

Ada相关的分布式 数据库管理系统Ada-DDBMS

徐泽同 (中国科学院数学研究所)

摘要
此文介绍了中国科学院科学数据库POREL小组在维也纳技术大学应用信息研究所消耗10人年,在VAX, UNIX上重新模拟实现的分布式数据库管理系统POREL;也介绍了中国科学院数学研究所Ada-DDBMS小组正在进行的VAX, VMS, DECNET上的Ada相关的分布式数据库管理系统Ada-DDBMS.

分布式数据处理及其核心分布式数据库管理系统因它能克服集中式数据处理及其核心集中式数据库管理系统的很多弱点,也因集中式数据库技术与计算机数据通讯技术的成熟,因而在过去十余年来国内外计算机科学界给予了此研究领域广泛的重视,投入了不少研究人员,花了不少钱,出了不少论文,也造了不少原型系统,而今据这方面的权威人士认为, DDBMS的研究是应走向实用化的时候了。我国地域辽阔,在离散的计算机应用初具规模之下,进而不失时机地走向大范围的联网的计算机应用,无论于哪个方面的现代化的发展,意义均是十分深远的。无论是POREL小组在外的工作,还

是Ada-DDBMS小组在内的工作,都是为了这一目的,我深望我们终会达此目的。

1. POREL简介

POREL是1977年以来西德计算机科学家E. J. Neuhold教授领导下研制的计算机非均质网上的关系型数据库管理系统。在1984年前的斯图加特时期,为研制此系统花了35个科学家人年,40个学生人年,花了450万西德马克,出了不少论文,其中博士论文七篇,硕士论文36篇;在PDP11, RSX11M上作程序近七万行,遗憾此系统过于复杂与庞

要关于处理站点所具有的推理的知识。以下提出的命题演算可以简化这种知识的表达和处理。从两种类型定义了四个新的知识操作符。

除了通常的结构化和量化的信息外, PIDDB中的数据对象如果没有其位置信息将不能被完整地描述。此处用布尔变量操作符, k 和 ϕ 来表达和测试这种信息:

k : $k \cdot \text{OBJ}$ 为真, 如果对象OBJ在站点 S_n 已知的(意味着OBJ可以由站点 S_n 拥有,也可以是站点 S_n 知道拥有OBJ的其它站点的标识)。

ϕ : $\phi \cdot \text{OBJ}$ 为真, 如果对象OBJ由站点 S_n 拥有。这两个操作符由推理规则 $\phi \cdot \text{OBJ} \rightarrow k \cdot \text{OBJ}$ 联系起来。

用这些操作符表示的断言的真实性与陈述该断言的位置无关。例如: $\phi \cdot \text{OBJ}$ 可能为真, 但 S_n 站点并不知道。OBJ既可为一数据, 也可是元数据对象, 还可是一个站点。如果OBJ是一个站点, 那么 ϕ 操作符将不适用, 这是由于在PIDDB结构中并不支持“站点间互相占有”这一概念。

第二类操作符是大写的 K 和 Φ 。它们定义了已知或拥有OBJ的站点集合。这一表示法简化了知识模式及约束的表达:

$$K(\text{OBJ}) = \{S_n, k \cdot \text{OBJ}\}$$

$$\Phi(\text{OBJ}) = \{S_n, \phi \cdot \text{OBJ}\}$$

[冯建华 贝来情译自“Proc. of the 13th VLDB Conf., Brighton, 1987, 增 辑 校”]

大，一直未能真正运行。

1985年元月至1987年元月，中国科学院科学数据库与维也纳技术大学进行合作，中方派出POREL小组到该校应用信息研究所 Neuhold处，在VAX，UNIX上重实现POREL。经中方POREL小组10个科学家几年的努力，完成了POREL用户服务USM，离网分析NUA，涉网分析NOA，关系基本机器RBM的修改；完成了进程通讯CS，事务管理TM的重写，实现了整个POREL在模拟环境下的联调。在此基础上通过了一批分布式数据库定义，查询，修改及分布计算的试例，使POREL计划在10年之后得到了一个可运行的原型；使POREL系统达到了八万余行Pascal+C程序的规模；为它的实用化开发创造了条件。

