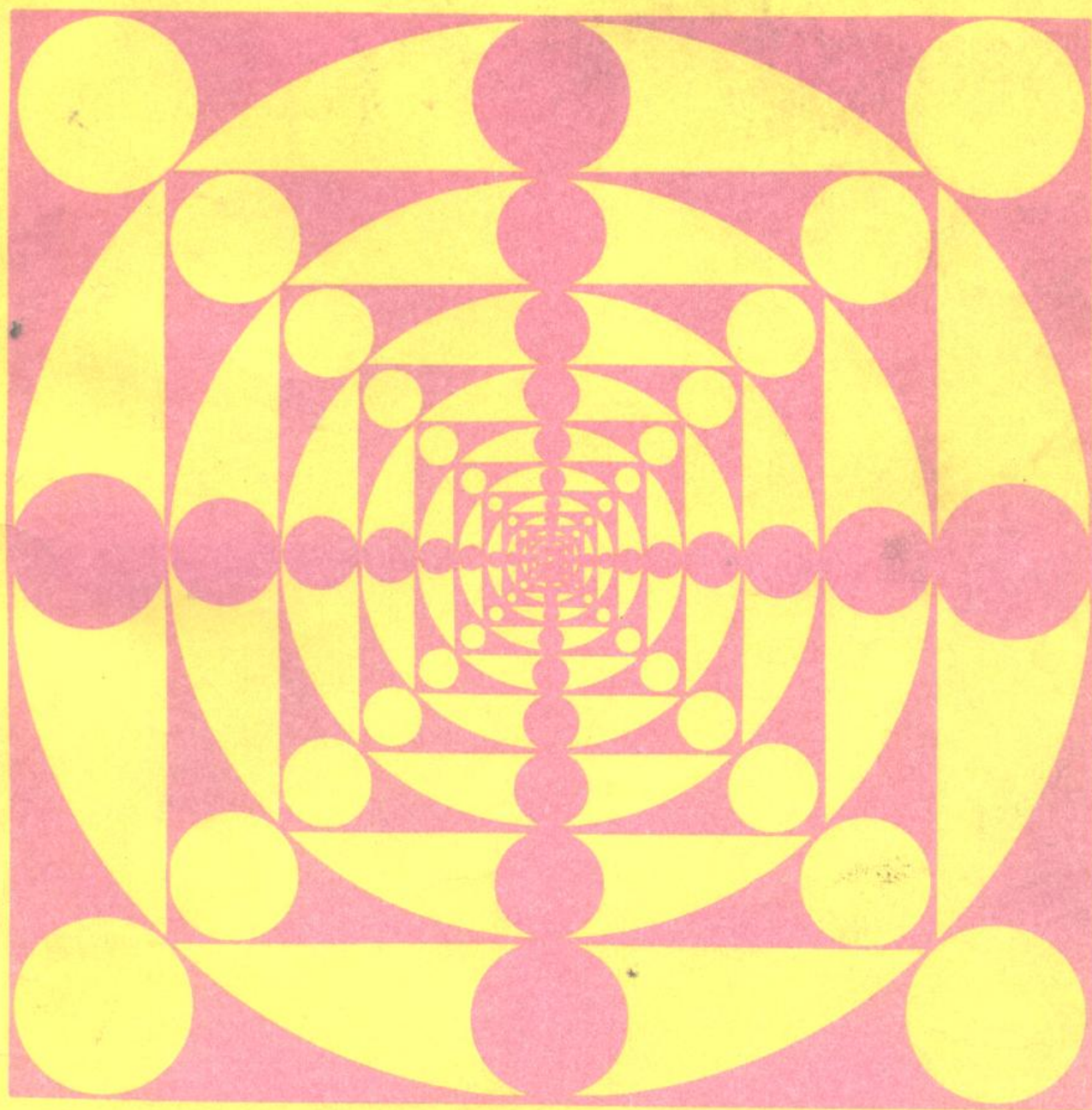


电子计算机应用系列教材

分布式数据库系统概论

李平安 张明 罗州平 编著



科学出版社

电子计算机应用系列教材

分布式数据库系统概论

李平安 张 明 罗州平 编著

科学出版社

1992

(京)新登字092号

内 容 简 介

本书是电子计算机应用系列教材之一。书中比较全面地介绍了分布式数据库系统的基本概念，并以国外的一个具体分布式数据库系统为例详细阐述了在微机局部网环境下分布式数据库系统的结构、数据描述语言、数据操纵语言，以及各部分的实现思想和方法。主要内容包括：MICROBE系统介绍、MIQUEL语言、优化、分解定位与并行处理、关系操作与存贮、分布式系统以及系统恢复。

