

面向最终用户编程的服务虚拟化及其应用

王术娟 梁剑萍 李吉桂

(华南师范大学计算机学院 广州 510631)

摘要 随着 Web 服务的快速发展,如何实现最终用户编程从而达到应用的按需构造是迫切需要解决的问题。本文以“数字博物馆通用建设平台”项目为应用背景,对服务虚拟化方法进行了尝试,重点研究了虚拟化的运作机制,并基于服务虚拟化的流程给出相应算法。

关键词 服务虚拟化,最终用户编程,服务组合,服务转换

Service Virtualization for End-user Programming and its Application

WANG Shu-Juan LIANG Jian-Ping LI Ji-Gui

(School of Computer, South China Normal University, Guangzhou 510631)

Abstract With the development of Web services, the need of applying End-User-Programming technologies to achieve just-in-time application construction is becoming more and more pervasive. This paper based on the project of Universal Digital-Museum Construct Platform, tries the method of service virtualization, the paper studies the mechanism of service virtualization, and gives the algorithm of service virtualization which is based on the flow chart.

Keywords Service virtualization, End-User-Programming, Service composition, Service transform

1 引言

随着面向服务的架构(SOA)被广泛接受,如何在动态、开放的计算环境下构造、部署和使用服务,如何构造和组织大粒度的业务级服务,如何能让最终业务用户自行组装出面向服务应用,从而更好、更灵活地构造出个性化的、能适应业务变化和满足即时需求的网络应用等问题备受关注。通过在面向服务的应用架构中引入业务端编程范型以允许最终用户自主构造应用,能够有效地满足最终用户的个性化需求以及实现对多变的业务需求的快速响应。然而,让最终用户真正可以自行编程,仍然有很多关键问题未得到完全解决,比如如何提供最终用户可理解、易使用的业务级抽象,如何提供易于最终用户使用的编程模型等等。本文研究的服务虚拟化方法及技术旨在探讨面向服务计算环境下,业务级抽象的落实和透明访问等问题,以尽可能保障业务级资源的抽象可用能力,从而支持最终用户编程。

2 支持最终用户编程的服务虚拟化问题

Web 服务是目前流行的服务实现方式,它所具有的良好互操作性使其非常适合成为最终用户编程的基本元素。但

Web 服务规范存在大量的技术规范,最终用户很难理解,让最终用户直接基于软件层 Web 服务构建应用非常困难。虚拟化机制的目标跨越业务领域和 IT 领域之间存在的鸿沟,将业务服务映射为软件层面的具体实现,即通过一个或多个合适的 Web 服务完成业务服务的功能。

分析常见的虚拟化文章,发现存在如下不足:

(1)文章仅针对虚拟化过程中涉及的某一个方面讨论,缺乏具体实施虚拟化的整体流程。

(2)缺乏支持最终用户编程的服务关联方法。

(3)对 Web 服务抽象、服务转换等虚拟化技术的研究还不完善。

本文针对以上问题,借鉴文[1]中提出的“中间相遇”的思想规划出虚拟化的运作机制,并基于虚拟化的流程给出相应算法。

3 服务虚拟化运作过程

服务虚拟化的目标是将最终用户端编程所需的业务服务映射为一个或多个 Web 服务,从而完成业务服务的功能。服务虚拟化的基本原理图如图 1 所示。

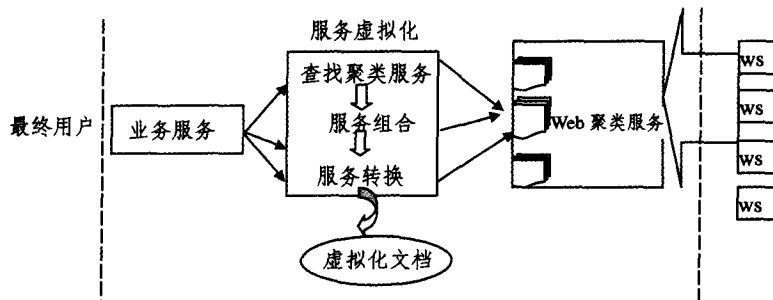


图 1 服务虚拟化原理图

首先对业务服务和 Web 服务都进行语义描述,然后通过聚类、组合、转换等操作以及服务匹配技术将具体 Web 服务关联到业务服务,并支持业务服务在运行时动态绑定到最优的 Web 服务去执行。具体的运作流程如图 2 所示。

