

# 一个面向电子政务的信息交换 与共享平台的设计与实现\*

莫 赓<sup>1,2</sup>, 钟 华<sup>1</sup>, 魏 峻<sup>1</sup>

(1. 中国科学院 软件研究所 软件工程技术中心, 北京 100080; 2. 中国科学院 研究生院, 北京 100049)

**摘 要:** 基于 J2EE、XML 等技术, 设计并实现了一个面向电子政务的信息交换与共享平台。对平台的总体结构、元数据定义与管理、数据交换模式等关键问题进行了论述, 最后给出了平台的模块实现。该平台已经应用于北京市信息资源管理中心, 实现了遥感信息、区县图层、企业基础信息等资源的交换与共享。

**关键词:** 信息交换; 元数据; 可扩展标记语言; 轻量级目录访问协议

中图分类号: TP393 文献标志码: A 文章编号: 1001-3695(2007)05-0156-04

## Design and Implementation of Information Exchanging and Sharing Platform for E-Government

MO Geng<sup>1,2</sup>, ZHONG Hua<sup>1</sup>, WEI Jun<sup>1</sup>

(1. Technology Center of Software Engineering, Institute of Software, Chinese Academy of Science, Beijing 100080, China; 2. School of Graduate, Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** This paper provided a design and implement of information exchange and share service platform for e-government based on J2EE, XML. It discussed key problems: the system structure of the platform, definition and management of meta-data, data exchange mode; and provided the implementation of platform function module. This platform had been applied in Beijing Information Resource Manage Center successfully, it realized data exchange and share of the following resources: remote sensing information, countries' map, enterprise basic information.

**Key words:** information exchange; meta-data; XML( extensibl markup language ); LDAP

### 0 引言

随着计算机技术发展和网络的普及, 电子政务近几年得到了深入广泛的发展, 各级政府和政府各部门建立了相应的业务系统, 收集了大量的信息资源。为了更好地为人民群众服务, 政府不断地追求更高更好的办事效率与质量。由于很多工作涉及到多个职能部门间的协同配合, 加强部门间的信息交流与共享是进一步提高政府办公效率的一个重要举措, 也是政府信息化建设面临的一个关键问题。

例如, 跨部门协同的并联审批业务和会计核算系统涉及到工商、税务、质监等多个部门。如果实现了信息共享, 就可建立合理的信息化流程, 形成信息化条件下的互连工作机制, 及时交换企业基础信息, 使相关政府管理部门全面、准确、及时地掌握企业基础信息, 提高政府部门的监管能力和执法水平。

不同应用系统之间的信息交换与共享的难度很大, 这主要是由于应用系统和信息资源的异构性造成的, 包括硬件平台、操作系统、数据库的异构性, 也包括信息资源在语法、语义方面的异构性, 还包括各职能部门在业务特点、工作方式上的差

异等。

为了使政府各部门之间方便地进行信息共享, 本文提出了一个基于 J2EE 的信息交换与共享平台。该平台作为政府部门之间信息共享的基础, 提供标准的接口, 可以使不同部门的应用系统之间能够相互交换信息。同时, 由于整合了各种信息资源, 还能够在此平台上开发新的业务系统, 如应急指挥、领导决策等, 以更好地发挥政府职能。

### 1 平台的体系结构设计

政府各职能部门通过信息交换共享平台进行业务数据资源的交换与共享。一方面各部门可向信息交换平台发布交换资源, 其他部门通过查询后获取已发布的资源; 另一方面, 部门之间通过信息交换平台直接交换资源。所有资源的发布、交换均由平台统一管理。在交换过程中, 某些信息可以落地, 即进入平台的中央数据库中。在平台的中央数据库的基础上可以进行信息的整合, 比如遥感信息、区县图层、企业基础信息等。这些落地的信息可以根据实际情况从无到有, 不断积累, 不断更新。这些资源对政务中新业务的开展是非常有用的。平台

收稿日期: 2006-03-10; 修返日期: 2006-05-27 基金项目: 国家“863”计划资助项目(2004AA112010); 国家“973”计划资助项目(2002CB312005)

作者简介: 莫赓(1980-), 硕士研究生, 主要研究方向为网络分布式计算和软件工程(mogeng@otcaix.iscas.ac.cn); 钟华(1971-), 硕士, 主要研究方向为网络分布式计算和软件工程; 魏峻(1970-), 硕士, 主要研究方向为中间件技术与基于中间件的软件开发方法。

