

语义Web服务组合若干 关键技术研究

何 丰 / 著



科学出版社

语义 Web 服务组合若干 关键技术研究

何 丰 著

科 学 出 版 社

北 京

内 容 简 介

Web 服务组合作为一种崭新的分布式计算模型,已成为面向服务的计算的核心技术。本书根据Web服务组合的生命周期,展开了对语义Web服务组合关键技术的研究,重点在Web服务和Web服务组合的形式化模型、Web 服务自动组合的推理算法、Web服务组合的正确性验证、Web服务组合的异常处理四个方面进行了有效的探索与尝试。同时开发了一个基于颜色Petri 网的Web服务组合原型系统CPN4WSC,对于推进动态服务组合技术的理论研究和实用化具有一定的理论价值与应用价值。

本书可作为理工科院校计算机专业研究生的参考书,也可作为Web服务组合研究和应用开发人员较快地了解相关领域知识及技术的参考资料。

图书在版编目(CIP)数据

语义Web服务组合关键技术研究/何丰著. —北京:科学出版社, 2012.5

ISBN 978-7-03-033983-6

I. ①语… II. ①何… III. ①网络编程—研究IV. ①TP393.09

中国版本图书馆 CIP 数据核字 (2012) 第 061218 号

责任编辑:韩卫军/责任校对:黄 嘉

责任印制:邝志强/封面设计:陈思思

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

四川煤田地质制图印刷厂印刷

科学出版社发行 各地新华书店经销

*

2013年9月第 一 版 开本:889×1194 1/32

2013年9月第一次印刷 印张:5 1/4

字数:130千字

定价:35.00元

前 言

Web 服务组合作为一种崭新的分布式计算模型，是 Web 上服务集成的有效机制。随着 Web 服务成为公认的实现服务的主流技术选择，这使得动态 Web 服务组合技术成为面向服务计算的核心技术。目前，许多国内外研究机构围绕动态 Web 服务组合技术展开研究工作，并取得了一些有价值的探索性成果。然而，Web 服务组合还有许多需要进一步研究的问题。根据 Web 服务组合的生命周期，展开动态 Web 服务组合关键技术研究，重点在于：Web 服务和 Web 服务组合的形式化模型，Web 服务自动组合的推理算法，Web 服务组合的正确性验证，Web 服务组合的异常处理。本书主要工作与贡献包括以下五个方面。

(1) 利用标签颜色 Petri 网 (LCPN) 方法对 OWL-S 描述的语义 Web 服务原子与复合进程的过程模型进行形式化建模，为建立动态的 Web 服务组合模型提供了有效的逻辑基础。然后，定义了一种基于层次颜色 Petri 网 (HCPN) 的 Web 服务组合模型 WSC_HCPN。通过自顶向下的服务分解方法支持层次化 Web 服务组合描述，并以图形方式表示 Web 服务间的逻辑依赖，清晰、直观地刻画了 Web 服务组合。

(2) 借助基于接口匹配的服务组合思想，将 Web 服务抽象成由输入和输出组成的实体，分析了各种服务依赖关系，建立了服务发布图。然后，借鉴了人工智能的思想，采用基于模糊推理 Petri 网的推理方法，通过反向、正向推理算法从服务发布图中提取出 Web 服务组合模型，实现了自动组合。由于采用了基于矩阵

运算的正向、反向推理算法，并且支持并行处理，提高了组合效率，适合于大规模、复杂情况下的 Web 服务自动组合。

(3) 提出一个基于 Petri 网的形式化验证技术，并对 Web 服务组合过程进行分析及验证，以保证服务组合的正确性和可靠性。此方法包含结构验证、可靠性验证两个部分。首先，利用转移矩阵 (Transitive Matrix) 分析方法对组合模型进行了可达性分析并检测是否有死锁，这种方法能有效地检验组合模型的结构正确性。然后，描述了基本 Web 服务的可靠性属性，将其引入到 Web 服务组合的模糊推理 Petri 网模型 FRPN 中以模拟和验证 Web 服务组合整体的可靠性。算例结果证实了该方法的有效性。

