

基于 SOA 的数据协同模型

易峥荣, 卜 炜, 葛序风, 刘颖娜

(华东计算技术研究所, 上海 200233)

摘 要: 传统的企业信息集成采用的是功能驱动的方法, 主要用于解决企业数据的分布性与异构等问题, 它们通常都忽略了业务逻辑而直接进入企业业务集成的层次。然而不在数据集成的层次中响应企业业务流程的处理要求, 就会直接限制系统的数据处理能力。该文基于用户的业务逻辑, 通过事件驱动数据在不同组织或系统间流动来实现数据的集成。这是一种松耦合的数据集成模式, 用户通过协同不同系统提供的能力来控制系统的集成行为及数据转换规则, 适用于数据共享需求变更频繁的应用集成系统。

关键词: Web 服务; ECA 规则; 计算机支持的协同工作

Data Cooperative Model Based on Service-oriented-architecture

YI Zheng-rong, BU Wei, GE Xu-feng, LIU Ying-na

(East China Institute of Computer Technology, Shanghai 200233)

【Abstract】 The traditional enterprise information integration is driven by function, which mainly solves the distributed and heterogeneous data integration problems, but without the consideration of the business logic, stepping strictly into the business integration level. If no actions are taken at the data integration level to response the enterprise business work flow, it restricts the system ability of sharing data. This paper provides a data cooperative model which is based on business logic of enterprise. It invokes the event which drives data flowing between different organizations. It is a loose coupling model and gives customer the ability to change functions and data transform rules. The system which changes a lot will benefit from the model.

【Key words】 Web service; ECA role; computer supported cooperative work

1 概述

近年来, 随着企业逐步实现职能优化, 企业的资金流、物流和生产效率都得到提高。然而职能部门主权的提高使得相互间的协作成本越来越高。为解决上述问题, 提出了业务流程^[1]的概念。企业以流程为中心, 通过优化的组织结构、灵活易变的流程设计来提高资源利用率, 从而降低成本。由于企业业务流程灵活多变, 因此对信息共享和数据流转的灵活性提出更高的要求。

传统的信息集成技术^[2]通过专用的、点对点的接口程序来实现信息共享和传递, 虽然能够解决集成中信息的分布性与异构的问题, 但忽略了与业务逻辑的关联, 从而无法适应企业业务流程频繁变更的需要。

本文针对当前企业信息化建设中的特点和需求, 采用 Web 服务技术, 引进基于事件-条件-动作(Event Condition Action, ECA)规则^[3]的事件驱动模型, 提出基于面向服务的架构(Service Oriented Architecture, SOA)的数据协同模型, 通过事件驱动数据在不同组织之间的流转, 达到信息共享的目的。该模型具有分布性、松耦合等特点, 提供可基于企业业务逻辑的流程表达和执行能力, 允许信息系统根据预定义的规则, 通过事件、触发条件和处理活动等驱动数据的流转, 易于实现数据共享协同的自动化, 提高生产效率。

2 研究背景

现在的企业信息集成方法一般分为 2 种: (1)数据仓库; (2)代理/中介(wrapper/mediator)方法^[4]。但是数据仓库由于集中式的一般缺点, 不适合分布的特点, 已经很少再被采用; 而 wrapper/mediator 方法既可以保留企业现有的信息系统不

变, 又有利于企业业务流程的执行, 因而被广泛采用。

Web Service 是一种以 SOAP 为轻量级传输协议、以 XML 为数据封装标准、基于 HTTP 的组件集成技术。Web Service 技术能有效地封装分布的数据源, 从而极大提高系统的延展性, 从根本上保护了业已存在的投资, 是一种比较理想的实现代理/中介方法的技术机制。

