、五、六章由李建中同志撰写,第四章由孙文隽同志撰写。李建中同志最后对本书进行了统编。由于作者水平有限,书中难免存在欠妥和不当之处,恳切希望同行和广大读者批评指正。

目 录

第一章 数据库机器和并行数据库系统概述	1
1.1 新一代数据库应用的特点	1
1.1.1 复杂数据对象	2
1.1.2 规则管理	2
1.1.3 数据模型中的新概念	3
1.1.4 巨型数据库	3
1.1.5 第三级存储器	3
1.1.6 长事务处理	4
1.1.7 版本与格局	4
1.2 Von Neumann 计算机系统结构的不适应性	4
1.2.1 数据存取问题	5
1.2.2 并行处理问题	5
1.2.3 处理器瓶颈问题	6
1.2.4 存储器瓶颈问题	6
1.2.5 I/O 瓶颈问题	7
1.2.6 总结	7
1.3 数据库机器	8
1.3.1 数据库机器的分类	8
1.3.2 具有数据库级索引的数据库机器	10
1.3.3 具有关系级索引的数据库机器	13
1.3.4 具有页索引级的数据库机器	15
1.3.5 数据库机器的问题和意义	16
1.4 并行数据库系统概述	18
1.4.1 并行数据库系统的进展	19
1.4.2 关系数据库系统的固有并行性	19
1.4.3 支持并行数据库系统的并行计算结构	21
1.4.4 并行数据库系统的软件结构	22
1.5 文献阅读导引	23
参考文献	23
第二章 并行计算机系统与并行算法	27
2.1 并行计算机系统的发展	27
2.2 并行计算机系统的分类模型	28
2.2.1 SISD 计算机系统	28
2.2.2 MISD 计算机系统	28
2.2.3 SIMD 计算机系统	29
2.2.4 MIMD 计算机系统	31

2.3	连接网络	32
2.3.1	动态连接网络	32
2.3.2	静态连接网络	35
2.3.3	静态连接网络的性能分析	39
2.4	计算机机群并行计算系统	42
2.4.1	机群并行计算系统的特点	42
2.4.2	机群并行计算系统的高速通信网络	43
2.4.3	机群并行计算系统的硬件环境	43
2.4.4	机群并行计算系统的软件环境	46
2.5	并行算法及其复杂性分析	49
2.5.1	并行执行时间	50
2.5.2	工作量	50
2.5.3	加速比	50
2.5.4	效率	51
2.5.5	伸缩性	51
2.5.6	算法复杂性测度函数的阶	51
2.6	文献阅读导引	52
	参考文献	53
第三章	并行数据库的物理存储方法	57
3.1	一维数据分布方法	57
3.1.1	Round-Robin 分布方法	57
3.1.2	Hash 分布方法	58
3.1.3	Range 分布方法	58
3.1.4	Hybrid-Range-Partition 分布方法	59
3.2	多维数据分布方法 CMD	62
3.2.1	基本概念	62
3.2.2	CMD 数据分布方法	64
3.2.3	CMD 方法的性能分析	66
3.2.4	单处理机上的数据组织	70
3.2.5	数据更新操作	71
3.3	基于错误校正码的多维数据分布方法	71
3.3.1	线性错误校验码	71
3.3.2	ECC 数据分布方法	73
3.4	其它多维数据分布方法	74
3.4.1	随机数据分布方法	74
3.4.2	一般化的 CMD 数据分布方法	74
3.4.3	BM 数据分布方法	75
3.4.4	FX 数据分布方法	75
3.4.5	基于 Hilbert 曲线的数据分布方法	76
3.4.6	BERD 多维数据分布方法	77
3.5	并行 B-树	78
3.5.1	大结点并行 B-树	78

