

## 面向动态工作流的 Web 服务组合模型研究

周 燕, 姜 浩

(东南大学 计算机科学与工程学院, 江苏 南京 210096)

**摘 要:**把 Web 服务与工作流相结合, 将 Web 服务作为工作流中的各项任务, 通过 Web 服务的动态组合技术, 建立面向动态工作流的服务组合模型。对 Web 服务描述进行了扩展, 提出了一个基于层次结构的服务发现模型, 支持 Web 服务开发与部署的分离, 并通过 Web 服务的动态查找, 对符合要求的可用服务器, 引入服务评价函数, 根据服务器服务份额、价格策略以及信任度等因素对服务器进行综合评定, 实现复合服务在运行时的动态部署与绑定。并通过工作流引擎对各项子任务的交互与完成进行相应控制, 最终实现 Web 服务的动态绑定模型的执行框架。

**关键词:**动态工作流; Web 服务组合; 组合模型

**中图分类号:** TP311

**文献标识码:** A

**文章编号:** 1673-629X(2008)01-0063-04

## Study on Dynamic - Workflow Oriented Web Service Composition Model

ZHOU Yan, JIANG Hao

(Dept. of Computer Sci. and Eng., Southeast University, Nanjing 210096, China)

**Abstract:** Combined with workflow, the Web services are assigned as the tasks in the workflow process. A dynamic - workflow oriented service composition model is established according to the dynamic composition technology. The description to the Web service is expanded, a hierarchy - based service discovery model is proposed. This model also supports the separation of exploration and disposition. The service evaluation function is introduced to the usable server by dynamic lookup of the Web service. The server is evaluated according to the service share of the server and the price strategy and the trust level in order to realize the dynamic disposition and binding of the combined service when running. Finally, the execution frame of the dynamic binding to the Web service is realized by the engine of the workflow which controls the interaction and the accomplishment of the subtask.

**Key words:** dynamic workflow; Web service composition; composition model

## 1 动态工作流模型

一个工作流包括一组活动(也称任务)以及它们之间的执行顺序关系、过程及活动的启动、终止条件、对每个活动的描述, 活动间的输入、输出数据<sup>[1]</sup>。一个工作流描述实际应用中的一个具体业务流程, 完成一个实际业务目标。

动态工作流系统是指组成工作流的任务组件在运行时才确定下来, 能够支持比较灵活的业务逻辑的实现, 在较短的时间内, 建立适应具体业务变化的动态工作流系统, 如: 适应变化的商业策略、组织策略, 能处理预见和不可预见的