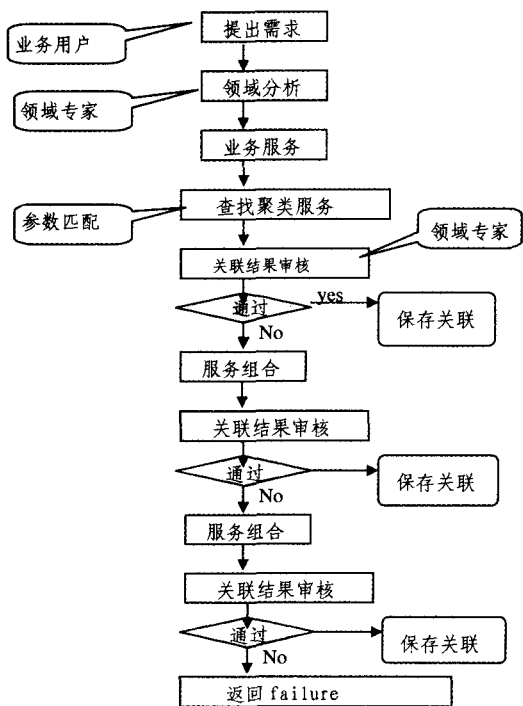


图 2 服务虚拟化流程图

业务端用户提出需求后,领域专家对需求进行领域分析,抽象成业务服务,首先采用参数匹配的方式将其与 Web 聚类服务进行匹配,对匹配结果进行审核。若满足,则保存结果,否则再对 Web 服务进行组合转换,重新匹配关联,审核结果。

例如,最终用户希望采用数字博物馆通用建设平台提供的藏品采编服务,如表 1 所描述。但是由于不同博物馆对数据采集所涉及的属性有所差别,平台只能提供最基本的特性,针对不同用户需要扩充输入选项以支持其个性化建设。因此,为了满足用户使用已有服务自行编程的目标,需要领域专家介入,对藏品属性进行分析,设定需要添加和删除的属性,采用虚拟化机制对服务进行转换,修改输出,提供给最终用户满意的选择。

表 1 藏品采编的服务描述

ServiceName	#藏品采编	Desc	藏品信息的采编
Input	#藏品类型	Ouput	#藏品属性

为了更清楚地描述整个关联过程,以及其中涉及的基本算法,给出如下定义和假设。

基本假设 服务提供者和领域专家基于相同的语义本体发布 Web 服务和业务服务。

定义 bs(业务服务)ws(Web 服务)WS(Web 服务集合)as(聚类服务)AS(聚类服务集合)

Boolean Clustered(ws, As):ws 是否已经聚类,若已经聚类结果为真,否则为假。

createAs(ws,as)根据 ws 生成 as

(1)服务聚类

根据服务输入输出参数关联的领域概念的相似程度对

Web 服务进行分组,从而抽象出聚类服务。

```

Public as Ws_Cluster(ws)
{
  If(存在 ws 属于 WS)and (clustered(ws, As)==false)
  {
    Foreach(asi 属于 AS)
    Sim=max{similarity(ws, asi)}
    If(sim>δ)
    {AS=As∪asi;}
  }
  Else
  {CreateAS(ws, asi);}
  Clustered(ws, AS)=true
  Return asi
}
    
```

说明:在 Web 服务中若有未聚类的服务,则寻找与之相似度最高的服务,将其加入。若没有符合要求的,就新建。Similarity(ws, as)在文[2]中有详细说明,δ 是阈值。

(2)服务匹配和绑定

将 Web 服务输入输出接口与业务服务功能概况侧面提供的输入输出信息进行匹配。若匹配成功,则表明该 Web 服务很有可能实现业务服务所描述的业务功能,在领域专家确认后与业务服务进行绑定。

具体步骤是:首先使用参数匹配建立关联,Web 服务模型基于 WSDL 规范描述,业务服务模型则采用业务服务规范去描述,为了解决这种异构元模型元素之间的转换映射关系,采取抽取本体概念的方法,该方法首先数据接口部分抽取本体概念,使用这些概念来存贮已有的匹配结果,这样可以充分利用已有的映射结果来提高匹配的成功率。

①对业务服务接口进行本体语义抽取,得到 $\phi(Binput, Boutput, Bnon_function)$

②对 as 接口进行本体抽取得到概念集 $\omega(Ainput, Aoutput, Anon_function)$

③针对每个 bs

```

Forall(as 属于 As )
{
  If compatible(bs. Binput, as. Ainput)>ε
  {
    If(compatible(bs. Boutput, as. Boutput)>ζ
    { If(compatible(bs. Anon_function, Anon_founction)>η
    Then dobinding
    }
  }
  else return o
}
    
```

说明:compatible 对参数匹配的结果进行评估,不同的参数匹配的衡量方式不同,因此对应不同的计算方法。当匹配结果都大于阈值时,专家确认对结果确认后执行绑定操作。

