

Web 服务核心支撑技术:研究综述*

岳 昆⁺, 王晓玲, 周傲英

(复旦大学 计算机科学与工程系,上海 200433)

Underlying Techniques for Web Services: A Survey

YUE Kun⁺, WANG Xiao-Ling, ZHOU Ao-Ying

(Department of Computer Science and Engineering, Fudan University, Shanghai 200433, China)

+ Corresponding author: Phn: +86-21-65643503, E-mail: kuny@fudan.edu.cn, <http://www.cs.fudan.edu.cn/wpl>

Received 2003-07-11; Accepted 2003-11-18

Yue K, Wang XL, Zhou AY. Underlying techniques for Web services: A survey. *Journal of Software*, 2004,15(3):428-442.

<http://www.jos.org.cn/1000-9825/15/428.htm>

Abstract: With the rapid development of e-business, web applications based on the Web are developed from localization to globalization, from B2C(business-to-customer) to B2B(business-to-business), from centralized fashion to decentralized fashion. Web service is a new application model for decentralized computing, and it is also an effective mechanism for the data and service integration on the web. Thus, web service has become a solution to e-business. It is important and necessary to carry out the research on the new architecture of web services, on the combinations with other good techniques, and on the integration of services. In this paper, a survey presents on various aspects of the research of web services from the basic concepts to the principal research problems and the underlying techniques, including data integration in web services, web service composition, semantic web service, web service discovery, web service security, the solution to web services in the P2P (Peer-to-Peer) computing environment, and the grid service, etc. This paper also presents a summary of the current art of the state of these techniques, a discussion on the future research topics, and the challenges of the web services.

Key words: Web service; service composition; semantic Web; service discovery; security; P2P; grid

摘要: 随着电子商务的迅速崛起,基于 Web 的应用模式迅速发展,Web 应用从局部化发展到全球化,从 B2C(business-to-customer)发展到 B2B(business-to-business),从集中式发展到分布式,Web 服务成为电子商务的有效解决方案.Web 服务是一个崭新的分布式计算模型,是 Web 上数据和信息集成的有效机制.Web 服务的新型构架,Web 服务的高效执行方式,Web 服务与其他成熟技术的有机结合以及 Web 服务的集成是解决现实应用问题的重要技术.从 Web 服务研究的不同侧面对其进行了综述,阐述了 Web 服务的基本概念,分析了当前 Web 服务的主要研究问题

* Supported by the National Natural Science Foundation of China under Grant No.60228006 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2002AA116020 (国家高技术研究发展计划(863)); the Fok Ying Tung Education Foundation under Grant No.81062 (霍英东教育基金青年教师基金).

作者简介: 岳昆(1979—),男,云南曲靖人,主要研究领域为数据挖掘,XML 数据管理;王晓玲(1975—),女,博士,讲师,主要研究领域为 XML 数据管理;周傲英(1965—),男,博士,教授,博士生导师,主要研究领域为数据挖掘,XML 数据管理,P2P 对等计算.

及其核心支撑技术,概括了 Web 服务中的数据集成技术、Web 服务的组合、语义 Web 服务、Web 服务发现、Web 服务安全、P2P(Peer-to-Peer)新型计算环境下的 Web 服务解决方案和网格服务等方面的研究内容,并对这些技术进行了总结,结合已有的研究成果,展望了 Web 服务未来的研究方向及其面临的挑战。

关键词: Web 服务;服务组合;语义 Web;服务发现;安全性;P2P;网格

中图法分类号: TP393 文献标识码: A

近年来,随着电子商务的迅速崛起,Web 应用从局部化发展到全球化,从 B2C(business-to-customer)发展到 B2B(business-to-business),从集中式发展到分布式。Web 服务作为一种新兴的 Web 应用模式,是一个崭新的分布式计算模型,是 Web 上数据和信息集成的有效机制。从电子商务应用领域来看,复杂的应用连接和程序代码造成了电子商务应用的高维护代价和更新代价,而 Web 服务正好能够解决这一问题,成为目前应用环境中最为合理的解决方案^[1]。从 Web 服务中的支撑技术来看,很多关键问题有待解决,具有广阔的研究空间,但同时也存在很多挑战。

随着 Web 应用从集中式向分布式的迅速发展,以提高电子商务系统的健壮性、高性能计算能力为出发点,为了解决现实 Web 应用中“应用到应用(application-to-application)”及“点对点(peer-to-peer)”的核心问题,使当前 Web 应用适应全球化和复杂商务处理的需求,进行 Web 上已有的网络计算组件的集成,基于现有协议提高 Web 应用的互操作能力及服务质量,研究 Web 服务核心支撑技术具有重要的意义^[2-4]。

文献[1,3,5]指出了 Web 服务的基本构架,包括 Web 服务提供者(service provider)、服务请求者(service requester)和服务代理(service broker)以及它们之间的通信、XML 格式的消息传递机制等。正是基于该架构,对 Web 服务核心支撑技术的研究和应用才得以迅速开展。同时,源于 Web 服务的现实应用背景,针对现存的主要问题,学术界有不同侧面的研究工作和相应的技术成果。例如,Web 服务中的数据集成^[6-16];Web 服务的组合^[17-19];基于分布式数据库的特征研究提高 Web 服务质量的有效解决方案^[20];语义 Web 在 Web 服务中的应用^[21-25];高效的 Web 服务发现技术和架构^[26-29];针对 Web 服务应用中的安全及隐私问题,适合 Web 服务的安全和加密策略^[20,30-34];P2P 环境中 Web 服务的有效表示和执行^[35,36];网格计算与 Web 服务的有机结合^[37-40]等。

本文首先概述 Web 服务的基本架构及特征,然后从多种角度概括 Web 服务的不同描述和定义,接着分析当前 Web 服务的主要研究问题及 Web 服务的核心支撑技术,最后总结全文,指出 Web 服务研究面临的挑战,并展望未来工作。

1 Web 服务基本概念

1.1 Web 服务架构

文献[1,5]从面向应用的角度,描述了 Web 服务的基本架构。该架构由 3 个参与者和 3 个基本操作构成。3 个参与者分别是服务提供者、服务请求者和服务代理,而 3 个基本操作分别为发布(publish)、查找(find)和绑定(bind)。Web 服务基本架构如图 1 所示。服务提供者将其服务发布到服务代理的一个目录上;当服务请求者需要调用该服务时,他首先利用服务代理提供的目录去搜索该服务,得到如何调用该服务的信息;然后根据这些信息去调用服务提供者发布的服务。当服务请求者从服务代理得到调用所需服务的信息之后,通信是在服务请求者和提供者之间直接进行,而无须经过服务代理。Web 服务体系使用一系列标准和协议实现相关的功能,例如:使用 WSDL(Web service description language)来描述服务,使用 UDDI(universal description, discovery, integration)

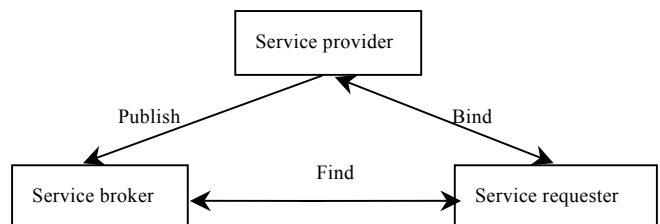


Fig.1 Architecture of Web services

图 1 Web 服务架构

使用 WSDL(Web service description language)来描述服务,使用 UDDI(universal description, discovery, integration)

来发布、查找服务,而 SOAP(simple object access protocol)被用来执行服务调用^[41]。

在 Web 服务架构的各模块间以及模块内部,消息以 XML 格式传递.其原因在于,以 XML 格式表示的消息易于阅读和理解,并且 XML 文档具有跨平台性和松散耦合的结构特点;从商务应用的角度看,从 workflow 到查询数据库,直到同贸易伙伴交换信息,XML 格式表示的消息封装了词汇表,可以同时在该行业组织内部和外部使用;它还有较好的弹性和可扩展性,允许使用附加的信息,并且 XML 标签提供了可访问的进程入口,从而可强化商业规则,并且增强了互操作性^[42],为信息的自动处理提供了可能。

1.2 Web服务的不同描述

Web 服务具有广泛的适应性和应用背景,而且 Web 服务的很多相关问题仍处在研究过程中,学术界从不同的侧面对 Web 服务有不同的描述,限于篇幅,我们概括了如下几种有代表性的 Web 服务描述。