各种事件, 能支持在运行时对过程模型的动态修改。

一个典型的动态工作流模型<sup>[2]</sup>如图 1 所示。

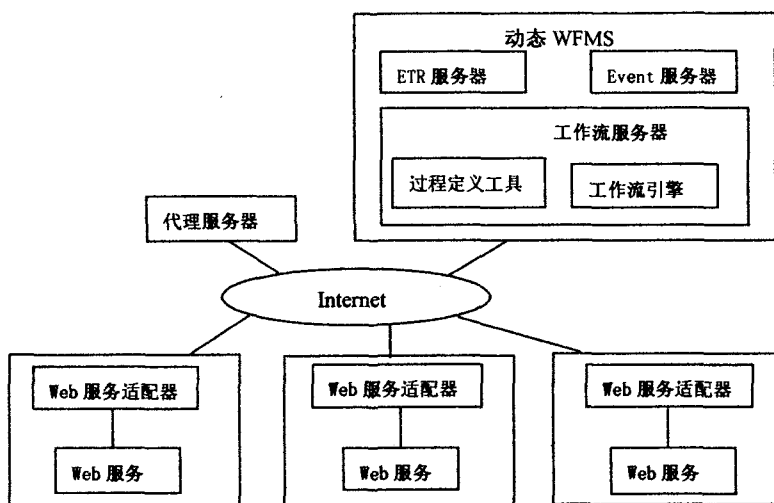


图 1 动态工作流系统模型

在此模型中, 对提供 Web 服务的不同企业, 统一安排服务适配器对 Web 服务进行封装, 在 Internet 上

收稿日期: 2007-03-22

**作者简介:**周 燕(1983-), 女, 江西萍乡人, 硕士研究生, 研究方向为工作流技术、Web 服务技术; 姜 浩, 副教授, 研究方向为工作流技术、Petri 技术、Web 服务。

通过 SOAP 和 HTTP 协议进行访问,这样,Web 服务就可以通过不同的开发方法来实现,而不用担心其它应用系统无法在 Internet 上对它进行访问。此模型中的代理服务器,用于作为服务的注册中心,提供服务的查找功能。 workflow 服务器包括两个功能组件:过程定义工具和工作流引擎。过程定义工具负责定义业务过程中的各项子任务,对业务过程进行模块化,并通过对相关服务组件的集成来实现;工作流引擎则负责在运行过程中,动态绑定相关服务的实现,对服务组件的运行进行调度。在该模型中,Event 服务器和 ETR(Event - Trigger - Rule)服务器的采用则使 workflow 系统具备了一定的主动性和自适应性。

## 2 基于层次结构的服务发现

采用基于层次结构的服务发现模型<sup>[2]</sup>,主要包括四个功能模块:服务用户、服务提供者、服务代理以及服务管理者,如图 2 所示。

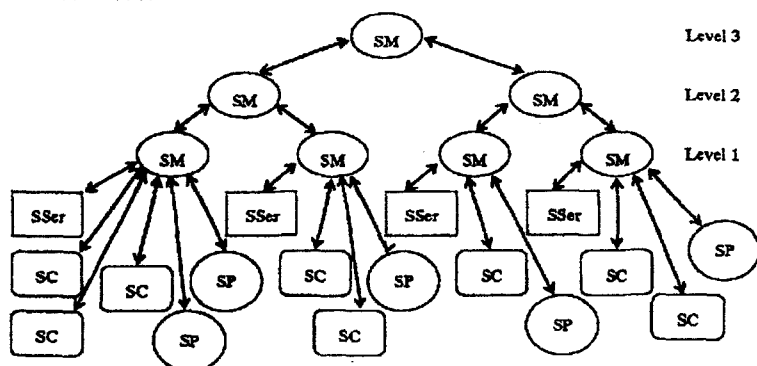


图 2 基于层次结构的服务发现模型

### (1) 服务用户。

服务用户 (Service Client, SC) 作为该模型中的服务使用者,向服务代理提出服务查询及使用请求。由于该层次模型通过域的方式进行组织,服务用户只向所在域的代理发出服务请求。这种服务请求采用 XML 消息格式,描述了服务请求的属性,如:服务名、服务目标、性能要求以及与服务相关的属性。

### (2) 服务提供者。

服务提供者 (Service Provider, SP) 在 Internet 上提供应用级分布式服务,它需要将它所能提供的服务注册到它所在域的服务代理。注册信息中除了服务的基本信息之外,还包含服务的安装配置信息,可以用来将服务部署在其它的服务器上运行;与此同时,还包含了服务的安全授权信息,描述了有权使用该服务的服务用户。通过向服务代理注册,服务提供者可以很容易地加入到该服务代理所在的域。

### (3) 服务代理。

服务代理 (Service Manager, SM) 作为层次模型中

的核心模块,是服务用户与服务提供者之间的桥梁。在每一个域中只有一个服务代理。服务代理以层次结构组织,在最高层上,会有一个服务总代理。一个父亲代理可能会拥有多个儿子代理,拥有同一个父亲代理的服务代理可以看作是兄弟关系。服务代理会维护一张本地服务表(如:UDDI),记录了所在域的服务提供者所提供的服务信息。在执行服务查找时,服务的查询信息会在服务代理之间传递。例如:某个域中的服务代理在本域中未找到相关服务时,会将服务查询信息提交到它的父亲代理,然后转到其它的域中进一步查询。服务提供者注册到服务代理上的服务信息会存储到服务管理者处,进行统一管理。

### (4) 服务管理者。

服务管理者 (Service Server, SSer) 储存了服务提供者所注册的服务描述信息。为了在分布式环境下实现服务分配和执行的灵活性和可扩展性,对服务描述进行了扩展,包括了服务安装运行的资源配置描述信息,并采用动态复制技术<sup>[3]</sup>,可以将服务的副本运行在多个宿主服务器上,为不同的用户需求提供服务。除此之外,服务管理者还提供了服务设计模块,可以按照一定的标准和规范,建立独立的服务,这种服务并没有真实地运行在某个服务器上,而是可以根据需要,动态部署在特定的服务器上运行。这样,服务的来源不仅仅是服务提供者所注册的服务信息,也可以是专门的服务设计机构提供的独立服务,通过这种方式,增加了系统的灵活性和可扩展性。

