

基于 ESB 与 Agent 的共享数据中心同步模型研究

李映红¹ 何 婧² 沈 立¹ 何丽波³ 范博文³

(云南省无线电监测中心信息科 昆明 650228)¹ (云南大学软件学院 昆明 650091)²

(云南大学信息学院 昆明 650091)³

摘 要 为了解决云南省无线电监测中心在信息化建设的过程中如何保持应用系统之间数据一致性的问题,对数据同步及其相关技术进行研究。通过分析 and 总结数据同步研究领域的研究方向和内容,提出了基于 ESB 以及 Agent 实现数据同步的方法,并采用包含共享数据中心的模式,最终提出了一个数据同步系统的新架构。

关键词 数据同步, ESB, Agent, 共享数据中心, 同步策略

中图分类号 TP393.09 **文献标识码** A

Data Synchronization of Data-sharing Center Based on ESB and Agent

LI Ying-hong¹ HE Jing² SHEN Li¹ HE Li-bo³ FAN Bo-wen³

(Information Department of Yunnan Radio Detection Center, Kunming 650228, China)¹

(National Pilot School of Software, Yunnan University, Kunming 650091, China)²

(School of Information Science and Engineering, Yunnan University, Kunming 650091, China)³

Abstract Radio monitoring center of Yunnan province has the process of information construction. In order to solve the problem of how to keep data consistency among application systems in this process, some research on data synchronization and relevant technologies were done. Based on analyzing and summarizing the research directions and content of data synchronization, a method was implemented based on ESB and Agent. Finally, a new architecture of data synchronization was implemented with the technique of data-sharing center.

Keywords Data synchronization, ESB, Agent, Data-sharing center, Synchronization strategy

云南省无线电监测中心在信息化建设的过程中,各类无线电管理业务系统彼此孤立,业务流程容易中断,业务数据“碎片化”,信息资源的使用效率整体偏低,这些“信息孤岛”无论在使用上还是在维护上,都会为实现数据的统一化增加一定的工作量,造成了重复开发和建设资金的大量浪费。为了实现各个系统之间的自动数据同步,将各个系统间的信息真正整合起来,实现系统间的信息互动,提出了共享数据中心方案。共享数据中心可以解决大部分“信息孤岛”的问题,保证在多个应用系统中同类数据的一致性,同时也减轻了系统管理人员对数据同步处理的工作量,使得多个现有系统之间的数据整合成为可能^[1-4]。

建立共享数据中心,必须考虑到异构数据库的同步问题。目前市场上基于异构数据库进行数据同步的产品有很多,这些平台虽然能够很好地实现数据交换,但是费用比较高。第三方的应用系统直接访问数据库,在数据的安全上存在一定的问题,给第三方的应用开发增加了很大的工作量,也不是最好的解决方法^[5,6]。使用 SOA 架构中的 ESB 技术进行数据同步,不仅是国内外大多数企业进行数据同步的主要方式,同时也便于应用到基于 SOA 的无线电监测网数据交换平台中。

本文主要研究如何利用基于 ESB 以及 Agent 的共享数

据中心技术实现各个现有应用系统之间的数据同步,以解决目前云南省无线电监测中心异构信息系统之间的数据共享和数据交换问题。

1 相关技术

1.1 ESB

企业服务总线(Enterprise Service Bus, ESB)是面向服务体系架构(Service-Oriented Architecture, SOA)中的一个重要的概念,是传统中间件技术与 XML、Web 服务等技术相结合的产物。它基于开放的标准,为企业之间不同的应用程序提供了一个准确、高效和安全的传递环境。它提供的解决方案成本更低,能够消除不同系统之间的技术差异,以便各个已有系统可以协调运作,实现不同服务之间的通信与整合。

在一个复杂的企业计算环境中,如果服务提供者(Service Provider)和服务消费者(Service Consumer)之间采用直接的端到端的交互,那么随着企业内应用程序的增加和复杂度的提高,最终应用程序之间的关联会逐渐变得非常复杂,形成一个网状结构,这将带来昂贵的系统维护费用,同时也使得 IT 基础设施的重用变得困难重重。ESB 提供了一种基础设施,消除了服务消费者与服务提供者之间的直接连接,使得服务