(4) 在对 Web 服务的事务类型、事务状态以及事务依赖进行形式化描述的基础上，提出了一种 ECA 规则驱动的方法以处理服务组合中的可预测异常，建立了相应的 Petri 网模型。模型的优点是具有可扩展性以及适应性。然后对于不可预测异常，本书提出了一个层次式的异常恢复算法以保证服务组合的语义原子性。该算法依据层次间的事务依赖和服务的各种不同恢复特性进行层次式的向上规约处理，动态确定补偿终止点。此算法主要目的是有效减小补偿域。

(5) 开发了一个基于颜色 Petri 网的 Web 服务组合原型系统 CPN4WSC。上述方法和算法的适用性和有效性在原型系统 CPN4WSC 中得到了验证。

本书目标是使用 Petri 网形式化方法，对 Web 服务自动组合研究领域中的几个关键技术问题进行探索与尝试，希望对推进 Web 服务组合技术的理论与应用研究具有一定的参考价值。

本书的研究工作大都是在我的导师乐嘉锦教授指导下完成的，他严谨求实的治学态度、勤奋朴实的工作作风、广博的知识使我于耳濡目染中受益匪浅。本研究工作期间得到了宁夏自然科学基金项目“可信 Web 服务组合关键技术研究”（项目编号：NZ0692）的经费资助。

本书的出版得到了浙江嘉兴学院出版经费的资助，科学出版

社为本书的出版给予了大力支持，在此一并表示衷心感谢。

由于作者水平有限，加之 Web 服务组合方面研究与应用的飞速发展，书中难免存在缺点与不妥之处，恳请广大读者和专家批评指正。

何 丰

2013 年 9 月于浙江嘉兴学院

目 录

前言

第 1 章 绪论	(1)
1.1 研究背景	(1)
1.1.1 Web 服务	(1)
1.1.2 语义 Web 服务	(3)
1.1.3 Web 服务组合	(5)
1.2 研究目标、思路及主要内容	(12)
1.2.1 研究目标	(12)
1.2.2 研究思路及主要内容	(14)
1.3 本书结构	(17)
第 2 章 基于 Petri 网的 Web 服务及服务组合建模	(19)
2.1 Petri 网的基本概念	(21)
2.2 OWL-S 进程	(26)
2.2.1 OWL-S 原子进程的 Petri 网描述	(30)
2.2.2 OWL-S 复合进程的 Petri 网描述	(33)
2.3 基于颜色标签 Petri 网的 Web 服务组合模型	(38)
2.4 本章小结	(43)
第 3 章 基于 Petri 网的语义 Web 服务组合	(44)
3.1 Web 服务组合中的依赖关系	(47)
3.2 基于模糊推理 Petri 网的 Web 服务 Petri 网 描述	(50)
3.3 基于 Petri 网的 Web 服务依赖关系图	(57)

3.4	基于模糊推理 Petri 网的 Web 服务组合	(60)
3.4.1	基于推理 Petri 网推理的基本概念	(62)
3.4.2	基于推理 Petri 网的 Web 服务组合两 阶段算法	(63)
3.5	基于推理 Petri 网的 Web 服务组合两阶段 算法示例	(65)
3.6	本章小结	(71)
第 4 章	可信 Web 服务组合的形式化验证	(72)
4.1	Web 服务组合的行为特性验证	(75)
4.2	Web 服务组合的可靠性验证	(85)
4.3	实验结果与分析	(95)
4.4	本章小结	(97)
第 5 章	Web 服务组合事务模型及异常恢复	(98)
5.1	Web 服务组合中 Web 服务的活动特性及 事务模型	(101)
5.1.1	Web 服务的活动特性	(101)
5.1.2	Web 服务组合的事务模型	(105)
5.2	Web 服务组合异常恢复的基本概念	(108)
5.3	基于 ECA 规则的 Web 服务组合异常恢复	(112)
5.3.1	基于 ECA 规则的 Web 服务组合异常 恢复模型	(112)
5.3.2	基于 ECA 规则的 Web 服务组合异常 恢复实例	(116)
5.4	基于层次式恢复算法 WSCRecovery 的 Web 服务组合异常恢复	(119)
5.4.1	基于层次 Petri 网的 WSCRecovery 异常 处理补偿模型	(119)
5.4.2	Web 服务组合异常处理恢复的 WSCRec- overy 算法	(122)
5.4.3	WSCRecovery 算法正确性分析	(126)

