

- 版权注意事项：1、书籍版权归著者和出版社所有；
- 2、本PDF仅用于个人获取知识，进行私底下知识交流；
- 3、PDF获得者不得在互联网以任何目的进行传播；
- 如有需要，请尽量购买正版实体书！支持书籍作者！！

Web服务与 数据交换关键技术 研究

WEB FUWU YU SHUJU JIAOHUAN
GUANJIAN JISHU YANJIU

马晓轩 / 著



中国环境出版社

Web 服务与数据交换 关键技术研究

马晓轩 著

中国环境出版社·北京

图书在版编目(CIP)数据

Web 服务与数据交换关键技术研究/马晓轩著. —北京:
中国环境出版社, 2017.3
ISBN 978-7-5111-2839-3

I. ①W… II. ①马… III. ①Web 服务器—研究
②数据交换—研究 IV. ①TP393.092.1②TN919.6

中国版本图书馆 CIP 数据核字 (2017) 第 040145 号

出版人 王新程
责任编辑 侯华华
责任校对 尹芳
封面设计 宋瑞



更多信息, 请关注
中国环境出版社
第一分社

出版发行 中国环境出版社

(100062 北京市东城区广渠门内大街 16 号)

网 址: <http://www.cesp.com.cn>

电子邮箱: bjgl@cesp.com.cn

联系电话: 010-67112765 (编辑管理部)

联系电话: 010-67112735 (第一分社)

发行热线: 010-67125803, 010-67113405 (传真)

印 刷 北京中献拓方科技发展有限公司
经 销 各地新华书店
版 次 2017 年 3 月第 1 版
印 次 2017 年 3 月第 1 次印刷
开 本 880×1230 1/32
印 张 5
字 数 115 千字
定 价 26.00 元

【版权所有。未经许可请勿翻印、转载, 侵权必究】

如有缺页、破损、倒装等印装质量问题, 请寄回本社更换

前 言

传统客户/服务器的分布式计算模式主要解决的是部门或企业内部的应用，关注的是局部应用，建立的应用系统多数是为完成某种需求而建立，因此是孤立封闭的。即使在同一个部门或者企业内部，应用系统相互之间也常常因为没有通信而导致信息无法共享，使得各个 IT 资源成为一个个信息孤岛。随着互联网软件技术及其应用的迅速发展，如何提供一个统一开放的交互环境，使得各个应用实体之间能够相互发现、了解各自所提供的服务，并将这些应用低代价、方便地连接在一起，实现开放式网络环境中的互联、互通、互操作的目标，是当前互联网环境下所面临的一个重要问题。

想要有效地整合现有的应用系统并实现信息共享，关键在于实现这些应用系统间的数据交换。数据交换是网络环境下分布式应用的共性、基础性和关键性的需求，它用于解决信息化过程中数据资源的互操作问题，即解决不同的异构系统之间的

数据资源的整合和共享。Web 服务的目的和作用提供一种国际统一的规范和技术,进行 Internet 上各种软件应用的统一功能描述和功能共享,为功能整合集成和信息交换处理提供实现基础。本书分析了大规模数据交换在 Web 应用中所面临的问题,针对数据交换平台框架、数据交换的可靠性、数据交换的效率以及数据服务中心等一系列关键技术展开研究,主要的研究结果如下:

(1) 设计了基于 Web 服务的数据交换框架。分析了使用 Web 服务进行数据交换的必要性以及数据交换的应用模式。提出基于 Web 服务的数据交换平台框架,并由此提出基于 Web 服务的大规模数据交换的若干关键技术。

(2) 提出了基于 Web 服务的数据交换可靠消息机制。通过分析数据交换的可靠性因素,从数据交换消息层定义了消息可靠性,讨论了消息可靠性的参数,提出一种可靠的消息传输机制,主要从消息确认、消息重复、消息丢失三个方面来确保数据交换的可靠性,并从可靠消息处理机和可靠报文结构两个方面阐述了该机制的实现。

(3) 提出了两种改进基于 Web 服务的数据交换传输效率的机制。首先分析了影响 Web 服务性能的因素,提出了 Web 服务性能的优化方案,在此基础上提出了基于 Web 服务的类

