

基于 MAS 的 RFID 物联网数据集成中间件研究与实现

陈毅红^{1,2}, 冯全源¹, 校景中²

(1 西南交通大学 信息科学与技术学院, 四川 成都, 610031; 2 西南民族大学 计算机科学与技术学院, 四川 成都 610041)

摘 要: 提出 RFID 数据集成中间件概念. 为了不改变现有各数据库局部模式, 提供各数据库的全局模式, 将 MAS 和分布式数据技术结合, 提出基于 Agent 的关系数据库访问技术的联邦模式虚拟集成法. 该虚拟集成法基于本体构建分布式数据库; 针对网络环境, 基于 CIAgent 开发包研究设计分布式智能代理 DIAgent, 用 KQML 语言描述 DIAgent 间协作通信过程, 由多个协作 DIAgent 构建 GDBMS. 实现了油田企业数据集成, 系统运行稳定可靠.

关键词: 物联网; RFID; Multi-Agent; 分布式数据库; 数据集成; DIAgent

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2012)04-0001-05

Research and Implement of Data Integration Middle Ware in Internet of Things Based on RFID Using MAS

CHEN Yi-hong^{1,2}, FENG Quan-yuan², XIAO Jin-zhong²

(1 School of Information Science & Technogy, Southwest Jiaotong University, Chengdu 610031, China;

2 School of Computer Science & Technogy, Southwest University of Nationalities, Chengdu 610041, China)

Abstract: In order not to change the existing local database models, providing the global model of databases, by combining the MAS technology and distributed data technology, The virtual integration way based on ontology agent is proposed, the way is relationship database accessing technology, and it is also federal mode. Based on Ontology the global model of distributed database is constructed in the way. In the network environment, The DIAgents are design based on CIAgent SDK, the communication between DIAgents is described by KQML language, GDBMS are composed of cooperative DIAgents. Using these technologys, data integration of the oilfield enterprise is realized, the system operation is stable and reliable.

Key words: internet of things; RFID; Multi-Agent; distributed database; data integretion; DIAgent

1 引言

数据集成核心是将相互关联的分布异构数据库集成在一起, 使用户能够以透明方式访问这些数据库^[1]. 本文提出用数据集成中间件技术来解决这个问题, 根据数据集成是否存在数据移动, 可将数据集成分为物化集成法和虚拟集成法^[2], 或称数据复制法和模式集成法^[3]. 物化集成法是通过将各个局部数据按照一定规则转移到中心库中, 实时性差; 虚拟集成法具有实时一致性和透明性好^[3], 适用于被集

成的系统规模大、数据更新频繁数据实时一致性要求高的情况.

传统虚拟集成法所研制的系统采用模块式结构, 系统构建复杂, 不便于系统扩展和升级, 为此需要寻找系统构建新技术^[4]. 由多个数据库来构建具有全局模式本质上是一种分布式数据库, 数据库管理系统具有智能特点^[5], MAS 分布式人工智能技术^[6]和分布式数据库间均具有并行性、系统结构相似、同源于分布式技术, 将分布式人工智能与分布式数据库相结合, 具有优点: ①通过智能 Agent 方式

收稿日期: 2011-07-01; 修回日期: 2011-08-31

基金项目: 国家自然科学基金重大项目(60990320, 60990323); 国家自然科学基金一中物院联合基金资助项目(10876029); 国家高技术研究发展计划(“八六三”计划)资助项目(2009AA01Z230); 中央高校专项科研基金(11NPT02)

来实现分布式数据库管理系统,发掘分布式数据库中固有智能,提升系统性能;②克服传统模块式结构所带来系统复杂性高的缺点,为此提出基于 MAS 的数据集成中间件研究 RFID 物联网数据集成,以实现一个跨平台、实时性和透明性好的物联网 RFID 远程数据集成系统,中间件技术有利用企业应用系统只关注高层应用逻辑,而无需要关注低层多点数据采集系统差异^[2]。