的业务流程如图 1 所示。

信息交换共享平台主要由元数据管理、交换服务、目录服务、安全服务和平台管理等模块组成。平台的组成结构如图 2 所示。下面对元数据管理、交换服务、目录与安全服务进行更细致的描述, 而平台管理模块负责各模块的配置, 协调模块之间的交互, 管理平台日志, 监控整个平台运行等。基于平台的这些支撑模块, 就可以整合已有的业务系统和开发新的业务系统, 满足政务信息交换与共享的需要。

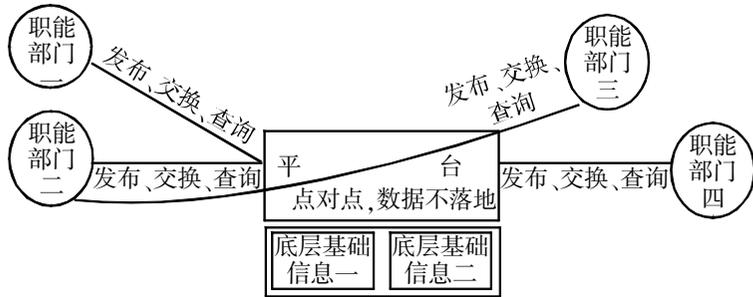


图 1 信息交换与共享的业务流程

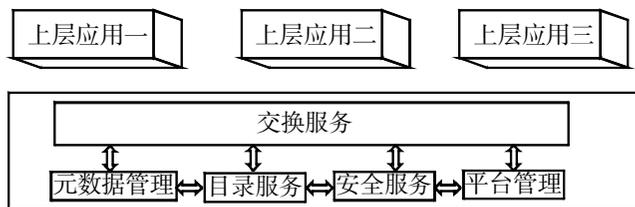


图 2 平台模块结构图

### 1.1 元数据定义与管理

元数据是对信息资源的结构化描述, 其作用为描述信息资源或数据本身的特征和属性, 就像图书馆里的书和书卡一样。书是具体的数据, 书卡是元数据, 每本书均有一张书卡与它对应, 上面描述了书的名称、作者、出版社、分类等相关信息。

平台基于元数据来存储各个职能部门的信息资源和相关其他信息。一般来说, 不同的职能部门有自己管理资源的系统和方式, 多是基于关系型数据库, 使用大量的表来存储数据。表之间的关系错综复杂, 难以管理, 可扩展性不强。在不同的职能部门之间交换数据时, 格式的统一也是一个问题。本文采用元数据来描述平台中涉及到的所有数据。元数据管理就成为平台管理的基础, 也是实现平台其他核心模块的基础。

平台中元数据的 Schema 分为两部分: 核心元数据, 采用的是都柏林的元数据的 Schema, 这部分描述的是一些资源的普通信息, 由平台管理者维护。行业元数据。这部分在不同的行业就会有不同的 Schema, 在该系统中这部分能进行动态改变, 即允许用户注册新的行业元数据 Schema, 也可以修改已注册的行业元数据 Schema。这部分信息根据行业的不同而差异很大, 由资源提供者自行维护。元数据管理模块通过该 Schema 对整个系统的元数据进行分级管理。

一个电子政务元数据的 Schema 如下:

```

schema targetNamespace = "http://localhost/share.xsd"
xmlns: me = "http://localhost/share.xsd"
xmlns = "http://www.w3.org/2001/XMLSchema"
elementFormDefault = "qualified"
complexType name = "MetaType"
sequence
  element name = "resTitle" type = "string" /
  element name = "PubDate" type = "date" /

```

```

  element name = "abstract" type = "string" /
  element name = "descKeys" type = "me:resKeys" maxOccurs = "unbounded" /
  element name = "spatDom" type = "string" minOccurs = "0" /
  element name = "Period" type = "me:resPeriod" minOccurs = "0" /
  element name = "dataLang" type = "string" maxOccurs = "unbounded" /
  element name = "rescat" type = "me:resCategory" maxOccurs = "unbounded" /
/sequence
/complexType
complexType name = "resKeys"
sequence
  element name = "keyword" type = "string" maxOccurs = "unbounded" /
  element name = "thesaName" type = "string" minOccurs = "0" /
/sequence
/complexType
complexType name = "resPeriod"
sequence
  element name = "begDate" type = "date" /
  element name = "endDate" type = "date" /
/sequence
/complexType
complexType name = "resCategory"
sequence
  element name = "catecode" type = "string" /
  element name = "catestid" type = "string" /
/sequence
/complexType
element name = "Meta" type = "me:MetaType" /
/schema