(3)组合和转换操作

当服务匹配不成功时,领域专家可借助于组合和转换操作来进行 Web 服务与业务服务的关联。Web 服务的动态组合是一项很复杂的工作。首先,Web 服务由不同的组织使用不同的概念模型开发,其运行方式不可能事先预知;其次,Web 服务能被动态地创建和更新,组合的系统必须能够动态地检测并自适应于这种变化。只有实现了自动组合,才有可能实现 Web 服务的有效重用。组合虚拟化是对一组具有控制逻辑协同关系的服务资源进行虚拟化,为这一组服务资源提供单一的使用接口,并对使用者屏蔽服务资源间具体的协同关系。

本文所采用的方式是基于本体的接口组合方法,由于服务可以抽象成由输入和输出构成的实体,Web 服务采用本体来描述,因此只要考虑本体中类和属性的匹配即可。在领域本体的支持下,首先找出服务请求的后继服务,然后找到该服务的后继服务,以此类推,直到找出的每个服务的后继服务是

服务请求,所得到的这个服务组合就是结果。由于有的服务可能存在多个后继服务,因此,这些语义关联的服务构成了一

个 Web 服务组合,参考文[3],可得到 Web 服务组合图(如图 3 所示)。

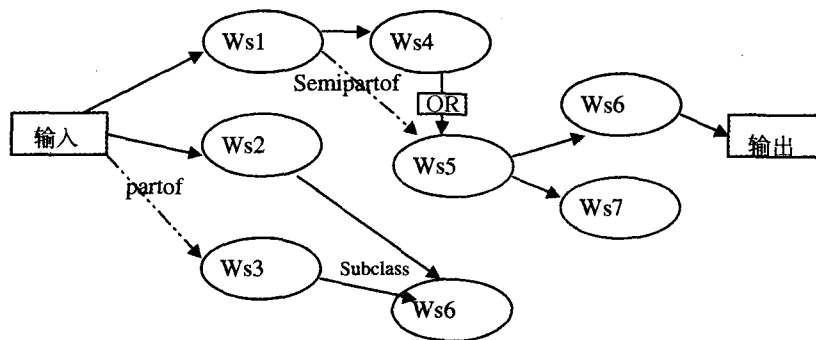


图 3 Web 服务组合图

因此,对于一个服务请求,对应的服务组合图中,从起点到终点的每条路径都是一个满足该服务请求的服务组合。为了得到最优的服务组合,可以采取“最优路径”算法。

Web Services 仍然是一个发展中的技术,其体系结构、具体技术的实现等都有待进一步完善。如 Web Services 中的服务发现、服务组合、可靠性、互操作性、事务性和性能优化等都存在一些语义上的问题,给 Web Services 的使用带来一定的困难。本体可以提供丰富的语义信息,为了解决语义问题,提出了语义 Web 服务的思想,即对语义 Web 服务中服务的属性、功能、接口和作用都以统一的机器可理解的形式进行编码,这种在 Web 服务中添加机器可处理的语义信息的方法解决了 Web Services 语义不足的问题,使得应用更加智能,更加接近人类的概念思维。

基于本体的思想解决信息网格中数据获取问题一直是研究热点。本文提出的基于本体的 Web 服务动态组合转换能够较大程度复用已有服务,生成优化的路径,能够有效解决信息网格中服务请求多样性和信息关联问题,同时保证了集成和重构服务的质量及效率。

4 应用

数字博物馆通用建设平台目标是创建一个博物馆数字化的通用平台,面向业务领域服务建模的数字博物馆通用建设平台设计方法,为博物馆(业务)人员在文博业务服务的支撑框架下构建和调整应用带来了很大的方便,使数字化博物馆建设用支持最终用户编程的方法实现变得简单。参考文[4]中业务服务规范建模的标准,建模过程如图 4 所示。

