

# 覆盖网络中一种公平负载均衡 QoS 路由算法

曹雪松<sup>1),2)</sup> 胡瑞敏<sup>1)</sup> 王朝萍<sup>1)</sup>

<sup>1)</sup>(武汉大学计算机学院 武汉 430072)

<sup>2)</sup>(中国通信服务集团天讯瑞达通信技术有限公司 广州 510623)

**摘 要** 覆盖网络通过选择资源消耗代价较低的应用层服务路径实现分布式网络应用的负载均衡,但传统方法未考虑局部负载较重链路中存在的资源瓶颈问题,导致系统吞吐量的下降. 针对该问题提出一种公平负载均衡 QoS 路由算法,该方法引入表征路径瓶颈程度的资源公平指数,依据代价模型建立新的负载均衡效用函数,并采用自适应的系统负荷状态加权方法修正 QoS 约束项权重. 实验证明该算法相对传统算法性能更优,有效提高了服务成功率,并降低了负载分布的波动程度.

**关键词** 覆盖网络; QoS 路由; 负载均衡; 分布式系统

**中图法分类号** TP309 **DOI 号**: 10.3724/SP.J.1016.2011.01650

## A Fair Load-Balancing QoS Routing Algorithm in Overlay Network

CAO Xue-Song<sup>1),2)</sup> HU Rui-Min<sup>1)</sup> WANG Zhao-Ping<sup>1)</sup>

<sup>1)</sup>(School of Computer, Wuhan University, Wuhan 430072)

<sup>2)</sup>(Tisson Regaltec Communications Technologies Co., Ltd., Chinese Comservice, Guangzhou 510623)

**Abstract** Overlay network can achieve load balancing of distributed system by choosing lower-cost application-level paths. However, traditional methods didn't take into account the resource bottleneck of partial heavy load paths, so as to decreasing the throughput. Aiming at this problem, a fair load balancing QoS routing algorithm is proposed. The algorithm introduces a fairness resource index into the cost model, and establishes a new load-balancing utility function. At the same time, an adaptive weighting method based on system load status is used to amend the weight of QoS constraints. Experiments show that the algorithm is relatively better performance of the traditional algorithm, to effectively improve service success rate and reduce the fluctuations degree of load distribution.

**Keywords** overlay network; QoS routing; load-balancing; distributed system

## 1 引 言

覆盖网络<sup>[1]</sup>(overlay network)又称重叠网络,是指通过在基础网络的关键位置上设立智能的节点,以单播路由方式连接起来构成的虚拟应用层网

络,这种网络位于用户终端与物理网络之间,可以用于实现特定的应用. 相对于底层基础网络而言,覆盖网络层能够监测底层网络上的路由走向、传输延时、流量分布、拥塞状况以及异常故障等状态信息. 覆盖网络层一方面依靠覆盖节点(overlay node)相互交换这些信息,另一方面基于这些信息实施智能的控

收稿日期:2009-08-20;最终修改稿收到日期:2011-07-25. 本课题得到国家“九七三”重点基础研究发展规划子项目(2009CB320906)、国家自然科学基金(60772106)、武汉市科技攻关项目(20063003109-2)和广东省教育部产学研结合项目(2006D90104005)资助. 曹雪松,男,1979年生,博士,中国计算机学会(CCF)学生会会员,主要研究方向为多媒体网络通信协议、高性能分布式多媒体系统、覆盖网络、服务质量 QoS. E-mail: caoxs1979@gmail.com; caoxs0207@163.com. 胡瑞敏,男,1964年生,博士,教授,博士生导师,主要研究领域为多媒体网络通信、多媒体信息处理和编码. 王朝萍,女,1981年生,博士,主要研究方向为多媒体网络通信系统、覆盖网络、服务质量 QoS.

制和管理手段,如路由探测选择、流量控制以及内容检索、分类和管理等等.覆盖网络的优点在于不改变基础网络架构而有效提高分布式网络应用的服务质量(Quality of Service, QoS).

覆盖网络 QoS 路由是覆盖网络技术的重要研究方向.由于传统的 TCP/IP 网络只能提供尽力而为的服务模式,不能有效保障分布式应用的传输质量,而在应用层上重新构建可靠的覆盖路由网络,就可以提供 QoS 保障的传输控制机制,这种技术被称为覆盖网络的 QoS 路由<sup>[2]</sup>.在这种体系结构中,覆盖网络是由应用层的服务器主机节点构成,这些节点不仅能够传递和转发数据流,而且还具有特殊的计算能力和存储能力,通过互相协作实现复杂的服务组合.因此覆盖网络能够使用路由选择方法有效解决分布式应用的资源调度问题<sup>[3]</sup>.

负载均衡的资源调度是分布式系统应用研究中的关键问题,因为网络环境的差异性,不同区域业务吞吐量的波动以及不合理的资源分配策略都会导致网络负载的不均衡,进而引起网络出现拥塞现象,降低系统的整体吞吐量,目前该问题在 IP 网络上仍没有很好的解决办法.而覆盖网络 QoS 路由技术可以提供基于应用层部署负载均衡资源调度策略的方法,有助于分布式系统实现公平的资源分配,因此成为了当前的研究热点.

经典的覆盖网络负载均衡 QoS 路由算法主要有 QUEST<sup>[4]</sup>、PBLCP<sup>[5]</sup>、RBLCP<sup>[6]</sup>等,这些算法的基本原理均是推导表征覆盖网络传输、计算资源的消耗代价和的启发式效用函数,基于最短路径算法搜索最优负载均衡服务路径,有利于选出整体资源消耗较低的路径.然而该类方法仍存在一个问题:未考虑局部负载较重的链路引发的资源瓶颈问题,这种瓶颈仍会导致系统发生拥塞,严重时会导致系统吞吐量和负载均衡程度.针对该类问题,本文重点研究了路径瓶颈程度的计算方法,并将其引入传统路由模型,提出一种更加公平的负载均衡覆盖路由算法,实验证明该算法能够进一步提高网络系统的吞吐量及负载均衡程度.

本文第 2 节介绍覆盖网络 QoS 路由技术相关的国内外研究现状;第 3 节通过原理性分析和实验结果验证资源瓶颈问题及其影响;第 4 节针对该问题提出解决方案,详细阐述本文核心算法;第 5 节给出仿真实验的结果及分析;最后第 6 节阐述本文的研究结论.