本书可作为计算机培训班的教材，也适宜从事研制、开发数据库系统和分布式数据库系统的科技人员及大专院校有关专业的师生参考。

JS46/68
22

电子计算机应用系列教材 分布式数据库系统概论

李平安 张 明 罗州平 编著

责任编辑 李 立

科学出版社出版

北京东黄城根北街16号

邮政编码：100707

北京市华星计算机公司激光照排

天津市静一胶印厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1992年8月第 一 版 开本：787×1092 1/16

1992年8月第一次印刷 印张：12 3/4

印数：1—5 700 字数：288 000

ISBN7-03-001354-9/TP. 89

定价：10.50元

“电子计算机应用系列教材”主持、组织编著单位

主持编著单位:

国务院电子信息系统推广应用办公室

组织编著单位:

广东、广西、上海、山东、山西、天津、云南、内蒙古、
四川、辽宁、北京、江苏、甘肃、宁夏、江西、安徽、
河北、河南、贵州、浙江、湖北、湖南、黑龙江、福建、
新疆、广州、大连、宁波、西安、沈阳、武汉、青岛、
哈尔滨、重庆、南京等 35 省、市、自治区、计划单列市

电子振兴

计算机领导小组办公室

科技工作

“电子计算机应用系列教材”联合编审委员会名单

(以姓氏笔划为序)

主编审委员:

王长胤* 苏世生 何守才 陈有祺 陈莘萌* 邹海明* 郑天健
殷志鹤 童 颖 赖翔飞 (有“*”者为常务主编)

常务编审委员:

于占涛 王一良 冯锡祺 刘大昕 朱维华 陈火旺 陈洪陶 余 俊
李 祥 苏锦祥 佟震亚 张广华 张少润 张吉生 张志浩 张建荣
钟伯刚 胡秉光 高树森 徐洁盘 曹大铸 谢玉光 谢育先 韩兆轩
韩培尧 董继润 程慧霞

编审委员:

王升亮 王伦津 王树人 王振宇 王继青 王翰虎 毛培法 叶以丰
冯鉴生 刘开瑛 刘尚威 刘国靖 刘晓融 刘德镇 孙令举 孙其梅
孙耕田 朱泳岭 许震宇 何文兴 陈凤枝 陈兴业 陈启泉 陈时锦
邱玉辉 吴宇尧 吴意生 李克洪 李迪义 李忠民 迟忠先 沈林兴
肖金声 苏松基 杨润生 芮福德 张志弘 张银明 张 勤 张福源
张翼鹏 郑玉林 郑 重 郑桂林 孟昭光 林俊伯 林钧海 周俊林
赵振玉 赵惠溥 姚卿达 段银田 钟维明 袁玉馨 唐肖光 唐楷全
徐国平 徐拾义 康继昌 高登芳 黄友谦 黄 侃 程锦松 楼朝城
潘正运 潘庆荣

秘书组:

秘 书 长:胡茂生
副秘书长:何兴能 林茂荃 易 勤 黄雄才

序

当代新技术革命的蓬勃发展,带来社会生产力新的飞跃,引起整个社会的巨大变革.电子计算机技术是新技术革命中最活跃的核心技术,在工农业生产、流通领域、国防建设和科学研究方面得到越来越广泛的应用.

党的十一届三中全会以来,我国计算机应用事业的发展是相当迅速的.到目前为止,全国装机量已突破 30 万台, 16 位以下微型计算机开始形成产业和市场规模,全国从事计算机科研、开发、生产、应用、经营、服务和教学的科技人员已达十多人,他们在工业、农业、商业、城建、金融、科技、文教、卫生、公安等广阔的领域中积极开发应用计算机技术,取得了优异的成绩,创造了显著的经济效益和社会效益,为开拓计算机应用的新局面作出了重要贡献.实践证明,人才是计算机开发应用的中心环节.我们必须把计算机应用人才的开发与培养放在计算机应用事业的首位,要坚持不懈地抓住人才培养这个关键.

从目前来看,我国计算机应用人才队伍虽然有了很大的发展,但是这支队伍的数量和质量还远不适应计算机应用事业发展的客观需要,复合型人才的培养与教育还没有走上规范化、制度化轨道,教材建设仍显薄弱,培训质量不高.因此,在国务院电子信息系统推广应用办公室领导、支持下,35 个省、市、自治区、计划单列市计算机应用主管部门共同组织 118 所大学和科研单位的 400 多位专家、教授编写了全国第一部《电子计算机应用人才培养大纲》以及与之配套使用的“电子计算机应用系列教材”,在人才培训和开发方面做了一件很有意义的工作,对实现培训工作规范化、制度化将起到很好的推动作用.