3.5.2 基于记录分布的并行 B-树	80
3.5.3 基于树结点分布的并行 B-树	82
3.6 简单并行 GRID 文件	86
3.6.1 SPGRID 文件结构的初始构造算法	86
3.6.2 动态数据平衡算法	88
3.7 两级并行 GRID 文件	90
3.7.1 逻辑并行 GRID 文件	91
3.7.2 物理并行 GRID 文件	91
3.7.3 并行 GRID 文件结构的构造和数据加载算法	92
3.8 文献阅读导引	101
参考文献	101
第四章 并行数据操作算法	104
4.1 并行排序算法	105
4.1.1 基于合并操作的并行排序算法	105
4.1.2 基于比较-交换操作的并行排序算法	108
4.1.3 基于数据划分的并行排序算法	110
4.2 并行选择、投影和集合操作算法	113
4.2.1 并行选择算法	113
4.2.2 并行投影算法	115
4.2.3 并行集合操作算法	117
4.3 并行连接算法	120
4.3.1 并行嵌套循环连接算法	120
4.3.2 并行排序合并连接算法	122
4.3.3 并行 Hash 连接算法	124
4.4 抗数据偏斜的并行连接算法	128
4.4.1 简单抗数据偏斜并行连接算法	129
4.4.2 并行 Tuple-Interleaving-Hash 连接算法	130
4.4.3 并行 Adaptive-Load-Balancing-Hash 连接算法	131
4.4.4 扩展的并行 Adaptive-Load-Balancing-Hash 连接算法	135
4.5 基于 CMD 并行存储结构的并行连接算法	137
4.5.1 基本概念	138
4.5.2 CMD-Join 算法	139
4.6 基于并行 B-树存储结构的并行连接算法	144
4.6.1 基于 RANGE 划分策略的并行 B-树连接算法	144
4.6.2 基于 Hash 划分策略的并行 B-树连接算法	150
4.7 文献阅读导引	152
参考文献	153
第五章 并行查询优化处理的数据流方法	157
5.1 顺序数据库操作算法	157
5.1.1 选择操作的实现算法	157
5.1.2 笛卡尔乘积的实现算法	159

5.1.3	连接操作的实现算法	161
5.1.4	投影操作的实现算法	164
5.1.5	集合的并、交、差实现算法	165
5.2	数据库操作结果大小的估计	167
5.3	简单并行数据流方法	171
5.3.1	简单并行数据流图	171
5.3.2	处理机分配算法	172
5.3.3	并行查询计划的调度执行算法	173
5.3.4	处理机之间的通信	175
5.3.5	改进的处理机分配和查询计划调度执行算法	176
5.4	基于数据划分的并行数据流方法	178
5.4.1	DPBPDF 方法概述	178
5.4.2	DPBPDF 方法的查询处理器	179
5.4.3	执行顺序图的构造算法	181
5.4.4	数据操作划分算法	183
5.4.5	分解与合并操作连接算法	187
5.5	文献阅读导引	191
	参考文献	191
第六章	并行数据库查询优化方法	192
6.1	引言	192
6.2	两阶段查询优化方法	193
6.2.1	以最小化查询执行时间为目标的方法	193
6.2.2	以充分利用系统资源为目标的优化算法	198
6.3	基于左线性树的查询优化方法	201
6.3.1	查询执行计划模型和查询执行计划空间	202
6.3.2	左线性树查询执行计划的复杂性模型	203
6.3.3	查询优化算法	206
6.3.4	存储空间限制问题	207
6.4	基于右线性树的查询优化算法	208
6.4.1	查询执行计划模型和查询执行计划空间	208
6.4.2	右线性树查询执行计划的复杂性模型	208
6.4.3	查询优化算法	210
6.5	基于片段式右线性树的查询优化方法	211
6.5.1	查询执行计划模型和查询执行计划空间	211
6.5.2	SRD 树查询执行计划的复杂性模型	212
6.5.3	基于 SRD 树的查询优化算法	213
6.6	基于浓密树的查询优化算法	216
6.6.1	查询执行计划模型和查询执行计划空间	216
6.6.2	查询优化算法	217
6.7	基于多重加权树的查询优化方法	221
6.7.1	查询执行计划模型和查询执行计划空间	221