## 2 相关研究现状

覆盖网络负载均衡 QoS 路由的目标<sup>[3,7]</sup>就是在覆盖网络中找到一条合适的服务组合路径,其接入业务所花费的资源代价相对较小,并且能够充分地利用系统的“空闲”资源,这样在不断接入用户请求的过程中,负载就能均衡分布在网络空间内部,从而使网络资源的利用率达到最大.覆盖网络中的资源类型可以抽象为两类:传输资源和计算资源,一般可用链路的传输带宽(BR)和节点的计算能力(CR)表示,目前的负载均衡 QoS 路由算法都是以优化这两类资源的分布为目标,求解最优的 QoS 服务路径.QUEST<sup>[4]</sup>是较早提出的多 QoS 优化目标覆盖路由算法,它证明该类问题属于 NP-Complete 问题.该算法首先定义了多 QoS 目标优化的代价函数,根据问题目标的侧重程度可对代价函数进行加权调节,然后以最短路径 Dijkstra 算法<sup>[8]</sup>求解最优路由结果.QUEST 针对负载均衡目标优化采用了两种 QoS 度量约束项:

(1) 链路  $l_{i,j} = (S_i, S_j)$  的带宽消耗率  $f_B(l_{i,j}, B_i)$ ,  $B_i$  为服务请求的带宽资源需求约束;

(2) 节点  $S_j$  的计算资源的消耗率  $f_c(l_{i,j}, C_i)$ ,  $C_i$  为服务请求的计算资源需求约束;

设  $CR_j$  和  $BR_{i,j}$  为节点  $j$  和链路  $l_{i,j}$  的可用资源,得到链路启发式代价函数  $C(l_{i,j})$  为

$$C(l_{i,j}) = f_B(l_{i,j}, B_i) + f_c(l_{i,j}, C_i) = C_i / CR_j + B_i / BR_{i,j} \quad (1)$$

路径  $P$  的代价通过链路代价和计算,其总体消耗代价函数  $Cost(P)$  定义为

$$Cost(P) = \sum_{l_{ij} \in P} C(l_{i,j}) = \sum_{l_{ij} \in P} \left( \frac{C_i}{CR_j} + \frac{B_i}{BR_{i,j}} \right) \quad (2)$$

2005 年 Li 等人<sup>[5]</sup>提出了一种基于传输资源和计算资源残余比例的负载均衡路由算法 PBSP(Proportional Bandwidth Shortest Path),比 QUEST 能更加精确表征资源利用程度,算法改用了节点和链路的残余可用资源的比例作为选择权值,并通过乘法混合度量生成代价函数,路径的代价仍通过所有链路的代价和计算,其代价函数  $Cost(P)$  如下:

$$Cost(P) = \sum_{l_{ij} \in P} \frac{CR_j}{CR_j - C_i} \times \frac{BR_{i,j}}{BR_{i,j} - B_i} \quad (3)$$

2007 年应必娣等人<sup>[6]</sup>在 PBSP 的基础上又提出了一种改进的路由算法 RBLCP.该算法引入了全局服务资源分布的约束因子,可以基于整体资源分布状况动态对带宽和计算资源代价函数加权,这样算法能够根据两类资源的瓶颈情况作自适应的负载

均衡( $\alpha$  值为权重因子),其代价函数  $Cost(P)$  如下:

$$Cost(P)=\sum_{l_{ij}\in P}\frac{1}{\alpha\times(CR_j/C_t)+(1-\alpha)\times BR_{i,j}/B_t}\tag{4}$$

3 问题分析

由第 2 节研究现状可知,传统的覆盖路由算法采用了表征资源消耗代价的变换函数作为负载均衡启发式效用函数,这些方法能够尽量选出整体资源饱和度<sup>[9]</sup>(实际承受负载与最大承受负载比值)较低的空闲服务路径,但仍然存在一定的非公平性,这主要是因为它们没有考虑局部负载较重链路的资源瓶颈问题.服务路径是由多段覆盖链路组成,链路和节点的能力也会存在不均衡的情况,虽然叠加的总体消耗代价低,但是单端链路或节点的资源消耗率仍然可能较高,因而其出现拥塞的概率也会相对较高,一旦出现异常不但会影响到运行期内的服务会话,而且可能导致局部网络瘫痪,反而降低系统吞吐量.实验 1 的数据表明传统覆盖路由算法在高业务流量下仍存在较严重的资源拥塞现象.

**实验 1.** 基于 NS2 平台<sup>[10]</sup> 模拟大规模持续流式服务请求接入覆盖网络系统的场景,具体实验方法参见第 5 节.实验统计了在不断递增的业务负荷量(20~200 unit/s)情况下通过 PBSP、RBLCP 两种路由算法调度服务所产生的节点吞吐量分布情况.图 1(a)、(b)描述了当业务总量达到 100 unit/s 时所

有物理节点的吞吐流量.结果表明当并发业务流量较大的时候,资源分布不均衡的状况趋于明显,并出现了若干可能会导致局部网络拥塞的流量峰顶节点.

图 2 描述了业务量不断上升过程时的服务请求接收成功率,结果表明当系统业务量不断上升直至超负荷情况下,服务成功率也随之陡降(80 unit/s 时出现过载点)这种现象证明了网络拥塞造成的资源瓶颈问题会严重降低系统的服务吞吐量.

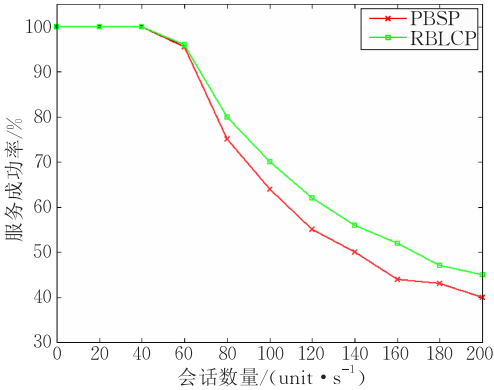


图 2 服务成功率的不断降低

4 公平负载均衡 QoS 路由算法

由第 3 节理论分析和实验验证可知,传统算法存在无法排除局部路径瓶颈的问题,针对该问题本文提出一种公平负载均衡 QoS 路由算法 F-LBQAR (Fair-Load Balance QoS Aware Routing),在路由模型中引入路径资源瓶颈程度约束因子,尽量避免选择瓶颈链路,能够在一定程度上降低网络拥塞的出现概率.本节给出具体算法的技术路线.

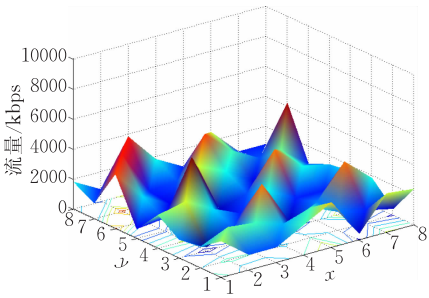
4.1 服务资源消耗的代价模型

传统算法通过求取链路的资源消耗率总和并计算服务资源消耗代价,F-LBQAR 仍采用类似的模型,所不同的是本算法用路径的消耗代价权值代替链路的消耗代价权值,这种方法分开考虑传输资源和计算资源的约束因子,更符合代价特征量的物理意义,并有利于针对不同资源 QoS 需求做独立的加权.