服务用户在向所在域的服务代理提交服务查询请求后,服务代理会根据服务的相关信息,在本地服务描述库中(UDDI)查找:

①如果在本地服务描述库中(UDDI)没能查找到相关服务信息,服务代理会将服务查询请求提交给它的父亲代理,再由父亲代理传递给其它的儿子代理进行查找,以此类推,直到查找到相应服务,否则,若不存在服务用户所需要的服务,则返回服务不存在的相关提示信息;

②如果在本地服务描述库中(UDDI)查到相关服务信息,则服务代理根据该服务的安全授权信息,判定服务用户是否具有使用权限,当服务用户通过安全验证之后,可以建立服务用户与服务提供者之间的联系,进行进一步的操作(如:协商等)。通过层次结构来组织 Web 服务,可以优化服务查找的效率,但在某些节点上(如:根节点)可能会存在一定的通信开销,造成一定的性能瓶颈。

### 3 面向动态工作流的服务组合

一个面向工作流服务可以看作是一种复合服务,它由若干个已存在的子服务通过服务组合技术构建而成,以完成特定功能的合成服务。这种复合服务的执行可看作是一个面向工作流的过程,通过构建它的子服务间的复杂交互来动态完成。这些子服务会被分配到异构的、地理上分布的服务器上运行,而整个运行过程可能会是一个比较长时间的过程。在此过程中,服务执行所需的资源属于不同的个体或组织,拥有各自不同的策略、访问方式以及价格模型等。

由于服务组合是在一种高度动态的环境下进行,对于组成复合服务的各子服务,需要进行动态的查找与即时的集成。在此处模型框架中,服务组合的过程可分为两种形式:一种是静态的服务 XML 描述的合成,是一个复合服务和其中一个子服务的 XML 描述(该复合服务称为“travel”,子服务为“airline”);一种是动态运行时的服务器绑定合成。前者根据服务的内在逻辑关系,按照一定的逻辑规则和时序关系以 workflow 管理方式将各子服务的 XML 描述以某种组织方式组合在一起,构成复合服务的 XML 描述,用于指导动态的服务绑定合成过程,这里,需要在复合服务的描述中加入对各个子服务的控制过程描述,以保证各个子服务按照工作流过程模型的执行顺序完成。后者则是在运行时,根据前者静态描述,动态查找和定位各子服务,通过事件驱动和消息传递机制实现数据交换,将各子服务动态绑定运行在可用的服务器上,进行实际的服务合成。复合服务建立后,可作为一般服务,储存于服务描述库中,提供给用户查找、发现和使用。

#### 3.1 Web 服务的动态绑定与执行的模型框架

动态复合服务的执行是一个工作流过程,每一步的服务绑定都会根据当前可用服务器的负载及所需开销情况动态实现以达到最佳服务质量(QoS)。如图3所示,在模型框架中,对注册的服务通过层次结构来组织<sup>[2]</sup>,服务器状态收集器负责收集本地已注册服务器

运行时的负载情况,以供服务调度器(service scheduler)来动态选择本地可用服务器进行服务绑定与执行。对本地域系统中每个已注册的服务器(由服务提供者提供),都会有相关的服务器描述,储存于服务器描述库(server description repository)中。在对服务器的描述中,记录了服务器(也可称作服务提供者)当前可用的服务份额以及所采用的价格策略等相关信息。此外,在模型框架中,会对不同的服务器进行信任度评价,这些信息可用来作为相关参数计算服务评价函数<sup>[2]</sup>。

#### 3.2 动态服务绑定与执行模型框架分析

服务评价函数(service evaluation function)是对服务器可用服务份额(available Service share)、价格策略(pricing policy)以及信任度(trust)等因素的综合评定。

##### 3.2.1 服务份额

服务份额是服务器资源在容量(capacity)上的一种抽象表示,如 CPU 的数目、缓存的大小、磁盘空间以及主存配置等。

##### 3.2.2 价格策略

价格策略是指服务器在一定条件下对于运行特定服务的要价,包括价格折扣信息以及对于每服务份额的报价信息等。

在 Internet 环境下,对同一个服务,会有多个服务提供者提供不同的实现方式,从而会有不同的报价。即使对于某个服务提供者而言,在不同的运行环境和条件下,对同一个服务也会有不同的报价,这取决于服务提供者所制定的价格策略。价格策略由服务提供者根据所提供服务的性质及开销来决定,是一组类似于“if <condition> then <pricing>”的规则,表示在不同的条件因素下,会有不同的报价结果。