SOA 是一种面向服务的、松耦合架构, 能够方便地将 Web 服务整合集成, 以提供系统级的数据集成和转换, 具备很好的延展性和灵活性。

事件驱动模型^[5]以事件为触发, 根据条件判断采取什么动作, 也叫基于 ECA 规则的事件驱动模型, 它具有及时高效、灵活的特点。采用基于 ECA 规则的事件模型能很好地表达企业业务流程, 满足企业数据协同时效性要求。基于 ECA 规则的数据协同模型具有很强的表达能力, 能表达实际业务中的多种情形, 可以大大增强数据共享的适应性。

本文提出基于 SOA 的数据协同模型, 基于 Web 服务实现分散信息的集成和封装, 它以 SOA 为系统架构, 通过 ECA 规则事件驱动数据的流动共享^[6], 具有松耦合、分布、灵活的特点。同时, 模型将流程表示和应用程序分开, 既满足了企业信息共享的需求, 又满足了企业业务流程的灵活性表达和执行。

3 基于 SOA 的数据协同模型体系结构

基于 SOA 的数据服务协同模型的目标是将分布在网络

作者简介: 易峥荣(1982 -), 男, 硕士研究生, 主研方向: 分布式计算; 卜 炜, 高级工程师; 葛序风、刘颖娜, 工程师

收稿日期: 2008-06-28 **E-mail:** linbirg@gmail.com

中的各类独立数据服务汇集起来,提供面向流程的按需数据服务协同能力,满足企业环境下用户对数据灵活共享的要求。

其系统结构如图 1 所示。

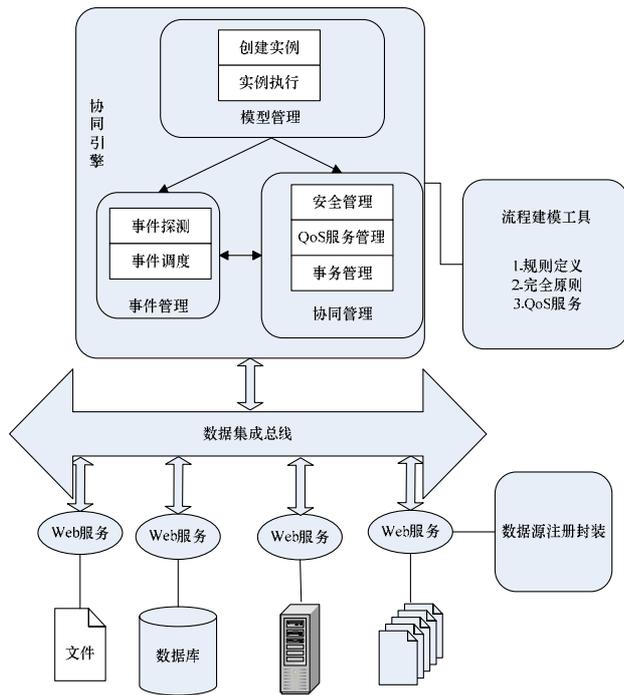


图 1 系统结构

系统以 Web 服务作为统一的数据接入接口,即通过一个数据源封装注册管理工具,将企业中各种异构、分布的数据源封装成标准的 Web 服务,用以屏蔽数据源之间的差异,提供统一的数据访问方法。这些 Web 数据服务通过数据集成总线勾连起来,形成一个松耦合的集成平台,为各个数据源提供信息共享的通道。

服务协同引擎是数据服务协同系统的中枢,用于实现数据流模型解释识别、转换为自定义模型,并映射为对应的数据流实例,与集成总线结合,提供完整的运行环境。它通过流程建模工具定义的流程描述文件,将各个 Web 服务组织起来,驱动数据在系统中的共享和流转。