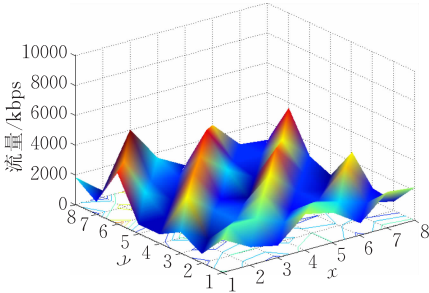
本节首先定义针对路径接入服务后的资源消耗代价权值的计算模型,该模型分为两部分:一是传输服务代价:以链路流量消耗率的总和表示;二是计算服务代价:以节点计算能力消耗率的总和表示,代价函数定义如下.

**定义 1.** 路径传输资源消耗代价函数.

设某到达业务请求  $T$ ,其传输资源的 QoS 需求为  $B_t$ ,假定覆盖网络系统能够提供服务路径  $P$  接受



(a) PBSP算法



(b) RBLCP算法

图 1 负载不均衡现象

业务  $T$  的服务请求, 则  $T$  对  $P$  的传输资源消耗代价  $\Delta B(P)$  为

$$\begin{aligned} \Delta B(P) &= B_t / BR_{0,1} + \lambda_1 B_t / BR_{1,2} + \\ &\quad \lambda_1 \lambda_2 B_t / BR_{2,3} + \cdots + \prod_{j=0}^{t_p-1} \lambda_j B_{t_p-1, t_p} \\ &= B_t \cdot \sum_{i=1}^{t_p} \left( \prod_{j=0}^{i-1} \lambda_j / BR_{i-1,i} \right) \end{aligned} \quad (5)$$

$BR_{i,j}$  为路径  $P$  上每跳链路  $l_{i,j}$  的可用带宽资源存量,  $\lambda_1, \dots, \lambda_i$  为路径  $P$  每一跳节点  $i$  对流量的影响系数(节点输出流量  $B_{\text{out}} = \lambda \times B_{\text{in}}$ ,  $\lambda_0 = 1$ ),  $t_p$  为当前路径跳数。

**定义 2.** 路径计算资源消耗代价函数。

设某到达业务请求  $T$ , 其计算资源的 QoS 需求为  $C_t$ , 假定覆盖网络系统能够提供服务路径  $P$  接受业务  $T$  的服务请求, 则  $T$  对  $P$  的传输资源消耗代价  $\Delta C(P)$  为

$$\Delta C(P) = C_t / CR_0 + C_t / CR_1 + C_t / CR_2 + \cdots + C_t / CR_{t_p}$$

$$= \sum_{i=0}^{t_p} \frac{C_t}{CR_i} \quad (6)$$

$CR_0, CR_1, \dots, CR_{t_p}$  为路径  $P$  上每跳节点  $i$  (包括源节点和目的节点) 的可用计算资源存量,  $t_p$  为当前路径跳数。

由以上定义可知,  $\Delta B(P)$  和  $\Delta C(P)$  两个代价约束项是相互独立的参数因子, 因此可通过加法混合度量求取路径的整体代价函数, 如下定义。

**定义 3.** 路径资源消耗代价函数

$$Cost(P) = \Delta B(P) \oplus \Delta C(P) \quad (7)$$

将式(6)和(7)代入式(5)得式(8)

$$Cost(P) = B_t \cdot \sum_{i=1}^{t_p} \left( \prod_{j=0}^{i-1} \lambda_j / BR_{i-1,i} \right) + \sum_{i=0}^{t_p} \frac{C_t}{CR_i} \quad (8)$$

## 4.2 服务资源利用率公平指数

为了降低服务路径的瓶颈程度, 本文引入了一种基于服务资源利用率公平指数的权值约束路由方法, 该方法通过计算服务路径中所有链路和节点的资源利用率的公平性来衡量路径中是否存在瓶颈。假设存在序列  $x_1, x_2, \dots, x_n$ , 则衡量序列均衡度的公平指数函数  $f(x)$  的定义<sup>[11]</sup> 如下:

$$f(x) = \frac{(x_1 + x_2 + \cdots + x_n)^2}{n \cdot (x_1^2 + \cdots + x_n^2)} = \frac{\left( \sum_{i=1}^n x_i \right)^2}{n \cdot \sum_{i=1}^n x_i^2} \quad (9)$$

从原理上分析可知, 公平指数值  $f(x)$  当  $x_1 = x_2 = \cdots = x_n$  时  $f(x)$  取最大为 1, 而如果序列  $A_i$  间的差值波动越高, 则其值越小, 无限趋近于 0, 因此

$f(x)$  的值域为  $(0, 1]$ 。如果采用数值序列的均方差计算均衡度, 则情况正好相反, 取值越大则波动越高, 值域范围为  $[0, 1)$ , 但是其值有可能出现为 0 的情况, 这种情况并不利于作除法系数, 所以采用公平指数计算均衡度的方式更理想。

由以上定义可知, 可以通过计算服务路径上所有节点和链路的资源利用率公平指数评价路径的负载均衡程度。假设采用资源饱和度<sup>[9]</sup> 表征利用率, 则如果路径上部分链路和节点的饱和程度较高(表明链路和节点的负荷较大), 会造成该指数值偏大, 如果饱和程度越均衡的话, 则该指数值就越小(但高负荷下也是如此)。因此将该指数作为最短路径权值的正则约束项, 可减少出现瓶颈的概率, 具体计算方法如下。

**定义 4.** 路径传输资源的利用率公平指数。

设某到达业务请求  $T$ , 其传输资源的 QoS 需求为  $B_t$ , 假定覆盖网络系统能够提供服务路径  $P$  接受业务  $T$  的服务请求, 则  $P$  接入  $T$  后的传输资源公平指数  $f_B(P)$  计算如下:

$$B(l_{i,i+1}) = \frac{BR_{i,i+1}^{\text{total}} - (BR_{i,i+1} - B_t)}{BR_{i,i+1}^{\text{total}}},$$

$$l_{i,i+1} \in P, i=0, 1, \dots, t_p-1 \quad (10)$$

$$f_B(P) = \left( \sum_{i=1}^{t_p} B(l_{i,i+1}) \right)^2 / t_p \cdot \sum_{i=1}^{t_p} B(l_{i,i+1})^2 \quad (11)$$

$B(l_{i,i+1})$  为路径  $P$  上链路  $l_{i,i+1}$  接入请求后的传输资源负荷饱和度函数,  $BR_{i,i+1}$  为路径  $P$  上链路  $l_{i,i+1}$  的可用带宽资源存量,  $BR_{i,i+1}^{\text{total}}$  为路径  $P$  上链路  $l_{i,i+1}$  的可用带宽资源总量,  $t_p$  为当前路径跳数。该指数与链路的传输负荷波动成反比线性关系, 当波动程度越高时, 指数取值也会越小, 表明路径中存在瓶颈链路的概率越大。

