

# 基于 CIAgent 技术的分布式数据库研究与设计

吴梅, 曹谢东, 曹蕾, 张琴

(西南石油大学, 四川 成都 610500)

**摘要:**文中为解决计算机系统本身存在一些问题,找到新的解决方案,把不同的计算机技术相融合,介绍了分布式数据库技术及多 Agent 技术,并研究了二者结合的可行性及关键技术。然后分析了采输处数据库集成需求,按照自底向上设计了同构异质分布式数据库,设计了多 Agent 系统结构,在系统实现中设计了包括支持网络分布环境的通信模块的通用 Agent (DIAgent),使得此 DIAgent 能够在网络环境下与其它 DIAgent 交互,并在此基础上设计了全局应用。这为企业提供了新的数据库应用平台。

**关键词:**分布式数据库;多 Agent;分布式查询;分布式事务;DIAgent

**中图分类号:**TP311.13

**文献标识码:**A

**文章编号:**1673-629X(2006)11-0050-04

## Research and Design on Distribute Database System Based on CIAgent Technology

WU Mei, CAO Xie-dong, CAO Lei, ZHANG Qin

(Southwest Petroleum University, Chengdu 610500, China)

**Abstract:** For solving computer system itself to have some problems, finding the new solution, fused different computer technology, introduced the distributional data bank technology and the multi-agent technology, and studied the two union feasibility and the essential technology. Then analyzed has picked the losing place database integration demand, has upwardly designed the isomorphism neterogeny distributional database from the bottom, has designed the multi-agent system structure, has designed in the system realization including support network distribution environment connection module general agent (DIAgent), enabled this DIAgent to be able under the network environment and other DIAgent interactive. And has designed the overall situation application in this foundation. This provided the recent database for the enterprise to apply the platform.

**Key words:** distribute database; multi-agent; distribute query; distribute transaction; DIAgent

### 0 引言

计算机系统本身存在一些问题尚未解决或解决方案不太理想,要解决这些问题,可以把不同的计算机技术相融合,可以找到新的解决方案;可能产生新的技术领域,满足新的需要。比如:数据库技术与分布式处理技术相结合,出现了分布式数据库系统,满足企业由多个下级组织组成实际情况对数据库技术需要;数据库技术与并行处理技术相结合,出现了并行数据库系统;数据库技术与多媒体技术相结合,出现了多媒体数据库系统;分布式技术与人工智能技术相结合产生了分布式人工智能;数据库技术与模糊技术相结合,出现了模糊数据库系统等,这些新技术的出现满足了实际需要。文中提出了采用分布式人工智能中的多代理系统(MAS, Multi-Agent System)与分

布式数据库技术相结合办法,集成一个企业的多个现有数据库而成为一个新的同构异质分布式数据库,在不影响原来的数据库应用的条件下,提供全局模式来满足企业的新需要,为油田的信息集成提供了科学合理的方法。

### 1 多 Agent 技术

分布式人工智能(DAI)主要研究逻辑上或物理上分散的智能系统如何并行地、相互协作地实现问题求解,主体计算(agent-based computing)不仅成为人工智能的热点,而且可能成为下一代软件开发的重要突破点。

多 Agent 系统是由多个 Agent 组成的集合,Agent 之间以及 Agent 与环境之间,通过通讯、协商与协作来共同完成单个 Agent 所不能解决的问题。多 Agent 系统的研究领域包括许多方面,比如系统的组织结构,Agent 间的通讯,Agent 间的协商与协作,Agent 系统中的机器学习,Agent 系统的动态特性等。这些方面与多 Agent 系统的性能直接相关<sup>[1,2]</sup>。

#### 1.1 CIAgent 开发平台分析

Joseph P. Bigus 和 Jennifer Bigus 研制了一个构建智能

收稿日期:2006-02-28

基金项目:国家十五科技攻关项目(2004BA616A-1-01)

作者简介:吴梅(1982-),女,四川仁寿人,硕士研究生,研究方向为人工智能与分布式数据库系统;曹谢东,硕士生导师,教授,研究方向为人工智能与分布式数据库系统。

Agent 环境 CIAgent (Constructing Intelligent Agents)。CIAgent 全部采用 Java 语言编写,可以在 JDK1.1 环境下运行。CIAgent 智能主体环境具有下列特点:

- a. 很容易在已有的 Java 应用系统中加入主体。
- b. 提供生成主体的图形构建工具。JDK1.1 包含 java.beans 软件包,具有可视建造工具,并且可以嵌套定义。
- c. 主体能支持相当强的事件处理能力。
- d. 可以将领域知识加到主体知识库中。
- e. 主体具有一定的学习功能。
- f. 多主体是永久的,可以随时将它写入文件或装入运行。

本开发工具包通过提供一个事件处理模型,即将产生事件的事件源 CIAgent 与对事件作出具体动作的实体的监听器分开,使得系统结构明晰。CIAgent 通过设置定时器类来实现对异步事件和自治的时钟事件的周期性处理。

本开发工具的一个最大优点是源代码是开放的,这是选用它的最主要原因,源代码开放,可以通过深入分析源代码,确保在应用开发时正确地使用开发包,同时为用户扩展它的功能提供了方便。本开发工具存在的一个问题是:只提供了在同一计算机中的多 Agent 之间的通过事件方式进行交互,没有提供网络环境下多 Agent 之间进行交互。另外,有基于 Java 的开发包,利用集成异种平台上的数据库系统。

## 1.2 面向 Agent 技术开发方法学

软件开发方法的发展与编程语言的发展相辅相成,经历了面向过程的方法、面向实体的方法、面向对象的方法 3 个主要阶段,目前正在面向 Agent 的开发方法发展,但此方法尚不成熟,面向 Agent 的开发方法从总体上具有以下几特点:Agent 是并发执行的、自治的实体;Agent 是用信念、目标等心智状态概念编制面的认识系统;Agent 是具有逻辑规范的推理系统;Agent 采用语言行为协议和消息传送方式相互通讯<sup>[3]</sup>。

## 2 分布式数据库

定义 DDBS 是物理上分散逻辑上集中的数据库系统,系统中的数据分布存放在计算机网络的各个场地的计算机中,每一场地都有自治处理(即独立处理)能力并能完成局部应用,而每一场地也参与(至少一种)全局应用,程序通过网络通信子系统执行全局应用。

一个分布式数据库中全局模式有多个全局关系,每个全局关系按分片模式进行分片为多个逻辑片段(一对一或一对多),每个逻辑片段按冗余策略分配到物理片段,如果是冗余的则是一对多,如果不冗余则是一对一。因此构成了分布式数据库的 3 种关系:全局关系、逻辑模式、全局模式(由所有全局关系表示)、分片模式(所有逻辑片段表示)、分配模式(所有物理片段表示)。也由此形成了 3 种数据库:全局数据库 GDB、逻辑数据库 LgDB、物理数据库 PDB。二种模式定义:分片模式 FS、分配模式 AS。故可得

分布式数据库中的查询涉及 3 种查询:对应于 GDB 的查询是用户查询 QU、对应于 LgDB 的查询是逻辑查询 QL、对应于 PDB 的查询是物理查询 QP。全局查询转化为逻辑查询最后转化为物理查询,这个过程涉及到全局分布查询为逻辑查询及对逻辑查询优化,逻辑查询转换为物理查询,物理查询提交给集中式数据库管理系统中,由其优化物理查询并执行。如图 1 所示。分布式数据库系统中,分布查询处理和优化、分布事务的恢复、并发控制等技术比集中式数据库更加复杂,是分布式数据库系统重要的技术内容。分布式数据库管理系统对分布式查询的处理是把用户提交的简单查询或复合查询按照一定算法转换为物理查询(子查询)<sup>[4,5]</sup>。

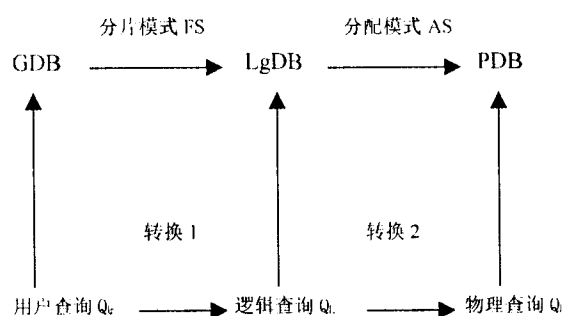


图 1 分布式数据库技术

## 3 多 Agent 技术与分布式数据库管理技术相结合

### 3.1 可行性分析

传统的异数据库集成的动态方法是通过中间件技术完成的,这种技术是以模块化结构实现的,文中以多 Agent 系统结构代替模块结构来管理分布式数据库,后者具体与前者相比有很大优点:模块式结构中各模块间的关系错综复杂、修改维护不方便、模块间耦合度相对较大,而后者关系比较清晰、Agent 间耦合度很小、易于维护。可行性分析如下:智能相似性、复杂性一致、系统结构类似性、并行性一致、同源源于分布式技术。

