

基于 Web Service 的网络构件研究及应用

肖 刚, 陆佳伟, 张元鸣, 高 飞

(浙江工业大学信息学院, 杭州 310014)

摘 要: 网络构件是一种新的分布式软件模型, 是目前的研究重点。该文研究了网络构件的3个组成要素: 物理域, 个体域和整体域, 给出了其形式化定义, 分析了网络构件的基本特征, 提出了满足这些特征要求的基于 Web service 的网络构件系统构架, 并对其作了介绍, 结合具体实例给出了一个基于网络构件的应用系统。

关键词: 网络构件; Web service; 系统架构; 分布式计算

Research and Application on Network Component Based on Web Service

XIAO Gang, LU Jia-wei, ZHANG Yuan-ming, GAO Fei

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310014)

【Abstract】 Network component is a new distributed software component model and a hotspot research. This paper studies the structure of the network component, which consists of physics domain, individual domain and whole domain. The formalized definitions of the components are given, and whose basic characteristics are analyzed. To satisfy these characteristics, the system architecture for network component based on Web service is put forward. An application of the architecture system is given, which is composed of the network components.

【Key words】 network component; Web service; system architecture; distributed computing

1 概述

近几十年来, 软件技术经历了一系列重大变化和发展, 构成软件的软件实体粒度不断增大, 软件运行平台能力不断增强, 软件建模方式渐渐符合人类的思维方式, 软件技术的应用范围也不断增大。随着 Internet 的出现, 计算机技术产生了又一次深刻变革, 由于 Internet 具有无统一控制的分布性、单一节点的自治性、使用方式的灵活性等特征, 因此计算机硬件平台从集中封闭走向开放, 而作为计算机系统核心的软件系统, 必将是面向 Internet、开放且具有独立性、主动性和自适应性的一种软件形态。

网络构件是 Internet 开放、动态和多变环境中软件系统基本形态下的构件。传统的软件形态是基于实体驱动的, 而新的网络环境下的软件形态是基于 Internet 的协同驱动的。因此, 传统的软件理论、方法和技术等在处理网络构件问题时遇到了一系列的挑战, 难以适应或者不能适应开放、动态、无统一控制、自治的 Internet 平台。

网络构件是目前软件领域的研究热点, 研究主要集中在以下几方面: (1) 网络构件的软件体系语言(FSSCL)^[1], 它用来描述分布式网络构件与系统的耦合方式, 不关心软件的具体实现细节。国外提出了AT-EJB体系结构和HA-EJB体系结构, 国内浙江大学提出了CCCA和DFDM这2种并行算法^[2], 目的都是实现一个高性能、高可靠性的网络构件系统。(2) 对于网络构件间平台异构性和互操作问题, 利用XML存储数据、描述功能配置; 利用SOAP(simple object access protocol)以HTTP 承载消息、XML 格式化消息来实现异构系统的通信^[3]。(3) 结合Agent的网络构件研究, 如使用移动Agent技术实现一个高度智能和自治的网络构件。北京大学的杨芙清院

士则提出了网构软件的概念^[4-5], 从理论上对网构软件进行了阐述。

2 网络构件及其基本特征

2.1 网络构件的形式化定义

Internet 作为一个技术复杂的网络平台, 促使许多新应用领域出现。网络构件把软构件技术与 Internet 平台相结合, 为软构件在新的环境下提供了理论和技术, 其目标是充分发挥分布式计算的潜力, 使软构件具有较高的抽象层次以利于复用, 网络构件一经开发和部署, 无须本地安装, 即可供别的系统远程调用, 如同使用电力一样, 可以方便地使用网络上各种软构件资源。

网络构件概念概括了在 Internet 开放、动态和多变的网络环境下软件技术的一种发展趋势。根据定义, 网络构件包括3个基本构成要素: 物理域, 个体域和整体域, 它们构成了一个基本的网络构件, 见图1。

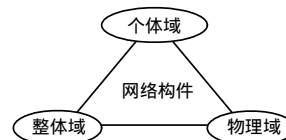


图1 网络构件3要素

基金项目: 浙江省自然科学基金资助项目(Y105200); 浙江省科技计划基金资助重点项目(2006C21003)

作者简介: 肖 刚(1965 -), 男, 教授、硕士, 主研方向: 软件复用, 信息集成管理, 电子商务, CAD/CAM; 陆佳伟, 硕士研究生; 张元鸣, 硕士、助教; 高 飞, 讲师、博士