```

除了元数据的定义之外, 元数据管理的具体功能包括元数据的采集、检索、发布, 以及元数据的增、删、改等。元数据的采集有两种方式: 一种是根据用户指定的网站进行信息采集, 搜索指定栏目, 依据预定的标准生成相应的元数据并入库; 另一种是根据关系数据库中的字段生成相应的元数据。元数据的检索主要包括关键字检索和组合检索。元数据发布可以让用户按照某标准发布元数据, 平台根据分类将该元数据入库。

### 1.2 数据交换

整个系统的数据流在数据交换模块内实现。通过对政务业务的分析, 平台上的数据交换模式主要有三大类, 即点到点、点到多点和共享。这里的共享指的是各职能部门把自己的数据上传到中央数据库中, 其他职能部门可以从中央数据库中获取需要的信息。为了实现这三类交换, 笔者提供了三种主要的交换模式:

#### (1) 基于消息/事件驱动的数据交换模式

基于消息/事件驱动的方式进行数据交换时可以按照如下描述进行: 运行于职能部门 A 节点上的数据库接口适配器 DBOutAdapter 感知出所监控的数据表格发生了数据增量(插入新的数据), 增量数据被抽取出来并转换成 XML 格式文件, 经内容路由组件分发到不同职能部门, 由数据转换组件(XSLT Transformer)转换成不同格式的 XML 文件, 并插入位于职能部

部门 B 或 C 中的异构数据库中。发生数据更改或删除时,数据同步原理一样。其交换模式示意图如图 3 所示。

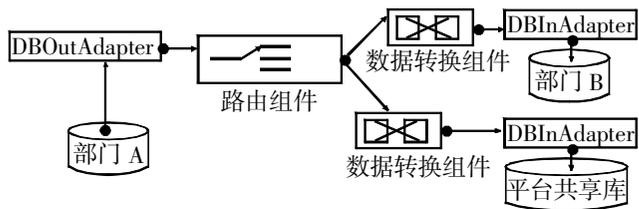


图 3 基于消息/事件驱动的交流模式示意图

(2) 基于请求/响应的交换模式

基于请求/响应的方式进行数据交换时可以按照如下描述进行:各职能部门之间要进行数据访问时,发送部门的应用程序(如目录服务)通过使用 API 接口函数将数据打包后发送到数据交换平台,通过平台路由和数据格式转换,转换为下一个部门可以识别和接收的形式,运行于接收部门节点服务器上的接口服务组件(Adapter)从数据库中获得数据并返回。

(3) 发布/订阅的交换模式

发布/订阅的交换模式的流程是:职能部门可以把自己的数据信息以发布的方式发布到平台,其他职能部门再通过订阅的方式获取平台上发布的信息。其交换模式如图 4 所示。

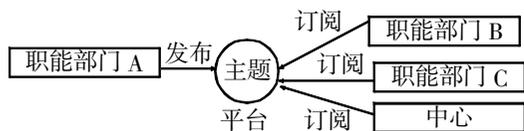


图 4 发布/订阅交换模式示意图

1.3 目录与安全服务

目录服务主要包括管理数据信息、安全信息、Web 服务信息等主要内容。其中数据信息主要指通过元数据描述的数据信息。笔者将元数据通过树状形式表现出来,使得浏览、查找更有组织,层次清晰。每个节点均有自己的属性。这里的属性是结构化的数据,能够存储用户、操作权限等相关信息。用户可以通过节点的属性进行查找,效率极高。而安全信息是用户、权限等信息,安全服务基于 RBAC,通过安全信息的设置给角色分配权限,给用户分配角色。

安全服务中的权限具有多样性,不仅包括平台上的具体操作,还包括资源的可见性、元数据指标的可见性和数据有关的权限信息。通过用户、角色、权限三者形成三级管理方式,使得安全部分伸缩性很大,易于扩展和管理。Web 服务的管理使平台用户可以注册自己的 Web 服务,也可以在平台上查找已注册的 Web 服务。

目录服务中不同用户的资源目录树是可以定制、修改的,具有很好的个性化。每个目录分支的信息内容、状态是动态更新的。用户可以定制自己关心的信息,而不是大而全地包括系统所有可用信息的目录树。用户面对自己的资源目录树时能一目了然地知道哪些分支有更新,就能快速地掌握新的数据。

2 平台实现

在平台的实现中,元数据部分采用 XML 格式进行存储,而不是一般的关系数据库中的关系表。元数据本身有一些复杂的数据结构,如结构体等。如果用关系表实现的话,结构复杂一点的 Schema 就会要求使用多个表来表示,表之间用外键关

