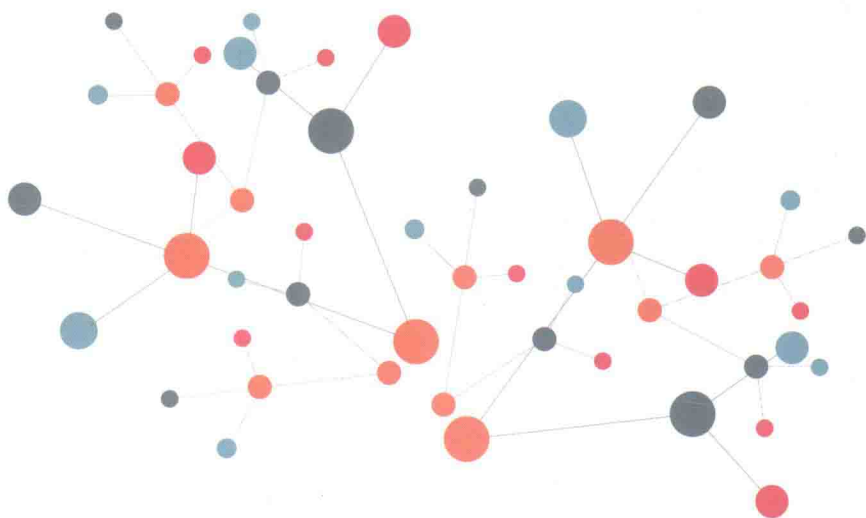


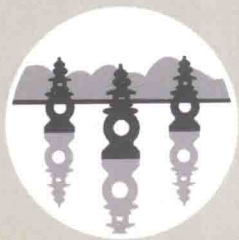
浙江省高校人文社科重点研究基地“管理科学与工程”项目资助出版



服务计算

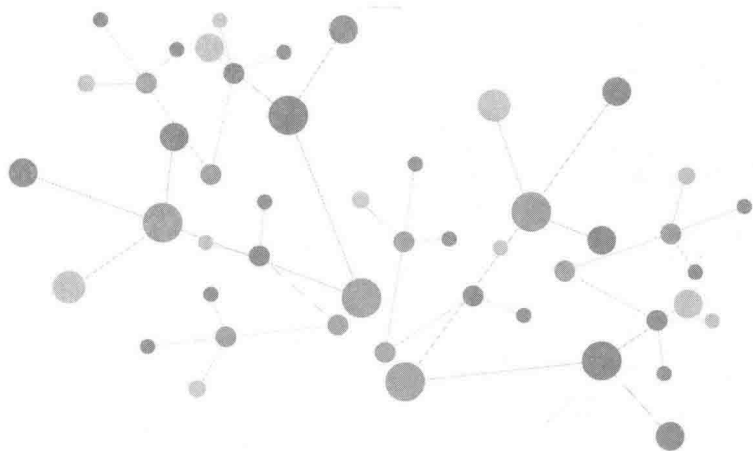
服务管理与服务组合流程

杨 洁 | 著



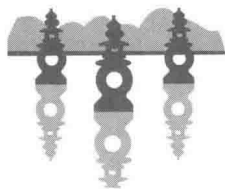
清华大学出版社





服务计算

服务管理与服务组合流程



清华大学出版社
北京

内 容 简 介

本书在全面介绍服务计算中的基本概念、关键问题、与云计算的关联的基础上,着重介绍在SOA 架构下 Web 服务研究中包含的两大关键问题——服务管理与服务组合,主要包括 Web 服务架构与标准、服务的分类管理、服务检索、服务选择、服务推荐等问题。

全书共分 10 章:第 1~3 章为对服务计算的基础知识的介绍,着重介绍服务计算的基本概念及关键问题,Web 服务中已成熟的架构与标准,以及 Web 服务与语义之间的关联;第 4 章为服务计算中关键问题——服务管理的介绍,并附有服务分类管理的各算法的实现过程;第 5~8 章为服务计算中另一关键问题——服务组合的内容,着重介绍集中式服务组合中包含的内容:服务检索、服务选择、服务推荐,并且每一章后都有相应的研究方法,以及在 Protégé 平台下服务组合的详细实验步骤;第 9、第 10 章为服务计算中的扩展内容,第 9 章为将 Agent 技术应用于服务组合问题的研究方法,第 10 章为云计算中涉及的服务安全问题。全书提供了各部分研究内容的实现过程以及相应的应用实例。