收稿日期: 2007-01-22 **E-mail:** viivan@163.com

(1)在物理域上,网络构件是分布在 Internet 上支持分布式计算的松耦合构件,它的组成部分可以分布在网络的各个节点上,是可移植和跨平台的。

(2)在个体域上,网络构件具有自主性和自适应性,能根据需求的不同自发调整。当网络构件被其他软件实体搜索并请求时,会自动执行相应任务,即它们可以感知外部环境。

(3)在整体域上,网络构件能与周边的软件实体进行交互,支持静态连接和动态连接,且交互具有一定的屏蔽性,即受托任务的软件实体无须知道实现的细节,甚至可以在不知道该网络构件存在的情况下对其进行调用。网络构件在发生更新时还会自发通知与其作用的对象。

综上所述,本文对网络构件作了如下定义。

定义 1 用 IC 表示网络构件,那么 $IC = \langle P, S, E \rangle$, 其中, P 是 IC 在物理域上的要素集合; S 是 IC 在个体域上的要素集合; E 是 IC 在整体域上的要素集合。

定义 2 P 是物理域要素集合,定义 $P = \langle Sys, PD \rangle$, 其中, Sys 表示 IC 所属的系统; PD 表示 IC 的物理分布状态。另外, $PD = \langle f, PS \rangle$, f 是一个标识位, $f=0$ 表示网络构件分布于其他网络的各个节点上, $f=1$ 表示网络构件分布于本节点上; PS 是存储网络构件的路径集合,即

- (1) $IC \in Sys$
- (2) $f=0 \Rightarrow PS(IC_i, PIP)$; $f=1 \Rightarrow PS(PIP)$
- (3) $f=1 \Rightarrow PS(PIP) \Rightarrow PS(PIP) = Sys.PIP$

其中, IC_i 代表 IC 的各个组成部分; PIP 为物理分布地址。

定义 3 S 代表了 IC 的独立性和自主性, $S = \langle Inter_i, Fun_i, Env_i \rangle$, 其中, $Inter_i$ 为 IC 的接口; Fun_i 为 IC 的功能; Env_i 为 IC 接口所需的运行环境。在 S 内必然满足以下性质:

$$\forall Inter_i \in IC \Rightarrow \exists Fun_i \in IC \vee \exists Env_i \in IC$$

IC 中的每个接口都有相应的功能实现和运行环境,即当请求发送过来时, IC 会根据请求者提供的参数调用不同的接口完成正确的任务。

定义 4 E 中包含着 IC 周边联系的对象,即 IC 在新增或更新时需要自发通知的对象、 IC 执行任务时需要协调的对象等。 E 由 $Agent, OSys_i, Sys$ 这 3 个元素集合构成。

(1) $Agent$ 是一个网络构件代理(如注册中心),如果它与 IC 发生交互,则每个 IC 都将在该 $Agent$ 中存在一个映射关系,即 $\forall IC_i \Rightarrow \exists AgentIC_i \in Agent$, 表示 $AgentIC_i$ 为 IC 在 $Agent$ 中的映射;

(2) $OSys_i$ 代表一批与 IC 交互关系密切的系统,它们无须通过 $Agent$ 就知道如何调用 IC 。每个与 IC 交互的 $OSys_i$ 中都有 IC 的映射,即 $\forall OSysIC_i \in OSys_i \Rightarrow IC$, 表示 $OSysIC_i$ 为 IC 在 $OSys_i$ 中的映射;

(3) Sys 代表包含 IC 的系统本身,而每个 IC 在更新时必须通知其所属的系统。

2.2 网络构件的基本特征

由于传统软构件的软件技术体系本质上是一种静态和封闭的框架体系,在构建的过程中,它的基本功能和基本结构是按照系统所要解决的问题和问题领域的特征、通过系统分析与设计逐步确定的^[4],程序实体间传统的协同方式是过程调用和对象引用,由于其时间和空间耦合的特征,过程调用和对象引用通常只适用于程序实体功能固定、位置固定、协同逻辑固定的静态封闭世界,难以满足开放、动态和多变的 Internet 环境对时间或空间的紧耦合或松散耦合等多种协同方式的需要。

网络构件是基于 Internet 的一种软构件,它既是传统软构件的自然延伸,又有别于集中封闭式环境下产生的软构件,其具有以下 3 个特征:(1)通用性:由于网络构件的目标是支持分布式计算,且作用于整个 Internet 上,因此网络构件必须能方便地跨平台、跨语言、跨硬件地被其他异构系统调用。(2)重用性:以尽可能支持重用为目的,自适应地支持用户不同的需求。(3)易用性:为网络构件的使用者提供一种便捷的访问手段,使其无须了解网络构件实现的内部细节既可方便地调用网络构件提供的服务,以组装自己的系统。

