

网络 QoS 控制的效用模型及优化^{*}

刘 丽¹ 杨 扬¹ 刘松涛²

(北京科技大学信息工程学院 北京 100083)¹ (秦皇岛港务局 秦皇岛 066002)²

摘 要 基于经济模型研究网络 QoS 控制的资源分配问题,给出多 QoS 属性多任务的资源分配优化的效用模型,通过定义任务的多 QoS 属性效用模型来描述动态需求和偏好,计算需要提供 QoS 的资源分配。用多属性优化方法,求解优化多 QoS 属性效用函数。

关键词 网络,服务质量(QoS),效用模型

Computing Economics-based Grid QoS Control Model

LIU Li¹ YANG Yang¹ LIU Song-Tao²

(Information Engineering School, University of Science and Technology, Beijing 100083)¹

(Qinhuangdao Port Group Co., LTD, Qinhuangdao 066002)²

Abstract This paper applies the economy methodology to solve the quality of service(QoS) guarantees problem, and presents utility model expressed in multi QoS attributes, which can be used to dynamically allocate resource in the Grid. Multi-attribute optimization algorithm is used to optimize objectives of utility function for multiple dimension QoS.

Keywords Grid, Quality of Service(QoS), Computing economics model

1 引言

QoS 作为网络资源管理研究的重要内容之一,逐渐成为国内外研究的焦点。Ian Foster 提出判断网络的三个准则之一就是“提供非凡的服务质量”^[1]。开放网络体系结构(Open Grid Service Architecture, OGSA)的目标就是通过跨分布式异构平台管理资源来交付无缝的服务质量。QoS 在多媒体和网络通信领域中得到广泛的研究。多媒体 QoS 研究关注于如何以一定级别的服务质量保证来存取和传输媒体文件。网络 QoS 主要关注于如何提供更可靠的网络传输服务,满足应用的传输延时、延时抖动、分组丢失等质量保证。网络是运行在互联网之上的,为满足高性能网络应用端到端 QoS 保证的要求,除了网络层次外,还需在网络系统结构上实现相应的 QoS 保证,网络 QoS 问题的焦点在于网络能够提供端到端 QoS 的程度,而不是仅提供网络 QoS,目前的大多数网络实现都只能提供部分的 QoS 功能。

用户或应用为了满足自己的 QoS 需要,总是希望通过竞争服务资源获得最好的服务质量,这样造成资源的无效使用,因此需要研究经济的激励机制。本文在 QoS 相关研究基础上,针对 QoS 管理和控制问题,结合经济学模型^[2,3],给出多 QoS 属性多任务的资源分配优化的效用模型,通过定义任务的多 QoS 属性效用模型来描述动态需求和偏好,计算需要提供 QoS 的资源分配。进一步用多属性优化方法,求解优化多 QoS 属性效用函数。

2 相关研究

网络 QoS 研究还处于基础研究阶段,主要是针对网络 QoS 体系结构的研究。代表性研究包括:基于 Globus 的预留和分配体系结构(GARA)^[4],以及网络 QoS 管理体系结构(G-QoSM)^[5]。GARA 是 Ian foster 等人提出的网络 QoS 体

系结构,目标是为请求预留资源的客户端或应用程序能从资源管理者处获得特定的 QoS 提供保障。G-QoSM^[5]是英国 Cardiff 大学 Rashid 等人提出的 OGSA 环境下进行资源预留和自适应调节的面向服务 QoS 管理体系结构。提供基于 QoS 属性的服务发现;建立服务等级协议(Service Level Agreement, SLA),在中间件层和网络层提供 QoS 保证机制,为分配的资源提供自适应调节。Keahey 等提出了虚拟应用服务(Virtual Application Service, VAS)的体系结构^[8]来管理计算网络中的 QoS, VAS 是扩展了的网格服务,主要目标是使有特定最终期限约束的实时服务更易得到执行。

国内也开始重视网络 QoS,出现了一些研究文章^[6,7],许多网格平台开始加入对 QoS 的支持,建立起性能监控平台,更多的研究也在进行。

3 网络 QoS 控制的效用模型

基于经济的网络 QoS 控制把提供 QoS 保障的网格服务看作服务市场, QoS 的生产者是网格服务提供者代理, QoS 的消费者是用户或应用的需求代理,交易对象是具有一定 QoS 级别的服务,由网格服务提供者以资源为生产要素生产出不同级别的 QoS 产品。用户或应用对 QoS 的偏好由效用函数确定。优化目标是通过资源分配提供给用户 QoS 保障,使得整体效用最大化。