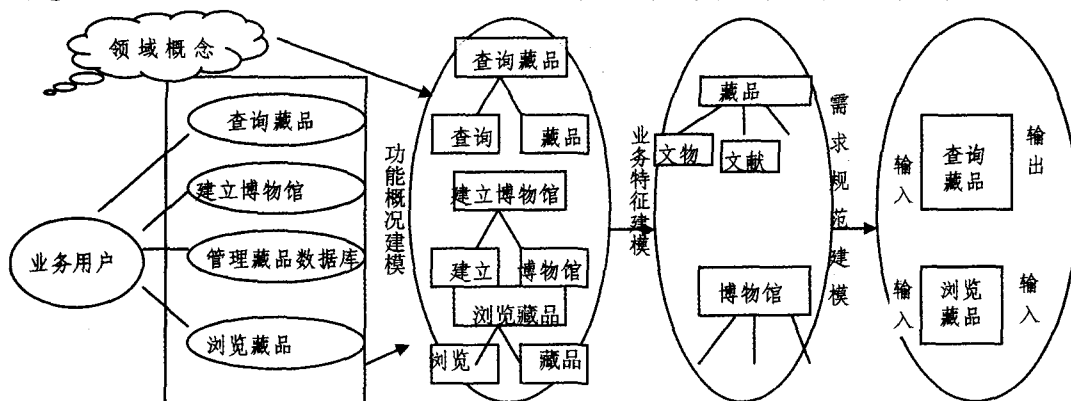


图 4 业务服务建模

(1) 构建领域概念规范

通用的数字化博物馆,使得最终用户可以利用平台顺利完成自己的博物馆的创建。具体实现为一组实体本体和动作本体,其中实体本体的概念有:藏品、展览馆,动作本体的概念有:建库、查询、浏览、管理。

(2) 业务服务规范构建

业务服务规范是通过业务逻辑领域的动作分析得到的。业务服务模型的构成可分为功能概况侧面、业务特征侧面、需求规范侧面和实现侧面。

①功能概况侧面通过描述服务涉及的动作及动作作用对象,以一种直观的方式对业务服务进行简单抽象:如在通用平台中,查询藏品,动作为查询,对象是藏品。

②业务特征侧面通过为每个业务服务创建一个特征模型,供业务用户了解业务服务的详细信息,并进行服务的个性化配置。与藏品有关系的概念有:藏品基本信息、展览馆、相关文献

等,从而又可进一步得到朝代、国别、名称等,如图 5 所示。

(3)需求规范侧面通过业务用户对特征模型的配置,实现对服务需求规范的配置,以表达不同业务用户对业务服务的接口信息和非功能属性的需求,提供 Web 服务个性化选择的依据,建立需求规范侧面各项概念与以上特征概念间的关系。具体包括输入输出接口、非功能属性和前提。

(4)实现侧面提供了业务服务与 Web 服务进行绑定所需的信息。该侧面主要描述业务服务执行时所需的信息,具体包括与该业务服务关联的 Web 服务和对这些 Web 服务进行的虚拟化操作。

例如:数字博物馆通用建设平台提供了查询藏品的服务,要求用户输入藏品名称或分类或朝代提供精确或模糊的查询,返回结果包含藏品的图像、藏品的详细文字介绍以及相关论文的名称等信息。首都博物院网站提供了资料检索服务,通过输入文章或书名,返回内容。上海博物馆网站提供检索

可以返回部分展品的虚拟展览,如表 2 所示。

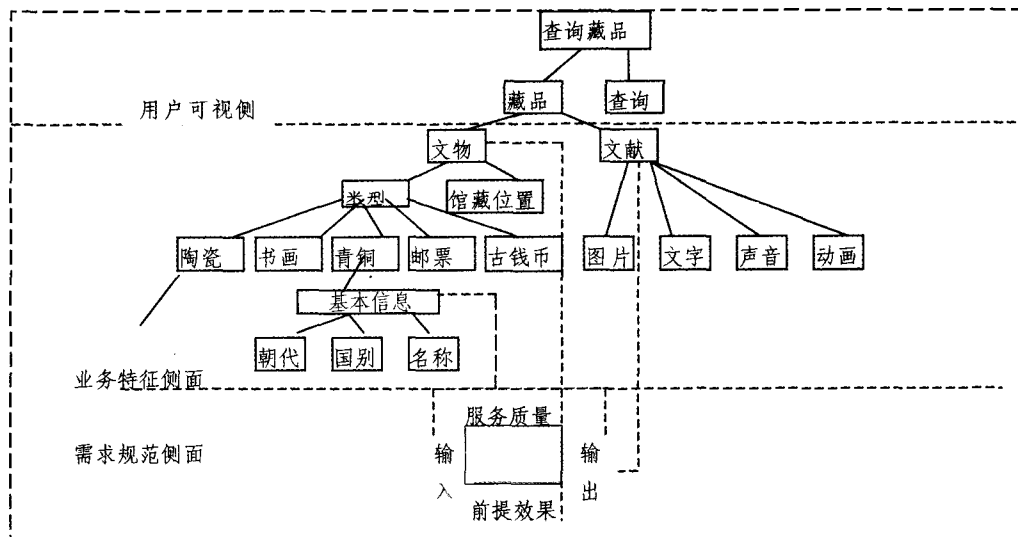


图 5 查询藏品业务服务

表 2 服务列表

Web 服务	输入	输出	提供者
# 查询藏品	# 藏品名称 Or # 藏品分类 or # 藏品朝代	# 藏品文字信息 # 藏品图像 # 藏品相关论文目录	通用平台
# 资料检索	# 文著名称	# 文著内容	首都博物院
# 虚拟展览	# 藏品名称	# 三维动画	上海博物馆