目前支持软构件模型的主流技术有 EJB, CORBA, COM+ 等,其共同特点是:(1)对系统的同构性要求较高;(2)使用的远程调用协议都有自己的端口,当需要跨越防火墙时,请求很可能被防火墙拒绝;(3)基于这些技术构成的软构件之间的互操作问题难以得到有效解决。因此,采用传统中间件技术解决网络构件时,仍然有诸多不足之处。本文则提出了基于 Web service 技术的开发方案。

3 基于 Web service 的网络构件

Web service 是一种分布式计算体系结构,用于跨平台的程序和程序之间的通信服务,是 Web 上数据和信息集成的有效机制。Web service 技术具有强大的分布式计算能力^[6~7]:

(1)能够解决以前中间件跨防火墙通信所带来的问题,使网络构件具有广泛的通用性,真正实现跨越整个 Internet。

(2)通过对业务逻辑的有效封装、发布、查找和绑定机制,将提供者所生产的服务通过 UDDI 注册并发布,供请求者选择重用。服务可对不同粒度的应用逻辑进行封装,实现了系统级的较大粒度重用,提高了软件开发和部署的效率。

(3)使用 SOAP 让所有的构件模型、开发工具、程序语言和应用系统能够轻易地进行互操作,不会因为平台、操作系统以及构件模型的升级而无法使用。因此,Web service 能较好地体现网络构件的 3 个本质特征。

Web service 在网络上已经获得了广泛的应用。开放网络服务架构(open grid services architecture, OGSA)以网格计算和 Web service 技术为支撑,将资源、信息和数据统一起来,有利于实现灵活、一致、动态的共享机制。它是 Web 服务的扩展。图 2 展示了 OGSA 网络架构中的各个逻辑层。基础层包括高度可用的、虚拟的、自治的逻辑和物理资源,它们可以在网格中共享;中间层表示对基础层中定义的资源的功能的逻辑抽象;应用层内含一个或多个基于网格架构的面向服务的应用程序^[8~9]。

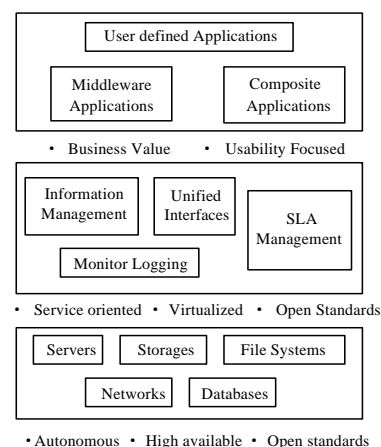


图 2 网络架构的概念视图

4 基于 Web service 的网络构件系统架构

在 OGSA 的架构中,所有网络资源(包括各种软硬件资源)都被建模为服务。而网络构件是让请求者访问网络中软件系统提供的功能,访问的主要内容是 HTML 文件或 XML 文件及系统底层数据。网络构件的软件构架不像以往那样将系统的运行开销全部由本地系统承担,而是将系统和计算分离,将真正的计算工作提交给服务器端构件完成,本地构件仅承担一些发送请求的工作(如发送调用请求或查询服务请求)。网络构件的架构和 OGSA 虽然服务的形式有所不同,但工作原理却有诸多相似之处。

借鉴 OGSA 的架构模型、结合网络构件本身的特点所形成的网络构件架构如图 3 所示。

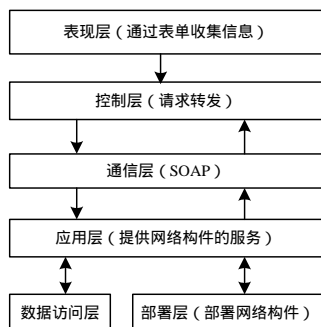


图 3 基于 Web service 的网络构件架构

整个模型由 6 层构成,从上到下依次是表现层、控制层、通信层、应用层、数据访问层和部署层。

表现层和控制层位于请求服务的系统内,表现层对用户界面提供支持,收集用户提交的请求信息。控制层对请求进行处理,负责转发请求到注册中心或远程请求调用网络构件。

通信层负责请求者和网络构件系统间的通信问题,而整个体系结构最基础的支柱是 XML 消息的传递。使用基于 XML 的 SOAP 协议进行通信,可以方便地在网络应用程序之间进行结构化数据交换。