从功能的角度描述 Web 服务,文献[3]认为 Web 服务基于 TCP/IP,HTTP,XML 等规范而定义,具备如下功能:Web 上链接文档的浏览、事务的自动调用、服务的动态发现和发布.文献[41]从组成框架及实现目标的角度描述了 Web 服务,认为 Web 服务作为一种网络操作,能够利用标准的 Web 协议及接口进行应用间的交互.文献[21,22,26,43]从语义的角度描述了基于语义 Web 的服务,认为 Web 服务是语义 Web 的一种应用,由于考虑了语义信息的描述及表示,Web 服务能够更准确地被执行,服务组合(service composition)能够按所期望的目标进行.文献[37]从网格计算(grid computing)的角度指出 Web 服务能用于 Web 上的资源发现、数据管理及网格计算平台上异构系统的协同设计,提出了网格服务的新概念.文献[44]从信息检索的角度提出了在包含了分布策略和路由信息的电子文档之上进行分布式文档检索的 Web 服务。

从另一方面看,针对不同的应用背景,Web 服务的应用对象也不同,目前被广泛应用的 Web 服务可分为如下 4 类^[1]:面向企业应用(business-oriented)的服务、面向消费者(consumer-oriented)的服务、面向设备(device-oriented)的服务和面向系统(system-oriented)的服务.尽管对 Web 服务进行描述的出发点或应用类型不同,但是它们均具有如下共同特征:(1) 应用的分布式.为适应网络应用中分布式的数据源和服务提供者,分布式的服务响应、松散耦合是 Web 服务必须具备的特征.在应用中,服务请求者不必关心服务提供者的数据源格式是什么,某一服务请求需调用哪些服务,服务请求在 Web 上怎样被执行等,即 Web 服务对用户具有分布透明性.(2) 应用到应用的交互.在分布式的环境中,若采用集中控制方式,服务器有较大的负荷,并且系统不具有健壮性.因此应用到应用的交互,使得 Web 服务更具可伸缩性.(3) 平台无关性.Web 服务的界面、跨 Web 服务的事务、workflow、消息认证、安全机制均采用规范的协议和约定;由于 Web 服务采用简单、易理解的标准 Web 协议作为组件接口和协同描述的规范,完全屏蔽了不同软件平台的差异,因此具有可集成能力^[1]。

1.3 Web服务研究的主要问题及其核心支撑技术

与学术界的研究相比,工业界注重的是 Web 服务规范和协议的标准化问题,例如:SOAP 协议,UDDI,WSDL 等.本文不详细介绍 W3C 制订的有关协议,我们着重从学术界的角度来探讨 Web 服务研究的主要问题、相关的解决方案和核心支撑技术。

从 Web 服务的技术层面看,为了满足建立在资源动态变化之上业务日益复杂的商务应用的需求,Web 服务有效地利用现有的 Web 数据集成^[6-10]、数据过滤^[11-13]、服务组合等技术,以进行 Web 上分散资源的集成,并在其中应用访问控制^[34,45]、事务机制^[10,14,15,16,46],以保证服务响应的可靠性及服务组件的协同工作.从 Web 服务的现实应用层面看,实际应用中客户面对表现形式和复杂性都可能不相同的服务,如何正确、高效地找到自己想要的服务至关重要,这正是 Web 服务发现的任务.目前,Web 服务的搜索与发现采用的普遍方法是由客户使用搜索引擎找到服务,或者在相关的 Web 页面中考查其是否满足服务请求的要求.因此,Web 服务发现的自动化,基于语义和约束进行 Web 服务的查找及发现,特定应用环境中的 Web 服务发现技术都是重要的研究方向^[27-29]。

如上所述,Web 服务使分布在网络中的资源构成了一个虚拟的计算机系统,资源的重用和协同工作就涉及到如下两个重要的问题,其一是复杂服务的组织和描述问题.从简单的 B2C,B2B 应用到复杂的 Web 服务,共享上下文(shared context)和服务组合(service composition)是一个重要的研究课题^[3].现实中的应用一般都非常复杂,为了分散和简化应用逻辑,提高服务可重用性,单个 Web 服务都不可能做得非常复杂,因此现实中复杂服务的应

用需要组合多个简单的 Web 服务。其次,Web 应用是基于不同方式的异构系统而建立的,为了将松散耦合的、分散的各类 Web 服务有机地组织成一个可用的系统,其中服务的集成需要组合异构的服务。相应地,服务组合描述语言的设计也是一个重要的研究分支^[47]。其二是充分利用现有的 IT 设备,并进行资源的按需配置的问题。如何使具有分布式结构的应用系统能够在开放式环境下提供较高质量的服务,缩短 Web 服务的响应时间,提高系统的健壮性,是直接影响商务应用的因素^[1,3,20]。针对这些问题,在 P2P 环境下实现 Web 服务是一种有效的解决方案,在这方面目前已经有大量的研究成果及相应原型系统^[35,36]。同样,为了提高 Web 服务协同工作的效率,提高网格计算能力,将 Web 服务与网格计算进行有机的组合,构建网格服务也是一种理想的解决方案^[38-40]。

在服务的组合和协调工作中,首先需要保证服务执行的质量和可靠性,其次需要保证服务执行中的安全性。前者是 Web 服务的质量控制问题。因为 Web 服务的描述缺乏足够的语义信息,Web 搜索、Web 服务发现、Web 服务的组合具有不确定性,为了能够根据用户提供的信息更加准确地描述并执行 Web 服务,考虑更加丰富的语义和上下文信息,在 Web 挖掘中充分利用本体论(ontology)领域模型以提高 Web 服务的质量成为另一个重要而又富有挑战性的研究热点^[3,5,22]。而后者则是 Web 服务的访问控制问题。Web 服务的分布透明性要求 Web 服务的自动集成,这也是 Web 服务研究的主要问题之一。为了达到这一目的,不同的 Web 服务提供者必须为其他模块或组件提供可供信息存取的接口,并且基于分布式的环境实现开放,标准的商务逻辑,为了保证各服务提供者的数据安全,必须进行安全控制,从而使“安全保证”成为 Web 服务中一个自动完成的部分。然而目前 Web 服务中普遍采用的是 Internet 和 WWW 的安全机制,现有的这些机制本身与 Web 服务之间就存在较大的技术鸿沟,因此,适合于 Web 服务的访问控制及其数据加密策略也成为 Web 服务中有待进一步研究的支撑技术之一^[20,30,31]。

总之,利用这些支撑技术,Web 服务的研究具有面向应用、高综合性和高集成度的特点,注重总体系统结构规划、高效服务执行算法和有效集成策略的探索。Web 服务及其支撑技术研究过程中的几个主要问题是:(1) 应用背景是什么,研究通用的方法还是实现特定的 Web 服务;(2) 如何定义和表示 Web 服务;(3) 以怎样的 Web 服务系统结构作为研究的平台和基础;(4) Web 服务的粒度怎样,如何进行服务组合;(5) 建立怎样的代价模型以评估 Web 服务组合的代价,如何权衡影响系统性能的各个因素,使得系统性能总体最优;(6) 语义在服务的研究或实现过程中如何表示、处理和利用;(7) 根据智能化、自动化和高效性的要求,怎样基于语义和本体论发现 Web 服务;(8) 如何进行高效的服务质量管理,以满足用户提出的服务请求;(9) 在安全性的特定要求下,采用怎样的 Web 服务安全、认证和加密策略。针对上述问题,本文将分别对 Web 服务的核心支撑技术进行概述和分析。

2 Web 服务中的数据集成

2.1 Web 信息集成技术

Web 数据源分散、动态变化的特点使得 Web 上的数据集成比基于数据库的数据集成更复杂,应用范围更广。以传统的信息集成技术作为基础,如数据格式的转化及统一、数据模式的动态调整^[49]、索引创建等,文献[8,9]分别提出了 Web 数据集成的新方案 ARAIADNE 和 InfoSleth,它们具有动态数据集成、基于代理技术进行查询计划的协调和数据分发等特征。首先,ARAIADNE 和 InfoSleth 均通过构建信息代理(Agent)来进行 Web 数据的抽取,查询执行和信息集成,包括用户代理、任务执行代理和资源代理。其次,领域模型(domain model)或本体论(ontology)被用来描述数据和资源的特征,获取数据的模式,从而为数据的集成提供统一的数据表示。ARAIADNE 以层次的方式对 Web 数据建模,并对其创建索引,以解决 Web 页面的定位问题,也基于此进行查询的处理和优化;InfoSleth 提出了基于内容(content-based)的数据分发技术。

鉴于上述方法,文献[6,7,10]提出了通过动态 Web 服务组合进行数据集成的技术,简化了 ARAIADNE 和 InfoSleth 中代理的定义及实施策略,并对查询计划进行了规范化。其中服务组合的上下文协同转换以及 Web 资源的选取策略均是有待进一步改进的方面。本文第 3 节将对 Web 服务组合技术作相应的概述。

2.2 Web 服务中的数据过滤

Web 服务中的数据过滤包括 Web 服务请求的过滤和 Web 服务描述的过滤两方面,这些信息都是用 XML 格式表示的,因此,XML 文档的一些过滤技术^[11,12]可应用到 Web 服务中。例如,文献[11]提出了事件驱动的 XML

数据在线过滤技术,以 XPath 查询表达式作为过滤条件(如//stock/symbol[text()>15]),由此建立 XTriE 索引,过滤

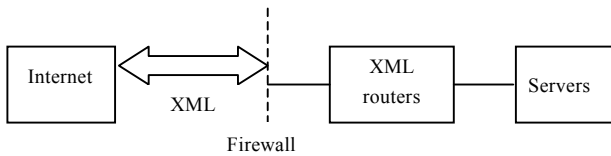


Fig.2 The system architecture of SOAP message filtering

图2 SOSAP 消息过滤系统架构

的过程就是 XML 文档与 XPath 索引树匹配的过程.将该技术应用到 Web 服务请求的过滤中,文献[13]提出了基于 XTriE 进行 SOAP 消息过滤的技术.其基本框架如图 2 所示.其中,XML Router 对服务器接受到的 XML 消息进行过滤,并将其发布到服务器上,其作用类似于防火墙.该方法将 XML 格式的 SOAP 消息请求表示为

树形结构,在系统设计时将相关的 XPath 过滤条件构建为 XTriE 索引树,SOAP 消息的过滤过程即 XML 与 XTriE 树形结构数据的匹配过程.此外,文献[13]也提出了并行 XTriE,层次 XTriE,XTriE 聚类等 SOAP 消息过滤的优化策略.