**定义 5.** 路径计算资源的利用率公平指数。

设某到达业务请求  $T$ , 其传输资源的 QoS 需求为  $C_t$ , 假定覆盖网络系统能够提供服务路径  $P$  接受业务  $T$  的服务请求, 则  $P$  接入  $T$  后的计算资源公平指数  $f_c(P)$  计算如下:

$$C(n_i) = \frac{CR_i^{\text{total}} - (CR_i - C_t)}{CR_i^{\text{total}}}$$

$$n_i \in P, i=0, 1, 2, \dots, t_p \quad (12)$$

$$f_c(P) = \left( \sum_{i=0}^{t_p} C(n_i) \right)^2 / (t_p + 1) \cdot \sum_{i=0}^{t_p} C(n_i)^2 \quad (13)$$

$C(n_i)$  为路径  $P$  上每跳节点  $C(n_i)$  (包含源节点和目的节点) 的计算资源饱和度函数,  $CR_i$  为节点  $n_i$  的可用带宽资源存量,  $CR_i^{\text{total}}$  为  $n_i$  的可用计算资源

总量. 该指数与节点的计算负荷波动成反比线性关系, 当波动程度越高, 指数取值也会越小, 表明路径中存在瓶颈节点的概率越大.

单纯使用该指数作路径选择只能选出均衡度高的路径, 但无法降低路径的整体消耗代价, 因此将式(11)、(13)引入式(7), 在服务代价模型的基础上加入瓶颈约束因子, 通过乘性混合度量计算服务路径的负载均衡度, 这样就能够获得瓶颈概率小且资源消耗少的服务路径, 负载均衡效用函数的计算方法如下.

**定义 6.** 服务路径的资源负载均衡效用函数

$$w(P) = \frac{\Delta B(P)}{f_B(P)} \oplus \frac{\Delta C(P)}{f_C(P)} \quad (14)$$

### 4.3 基于负荷加权的负载均衡启发式函数

式(14)描述了服务请求接入后对系统及服务路径负载均衡状态的影响情况, 因此根据影响程度的轻重就可决定路由的选择. 分析该公式可知,  $\Delta B(P)/f_B(P)$  因子对提高传输资源的负载均衡程度有利, 而  $\Delta C(P)/f_C(P)$  因子对提高计算资源的负载均衡程度有利, 两者影响权重完全相等. 而在实际情况中, 覆盖网络在特定运行周期内不同类型资源的负荷状态和瓶颈程度往往并不相同. 显然, 如果某类资源的瓶颈情况或者负载更重, 应当更倾向于优化该类资源的分配, 因此本文考虑采用系统负荷分布状况对两类因子进行加权处理, 修正模型的精度.

#### 4.3.1 系统负荷状况评价方法

考虑到 F-LBQAR 算法的目标是降低系统资源饱和度的波动, 尽量提高整个网络的负载分布的均衡性, 因此本文采用了基于全局资源公平指数的加权方法, 通过计算系统总体传输和计算服务资源的波动程度来得到权重值. 式(15)、(16)可计算出全局计算资源饱和度的公平指数  $G_c$  和传输资源饱和度的公平指数  $G_b$ , 这两个 QoS 指标能够对当前系统的负载均衡程度进行量化评价:

$$G_c = \left( \sum_{i=1}^k C(n_i) \right)^2 / k \cdot \sum_{i=1}^k C(n_i)^2 \quad (15)$$

其中  $C(n_i) = (CR_i^{\text{total}} - CR_i) / CR_i^{\text{total}}$ ,  $n_i \in G$ .

$$G_b = \left( \sum_{j=1}^m B(l_j) \right)^2 / m \cdot \sum_{j=1}^m B(l_j)^2 \quad (16)$$

其中  $B(l_j) = (BR_j^{\text{total}} - BR_j) / BR_j^{\text{total}}$ ,  $l_j \in G$ .

$k$  和  $m$  分别是图  $G$  的节点数和链路数,  $G_c$  和  $G_b$  能够表征当前资源饱和度的波动程度, 两个 QoS 参数越大, 则表明接入服务后系统的资源平均饱和程度越均衡, 瓶颈程度越轻, 而反之两个 QoS 参数越小, 则表明接入服务后系统的资源平均饱和程度越不均衡, 瓶颈程度越严重, 注意  $G_c$  和  $G_b$  的值域为  $(0, 1]$ .

#### 4.3.2 加权负载均衡启发式函数

本节基于  $G_c$  和  $G_b$  的比例关系设计负载均衡的权重因子的计算方法, 设定  $\alpha$  为计算资源负载均衡的加权判决因子,  $\beta$  为传输资源负载均衡的加权判决因子, 则  $\alpha, \beta$  的计算公式如下:

$$\alpha = 1 - \frac{G_c}{G_c + G_b} = \frac{G_b}{G_c + G_b} \quad (17)$$

$$\beta = 1 - \alpha = \frac{G_c}{G_c + G_b} \quad (18)$$

由于  $G_c$  和  $G_b$  值越大表明资源分布越均衡, 那么权重应倾向于值越小的那一类型资源, 因此实际上式(17)和(18)采用的是反比例计算.

基于式(17)和(18)对式(19)进行线性加权, 重新定义负载均衡效用函数为

$$\begin{aligned} w(P) &= \beta \times \frac{\Delta B(P)}{f_B(P)} \oplus \alpha \times \frac{\Delta C(P)}{f_C(P)} \\ &= \frac{G_c}{G_c + G_b} \times \frac{\Delta B(P)}{f_B(P)} \oplus \frac{G_b}{G_c + G_b} \times \frac{\Delta C(P)}{f_C(P)} \end{aligned} \quad (19)$$

由以上定义可知,  $w(P)$  值越小则路径负载均衡程度越高.

### 4.4 F-LBQAR 算法流程

本节将描述 F-LBQAR 算法的具体流程, 算法中将采用 Dijkstra 算法<sup>[8]</sup>搜索最短路径, 算法采用了邻接矩阵方法, 因此其计算复杂度为  $O(n^2)$ , 与 QUEST、PBSP 及 RBLCP 相同, 具体步骤如下.