6.7.2 多重加权树并行查询计划的复杂性模型	225
6.7.3 并行查询优化算法	229
6.8 文献阅读导引	243
参考文献.....	244

第一章 数据库机器和并行数据库系统概述

随着计算机应用领域的迅速扩大,数据库规模日益增长,用户查询越来越复杂,对数据库管理系统性能(处理能力和速度)的要求越来越高。基于 Von Neumann 计算机的数据库管理系统已经难以适应迅速增长的性能要求,提高数据库管理系统性能已经成为急需解决的课题。长期以来,人们一直在探索冲破 Von Neumann 计算机系统结构的束缚,建立高性能数据库系统的新路。70 年代中期到 80 年代中期,数据库研究者在数据库机器方面开展了大量的研究,主要集中在关系代数操作的并行化和实现这些操作的专用硬件的设计。由于磁盘和处理器之间速度差别的迅速增长以及数据库机器硬件设计的高开销,这方面的研究受到了很大限制,数据库机器的研究未能达到人们所希望的目标。最近十几年来,并行计算机系统的发展十分迅速,很多商品化的高性能并行计算机系统已经问世。并行计算机系统的出现为高性能数据库管理系统的实现带来了希望。数据库研究者开始把注意力转向以并行计算机为基础的数据库系统(称为并行数据库系统)的研究。并行数据库系统试图通过充分利用多处理机和多磁盘的并行性提高处理能力和速度。目前,并行数据库系统已经成为一个新的数据库研究领域,并取得了一些引人注目的进展。几个建立在通用并行计算机系统上的并行数据库原型系统已经取得成功。很多研究者认为,在具有数千个处理器的并行计算机上建立高性能数据库系统是可行的。并行数据库系统有希望成为未来的高性能数据库系统。

本章首先讨论新一代数据库应用的特点和 Von Neumann 计算机系统结构的不适应性,然后综述有关数据库机器的主要研究成果,最后概述并行数据库系统。

1.1 新一代数据库应用的特点

从 60 年代末期开始,数据库系统已经走过了 30 年的历程,经历了两代演变。第一代数据库系统是层次与网络数据库系统。第二代数据库系统是关系数据库系统。30 年来,人们主要致力于第二代数据库系统的理论研究和系统开发,取得了辉煌的成就。完整的关系理论的确立标志着关系数据库系统理论趋于完善。商品化的 DB2、INGRES、ORACLE、SYBASE 等关系数据库管理系统的出现,标志着关系数据库管理系统已经接近顶峰。第一代和第二代数据库系统的设计目标源于商业事务处理。多年来,这两代数据库系统主要用于事务处理。80 年代以来,随着计算机应用领域的迅速扩大,新的数据库应用,如工程数据库应用、地理信息系统、多媒体信息系统、科学与统计数据应用等层出不穷。我们称这些应用为新一代数据库应用。新一代数据库应用对数据库系统的各个方面都提出了新的要求。以下我们称满足新一代数据库应用要求的数据库系统为新一代数据库系统。本节讨论新一代数据库应用的主要特点及其对新一代数据库系统的要求。

1.1.1 复杂数据对象

新一代数据库应用需要存储和使用很多新型数据对象,这些新的数据对象多数都是规模庞大的复杂数据对象。例如,多媒体信息系统需要存储和管理图像、正文、声音等复杂数据对象。科学与统计数据库应用系统需要存储和管理大型矩阵、数组和向量等复杂数据对象。软件工程数据库应用系统需要存储和管理各种级别的程序版本和文档等复杂数据对象。化学数据库应用系统需要存储和管理诸如蛋白质结构一类的复杂数据对象。如此众多的新型数据对象,对新一代数据库系统提出了一系列的新要求,如:

(1) 直接支持图像、声音、文本、矩阵、向量、时间序列等复杂数据对象。

(2) 支持新一代数据库查询语言。新一代数据库查询语言能够直接支持基于复杂数据对象的查询。相应地,新一代数据库系统必须包括新一代数据库查询语言的编译和优化处理技术。

