

面向服务级别协议的核实与规划服务 质量系统框架研究

田红波, 董小社, 赵晓晔, 刘菲菲, 张兴军
(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对现有服务质量(QoS)保障服务系统模式效率低、可扩展性差、QoS自动控制不足等问题,提出一种面向服务级别协议的核实与规划服务质量系统框架.该框架包括服务规划、服务核实和服务执行3个功能模块,可有机地融合服务执行管理的系统功能和系统服务,并通过一种新的服务自动组合模式实现了系统监控与管理功能的优化,实现了QoS自动控制,由此解决了系统的可扩展性问题.实例与仿真实验结果表明,核实与规划服务质量的系统框架能够自适应地组合执行管理系统功能,使之适应服务执行的动态变化,满足自动组合带来的功能需求,最后达到了QoS自动控制的目的.

关键词: 核实与规划;服务组合;QoS自动控制

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2010)06-0001-05

Service Level Agreement Driven Verification and Programming Framework for Quality of Service

TIAN Hongbo, DONG Xiaoshe, ZHAO Xiaoyi, LIU Feifei, ZHANG Xingjun
(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A service level agreement driven system framework with quality of service (QoS) verification and programming is proposed to address the problems of the existing QoS-guaranteed service systems, such as low efficiency, bad scalability, and inadequate QoS automatic control. The framework is composed of Service Programming, Service Verification and Service Execution, and integrates functions of executive management with system services. The optimization of system monitoring and system management, and automatic QoS control are achieved through an automatic combination of services so that the problem of bad scalability is solved. Examples and experimental results show that the system framework for verification and programming of QoS can adaptively integrates functions of executive management for the dynamic needs of service composition, executive management and automatic QoS control.

Keywords: verification and programming; service composition; automatic QoS control

网络作为未来面向服务计算的基础设施,其目标是对开放网络环境中资源的有效管理和服务的有效聚合,是实现面向服务计算环境中高质量 QoS (Quality of Service) 的保障,因此 QoS 在网络环境中发挥着基础性的作用,已成为了许多科研团体的

研究热点.

针对如何保障 QoS,文献[1]基于 QoS 的资源管理把网络系统分为应用层、系统层和资源层.从3个逻辑层角度研究基于 QoS 的资源管理的方法具有普遍意义,但是系统层涵盖的内容过于庞大,没有

很好地区分资源管理与服务管理,不利于系统层管理的研究和设计.文献[2]提出了以 WSLA 为驱动并能够为用户提供不同等级的 Web 服务的自动管理系统,但由于系统的模块功能定义不明确,不利于系统扩展及开发,影响系统运行效率.文献[3-4]提出了一个服务执行管理系统,该系统能够协调应用服务与网络服务之间的合作,为用户提供 QoS 保障的服务.文献[5]提出了一种基于智能 Agent 的面向服务级别协议(SLA)的管理框架,该框架能够利用 Agent 交互来实现 SLA 的自动管理.文献[6]提出了一种基于智能 Agent 模块的 QoS 诊断框架,但是 Agent 组织模型、形成方法以及规则的构造过程极为复杂,使得基于 Agent 的 SLA 管理框架难以推广和使用.这些已有的保障 QoS 服务系统主要由各服务提供商独自提供,缺乏统一的系统描述,其兼容性较差,很难有机地组合成一个更大的有效系统来满足跨越服务组合的监控及其管理需求.针对以上问题,本文提出一种面向 SLA 的核实与规划服务质量的系统框架,该框架有机融合执行管理系统功能和系统服务,通过一种新的服务自动组合模式来实现系统监控与管理功能的优化及 QoS 自动控制,具有良好的可扩展性.

1 核实与规划服务质量的系统框架

核实与规划服务质量的系统框架为服务提供、服务协商、服务监控等提供一个自动服务管理运行系统,以一种新的服务自动组合模式来核实与规划服务质量,实现 QoS 自动控制,为用户提供有 QoS 保障的服务,满足 On-Demand 和透明的使用服务需求,其功能模块包括服务规划 P(Programming)、服务核实 V(Verification)和服务执行 E(Execution).

1.1 核实与规划服务质量系统框架功能模块定义

服务规划 依据用户 QoS 和系统 QoS 对服务组合、协商和执行,以及对意外事件处理进行规划,确保用户和系统需求得到最大满足.

