

文章编号:1001-9081(2008)07-1831-04

# 基于公平服务多维加权轮询调度的 IPTV 组播流量控制策略

张 鹏,俞柏锋,庞 琳,张兴明

(国家数字交换系统工程技术研究中心, 郑州 450002)

(zhangpeng19810916@gmail.com)

**摘 要:**根据大规模接入汇聚路由器(ACR)的总体技术规范,IPTV 组播复制在 ACR 的以太网分合路器(EMD)中进行,而 IPTV 组播业务的增长会对其他业务造成冲击,因此需要对 IPTV 组播流量进行控制,以保证其他业务的带宽分配,针对 IPTV 流量控制中的关键环节——队列调度,提出了一种公平服务多维轮询调度策略(FSMWRR),性能仿真的结果表明该调度策略能够满足 IPTV 业务的低时延、时延抖动的需求,同时保证各个业务带宽分配的公平性。

**关键词:**IPTV 组播复制;队列调度;加权轮询调度

**中图分类号:**TP393.01 **文献标志码:**A

## Scheduling scheme of IPTV multicast traffic control based on fair service and multidimensional weighted round robin

ZHANG Peng, YU Bai-feng, PANG Lin, ZHANG Xing-ming

(National Digital Switching System Engineering and Technological R&D Center, Zhengzhou Henan 450002, China)

**Abstract:** According to the general technology criterion of Access Convergence Router (ACR), the IPTV multicast copy is processed in Ethernet Multiplexer-Demultiplexer (EMD). Because the increase of the IPTV multicast service will impact other service, the IPTV multicast traffic should be controlled to ensure the bandwidth allocation of other service. To solve the key portion - the queue schedule of IPTV multicast traffic, a scheme of Fair Service and Multidimensional Round Robin was presented, the result of simulation indicate that Fair Service and Multidimensional Weighted Round Robin (FSMWRR) can achieve the good delay and jitter performance for IPTV service and provide fair bandwidth allocation for all service.

**Key words:** IPTV multicast copy; queue schedule; Weighted Round Robin (WRR)

### 0 引言

在 IPTV 组播复制时,需要根据不同用户的定制情况将 IPTV 组播报文缓存到用户对应的队列中,通过某种队列调度方式对 IPTV 组播报文的分组和其他业务的分组进行调度输出,目前通常采用的队列调度方式有: FIFO (First-In-First-Out), PQ (Priority Queuing), WRR (Weighted Round Robin) 和 WFQ (Weighted Fair Queuing) 等。其中 FIFO 严格根据报文的达到顺序调度每个报文输出,因此这种调度方式无法区分业务的优先级,不能保证高优先级的业务的服务质量;而 PQ 是针对关键业务应用设计的,在队列调度时,PQ 严格按照优先级从高到低的次序,优先调度输出较高优先级队列中的报文,当较高优先级队列为空时,再输出较低优先级队列中的报文。这样,将关键业务的报文放入较高优先级的队列,将非关键业务的报文放入较低优先级的队列,可以保证关键业务的报文被优先传送,非关键业务的分组在处理关键业务数据的空闲间隙被传送,但当较高优先级队列中长时间有报文存在,那么低优先级队列中的报文将可能一直得不到服务,因此 PQ 方式无法保证各个业务带宽分配的公平性。为了克服这种不公平性,文献[1]提出了 WRR,在 WRR 算法中,队列被划分为不同的服务类别,不同服务类别的流进入不同的队列中,每个队列按一定的准则轮询接受服务,因为 WRR 可以为每个队