2.3 Web 服务中的事务处理

Web 服务提供的是相互独立的网络平台上应用的连接和信息的集成,在这样的分布式环境中,Web 应用需要相互协同工作并保持一致,得到可靠的结果和输出.自然地,事务机制可解决这些问题,而 Web 服务自身具有不同于传统数据库的特征,Web 服务中的事务机制具有以下主要特点:(1) Web 事务比传统事务更松散,更灵活,更复杂,并不严格地遵循传统事务 ACID 原则;(2) 商务处理中有很多长事务,也存在可选的子事务;(3) 事务涉及到的实体是分布在网络中不同位置、不同平台上的服务参与者;(4) 服务组合中需要事务机制来保证其协调工作;(5) 事务对于服务质量和计算的可靠性起着重要的作用.

目前,Web 服务中的事务处理策略通常是扩展已存在的事务处理技术,仍然有待进一步研究,用于 Web 服务的事务模型主要有:Business Transaction^[50],WS-Transaction^[14]和 Activity Service^[46],其中 WS-Transaction 建立在 SOAP,WSDL 等 Web 服务的标准之上.文献[15]在满足传统事务 ACID 性质的基础之上引入协调器(coordinator),扩展了现有的事务处理技术,提出了适用于 Web 服务的协同工作框架及协议.协调器实际上是用 XML Schema 定义的协调上下文,包括 ID、上下文的有效期、协调类型等元素.该方法为 Web 服务中的事务定义了共同的模型,为服务所需创建的事务提供了必须遵循的模式.此外,文献[16]针对服务组合,基于 CORBA 和 J2EE 对象通信模型,提出了对象事物(object transaction)的解决方案.

3 Web 服务组合

3.1 Web 服务组合问题的提出

一方面,异构系统中运行在不同平台之上的 Web 服务可能是以不同的方式创建、用不同程序语言实现、由不同供应商提供的,服务的请求需要根据特定的应用背景和需求进行合理的服务组合;另一方面,服务组件(service component)或基本服务(elementary service)不可能很复杂,这也是服务可重用的必要条件之一.因此需要按照一定的粒度进行 Web 服务的组合.文献[3]提出了复杂 Web 服务(complex Web service)的解决方案:共享上下文及 Web 服务组合.任何与特定应用主题相关的 Web 服务之间总是具有某些直接或间接的联系,它们具有一些共同的上下文及应用背景,这样的一系列元数据称为共享上下文.例如,相同的本体论、词典(thesauri)及 XML 信息交互格式等.同样,上下文也可以是执行特定 Web 服务应该具备的先决条件或系统中与该应用相关的状态信息.基于共享上下文,服务组件及基本服务间的关联能对服务的组合进行指导和监督.

总结上述方法,在共享上下文的环境中进行服务的组合需要解决以下几个主要问题:(1) 对于与组合服务相关的各服务组件和基本服务,怎样定义它们之间的逻辑及时序关系,以实现复杂 Web 服务执行的自动化;(2) 怎样实现服务组件和基本服务之间的动态交互、协调及状态保持,以保证 Web 服务执行的有序性;(3) 怎样保持语义信息,怎样验证和测试组合 Web 服务,以确保 Web 服务执行结果的正确性.

3.2 从Web服务组件到组合Web服务

文献[18,19]以动态 Web 服务组合为出发点,独立于分布式的系统架构,基于分布透明性的基本要求,直接针对服务本身提出了 Web 服务组合及管理的有效策略:运行时(runtime)服务的动态创建.其中提到的基本服务,是外部用户可以引用的最小服务单位^[17,36].复杂 Web 服务可以由服务组件和基本服务构成.服务由服务组件构成,与服务组件相比,服务可由外部用户引用^[19];而各服务组件封装了相应的服务功能和适合的数据,它们之间相对独立,其属性包括组件的功能描述及操作系列、组件间的操作约束和依赖信息.动态 Web 服务组合是在运行时从一系列服务组件中动态创建满足特定应用需求的新服务的过程,该过程包括 3 个阶段:组合前服务组件的定位及选择、动态服务组合、组合后的新服务注册.同时,文献[19]提出了由服务组件动态创建组合服务(composite service)的 3 种不同技术:(1) 创建组合服务接口;(2) 基于管道和过滤(pipe-and-filter)机制创建新的组合服务;(3) 按照需求通过服务组件的动态组合创建新的组合服务.正因为服务是动态创建的,服务的设计阶段只需考虑构成这些服务的服务组件^[18],而且,组合服务的动态生成可以有效地提高商务软件的敏捷性、灵活性及可用性^[24].

在分布式的环境下,不同的服务组件可能提供相同的功能,若构成某服务的组件之一在功能上发生了变化,为了使该服务的执行不受影响,当该服务被请求时需要进行服务组件的动态切换,将对它的调用动态切换到与之具有相同功能的另一服务组件.服务组件之间动态切换问题的解决方案直接影响了整个 Web 服务系统的有序性和可再现性.此外,如何管理服务组件、组合服务与服务组件之间的对应关系、服务组件的定义及服务组合中所必需的应用逻辑与 workflow 管理,也会成为上述方法的瓶颈问题.

针对上述问题,文献[24]从 Web 服务管理的角度提出了对服务提供者进行操作控制的动态调整算法,并扩展了 WSDL,使之能够描述不同类型服务之间的约束、依赖、认证策略等;也提出了 WSOL(Web service offering language),形式定义了组合服务的相关约束,包括前提条件(pre-conditions)和组合后条件(post-conditions);还提出了用于动态服务组合的服务组件动态切换(dynamic switching)机制.同样,文献[11]从更高的层次,以基本服务作为构成复杂服务的最小单位,基于状态保持机制提出了查阅状态表的解决方案.当利用此方法时,在服务设计阶段,对所有提供相同功能的基本服务建立服务名到期对应例程(routine)间的映射,同时分别对每个基本服务建立状态表(state chart).这样,完全通过查阅状态表进行服务的动态组合,组合服务的管理,执行与基本服务间的自动切换.该算法形式化地描述了服务组合所需要的约束和时序信息,并且服务组合的方法更具可操作性.

3.3 组合Web服务的描述及实现

文献[47]分析了目前 Web 服务设计语言(如 Java)及各种数据源之间存在的不匹配问题,以设计一种非过程化,专门用于 XML 应用和 Web 服务的描述性高级语言为目标,提出了以 Web 服务的定义和组合为核心的 XML 编程语言 XL.基于 XML 查询语言 XQuery 的表达式,XL 定义了赋值(assignment)、表达式(expression)、序列(sequence)和并行(parallel)等 Web 服务的组合模式,并且将每个 Web 服务用一个 XL 程序来表示,XL 程序之间的调用和交互通过 SOAP 消息的传递来实现.从而利用上述的 Web 服务组合模式和 XL 定义的事务处理机制将服务有机地组合起来.基于这样的描述语言和平台,文献[48]实现了用 XL 描述的 Web 服务原型系统.

此外,基于 workflow 的基本思想,文献[17]提出实现组合 Web 服务的一种解决方案,并给出了保证组合 Web 服务正确调度的理论依据.该方法通过服务社区(service community)、组合服务、基本服务和异构服务提供者的层次结构,有效地描述并解决了增量式的 Web 服务发布及注册问题.该层次结构如图 3 所示.