服务核实 作为对服务以及系统状态的监控操作集合,从 3 个不同角度进行定义.

(1)核实内容分为用户 SLA 和系统 QoS. 用户 SLA 用来表示系统能够为用户提供的有 QoS 保障服务的能力,而系统 QoS 则表示系统能为系统资源提供的有 QoS 保障服务的能力. 用户 SLA 与系统 QoS 有时存在不可调和的矛盾,因此服务提供商需要在用户需求与资源提供商需求之间权衡.

(2)核实方式分为确定性核实与非确定性核实.

确定性核实是对系统准确状态数据的核实,而非确定性核实则为对系统随机数据进行核实.

(3)核实周期可划分为服务执行前核实、服务执行中核实和服务执行后核实.

服务执行 按照规划的要求受理用户提交的作业,管理和执行服务并管理任务队列,例如等待队列、运行队列和挂起队列.

在图 1 所示的核实与规划服务质量框架中,用户根据自己的需求与服务规划签订 SLA 并提交作业. 服务规划受理用户的请求,根据 SLA 的要求规划服务执行策略. 为了确保规划的效率,服务规划在规划策略的过程中需要得到服务核实数据支持,依据服务核实得到的参数对服务进行规划以及在运行时进行动态优化调整. 服务核实响应服务规划的请求,对任务执行以及系统状态情况进行核实,为规划提供核实结果以及为用户提供粗粒度的核实数据. 服务执行接受服务规划的命令调用资源,按照规划的要求受理任务并执行任务. 服务执行在执行任务的过程中接受服务核实的监控.

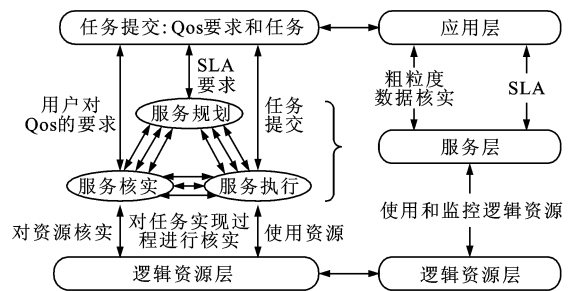


图 1 核实与规划服务质量的系统框架

1.2 核实与规划服务质量系统框架的自构建

服务组合 协调和组织多个服务构造新的服务或应用系统的技术,其目标是提高服务组件及基本服务的可重用性和利用率,减少新应用的开发时间和费用. 与传统定义不同,服务组合对象包括服务规划组合、服务核实组合、服务执行组合和 SLA 组合.

(1)服务规划组合是多个服务规划功能模块构造造成 1 个新的服务规划功能模块的过程,可描述为 $P=p_1 \wedge p_2 \wedge \cdots \wedge p_n$;

(2)服务核实组合是多个服务核实功能模块构造造成 1 个新的服务核实功能模块的过程,可描述为 $V=v_1 \wedge v_2 \wedge \cdots \wedge v_n$;

(3)服务执行组合是多个服务执行功能模块构造造成 1 个新的服务执行功能模块的过程,可描述为 $E=e_1 \wedge e_2 \wedge \cdots \wedge e_n$;

(4)SLA 组合是多个 SLA(设为 s_i)构造造成 1 个

新 SLA(设为 S)的过程,可描述为 $S = s_1 \wedge s_2 \wedge \cdots \wedge s_n$.

QoS 自动控制 自动构建服务满足用户需求,同时自构建与之配套的监控管理系统,以实现 QoS 自动控制,其模式分为 SLA 组合和 PVE 组合.

(1)SLA 组合:原子服务核实、原子服务规划和原子服务执行构造成一个 s_i ,这个过程称为功能匹配组合,然后通过 s_i 的组合来实现 S ,这个过程称为功能增强组合,可描述为,如果 $v_i, p_i, e_i \vdash s_i$,则 S .

(2)PVE 组合:原子服务核实、原子服务规划和原子服务执行首先通过功能增强组合构成功能更强大的服务核实、服务规划和服务执行,然后通过功能匹配组合直接完成用户 SLA,可描述为 $P = p_1 \wedge p_2 \wedge \cdots \wedge p_n, V = v_1 \wedge v_2 \wedge \cdots \wedge v_n, E = e_1 \wedge e_2 \wedge \cdots \wedge e_n \vdash S$.