**F-LBQAR 算法.** 公平负载均衡 QoS 路由算法.

输入: 业务请求  $T$ , 传输资源的 QoS 需求为  $B_t$ , 计算资源的 QoS 需求为  $C_t$ , 覆盖网络  $G$

输出: 从  $G$  找到一条从节点  $K$  到节点  $M$  的单播服务路径

1. 根据  $T$  的资源 QoS 约束, 剪除  $G$  所有无法满足  $B_t$  和  $C_t$  的节点与链路, 获取备选网络  $G'$ , 如果目标节点  $K, M \notin G'$ , 则算法结束, 拒绝业务  $T$ ;

2. 基于 Dijkstra 算法开始最短路径搜索;

3. 初始化  $K$  到所有节点对间的代价及路由信息, 设  $\forall t \in G'$ , 有  $DISTANCE[t] = \infty$ ,  $ROUTE[t] = -1$ ,  $DISTANCE[]$  是距离矩阵表,  $ROUTE$  是  $K$  节点路由表;

4. 设定根节点  $ROOT = K$ , 搜索节点  $SN = K$ , 备选搜索集合  $B\_SERARCH = \{G' - K\}$ , 开始最短路径迭代搜索;

5. 设定  $NB$  为  $SN$  的邻居节点集合, 则有  $\forall$  节点  $S \in NB$ ,  $\exists$  链路  $SN \rightarrow S$ , 且  $\exists$  路径  $PATH(ROOT, S)$ ;

6. 逐次计算  $ROOT$  至  $NB$  中节点的路径  $PATH$  的传输资源和计算资源代价值  $\Delta B(P)$  和  $\Delta C(P)$ ;

7. 逐次计算  $ROOT$  至  $NB$  中节点的路径  $PATH$  的传输资源和计算资源公平指数  $f_B(P)$  和  $f_C(P)$ ;

8. 计算  $G'$  的全局公平指数  $G_c, G_b$  及加权系数  $\alpha$  和  $\beta$ ;

9. 逐次计算  $ROOT$  到所有  $S \in NB$  路径的负载均衡权

值  $w(P)$ , 如果  $w(P) < \text{DISTANCE}[S]$ , 则更新距离矩阵  $\text{DISTANCE}[S] = w(P)$  及路由表  $\text{ROUTE}[S] = SN$ ;

10. 如果  $\exists N \in B\_SEARCH, \text{DISTANCE}[N] \neq \infty$ , 且为目前到  $ROOT$  的最小距离值, 则取  $SN = N, B\_SEARCH = B\_SEARCH - N$ , 并跳转到步 5, 否则到步 11;

11. 如果  $\text{DISTANCE}[M] = \infty$ , 则搜索失败, 拒绝业务  $T$ ;

12. 如果  $\text{DISTANCE}[M] \neq \infty$ , 则接受业务  $T$ , 并根据  $\text{ROUTE}[]$  回溯建立  $K \rightarrow M$  的服务路径。

## 5 仿真实验分析

### 5.1 实验环境

本文通过 NS2<sup>[10]</sup> 仿真工具模拟覆盖网络系统的资源调度场景, 验证 F-LBQAR 算法与传统算法间的性能差异。仿真方法采用 NS2 平台生成模拟的物理拓扑, 然后在此拓扑基础上通过节点间建立逻辑的应用连接构建覆盖网络空间, 最后部署应用层路由协议实现相关的算法, 具体步骤分为两步:

(1) 覆盖空间与物理拓扑间的虚映射。首先构建一组三元向量  $G = \{\text{OverlayGraph}, \text{OverlayNode}, \text{OverlayLink}\}$  描述覆盖网络空间的 3 种实体对象, 对应分别为覆盖拓扑对象、覆盖节点对象及覆盖链路对象, 并且通过继承 NS2 中的 NSObject 对象, 使这些元素能够在 TCL 实例化; 然后建立 NS2 中的物理拓扑空间与覆盖对象间的映射关系, 由主动选出的物理节点一一对应生成若干 OverlayNode 对象, 并根据底层路由表采用邻近邻接法<sup>[12]</sup> (即使得两个覆盖节点间的覆盖链路不存在其它覆盖节点) 构建出覆盖网络拓扑, 生成相应的 OverlayLink 对象; 最后通过 OverlayNode 和 OverlayLink 组合形成覆盖拓扑对象 OverlayGraph。

(2) 实时 Agent 监测。生成覆盖网络后, 仍需要采用端到端的测量方法<sup>[13-14]</sup> 获取覆盖网络中的各种网络性能参数, 首先继承 NS2 中的 Agent 对象生成 OverlayAgent 对象, 该对象可放置在任一 OverlayNode 上, 并通过向邻近的节点发送探测包来评估和交换 OverlayLink 和 OverlayNode 的性能, 比如传输延时、路径可用带宽等, 所有实时性能数据和通信连接信息可以存放在 OverlayGraph 的性能数据库中, 根据这些信息 OverlayGraph 对象可以实时计算出全局路由表。

对于业务通信的模拟则改进了 NS2 的流量发生机制。本文设计了 OverlayTraffic 对象可以依照 OverlayGraph 中的路由表建立多个 OverlayNode 节点对间的流 Application (NS2 中的 Application

对象) 形成应用层连接式的服务路径, 并且基于服务路径对象 OverlayPath 存储已建立的路由信息, 随时进行路径的拆除、中断、更新等操作。

### 5.2 仿真实验结果分析

本实验共分成两组: 衡量算法在网络规模和业务量增长的情况下对系统吞吐量的影响 (5.2.1 节实验 2); 衡量算法在业务量增长的情况下对系统负载波动程度的影响 (5.2.2 节实验 3)。

#### 5.2.1 服务成功率及吞吐量测试

**实验 2.** 模拟海量 QoS 服务请求不断到达的场景, 通过服务成功率和吞吐量等 QoS 指标验证不同的负载均衡覆盖路由算法对系统吞吐量的影响。实验环境基于 GT-ITM<sup>[15]</sup> 分别生成 4 种随机分布物理网络, 网络大小分别为 50 个节点 ( $scale = 100 \times 100$ )、100 个节点 ( $scale = 100 \times 100$ )、150 个节点 ( $scale = 100 \times 100$ )、200 个节点 ( $scale = 100 \times 100$ ); 节点间存在一条双向链路为 0.02 概率, 每条链路的带宽为 2 Mbps, 延时为 200~1000 ms; 选择 30% 的物理节点为覆盖节点, 每个节点最大计算能力为 20 unit/s; 通过 OverlayTraffic 模拟到达率为 20~400 unit/s 的 CBR 业务流, 平均流量为 300 kbps, 分别测试 QUEST、PBSP、RBLCP 及 F-LBQAR 算法的 QoS 服务成功率 (QoS-Service Success Rate, QSSR)<sup>[5]</sup>:

$$QSSR = \frac{\text{请求被接受的会话总数}}{\text{到达的服务会话总数}} \quad (20)$$

实验采用了度稀疏的物理网络 (0.02 概率) 构建覆盖网络, 因此受限的网络资源导致拥塞较易出现, 从图 3 的数据中明显观察到系统逐渐过载, 成功率逐渐下降的过程。分析实验结果可知, 随着业务流量的不断增长, 资源调度的 QoS 服务成功率也在不断下降, 这表明网络出现拥塞而超载。由于通常网络承载负荷的能力随着网络规模的增长而提高, 因此图 3(a) 中的 50 个节点网络过载点出现较早, 在 80 unit/s 负荷左右, 而图 3(d) 中的过载点则出现较晚, 在 260 unit/s 负荷左右。从该实验结果可以看出, 无论在何种拓扑下, 相对于传统算法, F-LBQAR 算法能够在系统过载的情况下 QSSR 提高了 1%~5% 左右, 特别对小规模的网络更加明显, 最高能提高 5% 左右的成功率 (图 3(a), 120~200 unit/s), 该结果说明 F-LBQAR 算法比传统算法更能适应恶劣的网络环境, 有利于提高网络系统的吞吐量。