《电子计算机应用人才培养大纲》和“电子计算机应用系列教材”贯穿了从应用出发、为应用服务,大力培养高质量、多层次、复合型应用人才这样一条主线.这部培训大纲总结了近几年各地计算机技术培训正反两方面的经验,提出了计算机应用人才的层次结构、不同层次人才的素质要求和培养途径,制定了一套必须遵循的层次化培训办学规范,编制了适应办学规范的“课程教学大纲”.这部培训大纲为各地方、各部门、各单位制定人才培养规划和工作计划提供了原则依据,为科技人员、管理人员以及其他人员学习计算机技术指出了努力方向和步骤,为社会提供了考核计算机应用人才的客观尺度.“电子计算机应用系列教材”是培训大纲在教学内容上的展开与体现,是我国目前规模最大的一套计算机应用教材.教材的体系为树型结构,模块化与系统性、连贯性、完整性相兼容,教学内容注重实用性、工程性、科学性,并具有简明清晰、通俗易懂、方便教学、易于自学等特点,是一套很好的系列教材.

这部培训大纲和系列教材的诞生是各方面团结合作、群策群力的结果,它的公开出版和发行,对计算机应用人才的培训工作将起到积极的推动作用.希望全国各地区、各部门、各单位广泛运用这套系列教材,发挥它应有的作用,并在实践中检验、修改、补充和完善它.

和科技人员继续工程教育等制度相结合,逐步把计算机应用人才的培训工作引向规范化、制度化轨道,为培养和造就大批高素质、多层次、复合型计算机应用人才而努力奋斗,更好地推动计算机应用事业向深度和广度发展.

李祥林

1988年10月17日

前 言

分布式数据库系统是最近几年才发展起来的一门新技术，它主要研究在计算机网络上如何进行数据库处理。这是当今计算机世界最引人入胜的领域之一，也是今后若干年内计算机科学研究和应用的主要课题。

与集中式和分散式数据库系统相比，分布式数据库系统具有很多突出的优点。比如，提高了系统的可靠性，加快了数据的存取速度，降低了通讯费用，扩大了数据库的容量等。正是由于这些优点，使得分布式数据库系统可以广泛地应用于社会的各个领域，如军事指挥自动化系统、铁路系统、银行系统、商业系统、大型厂矿系统、办公自动化系统以及其他大规模的数据资源系统等。

在国外，世界著名的四大分布式数据库系统已经投入运行，并在其上发展了各种应用软件。另外还有一批铁路、银行、商场、大型厂矿企业专用的分布式数据库系统业已交付使用，而且产生了良好的社会效益和经济效益。因此，从事这方面工作的人越来越多。一些发达国家甚至为发展分布式数据库系统制订了先行计划，建立了全国性的协调组织。

在国内，尽管这方面的工作起步较晚，但是不少单位都在积极创造条件，根据自己的环境和条件来研制各种类型的分布式数据库系统。一些大学也开始为研究生和本科生开设这门课程或讲座。建立分布式数据库系统已列为我国“八五”期间的重点攻关项目。

本书比较全面地介绍了这一领域的基本概念及发展概况。同时以法国 IMAG 研究中心的一个分布式数据库系统——MICROBE 系统作为实例，详细地阐述了分布式数据库系统的结构、数据描述语言与数据操纵语言以及各部分的实现思想和实现方法。此外，还介绍了其它系统所采用的方法。本书的特点是：方法+算法，理论+实际。读者阅读完这本书以后，不仅对分布式数据库系统有一个整体的概念，还可以在理论的指导下，自己动手设计一个分布式数据库系统。

本书共十一章。第一章概述了分布式数据库系统的发展历史，介绍了一些基本概念，给出了分布式数据库系统的一般结构、实现方法、优点及评价标准，最后简要地介绍了 SIRIUS 计划。第二章介绍了 MICROBE 系统的实现环境、系统结构，以及大致工作过程。第三章介绍 MICROBE 系统所采用的高级关系数据库语言——MIQUEL 语言，并通过两个具体的应用实例说明如何使用这种语言。从第四章开始，详细地阐述了 MICROBE 系统各部分的实现思想和实现方法，其中第四、五、七、八章是关于集中式关系数据库系统的内容，第六、九、十、十一章是关于分布式数据库系统的内容。读者可以按目录的顺序阅读，因为这正好是一个用户请求经过系统各结构层的顺序。读者也可以跳过第六章，先阅读第七、八章，然后再集中阅读第六、九、十、十一章，因为这样做可以先了解集中式系统，然后再了解分布式系统。MICROBE 系统的实现阶段也是这样划分的。第六章介绍了目前所采用的几种分解定位与并行处理的方法，重点在于介绍 MICROBE 系统采用