3.1 QoS 需求模型

假设对于每一个任务集合中的一个任务 t_i , 都有 QoS 需求集合 $Q_i = (Q_{i1}, \dots, Q_{id_i})$, ($1 \leq i \leq n, d_i = |Q_i|$), 同时用效用函数表示用户从完成任务中获得的 QoS 满意程度,效用函数 $u_i: Q_i \rightarrow R^+$, R^+ 为非负实数空间。任务集中所有任务或应用的总效用为 $\sum_{i=1}^n u_i(Q_i)$, 系统优化目标是最大化总效用,表示为: $\max \sum_{i=1}^n u_i(Q_i)$, s. t. $p_i Q_i \leq w^i$, 其中 w^i 表示第 i 个请求代

^{*} 本课题得到国家自然科学基金(No. 90412012)资助。刘 丽 副教授,高级工程师,研究方向为网络资源管理。杨 扬 教授,博士生导师,研究方向为计算机网络和图像处理。

理的预算。用 u_{ij} 表示用户对获得的 Q_{ij} 的满意程度, 用户 i 的总效用表示为 $u_i = \sum_{j=1}^{d_i} \omega_{ij} u_{ij}(Q_{ij})$, 其中, ω_{ij} 为任务 t_i 的不同 QoS 属性的权重。

3.2 QoS 生产模型

服务提供者根据计算能力、用户对服务质量的需求和偏好及服务成本, 来组织和调度资源。对某个任务的资源分配是服务提供者的生产要素, 以一定的生产函数生产 QoS 产品, 服务的成本决定于需要多少资源、使用多长时间、服务等级。该生产函数将资源分配和 QoS 联系在一起, 描述了系统的性能模型。定义服务提供者对任务 t_i 提供第 k 个 QoS 级别的服务所需要的资源分配为 $r_i = (r_{i1}, \dots, r_{im})$, 定义生产函数 $h_i: r_i \rightarrow Q_i$, 表示给定资源组合, 可以提供什么级别的服务。又根据 Q_i 和效用的关系, 得到资源分配和效用的关系, 表示为 $h_i(r_i) = \{q \in Q_i \mid \arg \max u_i(q)\}$ 资源分配、QoS 和效用三者关系如图 1 所示。

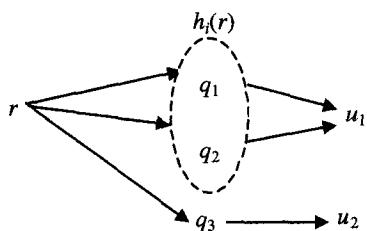


图 1 资源分配、QoS 和效用关系示意图

从图 1 中可以反映出, 给定一组资源, 生产者可以提供一定级别的服务能力(QoS), 通过获得服务能力, 实际用户或应用可以获得一定的效用。

4 效用驱动的目标优化

利用上面给出的效用模型, 用时间、预算和服务可用性的联合效用函数来选择服务完成资源分配。假设有 n 个任务, m 个服务节点, 对于任意任务 $t_i (1 \leq i \leq n)$ 在第 j 个服务节点 $S_j (1 \leq j \leq m)$ 的效用包含 3 方面属性: 其时间效用表示该任务在 S_j 上完成时间 FT_{ij} 的效用值, 用 $u_1(FT_{ij})$ 表示; 预算效用表示 S_j 完成该任务所需的服务费用 C_{ij} 的效用值, 表示为 $u_2(C_{ij})$; 服务的可用性效用函数表示 S_j 正常工作的效用, 表示为 $u_3(A_{vij})$, 为任务 t_i 构建决策矩阵 $A = (a_{jk})_{m \times d}$, 其中 m 表示可提供服务的服务器数量, d 表示用户关心的 QoS 属性的维数, 通过前面的论述, 知 $d = 3$, $a_{j1} = u_1(FT_{ij})$, $a_{j2} = u_2(C_{ij})$, $a_{j3} = u_3(A_{vij})$ 。

在时间、花费和可用性 3 方面属性中, 时间和花费为成本型属性, 即数值越低对用户越有利; 而可用性为效益型属性, 即数值越高对用户越有利, 故可对矩阵 A 进行如下规范化处理从而得到规范化矩阵 $R = (r_{jk})_{m \times 3}$ 。

$$r_{jk} = \frac{\min(a_{jk})}{a_{jk}}, (j=1, \dots, m) (k=1, 2)$$

$$r_{jk} = \frac{a_{jk}}{\max_j(a_{jk})}, (j=1, \dots, m), (k=3)$$

假设属性权重向量为 $\bar{\omega} = (\omega_1, \omega_2, \omega_3)$, 并满足单位化约束条件 $\sum_{k=1}^3 \omega_k^2 = 1$, 则表示对服务提供者选择方案的综合评价的效用函数可定义为:

$$U_{ij}(\omega) = \omega_1 u_1(FT_{ij}) + \omega_2 u_2(C_{ij}) + \omega_3 u_3(A_{vij}) = \sum_{k=1}^3 r_{jk} \omega_k$$