2 数据集成中间件系统模型

物联网系统中有 n 个 RFID 局部数据库系统,数据集成系统 $RFID_DIS=\{GDBMS, LDBS\}$, $GD-BMS$ 表全局数据库管理系统, $GDBMS=\{Task_DI-Agent, TranParse_DI-Agent, LocalTrans_DI-Agent, ANS_DI-Agent, Facilitator_DI-Agent, Ontology_DI-Agent\}$, $LDBS$ 是 RFID 局部数据库系统集合, $LDBS=\{LDBS_k=\langle LDB_k, LQ_DI-Agent_k, LT_DI-Agent_k \rangle \mid k=1, 2, \dots, n, LDB_k \in LDB, LQ_DI-Agent_k \in LQ_DI-Agent, LT_DI-Agent_k \in LT_DI-Agent\}$, 局部数据库集合 $LDB=\{LDB_k \mid k=1, 2, \dots, n\}$, 局部查询智能主体集合 $LQ_DI-Agent=\{Local_Query_DI-Agent_k \mid k=1, 2, \dots, n\}$, 局部事务智能主体集合 $LT_DI-Agent=\{LocalTrans_DI-Agent_k \mid k=1, 2, \dots, n\}$ 。

智能本体 $Ontology_DI-Agent$ 中的知识库 $Knowledge\ DB=\{gtfs, ltfs, ms\}$, $gtfs$ 为全局表集合 $gtfs=\{gt_i \mid i=1, 2, \dots, r\}$, $gt_i=\{gf_{ij} \mid j=1, 2, \dots, i_u\}$, 全局字段无歧义,即字段变量 gf_{iv} 和 gf_{ju} 值域分别为 $D(gf_{iv})$ 和 $D(gf_{ju})$, gf_{iv} 和 gf_{ju} 变量名表示为 $N(gf_{iv})$ 和 $N(gf_{ju})$, 当 $N(gf_{iv})=N(gf_{ju})$ 时,则 $D(gf_{iv})=D(gf_{ju})$, $i \neq j$, $gtfs$ 为所有全局表字段集合,则 $gtfs=\bigcup_{i=1}^r gt_i$; $ltfs$ 为局部表集合, $ltfs=\{lt_i \mid i=1, 2, \dots, s\}$, 局部表 $lt_i=\{lf_{ij} \mid j=1, 2, \dots, i_v\}$, $ltfs$ 为所有局部表字段集合, $ltfs=\bigcup_{i=1}^s lt_i$, 注如局部表字段有歧义时,通过在字段名前加前缀,从而消除歧义; ms 为局部表字段与全局表字段映射集合, $ms=\{\langle gf, lf \rangle \mid \langle gf, lf \rangle \in gtfs \times ltfs\}$ 。

定义 1 RFID 数据集成中间件指集成多个 RFID 数据采集系统数据库,向上层应用程序提供全局视图,从而屏蔽低层 RFID 数据采集系统差异,以达到系统透明性、可扩展性的目的。

3 系统总体架构

RFID 数据集成系统架构如图 1 所示,系统由两大部分组成:企业 RFID 数据采集系统,由局部数据库 $LDBS$ 和应用系统 $Application$ 组成;MAS 系统,由一系列具有网络通信能力相互协作的分布式智能代理 $DI-Agent$ 组成。