根据现实中商务应用的不同种类,将 Web 服务也按其行业的相关性划分成不同的类别,即这里的服务社区.它是一个虚拟的概念,是表示不同类别服务的视图.组合服务及一部分基本服务被注册到服务社区上,使得基本服务发生的变化只影响到所属服务社区的注册信息,由此可实现 Web 服务功能变化时注册信息的局部化更新.此外,服务的执行实际上是由基本服务的通信、协调和执行来完成的.该方法通过状态定义和检测的方式有效地解决了组合 Web 服务执行中存在的如下问题:组合服务的执行需要调用哪些基本服务以及相关的基本服务怎样按预期的顺序自动执行.具体而言,该方法预先定义了特定应用的服务(包括组合服务和基本服务)以及执行它所需具备的系统初始状态(preconditions)和执行后的状态(postprocessings),这些状态与服务间的对应关系构成了状态路由表(routing table);通过当前状态检测查找状态路由表,以确定是否可以结束当前执行的服务、同

时应该启动哪些服务,基于并发策略实现了组合 Web 服务执行过程中的合理调度.

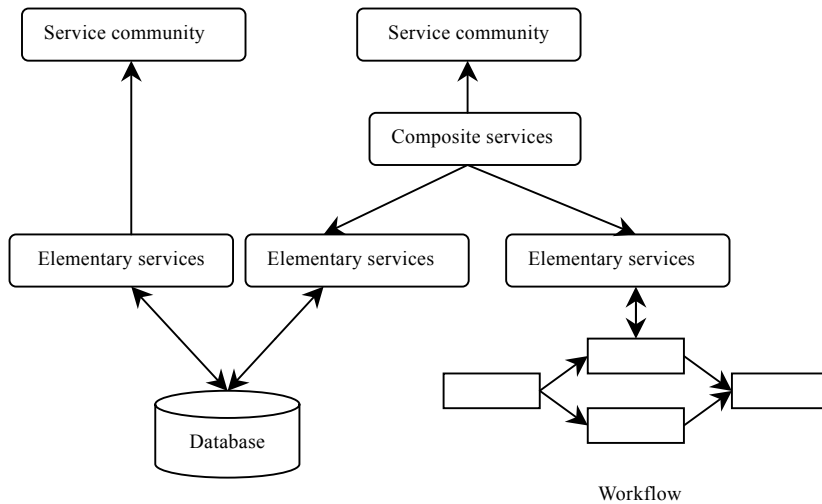


Fig.3 Hierarchical structure of composite Web services

图3 组合 Web 服务层次结构

总而言之,Web 服务的组合为复杂的 Web 应用提供了有效的解决方案.从现实应用来看,Web 服务的组合可以实现虚拟社区中软硬件的共享;从 Web 服务本身来看,Web 服务的组合实现了组合服务的动态生成,提高了服务组件及基本服务的可重用性和利用率,减少了系统的开销.但是其中存在的诸如服务组合粒度、服务组件或基本服务之间的通信方式及其优化、服务的聚类、服务的社区化分类管理、组合服务的有效性验证及安全等问题,有待进一步的研究和探索.

4 基于语义的 Web 服务

4.1 语义 Web

随着 Web 应用的迅速发展,Web 应用面临着自动化及语义保持的问题.对于用户提交的应用请求,如何根据语义信息执行分布在 Web 上的相关服务,并自动地进行这些服务间的切换,是语义 Web 被提出的动机所在.使机器可存取 Web 数据并实现处理的自动化是语义 Web 的目标^[15].就语义 Web 本身而言,我们可以认为它是目前 Web 应用的扩展.它能够描述一定的语义,使计算机及人类能够更好地协调合作.从语义 Web 的表示方式来看,可以认为它是基于资源描述框架(resource description framework,简称 RDF)和元数据(metadata)对 WWW 上数据的抽象表示^[26,51],是本体论领域模型的具体表示和应用实例^[52,53].而本体论是一组概念及这些概念间关联描述的集合,它描述了包括客观事物及它们之间联系的领域知识.为了表示语义信息,DAML 扩充了 WSDL.它是基于语义 Web 的服务描述语言,建立在 XML 和 RDF 的基础上,为机器提供了读取数据以及对数据进行解释和推理的能力,使得面向人的 Web 转换到了语义 Web^[16].在 Web 应用中有效利用本体论,可以在一定程度上实现语义 Web 信息处理的自动化,例如,提高 Web 搜索的准确性,提高 Web 服务质量.目前,语义 Web 研究的问题主要包括:Web 语义的提取和表示,Web 应用处理过程中的自动规则推理及知识管理,以及有效利用语义信息进行 Web 服务处理和 Web 挖掘,提高 Web 发现的准确性和智能化.

4.2 组合 Web 服务中的语义保持

如果在 Web 服务中利用了语义信息,服务的执行就更能体现用户预期的目标和限制条件,能够得到更精确的结果,从而提高 Web 服务执行结果的准确性;从组合 Web 服务的角度来看,基于语义信息进行 Web 服务的组合能够有效地利用知识表示及推理来指导和监督服务的组合.例如,可以通过规则的推理来确定满足条件的 Web 服务,从而确定合理的 Web 服务调用顺序和工作流.

在现实应用中,服务是否被执行需要通过对一系列条件的判断来确定,同时也要求 Web 服务遵循预先定义的语义而被执行,并需要判断服务执行的结果是否符合所期望的目标.基于文献[20]提出的 Web 服务本体论描述语言 DAML-S 的 Web 服务描述能力,文献[22]用一阶逻辑语言定义了 DAML-S 相关子集的语义,提出了基于推理规则自动执行组合 Web 服务的解决方案.即当任何服务组件或者基本服务被调用时,将相关的当前状态及输入信息构成的初始条件用一阶谓词的形式表示为逻辑表达式,通过对该表达式真值的判断来确定该服务是否可以被调用.例如,服务组件或基本服务 a 在系统环境 s 中执行的前提条件表示为以下逻辑表达式: $\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n \wedge \text{input}(\varphi_1, s) \wedge \text{input}(\varphi_2, s) \wedge \dots \wedge \text{input}(\varphi_m, s)$,其中, $\pi_i (i \in [1, n])$ 为当前的状态信息, $\varphi_j (j \in [1, m])$ 为输入信息,只有当该逻辑表达式为真(true)时, a 才可以被调用.

利用上述方法,在服务设计阶段就可以根据实际应用中的语义信息定义各个服务组件或基本服务应该满足的前提条件,并以逻辑表达式的形式表示.组合服务在执行过程中,对于当前状态和输入信息,通过判断该逻辑表达式的真值来确定应该被调用的服务组件或基本服务.显然,采用上述方法可以使服务组件或基本服务的调用和执行能够按预期的目标被调度,原先定义的语义信息可得以保持.但是,这是一种静态的语义保持方法,不利于语义信息的动态扩展及 Web 服务的动态更新.

4.3 语义Web和Web服务的结合

Web 上的本体论研究包括语义 Web 的研究和 Web 服务的研究,其共同目标都是通过利用 Web 上人和机器都能够存取的内容,创建智能自动服务及商务处理基础设施.考虑以上二者的结合,实现功能互补是一种自然的选择^[23,25].首先,用于描述 Web 服务 WSDL 并不能很好地表达 Web 服务的语义信息;并且所描述的是静态的 Web 服务,不包含任何有关服务执行过程的信息,而实际应用中要求服务的执行具有动态性,并且能够达到自动化.其次,从语义 Web 的角度看,一系列本体论构造算子(ontology construct)能够通过 Agent 使 Web 服务自动地被发现、激活及执行,并且本体论为服务及服务间关系的描述提供了强有力的手段.文献[23]提出了在 WSDL 中加入以 XPath 语法形式表示的语义标注,从而对 WSDL 进行扩充.其语义的表示完全基于 XML 规范,与原 WSDL 进行了无缝结合,使 Web 服务的描述能够体现“实际值”(actual value)条件限制,服务的执行更符合预期目标.

总之,在 Web 服务中有效利用本体论领域模型进行服务的概念建模,可以指导 Web 服务应用的设计;在 Web 服务中有效利用语义信息,进行 Web 服务和语义 Web 的有机结合可提高 Web 服务的质量;语义 Web 服务是语义 Web 和 Web 服务的结合,可为 Web 服务的发现、执行、解释和组合的自动化提供有效的支持.其中,分布式环境下进行语义的提取、语义的表示、探索适合 Web 服务的知识推理方法等方面都是有待进一步研究的开放课题.

5 Web 服务发现

5.1 Web服务发现问题的提出

基于 Web 的商务应用为客户提供了不同类型的 Web 服务,而这些服务可能具有不同的形式,并且它们的复杂程度不相同,例如面向软件(software-oriented)的服务、面向硬件(hardware-oriented)的服务、移动(mobile)服务等.所谓 Web 服务发现,就是客户以某种方式在这些不同类型的 Web 服务中找到其想要的服务,以执行 Web 服务请求.Web 服务发现是 Web 服务系统架构中的一个重要部分,UDDI 是其中一种解决方案.而传统的服务发现技术是通过精确匹配实现的,不能较好地支持基于概率和语义约束的模糊匹配,使得服务执行的整个过程受到影响,而且语义冲突问题(对于同一服务在不同站点上实现了不同接口)也没有可行的解决方案.