我们将需要解决的问题归结为属性权重完全未知且属性

值为实数的多属性决策问题。多属性决策问题的求解一般都与属性权重有密切关系, 因为权重的合理性直接影响着多属性决策排序的准确性, 本文运用离差最大化进行多属性决策和排序, 即若所有的服务提供者属性 u_k 下的属性值差异越小, 则说明该属性对方案的选择与排序所起的作用越小; 反之, 如果属性 u_k 能使所有服务提供者的属性值有较大的差异, 则说明其对排序比较将起重要的作用。因此, 从对各方方案进行排序的角度考虑, 属性值偏差越大的属性应该赋予越大的权重。对于属性 u_k , 用 $V_{jk}(\omega)$ 表示任意服务提供者 S_j 与其它服务提供者 $S_{j'}$ 之间的离差, 可以定义为:

$$V_{jk}(\omega) = \sum_{j'=1}^m |r_{jk} \omega_k - r_{j'k} \omega_k|$$

令 $V_k(\omega) = \sum_{j=1}^m V_{jk}(\omega) = \sum_{j=1}^m \sum_{j'=1}^m |r_{jk} - r_{j'k}| \times \omega_k$, 则 $V_k(\omega)$ 表示对属性 u_k 而言, 所有方案与其它方案的总离差。根据前面的分析, 加权向量 ω 的选择应使所有属性对所有方案的总离差最大。于是, 求解权重向量 ω 等价于求解如下最优化模型

$$\begin{cases} \max V(\omega) = \sum_{k=1}^3 \sum_{j=1}^m \sum_{j'=1}^m |r_{jk} - r_{j'k}| \times \omega_k \\ \text{s. t. } \omega_k \geq 0, k=1, \dots, 3, \sum_{k=1}^3 \omega_k^2 = 1 \end{cases}$$

求解最优化模型, 作拉格朗日(Lagrange)函数并求其偏导数, 并令

$$\begin{cases} \frac{\partial L}{\partial \omega_k} = \sum_{j=1}^m \sum_{j'=1}^m |r_{jk} - r_{j'k}| + \lambda \omega_k = 0, k=1, \dots, 3 \\ \frac{\partial L}{\partial \lambda} = \sum_{k=1}^3 \omega_k^2 - 1 = 0 \end{cases}$$

由此可得到

$$\omega_k = \frac{\sum_{j=1}^m \sum_{j'=1}^m |r_{jk} - r_{j'k}|}{\sum_{k=1}^3 \sum_{j=1}^m \sum_{j'=1}^m |r_{jk} - r_{j'k}|}, k=1, \dots, 3$$

将求得的 ω_k 带 $U_{ij}(\omega) = \sum_{k=1}^3 r_{jk} \omega_k$ 入, 可求出相应的 $U_{ij}(\omega)$, 利用其对各方方案进行排序便可求出最适合的服务提供者, 从而完成服务选择和资源分配。

结论 提出了基于计算经济模型的网格 QoS 控制机制, 并基于生产——消费模型给出多 QoS 属性优化效用模型, 模型中用户的 QoS 偏好由效用函数描述, 服务提供者以其拥有的资源为生产函数的要素, 生产出不同级别的 QoS 产品, 用多属性优化方法求解效用函数, 效用模型可以量化 QoS 并指导资源分配。

参考文献

- 1 Foster I, Kesselman C. The Grid: Blueprint for a Future Computing Infrastructure [M]. USA: Morgan Kaufmann Publishers, 1999
- 2 张金水. 数理经济学-理论与应用[M]. 北京: 清华大学出版社, 1998
- 3 武康平. 高级微观经济学[M]. 北京: 清华大学出版社, ISBN 7-302-04171-7/F302, 2001
- 4 Foster I, Roy A, Sander V. A quality of service architecture that combines resource reservation and application adaptation [R]. In: 8th Int'l Workshop on Quality of Service(IWQoS'00), June 2000
- 5 Al-Ali R, Amin K, Laszewski G, et al. An OGSA-based Quality of Service Framework [C]. In: Proc. of the Second International Workshop on Grid and Cooperative Computing(GCC2003), 2003
- 6 金海, 陈汉华, 等. CGSP 作业管理器合成服务的 QoS 优化模型及求解[J]. 计算机学报, 2005, 28(4): 578~588
- 7 朱穗晖, 都志辉, 林伟坚. 网格环境合同计算元服务的设计与应用[J]. 计算机学报, 2005, 28(4): 486~494
- 8 Keahey K, Motawi K. The Taming of the Grid: Virtual Application Service, Argonne National Laboratory Technical Memorandum, 2003(262)