如图 2a 所示,在 SLA 组合中,服务提供的原子服务 p_i, v_i 和 e_i 首先构成原子 s_i ,然后 s_i 有机组合实现用户 SLA;如图 2b 所示,在 PVE 组合中, p_i, v_i 和 e_i 首先分别组合成 P, V 和 E ,然后 P, V 和 E 构成 S .

随着面向服务计算技术的广泛应用,提供同一功能的服务越来越多,QoS逐渐成为影响服务组合

优化结果的主要因素^[7-9]. QoS 一般包括费用、时间、成功率等指标,这些指标往往存在矛盾,很难同时得到满足,例如服务时间越短,往往费用越高,所以系统需要对这些指标进行权衡.由于服务执行开销时间和费用一般是用户最关心的指标,因此本文以这两类 QoS 指标为参数,分别给出 SLA 组合与 PVE 组合的数学表示.

SLA 组合的数学表示为

$$\sum_{i=1}^n \min f(s_i) = \sum_{i=1}^n \min f(p_i + v_i + e_i)$$
$$\text{s. t. } \sum_{i=1}^n C(p_i + v_i + e_i) < C_0 \quad (1)$$

PVE 组合的数学表示为

$$\min f(P) + \min f(V) + \min f(E) =$$
$$\min \sum_{i=1}^k f(p_i) + \min \sum_{i=1}^m f(v_i) + \min \sum_{i=1}^n f(e_i)$$
$$\text{s. t. } \sum_{i=1}^k C(p_i) + \sum_{i=1}^m C(v_i) + \sum_{i=1}^n C(e_i) < C_0 \quad (2)$$

式中:函数 f 和 C 分别表示完成服务所需时间开销和费用; C_0, k, m 和 n 分别表示用户能够接受的最大费用、服务组合所需服务规划数量、服务核实数量和服务执行数量.

SLA 组合和 PVE 组合的区别在于两者执行功能匹配组合和功能增强组合时先后次序不同.前者先执行功能增强组合后执行功能匹配组合,而后者与之相反.这两种组合的求解过程为系统最优化问题(NP 问题^[10]),依据约束条件 C_0 的不同,又可分为无限约束条件问题(C_0 为无穷大)和有限约束条件问题(C_0 为有限定值).为了研究 SLA 组合和 PVE 组合的实现特性,本文重点研究 C_0 为无穷大时两种组合的求解算法, C_0 为有限定值情况下两种组合的求解算法将是我们下一阶段的研究内容.

SLA 组合求解算法的计算过程如下.

步骤 1:检查用户提交的服务集,如果不为空,则转到步骤 2,否则转到步骤 3.

步骤 2:计算每个 s_i 的最小执行时间,找出执行时间最小的 s_i 并记录该 s_i 中的 p_i, v_i, e_i ,然后转到步骤 1.

步骤 3:累计记录中的 p_i, v_i, e_i ,并计算所有 s_i 的执行时间总和.

PVE 组合求解算法的计算过程如下.

步骤 1:检查用户提交的服务集,如果不为空,转到步骤 2,否则转到步骤 3.

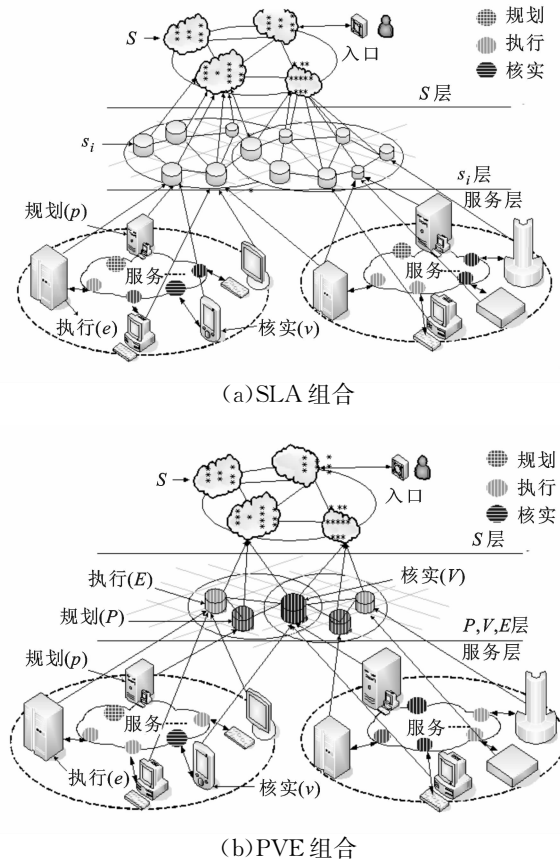


图 2 QoS 自动控制模式

步骤2:依据用户提交的服务类型,分别找出执行时间最小的 P 、 V 、 E ,并记录其中的 p_i 、 v_i 、 e_i 以及执行时间,然后转到步骤1.