的动态分布式的分解定位与并行处理的方法。这是世界上第一个通过分布式执行系统（见第九章）实现这种算法的系统。由于采用了这种算法，因而提高了系统的效率，增加了系统的实用性。这也是 MICROBE 系统对分布式领域所作的最大贡献。附录 A 介绍了目前国际上几个较有影响的分布式数据库系统的简单情况。附录 B 介绍了作者在 3+ 网络上采用最流行的数据库管理系统 DBASE III 研制的一个分布式数据库系统，目的是为了让读者能够举一反三，学以致用。

本书在编写过程中，多次得到 MICROBE 系统小组负责人 G. Sergeant 博士的帮助。本书部分章节，是法国国家信息与自动化研究所 (INRIA) 科学顾问、巴黎第六大学信息系 G. Gardarin 教授为研究生开设分布式数据库系统课程的讲稿。我国著名的数据库专家萨师煊教授在百忙中审阅了全稿，并且提出了许多宝贵意见。杜小勇同志协助萨教授审稿。在此，作者向他们表示衷心的感谢。

作者

目 录

第一章 分布式数据库系统概述	(1)
1.1 引言	(1)
1.2 基本概念	(2)
1.3 一般结构	(9)
1.4 实现方法	(10)
1.5 优点及评价标准	(11)
1.6 SIRIUS 计划简介	(12)
第二章 MICROBE 系统简介	(14)
2.1 MICROBE 系统的工作环境	(14)
2.2 MICROBE 系统的结构与工作过程	(14)
第三章 MIQUEL 语言及实例	(17)
3.1 MIQUEL 语言的语句	(17)
3.2 MIQUEL 语言的语法	(20)
3.3 应用实例	(24)
第四章 翻译程序	(33)
4.1 关系代数	(33)
4.2 翻译程序的结构与工作过程	(34)
4.3 翻译程序实现的基本思想与方法	(36)
4.4 翻译程序的一些约定	(48)
4.5 规则 I 与规则 II	(48)
第五章 优化	(55)
5.1 关系代数的表达式定义及符号表示法	(55)
5.2 优化实例	(61)
5.3 SQUIRAL 方法	(63)
5.4 PRTV 系统的优化	(66)
5.5 MICROBE 系统的优化	(70)
5.6 MICROBE 系统优化程序的结构	(75)
第六章 分解定位与并行处理	(77)
6.1 引言	(77)
6.2 分解定位与并行处理的一般方法	(77)
6.3 从分布式角度看 MICROBE 系统的总体结构	(80)
6.4 MICROBE 系统的分解定位与并行处理	(80)
6.5 动态程序设计技术在动态分解定位与并行处理算法中的应用	(87)
第七章 关系操作	(91)
7.1 解释执行的方法	(92)
7.2 关系操作的分类	(94)

7.3	关系操作的实现	(95)
第八章	关系存贮	(106)
8.1	几个基本概念及其结构格式	(107)
8.2	关系存贮的实现	(110)
8.3	活动指针关系与自变量关系	(122)
8.4	使用关系存贮的基本操作的例子	(123)
第九章	分布式执行系统	(125)
9.1	分布式计算机系统	(125)
9.2	网络的结构及其功能层	(130)
9.3	分布式执行系统	(133)
第十章	分布式系统 CPR	(141)
10.1	CPR 概述	(141)
10.2	CPR 的设计思想	(144)
10.3	CPR 的实现方法	(145)
10.4	CPR 的用户界面	(154)
第十一章	系统恢复	(164)
11.1	SIRIUS-DELTA 的系统结构	(164)
11.2	事务处理的有效性	(165)
11.3	联系中断的处理	(166)
11.4	系统恢复中用到的几个文件	(167)
11.5	局部恢复	(168)
11.6	总体一致点	(169)
11.7	总体恢复	(171)
附录 A	国际上有影响的分布式数据库系统	(173)
附录 B	分布式数据库系统 DDBASE	(175)
B.1	DDBASE 系统概述	(175)
B.2	DDBASE 系统的结构和工作过程	(175)
B.3	DDBASE 的系统命令	(176)
B.4	系统实现梗概	(188)
	参考文献	(192)