5.5 本章小结	(128)
第 6 章 Web 服务组合辅助工具 CPN4WSC 原型	(129)
6.1 基于 CPN4WSC 的 Web 服务组合系统模型	(130)
6.2 Web 服务组合辅助工具 CPN4WSC 原型系统	(132)
6.2.1 CPN4WSC 的设计	(133)
6.2.2 CPN4WSC 的功能模块	(136)
6.3 本章小结	(139)
第 7 章 结束语	(140)
7.1 主要工作和创新点	(140)
7.2 工作展望	(142)
参考文献	(143)

第1章 绪 论

近年来,面向服务的计算(Service-Oriented Computing, SOC)无可争议地成为计算机软件领域的研究热点,而随着 Web 服务(Web Service)标准与技术的不断成熟,Web 服务业已成为 Internet 上一种最为重要的计算资源,并极大地推动了 SOC 的发展与应用,为工业界和学术界的研究者所关注。SOC 的研究涉及许多研究领域,其中重要的一个课题就是 Web 服务及其组合技术,它已经成为 SOC 的核心技术,是松散耦合环境下分布式软件集成技术发展的新方向。本章首先介绍基于语义 Web 的服务自动组合问题的研究背景,概述相关的概念和研究现状;然后分析该问题涉及的主要关键技术并综述相关研究的现状;最后提出本书的研究目标并概括本书的主要工作,说明本书的整体结构。

1.1 研究背景

Web 服务组合是将多个 Web 服务按照一定的逻辑顺序组织起来构成新的、功能更丰富的面向服务的 Web 软件。在业务需求驱动的促使下,语义 Web 服务组合成为 Web 服务领域研究的重点,本书的工作正是围绕这一具有深刻技术背景和广泛应用前景的热点研究问题而展开的。

1.1.1 Web 服务

Web 服务是一种新的计算模式^[1],其基本思想是提供基于

Internet 的与开发环境无关的软件服务。它能较好地解决 Internet 上的数据相互孤立的问题,只要服务遵循 Web 服务体系结构,它们就可以相互调用,从而为基于 Internet 的数据集成奠定技术基础。

从外部使用者的角度而言,Web 服务是一种部署在 Web 上的对象/组件,它以标准的方式发布到 Web 站点,其他 Web 应用程序可以方便地找到并使用它们,从而协同完成某个商业逻辑^[2]。换言之,Web 服务是由企业发布的、完成其特定商务需求的在线应用服务,其他公司或软件可以通过 Internet 访问和使用该项服务。Web 服务具备以下特征。

(1) 完好的封装性。对于使用者而言,他能且仅能看到该 Web 服务提供的功能列表。

(2) 松耦合。Web 服务为企业应用提供了一个灵活的、松散耦合的分布式计算环境。同时,Web 服务屏蔽了底层的企业应用平台,具有良好的平台无关性,并且通信数据都是标准的 XML 数据,具有良好的语言无关性,为企业应用集成提供了一个便捷而快速的途径。

(3) 使用标准协议。Web 服务是构建在 XML 上的应用,其使用的协议也是开放的标准协议;对 Web 服务功能的描述采用标准的 Web 服务描述语言(WSDL);服务之间的通信采用 SOAP(简单对象访问协议),而 SOAP 本身是基于 XML、HTTP 和 SMTP 等标准协议的;在整个 Web 服务的体系结构中,Web 服务代理中心采用的是 UDDI(统一描述、发现和集成协议)。

(4) 高度可集成能力。由于 Web 服务采用简单的、易理解的标准 Web 协议作为组件界面描述和协同描述规范,完全屏蔽了不同软件平台的差异,无论是 CORBA(公共对象请求代理体系结构)、DCOM(分布式组体对象模式),还是 EJB(用于开发基于组件的企业多重应用标准)都可以通过这一种标准的协议进行互操作,实现了在当前环境下较高的可集成性。