3.1 服务协同引擎

服务协同引擎是数据服务协同系统的中枢,用于实现数据流模型解释识别并转换为自定义模型,为数据流实例提供运行时的执行环境,包括实例的创建、激活、挂起、终止,并按定义流程和数据信息导航推进实例。它由 3 个模块组成,即模型解释、事件管理、协同管理。模型解释负责读入由模型建模工具生成的模型描述文件,解释并创建流程实例,负责实例的运行时的状态维护、异常处理等,模型解释部件是引擎中的基础部分,为系统的日常运行提供基础支持部分;事件管理负责探测系统中事件的发生、消息的传递和事件调度执行,事件探测部件是系统得以有效运转的基石;协同管理负责系统的安全管理、QoS 服务、事务管理等,协同管理部件协调用户和系统之间的需求要求,它保证了系统的有序高效运行。

服务协同模型的核心机制是 ECA 规则处理机制。

定义 1 一个 ECA 规则是一个三元组,定义为 $R=(E,C,A)$ 。其中, E 为激活该规则的事件; C 为条件集,用以反映环境中的不同情况; A 为操作集(或动作集合)。

C 在这里通称为对象变量。对象变量是数据流中涉及的影响工作流路径选择的特性或参数,简称变量。变量以 x_i 来表示,变量集合记为 $X=\{x_1,x_2,\dots,x_k\}$ 。

变量 x_i 的取值范围为域,用符号 D_i 来表示。域有多种多样的形式,例如连续数域,速度的范围为 $(1.0\text{ m/s}\sim 1.2\text{ m/s})$,离散域,如颜色{红,黄,蓝}。

条件 ci 是 $X_i=\{x_{i1},x_{i2},\dots,x_{ik}\}\subseteq X$ 上的 k 元关系,是笛卡尔积 $D_{i1}\times D_{i2}\times\dots\times D_{ik}$ 上的子集。条件 ci 是对变量之间关系的限制。条件集合记为 $C=\{c_1,c_2,\dots,c_j\}$ 。

操作可以表示为 $A=(E,Ta,Ob)$ 。 E 为触发操作的事件; Ta 为操作的类型; Ob 为操作的对象。

事件是一个函数,它将时间映射为布尔值,可以表示为 $E:T\rightarrow\{True, False\}$

其中, T 称为事件 E 的发生时间。

事件可以分成原子事件和复合事件。原子事件是那些可以被工作流管理系统直接探测到的事件。原子事件可以分为绝对时间事件、活动状态变化事件和操作事件。

为了表达数据流执行中的复杂逻辑,原子事件可以复合。如常用的 AND, OR 等复合算子,可以用这些算子定义更复杂的事件,或者叫复合事件。

在 ECA 规则的基础上,定义了数据协同模型。一个数据协同的运行实例是一个大的活动,活动用 p 来表示。每一个活动有其内在的运行机制,它依据特定的输入产生特定的输出,用流关系来描述活动的这种特性。

定义 2 设 p 为某一活动,活动流关系可以定义为 $F=I,O,S,R_s$ 。其中, I 为活动的输入对象集合; O 为活动的输出对象集合; S 为活动的状态集合; R_s 为活动的状态转换规则集合,它是一组 ECA 规则。

活动的缺省状态有:等待(W),开始(B),挂起(H),执行(E),完成(C),超时(T),放弃(A),图 2 所示为活动状态图。

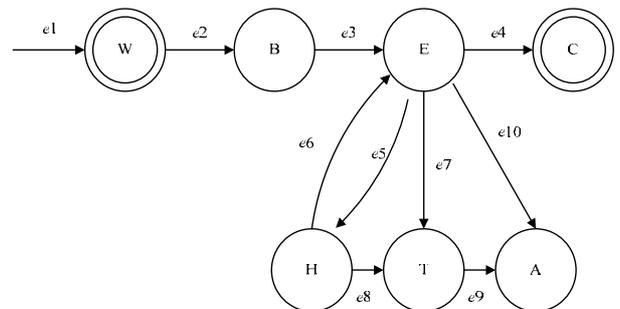


图 2 活动状态图

各个输入对象和输出对象也有其相应的状态。

活动可以用五元组 $p=(ID,F,Ac,T)$ 来表示。其中, ID 为活动的全局唯一的标识; F 为活动的流关系; Ac 为活动的执行者属性; T 为活动的时间属性。