第一章 分布式数据库系统概述

1.1 引言

分布式数据库系统是当今计算机世界最活跃的领域之一，它的发展越来越受到人们的重视。正如第二届国际分布式数据库会议总主席 H. J. Schneider 所指出的那样：“发展分布式系统，特别是以分布式数据库作为该系统的核心，将成为今后 10 年计算机科学的主要方向之一。”事实已经证明，在计算机科学领域里，70 年代是数据库系统的年代，80 年代则是分布式系统和分布式数据库系统的年代。

分布式数据库系统是计算技术与通讯技术迅速发展的产物。一方面，由于计算机功能增强，成本下降，使得不少的办公室、实验室，甚至个人用户都可以具有小型或微型计算机，从而有利于数据的分散处理。另一方面，由于通讯技术的迅速发展，出现了各种计算机网络，降低了数据传输费用。而计算机局部网络的广泛应用，则为分布式数据库系统的出现提供了实现的可能性。

分布式数据库系统的研制成功，经历了一个艰难而曲折的过程。最初由 IBM, UNIVAC, CII/HB 等设计单位提出、在 ARPA, CYCLADES 计划中发展起来的技术，只实现了分布式系统的信息通讯。关于分布式环境中的信息存贮、分配、交换以及处理等问题均没有得到妥善的解决。这说明在集中式数据库系统中所使用的传统技术，对于解决上述问题已经不再适用。因此，简单的移植是不够的，必须根据分布式系统的特殊要求，对实现技术作重新的考虑。

自 1975 年以来，美国、法国、联邦德国等国家的许多大学、科研单位都在研究分布式数据库领域里的各种问题，力图实现分布式数据库系统。随着对各种理论问题研究的进展，以及人们对分布式数据库系统认识的深化，一些国家的政府也先后拟订了发展这一领域的先行计划，如美国的 SDD-1 计划、法国的 SIRIUS 计划、联邦德国的 POREL 计划等。我们在本章第六节将简要地介绍 SIRIUS 计划，它负责协调全法国在这一领域里的各项研究工作。

自 1977 年法国 IRISA 研究中心公布第一个分布式数据库系统以来的 10 年间，已经有几十个系统投入了运行。其中最著名的有柏林工业大学的 VDN 系统、斯图加特大学的 POREL 系统、IBM 公司 San Jose 实验室的 k* 系统，以及法国 INRIA 的 SIRIUS-DELTA 系统，另外还有 SDD-1 系统和 POLYPHEME 系统，它们代表了分布式数据库领域里的某些开创性的工作。

为了便于读者了解以后各章的内容，本章第二至五节将顺序介绍分布式数据库系统的基本概念、一般结构、实现方法，以及分布式数据库系统的优点及评价标准。

1.2 基本概念

分布式数据库系统正以很快的速度发展着，然而它的某些方面尚处于理论探讨和实验运行的阶段，因此很难对该领域里的基本概念做严格、简明的概括。事实上，由于人们看问题的角度不同，采用的方法不同，从而产生了各种各样的关于分布式数据库的定义。下面，我们仅对该领域里的基本概念给出一个轮廓性的描述。

1.2.1 分布式数据库管理系统

分布式数据库管理系统是一组软件，它是分布式数据库系统的核心，它负责实现诸如分布式数据库的建立、查询、更新、复制、维护等功能。

1.2.2 匀质系统与异质系统

如果一个分布式数据库系统的所有节点只允许相同类型的数据模型，则称该系统为匀质系统。如果所有的节点又是同类型的机器，则称该系统为完全匀质系统。在一般情况下，如果没有特别声明，都是指匀质系统。相反，如果一个分布式数据库系统的每一个节点都允许不同类型的数据模型，那么则称该系统为异质系统。一般来说，异质系统所采用的机器类型也是不相同的。

1.2.3 专用系统与通用系统

专用系统是指为某一个特定的应用而专门研制的分布式数据库系统。专用系统的管理对象、信息流程都比较明确，系统效率高，实现起来比较容易。根据目前分布式系统发展的情况来看，研制各种类型的专用系统仍然是今后的主要发展方向。

通用系统则相反，它是为某一类应用而研制的分布式数据库系统。它对自己的管理对象、信息流程均不十分明确，只能按照一般的思想和方法来构造，因此系统效率低，实现起来比较困难。目前实现的通用系统大都处于理论探讨和实验运行阶段，还没有成熟的商品进入市场。