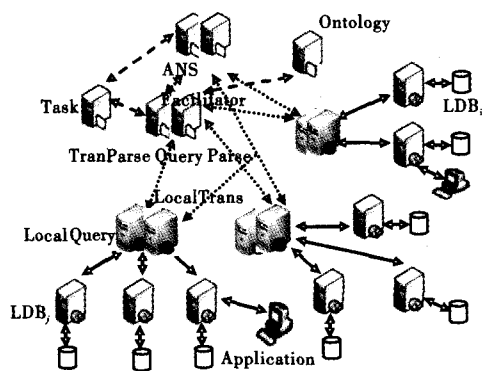


图 1 数据集成系统架构示意图

MAS 系统构建分布式数据库管理系统的 $GD-BMS$, 系统主要功能是全局数据查询和全局数据更新,数据更新新涉及到多个局部数据库,需要用分布式事务来完成数据更新操作,实现分布式事务的原子性。MAS 系统由多个分布式智能代理组成:任务 $Task_DI-Agent$, 简称为 tda ; 事务解析 $TranParse_DI-Agent$, 简称为 tpd ; 查询解析 $QueryParse_DI-Agent$, 简称为 qpq ; 局部查询 $LocalQuery_DI-Agent$; 局部事务 $LocalTrans_DI-Agent$; 地址服务 $ANS_DI-Agent$; 功能服务 $Facilitator_DI-Agent$; 本体 $Ontology_DI-Agent$; 本节剩余部分说明各 $DI-Agent$ 共同点,特定 $DI-Agent$ 构成和功能在第 4 节详细说明。

$CI-Agent$ 全用 Java 语言编写,可在 $JDK1.1$ 环境下运行, $CI-Agent$ 智能主体环境特点:很容易在已有的 Java 应用系统中加入 agent; $JDK1.1$ 包含 $java$ 、 $beans$ 软件包; agent 能支持相当强的事件处理能力;可以将领域知识加到主体知识库中;主体具有一定的学习功能;多主体是永久的,可随时将它写入文件或装入运行。

分布式智能代理 $DI-Agent$ 在 $CI-Agent$ 开发包基础上研发的具有网络功能的分布式智能 Agent。 $CI-Agent$ 无网络功能,不适用于网络分布计算环境,对开发包的源代码进行分析,充分吸取其中的多 Agent 技术,然后对其加以继承,增加网络通信模块后,以 $DI-Agent$ (Distribute Intelligence Agent) 为基

础对各 Agent 命名,DIAgent 是 CIAgent 子类。

DIAgent 采用事件机制消息通信和对消息处理,各 DIALOGent 由通信模块 Communicator、消息队列 Message Queue、事件监听器 Event Monitor、事件处理器 Event Processor、KQML 消息组装器 Message Assemble 组成,事件处理器构成 DIALOGent 逻辑部分。消息接收实现原理:在 bean 类(CIAgent 是 bean 的子类)设置一个属性,当过滤器收到 KQML 消息后,存放在该属性中,产生一个属性改变事件,对应此事件,有一个事件监听者,在监听者接口中,有处理 KQML 消息事件处理器方法(即所谓注册函数),该方法的一个事件参数,通过此参数,可得到消息内容。

通信消息:DIAGENT 通信过程是消息传送的过程,用 KQML 语言表示各 DIALOGent 间交互的消息。原语格式为:Performative_name(sender:, receiver:, from:, to:, in_reply_to:, reply_with:, language:, ontology:, content), KQML 被分为分为三层:内容层、消息层、通信层,内容层由 content 域表示,内容层表示的涵义由本体 Ontology_Dialogent 解释,通过本体实现全局模式与局部模式间映射,从而达到数据集成目的,内容层可以传送查询和更新命令、命令执行结果;消息层即整个 KQML 语句;而通信层即整个 KQML 消息的传送,但通信层要依赖 Sender 和 Receiver 两个域,Sender 属性指明 KQML 消息发送者,Receiver 指明 KQML 消息接收者。

网络通信模块:为满足 DIALOGent 间通信需要,通信模块具有如下特征:①网络通信模块支持与外界双向通信,与外界的通信接口应该有收发两个端口;②网络通信模块支持与宿主双向通信,与宿主的通信接口应该有两个端口;③网络通信模块内具有队列,保证通信的可靠性,防止拥塞而掉包发生;④网络通信模块作为服务方接收信息时,应支持多客户访问。本模块特点是各部分高度并行运行,保证在有多个 DIALOGent 的条件下通信要求,结构见图 2。

4 DIALOGent 智能主体设计与实现

本体 ontology_Dialogent:Ontology_Dialogent 由关于数据库元数据的知识库 Knowledge DB、消息组装模块 Message Assemble、消息队列 Message Queue、事件监视器 Event Monitor、事件处理器 Event Processor、通信模块 Communicator 组成,系统结构如图 3 所示。功能:知识库管理;为其他 DIA-