李映红(1981—),男,硕士,工程师,主要研究方向为无线电监测与管理信息化;何 婧(1978—),女,博士生,讲师,主要研究方向为云数据管理、数据挖掘、知识发现;沈 立(1980—),男,硕士,工程师,主要研究方向为无线电管理信息化;何丽波(1982—),女,博士生,主要研究方向为云计算、语义网;范博文(1988—),硕士生,主要研究方向为计算机应用技术、信息安全, E-mail: fanbw-1988@163.com(通信作者)。

消费者与服务提供者之间进一步解耦^[7],如图1所示。

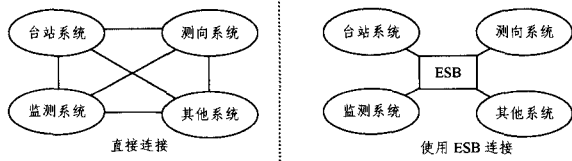


图1 ESB 示例图

ESB 的上述特点,使得它非常适用于现有的云南省无线电监测中心异构信息系统的整合过程中。整个云南省无线电监测中心在信息化建设的过程中,共享数据中心的 ESB 通过 Oracle 的 OSB 组件来实现,主要负责监控各系统数据由接口进入,经过清洗、存储最后分发到其他应用系统的整个业务编排。通过 OSB 实现业务的灵活组合处理。

1.2 Agent

数据库同步更新技术根据它的两步物理过程的差异形成了两种同步更新方式:紧凑一致性方式(Tight Consistency),要求修改过程和复制过程同时进行,即数据所有拷贝在任何时间都应保持一致;松散一致性方式(Loose Consistency),允许修改过程和复制过程可异步进行,两者之间存在一个时间延迟^[8]。紧凑一致性方式虽然能够保证数据一致的实时性,但是存在远程传输速率低、相应时间长以及费用开销大等不足;而松散一致性方式中,必须考虑修改于复制过程中存在的时间延迟,如果想要保持数据的每个拷贝之间都一致,那么修改必须停下来等到复制过程完成,无法保证需要实时性数据请求的时间要求。

综合上述两种方式的利弊,引入 Agent 的概念。它拥有智能 Agent 的自主性、反应性、主动性和交互性^[9,10],可以根据数据的实时性要求以及不同系统的请求,在不同的同步更新方式之间进行自主切换。在本次研究中,Agent 的作用还包括智能控制数据的清洗以及发布,这部分功能会在下节中详细描述。

2 数据同步模型

2.1 组件描述

传统的 EAI 实现方式是早期保持数据同步的一种方法。这种方式需要各个系统对与之交互的系统都要提供相应的数据接口,开发难度以及工作量较大,接口依赖性高,实现周期不可控^[11,12]。在本模型中使用共享数据中心技术,通过数据集成的方式,将所有数据以并集方式存放在共享数据中心。如果发生数据变化,首先将相关数据的改动发送到共享数据中心,再由共享数据中心整理后分发到各个系统并更新数据。

ESB 的存在以及 XML 语言的使用,屏蔽了各独立系统之间的异构性。因此 ESB 是应用层之间以及应用层与代理层交互的桥梁。引入 Agent 概念,是为了自主地控制在数据同步的各个阶段所采用的策略,使多个独立的系统集成成为一个一体的协同系统平台,并在这个系统平台中实现数据同步,如图2所示。

同步监听器会监听应用程序的存取数据库操作,并将所有数据库字段的 XML 信息存入数据字段表中。数据字段表中存放每个系统的数据集合,并将所有系统的数据字段表以并集的方式整合,作为共享数据中心的数据集合。若应用系统有存取数据库操作,那么数据也会在同步监听器的控制下,通