表 3 藏品出入库服务

Name	# 藏品入库	Desc	藏品信息的入库
Input	# 藏品名称或# 藏品类别	Ouput	# 藏品入库情况

若用户希望使用这个平台得到藏品的专著资料,以及该藏品的三维效果图,领域专家进行分析后发现独立的查询都满足不了用户的需求。因此可以在采用虚拟化机制,将三个 Web 服务进行组合转换,返回给用户一个包含所需信息的界面。

通用平台中的藏品出入库服务,其描述如表 3。藏品入库时首先应该先检索库中是否已有该物品。若没有,则藏品

入库,否则按情况修改。需要调用查询藏品的服务。因此可以使用服务虚拟化的方法将二者组合起来,提供给用户藏品入库的接口,屏蔽处理细节。

结束语 服务虚拟化机制的目的是利用已有的 Web 服务实现业务服务功能,保证业务服务能透明落实到 Web 服务,从而支持最终用户编程。本文通过研究虚拟化的运作机制,并对服务虚拟化相关方法进行尝试,结果表明,采用服务虚拟化可以有效降低最终用户构建应用的时间和难度。后续工作将深入研究基于虚拟化方法建立关联时的效率问题。

参考文献

- 1 赵卓峰,韩燕波,喻坚,等. 一种支持业务用户编程的服务虚拟化技术-VINCA 聚合服务机制. 计算机研究与发展,2004,41(12):2224~2230
- 2 Fan W, Wu Z, Yang J. Approximate Common Structures in XML Schema Matching. In: Proceedings of Web-Age Information Management Conference. Hangzhou, China, 2005
- 3 顾宁,刘家茂,柴晓路,等. Web Services 原理与研发实践. 北京:机械工业出版社,2006
- 4 林海略,刘晨,王建武,等. 面向业务领域的服务建模方法及支持框架. 信息技术快报,2006,4(3):10~17

(上接第 106 页)

节点的上传负载。

表 2 MSB&RT 算法备份性能测试结果

	主机 1	主机 2	主机 3	主机 4	主机 5
下载平均速度(kB/s)	—	4983.72	5632.96	6441.38	441.83
上传平均速度(kB/s)	757.91	243.85	339.03	406.95	220.67
总上传量/M	878.28	282.58	392.87	471.60	22.67

实验仅备份 512M 数据,在广域网环境中就花去 1 个半小时。如若网络条件更差的话,对于海量数据的消耗时间将是不可容忍的,异地多点数据备份将成为不可能完成的任务。使用 MSB&RT 算法,在现有网络环境下,即可以快速完成网络异地多点备份,而无需花费高昂的代价来建立专线网络。

结论 MSB&RT 算法实现了快速的网络异地多点备份和恢复。在备份过程中,源数据节点向各副本节点并行传送编码包,与此同时各副本节点之间也互相传送互补的编码包;在数据恢复传输过程中,多副本节点同时向原数据节点传送

相异编码包,提高数据恢复的速度及各个副本节点的利用率。MSB&RT 传输方法结合 LECC 编码,使得容灾系统在现有条件下,既能实现快速的海量数据备份和恢复,又能有效地提高数据的安全性和可用性,具有很高的实用价值。

参考文献

- 1 Bhattacharyya S, Kurose J F, et al. Efficient rate-controlled bulk data transfer using multiple multicast groups [J]. In: Proc. IEEE INFOCOM, 1998. 1172~1179
- 2 Byers J, Luby M, Mitzenmacher M. A Digital Fountain Approach to Asynchronous Reliable Multicast [J]. IEEE Journal on Selected Areas in Communications, 2002,20(8):1528~1540
- 3 Byers J, Luby M, Mitzenmacher M. Accessing multiple mirror sites in parallel: Using Tornado codes to speed up downloads [J]. Proc IEEE INFOCOM, 1999, 1: 275~283
- 4 万淑超,金蓓弘,黄宇. P2P 平台的关键技术[J]. 计算机科学,2005,32(6):21~24
- 5 张铁军,张玉清,战守义. Peer-to-Peer 典型应用安全需求分析[J]. 计算机工程,2005,31(20):56~58
- 6 李钟华,李伟华,武鲁. 网络灾备的传输补偿技术研究[J]. 计算机科学,2005,32(12):58~60