列分配不同的归一化权值,因此它能为各个业务提供有区别的服务。权值代表各业务所占的输出链路带宽的比例,队列  $i$  对应的权值为  $W_i$ ,且  $\sum_i W_i = 1$ 。与此同时每个队列还要维护一个计数器,每次轮询时只有计数器非零的队列才能输出与其权值相应数量的分组。每个计数器的初值为各自的权值,每发送一个分组计数器就减 1,当所有队列的计数器都为零时,则重置计数器初值为各自的权值。WRR 调度机制顺次为每个队列服务,允许分配了高比例带宽的队列在一次服务时调度多个分组,或者限定了每次服务一个队列只传送一个分组,但在一次轮询服务中可多次访问分配到高比例带宽的队列,因此 WRR 可确保低优先级业务始终能得到额定的服务资源,但由于轮询算法的本身的特点,使得 WRR 无法保证高优先级业务低时延的要求。为解决 PQ 和 WRR 各自的缺点,文献[2]结合 PQ 和 WRR 的优点提出了混合调度算法 PQWRR,此算法用来保证高优先级业务的低时延要求和不同业务带宽分配的公平性,但通过对 PQWRR 的分析,其在不同业务的带宽分配方面仍有不足,尤其是在负载高的情况下,无法保证低优先级业务的输出带宽,针对此问题,本文提出了一种公平服务多维轮询调度策略,其基本思想是对进入虚拟优先级队列 (Virtual Priority Queue, VPQ) 的不同优先级业务流进行多指针的加权轮询调度,从而决定不同用户的不同业

收稿日期:2008-01-23;修回日期:2008-03-28。 基金项目:国家 863 计划项目(2004A103130)。

作者简介:张鹏(1981-),男,北京人,硕士研究生,主要研究方向:IP 宽带网络; 俞柏锋(1982-),男,浙江余姚人,硕士研究生,主要研究方向:视频编解码、嵌入式开发; 庞琳(1982-),女,硕士,山东济南人,主要研究方向:宽带信息网络; 张兴明(1963-),男,河南新乡人,教授,主要研究方向:IP 宽带网络。

务的分组输出。

### 1 FSMWRR 的基本原理

FSMWRR 的原理图如图 1 所示,用户的缓冲队列采用 VPQ 方式,每个用户不同的 VPQ 为  $VPQ(n,k)$ ,其中  $n$  为不同用户,  $1 \leq n \leq N, k$  表示用户订制的不同业务,  $1 \leq k \leq K$ 。该算法采用多指针加权调度方式,用户业务指针  $usp(k)$  根据每类业务所分配的带宽对应的权值进行加权轮询,轮询频率为  $\mu_1$ ; 用户带宽指针  $ubp_k(n)$  根据每个用户总带宽所对应的权值进行加权轮询,轮询频率为  $\mu_2$ , 如果某一用户当前所使用的业务带宽超出了该业务的带宽限制,但没有超出该分配给该用户的总带宽,那么该算法仍允许该业务顺利调度输出,此时处理机将会发出针对指针  $ubp_k(n)$  轮询频率  $\mu_2$  的更新指令,将其改为与此业务带宽匹配的频率进行轮询。在调度不同用户的不同业务分组到输出链路时,先检查指针  $usp(k)$  所指向的某一业务,再根据  $ubp_k(n)$  对此业务中的不同用户进行加权轮询,从而确定某一  $VPQ(n,k)$ , 若此  $VPQ(n,k)$  非空,则调度该队列中的分组输出,否则服务下一个队列。