型映射及基于滑动窗口的 SOAP(Simple object access protocol, 简单对象访问协议) 消息传输等两种改进数据交换的传输效率的机制。针对 Web 服务的性能优化需求, 基于 WSDL (Web services description language, Web 服务描述语言) 文档校验机制, 提出了一种可扩展的高效类型映射机制; 基于滑动窗口的 SOAP 消息传输机制通过将数据切片, 实现对大数据和可变长度数据的传输优化, 并通过设计滑动窗口协议来进一步提高 SOAP 消息的传输速度; 提出了基于应用层的拥塞控制算法, 并对影响传输效率的两个主要因素(滑动窗口大小和数据切片长度)进行了测试。

(4) 设计了基于分布式信息存储模型的数据服务中心。在对数据交换中数据服务的可扩展性、发现效率和有效性进行分析的基础上, 设计了一个分布式信息存储模型, 讨论了注册中心的组织, 提出了信息动态注册与更新机制, 设计了一个分布式信息搜索算法, 并设计实现了基于 UDDI (Universal description, discovery, and integration, 通用描述发现和集成) 的数据服务注册中心 DSC。针对数据服务中心的安全及管理等问题, 对原有的 UDDI 协议进行了一些扩展, 提出了发布者声明关联匹配算法, 发布者声明关联匹配算法能够描述注册应用之间的关联关系。

(5) 设计并实现了基于 Web 服务的数据交换平台。提出了一个数据交换协议,在此基础上,设计实现了基于 Web 服务的数据交换平台,讨论了其系统的总体结构,阐述了其中的关键机制如事务、适配器等,并且讨论了数据交换平台在电子政务中的应用部署。

目 录

第 1 章 概 论	1
1.1 异构系统数据交换面临的挑战	1
1.2 相关技术	3
1.2.1 数据交换技术的发展过程	3
1.2.2 Web Services 技术	5
1.2.3 消息中间件	10
1.2.4 国内外的数据交换平台产品现状	12
1.3 数据交换关键技术	16
第 2 章 基于 Web 服务的数据交换平台框架	19
2.1 数据交换与 Web 服务	19
2.1.1 数据交换概述	19
2.1.2 Web 服务分析	22
2.1.3 基于 Web 服务进行数据交换的必要性	24
2.2 数据交换应用模式分析	25
2.2.1 部署结构	26
2.2.2 信息总线	28
2.3 基于 Web 服务的数据交换平台框架	29

2.3.1	总体设计	29
2.3.2	数据交换功能	31
2.4	本章小结	32
第 3 章	数据交换可靠性研究	33
3.1	消息传输的可靠性分析	33
3.1.1	传统的 TCP 协议的可靠性	33
3.1.2	JMS 消息队列可靠性	34
3.1.3	Web 服务可靠性	35
3.2	数据交换的可靠性	40
3.2.1	可靠消息定义	40
3.2.2	可靠性参数	41
3.2.3	可靠性因素分析	42
3.3	数据交换可靠消息机制	44
3.3.1	可靠消息机制目标	44
3.3.2	数据交换可靠消息机制的设计	45
3.3.3	数据交换可靠消息机制的实现	52
3.4	本章小结	55
第 4 章	数据交换效率研究	56
4.1	Web 服务传输性能分析	56
4.1.1	Web 服务性能现状分析	58
4.1.2	影响 Web 服务性能的因素	60
4.1.3	Web 服务性能优化方案	64
4.2	基于 Web 服务的类型映射机制	68

4.2.1	Web 服务与类型映射	69
4.2.2	基于 Web 服务的类型映射机制	71
4.2.3	WSTM 系统的设计与实现	75
4.2.4	实验结果及分析	79
4.3	基于滑动窗口的 SOAP 消息传输机制	82
4.3.1	SOAP 滑动窗口工作原理	82
4.3.2	滑动窗口消息结构	83
4.3.3	滑动窗口工作流程	84
4.3.4	应用层拥塞控制算法	85
4.3.5	影响 SOAP-SW 传输效率的主要因素	87
4.4	本章小结	89
第 5 章	基于分布式信息存储模型的数据服务中心	91
5.1	数据服务分析	91
5.2	分布式信息资源存储模型	93
5.3	注册中心的组织	95
5.3.1	注册中心的加入	96
5.3.2	注册中心的退出	96
5.4	信息资源动态注册与更新	97
5.5	分布式信息资源搜索算法	98
5.6	DSC 的设计与实现	102
5.6.1	通用描述发现集成协议	103
5.6.2	系统的总体结构设计	106
5.6.3	DSC 客户端	108
5.6.4	DSC 服务器	109