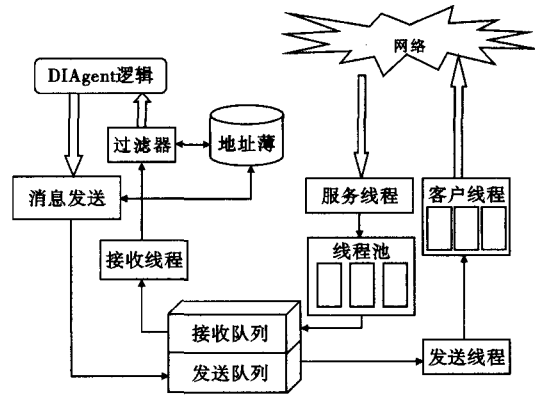


图2 DIALOGent 网络通信模块

gent 提供数据库元数据解析信息,当其他 DIALOGent 向 ontology_Dialogent 发出元数据解析请求,从将全局数据库表名和字段名映射为各局部数据库表名及字段名,从而实现分片透明性和位置透明性。

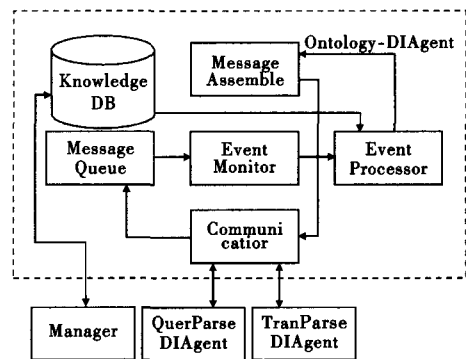


图3 本体 ontology_Dialogent 结构

任务 Task_Dialogent:接收用户提交的全局 SQL 语句,分析是事务还是查询,然后传送到 QueryParse_Dialogent 或 TranParse_Dialogent 处理。如果是查询,则要求用户查询的结果返回到任务 DIALOGent,如果是事务,则要求将事务处理情况返回到 Task_Dialogent。

功能服务 DIALOGent 和地址服务 DIALOGent:Facilitator_Dialogent 由消息组装模块 Message Assemble,消息队列 Message Queue,事件监视器 Event Monitor,事件处理器 Event Processor,通信模块 Communicator 组成,Active checks 模块功能是检测各 LocalQuery_Dialogent 和 LocalTran_Dialogent 活跃状态,Function 为记录各 DIALOGent 功能库。其功能是接收其它 DIALOGent 发来的注册 KQML 消息,消息内容有 DIALOGent 名称及功能,用 Register 原语表示消息,在 content 域是其功能描述;还要接收其它 DIALOGent 发来的查询,查询能够完成某种功能 DIALOGent 的名称,然后返回此 DIA-

gent 的名称, KQML 消息采用 KQML 原语 Recommend_one.

地址服务 ANS_DIAgent 结构与 Facilitator_DIAgent 相似, 有两个功能, 接收其他 DIAgent 发来的注册 KQML 消息, 包括的内容为 DIAgent 名称及访问地址, 访问地址由 IP 和 port 构成, 用 Register 原语实现, 在 content 域是其地址; 接收其他 DIAgent 发来查询 DIAgent 的物理地址消息, 然后返回此 DIAgent 地址.

查询解析 QuerParse_DIAgent: QuerParse_DIAgent 结构见图 4, 事件处理器 Event Processor 由查询分解 Query_Decomposition 模块和查询优化 Query_Optimization 模块组成, 心跳模块 heartbeat 用于定时向地址服务 ANS_DIAgent 发送表示其活性的消息. 功能: 系统启动后在 Facilitator_DIAgent 和 ANS_DIAgent 注册, 随后可接收 Task_DIAgent 提交的数据库查询 KQML 消息, 通过查询分解算法和查询优化算法^[7], 将其分解为一系列子查询, 也叫局部数据库查询, 向 Facilitator_DIAgent 查询能够各子查询的 DIAgent 的名字, 收到该完成子查询的 DIAgent 后, 如果通信模块地址簿中有该 DIAgent 的地址, 则直接向目标 LocalQuery_DIAgent 发送查询命令; 如果地址簿中没有地址, 则向 ANS_DIAgent 发送查询该 DIAgent 的地址消息, ANS_DIAgent 在其地址库中找到, 则返回其地址, 向目标 LocalQuery_DIAgent 发送查询命令; 然后等待各 LocalQuery_DIAgent 返回查询结果, 收到所有结果后, 综合子查询数据集 data set, 向 Task_DIAgent 返回结果, 全局查询解析过程的交互图如图 5 所示.