1.2.4 数据的透明性

数据的透明性是指用户和用户程序不需要知道分布式数据库里的数据位于何处（其中包括位于不同的节点上），因为这种位置信息是由系统统一保留，统一进行解释的。数据的透明性是分布式数据库系统的主要目标之一。

1.2.5 副本的透明性

为了改善分布式数据库系统的性能，增加可用性，对于一个给定的逻辑数据库，除了在网络的一个节点上保留其物理存贮（称之为正本）以外，还应该在其它节点上有多处重复存贮（称之为副本）。这样一来，用户可以操作本地的拷贝以代替同远程的通讯。另外，网络中只要有一个拷贝可用，就可以对该数据库进行存取，同时也可以对系统进行恢复。这种特性正是分布式数据库系统的潜在优点。

在具有副本的系统中，数据的存取细节显得十分复杂。用户希望数据的定位和维护副本的工作也由系统完成，而不必由用户来处理，这就是所谓副本的透明性。

由于有了数据的透明性和副本的透明性，在用户看来，分布式数据库系统就像传统的集中式数据库系统一样，当用户考虑一个逻辑数据库时，不必关心其物理存贮在哪里以及存贮了几次。

1.2.6 数据分段的类型

数据分段的类型反映了分布式数据库中的数据在网络中的分布种类。下面，我们通过一个具体的例子介绍数据分段的类型。为了叙述的方便，而且考虑到不失一般性，我们以关系模型为例。

假定有一个分布式数据库，它的概念模式为：

职工关系

P (职工编号、姓名、工资、部门名称、子女数、企业单位)

企业关系

E (企业单位、地址、邮政编码、电报挂号)

部门关系

D (部门名称、职工人数、负责人姓名)

如果在局部数据库中，只具有分布式数据库中的部分数据，那么可以想象，这些关系可以具有很多种分布情况。概括起来，可以分为两大类，即均匀分布和非均匀分布。

1. 均匀分布

我们可以这样设想，如果每一个部门都对应着一个局部节点，在这些局部节点上都有一个局部数据库，在该局部数据库中具有这个部门的分布式数据库的全部数据。那么，这时局部数据库的概念模式与分布式数据库的概念模式是一致的。这种情况，称之为均匀分布。

2. 非均匀分布

对于一些比较简单的系统来说，数据采取均匀分布的方法就可以了。但是对于那些比较复杂的系统，为了增加系统的实用性，必须支持数据的非均匀分布。所谓非均匀分布，就是局部数据库的概念模式与分布式数据库的概念模式不一致。这种思想比较合乎实际。例如，对于一个分布式数据库，绝大部分用户仅仅是使用其中的一部分数据，或者是由于存在着保密数据的问题，而禁止一般用户对总体数据进行访问等等。这些情况都需要使用非均匀分布。

有几种类型的非均匀分布：

● 在关系一级进行分布，即全体数据是分布的，但是每一个关系却完整地存放在一个节点上。例如，

职工关系 P 在节点 1 上；

企业关系 E 在节点 2 上；

部门关系 D 在节点 3 上。

●对总体关系进行水平分段的分布,即总体关系是通过在局部关系之间进行 UNION 操作来实现的.例如,把职工关系 P 水平划分成 X, Y, Z 三个局部关系,并且

X 在节点 1 上;

Y 在节点 2 上;

Z 在节点 3 上.

则 $P = (X \text{ UNION } Y) \text{ UNION } Z$.

●对总体关系进行垂直分段的分布,即总体关系是通过在局部关系之间进行 JOIN 操作来实现的.例如,把职工关系划分为:

P_1 (职工编号、姓名、工资、企业单位);

P_2 (职工编号、子女数、部门名称).

其中,

P_1 在节点 1 上;

P_2 在节点 2 上.

则 $P = \text{JOIN} (P_1, P_2)$.

●前面列举的几种分布的混合型分布,例如,把职工关系 P 划分为三个关系 P_1, P_2, P_3 , 其中,

当企业单位= X 时,为关系 P_1 ;

当企业单位= Y 时,在职工编号上有关系 P_2, P_3 .

即

P_1 (职工编号、姓名、工资、部门名称、子女数、企业单位);

P_2 (职工编号、工资、部门名称、企业单位);

P_3 (职工编号、姓名、子女数).

并且

P_1 在节点 1 上;

P_2 在节点 2 上;

P_3 在节点 3 上.