部署层和应用层是整个网络构件架构的核心。部署层用 XML 与 Web 服务描述语言(WSDL)为网络构件指定标准的接口、行为与交互,将网络构件的功能描述为一种服务。部署层的重心在于网络构件的建立,主要包括 2 方面:

(1)注册中心的建立

在应用范围内建立一个私有环境下基于 UDDI 标准规范 XML 格式的注册表服务器。系统把自己的网络构件以服务的形式发布到 UDDI 注册表上,UDDI 根据这些信息生成一个 WSDL 文档并将相关信息写入自带的固有 XML 数据库中。

(2)网络构件的编码和注册

在系统内首先定义网络构件需要提供的功能,包括该构件的用途、对输入参数的要求、输出结果的形式、数据的使用范围等。在对该构件进行基本编码后,使用 WSDL 语言在 XML 文档中对其进行描述。WSDL 文件适用于说明可从不同平台、以不同编程语言访问的网络构件接口。在 WSDL 文件中定义了编写使用网络构件的程序所需的全部内容。客户会根据 WSDL 描述文档生成一个 SOAP 请求消息来访问构件。

应用层调用各种已定义的服务来组合构造各种应用,提供网络构件供别的系统调用。网络构件无须下载即可使用。无论是发送服务请求的系统还是用户都不需要知道实现细节。整个过程的处理如图 4 所示。

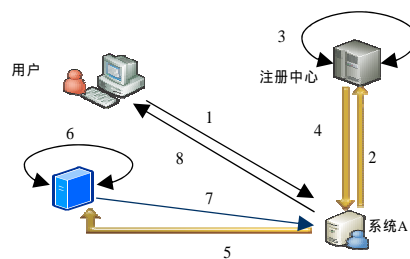


图 4 网络构件的使用

用户登录到系统 A,向其发出检索某批数据的请求(步骤 1)。由于系统 A 的数据库内不存在此数据,因此向注册中心发出检索服务的请求(步骤 2)。注册中心检索自己内部的服务,发现某个系统的网络构件提供符合要求的服务(步骤 3),于是向系统 A 发送以 WSDL 语言描述的服务(步骤 4)。系统 A 解析服务描述,找到提供该网络构件的系统的位置,然后按照服务描述的规范要求向该系统发送服务请求(步骤 5)。网络构件系统根据提供的请求参数在系统中进行计算(步骤 6),然后将计算结果返回到系统 A(步骤 7),系统 A 再将结果集反馈给用户(步骤 8)。采用网络构件后客户端被屏蔽,客户直观上会认为数据是由系统 A 提供的,而实际上系统 A 仅仅是提供一个请求,真正的工作交由网络构件完成。

如果对一些关系紧密的系统使用注册中心的方式来调用网络构件,会在一定程度上降低复用的效率,因此,允许这些系统直接向构件所在处发出请求进行调用。这种紧耦合关系可以通过 RPC 方式实现,以提高网络构件被调用的效率。

5 系统的实现

由于在开发浙江工业大学教师数据中心的实践过程中考虑到教师的许多信息(如教学情况、发表论文、科研工作等)需要被全校范围内的系统使用,因此利用 Web service 技术开发网络构件,通过网络构件封装数据、提供服务,以供其他系统调用。Web service 平台是一套标准,能使应用程序在 Web 上轻松实现互操作^[10],系统总体架构如图 5 所示。

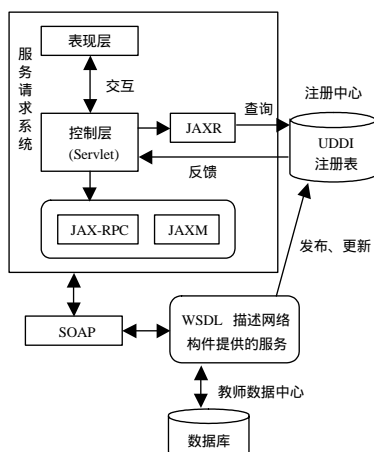


图 5 浙江工业大学教师数据中心总体架构

用户通过表现层提交对数据的请求,服务请求系统接收用户的数据并在控制层通过 Servlet 等方式过滤数据,如果发现请求的数据非本系统所有,则通过 JAXR 查询注册表,在得到注册中心的反馈后,再由 JAXM 发送 SOAP 消息到具有网络构件的系统,或者服务请求系统直接通过 JAX-RPC 发送 SOAP 消息调用网络构件提供的服务。

(下转第 107 页)