价格策略分为全局策略和个体策略。全局策略是指服务提供者对所提供的某类(或所有)服务具有统一的价格策略,此时,价格策略的定义可以与特定的服务分离开来,单独设定;个体策略是指服务提供者对所提供的服务都有各自的价格策略,此时,价格策略可以通过 XML 描述<sup>[4]</sup>,定义在服务的 WSDL 描述中,作为服务的某个属性。

##### 3.2.3 信用度

随着广域网规模的不断壮大,越来越多的互不认识的企业通过 Internet 进行电子商务(13213)或组织业务生产(Virtual Enterprise)。在企业的合作过程中,信用度变得非常重要。无论是作为买方,还是卖方,都希望能有较好的信誉,吸引更多的企业与之合作,进行交易。

在 Internet 环境下所实现的复杂业务过程中,基

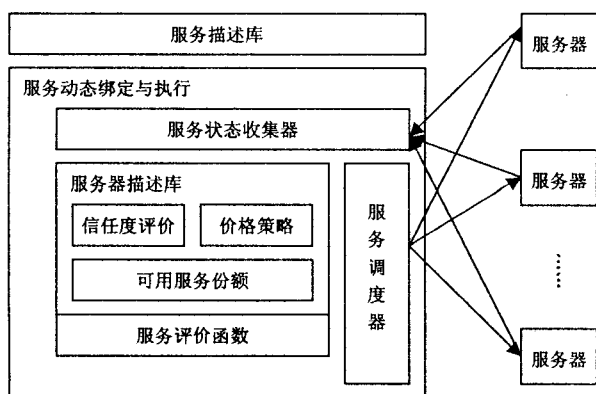


图3 动态服务绑定与执行

于信任的操作会占有相当大的比重,例如:一个服务消费过程需要通过一些基于信任的操作来建立对潜在服务供应商的信任,如会使用身份证明、信任书或者第三方机构的推荐信等相关信息来建立某个服务供应商的信誉档案。在选择服务供应商时,往往会根据与该供应商的历史交易记录来判定对该供应商的信任度。

文中,对信用度的定义如下<sup>[5]</sup>:“信用度作为一种抽象概念,是指实体的行为表现与所承诺的之间的匹配程度,信用度不是一成不变的,取决于一个实体的实际行为表现,且仅仅适用于给定时期的特定范围内。”也就是说,信用度是一个动态值,它的取值范围会从极信任(100%)到极不信任(0%)。信用级别建立在对过去行为表现的评价上,是由历史经验来判定的。信用度本身在一定程度上可以传递,例如,作为企业 A 的合作伙伴 B (B 在 A 处的信用度较好),可以将它所信任的企业 C (C 在 B 处的信用度较好,但 C 与 A 可能没有合作的历史记录)推荐给 A,从而在一开始就使 C 在 A 处的信用度较好。信用度具有一定的周期,往往随着时间的变化而变化。例如:企业 A 在企业 B 处的信用度与一年前的值会有所不同,可能会下降,也可能上升。

#### 3.2.4 动态绑定与执行分析

由于在模型中支持了 Web 服务开发与部署的分离,因此,在实际情况下,可能会存在某些服务还未运行在任何服务器上,对于这种没有处于运行状态的 Web 服务,可以在服务查找的过程中,根据服务描述中对服务运行环境的描述,选择符合服务运行要求的服务器进行动态部署和运行。此时,所找到的服务器必须具备运行该服务所需要的各种条件。

因此,我们所支持的服务查找有两种类型:

1) 查找符合服务运行要求的可用服务器进行服务的动态部署和运行;

2) 查找符合功能要求的服务(服务提供者)来提供相应服务,此时,不需要对服务进行部署和运行。

在 Internet 环境下,服务查找的结果是往往会有多个 Web 服务符合用户的要求,这时,就需要根据一定的选择策略从中选择出最符合用户需求的 Web 服务。这种策略反映在用户需要设定个性化参数,对服务的某些参数指标值进行进一步限制,服务提供者也可以根据实际情况动态调整服务的相关参数指标值和相应的价格策略,尽可能地接近和满足用户的需求,这个过程就是动态绑定的过程。