本书适合作为高等院校计算机、软件工程、信息管理与信息系统专业的高年级本科生、研究生的教材,同时也可供对服务计算比较熟悉并且对 Web 服务架构有所了解的开发人员、广大科技工作者和研究人员参考。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

服务计算:服务管理与服务组合流程/杨洁著. —北京:清华大学出版社,2017
ISBN 978-7-302-46466-2

I. ①服… II. ①杨… III. ①互联网络—网络服务器 IV. ①TP368.5

中国版本图书馆 CIP 数据核字(2017)第 026444 号

责任编辑:刘向威 李 晔

封面设计:文 静

责任校对:梁 毅

责任印制:刘海龙

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:保定市中国画美凯印刷有限公司

经 销:全国新华书店

开 本:155mm×235mm 印 张:17.5

字 数:270 千字

版 次:2017 年 6 月第 1 版

印 次:2017 年 6 月第 1 次印刷

印 数:1~1000

定 价:49.50 元

产品编号:067513-01



前言

服务计算为一门新兴的交叉学科,包含了多种突破性的技术。本书将其划分成两部分内容,分别为服务管理与服务组合。事实上,服务管理研究的目的是为服务组合过程提供技术支撑。对服务有效的分类管理有助于用户或系统快速、精确地查找服务。

服务组合中的服务是指 Web 服务,而 Web 服务的实现理念是基于面向服务的架构的,发布于互联网上的 Web 服务,可被互联网上的任何机构和人所调用。这些 Web 服务通常具有一些简单的功能,其本质是软件模块。单独的一个服务是无法对复杂问题求解的,此时需要将服务进行组合以完成对复杂任务的求解。因此,服务之所以具有重要的研究价值,是因为它可以被组合形成流程,用于复杂问题求解。最理想的目标是在求解问题时,从互联网上调用服务,将这些服务按逻辑操作组合建立形成流程。这些操作只需通过图形化界面,将它们建立连接并运行,不理解服务技术背景的人也可以完成。但是现在的服务组合技术的发展还远没有达到这种阶段。随着 Web 服务的发展,流程中另一重要的要素——人与工作流的交互也逐渐受到了重视,如 2007 年 Web Services Human Task(WS-HumanTask)规范的发布,它将与工作流交互的人需完成的任务定义为人工任务,包含了服务流程与人的交互。

本书以服务流程为研究对象,以服务的已有技术、标准为研究基础,通过应用相关的理论与技术,主要包括语义、本体、分类算法、TF-IDF 算法、命题库、粒子群优化算法、遗传算法、Agent 等,分阶段地探讨了服务

计算的关键问题——服务组合,以帮助实现服务流程。

本书的特色在于:将服务组合划分成多个阶段并分别做了介绍,且每部分的内容都有相应的提出的方法、实验与结论,分别为服务检索、服务选择、服务推荐,以及应用 Agent 技术构建服务组合流程的方法。在服务检索中,为了检索到满足用户功能需求的服务并提高服务检索的效率与准确率,在同一服务类别中,提出利用相似度建立服务间关系模型,并对个体服务的描述元素的内容进行语法结构分析,通过语法结构以及语义相似度以完成查询内容与检索服务之间的匹配。在服务推荐中应用用户组建形成的社会网络预测候选服务的信任值,为用户推荐满足其非功能需求的服务,还介绍了 Protégé 平台支持的服务解析与组合的实现过程。Agent 技术与服务的结合,则应用了 BDI 模型于分布式服务组合流程的构建。最后是对云计算模式下的未来发展趋势——多云架构的介绍。