1.1.2 语义 Web 服务

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。但是,现有的 Web 服务缺少明显的语义,Web 服务之间不能很好地理解其相互传递的消息,因而不能实现 Web 服务发现、执行和组合的自动化。

语义 Web (semantic Web)是由 WWW 的创始人 Berners Lee 在 2001 年正式提出的概念^[3]。语义 Web 研究的主要目的就是扩展当前的 WWW,使得网络中所有信息都是具有语义的,是能够被计算机理解和处理的,便于人和计算机之间的交互与合作。因而其研究的侧重点就是如何把信息表示为计算机能够理解和处理的形式,即带有语义。

语义 Web 服务 (semantic Web Service)是语义 Web 和 Web 服务的结合,是基于语义的智能化的 Web 服务^[3]。Web 服务以服务为目标,语义 Web 以计算机可理解性和可处理性为目标。综合这两方面的优点,在 Web 服务中加入语义的支持,以服务为导向,进行带有语义的服务描述和推理的研究,构建一种带语义的服务环境。基于语义的 Web 服务可以做到:

- (1) 明确描述和推理事务之间的联系和规则;
- (2) 明确描述 Web 服务所执行的任务,实现 Web 服务的自动发现;
- (3) 在 Web 服务的执行过程中进行监控,并随时自动进行调整,实现自动执行;
- (4) 将简单 Web 服务组合成为复杂 Web 服务,并实现自动组合。

目前实现语义 Web 服务的主要方法是利用本体来描述 Web 服务,然后通过这些带有语义信息的 Web 服务,来达到服务的自动发现、调用和组合的目的。语义 Web 和 Web 服务是语义 Web 服务的两大支撑技术,OWL-S (Ontology Web Language for Serv-

ice)是用 OWL 语言描述的语义 Web 服务的本体^[4],是连接两大技术的桥梁。

OWL-S 的本体结构包括如下三部分:服务轮廓(service profile)、服务模型(service model)和服务基点(service grounding)。使用 OWL-S 的 Web 服务架构如图 1-1 所示^[5],让 OWL-S 结合已有的 Web 服务构架(SOAP、WSDL、UDDI)可以实现 Web 服务的发现、执行和组合的自动化。

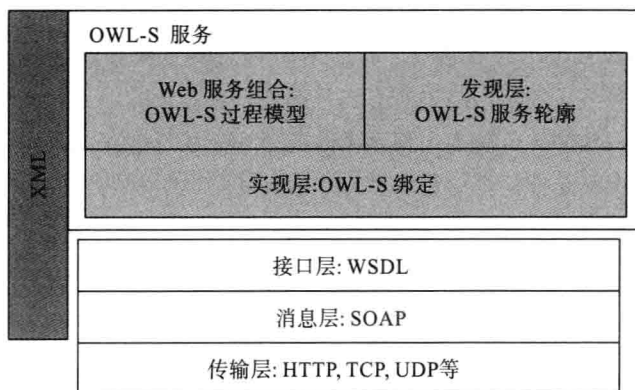


图 1-1 应用 OWL-S 的 Web 服务体系结构

语义 Web 服务的体系结构如图 1-2 所示^[3],它总共有三种基本的角色:服务提供者(provider)、服务请求者(requester)和服务注册/匹配中心(OWL-S/UDDI matchmaker)。其中由服务提供者发布自己的服务(service profile),并对服务请求进行响应;服务注册/匹配中心用来注册已经发布的服务,并提供语义搜索服务;服务请求者则通过服务注册/匹配中心查找/匹配所需的服务,然后使用该服务。

语义 Web 服务研究的主要内容就是 Web 服务的自动化。具体来说,它包括自动的 Web 服务发现、执行、组合及互操作和执行监控。为了实现语义 Web 服务,必须在语义 Web 的基础上对服务进行标注,即语义 Web 服务标记。它的内容主要包括:

(1) 服务性质和能力的说明性信息,这些信息主要用于自动

的服务发现；

(2) 服务的说明性 API，主要用于服务的自动执行；

(3) 服务的前提和结果的说明性信息，这些信息主要用于服务的自动组合；

(4) 用户约束和偏好的标记。

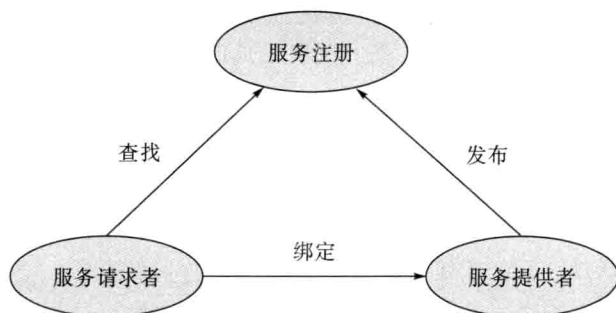


图 1-2 语义 Web 服务的体系结构

1.1.3 Web 服务组合

虽然 Web 服务给应用程序集成提供了方便，但是单个服务提供的功能有限，无法满足实际的需要。因此，只有对现有的服务进行组合，形成新的 Web 服务以提供更多的功能。例如，一个旅行服务通常是由订票服务、旅馆服务以及银行支付服务等组合而成的。随着 Web 服务的发展，Web 服务组合已经成为当前的研究热点。

所谓 Web 服务组合，指的是通过服务查找以及服务之间的接口集成，将多个自治的 Web 服务根据应用需要进行组合，从而提供新的、功能更强的 Web 服务。从“粒度”意义上讲，Web 服务组合是对 Web 服务进行更大规模的封装，并把该封装的结果作为一个 Web 服务展现给外界；从“顺序”意义上讲，Web 服务组合是一个满足用户需求的 Web 服务调用序列^[6]。

目前，已经有大量的研究者对 Web 服务组合进行了相当广泛和深入的研究并取得了许多研究成果。国内外学者对此进行了综

述^[7~14]，并从不同的角度对组合方法进行了分类。综合各种分类方法，本书根据 Web 服务组合方法所依赖的技术基础，将其归纳为基于业务过程驱动的组合方法和基于人工智能(AI)理论的 Web 服务组合方法，下面将对这两类方法进行介绍。

1. 基于业务流程的 Web 服务组合方法

这类服务组合方法是以业务流程为基础，将业务流程中的每一个环节分别选择和绑定 Web 服务，从而形成一个流程式的服务组合。这类服务组合的内部结构、服务之间的交互关系和数据流等全都受控于业务流程。基于业务流程的 Web 组合方法使用与经典工作流建模方法相类似的模型来描述组合服务，服务组合的功能被分解为由各个 Web 服务完成的活动，并通过描述流程中的活动以及活动之间的依赖关系刻画服务组合的实现逻辑。由于基于流程的服务组合方法直观地反映了组合服务的执行过程，并且易于实现相应的实际系统，因此在当前服务组合研究项目和原型系统中得到了广泛的应用。

(1) 基于工作流(workflow)的动态服务组合。Web 服务组合过程和工作流之间具有很强的相似性，因此，具有自适应性、跨企业集成的工作流技术同样适用于 Web 服务组合。此外，动态的工作流技术也提供了自动将抽象的节点同具体的服务或资源绑定的方法。组合 Web 服务的业务流程描述语言是业务流程集成的基础。到目前为止，已经出现了不少具有一定特色、但又有类似之处的语言，如 Web 服务的业务流程执行语言(Business Process Execution Language for Web Services, BPEL4WS/BPEL)^[15]、业务流程建模语言(Business Process Modeling Language, BPML)^[16]、W3C 的 Web 服务编排定义语言(Web Services Choreography Definition Language, WS-CDL)^[17]。这些语言所关注的应用领域不同，描述能力也有差别，但都提供了一种规范化的业务流程描述方式。

BPEL4WS 是专为整合 Web 服务而制定的一项规范标准。BPEL4WS 流程是一个流程图，用来表达特定业务的处理逻辑和算法，流程的每一步称为一个活动。BPEL4WS 的作用是将一组

现有的 Web 服务整合起来,从而定义一个新的 Web 服务。它能够将 Web 服务调用、操作数据、抛出故障或终止一个流程等不同活动连接起来,从而创造出复杂的流程。这些活动可以嵌套到结构化活动中,结构化活动定义了活动的运行方式和控制流程。但 BPEL4WS 不支持在应用运行时调整流程模型。