图1展示了POREL的逻辑结构。其中USM是数据库系统和用户的接口。它可从用

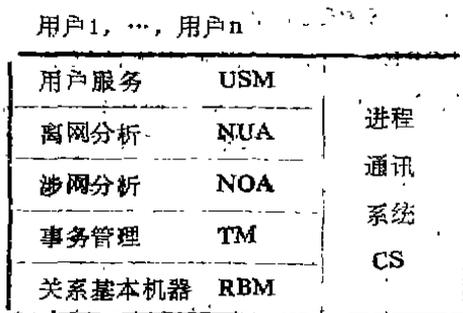


图1 POREL的逻辑结构

户接受关系数据库语言RDBL写的事务，其后启动NUA和NOA工作，并把加工结果送向TM，它还可从用户接受P-RDBL事务，即嵌有RDBL事务的Pascal程序，这将启动Pascal预编译工作，进而引起NUA，NOA及Pascal编译工作；它还管理用户数据的输入输出，并把中间出错信息，或者把计算结果向用户输出；它还以多种方式管理用户作业。它允许单结点上多用户使用数据库系统。

NUA从USM或宿主语言预处理处接受RDBL事务，经不涉及网络知识的分析，产生关系机器语言RML的事务；

NOA接受NUA交来的RML事务，并涉

及网络知识的分析，产生出从优送到网上适当结点去执行的RML子事务，此种优属局部优。NOA产生的结果在USM控制下入库以备多种方式执行；

TM启动并监控NUA，NOA分析过的RML事务的网上执行，它同时要保证上述事务执行的并发性和完整性；

RBM是POREL的执行部份，它在TM的监控下执行RML子事务，并兼管目录管理工作；

CS负责POREL各分布进程间的信息和数据的传送，它包括局部进程通讯和远程进程通讯。

2. POREL计划不足之处

POREL因它是美欧早期几大知名DDBMS系统之一，国内对于它的介绍有不少文章。作为POREL重实现的主要作者，尊重科学工作实事求是的精神，我想指出它的不足之处。因我们知其不足而踏实工作，才会为它创造一美好未来。

1. POREL的研制由来已久，因初期对DDBMS的研究尚不成熟，因而它的初期设计太粗，好多关键之处均是前进一步改进一步地作了过来；再有POREL的不少目标设计均太大，太大作不到则空。比如POREL的用户语言RDBL只实现了设计功能的40%，且不合SQL标准；宿主语言先定为FORTRAN，后改为Pascal，但均未实现；数据只作了按关系分布而不是按子关系的水平分布；事务的执行在一结点内并未作到并行，而只是外部并行，内在串行；网际恢复未作好；死锁控制尚待严格；处理关系代数的功能多有原型之态；远程进程通讯只是模拟，尚不可能适应于非均质网等；

2. 先后介入POREL研究已35年，除Neuh-old外，多数人川流不息，他们很多精力忙于学习，书写论文，成名后即离去；于系统实现的实际工作重视不够，甚致不重视。这样POREL的论文往往脱离POREL的实际，文不符实，实是POREL研究的一大遗憾；

3. 无论是在斯图加特用PDP11，RSX11M；还是在维也纳用VAX，UNIX，支持POREL研究的物质条件均不足，因而不得不多次模拟网络环境。

从上可以看出, 经10年余85个人年的努力, POREL已是一个可运行的原型系统, 但又可看到, 它有很多不足之处。对于我们经艰苦努力而获得的上述怪物, 我们应将其束之高阁呢? 还是进一步努力将它驯服, 让它更好为我们服务呢? 我想正确的当然后者! 这就是我及我的同事们, 在重实现 POREL 的基础上走向Ada相关的分布式数据库管理系统Ada-DDBMS的依据。