### 3.2 关键技术及难点

(1)数据库集成方面的技术,只能采用自底向上方式集成现有的局部数据库,由于局部数据库名称与全局名称、局部数据库之间名称可能冲突,可用数据字典解决。

(2)目前的多 Agent 开发工具,大多不支持网络分布式环境,所以必须把单机环境的多 Agent 开发工具移植到网络环境下。要解决这个问题,最好的方式是采用提供源代码的开发工具,通过分析其工作原理,在其中加入网络通信模块。

(3)分布式查询处理方法,由于集成现有的数据库,采用传统的分布式数据库查询处理算法不太合适。因为分布式数据库的分片分配模式不同,所采用的查询处理算法不同,必须根据所集成的分布式数据库具体特点,选择简单适用的算法。具体原则是不同种类的查询采用不同的处理方法。

#### 4 生产运行管理系统设计与研究

采输处各采气队自行开发生产管理数据库应用系统,满足了当时需要。但随着管理信息化进步,要求能够通过计算机网络系统及时地了解整个采输处各采气队的生产情况,并需要在此基础上开发新的全局

应用系统,如全处天然气生产报表系统、查询所有队的生产数据、更新基础数据等。如果重新开发一套全新的系统来替换原有的生产数据库系统,则存在如引言中所述的许多问题。为此,要求在保留原各采气队的生产数据库系统前提下,开发具有全局模式的分布式多数据库系统,并在全局模式基础上开发生产运行管理系统来满足用户需要。

下面运用分布式数据库技术与多 Agent 技术相结合来解决动态集成采输处现有数据库系统,并进行研究。

##### 4.1 实现技术分析

由于各采气队现有的数据传输系统必须继续使用,实现各数据库之间数据通信。各采气队均采用关系数据库系统,只是数据库产品不同而已;同时,由于采输处对现场需要数据作了统一规定,使得各采气队自行开发的数据库表结构很相近,这为采用同构异质分布式数据库系统来完成全公司数据的集成提供了极大方便,设计分布式数据库时适宜采用自底向上方法,在此基础上开发分公司的全局数据库应用。

对各采气队数据库的静态数据进行统一修改,例如修改或增加某些采集量的取值范围,要求各个采气队的这些数据必须一致,由此需要事务方式更新。由于各采气队的数据库系统不一致,无法利用操作系统(Windows 的 MTC)的事务处理机制完成数据库的事务处理,只有开发事务处理的模块。

在采输处,可能需要查询一些动态和静态数据,与集中式数据库的查询不完全一样,涉及到采用分布式查询。

##### 4.2 分布式数据库设计

###### 4.2.1 全局模式设计

由于是集成的数据库为一个分布式数据库,所以采用自底向上的设计方法。通过分析各采气队原有的数据库可知,它们都是关系数据库,且是同构的,所以,全局数据库模型采用关系型数据库。构造的全局模式图如图 2 所示,图中只画出了部分属性。图形符号箭头表示外键依赖关系。

###### 4.2.2 多 Agent 体系结构设计

多 Agent 系统结构需按照多 Agent 系统组织结构并结合实现分布式数据库管理的要求来设计。由需求分析的实现技术分析可知,需要实现分布式查询、分布式事务功能。