联。但是如果 Schema 发生改变后,有可能原来的表就不能沿用,可扩展性不是很好。而 XML 本身是一种完备的结构化数据。XML Schema 是用来描述一类 XML 文件的 XML 文件。XML Schema 定义了 XML 的文档结构,与 DTD 类似。但是由于 XML Schema 支持数据结构和命名空间等多种功能,故比 DTD 功能更强大,可扩展性更强。

在数据库方面笔者采用了 Oracle 9i 数据库的 XDB。首先,根据元数据的 Schema 生成相应的 XML Schema,并注册到 XML 数据库中,就能依据这个 XML Schema 来组织元数据。每个元数据均作为一个 XML 文件进行处理。在插入数据库时用 XML schema 进行验证其合法性。由于每一条元数据均是一个 XML 文档,元数据的查询不能像查询一般的关系表那样做。在进行元数据查询时,笔者采用了 XPath 语言。XPath 是用于在 XML 文档中导航元素和属性的语言。XPath 用路径表达式在 XML 文档中导航。这种语言本身包含一个标准的内建函数库。XQuery 和 XPointer 均是基于 XPath 表达式基础之上的。平台中的业务逻辑部分包括 ControlCentre、RequestHandler、MBean、AddHandler、SearchHandler、EditHandler、ShowHandler 七个类。

(1) ControlCentre 是一个 HttpServlet 的子类,用于接收用户请求并转发到相应的业务逻辑处理类;

(2) RequestHandler 是一个接口,它声明了每个业务逻辑处理的通用方法;

(3) MBean 类记录元数据的信息,与存储元数据的 XML 形成一个对应关系;

(4) AddHandler 类实现了 RequestHandler 接口,用于完成元数据的添加功能;

(5) SearchHandler 实现了 RequestHandler 接口,用于完成元数据的查询功能;

(6) EditHandler 类实现了 RequestHandler 接口,用于完成元数据的编辑功能;

(7) ShowHandler 类实现了 RequestHandler 接口,用于完成元数据的显示功能。

扩展性方面,当元数据的 Schema 发生改变时,只需改变相应的 XML Schema 即可。政府职能部门在批量向平台导入元数据时,采用 XML 格式显然非常不方便。本文实现了 XML 与 Excel 的互相转换模块。一般的职能部门大多喜欢使用 Excel 整理数据。在用户导入数据前根据其 Schema 生成相应的 Excel 模板供用户使用。这样用户提供给平台的实际上是一个 Excel 文件。平台在检测数据的合法性之后再将其转换成相应的 XML 文档。

交换服务结构比较复杂,包括消息服务器、业务流程管理器、部署管理器、服务组合器、服务和安全管理器,以及种类繁多的适配器等,并实现了上面提到的三种交换模式。交换服务提供了数据描述、传输、安全加/解密、转换、汇总、分发、转发、对等交换、同步、上传/下载、组装等功能服务,实现数据的高效传输,并可防止数据的丢失、重传,简化应用系统的开发。交换服务为相关的应用提供统一的数据共享途径,为每一个参与数据共享的应用提供访问其所需数据的接口。由于交换体系隔

离了数据传输与数据访问和数据处理,提高了开放平台信息资源的安全性。

在整个交换服务中,与应用紧密相关的功能实现类主要集中在文件交换资源子模块。此模块主要是实现文件资源描述信息的获取,文件到本地发送目录的拷贝,发送完毕后反馈信息的获取和显示。最主要的类有 FileInfoGet、FileCopy、FeedBackProcess、PropertySetting、MessageSender、InteractionSetting、FileReceiving、RequestSender、RequestVerification、RequestTransmitting、RequestRuleSetting 等。

(1) FileInfoGet 负责获取并存储用户在发布文件时所填写的有效参数信息。

(2) PropertySetting 负责获取用户配置组件服务参数的接口,通过此接口完成对 FioranoESB 所提供服务的配置。

(3) FileCopy 负责实现将源文件传送至发送目录,由 FioranoESB 提供的文件传输组件来完成到达目的位置的传送。

(4) InteractionSetting 完成自动传输方式时,对服务所提供的时间规则的设置。

(5) FileReceiving 完成文件到达时向接收方显示提示信息,内容包括整个资源的描述信息,用户可以通过选择接收和拒绝来完成文件的保存和丢弃。

(6) MessageSender 完成向服务器提交信息的功能,信息包括服务配置信息和业务流程相关信息。

(7) FeedBackProcess 负责将传输完毕后的信息提取出来,根据信息类型加以区分,然后以一定的格式显示给用户。

(8) RequestSender 实现请求/响应方式中请求信息的生成和发送。

(9) RequestVerification 用于资源拥有方的请求格式与获取的请求信息对照,判定是否给予传送。

(10) RequestTransmitting 对不符合请求验证的请求信息发送至服务器存储。

(11) RequestRuleSetting 实现用户对自动响应请求的服务设置响应规则。

交换服务支持的数据类型包括关系型、结构化、XML、消息、文本数据、Web 数据、矢量地图、影像、视频、音频等。传输过程中均将其按照相应规则转换成 XML 格式的数据,提高了异构平台系统间的数据兼容性。对关系数据库而言,能够通过简单设置来完成不同数据库、不同表、不同字段之间的映射转换,而且能够实现在映射过程中对映射结果进行多种运算操作的功能,从而解决了异类数据库之间的数据转换、整合的问题。