3. 程序设计语言Ada

Ada是1973年到1983年、美国国防部领导之下耗资五亿美元, 调动世界上数以千计的计算机科学家设计出的80年代的公共程序设计语言。1983年形成ANSI Ada, 之后被接受为ISO Ada。1983年以后美国国防部又领导了Ada工程, 致力于Ada语言及环境的实现, 推动Ada的应用。现外界已有上百个Ada编译可供使用, Ada早已从研究课题走向了课堂, 走入了应用。Ada已走入了商业领域, 已形成数十亿美元的产值。1983年10月我国计算机科学界经认真讨论, 主张接受ANSI Ada, 而不是什么子集。其后经国防科工委的重视, 又形成了共性软件研究计划, 主要致力于Ada语言及环境的实现及应用。现在我国汉译Ada军标即将出版, 引进的Ada编译已完成多种适合国情的改造而走向应用, 国产的Ada编译也快成功, Ada在国内走向应用已是事实。

4. Ada相关的DDBMS

据我们对Ada-DDBMS的实用需要, 依我等昔日研制POREL, 翻译ANSI Ada并制作Ada军标的实践, 在VAX、VMS、DECNET局域网及远程网上研制Ada相关的分布式数据库管理系统是顺理成章之事。此项目的含义有四: 1. 在VAX, VMS, DECNET均质网上作出可供实用的DDBMS; 2. 将Ada作成宿主语言; 3. 增加DDBMS的逻辑接口, 作成演绎分布式数据库管理系统; 4. 将Ada作为上述系统的书写语言。上述四项其核心为第1项, 具体作法是: a. 将在维也

纳作成的POREL原型系统从UNIX移植到VMS。之所以作此, 是因为我们有广用的VAX, VMS, DECNET, 而现在只有从VMS才能方便地使用DECNET网; b. 在VMS, DECNET局域网及远程网的环境下将POREL模拟的通讯系统作成实用系统; c. 依SQL标准修正RDBL语言, 使RDBL与SQL相容。这要修改POREL的NUA; d. 改造NOA, 实现真正的子关系级的水平分布; e. 修改POREL的TM, 使其在局部结点内对于事务操作真正并发执行, 改进网际恢复功能, 改进死锁检测功能; f. 将POREL的RBM中的目录管理分离出来, 作成一个独立进程, 将RBM作成对多个子事务并发生成及执行, 完善关系代数计算功能与宿主语言共存的计算功能; 第2项主要的工作为: a. 设计实现Ada预编译, 其功能是将Ada+RDBL程序分离出RDBL事务, 经NUA, NOA加工后形成A-RML交Ada编译处理; b. 设计实现宿主语言支持, 在它的支持下使A-RML得以执行。第3项的工作是借助已有的逻辑程序设计语言的实现工作, 作一个与上述DDBMS的接口, 这样演绎处理是集中式的, 数据处理是分布式的。第4项是把Ada作为上述系统的书写语言。南京大学作DDBMS的同行曾以C+系统调用+通讯原语作为分布式数据库管理系统实现语言。其实Ada(非分布的Ada)已有C+系统调用功能, 故Ada+通讯原语, 或者分布Ada编译支持的Ada已是一个很好的DDBMS实现语言。如用Ada写DDBMS, 则可省去局部进程通讯子系统, 省去对OS的不少调用(其中包括实时调用); 如用分布的Ada写DDBMS, 则可省去整个通讯系统。这是功能上的好处。另一方面, 无论是C还是Pascal语言本身, 因无环境可言, 在实现DDBMS时都困难的。我们在UNIX上重实现POREL时, 利用了UNIX的环境。而Ada本身有着良好的环境, 因而在作DDBMS时, 无论编辑、编译、连接、调试、繁杂的管理等等都会得到巨大的好处。但是现实

POREL原型已由8万余行 Pascal+C 程序写成,如现去改为Ada重写,则不合算。若上述计划成功,真需要用Ada重写时,我们会利用Pascal到Ada的翻译(这已由中国科学院软件研究所实现)将大量的Pascal程序自动译为Ada,再作适当的手工修改而达到。图2介绍了Ada-DDBMS的进程结构。