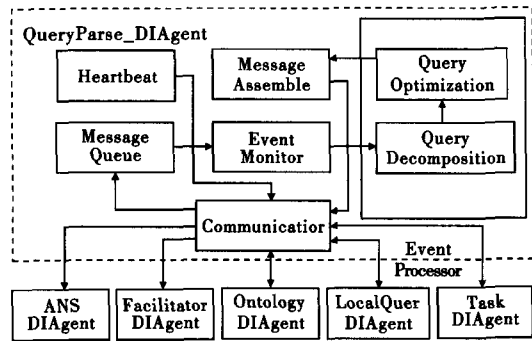


图 4 QueruParse_DIAgent 结构

事务解析 TransParse_DIAgent: TransParse_DIAgent 结构与 Quer Parse_DIAgent 相似, 其中事务处理器由事务分解 Transaction Decomposition 和事务协调 Transaction Coordination 两个模块组

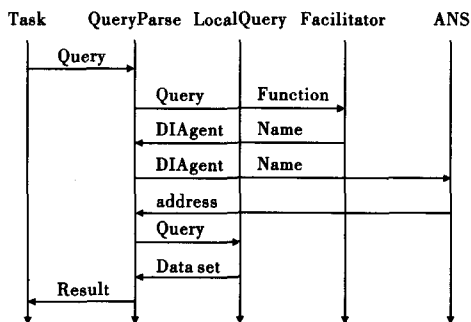


图 5 全局查询 DIAgent 交互图

成, 该 DIAgent 启动后向 Facilitator_DIAgent 和 ANS_DIAgent 注册, 接收 Task_DIAgent 提交的全局事务, 分解为子事务, 通过向 Facilitator_DIAgent 查询能够完成所有子事务所有 TranParse_DIAgent, 得到 TranParse_DIAgent 名称, 向所有局部事务的 LocalTran_DIAgent 发送开始提交请求命令, 启动计数器, 与其他 DIAgent 交互过程见图 6.

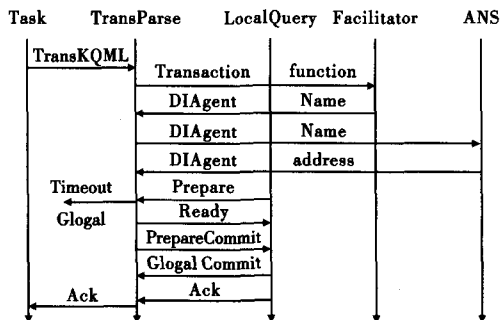


图 6 全局事务 DIAgent 交互过程

局部查询 DIAgent 和局部事务 DIAgent: 局部查询 LocalQuery_DIAgent 事件处理器由局部查询命令的组装器 Query Assemble 和缓存查询结果的数据队列 Result Queue, 消息队列 Message Queue 用于支持多用户查询, 与其他 DIAgent 交互. 局部事务 LocalTrans_DIAgent (结构见图 7) 的事件处