在分配子服务的过程中,有两种情况服务调度器(service scheduler)需要处理:

(1) 服务已经运行在某些服务器上(即:存在多个

服务器同时运行同一个服务);

(2) 服务没有运行在任何一个服务器上。

对于第一种情况,服务调度器会根据不同服务器的价格策略所计算出来的服务评价函数值来选择可用服务器提供服务,实现服务的动态绑定。

对于第二种情况,服务器调度器首先选择拥有运行服务所需份额的服务器,然后根据这些服务器对每份额的报价计算出服务评价函数值,根据服务评价函数值来选择最佳服务器,采用动态复制技术,将服务副本部署和运行在该服务器上,实现服务的动态部署与运行;与此同时,该服务副本的相关信息会注册到 UDDI,提供服务发现功能。复合服务的执行是一个工作流程,每一步的服务分配都会根据服务评价函数动态选择具有最优性价比的可用服务器来提供或运行特定子服务。在整个执行过程中,都会由 workflow 引擎根据复合服务的动态描述进行协调与控制,对每个子服务的运行进行调度,保证各个子服务间的交互与执行能够顺利进行。

## 4 结论与展望

服务合成是在一种高度动态的环境下进行,对于组成复合服务的各子服务,需要进行动态的查找与即时的集成。动态 workflow 管理系统能够在运行时修改它的过程模型以适应动态的商业需求和一些异常情况。文中着重讨论如何利用动态 workflow 模型,把 Web 服务与 workflow 相结合,将 Web 服务作为 workflow 中的各项任务,以动态 workflow 方式组合起来。在对当前服务组合模型进行分析的基础上,根据 workflow 过程模型中所定义的各项任务之间的业务逻辑与业务规则,建立面向动态 workflow 的服务组合模型,对 Web 服务描述进行一定的扩展,从而支持 Web 服务开发与部署的完全分离,并通过 Web 服务的动态查找,对符合要求的可用服务器(服务提供者),引入服务评价函数,根据服务器服务份额、价格策略以及信任度等因素对服务器进行综合评定,实现复合服务在运行时的动态部署与绑定;并通过 workflow 引擎对各项子任务(子服务)的交互与完成进行相应控制。

#### 参考文献:

- [1] Workflow M C. The workflow reference model. [WfMC 1003][R]. WfMC TC00 - 1003. Hampshire: Workflow Management Coalition, 1994.
- [2] Tang Jingfan, Zhou Bo, He Zhijun. Dynamic Economic - and QoS - Based Approach for Workflow - Oriented Distributed

(下转第 69 页)

$= \{[红], [黄], [蓝], [白], [黑]\}, [红] = \{u_1, u_3, u_8, u_9, u_{12}\}, [黄] = \{u_2, u_7, u_{10}\}, [蓝] = \{u_4, u_6\}, [白] = \{u_5\}, [黑] = \{u_{11}\}$ , 其中[红]为粒名,  $\{u_1, u_3, u_8, u_9, u_{12}\}$ 为相应的粒, 如果按其它属性分类, 依据上述描述, 以此类推。

表 2 多属性关系表

U	产品颜色 C	产品型号 M	产品价格 P
u <sub>1</sub>	红	B <sub>100</sub>	中
u <sub>2</sub>	黄	B <sub>200</sub>	贵
u <sub>3</sub>	红	B <sub>300</sub>	贵
u <sub>4</sub>	蓝	B <sub>400</sub>	中
u <sub>5</sub>	白	B <sub>100</sub>	便宜
u <sub>6</sub>	蓝	B <sub>100</sub>	贵
u <sub>7</sub>	黄	B <sub>500</sub>	中
u <sub>8</sub>	红	B <sub>200</sub>	贵
u <sub>9</sub>	红	B <sub>300</sub>	贵
u <sub>10</sub>	黄	B <sub>500</sub>	中
u <sub>11</sub>	黑	B <sub>300</sub>	中
u <sub>12</sub>	红	B <sub>200</sub>	贵