(3) 具有适应于复杂数据对象的存储结构、索引技术、存取方法以及在数据库与应用程序之间传递复杂数据对象的方法。

(4) 具有数据类型管理能力,为程序员提供构造复杂数据类型的机制。为此,新一代数据库系统需要提供灵活的类型系统。这样的类型系统一直是面向对象数据库系统的重要研究任务。目前已出现的数据类型管理系统具有一个共同的问题:类型的正确性检查是动态进行的,这种动态检查会使很多程序设计错误不能在编译时发现,而在程序运行时才能发现。为了给数据库应用程序提供一个类似于高级程序设计语言所提供的安全程序设计环境,我们需要把程序设计语言编译器的静态数据类型正确性检查技术应用到数据库管理系统。

1.1.2 规则管理

新一代数据库应用需要管理和使用大量规则。这些规则可以分为两种类型。一种是状态性规则,即型为“如果 A 真,则 B 真”的规则。另一种是行为性规则,即型为“如果 A 真,则执行 C”的规则。状态性规则不要求完成任何动作。行为性规则在条件成立时需要执行一个过程,完成一系列的动作。例如,在工程设计应用中,如果一个设计者对他的设计进行了修改,而这些修改影响到另一个设计者的设计,数据库系统则应该通知这个设计者。这个“通知”动作可以由行为性规则来实现。显然,工程数据库系统必须支持多种状态性和行为性规则的存储和处理。这些规则分别表示设计必须遵循的约束规则、特殊事件发生时必须激活的动作、系统内部的演绎推理知识等。规则处理已经引起了数据库界的极大注意。规则一直被用于触发器(Trigger)的表示、数据库中的知识发现和其它有关数据库推理机制的说明。状态性规则具有很多优点。这种规则只是逻辑地说明用户的要求,而不要求详细描述如何得到预期的结果。类似地,行为性规则提供了说明动作执行条件的逻辑机制。这种说明性的价值在关系查询语言中已经得到了充分的体现。在新一代查询语言中保持并扩充这种说明性是非常重要的。

传统的规则处理方法使用独立的子系统来完成规则的处理。这种方法不能支持数据库应用中的规则处理要求。例如,如果数据库管理系统不能直接处理规则,我们在工程设计例子中提到的有关“通知”动作的行为性规则就很难得到有效的支持。如何在数据库系

统中定义规则,如何有效地处理大型规则库,是需要深入研究的问题。虽然人工智能界已经作出了很大的努力,但是他们的研究都是在所有数据和规则皆留驻内存的假定下进行的。由于新一代数据库应用需要处理的数据和规则的数量极大,无法全部存储在内存,人工智能界提出的技术不能满足新一代数据库系统的要求。

在一个大型规则库中,增加一个规则可能引起规则库的不一致性,也可能引起混沌和不希望的结果,甚至引起所增加规则无休止的反复启动。我们需要研究开发大型规则库的调试和验证工具,也需要研究分解大型规则库的技术。

1.1.3 数据模型中的新概念

新一代数据库应用涉及很多新概念,例如数据的时间概念或历史数据概念、数据的空间概念等。这些概念都是目前的数据模型所不能支持的。我们需要从理论和实现两个方面研究如何把这些新概念扩充到现有的数据模型中,或者研究包括这些新概念的新数据模型。

1.1.4 巨型数据库

新一代数据库应用常常需要存取巨型数据库。巨型数据库的容量比目前最大的数据库还要高出几个数量级。大于 10^{12} 字节的数据库将成为很普通的数据库。目前的数据库管理系统不能有效地处理这样巨大的数据库。巨型数据库上的查询处理需要相当长时间,例如一个 10^{12} 字节数据库的一次扫描可能需要几天的时间。我们不能让一个用户在星期一早上提交一个实时查询,而在星期四晚上才能得到回答。下边,我们举两个例子来说明与巨型数据库有关的棘手问题。

例 1 在目前的数据库管理系统中为一个给定的关系建立索引,系统需要执行以下三个操作:(1) 对给定关系加锁;(2) 为给定的关系建立索引;(3) 释放在给定关系上所加的锁。如果按照这种方式为一个具有 10^{12} 字节的数据库建立一个索引可能需要几天的时间,在这几天内,该关系不能被使用。