5.2 主要的Web服务发现技术

Web 服务发现与传统的信息检索有很多相似之处,但前者具有更高的复杂性和实现技术的不成熟性,将有更广阔的应用领域,大多数 Web 服务发现技术都考虑有效利用语义信息和本体论,以服务发现的自动化和智能化为目标,采用信息检索中的某些评价标准来评价 Web 服务发现技术的性能,例如查准率(precision)和查全率(recall).目前 Web 服务发现的主要技术^[35]及其比较见表 1.

Table 1 Comparison of existing service discovery technologies

表 1 现有的服务发现技术比较

Service discovery technologies	Precision	Recall	Hardness
Keyword-Based	Low	High	Average
Frame-Based	High	Low	Average
Deductive retrieval	High	High	Hard

大多数服务查找和发现技术都是表 1 中“基于框架”的方法,如 UDDI。“演绎检索”方法是“基于框架”方法的进一步改进,所有的服务属性都以逻辑的形式表示;然而,服务的逻辑表示和服务的逻辑推理都具有很高的计算复杂性,使得该方法并不能应用到实际中.针对这些方法各自的不足,文献[27]提出了一种基于语义 Web,利用过程

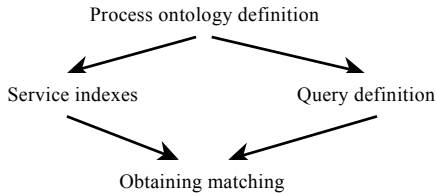


Fig.4 Architecture of service retrieval using process

图 4 利用过程本体论的服务检索架构

过程本体论^[21](process ontologies)的 Web 服务发现技术.该技术将服务的功能作为过程模型^[21](process model)定义了过程本体论,并将 Web 服务通过索引建立到本体论上(即用过程本体论表示服务).过程查询语言 PQL(process query language)用来定义查询,包括实体(entity)和联系(relationship)两类字句.服务的查找和发现过程就是过程模型中的过程本体论与查询中所用到的过程本体论相匹配的过程.其功能架构如图 4 所示.

与前面 3 种技术相比,这种技术中的本体论,查询都以 XML 格式定义,具有高查准率、高查全率和易实现的特点,但服务的过程本体论的定义直接影响了服务发现的最终结果.

此外,在特定应用背景下的服务发现技术方面,Dreggie^[28]和 GSD^[29]是两种具有代表性的方法.现有的服务发现技术不能有效地应用于移动自组网络(mobile Ad hoc network),Dreggie 用 DAML 表示服务的语义,并将其作为一种推理语言,不同于以往基于严格句法匹配的服务发现方法,Dreggie 用基于语义的 Prolog 推理引擎进行服务模糊匹配和不完全匹配.它依次解析 DAML 查询请求、服务本体论、来自服务提供者的 DAML 服务轮廓描述、DAML 联系规则,并装载到知识库(KB)中,再进行匹配,从而实现 Web 服务的发现.GSD(group-based distributed service discovery protocol)是一种用于 P2P 环境中的服务发现协议,在 P2P 环境下,每个 Peer 对其他 Peer 上的信息作缓存(caching),服务按照 DAML 中“类-子类”的层次关系进行分组,对于服务的请求,只需将它发送到语义相关的服务分组(group)而无须发送到所有 Peer 上.该方法避免了广播风暴,能有效利用网络带宽,保证了移动自组网络中服务发现的高效性.

总之,Web 服务发现的研究目标是服务发现的高效率和自动化,在服务发现技术中,利用语义描述和服务本体论是达到该目标的有效途径;而如何用 DAML 描述服务,如何将语义 Web、本体论和语义推理应用到 Web 服务的查找和匹配中是有待进一步研究的课题.从这个意义上说,Web 服务发现是语义 Web 在 Web 服务中的一种应用.特定环境中的 Web 服务发现技术是另一条研究线路,基于通用的方法,在新型网络环境(如移动通信,P2P)中实现高效的服务发现,需要将网络架构本身的机制及特点与 Web 服务的语义信息、本体论进行有机结合.

6 Web 服务安全

Web 应用要求 Web 服务具有分布透明性和自动集成的能力.为了达到这一目的,不同的 Web 服务提供者必须为其他参与集成的模块或组件提供可进行信息存取的接口,并且基于分布式环境实现开放的、标准的商务逻辑.为了保证各服务提供者的数据安全,必须在服务提供者的层次进行安全控制或者信息加密,从而使“安全保障”成为 Web 服务中一个自动完成的部分.

从共享上下文和 Web 服务组合的角度,文献[3]指出在简单服务集成的同时也需要提供可配置的入口(portal)以及可供服务组合的上下文,但这些信息可能是企业或个人的隐私,需要采取相应的认证、加密等措施,因此 Web 服务的安全是一个重要的研究方向.目前,Web 服务安全采用的是已有的 Internet 和 WWW 的一些安全策略,例如数字签名(digital signature)^[32]、XML 加密(encryption)技术和标准^[33]、访问控制(access control)技

术^[34,54]等,然而这些都是静态的措施,并且其本身和 Web 服务之间就存在技术上的鸿沟,因此,Web 服务的安全问题没有得到根本的解决.Web 服务是动态变化的,并且不同应用背景的服务所需要的安全措施可能完全不同.现有的一些安全技术只在一定程度上解决了特定系统的安全问题,在 Web 服务的集成中并不能发挥令人满意的作用.

6.1 访问控制技术

Web 应用的日益广泛使得 Web 页面不仅作为用户信息浏览的媒介,也作为访问邮件、金融数据、统计信息等多种个人机密信息的途径,访问控制技术的应用使得 Web 上的这些数据仅对某些用户可见,从而保证数据的安全性.就目前广泛应用的访问控制技术而言,包括如下两个基本组件^[57]:① 配置策略,用于定义哪些用户允许访问给定的对象、哪些用户禁止访问给定的对象;② 认证策略,用于判断用户或主机在允许访问的实体集中是否成功访问.基于这两个组件,目前主要的访问控制有 3 种方法,见表 2.

Table 2 Comparison of existing access control technologies

表 2 现有的访问控制技术比较

Access control technologies	Strategies	Primary characteristics	Reliability
Host-Based	Based on the network identity of the host that originates the client request	Unable to control the accesses aiming to specific users	Low
Basic authentication	Establish the special file storing the authentication of usernames and passwords	With the HTTP protocol,the usernames and passwords are traversed but unencrypted	Average
SSL/TLS-Compliant client certificates	X.509 client certificates, each certificate information client contains the mathematical algorithm, the certificate issuer, the validity period and certificate subject	Able to control the access of various users, the usernames and passwords are traversed and unencrypted	Very high

除了这 3 种较通用的方法之外,特定的应用背景中还有特定的访问控制方法.如文献[34]提出了基于带有验证授权(proof-carrying authorization,简称 PCA)的访问控制方法.它根据特定的应用逻辑,对表示为逻辑表达式的访问请求信息进行验证,主要用于 Web 页面的访问控制,这种控制对用户透明.

6.2 Web 服务安全架构和服务加密

首先,为了提高 Web 服务质量,缩短其响应时间,缓存和复本(replication)策略很自然地应用到 Web 服务中^[13].在这样的 Web 服务体系结构中,服务提供者保存了本地数据及其他服务提供者的数据副本,不但存在副本的一致更新问题,同时也存在由于服务提供者源数据在 Web 上的分散副本而引发的数据安全问题的.文献[20]提出了一种新型的 Web 服务中间件架构(如图 5 所示),每个服务提供者都包括安全子对象(security sub-object)模块,该模块检测提交的服务请求是否有效,是否被允许,并且以标准化的接口进行通信,保证了数据的安全.