本书的主要目标是在阐述服务计算的基础上,重点探讨服务管理与服务组合的内容及其实现技术。面向的读者是相关领域的研究人员、科技工作者、高年级的本科生与研究生。

本书出版得到了浙江省高校人文社科重点研究基地“管理科学与工程”项目(No. ZX140204304001/007)专著出版基金的资助,作者在此表示由衷的感谢。

本书在编写过程中参考了大量的国内外文献,在此感谢这些作者。

由于作者水平有限,书中难免存在一些不足之处,恳请读者批评指正。

编者

2016年9月



目录

第 1 章 服务计算	1
1.1 云计算	1
1.1.1 云计算的定义	1
1.1.2 云计算的特点	2
1.1.3 云计算的挑战	3
1.1.4 云计算与服务计算	4
1.2 服务计算	6
1.2.1 SOA 与 SOC	6
1.2.2 服务计算的关键问题	8
1.2.3 服务计算的挑战	9
1.3 服务管理	10
1.4 服务组合	13
1.4.1 服务组合的特征与分类	13
1.4.2 分布式服务组合	14
1.4.3 服务检索	14
1.4.4 服务选择与推荐	16
1.5 总结	18
本章参考文献	18

第 2 章 Web 服务架构与标准	21
2.1 Web 服务	21
2.1.1 定义	21
2.1.2 特性	22
2.1.3 实现平台	23
2.2 Web 服务流程的总体实现架构	25
2.3 Web 服务实现的支撑技术	27
2.3.1 XML	28
2.3.2 WSDL	29
2.3.3 SOAP	33
2.3.4 UDDI	35
2.3.5 RPC	36
2.4 Web 服务的描述语言	37
2.4.1 WSDL	37
2.4.2 OWL-S 与 WSML	38
2.5 服务质量	40
2.5.1 简单服务的 QoS 计算	43
2.5.2 组合服务的 QoS 计算模型	45
2.5.3 QoS 的应用场景	47
2.6 人件服务	49
2.6.1 定义	49
2.6.2 特性	50
2.6.3 描述方式	51
2.6.4 挑战	54
2.7 总结	55
本章参考文献	56
第 3 章 语义网与 Web 服务	57
3.1 语义网	57

3.2	本体	60
3.2.1	本体的定义	61
3.2.2	本体的建模元语	62
3.2.3	本体的描述语言	65
3.2.4	本体的建模工具	70
3.2.5	已有本体及其分类	72
3.2.6	本体的应用	73
3.3	本体和 Web 服务	74
3.3.1	Web 服务的本体语言	74
3.3.2	OWL-S 与 WSDL	76
3.4	语义相似性	80
3.4.1	语义和 Web 服务	80
3.4.2	语义相似性算法	81
3.4.3	语义在 Web 服务中的应用	86
3.5	总结	87
	本章参考文献	88
第 4 章	服务的分类管理	90
4.1	服务分类管理的提出	90
4.2	服务分类管理的过程	93
4.3	服务的分类步骤	95
4.3.1	服务的描述文档解析	95
4.3.2	服务预处理	97
4.3.3	服务的向量化	99
4.3.4	机器学习分类算法	102
4.4	服务分类实验	114
4.4.1	数据源	115
4.4.2	分类正确率	116
4.4.3	实验结果	116

4.5	实验结果分析	120
4.6	服务的输出预测	122
4.6.1	人件服务的输出预测	123
4.6.2	人件服务的输出预测结果分析	125
4.7	总结	126
	本章参考文献	126
第 5 章	服务检索	129
5.1	服务检索的提出	129
5.2	服务检索的过程	131
5.3	基于语义和语法的服务建模	133
5.3.1	基于相似度的服务间关系模型	133
5.3.2	个体服务的建模	138
5.4	服务与检索条件的匹配算法设计	143
5.5	服务检索实验	144
5.5.1	测量指标	144
5.5.2	实验步骤与数据	145
5.5.3	实验结论	146
5.6	总结	150
	本章参考文献	151
第 6 章	服务选择	153
6.1	服务选择的提出	153
6.2	服务选择的过程	155
6.3	服务选择中的 QoS	156
6.3.1	QoS 指标	156
6.3.2	信任值计算	157
6.4	服务选择问题的形式化	159
6.4.1	目标函数	159