在不同用户的缓冲队列中,VPQ 的个数是相同的,并且

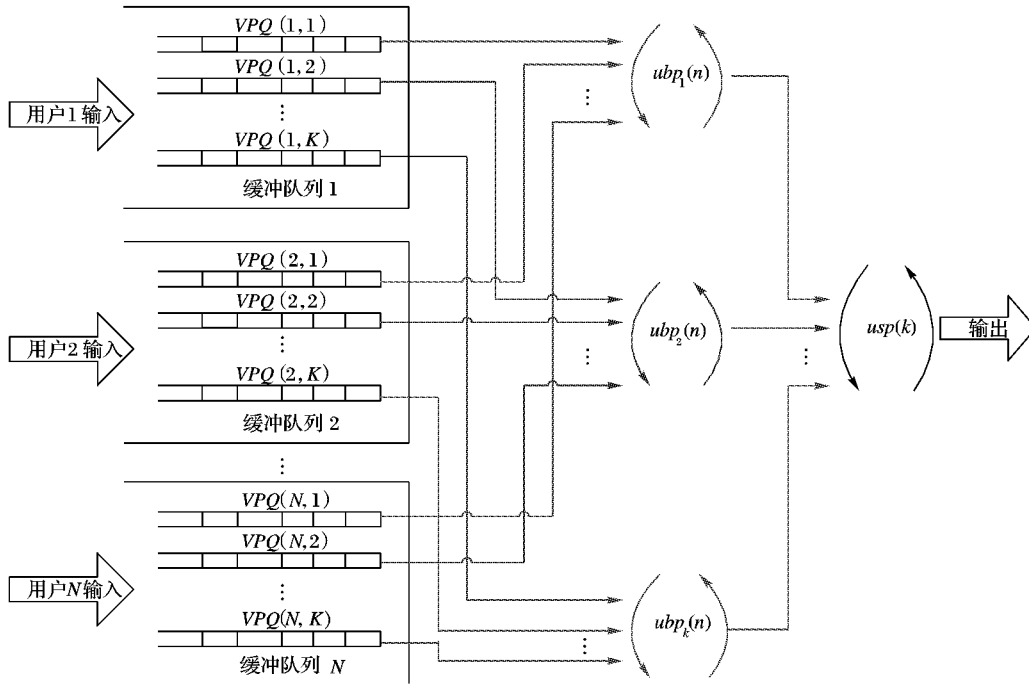


图1 FSMWRR 的原理

### 2 FSMWRR 的数学分析

为简化对 FSMWRR 算法的讨论,假设只有两个用户,他们各自订制了两类业务,第一类业务是对时延和丢包要求都很高的实时业务(如:IPTV 业务),第二类业务是对时延和丢包的要求都较低的业务(如:E-mail 业务),因此第一类业务的优先级高于第二类业务。第一个用户的第一、二类业务的分组到达时分别进入  $VPQ(1,1)$ 、 $VPQ(1,2)$ ; 同理,第二个用户的第一、二类业务分组到达时分别进入  $VPQ(2,1)$ 、 $VPQ(2,2)$ 。业务分组的输出通过指针  $usp(k)$  和  $ubp_k(n)$  联合调度实现,两个用户两类业务的 FSMWRR 调度算法示意图如图 2 所示。

FSMWRR 算法的数学分析具体如下:

1) 尽管到达 VPQ 的业务流的分布并不都是泊松分布,但类似的研究表明,利用泊松模型与用更复杂的模型(如流体

对应位置上的 VPQ 属于同一优先级业务,例如:第  $n$  个缓冲队列的第  $k$  个 VPQ,与第  $n+1$  个缓冲队列的第  $k$  个 VPQ 的优先级相同。其中一个用户的不同业务的优先级由高到低分别为  $1, 2, \dots, K$ 。指针  $usp(k)$  及指针  $ubp_k(n)$  的初始化是从最高优先级业务中最高优先级队列开始,选择优先级最高且非空 VPQ 的分组调度输出。而指针  $usp(k)$  和  $ubp_k(n)$  的更新则服从当前服务在下一个调度周期优先级最低的原则。例如:某一时刻指针  $ubp_k(n)$  指向  $VPQ(n,k)$ , 且调度了与其权值相匹配数量的分组,则  $n$  在下一个调度周期将更新为  $n = (n+1) \bmod N$ , 使得  $VPQ(n,k)$  在下次调度周期内的匹配优先级最低,从而保证用户不同业务的服务公平性。

因为 ACR 中的 IPTV 业务对时延及时延抖动的要求非常高,因此在 ACR EMD 中进行 IPTV 组播复制输出时,我们把最高优先级 VPQ 设置为低时延队列,而其他优先级 VPQ 与低时延 VPQ 的轮询遵循交替优先的原则,即指针  $usp(k)$  的轮询顺序为  $1, 2, 1, 3, 1, 4, \dots, 1, K$ 。交替优先级模式不仅保证了高优先级业务对时延的需求,而且其交替优先的服务顺序更进一步保证了 IPTV 业务及其他业务对带宽公平分配的要求。

模型,马尔可夫调制的 on-off 模型等)得到的结果是相似的,特别是在缓冲区容量有限的情况下<sup>[3-4]</sup>。因此,假设两类业务分组的到达过程都是泊松过程,且相互独立。针对缓冲队列的管理方式,我们采用一种区分优先级随机早期丢弃(Differentiated Priority Random Early Drop, DPRED)队列管理算法进行丢包。