图 4 的实验结果描述了在 100 个节点大小的网络平面上, 当系统负荷达到 100 unit/s 业务量时, 执行各类算法所构成的网络节点流量分布情况, 可以看出每种算法仍会产生出现拥塞的节点 (流量值较



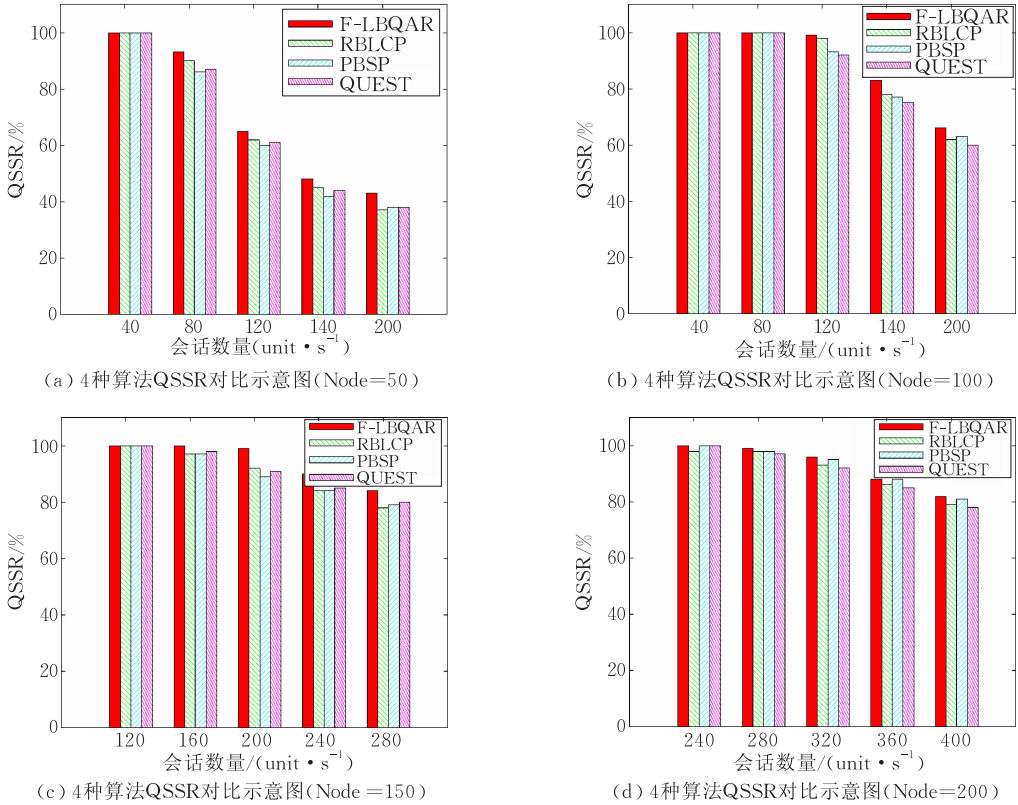


图 3 不同拓扑下 QSSR 测试实验结果

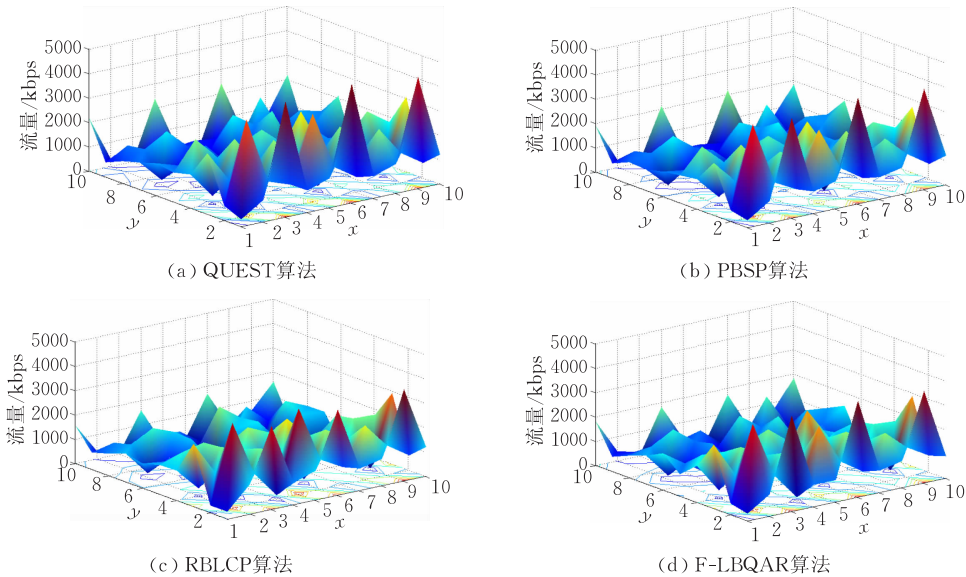


图 4 网络 size=100、100 unit/s 下各算法节点吞吐量分布情况

高的节点数),但执行 F-LBQAR 算法后,拥塞节点相对少于其它算法,负载能够被更加均衡地分布在网络平面上.表 1 统计了该情况下的流量峰值节点数目的分布情况,分析可知在这种实验环境下,QUEST 算法和 PBSP 算法出现 [5000 kpbs, 4000 kpbs]、[4000 kpbs, 3000 kpbs] 及 [3000 kpbs, 2000 kpbs] 流量区间上的高负荷节点概率要远高于

RBLCP 及 F-LBQAR,而 RBLCP 算法产生的在 [4000 kbps, 3000 kbps] 区间内的峰值节点数目略高于 F-LBQAR,但 [2000 kbps, 1000 kbps] 区间内节点数目要远高于 F-LBQAR,而 F-LBQAR 算法的峰值节点则主要集中在 [1000 kbps, 0 kbps] 部分,这说明 F-LBQAR 能够将负载更加均衡分布,即提高了网络资源的利用率,又同时降低了拥塞节点的数量.