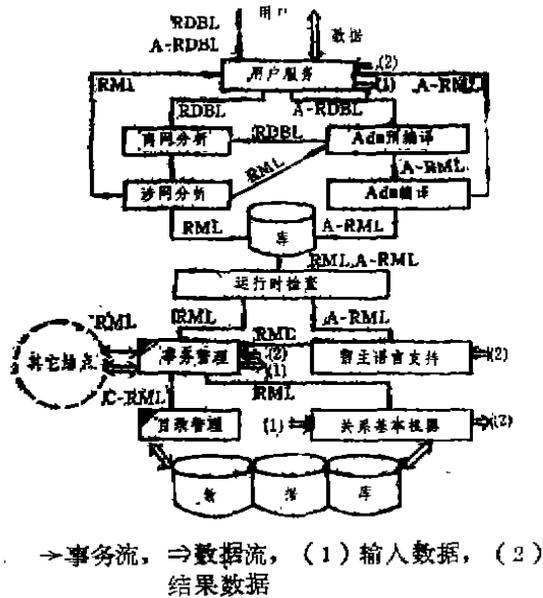


图2 Ada-DDBMS的进程结构

注:中国科学院科学数据库赴维也纳POREL小组由徐泽同、陈东岳、赵立平、王翰虎等人组成;数学研究所Ada-DDBMS小组由徐泽同、陈东岳、赵立平、韦梓楚、车克健、涂方登、金雅芬、涂健等人组成。

(上接第79页)

向对象的程序设计,以Smalltalk为背景,叙述其五个基本概念、特征和环境。每章均附有形式多样的习题和主要参考资料,供复习巩固和进一步深化提高之用。

在教材编写过程中,我们吸收了现担任教学工作的,前几年毕业的研究生和现正攻读博士学位的研究生参加。

五、其它

从教学目的与基本内容来看,本课程不仅适合于软件专业诸研究方向,而且近几年也开始引起计算机组织与系统结构等专业的重视,并陆续有研究生来选读。

随着本科生教学要求和质量的提高,本课程的

Ada-DDBMS研究计划,已得中国科学院院长特批基金及国防科工委共性软件研究计划支持。同时也得中国科学院数学研究所,北京天文台,科学数据库支持,正由Ada-DDBMS小组实施,计划工作三年。

参考文献

1. POREL Design Specification, Technical Report 78.5 University of Stuttgart.
2. Distributed Data Base Systems with Special emphasis toward POREL, 84, E.J.Neuhold.
3. POREL, a Distributed Data Base System, B. walter, E.J. Neuhold, 82, Recent Advances in DDBM. IEEE-CS Press.
4. 分布式数据库管理系统POREL的进程通讯系统CS, 徐泽同, 88.6初稿. 89.10修正投稿。
5. 分布式数据库管理系统POREL和程序设计语言Ada, 徐泽同, 88.10投《国防系统分析及软件》。
6. Ada导引/程序设计语言Ada参考手册, 袁崇义, 徐泽同译, 86.1, 科学出版社。
7. 中华人民共和国军用标准, 程序设计语言Ada参考手册, 徐泽同、袁崇义、王振宇、李辛、任承邦、李江月、孙潮义等译校, 国防工业出版社将出版。

内容或部分内容也可在高年级本科生中作为选修课开设,教学方法与研究生的相比,有所不同,如可多讲少讨论,我们认为,该课程内容对于即将参加工作的计算机系学生无疑是有益的。

主要参考文献

- [1] 国防科技大学研究生院, 程序语言概念与设计课程教学大纲, 1989.2.
- [2] 郭浩志王编, 程序设计语言概论, 国防科技大学出版社, 1989.7.
- [3] Ellis Horowitz, Fundamentals of Programming Languages, Springer-Verlag, 1987.
- [4] 郭浩志, 程序语言的特征与语言程序的教学, 全国高校软件经验交流会, 1983.12.