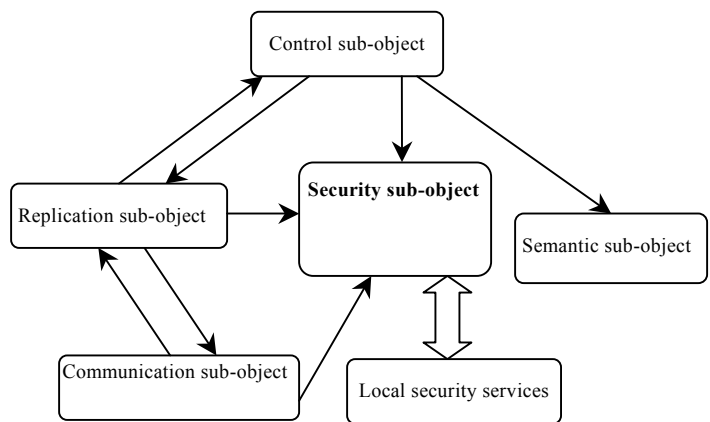


Fig.5 Web services middleware including security components

图 5 包括安全模块的 Web 服务中间件

其次,从 Web 服务安全本身来看,一方面,为了使包括安全模块在内的 Web 服务能够实现服务的集成,Web 服务的安全也需要标准化和规范化,不同的商务处理应该采用相同的 XML 安全标准;另一方面,网络环境中 Web 服务间需要进行消息的传递、信息的交互和数据的共享,同时也需要信息保密,而 Web 服务的分布式体系结构使得数据的安全和加密问题更加复杂.从现有的 Web 服务安全策略来看,基于特定的应用领域和服务架构,

利用成熟的数据分块和加密算法,将安全或认证信息加入到 XML 格式的消息文档中,可以实现 Web 服务消息的安全传输.此外,用户提交的请求中可能包括了其隐私信息,因此服务请求者一方也需要采取相应的信任和保密措施.针对上述几方面的问题,目前 Web 服务中采用的安全技术主要有以下几种^[30,31]:(1) 在客户端建立用户信任机制,执行服务时将相应的认证信息导入服务器;(2) 在 SOAP 消息头中加入针对特定应用的安全表示(token),则可从中提取认证、信任信息;(3) 在某个特定的应用领域内,对服务提供者的内部敏感数据进行加密,当其收到服务请求后直接在加密了的数据上进行相应的计算和处理,计算结果解密后返回给服务请求者;(4) 从服务请求者的角度看,请求者需要提交必要的输入数据.考虑到客户信息的安全性,客户将需要提交的服务请求信息进行分块,每次仅提交一个输入数据块,返回的结果对应于该请求,经过多次服务请求,在结果返回之后,由服务请求者进行各次服务执行结果的集成,从一定程度上保证了客户信息的安全.

综上,Web 服务安全架构的策略对 Web 服务系统的设计有较高的要求,并且,Web 服务的执行面临着 Web 服务基本模块与安全模块间的通信、安全模块中的服务信息随 Web 服务的一致更新等问题,适用于较大规模的 Web 服务应用,具有高可靠性,但通用性较差;服务加密的方法则侧重对服务本身的管理,可以利用或改进现有的数据加密算法,这种方法可应用到各种 Web 服务系统中,具有较好的适用性和较低的复杂度,但其可靠性不如前者.此外,考虑到在分布式的环境(如 P2P)下的 Web 服务组合,各服务组件或基本服务之间需要有信息交互和数据共享,以进行有效的调度.这些相对独立的服务组件或基本服务的安全仍然是具有现实意义的研究课题.

7 P2P 环境中的 Web 服务

P2P 是一种新兴的基于对等网络的架构,是计算机系统的结构从传统的集中式发展为浏览器/Web 服务器/数据库服务器的新模式.P2P 具有如下的特征^[55]:(1) 结点之间通过直接的交互而实现资源共享;(2) 资源分布在各个结点中,而不是集中在一个服务器中管理;(3) 节点具有动态性和即时性;(4) 纯粹的 P2P 系统没有任何集中控制机制,系统中各结点运行的 P2P 系统软件功能相同,各结点之间的交互对称.从 Web 服务提供者来看,它们分布在松散耦合的网络结点上,某些服务提供者相对于另一些服务提供者而言也是服务的请求者^[7],Web 服务和 P2P 计算环境本身具有较多的相似之处和共同特征,所以,在 P2P 计算平台上建立 Web 服务的是一种较理想的 Web 服务实现方案,可以有效利用 P2P 本身的优势高效地实现服务的集成及资源的自治.

原型系统 AXML^[7]是 P2P 平台上 Web 服务应用的一个实例.它采用的是“轻量级”的 P2P 架构,也就是说,系统的实现中对各 Peer 之间的交互作了大量的简化,并采用了相应的等价策略.AXML 中,一系列 XML 文档是 Web 服务的数据源;而作为分布在各 Peer 上的服务提供者,这些 XML 文档中嵌入了 Web 服务的调用请求,这些请求以特殊的 XML 标签标明,服务的请求中给出该服务的提供者的 Peer 标识;每个服务提供者对应一个 Peer,这些 Peer 之间通过 XML 格式的 SOAP 消息进行通信,实现了嵌入在 XML 文档中的 Web 服务请求的执行.

基于 P2P 计算环境,文献[17,36]提出了组合 Web 服务的解决方案.该方案有效利用了 P2P 计算平台的优点,提出组合服务的执行、基本服务的合理调度、协调通信和消息传递机制以及动态 Web 服务的增量式注册和发布策略,并实现了原型系统 SELF-SERV^[36].其中,协调器组件(coordinator)、容器组件(wrapper)及 XML 格式的状态路由表(routing-table)在 P2P 平台上以图 6 的方式进行通信.无论是组合 Web 服务还是基本 Web 服务,每个服务对应了一个协调器组件和一个容器组件,协调器组件实现了各服务间状态的通信,容器实现了相应 Web 服务的执行,“服务执行完成消息”被送回协调器以判断该协调器组件对应的

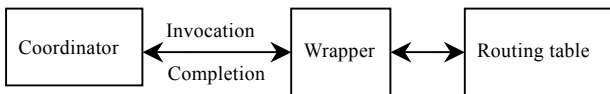


Fig.6 Communication of Web services on peers

图 6 Peer 上的 Web 服务通信

Web 服务是否还需等待其他 Web 服务的执行.此外,文献[36]提出了 P2P 环境下的一种 Web 服务社区化管理的基本框架.

总结上述几种解决方案,在 P2P 环境下实现 Web 服务应该考虑如下几方面的问题:(1) 定义在 Peer 上的服务及 Peer 上装载的服务数据源是什么;(2) Peer 上的 Caching 机制怎样充分发挥作用,如何对 Peer 上 Web 服务的备份信息进行一致更新;(3) 如何在 Web 服务中应用 P2P 相应的好机制解决 Peer 间的通信和路由、服务

注册和查找等问题;(4) 如何在 P2P 对等计算环境中进行 Web 服务基本架构中 3 个参与者的功能划分及分类管理;(5) 如何分类管理各个 Peer,并采用相应的消息分发机制,以避免由于服务请求发送而带来的广播风暴。

8 Web 服务和网格计算

8.1 网格和网格计算

网格(grid)是一种集成的资源和服务的环境,包括计算能力、数据信息和知识、软件、人等各种相关的资源和服务.网格的目的是利用互联网把分散在不同地理位置的电脑组织成一台“虚拟的超级计算机”,实现计算资源、存储资源、数据资源、信息资源、软件资源、存储资源、通信资源、知识资源、专家资源等的全面共享.传统互联网实现了计算机硬件的连通,Web 实现了网页的连通,Web 服务实现了程序和程序之间的共享,而网格试图实现互联网上所有资源的全面连通.网格希望用户在使用网格时,就如同现在使用电力一样方便地使用分布在网络上强大而丰富的各种资源^[38].

网格希望给最终用户提供与地理位置无关、与具体的设施无关的通用解决问题的计算能力,可以把网格认为是整个社会发展的基础设施的一部分,网格计算(grid computing)就是基于网格的问题求解,其目标首先是充分利用分布在网络上的 IT 设备,使 IT 设施及其所具有的计算、存储和软件等资源能够按需应变,随需所用;其次,要解决异构系统之间的集成问题,例如电子商务中各企业系统之间的动态信息交互问题.网格环境把网络上现有的计算资源、存储资源能看成一台虚拟的“超级计算机”,用户就像使用本地机群一样使用和管理网络上的各种资源.如同电网中的电压、频率一样,网格环境也提供了一系列的标准来解决异构、异种平台系统之间的集成问题.网格计算的基础是基于 XML 的消息传递.它用 XML 格式来包装各种数据,消息在各个系统之间进行信息交换^[38].

8.2 网格服务

网格计算的特点使其能够建立在 Web 服务的基础上,并利用 Web 服务提高网格计算的能力.为满足分布式计算中协同工作及实时计算的要求,提高网格计算能力,将 Web 服务与网格计算进行有机结合作为科学计算的中间件,网格服务(grid services)是一种理想的选择^[6].文献[45]对网格服务作了系统的定义,描述了开放网格服务体系结构(open grid services architecture,简称 OGSA),OGSA 把网格服务看做是一种特殊的 Web 服务,描述的是一个网格计算和 Web 服务相结合的计算环境,是一个全新的网格标准,它定义了网格服务的描述,服务实例的创建、发现和管理等所必须遵循的一系列的标准和规范.OGSA 为基于网格的应用定义了一个公共的、标准的体系结构,这个结构的核心是网格服务的概念.OGSA 是基于网格服务的分布式体系框架,它以服务为中心,强调的是服务实例的位置透明性和多协议绑定,支持对底层各种平台设施的集成^[38].

网格服务作为一种特殊的 Web 服务,文献[40]基于 OGSA 提出了一种开放网格服务架构实例,包括若干网格服务容器(container),这些容器位于自治、异构的应用服务器上,不同的容器对应基于不同需求的服务聚类,而为资源管理和共享构建了一个虚拟的环境.结构如图 7 所示.