表 1 流量峰值节点统计

吞吐量范围	节点数目			
	QUEST	PBSP	RBLCP	F-LBQAR
[5000 kpbs, 4000 kpbs]	3	3	0	0
[4000 kpbs, 3000 kpbs]	3	1	5	4
[3000 kpbs, 2000 kpbs]	15	12	4	4
[2000 kpbs, 1000 kpbs]	34	36	47	40
[1000 kpbs, 0 kpbs]	45	48	44	52
总数	100	100	100	100

5.2.2 负载均衡波动程度测试

实验 3. 模拟海量 QoS 服务请求不断到达的场景,测试系统链路和节点资源饱和度的波动情况,以验证不同的覆盖路由算法的负载均衡能力.实验通过链路的传输负荷(饱和度)和节点的计算负荷(饱和度)的波动率的方差(Residual Link Capacity Deviation,RLCD)、(Residual Node Capacity Deviation,RNCD)<sup>[5-6]</sup>两个 QoS 指标衡量算法对于系统的总体负载均衡程度的影响,定义如下:

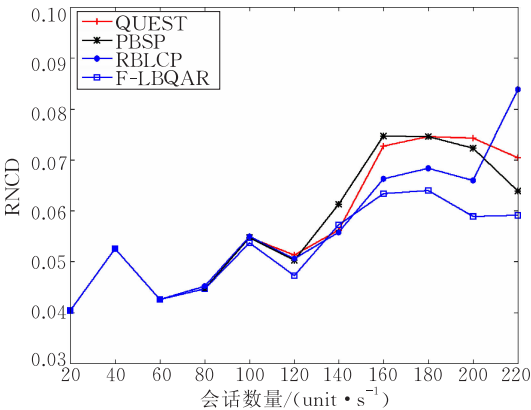
$$RNCD=\sqrt{\sum_{i=0}^k(C(n_i)-\bar{C})^2/k}$$
 (21)

$$RLCD=\sqrt{\sum_{j=0}^m(B(l_j)-\bar{B})^2/m}$$
 (22)

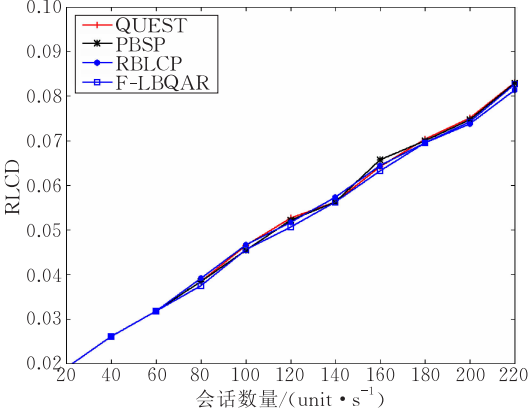
$C(n_i),B(l_j)$ 为网络中物理节点和链路的资源饱和度函数, $\bar{C},\bar{B}$ 为系统两类资源的平均饱和度, $k$ 和  $m$  分别是节点和链路总数.

实验首先基于 GT-ITM 分别生成 2 种随机分布物理网络,网络大小均为 100 节点;一种网络中节点间存在一条双向链路为 0.1 概率( $scale=10\times 10$ ),每条链路的带宽为 2Mbps,随机设定 30%的物理节点为覆盖节点,每个节点最大计算能力为 20 unit/s,该网络属于计算资源相对稀疏的网络拓扑;另一种网络节点间存在一条双向链路为 0.02 概率( $scale=100\times 100$ ),每条链路的带宽为 1Mbps,随机设定 30%的物理节点为覆盖节点,每个节点最大计算能力为 40 unit/s,该网络属于传输资源相对稀疏的网络拓扑;通过 OverlayTraffic 模拟到达率为 0~250 unit/s 的 CBR 业务流,平均流量为 300 kbps,分别测试 QUEST、PBSP、RBLCP 及 F-LBQAR 算法在不同拓扑及负荷量情况下的 RNCD 与 RLCD,实验结果如图 5、图 6 所示.

从图 5、图 6 的实验结果可以观察到,随着负荷量上升,所有资源饱和度的波动率也会呈上升趋势,

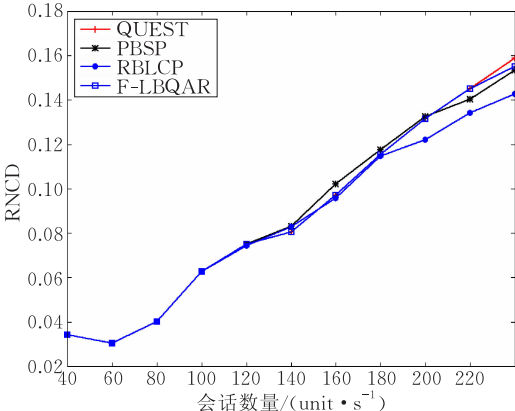


(a) RNCD变化曲线示意图

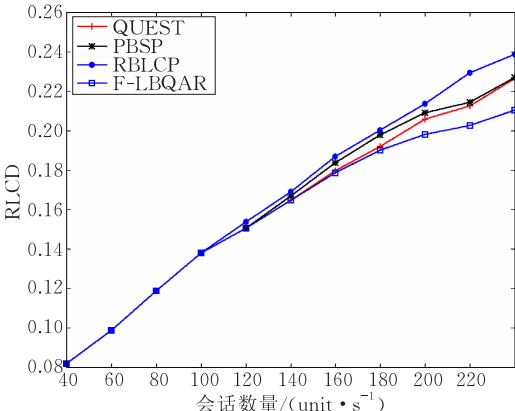


(b) RLCD变化曲线示意图

图 5 负载均衡波动率测试实验——100 点计算资源瓶颈网络



(a) RNCD变化曲线示意图



(b) RLCD变化曲线示意图

图 6 负载均衡波动率测试实验——100 点传输资源瓶颈网络



这说明负载会增大系统不均衡程度. 而当拥塞过载发生以后, 波动率反而会有所下降, 这是因为过载的服务节点和链路会拒绝新的请求, 导致负荷自然分布到空闲的资源上. 图 5 描述了计算资源相对稀疏网络上各种算法的均衡情况, 可以看出该网络上高负荷情况下计算资源出现明显拥塞(超过 140 unit/s), 而传输资源波动仍处在上升过程, 说明传输资源仍未饱和. 但图 5 结果说明在这种计算资源相对稀缺的网络环境下, 相对传统算法, F-LBQAR 算法能够在系统过载的情况下降低 3%~8%左右的 RNCD, 而 RLCD 值与其它算法基本保持在同一水平. 图 6 结果与图 5 相类似, 但由于节点拥塞会对多条链路造成影响, 因此传输拥塞(图 6(b))不如计算拥塞(图 5(a))明显, 然而其结果仍然表明 F-LBQAR 算法仍能够在系统过载的情况下降低 2%~6%左右的 RLCD (超过 160 unit/s), 而 RNCD 与其它算法基本保持在同一水平.

分析两种实验数据结果可以得出结论: F-LBQAR 相对传统算法具有更好地资源均衡分配能力, 特别是有利于优化系统紧缺资源的分布状况, 在高负荷业务环境及同样的吞吐流量下能够降低 2%~8%左右的资源饱和度波动率(RLCD、RNCD).