在这个例子中,职工关系 P 是下式的结果:

$$P = P_1 \text{ UNION } (\text{JOIN} (P_2, P_3))$$

通过这些例子,我们可以看到,一个分布式数据库可以是虚拟的.因为职工关系 P 并不存在,需要使用 UNION 和 JOIN 操作来实现.

1.2.7 分布函数与关系定位

分布函数可以描述分布式数据库的各总体关系存在的情况,并且给出总体关系的全体或者部分的实现规则.这样,就出现了关系定位的概念.下面介绍几种不同的定位方法:

●通过一个常数定位.例如把职工关系 P 定位在节点 1 上.

●通过一个在属性上的函数(布尔表达式)来定位.例如,当企业单位= X 时,职工关系 P 定位在节点 1 上,并称之为关系 P_1 ;当企业单位= Y 时,职工关系 P 定位在节点 2 上,并称之为关系 P_2 .

●通过一个邻近函数定位。例如，在同一个节点上，职工关系 P 的每一种情况都是部门关系 D 的对应情况。

1.2.8 分布式数据目录

分布式数据目录指出了分布式数据库文件在网络各节点上的分布情况。概括起来，分布式数据目录有以下几种形式。

1. 全局目录

这种形式要求至少在网络的一个节点的目录里存放所有的分布式数据库文件名，并且指出它们的物理地址及副本地址（如果有的话）。当用户访问某个文件时，直接扫描全局目录，然后确定应该访问哪台计算机。这种形式的主要优点是，查询文件非常方便。缺点是，当多个用户读写全局目录时，需要加锁处理。并且，当网络节点增多，特别是有多个全局目录时，效率就会明显地下降。另外一个更为严重的问题是，当存放全局目录的节点出现故障，或者全局目录受到破坏时，就无法访问数据库文件，整个分布式数据库系统就会处于瘫痪状态。

2. 局部目录

这种形式要求所有的节点都有一个局部目录。在这个局部目录里仅仅存放本节点的数据库文件名及其它局部目录的位置。

采用这种形式，访问本节点的数据库文件非常方便、迅速，但访问其它节点上的数据库文件时，就要增加用户的负担。

3. 全局局部目录

这种形式是全局目录与局部目录形式的组合，其含义是，每个节点只保留其自身的局部目录，另外还有一个中心节点，存放所有的局部目录。

关于分布式数据库的目录，除了上面介绍的三种主要形式以外，还有其它的组合形式，这里就不再一一介绍了。总之，每种形式各有其优缺点。至于一个具体的系统应该采用哪种目录形式，这要看系统的运行环境、应用对象，灵活地加以选用。

1.2.9 请求的分解定位

请求的分解定位是分布式系统和分布式数据库系统中的一个新概念。它指出在一个分布式环境中，请求如何以最优的方式被执行。当然，请求要根据所考虑的分布式类型，或多或少地加以变换。下面，我们首先通过一个具体的例子，来看看同一个请求，先介绍在 1.2.6 节中所说的各种分布式数据的情况下是如何进行分解定位的，然后再给出请求的分解定位的一般概念，最后叙述请求的分解定位的一般过程。

1. 例子

我们仍以上面给出的分布式数据库为例。假定，用户提出这样一个请求：
给出有三个以上子女（包括三个），且月工资低于 8000 法郎的职工的全部信息。

(1) 均匀分布的情况

这种情况比较简单. 首先, 系统把请求发送到所有的节点, 并且在那里加以执行. 接着, 把局部结果送回主节点 (我们把用户提交请求的节点称为主节点). 然后, 再在主节点进行 UNION 操作. 最后, 把结果提交给用户.

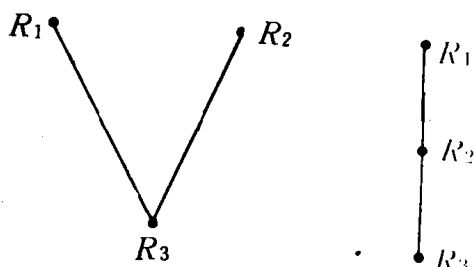
(2) 非均匀分布的情况

● 在关系一级进行分布. 由于请求只涉及到职工关系 P 的数据, 所以只需要把请求发送到节点 1, 并且在那里执行, 然后把结果送回主节点, 并提交给用户.

● 对总体关系进行水平划分的分布. 通过对请求的选取标准和分布的规则加以比较, 可以发现, 在这种情况下是不允许限制节点的. 这时, 必须把请求同时发送到节点 1, 2, 3, 并且在那里加以执行, 然后把结果送回主节点. 最后由主节点对这些结果做 UNION 操作, 然后提交给用户.