过 ESB 进入到数据交换区,并将应用系统变更的数据导入到数据交换区的相应增量表中。

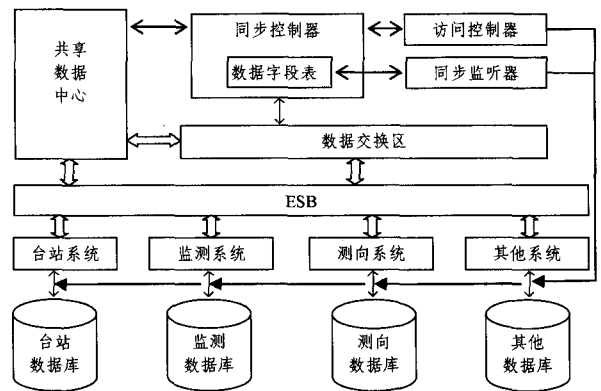


图2 基于 Agent 与 ESB 的共享数据中心同步模型

同步控制器主要有3个作用。第一个是控制数据清洗,即对数据交换区的相应增量表中的临时数据进行清洗,并存储到共享数据中心。数据清洗过程主要包括:字段类型匹配判断(通过映射逻辑转换后再进行判断)、字段长度限制判断、字段非空判断、惟一性检查(数据来源中含有重复数据)、冲突数据、插入数据操作检验和修改数据操作检验等。第二个是控制数据分发。控制器匹配各个系统的数据字段表,若其他应用系统(非提交数据的应用系统)数据库也包含该字段,则以数据库为单元整理出每个数据库的应同步数据字段并推送到应用系统各自的数据交换区当中。这个过程不用再进行数据清洗,只要进行简单的格式转换就可以了。第三个是控制数据传送,根据访问控制器提供的应用系统实时性要求,智能切换紧凑一致性同步与松散一致性同步,并可以根据每个应用系统的请求频率设定松散一致性同步的数据传送频率,提高整体效率。

访问控制器负责监听并控制应用系统对其数据库的读取操作。当应用系统有读取操作时,访问控制器会截取读取操作,将预读取的数据字段提交给同步控制器,根据同步控制器的返回信息,给予请求读取数据库的应用系统反馈信息;如果预读取数据不存在未完成的同步更新,则允许读取;否则需要等待同步控制器对其进行更新。访问控制器通过这种方式协助数据同步以确保其读取数据的准确性。

2.2 实现过程

通过模拟将台站系统中的合法频率等重要信息同步到其他应用系统的过程,来描述本模型的数据同步是如何实现的,如图3所示。

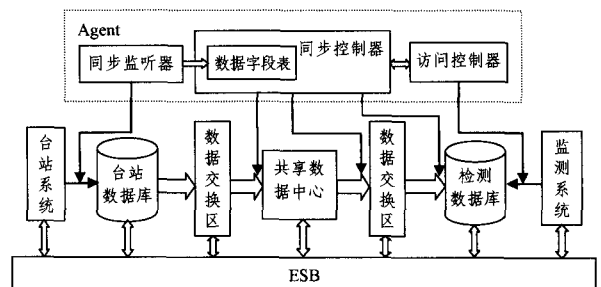


图3 基于 Agent 与 ESB 的共享数据中心同步模型实现过程

数据同步的过程主要分为4个步骤:数据复制、数据清洗、数据分发和数据传送。

(1)数据复制。当台站系统储存数据至数据库时,同步监听器复制其数据,通过 ESB 将数据存储到台站系统的数据交换区中,并修改相应的数据增量表。同步监听器同时更新台站系统以及数据共享中心的数据字段表。

(2)数据清洗。同步控制器通过匹配共享数据中心的数据字段表,将数据交换区中的数据按照数据清洗规则进行清洗,并存储到共享数据中心。

(3)数据分发。同步控制器将共享数据中心增量表中的数据匹配每个应用系统的数据字段表,并整理每个系统需要更新的数据再推送至相应应用系统的数据交换区中。

(4)数据传送。若应用系统没有对需要同步数据的访问请求,则按照同步控制器之前设定好的同步时间间隔将数据交换区中的数据统一同步传送至相应的应用系统数据库。这一过程也由同步控制器完成。如果有对为同步数据的访问请求,则通过访问控制器暂停对数据库的访问,直到数据同步完成。