5.6.5 DSC 管理控制台	113
5.7 本章小结	115
第 6 章 数据交换平台的设计与实现	116
6.1 引言	116
6.2 数据交换协议	118
6.2.1 问题分析	118
6.2.2 假设和说明	119
6.2.3 词汇和消息	119
6.2.4 原语和过程规则	120
6.3 WS-XP 的设计和实现	121
6.3.1 系统的总体结构设计	121
6.3.2 数据交换引擎	123
6.3.3 适配器机制	128
6.3.4 传输模块	131
6.3.5 事务机制	131
6.4 WS-XP 的应用部署	132
6.5 本章小结	136
参考文献	137

第 1 章

概 论

1.1 异构系统数据交换面临的挑战

随着计算技术，尤其是近年来网络技术和应用的不断发展，计算模式也在不断发生变化，从早期的单一主机、主机/终端模式演化到以构件技术和分布式对象技术为代表的客户/服务器模式。与此同时，大量的业务应用也开始向互联网迁移。由于传统客户/服务器的分布式计算模式主要解决的是部门或企业内部的应用，关注的是局部应用，建立的应用系统多数是为完成某种需求而建立，因此是孤立的、封闭的，即使在同一个部门或者企业内部，应用系统相互之间也常常因为没有通信导致信息无法共享，使得各个 IT 资源成为一个个信息孤岛，相互间壁垒森严，最终效率低下。随着社会信息化程度的不断提高，信息孤岛的制约作用越来越突出，导致信息技术不能更好地满足社会发展的需求。随着市场经济的发展，政府部门面临转变政府职能、提高行政能力、服务社会公众的挑战；而企业用户也随着业务规模的扩大和产品质量的不断提升，迫切需要提高业务运作的效率、快速用户响应，以提高用户的满意度。为适应这

些新形势的发展,迫切需要打破现有应用系统之间的信息隔离,整合和共享已有的信息资源,以提高应用系统间的信息共享和协同工作。如何提供一个统一开放的交互环境,使各个应用实体之间能够相互发现、了解各自所提供的服务,并将这些应用低代价、方便地连接在一起,实现开放式网络环境中的互联、互通、互操作的目标,是当前互联网环境下所面临的一大问题。

为有效地整合现有的应用系统并进行信息共享,关键在于解决这些应用系统间的数据交换。数据交换是网络环境下的分布式应用的共性、基础性和关键性的需求。数据交换用于解决信息化过程中数据资源的问题,即解决不同的异构系统之间的数据资源的整合和共享。

在大规模数据交换的过程中,数据交换双方可能处于同一个局域网内,也可能处于广域网上相距甚远的两个网络节点上,交换过程可能跨越多个网络自治域,不同自治域在网络状况、安全保护措施等方面都可能存在不同程度的差异;数据交换双方要交换的数据可能是文本类型的数据,如文本文件、XML (eXtensible Markup Language, 可扩展标记语言) 文件,也可能是图片、音频、视频、Word 文档等其他类型的文件,交换的数据可能小到只有几个字节,也可能是大到数十兆甚至上百兆数据量的多媒体数据文件。因此,由于计算机网络技术和应用的多样性,异构问题已成为当前应用系统集成的主要障碍,如何去解决异构系统间的集成以实现系统间的数据交换是资源整合的一个关键问题。

Web 服务的目的和作用是一种国际统一的规范和技术,进行 Internet 上各种软件应用的统一功能描述和功能共享,为功能整合集成和信息交换处理提供实现基础。通过使用 Web 服务技术,所有

的应用数据的封装都是标准化的，都是基于 SOAP 报文格式进行封装，并且使用 SOAP 消息进行传输；每个应用程序的接口描述都是使用 WSDL 来进行描述，公用数据信息使用通用描述发现集成协议来进行存储，从而使应用之间的维护、安全控制、信息过滤都变得非常容易。它不需要对遗留系统进行大的修改或重新开发。由于 Web 服务技术的出现，计算机系统之间实现统一的信息交换成为可能。

因此，基于 Web 服务技术，分析大规模数据交换所面临的问题，研究数据交换平台框架，并解决数据交换的可靠性、数据交换的效率以及数据的存储及共享等一系列关键技术，将具有重要的学术意义和应用价值。