● 对总体关系进行垂直划分的分布. 由于用户要求系统给出符合条件的职工关系 P 的全部信息, 所以就不能对组成职工关系 P 的任何部分的操作加以限制.

在这种情况下, 至少有如下图所示的两种可能的策略表示请求之间的同步.



第一种策略 (左图):

R_1 = 在 P_1 上按照给定的工资 (<8000 法郎) 进行选取得到的关系;

R_2 = 在 P_2 上按照子女数给定的条件 (≥ 3) 进行选取得到的关系;

$R_3 = R_1, R_2$ 在职工编号上进行连接得到的关系.

第二种策略 (上右图):

R_1 = 在 P_1 上按照给定的工资 (<8000 法郎) 进行选取得到的关系;

$R_2 = R_1$ 与 P_2 在职工编号上进行连接得到的关系;

R_3 = 在 R_2 上按照给定的子女数 (≥ 3) 进行选取得到的关系.

通过对上面两种策略的比较, 我们可以发现: 第一种策略很好地使用了并行处理的方法. 第二种策略采用了串行处理的方法, 请求的回答时间也就相应地增多.

● 混合分布. 正像垂直分布一样, 混合分布也不可能限制节点. 下面我们考查两种可能的策略.

第一种策略:

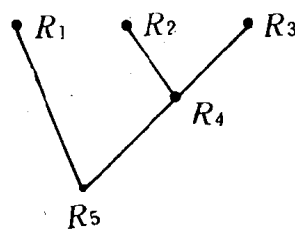
R_1 = 在 P_1 上按照给定的工资 (<8000 法郎) 和子女数的条件 (≥ 3) 进行选取得到的关系;

R_2 = 在 P_2 上按给定的工资 (<8000 法郎) 进行选取得到的关系;

R_3 = 在 P_3 上按照给定的子女数 (≥ 3) 进行选取得到的关系;

$R_4 = R_2$ 与 R_3 在职工编号上进行连接得到的关系;

$R_5 = R_1 \text{ UNION } R_4$



最后把结果提交给用户。

第二种策略：

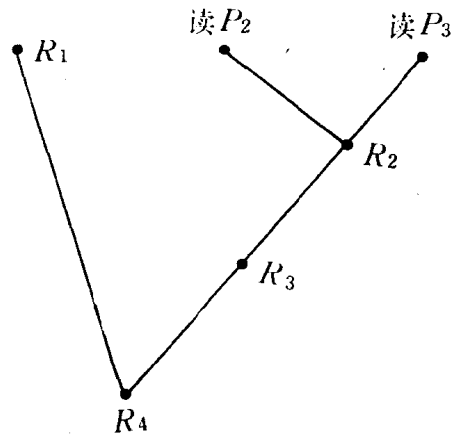
R_1 = 在 P_1 上按照给定的工资 (< 8000 法郎) 和子女数的条件 (≥ 3) 进行选取得到的关系；

R_2 = P_2 与 P_3 在职工编号上进行连接得到的关系；

R_3 = 在 R_2 上按给定工资 (< 8000 法郎) 和子女数 (≥ 3) 进行选取得到的关系；

$R_4 = R_1 \text{ UNION } R_3$ 。

最后把结果提交给用户。



2. 请求的分解定位概念

通过上面的例子，我们可以看到，同一个请求，由于使用不同的分布式数据类型，可以导致请求的各种不同形式的变换。这种变换过程，叫作请求的分解定位。

3. 请求的分解定位的一般过程

请求的分解定位可以分成两个阶段来进行。

第一阶段：把请求变换成涉及到局部数据的子请求的集合。为此，分布式数据库系统首先识别一个分布式数据库的全部总体关系，接着，用支持在分布式数据上的子请求的集合，来代替支持在总体关系上的总体请求。当然，最后还要负责构成总体关系的数据。

这一系列操作，可以在总体一级被压缩。采用的方法是，通过在请求的选取标准和数据的分布规则之间进行比较。

在这个阶段结束时，得到一份支持在局部数据上的子请求清单表。

第二阶段：负责优化，其目的是查看一下是否有比第一阶段更好的子请求清单表。采用的方法是树形变换技术。下面，我们通过两个具体的例子来说明。

〔例 1〕

按条件 C_1, C_2 对关系 R 进行选取操作，其中， R 垂直分布在 R_1, R_2 上，条件 C_1 支持在 R_1 的数据上，条件 C_2 支持在 R_2 的数据上。这里我们可以有下图所示的两种形式：