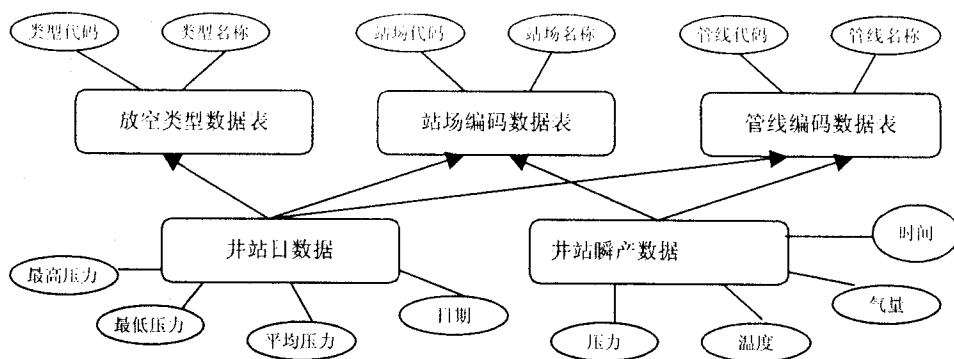


图 2 分布式数据库全局模式图

(1) 分布式查询过程: 由于生产数据表全局关系与分片后的逻辑片段是一对多,逻辑片段与物理片段也是一对多关系,故分布式查询处理需要通过水平分片原理实现全局查询转换到逻辑查询,而逻辑查询转换到物理查询是一对一的,根据数据字典映射即可。放空类型数据表、站场编码数据表、管线编码表这些全局关系与逻辑片段是一对一关系,而各采气队均有一份,所以逻辑片段与物理片段则是一对多关系,故分布式查询处理时,只是考虑通信代价、选择代价低的物理片段访问即可。用一个 DIAgent 接收用户查询(全局查询)要求,此 DIAgent 分析该查询并分解一个或多个物理查询,根据分布式查询的特点设计多个必要的物理查询,分别由一个 DIAgent 完成,这些物理查询的组合能够满足用户查询的要求。

(2) 事务处理具体过程设计: 数据更新涉及到分布式事务处理,实现数据一致性,分布式事务是由在多个结点执行的子事务组成的。由此用一个 DIAgent 接收用户提交的分布式事务,由它分析并把此事务分解为一个或多个子事务,提交到相应的数据库系统结点。传统的数据一致性的维护方法是采用两阶段提交机制,即假如该修改可以在本地提交,并且同时可以保证在所有该表的副本处提交,系统才可以接受该修改。但是这种方法事务容易进入阻塞状态,所以采用三阶段提交法,其过程是: 第一阶段,协调者向所有的参与者发“开始提交”报文,由每个参与者根据自己的情况进行投票,只有收到所有的参与者均赞成提交才进入第二阶段和第三阶段;第二阶段,协调者向所有的参与者发准备提交报文,参与者收到该报文后,若已经准备好提交,则回答“准备就绪”报文,否则进行夭折处理;第三阶段,当协调者收到所有的参与者“准备就绪”的回答,就向所有参与者发“提交”报文,此时每个参与者知道其它的参与者将“赞成”提交,因此它可以收到“提交”报文后进行提交。

本系统的多 Agent 组织结构适合采用行政管理型,所以有: 管理 DIAgent, 具有所管理的 DIAgent 名字、地址、能力; 处理分布查询的查询分解 DIAgent; 处理分布事务的事务协调 DIAgent。设计 DIAgent 系统结构如图 3 所示。具体包括本体 DIAgent 设计(ontologyDIAgent)、任务 DIAgent、功能解析服务 DIAgent、名字解析服务 DIAgent、查询解析 DIAgent、物理查询 DIAgent、事务分解 DIAgent、