DPRED 队列管理算法的基本思想是就为不同优先级 VPQ 中的分组设定不同的随机丢弃概率。如果有  $K$  类不同优先级的业务,则为这  $K$  类业务设置不同的随机丢弃概率  $p_k, k = 1, 2, 3, \dots, K$ 。优先级由高到低的 VPQ 有  $K$  个不同的最大阈值门限  $\max_{thk}, k = 1, 2, 3, \dots, K$ 。其中  $\max_{th1} \geq \max_{th2} \geq \dots \geq \max_{thK}$ , 阈值门限  $\min_{thk}$  为最大阈值门限  $\max_{thk}$  的一半。设  $t$  时刻队列  $k$  的长度为  $N_k(t)$  (包括正在输出的分组),新到达第  $k$  个 VPQ 的分组的随机丢弃概率为:

$$P_i = \begin{cases} 0, & N_k(t) \leq \min_{thk} \\ \frac{N_k(t) - \min_{thk}}{\max_{thk} - \min_{thk}}, & \min_{thk} < N_k(t) < \max_{thk} \\ 1, & \max_{thk} \leq N_k(t) \end{cases} \quad (1)$$

根据式(1)可得,当分组到达 VPQ 时,业务流  $k$  的到达强度  $\lambda$  为:

$$\lambda = \begin{cases} \lambda_k, & N_k(t) \leq \min_{thk} \\ \lambda'_k, & \min_{thk} < N_k(t) < \max_{thk} \\ 1, & \max_{thk} \leq N_k(t) \end{cases} \quad (2)$$

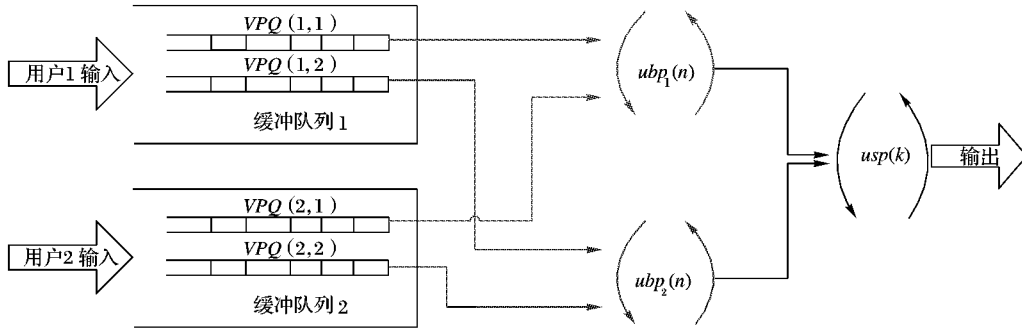


图 2 FSMWRR 调度算法示意图

设任意时刻  $t$  时,第  $k$  类业务的数据包个数为  $N_k(t)$ ,  $k = 1, 2$ 。令  $N_t = \{N_1(t), N_2(t)\}$ , 由于两类业务流的到达过程均为泊松过程,且服务过程服从参数为  $\mu_{kn}$  ( $k = 1, 2; n = 1, 2$ ) 的负指数分布,由参考文献[5]可知,  $\{N_t, t \geq 0\}$  为二维马尔可夫链。记  $P_{k,n}$  为稳态概率,由于缓存队列空间有限,所以其稳态分布必然存在。并设第一类业务和第二类业务的最小及最大阈值分别为:  $\min_{thk} = L_k, \max_{thk} = H_k$  ( $k = 1, 2$ ), 则参考文献[6] 两个用户两类业务所组成的系统中  $VPQ(1,1)$  和  $VPQ(2,1)$  系统状态转移图如图 3 所示。

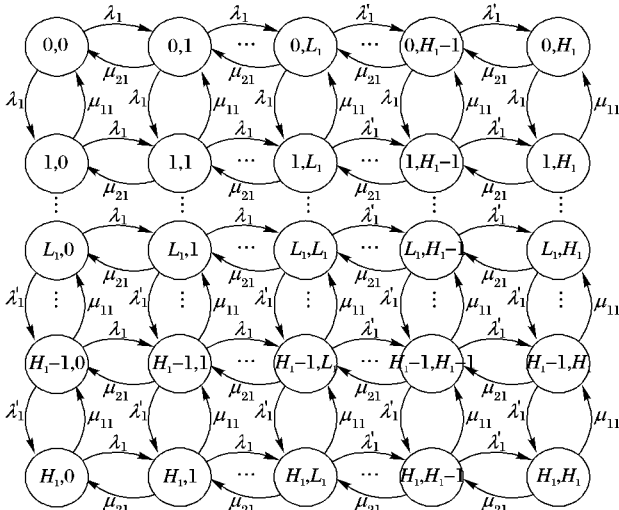


图 3  $VPQ(1,1)$  和  $VPQ(2,1)$  系统状态转移图

由状态转移图 3 得出的  $VPQ(1,1)$  和  $VPQ(2,1)$  系统的稳态概率方程组可知存在稳态解  $P_{k,n}$ , 则  $VPQ(1,1)$  和  $VPQ(2,1)$  的分组平均队长可表示为:

$$\begin{cases} L_{a(1,1)} = \sum_{n=0}^{H_1} k P_{k,n} \\ L_{a(2,1)} = \sum_{k=0}^{H_1} k P_{k,n} \end{cases} \quad (3)$$

当经过  $M$  次轮询后系统达到稳态时,记其丢包数为  $L_{d(k,n)}$ , 来到  $VPQ(k,n)$  中数据包总数为  $L_{t(k,n)}$ ,  $VPQ(k,n)$

其中  $\lambda'_k < \lambda_k$  ( $k = 1, 2$ )。两类业务流到达后,分别在各自的 VPQ 中排队。

2) 由于轮询时调度输出包的个数等于用户订制业务的总带宽乘以其相应的权值,因此设订制第  $k$  类业务的第  $n$  个用户的权值为  $C_{k,n}$ , 则订制不同业务的不同用户的服务速率为  $\mu_{k,n} = C_{k,n} \mu_2$  ( $k = 1, 2; n = 1, 2$ )。设订制第一类业务的第一个用户及第二个用户分别服从参数为  $\mu_{11}$  和  $\mu_{12}$  的负指数分布, 则订制第二类业务的第一个用户及第二个用户分别服从参数为  $\mu_{21}$  和  $\mu_{22}$  的负指数分布。

的平均队列长度为  $L_{a(k,n)}$ , 被调度输出的数据包个数为  $L_{s(k,n)}$ , 则:

$$L_{d(k,n)} = L_{t(k,n)} - L_{a(k,n)} - L_{s(k,n)} \quad (4)$$

因此  $VPQ(1,1)$  和  $VPQ(2,1)$  的丢包数为:

$$L_{d(1,1)} = L_{t(1,1)} - L_{a(1,1)} - L_{s(1,1)} = \lambda_1 \frac{M}{\mu_1} - \sum_{n=0}^{H_1} k P_{k,n} - C_{11} \mu_2 \frac{M}{\mu_2} = \lambda_1 \frac{M}{\mu_1} - \sum_{n=0}^{H_1} k P_{k,n} - C_{11} M \quad (5)$$

$$L_{d(2,1)} = L_{t(2,1)} - L_{a(2,1)} - L_{s(2,1)} = \lambda_1 \frac{M}{\mu_1} - \sum_{k=0}^{H_1} n P_{k,n} - C_{12} \mu_2 \frac{M}{\mu_2} = \lambda_1 \frac{M}{\mu_1} - \sum_{k=0}^{H_1} n P_{k,n} - C_{12} M \quad (6)$$

同理可得  $VPQ(1,1)$  和  $VPQ(2,2)$  系统,  $VPQ(1,2)$  和  $VPQ(2,1)$  系统,  $VPQ(1,2)$  和  $VPQ(2,2)$  系统的丢包表达式。

由此可以看出 VPQ 的丢包数与它所对应的权值相关, 即: 当业务分组的到达强度和轮循频率确定的情况下, 分配的权值增大, 丢包数减少, 从而分组输出时延降低, 因此, 对不同业务分配不同的权值, 通过多指针的加权轮询, 可以使 FSMWRR 满足高优先级业务低时延的性能要求, 保证了各业务之间公平地分配带宽。

### 3 性能仿真

本文从公平性和有效性两个方面评估 FSMWRR 的性能, 其中带宽分配比例衡量公平性; 有效性则由平均时延和时延抖动衡量。假设每个用户订制了三类业务, 按优先级有高至低分别为实时性业务(如: IPTV)、可靠性业务(如: 电子商务)和一般 Internet 性业务(如: 浏览网页), 三类业务的权值依次为 0.6, 0.3, 0.2。各业务的分组到达服从泊松分布, 每个分组长度为 64 字节, 它们的平均分组到达率分别为: 50%、40% 和 10%。性能仿真图如图 4、5 所示。