总之,将 Web 服务和网格服务相结合是提高网格计算能力的有效解决方案,以 Web 服务和相关标准为基础,网格计算更侧重对分布在网络上的资源的动态配置、组合和充分利用.因此,Web 服务的良好机制

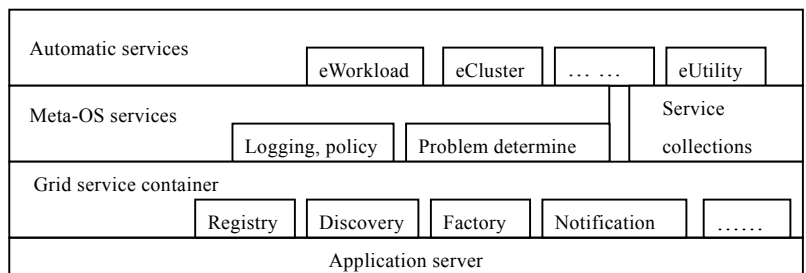


Fig.7 An open grid service architecture instance

图 7 一个开放网格服务架构实例

可以在网格服务中得以利用;网格服务适当地扩展了现有技术,从而使得 Web 服务的计算能力得以扩展.从网格服务的现实应用来看,首先,基于特定服务协议和商业规则的服务之间需要动态组合和配置, workflow 技术可以被

有效地利用到服务的处理过程中,以进行服务的调度和服务间的协调.其次,宽带技术的发展、分布式信息技术的出现、服务器和存储器的虚拟化,使得 IT 设备作为公共设施成为可能,“实时计算模型(on-demand)”和“公用计算(utility computing)”也是网格服务具有现实意义的研究方向.

9 总结与展望

随着电子商务的迅速崛起和 Web 应用的迅速发展,Web 服务成为电子商务的一种有效解决方案,是一种基于 Web 和网络体系结构的新兴应用模式.Web 应用和商务处理的网络化和全球化,信息处理、信息集成的自动化为 Web 服务的发展和研究提供了广阔的应用背景及市场.Web 服务不仅是一些协议的集合,也不是仅独立应用实体的集合,我们认为,Web 服务是一个集应用逻辑、商务智能、网络技术、 workflow 管理、知识表示、逻辑推理、安全保密和信息集成等技术为一体的新兴应用模式.工业界、学术界从不同的侧面对 Web 服务进行研究,从理论及实际应用的角度为其提出了一系列新兴的实施技术和改进策略.本文从 Web 服务的基本概念到 Web 服务的主要研究问题及其核心支撑技术,针对 Web 服务中的信息集成、Web 服务组合、语义 Web 服务、Web 服务发现、Web 服务安全、P2P 计算平台下 Web 服务的实现方案、网格服务等核心支撑技术作了全面的概述,既分析了存在的问题,也指出了研究的动机,既分析理论也结合现实应用,既概述实现技术和系统结构也总结各个问题间的联系,也作了解决同一问题的不同方法之间的比较,同时也指出了 Web 服务各支撑技术的研发中应解决的关键问题和未来的研究方向.

就 Web 服务核心支撑技术的研究而言,存在很多有待解决的开放问题.例如,在组合 Web 服务的实施方案中,服务组件或基本服务的定位、协调、通信及调用策略,服务执行结果的评估和正确性验证,高效服务质量管理策略的探索及服务质量代价模型的建立等;Web 服务组件间采用 P2P 计算平台进行通信和消息传递的新算法,系统的整体布局及优化,适合 Web 服务的加密算法及安全策略;Web 服务提供者异构数据源的管理和集成;基于本体论的 Web 服务概念建模和应用;有效利用语义 Web 的描述功能及元数据提高 Web 搜索^[29]及 Web 服务发现的性能等;基于用户定制、决策支持和客户关系及商务智能的 Web 服务管理;基于已有的 P2P 环境下关系模式动态转换、模式视图间语义关联等的表示方法和实现技术,在 Web 服务中进行服务的动态切换,组合服务视图的构建和组合服务中语义保持、约束的实现;基于预先定义的 DTD 从关系数据库到 XML 文档的发布,XML 数据集成,XML 文档间转换的中间件^[57,58]实现 Web 服务处理及相关 XML 文档的信息集成,Web 数据的转换,基于 Web 服务的中间件,并研究其优化技术;此外,利用本体论描述服务的结构和类型语义,把已有的用于描述和开发组合 Web 服务的概念模型和传统的 ER(entity-relationship)建模方法相结合,对 Web 服务架构,语义 Web 服务进行概念建模,使 Web 服务语义表示模型化,统一化;利用数据挖掘中的聚类方法对 Web 服务进行分类管理,利用关联规则挖掘的方法分析商务应用中 Web 服务间的关联及相关语义信息,从而更好地进行有监督的 Web 服务组合,实现基于商务智能的 Web 服务;利用 Web 搜索的现有方法,基于本体论和语义 Web 的元数据信息,实现更加准确和高效的 Web 搜索服务和 Web 服务发现.我们也正在从事这些方面的研发工作.

References:

- [1] Feisi Science and Technique Center of Research and Development. Application and Development of Java Web Services. Beijing: Publishing House of Electronics Industry, 2002. 190~196 (in Chinese).
- [2] Draluk V. Discovering Web services: An overview. In: Apers P, ed. Proc. of the 27th Int'l Conf. on Very Large Data Bases. Roma: Morgan Kaufmann Publishers, 2001. 637~640.
- [3] Tsur S. Are Web services the next revolution in E-Commerce? In: Apers P, ed. Proc. of the 27th Int'l Conf. on Very Large Data Bases. Roma: Morgan Kaufmann Publishers, 2001. 614~617.
- [4] Curbera F, Nagy WA, Weerawarana S. Web services: Why and how. In: Proc. of the OOPSLA 2001 Workshop on Object-Oriented Web Services. 2001.
- [5] Shmueli O. Architectures for internal Web services deployment. In: Apers P, ed. Proc. of the 27th Int'l Conf. on Very Large Data Bases. Roma: Morgan Kaufmann Publishers, 2001. 641~644.
- [6] Cabrera F, Copeland G, Freund T, Klein J, Langworthy L, Orchard D, Shewchuk J. Web services coordination. BEA Systems & IBM Corporation & Microsoft Corporation, 2002. <http://www-106.ibm.com/developerworks/library/ws-coor/>