BPML 是 BPML 组织的规范, BPML 规范的目标在于利用业务流程管理系统的技术,促成标准的电子商务流程管理。BPML 提供了一个表示业务流程和支持实体的抽象模型,描述了全面的控制流和数据流结构,通过补偿活动支持短事务和长事务,也支持异常和超时处理。

WS-CDL(早期称 WSCI)是一种基于 XML 的语言,抽象定义了交互信息的类型以及信息交互的序列和条件,关注于可观察的服务与用户间的交互。WS-CDL 利用编排定义不同的流程参与者之间的点对点的协作。一个编排描述指一个多方的合约,从全局的视点描述多个客户间的客观可观察的行为。

文献[18]从业务协作建模、流程执行控制建模、组合能力和异常处理等方面,对 BPEL4WS、BPML 和 WS-CDL 进行了比较。从中可以看出, BPEL4WS 满足流程建模、异常处理、交易管理等多方面的要求,既侧重 Web 服务的执行流程,又侧重 Web 服务的信息交互,是相对较好的服务组合语言。

工作流技术对 Web 服务组合的研究产生了很大的影响,典型的以工作流技术为基础的 Web 服务组合系统包括 METROR-S、eFlow、SCM 和 SELF2SERV 等;特别值得一提的是 LSDIS 实验室研究开发的 METROR-S (METROR for semantic Web Service)^[19]。该项目首先采用 BPEL4WS 提供的控制流结构来创建过程流,在过程中通过服务模板来表示每个服务需求,指定过程约束以达成优化;其次对给定的过程和服务模板,服务发现引擎将返回和该服务模板相匹配的一个服务集合,根据约束关系选择最优的待组合服务;最后在运行时,将抽象的过程和服务模板转换成一个可执行的 Web 过程,表明实现了基于工作流的语义 Web

服务的动态组合。该项目的关键特性是在 Web 服务组合过程的整个生命周期中使用了 OWL-S 对 Web 服务进行语义描述,因此,它是一个语义 Web 服务组合的工作流。

(2) 基于形式化方法的 Web 服务组合。基于 BPEL4WS 的 Web 服务组合方法是一种朴素的组合服务模型观,易于理解。但是它是非形式化的流程模型,建模理论基础比较薄弱。某些研究借鉴了 workflow 技术成果,使用较成熟的形式化工具对 Web 服务组合的业务过程进行建模,采用“替代”的方法将业务过程模型中的相关组件最终替换成为具体的 Web 服务,从而获得有效的 Web 服务组合执行方案。引入形式化工具提高了对 Web 服务组合进行模型性质分析和验证的能力,目前在 Web 服务组合研究领域常采用的形式化工具有 Petri 网、进程代数等。

Petri 网是一种并发的形式化的模型,已经被广泛使用在建模和验证事务中^[20]。因为 Petri 网的语义是形式化定义的,“安全、死锁、可达”等定义在描述并发、冲突、同步等流程现象上极具优势,并具有直观的图形化描述等特点,因此比较容易将 BPEL4WS 进程映射到一个 Petri 网从而得到一个形式化模型。这种方法不仅提供了一个模型,而且它允许验证技术和工具在 BPEL4WS 进程语境中使用 Petri 网,因此被较多地应用于对 Web 服务组合的建模和验证。

进程代数是一种描述和推理进程行为、对动态实体进行建模的形式化语言,具有严密的形式化语义。著名的进程代数是 Milner 的 CCS^[21] (Calculus of Communicating System)、Hoare 的 CSP^[22] (Calculus of Sequential Processes) 以及 Pi 演算^[23] 等。其中 Pi 演算支持通道名称传递,适用于描述结构动态变化的并发系统,成为进程代数在 Web 服务组合研究中应用的代表。从组合的观点看, Pi 演算提供了以顺序的、并发的和条件的执行形式以及组合,这些能够构造任意复杂组合的活动。像 Petri 网一样,进程代数是精密的和值得研究的形式化方法,它允许自动地验证其行为的一定属性,同时提供了关于强模拟的丰富理论,如能够