活动的执行者可以是某一远程的服务,也可以是本地的应用,甚至是具有自我决策能力的 Agent。活动的时间属性包括计划开始时间、周期、实际开始时间、实际周期。活动可以被分解,即 $p=(P,Rc,Rm)$, P 为子活动集合, Rc 为流程控制规则, Rm 为监控规则。被分解的活动称为复合活动。 Rc, Rm 为 ECA 规则。活动进行层次分解后,需要一定的机制来保证模型之间的一致性。

数据协同模型的另一个重要机制是事件探测机制。数据协同模型中的事件,如前面分析,分为原子事件和复合事件,

原子事件又可以分为绝对时间事件、活动状态变化事件和操作事件。这些事件从探测的角度来说,又可以分为3种类型:auto, pull, push。在系统中也对应着3种感应方式。auto型事件根据预先定义好的触发条件,自动触发任务的执行,如定时执行等。这类时间通常适用于数据的集中。pull型时间协同服务系统自动去感知相应信息系统中的触发条件,一旦条件满足就执行任务。这类事件适合于数据采集等任务。push型事件由单独的数据服务向协同服务系统发出触发条件,协同服务感知到请求后马上触发事件。这类时间往往应用于有事务触发的业务中。在定义每个事件实例的处理方式时使用了ECA规则,ECA规则负责建立事件和事件行为间的关联。

总之,协同模型定义了数据服务的7种活动(等待、开始、挂起、执行、完成、超时、放弃)状态,并通过特定的事件触发规则(往往对应于业务逻辑),将不同的数据服务勾连起来,实现业务逻辑相关的数据服务的协同。

3.2 数据源封装注册管理工具

数据源封装管理模块的目标是为底层异构数据源提供一套完整的注册流程,从而封装用户对数据源的物理数据访问。

该层将收集维护关于数据源访问的所有必要信息,以及描述数据语义和QoS的信息,并保证各数据源的动态增删过程中系统的稳定。

如图3所示,数据协同最底层的数据源是来自各地的异构数据源,包括关系型数据源、XML型数据源或Web Service数据源,数据源注册到系统时需给出一个元数据描述。另外数据源需要对每个局部类给出完整的数据访问方法:对于关系型数据源,提供SQL语句访问接口;对于XML数据源,给出XQuery语句的访问接口;对于Web Service数据源,通过指定的Web方法及参数给出访问接口。

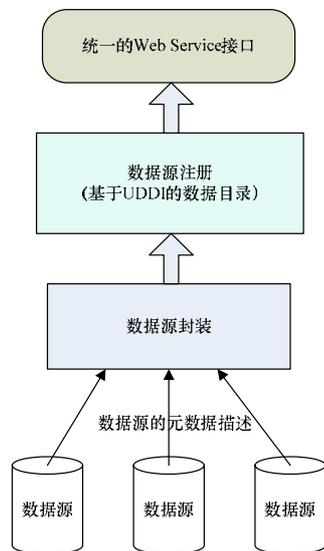


图3 数据源注册及封装

数据访问层封装了底层数据源的物理访问,所有非Web Service数据源都将在这里由一个统一的Web Service来控制数据访问,该服务将所有异构数据转换成笔者规定的基于XML的数据交换格式,而那些Web Service数据源,它们也将被要求以此格式提供数据。也就是说,上层只需通过调用一些Web Service接口即可获得统一格式数据,从而屏蔽底层异构数据访问方式的影响。