例 2 为了在系统发生故障时保证数据库的正确性,数据库管理系统的数据库恢复子系统需要经常把磁盘上的数据库复制到磁带上。当系统发生故障并修复后,数据库恢复子系统把磁带上的数据库最新副本恢复到磁盘上。把一个 10^{12} 字节的数据库复制到磁带上或从磁带上恢复到磁盘上,可能需要几天的时间。使用目前的数据库恢复技术,每当数据库恢复子系统把一个 10^{12} 字节的数据库复制到磁带上或从磁带上恢复到磁盘上,数据库管理系统就要在几天之内不能正常工作。

我们还可以举出很多例子来说明当前的数据库系统在很多方面都不能满足巨型数据库的管理要求。

1.1.5 第三级存储器

巨型数据库要求数据库管理系统在主存储器和辅助存储器(如磁盘)的基础上增加第三级巨型存储器,如机器人磁带库和机器人光盘库等。目前的数据库管理系统都要求数据存储在磁盘或内存中,磁带等大容量存储器只作为后备存储器使用。在新一代数据库系统中,巨型数据库只有少部分数据留驻在主存储器或磁盘上,大部分数据将存储在第三级存储

器上。在这种情况下,查询处理与优化将成为更复杂的问题。第三级存储器的存取速度要比主存储器和磁盘存储器慢。所以,查询优化处理程序所选择的查询执行策略必须避免数据在存储介质之间的频繁传递。此外,新一代数据库系统还必须优化数据记录在第三级存储器上的分布,降低第三级存储器中数据的选择和存取时间。在包括第三级存储器的数据库系统中,磁盘通常作为第三级存储器的读写缓冲器。新一代数据库系统需要具有三级存储器系统的高效缓冲处理技术。

1.1.6 长事务处理

我们称需要运行很长时间的事务为长事务,例如一个计算机辅助设计应用事务可能要持续几个小时甚至几天。传统数据库管理系统中的事务处理技术很难处理长事务。例如,如果我们使用两段锁技术实现并发控制,则当一个需要几天时间的长事务运行时,数据库中被这个事务加锁的数据项将在几天内不允许其它事务存取。如果这个事务需要存取的数据量很大,数据库系统的总体性能将极大地下降。新一代数据库系统需要新的长事务管理技术,如并发控制技术、数据库恢复技术等。

1.1.7 版本与格局

很多新一代数据库应用需要数据库对象(如关系、索引等)的版本和格局的概念。数据库对象的一个版本表示这个对象的一个状态。例如,关系 R 中在某个时刻的所有数据项的特定值是 R 的一个版本。一个数据库对象可以具有多个不同的版本。例如,在一个建筑设计数据库中,一个楼房的供电系统设计可能具有许多可供选择的设计方案,每个设计方案都是这个供电系统设计的一个版本。从而,供电系统设计这个数据库对象具有多个版本。

数据库中一般都存储很多相互有关的数据库对象。这些对象的全体描述了一个更高级的对象。例如,在上述建筑设计数据库中,一个楼房的设计包括这个楼房的供电系统设计、给排水系统设计、楼房的整体结构设计等对象。这些对象描述了“楼房设计”这个更高级的对象。数据库中相互有关对象的特定版本的集合称为它们所描述的高一级对象的一个格局。例如,一个供电系统设计的特定版本、一个给排水系统设计特定版本和其它有关设计的特定版本组合起来就形成了楼房设计的一个格局。这个格局表示了整个楼房的一种设计方案。

目前,虽然已经提出了一些有关版本和格局的数据模型,但是没有引起人们的注意。新一代数据库系统需要具有版本和格局概念的数据模型。此外,新一代数据库系统还需要具有较小时间和空间复杂性的版本管理算法和保证格局一致性的技术。

1.2 Von Neumann 计算机系统结构的不适应性

在 1.1 节我们已经看到,新一代数据库具有数据量大和结构复杂的特点,新一代数据库应用需要非常复杂的数据处理操作和高效率的事务处理能力。新一代数据库和新一代数据库应用需要由高性能数据库系统来支持。从 70 年代开始,人们就一直在探索实现高性能数据库管理系统的理论、技术和方法,但是,始终没有得到满意的结果,其原因在于所