步骤3:累计记录中的 p_i 、 v_i 、 e_i ,并计算所有 P 、 V 、 E 的执行时间总和.

1.3 核实与规划服务质量系统框架实例

本节以“旅行计划”Web 服务组合为例阐述核实与规划服务质量系统框架的实现过程及特性.为了完成计划,用户需要预定机票 s_1 、预定旅馆 s_2 和搜索景点 s_3 ,构造如图3所示的服务组合.

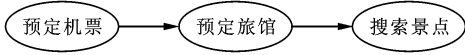


图3 服务组合实例

假设服务提供商提供如下服务

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 15 & 9 & 22 \\ 13 & 27 & 11 \\ 23 & 34 & 17 \end{bmatrix}$$

$$V = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} = \begin{bmatrix} 24 & 16 & 34 \\ 9 & 17 & 14 \\ 33 & 40 & 29 \end{bmatrix}$$

$$E = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} = \begin{bmatrix} 55 & 43 & 66 \\ 79 & 58 & 64 \\ 80 & 86 & 75 \end{bmatrix}$$

$$\begin{bmatrix} s_1 & s_1 & s_1 & s_2 & s_2 & s_2 & s_3 & s_3 & s_3 \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{21} & e_{22} & e_{23} & e_{31} & e_{32} & e_{33} \\ v_{21} & v_{13} & v_{12} & v_{32} & v_{13} & v_{31} & v_{32} & v_{22} & v_{21} \\ p_{13} & p_{11} & p_{32} & p_{33} & p_{22} & p_{31} & p_{12} & p_{22} & p_{21} \end{bmatrix}$$

采用 SLA 组合,得出服务结果集为 $S = s_1 \wedge s_2 \wedge s_3 = (p_{13}, v_{21}, e_{11}) \wedge (p_{22}, v_{13}, e_{22}) \wedge (p_{21}, v_{21}, e_{33})$,开销总时间为 302;采用 PVE 组合,得出服务结果集为 $P = p_{12} \wedge p_{21}$, $V = v_{21} \wedge v_{12}$, $E = e_{12} \wedge e_{22} \wedge e_{31}$,开销总时间为 246,其中 v_{21} 和 p_{12} 分别执行 2 次.

如以上实例所示,SLA 组合求解算法为 e_{22} 和 e_{33} 分别匹配了服务规划 p_{22} 和 p_{21} ,由于 p_{22} 和 p_{21} 属于同一类型服务,因此可以采用开销时间较小的 p_{21} 代替冗余 p_{22} 以获得更小时间开销.产生这种现象是因为 SLA 组合首先采用功能增强组合,然后实行功能匹配组合.这种组合模式使得系统在执行第 2 阶段功能匹配组合时,可更加灵活地选择服务核实 (v_i) 和服务规划 (p_i),从而保证在服务满意度较低

的情况下实现较高的服务组合成功率,但该组合模式会产生服务核实 (v_i) 和服务规划 (p_i) 冗余现象.与 SLA 组合求解算法不同,PVE 求解算法得到的服务结果集中没有类型重复的服务,从而可以消除冗余现象以获得时间开销的优化,如 PVE 组合的时间开销为 246,而 SLA 组合的时间开销为 302.但是,在 PVE 组合结果集中, v_{21} 和 p_{12} 分别执行了 2 次,这表示系统需要 v_{21} 和 p_{12} 分别能够匹配 2 种不同的 e_i .这种匹配方式对原子服务功能要求比较高,从而会降低服务组合的成功率.2 种组合模式最终会产生不同的服务选择结果集.

2 仿真系统及实验测试

2.1 仿真系统和改进 VPMIn-Min 算法

为了分析网格核实与规划服务质量系统框架在服务监控管理中的性能,本文运用该框架改进了文献[10]中网格模拟器,并实现了如图4所示的服务管理仿真系统.