1.2 相关技术

1.2.1 数据交换技术的发展过程

数据交换技术到目前为止大致经历了三个发展阶段：

- 第一个阶段，以电子数据交换（Electronic Data Interchange, EDI）技术为基础的数据交换。

EDI 是计算机、通信和现代管理技术相结合的产物。国际标准化组织（International Organization for Standardization, ISO）将 EDI 描述成“将贸易（商业）或行政事务处理按照一个公认的标准变成结构化的事务处理或信息数据格式，从计算机到计算机的电子传输”。而 ITU-T 将 EDI 定义为“从计算机到计算机之间的结构化的事务数据互换”。由于可以减少交易过程中的纸张使用量，EDI 被人们通俗地称为“无纸贸易”。

从 EDI 定义可以得出, EDI 包含了三个方面的内容, 即计算机应用、通信网络和数据标准化。其中, 标准化的工作是实现 EDI 互通和互联的前提和基础。目前, 在 EDI 标准上, 国际上最有名的是联合国欧洲经济委员会 (UN/ECE) 下属第四工作组 (WP4) 于 1986 年制定的《用于行政管理、商业和运输的电子数据互换》标准——EDIFACT (Electronic Data Interchange For Administration, Commerce and Transport) 标准。

EDIFACT 标准已被国际标准化组织 (ISO) 接收为国际标准。同时还有广泛应用于北美地区的, 由美国国家标准化协会 (ANSI) X.12 鉴定委员会 (AXCS.12) 于 1985 年制定的 ANSI X.12 标准。从目前的情况来看, EDIFACT 成为统一的 EDI 国际标准已是大势所趋。

- 第二个阶段, 以中间件技术为基础的数据交换。

中间件是基础软件的一大类, 属于可复用软件的范畴。中间件运行在操作系统、网络和数据库之上, 应用软件之下, 为处于自己上层的应用软件提供运行与开发的环境, 帮助用户灵活、高效地开发和集成日趋复杂的应用软件。中间件产品根据其作用可以分为远程过程调用、面向消息的中间件、对象请求代理和事务处理监控等。

目前用于开发中间件的主要技术有 OMG 的 CORBA (Common Object Request Broker Architecture, 公共对象请求代理体系结构), Sun 的 J2EE 和 Microsoft 的 DCOM (Distributed Component Object Model, 分布式组件对象模型)。

在中间件的使用过程中, 数据与传输紧密地耦合在一起, 不便数据的重用。同时, 由于不同的中间件产品采用不同的规范, 使得不同中间件之间的互操作就成了一个新的集成难题。要想解决这个问题, 就必须采用统一的开放标准, 使各个中间件产品依照这个

标准提供标准的接口并依据标准对接口进行描述,以方便不同中间件产品之间的相互调用。基于 XML 和 Web Services 数据交换技术可以解决这些问题。

- 第三个阶段,基于 XML 和 Web Services 技术的数据交换。

XML, Web Services 技术的出现,为人们寻找一种廉价、简单、有效的信息交换方式提供了技术基础。其中 XML 数据规范为运行于企业内部网络中不同节点的应用系统间进行数据交换奠定了基础; Web Services 又使这些应用系统相互连接并进行功能调用成为可能。

Web Services 采用 SOAP 协议可以基于 HTTP (Hyper Text Transfer Protocol, 超文本传输协议) 协议通信,由于 HTTP 在网络中广泛应用,为 Web Services 提供无限的互联成为可能,从而可以轻松地穿越绝大多数的防火墙,而与 XML 结合使其具有了数据交换的能力。采用基于 XML, Web Services 技术实现跨越网络的异构数据交换,也就成了理想的数据交换方式,使建设跨网络的企业应用集成成为可能。为了方便 Web 服务被描述和发现,还产生了 WSDL 和 UDDI,通过服务发现机制发现了一个 Web 服务之后就可以得到它的服务描述,有了这个描述就可以将这个功能集成到自己的应用系统中,而不用关心 Web 服务是如何构建的或是运行在什么平台之上。基于 Web Services 技术和 XML 数据结构的数据交换方法是当前数据交换平台最主要的研究方向。

1.2.2 Web Services 技术

消息服务为电子政务应用系统实现各种格式的信息从一个地方传递到另一个地方的功能服务。消息服务是一个软件,它以一种可靠的、异步的、松散耦合的、与语言无关的、与平台无关的方式在