数据目录主要存储各类数据服务的统一表示视图,包括

数据服务的功能信息(如服务的内容、目的、参与者等)、非功能信息(如安全因素、服务质量等)、服务过程逻辑(如接口信息、子组件描述)等,是实现数据服务协同的基础。

注册接入服务将分布在网络中的各类标准的数据服务添加相关的扩充信息后,形成统一的数据服务视图,并注册到数据目录中。

3.3 流程建模工具

流程建模工具提供了一个图形化和用户友好的方式构建数据协同流程。它通过提供图形提示和检查器,可以简化流程、合作伙伴链接和变更设置;可以根据运行条件和变量,快速地创建复杂的并行分支流逻辑;可以管理参与服务交互的伙伴服务、角色和端口类型,设计异常处理机制,而不必直接操作XML语句;可以直接利用WSDL数据创建流程图,提高了基于ECA驱动的数据服务流程的开发效率。

建模主要是详细描述具体的业务中数据的流程模型。它借鉴了WFMC提供的核心元素-过程的概念。通过过程对业务过程进行形式化描述,用来支持数据服务协同系统的建模和执行的自动化。这里的过程包括一系列活动和基于事件的活动间关系、过程的起始和终止信息以及有关个体的信息,如参与者、有关的应用和数据等。

为了提供拖放功能和良好的可视性,数据服务协同建模工具的数据流程建模设计采用了基于Applet的可视化B/S结构,主要完成过程中的活动(Active)和连接(Link)等核心元素的定义。整个流程建模用支持拖放的图标表示活动节点,箭头表示流程,使得用户可以随意安排界面。用户在完成模型的创建之后还可以通过单步调试来验证流程的正确性。

4 模型应用

笔者实现了本系统的3个基本核心软件:

- (1)数据源封装工具;
- (2)流程定义工具;
- (3)流程执行引擎。其中,数据源封装工具以Axis为SOAP客户端发布服务;流程执行引擎部署在服务器,对用户不可见,而用户可见的是流程定义工具,它提供拖放功能和良好的可视性。

下面以某房地产公司与银行合作开展售房业务的过程为例描述本模型的应用过程。该公司提供房源信息查看和自动售房业务。自动售房业务通过与银行合作完成,用户向银行提出房贷请求,由银行审核完成,将房款自动划拨房地产公司完成售房业务。

在这个过程中,客户首先利用pull类型事件向公司售房部提出查看房源信息请求,售房部利用auto类型事件自动从房源管理中心获得房源信息并推荐给客户;客户向售房部预订,并利用pull类型事件向银行房贷部申请房贷;银行利用auto类型事件向贷款人资格审查中心请求审查贷款人资格,得到认证后利用auto类型事件自动将房款划拨到房产公司;房产公司资产管理部门收到前款,可以利用auto型事件通知售房部向顾客出售楼房;用户最终购得需要的楼房。通过本系统,在房地产公司内部、银行内部以及房地产和银行之间达到信息的自动流转和共享,并通过自动的业务流程完成自动售房业务,实现了企业业务的自动、高效。同时,松耦合的系统结构为公司和银行开展其他的业务提供了便利。图4给出了房产公司和银行开展的业务流程示意图。图5是用流程建模工具设计的业务流程效果图。

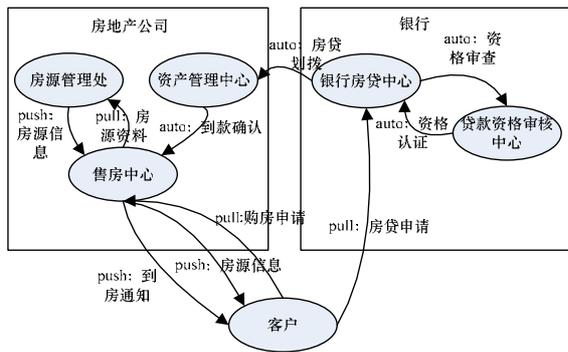


图 4 房产公司售房业务流程



图 5 由建模工具设计的业务流程效果图

5 结束语

随着企业组织各个独立职能部门权利的提升、职能的优

化、部门的效率得到提高，职能部门之间的合作会越来越频繁，而与之相对的协作成本却越来越高。再造企业的业务流程，通过高效的、灵活多变的业务流程的执行来提高效率是未来的趋势。这使得企业信息集成带来信息量大，共享变更频繁。