服务管理仿真系统分为 5 个功能模块:服务规划 Agent、服务执行 Agent、服务核实 Agent、服务产生模拟和服务请求模拟.服务产生模拟以随机概率的形式决定服务的加入或离开,以及动态调整服务的 QoS,并在此基础上建立服务队列.服务请求模拟以 Poisson 分布的到达率动态模拟用户提交的服务 QoS 请求,并建立 SLA 等待队列.为了使仿真系统更加贴近实际系统,仿真系统模拟的服务请求由不同数量的子服务请求组成,并且这些子服务请求有不同的 QoS 需求.服务核实 Agent 负责对服务动态变化情况以及服务执行情况监控.服务规划 Agent 根据服务核实 Agent 所核实的服务队列情况对 SLA 等待队列中的 SLA 进行分配,以及对服务执行进行管理.SLA 与服务匹配成功后形成匹配队列等待服务执行 Agent 的执行.服务执行 Agent 负责执行服务,并服从服务规划 Agent 的管理,以及接受服务核实 Agent 的监控.

为了方便地比较系统的性能,本文采用文献[10]中定义的用户满意度来表征综合 QoS 的高低

$$U_i = w_l l_i + w_s d_i + w_c c_i \quad (3)$$

式中: w_l 、 w_s 、 w_c 分别为逻辑资源 QoS (l_i)、安全 QoS (d_i) 和记账 QoS (c_i) 的权重; U_i 为该服务总的用户满意度.本文在基于 QoS 的 QoSMin-Min^[10] 的基础上增加了核实与规划过程,形成了核实与规划服务质量 Min-Min 算法 (VPMIn-Min),伪码如下: QoSMin-Min

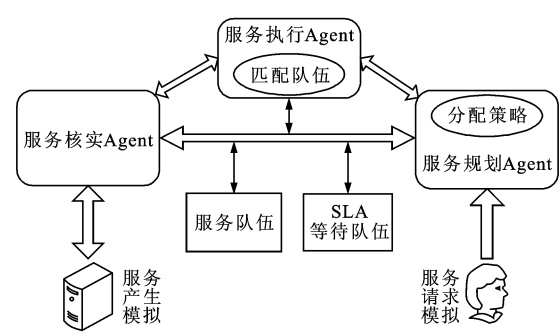


图 4 服务管理仿真系统

while 存在失效任务 t_i
 根据式(1)计算出 U_i
end while

根据 U_i 从大到小排列任务集合,生成从大到小的任务队列

QoSMin-Min

2.2 实验测试

为了测试核实与规划服务质量系统框架的性能,本文采用服务请求成功率^[10] (w)作为系统性能高低的衡量指标, w =得到满足的服务请求数总和/服务请求数总和。

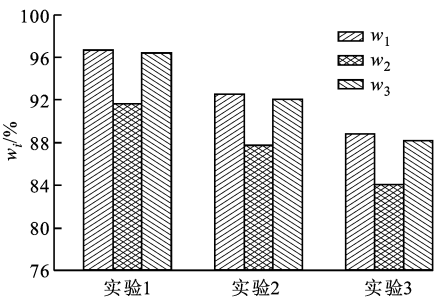
以文献[10]的网格资源管理仿真系统及其基于QoS的改进 Min-Min 算法为参照做了如下对比实验。

w_1 为文献[10]的网格资源管理仿真系统及其基于 QoS 的改进 Min-Min 算法下的服务请求成功率; w_2 为在文献[10]的网格资源管理仿真系统环境下,同时已执行的服务按 5%的概率失效时的服务请求成功率; w_3 为在服务管理仿真系统下,同时已执行服务按 5%概率失效时的服务请求成功率。

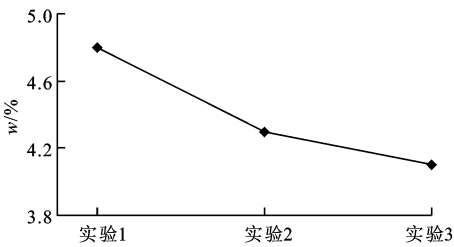
为了对比在不同的网格服务构成情况下核实与规划服务质量系统框架的执行管理性能,本文采用了 3 种不同的实验场景^[10]: 75%的网格服务的 QoS 水平大于其平均 QoS 需求;50%的网格服务的 QoS 水平大于其平均 QoS 需求;25%的网格服务的 QoS 水平大于其平均 QoS 需求。服务执行 Agent 以 5%的随机概率模拟服务失效,本文共模拟了 500 个网格服务节点和 2 000 个网格服务请求,并重复进行了 100 次对比实验,结果取其平均值。