## 6 总 结

基于覆盖网络环境解决分布式应用的负载均衡 QoS 路由是一个非常重要的研究方向, 而传统的覆盖网络 QoS 路由算法未考虑资源瓶颈对路由选择的影响, 容易造成系统的局部拥塞, 本文提出一种公平负载均衡 QoS 覆盖路由算法 F-LBQAR, 能够降低资源瓶颈引起的网络拥塞概率, 提高 QoS 服务成功率和系统吞吐量. 该方法相对传统算法具有两个创新之处:

(1) 在路由模型中引入了路径服务资源利用率的公平指数, 依据资源消耗代价模型推导负载均衡启发式函数, 有利于选出阻塞概率较小的服务路径, 且总体代价仍能保持相对较低水平;

(2) 基于全局系统资源的瓶颈程度计算路由约束项的权重因子, 有利于瓶颈资源自适应的优化分布. 实验结果证明相对传统算法, F-LBQAR 算法在高负荷情况下能提高 1%~5%服务成功率(QSSR), 降低 2%~8%负载波动(RNCD, RLCD), 该研究成果将有助于分布式系统开发高性能的资源服务调度策略.

F-LBQAR 算法目前仅考虑了覆盖网络端到端的单播服务路由, 尚未涉及覆盖组播路由的负载均衡问题, 本文未来研究工作重点将考虑覆盖组播路由情况下优化资源分布, 同时考虑节点度、组播树延迟等更多 QoS 参数的约束.

## 参 考 文 献

- [1] David Andersen, Hari Balakrishnan, Frans Kaashoek, Robert Morris. Resilient overlay networks//Proceedings of the 18th ACM Symposium on Operating Systems Principles (SOSP). Banff, Canada, 2001: 131-145
- [2] Jin J, Nahrstedt K. Source-based QoS service routing in distributed service networks//Proceedings of the IEEE International Conference on Communications 2004 (ICC2004). Paris, France, 2004: 2036-2041
- [3] Gu Xiaohui. SpiderNet: A quality-aware service composition middleware[Ph. D. dissertation]. University of Illinois at Urbana-Champaign, 2004
- [4] Gu Xiaohui, Nahrstedt K, Chang R N, Ward C. QoS-assured service composition in managed service overlay networks//Proceedings of the 23rd International Conference on Distributed Computing Systems. Rhode Island, USA, 2003: 194-201
- [5] Li Z, Mohapatra P. QRON: QoS aware routing in overlay networks. IEEE Journal on Selected Areas in Communications, 2004, 22(1): 29-40
- [6] Ying Bi-Di, Chen Hui-Fang, Zhu Sun-Bin, Qiu Pei-Liang. RBLCP: Improved QoS-aware routing algorithm for overlay networks. Journal of Zhejiang University, 2007, 41(4): 646-649(in Chinese)  
(应必娣, 陈惠芳, 朱孙斌, 仇佩亮. RBLCP: 一种覆盖网络的 QoS 路由改进算法. 浙江大学学报, 2007, 41(4): 646-649)
- [7] Li Wen-Zhong, Guo Sheng, Xu Ping, Lu Sang-Lu, Chen Dao-Xu. An adaptive load balancing algorithm for service composition. Journal of Software, 2006, 17(5): 1068-1077 (in Chinese)  
(李文中, 郭胜, 许平, 陆桑璐, 陈道蓄. 服务组合中一种自适应的负载均衡算法. 软件学报, 2006, 17(5): 1068-1077)
- [8] Dijkstra E W. A note on two problems in connexion with graphs. Numerische mathematik, Springer, 1959, 1(1): 269-271
- [9] Pan Yun, Yu Zhen-Wei, Wang Xing-Gang, Feng Yan-Li. Load balance problem on Overlay multicast routing. Journal of Electronics & Information Technology, 2007, 29(3): 739-742(in Chinese)  
(潘耘, 余镇伟, 王行刚, 冯烟利. Overlay 组播路由中负载均衡问题的度量. 电子与信息学报, 2007, 29(3): 739-742)
- [10] Xu Lei-Ming, Pang Bo, Zhao Yao. NS and Network Simulation. Beijing: Posts and Telecommunications Press, 2003(in Chinese)

(徐雷鸣, 庞博, 赵耀. NS 与网络模拟. 北京: 人民邮电出版社, 2003)

- [11] Jain R K, Chiu W, Hare W. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Digital Equipment Corporation: Technical Report DEC-TR-301, 1984
- [12] Nakao A, Peterson L, Bavier A. A routing underlay for overlay networks//Proceedings of the ACM SIGCOMM'03 Conference on Computer Communications. Karlsruhe, Germany, 2003; 11-18
- [13] Bovy C J, Mertodimedjo H T, Hooghiemstra G. Analysis of

end-to-end delay measurements in Internet//Proceedings of the Passive and Active Measurement Workshop-PAM'2002. Fort Collins, CO, USA, 2002; 1-8

- [14] Strauss J, Katabi D, Kaashoek F. A measurement study of available bandwidth estimation tools//Proceedings of the ACM SIGCOMM Internet Measurement Conference 2003. Miami, Florida, USA, 2003; 39-44
- [15] Calvert K, Doar M, Zegura E. Modeling Internet topology. IEEE Transactions on Communications, 1997, 35(6): 160-163



**CAO Xue-Song**, born in 1979, Ph.D..

His research interests include multimedia network protocol, high-quality distributed-multimedia system, overlay network, QoS.

**HU Rui-Min**, born in 1964, Ph. D. , professor, Ph. D.

supervisor. His research interests include multimedia network communication, multimedia signal processing and programming.

**WANG Zhao-Ping**, born in 1981, Ph. D. . His research

interests include multimedia network communication system, overlay network, QoS.

## Background

Current Internet network can only provide Best-effort Service, lack of scalability of resource scheduling mechanism. It is difficult to providing reliable QoS assurance for distributed multimedia application. The authors are committed to study network multimedia application such as video surveillance, video conference, and understand profoundly these areas are facing the issue of quality of service requirements. It is recognized that the network architecture must be changed in order to overcome the problem. Overlay network has become the hotspot of NGN technology for its advantage of improving distributed network application service of quality without changing fundamental network structure. This paper puts forward a fair load balancing QoS routing algorithm in overlay network. The goal is to provide high-performance resource scheduling mechanism for large-scale distributed application based on overlay network environment, to achieve load balancing of Network Computing Services. Simulation

results show that the algorithm can effectively improve system throughput and the degree of load balancing. The research results explore highly efficient and reliable QoS control strategy from the perspective of the application layer, to provide scientific ideas for designing large-scale distributed multimedia systems. It is of great theoretical significance. And it is also important for research and development of video surveillance systems, video conferencing systems and other distributed system products.

The research project is funding by National Basic Research Program (973 Program) of China (2009CB320906), National Natural Science Foundation of China (60772106), Scientific and Technological Project of Wuhan City and Education Research Plan of Guangdong Province (2006D90104005). The projects have focused on studying the theories and technologies including network multimedia application, network quality of service and so on.