在实际应用中, 可以根据实际情况, 计算属性重要度。在这里, 为方便, 假设直接以产品价格为主要属性, 构造出一棵多变量的树 T, 如图 1 所示。

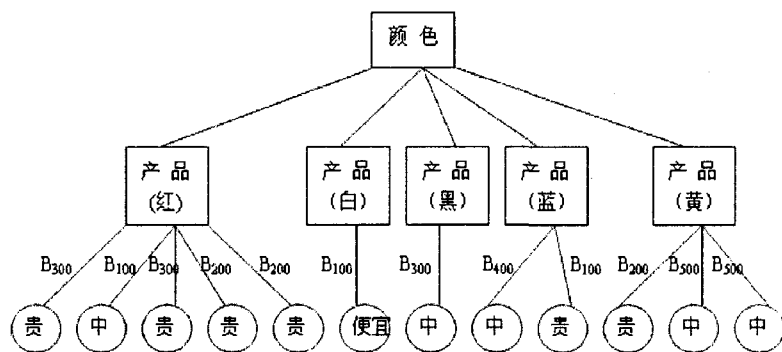


图 1 树 T

可以应用剪枝技术, 对树 T 进行剪枝, 在这里就不再赘述, 可以参考相关文献。对粒的描述可以按照此树结构及粒度形成的先后过程, 由连接树的根结点至叶子结点的路径分别给出, 对树 T 按照某一确定原则, 对其进行遍历, 获取相应规则, 如 if 颜色 = “红色” and 产品型号 = “B300” then 产品价格 = “贵”等。采用的硬件配置是 AMD Athlon64 处理器, 2400MHz,

512M 内存, 软件是 Windows XP 环境, 在软硬件环境相同的情况下, 应用 VC++ 6.0 进行编程, 进行有关对比实验, 与应用 ID3 算法进行相应实验, 实验结果一致, 另外, 应用多次属性约简, 数值计算相对变小, 说明该算法是行之有效的; 运算过程也相对简单。

## 4 结束语

基于粒度的知识获取方式是根据人的思维特性而提出的, 它与人认识客观世界的过程尽可能地保持一致, 是思维过程的自然反映。从粗糙集理论出发, 分析了知识粗糙性和知识粒度的关系, 即知识的粗糙度愈大对应的知识的粒度也愈大, 而得到的知识的信息量就愈小, 并提出了基于粒度原理获取粗糙性知识的一种算法。如何进一步提高粗糙性知识获取的效率, 还有待进一步深入研究。

## 参考文献:

- [1] 刘 清. Rough 集及 Rough 推理[M]. 北京: 科学出版社, 2001.
- [2] 王国胤. Rough 集理论与知识获取[M]. 西安: 西安交通大学出版社, 2001.
- [3] 苗夺谦, 范世栋. 知识的粒度计算及其应用[J]. 系统工程理论与实践, 2002(1): 48 - 56.
- [4] 苗夺谦, 王 珏. 粗糙集理论中知识粗糙性与信息熵关系的讨论[J]. 模式识别与人工智能, 1998, 11(1): 23 - 25.
- [5] 林镇彪. 知识粗糙性的粒度原理及其属性约简[J]. 武汉科技学院学报, 2006, 19(9): 31 - 34.
- [6] Liang J Y, Shi Z Z. The information entropy, rough entropy and knowledge granulation in rough set theory[J]. International Journal of Uncertainty, Fuzziness and Knowledge - Based Systems, 2004, 12(1): 37 - 46.
- [7] 梁吉业, 曲开社, 徐宗本. 信息系统的属性约简[J]. 系统工程理论与实践, 2001, 21(12): 76 - 80.
- [8] 顿毅杰, 张小峰, 孙 昊, 等. 一种基于粒度的规则挖掘方法[J]. 兰州理工大学学报, 2006, 32(1): 105 - 108.

(上接第 66 页)

- Services Allocation[C]//The Proceedings of the International Conference on Systems, Man and Cybernetics. Hague, Netherlands: [s. n.], 2004.
- [3] Byoung - Dar L, Jon B W. Dynamic Replica Management in the Service Grid on Hi Performance Distributed Computing [C]//10th IEEE International Symposium. San Francisco, CA, USA: [s. n.], 2001.
  - [4] Meng L, Krithivasan R, Sue S Y W, et al. Flexible Inter - en-

- terprise Workflow Management Using E - services[C]//on Advanced Issues of E - Commerce and Web - Based Information System. Proceedings of the 4th IEEE International Workshop, WECWIS. [s. l.]: [s. n.], 2002.
- [5] Azzedin F, Maheswatan M. Towards Trust - Aware Resource Management in Grid Computing Systems, on Security and Grid Computing[C]//Proceedings of First IEEE International Workshop. [s. l.]: Springer - Verlag, 2002.