本文提出了引入 ECA 规则的 SOA 数据协同模型，能满足这种频繁变更的共享的需求，提出并实现了原型系统中的基本功能，包括数据源封装注册管理工具、数据协同引擎和流程定义软件。但是功能还不完善，下一步主要研究和完善基于 QoS 的自动服务调度，Web 服务接入方式的安全控制，以及增加系统的事务处理能力等相关功能。

参考文献

- [1] 罗海滨, 范玉顺, 吴 澄. 工作流技术概述[J]. 软件学报, 2000, 11(7): 899-907.
- [2] Mcgovern D. Data Integration[J]. EAI Journal, 2002, 15(6): 53-60.
- [3] 徐正权, 王治国. 基于 ECA 规则的工作流过程建模[J]. 计算机工程与科学, 2006, 28(5): 105-108.
- [4] 宋宇宁, 尹建伟, 陈 刚, 等. 基于事件驱动的柔性数据集成研究[J]. 计算机应用, 2004, 24(12): 120-123.
- [5] Chiu D K W. Developing Workflow-based Information Integrated (WII) with Exception Support in a Web Services Environment[C]// Proc. of the 37th International Conference on System Sciences. Hawaii, USA: [s. n.], 2004: 80-90.
- [6] 郑白桦, 袁潜龙. 基于信息传送和事件驱动的分布式工作流系统[J]. 计算机科学, 1999, 26(12): 53-57.

(上接第 260 页)

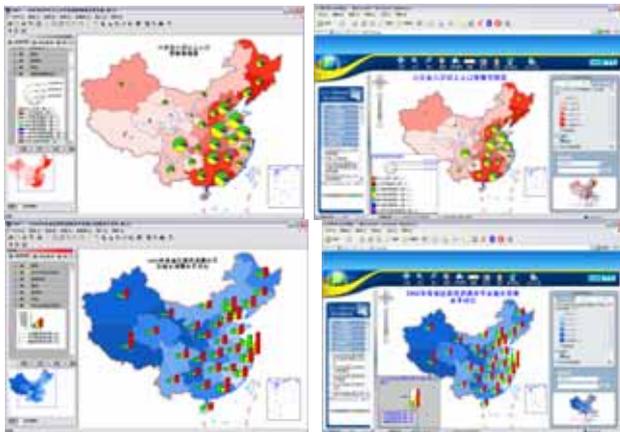


图 4 统计电子地图发布效果图

5 结束语

基于 Web 的统计电子地图发布系统设计是一个将统计电子制图、AJAX、WebGIS 和空间数据库技术相结合实现传统

桌面统计电子地图向网络统计电子地图快速、平滑转换的统计电子地图发布系统设计范例。随着网络技术的发展、信息化进程的加快，统计电子地图网络化将成为未来研究的一个热点。该设计针对传统桌面统计电子地图向网络统计电子地图转换过程中的关键问题提出了自己的解决方案，该方案具有一定的可行性和参考价值。

参考文献

- [1] 余卓渊, 王英杰, 苏 莹, 等. 中华人民共和国人口电子地图集系统设计[J]. 地球信息科学, 2005, 7(2): 23-29.
- [2] 苏 莹, 王英杰, 余卓渊, 等. 人口信息空间可视化系统设计研究[J]. 测绘科学, 2005, 30(3): 38-40.
- [3] 廖 克. 现代地图学的最新进展与新世纪的展望[J]. 测绘科学, 2004, 29(3): 5-9.
- [4] 伏玉琛, 周洞汝. Web GIS 的新技术[J]. 计算机工程, 2004, 30(3): 19-21.
- [5] 王英杰, 袁省勤, 余卓渊. 多维动态地学信息可视化[M]. 北京: 科学出版社, 2003.