6.4.2	QoS 约束条件	161
6.5	基于 QoS 与优化算法的服务选择模型	162
6.5.1	基于穷举搜索算法的服务选择求解	163
6.5.2	基于随机算法的服务选择求解	164
6.5.3	服务选择求解的优化算法——粒子群 优化算法	165
6.5.4	服务选择求解的优化算法——遗传算法	174
6.6	实验结论	179
6.6.1	服务选择优化算法的比较	179
6.6.2	信任在服务选择问题中的分析	180
6.7	总结	181
	本章参考文献	182
第 7 章	服务推荐	184
7.1	服务推荐的提出	184
7.2	基于信任网络的服务推荐	185
7.2.1	服务选择	186
7.2.2	信任及信任的度量	187
7.2.3	基于信任网络的服务推荐	189
7.3	信任网络模型	190
7.3.1	信任网络构成	192
7.3.2	网络特性	193
7.3.3	网络信任值计算	195
7.4	基于信任网络的服务推荐	200
7.4.1	服务推荐算法	201
7.4.2	用户间的间接信任值计算算法	202
7.5	服务推荐实验	203
7.5.1	实验环境	203
7.5.2	衡量标准	205

7.5.3 实验结果	205
7.6 总结	209
本章参考文献	210
第 8 章 服务组合	213
8.1 服务组合过程	213
8.2 服务组合实现	215
8.2.1 实验环境及开发工具	216
8.2.2 服务的统一分类管理	217
8.2.3 服务的操作解析	217
8.2.4 服务的组合实现	220
8.3 总结	228
本章参考文献	229
第 9 章 Agent 技术在服务组合中的应用	230
9.1 分布式服务组合	230
9.2 问题描述	231
9.3 Web 服务与 Agent	232
9.3.1 服务与 Agent 的特性	232
9.3.2 服务与 Agent 的结合	233
9.4 扩展的 BDI 模型	235
9.4.1 经典 BDI 模型	235
9.4.2 扩展的 BDI 模型	237
9.4.3 基于扩展 BDI 模型的架构	239
9.5 基于扩展 BDI 模型的服务操作与演化	241
9.5.1 服务的基本逻辑操作	243
9.5.2 流程动态演化操作	244
9.5.3 环境变化对操作的影响	247
9.5.4 定理与结论	248

9.6 实例分析	250
9.7 总结	254
本章参考文献	255
第 10 章 云计算中存在的问题	257
10.1 多云	257
10.2 多云平台下的研究方向	259
10.3 基于信任的多云安全机制	263
10.4 从多方面比较单云与多云	265
10.5 总结	267
本章参考文献	267

第 1 章

服务计算

服务计算是应用面向服务架构技术,消除了不同的商业机构、研究机构之间的鸿沟的革新性技术,且与云计算技术的发展形成了相互推动的关系。本章主要概括了服务计算的核心内容并对其形成总体认识,首先介绍了云计算的概念、云计算与服务计算之间的关联;其次,是对服务计算的核心技术 Web 服务的实现架构 SOA 的介绍,以及服务计算中的关键问题服务组合的认识;最后,介绍了围绕服务组合的主要研究方向:服务管理、服务检索、服务选择、推荐等。

1.1 云计算

云计算(Cloud Computing)概念的出现与崛起,吸引了一大批商业投资者的关注。它的理念拓宽了传统意义上对计算资源、硬件资源、软件资源等的认识。它的实现使用户可以随意地伸手到“云”中去获取资源,而不必受到本地资源的限制。“云”是无处不在的,随时、随地可扩展的,不受时间、空间的限制,且按使用收费。对它的使用可类比为,普通家庭里的水、电使用,只需向水、电公司按时缴纳一定的费用,便可随时、随地地使用水、电,而且收取的费用是低廉的。