- [7] Hansen M, Madnick S, Siegel M. Data integration using Web services. In: Bressan S, ed. Proc. of the VLDB 2002 Workshop Efficiency and Effectiveness of XML Tools and Techniques and Data Integration over the Web (EEXTT). Hong Kong: Springer-Verlag, 2003. 165~182.
- [8] Knoblock CA, Minton S, Ambite J, Ashish N, Muslea I, Philpot A, Tejada S. The ariadne approach to Web-based information integration. *Int'l Journal on Intelligent Cooperative Information Systems (IJCIS)*, 2001,10(1-2):145~169.
- [9] Bayardo J, Bohrer M, Brice R, Cichocki A, Fowler J, Helal A, Kashyap V, Ksiezyc T, Gale Martin, Nodine M, Rashid M, Rusinkiewicz M, Shea R, Unnikrishnan C, Unruh A, Woelk D. InfoSleuth: Agent-Based semantic integration of information in open and dynamic environments. In: Peckham J, ed. Proc. of the ACM SIGMOD Int'l Conf. on Management of Data. Tucson: ACM Press, 1997. 195~206.
- [10] Thakkar S, Knoblock CA, Ambite JL, Shahabi C. Dynamically composing Web services from on-line sources. In: Proc. of the 2002 AAAI Workshop on Intelligent Service Integration. 2002.
- [11] Chan CY, Felber P, Garofalakis M, Rastogi R. Efficient filtering of XML documents with XPath expressions. *VLDB Journal*, 2002,11(4):354~379.
- [12] Chan CY, Fan W, Felber P, Garofalakis M, Rastogi R. Tree pattern aggregation for scalable XML data dissemination. In: Proc. of the 28th Int'l Conf. on Very Large Data Bases. Hong Kong: Morgan Kaufmann Publisher, 2002. 826~837.
- [13] Felber P, Chan CY, Garofalakis M, Rastogi R. Scalable filtering of XML data for Web services. *IEEE Internet Computing*, 2003, 7(1):49~57.
- [14] Web Services Transaction Specification. <http://www-106.ibm.com/developerworks/library/ws-transpec/>
- [15] Cabrera F, Copeland G, Cox B, Freund T, Klein J, Storey T, Thatte S. Web services transaction. <http://www.infosys.tuwien.ac.at/Teaching/Courses/IntAppl/Papers/ws-transpec.pdf>
- [16] Mikalsen T, Rouvellou I, Tai S. Reliability of composed Web services from object transactions to Web transactions. In: Proc. of the OOPSLA 2001 Workshop on Object-Oriented Web Services. 2001.
- [17] Benattallah B, Dumas M, Sheng QZ, Ngu AHH. Declarative composition and peer-to-peer provisioning of dynamic services. In: Proc. of the 18th Int'l Conf. on Data Engineering. San Jose: IEEE Computer Society, 2002. 297~308.
- [18] Mennie D, Pagurek B. A runtime composition service creation and deployment and its applications in internet security, E-commerce and software provisioning. In: Proc. of the 25th Annual Int'l Computer Software and Applications Conf. (COMPSAC 2001). Chicago: IEEE Computer Society Press, 2001. 371~376.
- [19] Tosic V, Mennie D, Pagurek B. On dynamic service composition and its applicability to business software systems. In: Workshop on Object-Oriented Business Solutions (WOBS2001). 2001.
- [20] Steen MV, Tanenbaum A, Kuz I, Sips H. A scalable middle-ware solution for advanced wide-area Web services. *Distributed Systems Engineering*, 1999,6(1):34~42.
- [21] Burstein MH, Hobbs JR, Lassila O, Martin D, McDermott DV, McIlraith SA, Narayanan S, Paolucci M, Payne T, Sycara K. DAML-S: Web service description for the semantic Web. In: Horrocks, ed. Proc. of the Int'l Semantic Web Conf. Sardinia: Springer-Verlag, 2002. 348~363.
- [22] Narayanan S, McIlraith SA. Simulation, verification and automated composition of Web services. In: Proc. of the 11th Int'l World Wide Web Conf. WWW2002. Honolulu: ACM, 2002. 77~88.
- [23] Peer J. Bringing together semantic Web and Web services. In: Horrocks, ed. Proc. of the Int'l Semantic Web Conf. Sardinia: Springer-Verlag, 2002. 279~291.
- [24] Tosic V, Pagurek B, Esfandiari B, Patel K. On the management of compositions of Web services. In: Proc. of the OOPSLA 2001 Workshop on Object-Oriented Web Services. 2001.
- [25] Paolucci M, Kawamura T, Payne TR, Sycara K. Semantic matching of Web services capabilities. In: Horrocks, ed. Proc. of the Int'l Semantic Web Conf. Sardinia: Springer-Verlag, 2002. 333~347.
- [26] McIlraith SA, Son TC, Zeng HL. Semantic Web services. *IEEE Intelligent Systems*, 2001,16(2):46~53.
- [27] Klein M, Bernstein A. Searching services on the semantic Web using process ontologies. In: Isabel C, ed. Proc. of the Int'l Semantic Web Working Symp. (SWWS2001). Amsterdam: IOS Press, 159~172.
- [28] Chakraborty D, Perich F, Avancha S, Joshi A. Dreggie: Semantic service discovery for M-Commerce applications. In: Proc. of the 20th Symposium on Reliable Distributed Systems, Workshop on Reliable and Secure Applications in Mobile Environment. 2001.
- [29] Chakraborty D, Joshi A, Yesha Y, Finin T. GSD: A novel group-based services discovery protocol for MANETS. In: Proc. of the 4th IEEE Conf. on Mobile and Wireless Communications Networks (MWCN2002). 2002.
- [30] Boyens C, Günther O. Trust is not enough: Privacy and security in ASP and Web service environments. In: Manolopoulos Y, *et al.*, eds. Proc. of the 6th East European Conf. on Advances in Databases and Information Systems. Bratislava: Springer-Verlag, 2002. 8~22.

- [31] Thelin J, Murray PJ. A public Web services security framework based on current and future usage scenarios. In: Arabnia H, eds. Proc. of the Int'l Conf. on Internet Computing (IC2002). Las Vegas: CSREA Press, 2001. 825-833.
- [32] Digital Signature Web Service. <http://xml.coverpages.org/Entrust-WSI.pdf>
- [33] Security in Web Services. 2002. http://www.jssl.org/events/seminar/3dec02/JSSL20021203seminar-WYZ_Security.pdf
- [34] Bauer L, Schneider MA, Felteny EW. A general and flexible access-control system for the Web. In: Boneh D, ed. Proc. of the 11th USENIX Security Symp. San Francisco: USENIX, 2002. 93-108.
- [35] Abiteboul S, Benjelloun O, Manolescu I, Milo T, Weber R. Active XML: Peer-to-Peer data and Web services integration. In: Proc. of the 28th Int'l Conf. on Very Large Data Bases. Hong Kong: Morgan Kaufmann Publisher, 2002. 1087-1090.
- [36] Sheng QZ, Benatallah B, Dumas M, Mak EO-Y. SELF-SERV: A platform for rapid composition of Web services in a peer-to-peer environment. Proc. of the 28th Int'l Conf. on Very Large Data Bases. Hong Kong: Morgan Kaufmann Publisher, 2002. 1051-1054.
- [37] Globus OGSA Home Page. <http://www.globus.org/ogsa/>
- [38] Foster I, Kesselman C. The Grid: Blueprint for a new computing infrastructure. San Fransisco: Morgan Kaufmann Publisher, 1999.
- [39] Meta-Data Management of Web Services. <http://www-900.ibm.com/developerWorks/cn/webservices/ws-gt3/index.shtml>
- [40] Foster I, Kesselman C, Nick JM, Tuecke S. The physiology of the grid—An open grid services architecture for distributed systems integration. Open Grid Service Infrastructure WG, Global Grid Forum, 2002. <http://www.globus.org/research/papers/ogsa.pdf>
- [41] W3C Web Services Activity. <http://www.w3.org/2002/ws/>
- [42] Birbeck M, *et al.* Professional XML. 2nd ed. Wrox Press Inc, 2001.
- [43] The Semantic Web. <http://infomesh.net/2001/Swintro/>
- [44] Ghandeharizadeh S, Sommers F, Joisher K, Alwagait E. A document as a Web service: Two complementary frameworks. In: Chaudhri A, eds. Proc. of the XML-Based Data Management and Multimedia Engineering-EDBT 2002 Workshops. Prague: Springer-Verlag, 2002. 450-461.
- [45] Tuecke S, Czajkowski K, Foster I, Frey J, Graham S, Kesselman C. Grid service specification, 2002.
- [46] Procházka M. OMG Activity Service. <http://jotm.objectweb.org/doc/ActivityService020610.pdf>
- [47] Florescu D, Grünhagen A, Kossmann D. An XML programming language for Web service specification and composition. In: Proc. of the 11th Int'l World Wide Web Conf. Honolulu: ACM, 2002. 65-76.
- [48] Florescu D, Grünhagen A, Kossmann D, Rost S. XL: A platform for Web services. In: Franklin MJ, ed. Proc. of ACM SIGMOD Int'l Conf. on Management of Data. Madison: ACM, 2002.
- [49] Halevy AY, Ives ZG, Suciu D, Tatarinov I. Schema mediation in peer data management systems. In: Dayal U, ed. Proc. of the 19th Int'l Conf. on Data Engineering. Bangalore: IEEE Computer Society, 2003. 505-516.
- [50] OASIS Business Transaction Protocol, Committee Specification 1.0. 2002. <http://www.oasis-open.org/business-transaction/>
- [51] SWS Home Page. <http://swws.semanticweb.org>
- [52] Maedche A, Staab S. Ontology learning for the semantic Web. IEEE Intelligent Systems, 2001,16(2):72-79.
- [53] Bussler C, Fensel D, Maedche A. A conceptual architecture for semantic Web enabled Web services. SIGMOD Record, 2002,31(4): 24-29.
- [54] Tutorial: Access Control. <http://stronghold.redhat.com/support/docs-2.4.1/PDF/admin/tutorial.pdf>
- [55] Zhou A, Ling B. Peer-to-Peer system and its applications. Computer Science, 2002,29(8):200-202 (in Chinese with English abstract).
- [56] Berendt B, Hotho A, Stumme G. Towards semantic Web mining. In: Horrocks I, ed. Proc. of the 1st Int'l Semantic Web Conf. (ISWC2002). Sardinia: Springer-Verlag, 2002. 264-278.
- [57] Benedikt M, Chan CY, Fan W, Rastogi R, Zheng S, Zhou A. DTD-Directed publishing with attribute translation grammars. In: Proc. of the 28th Int'l Conf. on Very Large Data Bases. Hong Kong: Morgan Kaufmann Publishers, 2002. 838-849.
- [58] Zhou AY, Wang Q, Guo Z, Gong XQ, Zheng SH, Wu HW, Xiao JC, Yue K, Fan WF. TRES: DTD-Conforming XML to XML transformations. In: Halevy AY, ed. Proc. of the 2003 ACM SIGMOD Int'l Conf. on Management of Data. San Diego: ACM, 2003.

附中文参考文献:

- [1] 飞思科技产品研发中心. Java Web 服务应用开发详解. 北京: 电子工业出版社, 2002. 190-196.
- [55] 周傲英, 凌波. Peer-to-Peer 系统及其应用. 计算机科学, 2002, 29(8): 200-202.