服务管理仿真实验结果如图 5a 所示,当 w_1 中已执行的服务按 5%的概率失效时, w_2 出现明显下降,而在 w_3 中,当已执行的服务按 5%的概率失效时, w 能够得到有效的补偿以获得较高的值。在图 5b 中,随着网格服务的 QoS 水平大于其平均 QoS

需求的比例下降,服务管理仿真系统所提高的服务请求成功率呈现下降趋势,这是因为仿真系统采用核实失效服务请求后,需重新规划执行的补偿机制。



(a) 服务请求成功率对比图



(b) 改进的服务请求成功率

图 5 服务管理仿真实验结果

3 结 论

针对已有 QoS 保障服务系统模式效率低、可扩展性差、QoS 自动控制不足等问题,本文提出了核实与规划服务质量的系统框架,该框架可有机融合执行管理系统功能和系统服务,通过一种新的服务自动组合来实现系统监控与管理功能优化以及 QoS 自动控制,提高了系统的可扩展性。仿真实验结果表明,核实与规划服务质量的系统框架能够有效提高 QoS 监控与管理效率,方便 QoS 自动控制,其最大特点是能够自适应组合执行管理系统功能,以满足服务执行动态变化以及自动组合带来的功能需求。

参考文献:

[1] CHATTERJEE B S S, SYDIR M D J J, LAWRENCE T F. Taxonomy for QoS specifications [C]// Proceedings of 3rd Workshop on Object-Oriented Real-Time Dependable Systems. Newreport Beach, USA: IEEE, 1997:100-107.
[2] DAN A, DAVIS D, KEARNEY R, et al. Web services on demand: WSLA-driven automated management [J]. IBM Systems Journal, 2004, 43(1): 136-158.
[3] ANTONIOS L, KLEOPATRA K, VASSILIKI A, et

- al. Managing service level agreement contracts in OGSA-based Grids[J]. Future Generation Computer Systems, 2008, 24(4):245-258.
- [4] SHIRAZI H, COSMAS J, OWENS T J, et al. A QoS monitoring system in a heterogeneous multi-domain DVB-H platform[J]. IEEE Trans on Broadcasting, 2009, 55(1):124-131.
- [5] HE Qiang, YAN Jun, RYSZARD K, et al. Lifetime service level agreement management with autonomous agents for services provision[J]. Information Sciences, 2009, 179(15):2591-2605.
- [6] SHAH N, IQBAL R, IQBAL K. A QoS perspective on exception diagnosis in service-oriented computing [J]. Journal of Universal Computer Science, 2009, 15(9):1871-1885.
- [7] LIU Qin, JIA Xiaohua, WU Chanle. Optimal precomputation for mapping service level agreements in grid computing[J]. Future Generation Computer Systems, 2008, 24(8):845-859.
- [8] 陈彦萍, 李增智, 郭志胜, 等. Web 服务组合中基于服务质量的服务选择算法[J]. 西安交通大学学报, 2006, 40(8):897-900.
- CHEN Yanping, LI Zengzhi, GUO Zhisheng, et al. Service selection algorithm based on quality of service and its implementation for web services composition [J]. Journal of Xi'an Jiaotong University, 2006, 40(8):897-900.
- [9] 杨放春, 苏森, 李祯. 混合 QoS 模型感知的语义 Web 服务组合策略[J]. 中国科学, E 辑: 信息科学, 2008, 38(10):1679-1716.
- YANG Fangchun, SU Sen, LI Zhen. QoS-aware semantic web service composition[J]. Science in China: Series E, 2008, 38(10):1679-1716.
- [10] 伍之昂, 罗军舟, 宋爱波. 基于 QoS 的网格资源管理 [J]. 软件学报, 2006, 17(11):2264-2276.
- WU Zhi'ang, LUO Junzhou, SONG Aibo. QoS-based grid resource management [J]. Journal of Software, 2006, 17(11): 2264-2276.

(编辑 武红江 苗凌)