1.1.1 云计算的定义

对云计算的定义尚未有共同的认识,简言之,它是许多计算思想的综合体,包括分布式计算、并行计算、网络存储、虚拟化、效用计算等。以下

列出一些较权威机构所给出的定义：

(1) 云计算是一种新兴的商业计算模型。它将计算任务分布在大量的计算机构成的资源池上,使各种应用系统能够根据需要获取计算力、存储空间和各种软件服务。——刘鹏,《云计算》,电子工业出版社,2011

(2) 云计算是基于互联网的相关服务的增加、使用和交付模式,通常涉及通过互联网来提供动态易扩展且经常是虚拟化的资源。(Cloud computing is a style of computing in which dynamically scalable and often virtualized resources are provided “as a service” over the Internet.)——维基百科,(Wikipedia)

(3) 云计算是一种按使用量付费的模式,这种模式提供可用的、便捷的、按需的网络访问,进入可配置的计算资源共享池(资源包括网络、服务器、存储、应用软件、服务),这些资源能够被快速提供,只需投入很少的管理工作,或服务供应商进行很少的交互。——美国国家标准与技术研究院(National Institute of Standards and Technology, NIST)

1.1.2 云计算的特点

云计算的快速发展,为其他技术概念的发展提供了可能,如物联网,物体与物体之间像互联网一样进行信息交换、通信,需要使用到云计算的虚拟化技术;大数据具有数据规模大、多样化的特点,如果只利用本地的计算、存储资源,可能无法满足用户处理数据的需求,可租借云服务商提供的强大的计算资源。以下给出云计算概念的特点:

(1) 服务化。正如 1.1.1 节中定义(2)所给出的,在云计算的理念中,计算能力、存储能力不再作为物理资源提供给用户,而将其作为“服务”。因而有软件即服务(Software as a Service, SaaS)、基础设施即服务(Infrastructure as a Service, IaaS)、平台即服务(Platform as a Service, PaaS)、一切皆为服务(every Thing as a Service, TaaS)等概念陆续被提出,这些内容将在 1.1.4 节详细介绍。

(2) 大规模。云计算是使计算分布在大量的分布式计算机上,而非本地计算机或远程服务器中,因而也体现了云计算包含了分布式计算的

思想。

(3) 虚拟化。计算、存储资源不再仅依赖于本地的计算机,而是可以按需从云中获取。只需在本地终端登录便可向云服务商请求获取应用服务,因而体现了虚拟化、网络存储的思想。

(4) 可扩展性。本地的存储与计算资源是有限的,但云是无限大的,当所需的存储容量不够大时,可请求扩展,云中的计算、存储资源是可伸缩的。

(5) 廉价。与用户个人购买物理存储设备相比,向云服务商租借硬件资源是廉价的,将本地想要备份的资料随时随地上传到云中。

(6) 安全。在享受云计算带来的优势与便利的同时,它的安全问题也给用户带来了担忧——随时随地地上传资料于云中,用户设置的秘密易被攻击,或者云服务商作为黑客攻击的对象,给用户带来的损失是难以估量的。

1.1.3 云计算的挑战

云计算包含的内容也是极为广泛的,许多厂商纷纷推出了自己的云计算平台,如亚马逊(Amazon)公司通过亚马逊网络服务(Amazon Web Services, AWS)^①,将它的大规模计算基础设施租借给开发者;苹果公司提供的云服务 iCloud 等;以及国内的阿里云、百度云等。云计算在发挥自身带来的优势与便利的同时,也存在一些缺点,如网络延迟或者中断,会导致客户无法立即获取自己存储到云端的数据,像这样的缺点都可以克服。云计算面临的主要挑战为:

(1) 缺乏互操作。主流的云服务提供商之间,目前并不存在交互,也就是说云与云之间是相互独立的。不同的云分别属于不同的公司,因而云与云之间要达成交互不可避免的是云服务商之间要形成合作,也就会牵涉到其他的商业问题,如云服务商之间的博弈等。但是,云与云之间的互操作,能够进一步扩展云计算的发展,以及对它的有效利用,可以提高

^① <http://aws.amazon.com>.