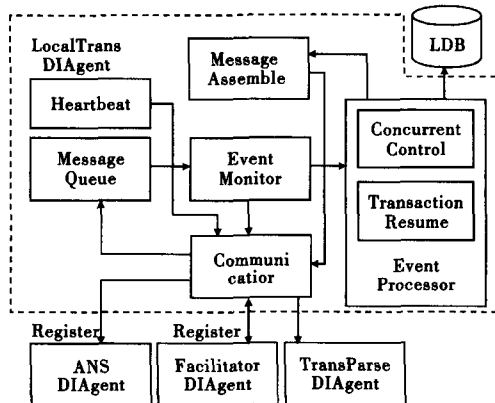


图 7 localTrans_DIAgent 结构

理器由事务恢复 Transaction Resume 模块和并发控制 Concurrent Control 模块组成。

5 应用实例

钻探公司有大量的钻井设备,存放这些设备库房有1个总部库房、10个井队库房、8个区域库房。为了便于管理,公司在三年前采用了RFID电子标签对设备出库和入库进行管理,大大提高了设备管理自动化水平。但随着信息化发展,公司需要将各库房数据集成在一些,以便对一公司设备使用情况进行管理,但又不能放弃现有数据采集系统,因此采用本文提出的数据库集成方案^[8]。系统基于JSP+Bean+JDBC技术,为四层分布式结构,以浏览器方式提供给用户使用,IE向Task-DIAgent提交任务,并由Task-DIAgent返回结果。通过JSP实现表示层Presentation layer与逻辑层通信;业务层business layer由DIAgent组成,并通过JDBC实现对数据层database layer的访问。运行软件环境为Windows 2003、IIS、JSP引擎为Resin。各数据采集点数据传输网络基于VPN方式连接。可按照库房、日期、设备类型、使用者等条件查询。支持对各库房基础数据更新,如设备分类基本信息、用户权限,测试表明系统运行稳定。

6 结束语

本文提出的基于MAS的数据集成中间件技术来构建全局系统,在RFID数据采集系统基础上,对应增加DIAgent即可,不会对现有系统造成影响。如果要新增加一个RFID数据采集系统到本系统中,只需要增加二个对应DIAgent,从而实现系统可扩展性。提出了具有网络功能的Agent概念——网络分布式智能代理,并将无网络环境CIAgent改造成有网络功能的DIAgent;采MAS技术减少了系统的耦合度,降低了系统开发难度;提出了DIAgent功能和地址注册机制,功能服务DIAgent和地址服务DIAgent中记录了各LocalTran_DIAgent和LocalQuery_DIAgent在线状态,实现了系统容错机制;系统实现采用Java技术,实现系统跨平台性。通过实例系统测试表明所研究和设计的数据集成系统的科学性和合理性,从而实现了基于Multi-Agent技术的数据集成中间件开发与应用。

参考文献:

- [1] Maurizio L. Data integration: a theoretical perspective [C]// Proc. Of the ACM SIGACT_SIGMOD_SIGART Symposium on Principles of Database Systems USA:Madison, 2002.
- [2] 王越,杨平利,杨梅. 一种分布异构数据库集成中间件的设计与实现[J]. 微电子学与计算机, 2009, 26(7):81-85.
- [3] 王行荣. 异构数据集成方法研究与实现[J]. 微电子学与计算机, 2006, 23(4):172-174.
- [4] Park Jinsoo. Schema integration methodology and toolkit for heterogeneous and distributed geographic database[J]. Journal of the Korea Industrial Information Systems Society, 2001, 6(3):51-64.
- [5] Amarnath Gupta, Ashraf Memon, Joshua TranInformation, et al. Mediation across heterogeneous government spatial data sources[C]// Proceedings of the 2002 annual national conference on digital government research. California, 2002:1-6.
- [6] 陈跃国,王京春. 数据集成综述[J]. 计算机科学, 2004, 31(5):48-51.
- [7] 陈毅红. 基于CIAgent技术的分布式数据库研究与设计[D]. 四川:西南石油学院, 2005.
- [8] 唐富年,姚莉,李金洋. 基于本体的关系数据库访问研究进展[J]. 小型微型计算机系统, 2011, 32(3):390-396.

作者简介:



陈毅红 男, (1972—), 博士研究生。研究方向为RFID、嵌入式系统、物联网、排队论。



冯全源 男, (1963—), 教授, 博士生导师。研究方向为RFID技术、集成电路设计、嵌入式系统、移动天线与智能天线系统、微波及毫米波技术、自适应信息处理、电磁兼容与环境电磁学、微波器件及材料。