同步类图如图 4 所示。

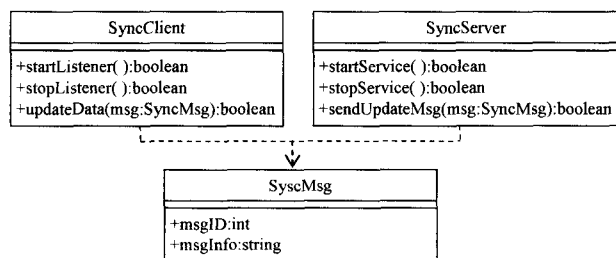


图 4 同步类图

结束语 基于 ESB 与 Agent 的共享数据中心同步模型,可以有效地解决云南省无线电监测中心当前的多系统多数据

库数据同步问题,降低耦合性,便于系统的更新与扩展。访问控制器在控制系统访问数据库的过程中可能造成的延迟对用户的影响,有待进一步研究。

参考文献

- [1] 黄序鑫, 聂瑞华, 罗辉琼, 等. 基于 SOA 的数据同步技术研究与应用[J]. 计算机工程与设计, 2009(14): 3338-3340
- [2] 邵欢庆, 康建初. 企业服务总线的应用[J]. 计算机工程, 2007, 33(2): 220-222
- [3] Gartner. Service-Oriented Architecture Scenario [R]. Technical report. 2003
- [4] 邢少敏, 周伯生. SOA 研究进展[J]. 计算机科学, 2008, 35(9): 13-20
- [5] 帅晓华. 基于 WebService 的应用系统数据同步的实现[J]. 决策与信息, 2008(11)
- [6] 邱宁, 吕何新, 冯淑娟. 基于 Web Service 的数据一致性维护方法[J]. 计算机工程与设计, 2006, 27(14): 2564-2565
- [7] 黄安安, 王丽芳, 蒋泽军, 等. 基于 ESB 的企业应用集成研究[J]. 微计算机应用, 2007, 28(9): 965-970
- [8] 李素萍. 分布式数据库的同步更新技术研究[J]. 中国科技信息, 2008(16): 94-95
- [9] 汪雄杰, 熊前兴. 基于 Mobile Agent 和 Web Service 的移动数据同步模型[J]. 计算机与数字工程, 2010, 38(4): 78-81
- [10] 蒋敏, 董亚波, 鲁东明. 一种基于 Agent 的数据同步方法[J]. 江南大学学报: 自然科学版, 2005, 4(6)
- [11] 张芬琦, 李仰东, 崔丽萍. 浅析 EAI 技术[J]. 山西通信科技, 2004, 25(2): 3-7
- [12] 肖万贤, 刘江宁. 企业数据集成模型的研究[J]. 计算机工程与科学, 2004, 26(5): 49-51

(上接第 374 页)

结束语 本文通过分析比较 Windows 的 GINA(图形标志与身份验证)登录模型和 Credential Provider(凭据提供)登录模型的安全性,采用 Windows 凭据提供模型结合 USB Key 第三方认证方式,设计并实现了基于 USB Key 的 Windows 凭据提供登录系统。这种登录系统提高了传统“用户名+密码”登录方式的安全程度,是符合当今技术发展趋势和时代要求的产物,具有成本低、易携带、专人专用、可靠性高等特点,是未来大数据时代数据安全、人性化使用、跨多应用密码存储的一个重要发展方向,且实现简单,灵活方便,具有良好的应用前景。

参考文献

- [1] 韩君. 基于 USB Key 的 Windows 身份认证与访问控制研究[D]. 武汉: 武汉大学, 2004
- [2] 时向卫, 李峥, 张少武. 基于 Usbkey 的 Windows 登录系统的设计与实现[J]. 计算机工程与设计, 2008, 10: 2483-2485
- [3] 万海山, 李国, 李大兴. 使用 USB KEY 控制 Windows 2000/NT 开机登录[J]. 计算机应用, 2003(S1): 254-255
- [4] Microsoft. MSDN Library[EB/OL]. [2014-4-2]. [http://msdn.](http://msdn.microsoft.com/library/)

microsoft.com/library/

- [5] Michael Muckin, Lockheed Martin. Windows Vista Security Internals [J/OL]. [2014-4-2]. <http://www.blackhat.com/presentations/bh-dc-09/Muckin/BlackHat-DC-09-Muckin-vista-security-internals.pdf>
- [6] 李春林, 王丽芳, 蒋泽军, 等. 基于身份认证技术安全体系的研究[J]. 微电子学与计算机, 2005, 04: 8-11, 14
- [7] 戴平阳. 指纹识别技术研究进展[J]. 厦门大学学报: 自然科学版, 2002, 06: 750-755
- [8] 卢朝阳, 张岗山, 刘琳. 指纹识别系统性能评价方法[J]. 西安电子科技大学学报, 2002, 6: 804-808
- [9] 吴教育, 曾东海. 基于指纹识别的网络身份认证系统[J]. 计算机技术与发展, 2007, 1: 170-171, 176
- [10] 史晓影. 基于动态口令身份认证系统的设计与实现[J]. 办公自动化, 2007, 2: 19-20
- [11] 刘宏, 李锦涛, 崔国勤, 等. 基于 SVM 和纹理的笔迹鉴别方法[J]. 计算机辅助设计与图形学学报, 2003, 12: 1479-1484
- [12] Microsoft. Smart card reference [DB/OL]. [2014-4-2]. <http://www.msdn.microsoft.com/library/>
- [13] 杨帆. USB KEY 体系研究与技术实现[D]. 武汉: 武汉大学, 2004