目录服务基于 LDAP 标准实现,为系统管理员及各级用户提供了标准的目录服务,便于实现系统的升级改造,降低了系统集成的成本,具有良好的可伸缩性和可扩展性。该目录服务支持四类 10 种操作:查询类操作,包括搜索、比较;更新类操作,包括添加条目、删除条目、修改条目;认证类操作,包括绑定、解绑定;其他操作,如放弃和扩展操作。用户管理子模块完成用户信息的管理(主要是针对用户权限的管理),实现用户安全认证、访问控制等功能。身份认证支持三种认证机制,即匿名、基本认证和 SASL(Simple Authentication and Secure Layer) 认证。

用户管理实现系统用户的添加、删除、修改、查询等功能。用户信息包括用户名(登录账号)、密码、用户单位和用户角色。系统根据用户所属的单位和用户的角色来综合评定用户使用系统资源的权限。其中,用户角色分为系统管理员、局管理员、领导、发布者、使用者,以及公务员。系统管理员具有最高权限,能够对整个系统平台的用户和资源进行管理;局管理员只能对本局的用户和资源进行管理,从而实现系统的二级权限管理。该模块主要涉及的类有 AddUser、QueryU、DeleteU、ModifyU、PowerTestImp、GetSysInfo 等。

### 3 结束语

2005 年 5 月开始,该平台在北京市信息资源中心进行了初步应用,实现了北京市电子政务信息交换共享平台,主要涉及遥感信息、区县图层、企业基础信息等资源。这些信息分布在不同的服务器上,存储的数据库类型也不尽相同。信息交换与共享平台较好地将这些资源整合起来。2005 年 11 月,平台成功地在北京市石景山区进行了部署,使用情况反馈良好。

基于信息交换共享平台还可以开发出一些新的政务应用,如领导决策系统、应急指挥系统等,并作为政府门户的支撑系统。例如,应急指挥即在出现重大疫情、灾情时需要各个部门紧密配合与支持,齐心协力才能处理好突发事件。这样必然要求各个部门之间信息共享,根据各部门的情况,把握全局,迅捷作出反应。

本文设计并实现了一个面向电子政务的信息交换与共享平台,探讨了平台的总体结构、元数据定义与管理、数据交换模式等关键问题,并给出了平台的模块实现。平台中的一些关键技术问题还需要进一步的研究、探讨,如数据冲突处理与一致性检查、海量数据的存储与检索、多部门间的数据流转等。

#### 参考文献:

- [1] ALEVY A, IVES Z, SUCIU D, *et al.* Schema mediation in peer data management systems: proc. of the 19th International Conference on Data Engineering (ICDE 2003) [C]. [S. l.]: [s. n.], 2003: 505-516.
- [2] HAAS L M, LIN E T, ROTH M A. Data integration through database federation[J]. **IBM Systems Journal**, 2002, 41(4): 578-596.
- [3] FAGIN R, KOLAITIS P G, POPA L. Data exchange: getting to the core: proc. of the 22nd ACM SIGACT SIGMOD SIGART Symp. on Principles of Database Systems [C]. [S. l.]: [s. n.], 2003: 90-101.
- [4] 史周军,叶晓俊.基于元数据的对象关系映射研究[J]. **计算机科学**, 2005(5): 95-97.
- [5] 李志义.谈电子政务中基于 XML 的元数据[J]. **电子政务**, 2005(7): 34-38.
- [6] 李少春,蒋泽军. J2EE 环境下基于 LDAP 的访问控制的设计与集成[J]. **微电子学与计算机**, 2005(3): 118-120, 124.
- [7] 李澜. LDAP 目录服务安全及研究现状[J]. **微电子学与计算机**, 2005(6): 48-52, 56.
- [8] 陈祖龙. ESB 奠定整合基础[J]. **软件世界**, 2005(7): 48-48.
- [9] 张繁,蔡家楣.电子政务系统中的数据交换和共享服务平台设计[J]. **计算机工程与应用**, 2003(7): 226-229.