云的安全性,也就是接下来要讨论的云计算的安全问题。云之间的互操作可发展成为多云(multi-cloud,亦为 inter-clouds, cloud-of-clouds)的思想。多云有着更好的安全性,可分散云端的数据、应用存储的风险以提高云使用的可靠性,以及拥有更好的容错能力等,这将是云计算未来的发展趋势。

(2) 安全问题。现有的云服务商主要集中于几家大型的 IT 公司,且用户将自己的一些资料备份于云端,但是,如果这些公司的服务器发生故障,或者遭到黑客的攻击,用户的私人信息就会存在被泄露的风险,产生的后果就会很严重,因为在云端存储了大量的客户个人数据;或者公司内部的人员违背了职业道德,私自将客户数据流出,等等,这样一些情况都不在客户的掌控范围之内,而且一旦发生,其后果是很严重的。在文献[1]中,找到了亚马逊 EC2 IaaS 云服务的漏洞,并描述了对它的攻击方法。在文献[2]中,亚马逊弹性计算云(Elastic Compute Cloud, EC2)的缺陷被发现,基于 XML 的 soap 交互协议被找到漏洞,在消息传递过程中的签名验证易被篡改。文献[3]描述了谷歌出现的云计算服务安全问题,对于信息的共享,比如与某用户共享某文档,那么曾经共享过信息的所有用户都能看到此文档,这导致 2009 年 3 月发生了用户大批数据泄露事件——用户保存在 Google Docs 上的文档在用户不知情的情况下被共享。

1.1.4 云计算与服务计算

根据云计算的定义,它是将资源以服务的形式提供的计算方式。根据资源的不同形式,服务可划分为以下几种形式:

基础设施即服务(Infrastructure as a Service, IaaS),是指客户通过互联网从云服务商处租借基础设施。例如,服务器租用。如果开发者或者研究人员需要 100 台服务器来完成一项实验,在实验期间便可以向云服务提供商租借,当完成实验后,便可关掉它,之后也不需要再为租借支出费用。亚马逊网络服务 AWS 上提供了大量的可用服务,它包含的基本服务有:

- 简单存储服务(Simple Storage Service, S3),用于在网络上存储数据,可与其他 AWS 产品配合使用,按使用时间长短、存储容量大小收费。
- 弹性计算云(Elastic Compute Cloud, EC2),可配置服务器运行大多数操作系统,是一项镜像服务。用户可以在几分钟内启动镜像服务,运行自己想要的操作系统、应用程序,弹性即是体现在这里,与利用虚拟光盘镜像安装操作系统类似,它也体现了接下来将要介绍的平台即服务的形式。
- 弹性 MapReduce (Elastic MapReduce, EMR),提供的 Hadoop (MapReduce 框架的实现)支持在多个动态可扩展的 Amazon EC2 的机器上分布式的处理数据,与其他 AWS 数据存储服务 S3、EC2 配合使用。EMR 是一种 Web 服务,可以可靠地处理大数据且简化了大数据的处理方式。

平台即服务(Platform as a Service, PaaS),是指将服务器平台、开发环境、操作系统等作为一种服务,以 SaaS 的模式提交给用户,可将其理解为 SaaS 模式的一种应用。如上所述的 AWS EC2 便是 PaaS 的一个实例。

软件即服务(Software as a Service, SaaS),是一种通过互联网提供软件的模式,用户无须购买软件,而是向云服务商租用软件,降低了用户的软件使用成本;用户无须对软件进行维护,服务提供商会全权管理和维护软件,降低了用户的维护成本。而且软件厂商还向用户提供软件的离线操作和数据存储服务,用户可以随时随地使用其订购的软件和服务,节约了用户的本地存储的成本。在 SaaS 的模式下,Web 服务(Web Service)是一种更简单、更灵活也更实用的 SaaS,基于 Web 服务延伸出了服务计算问题,1.2 节将详细介绍。

一切皆服务(everyThing as a Service, TaaS),是指所有的事物都通过互联网作为服务提供给用户。IaaS、PaaS、SaaS 是将基础设施、平台、软件放到互联网上,TaaS 则是人们工作、生活所需要的所有事物都放到互联网上,作为“服务”提供给用户。当然,这一切还依赖于网络设施与速