事务协调 DIAgent、子事务 DIAgent 等的设计。

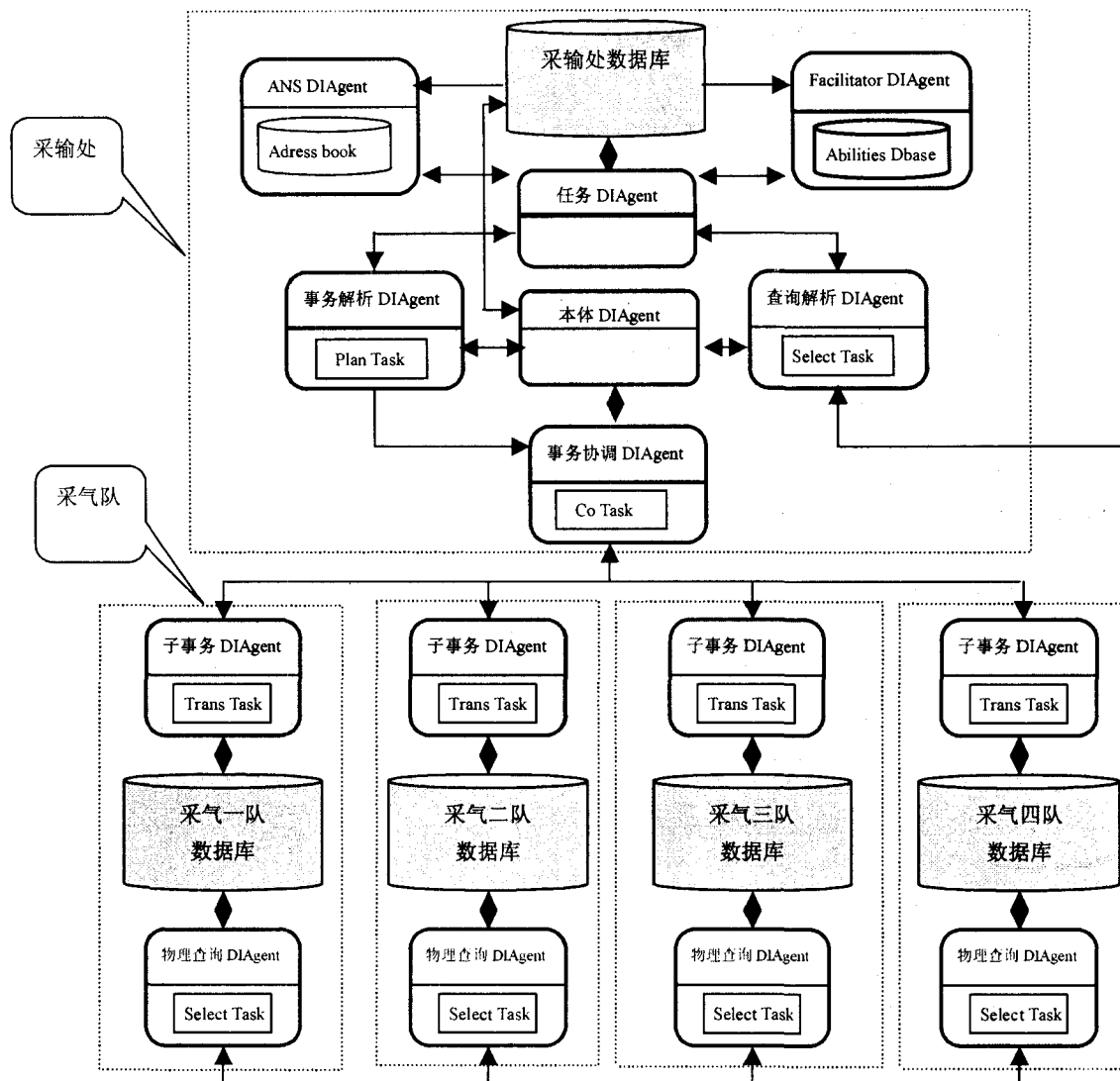


图 3 多 Agent 系统结构

### 5 结 论

文中笔者的创新观点是:通过研究多 Agent 技术与分布式数据库技术的特点,结合课题,提出了用多 Agent 技术集成现有数据库为一个具体全局模式的分布式数据库,为企业提供了新的数据库应用平台。通过文中的分析可以看出:所提出的集成方法为数据库系统集成提出了一个新思路,虽然目前还没有得到普及,但其本身所具有的优越性和强大生命力决定了它的发展前景必然十分广阔。

#### 参考文献:

[1] Carver Z, Cvetanovic Z, Lesser V. Sophisticated cooperation in

FA/C distributed problem solving systems: Proceedings of the 9th national conference on Artificial Intelligence[C]. Anaheim, CA: [s. n.], 1991: 191 - 197.

[2] Bond A H, Gasser L. An analysis of problem and research in Distributed Artificial Intelligence[C]//Bond A H, Gasser L. Reading in Distributed Artificial Intelligence. San Mateo, CA: Morgan Kaufmann, 1995.

[3] 李 英. 多 Agent 系统及其在预测与智能交通系统中的应用[M]. 上海: 华东理工大学出版社, 2004: 31 - 33.

[4] 萨师焯, 王 珊. 数据库系统概论[M]. 第 3 版. 北京: 高等教育出版社, 2000.

[5] 毛国君. 高级数据库原理与技术[M]. 北京: 人民邮电出版社, 2004.

(上接第 49 页)

[3] Yoshida Y. A continuous-time dynamic fuzzy systems, (1): A limit theorem[J]. Fuzzy sets and systems, 2000, 113(3): 453 - 460.

[4] 李凡长, 刘贵全, 余玉梅. 动态模糊逻辑引论[M]. 昆明:

云南科技出版社, 2005.

[5] Reiter R. A logic for default reasoning[J]. Artificial Intelligence, 1980(13): 81 - 132.

[6] 张继负, 刘 静. 表示和处理缺省知识的多维模糊推理方法和算法[J]. 小型微型系统, 2001, 22(7): 893 - 896.