图 4 表明, 在三类业务的带宽分配的公平性方面, FSMWRR 优于 PQWRR; 从图 5 的仿真结果可以看出在 IPTV

组播分组平均时延和时延抖动上,FSMWRR 的性能接近于 PQWRR。

#### 4 结语

针对 ACR 中 IPTV 组播流量控制问题,本文提出了一种基于公平服务多维指针加权轮询调度的策略,通过与

PQWRR 的性能比较,仿真结果表明:在不同业务的带宽分配的公平性方面,FSMWRR 优于 PQWRR;在 IPTV 组播复制输出的平均时延和时延抖动上,FSMWRR 与 PQWRR 的性能相当,因此,在 IPTV 组播复制时,FSMWRR 能够为 IPTV 业务提供较好的时延、时延抖动性能,同时保证了各个业务的带宽分配更为公平。

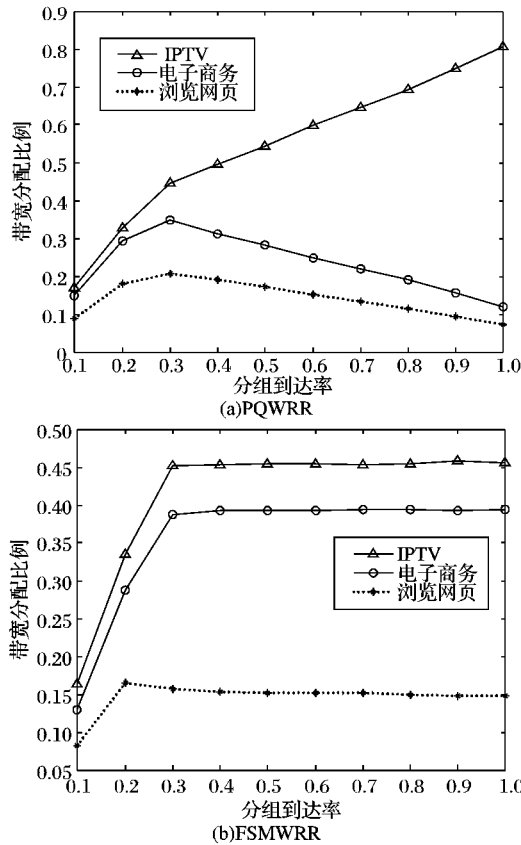


图 4 带宽分配比较

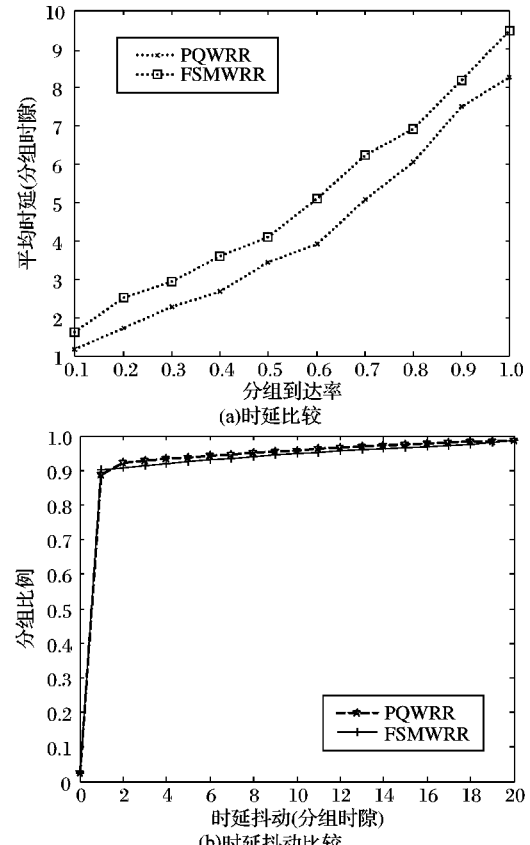


图 5 时延、时延抖动比较

#### 参考文献:

- [1] KATEVENIS M, SIDIROPOULOS S, COURCOUBETIS C. Weighted round-robin cell multiplexing in a general-purpose ATM switch chip [J]. IEEE Journal on Selected Areas in Communications, 1997, 9(8): 1265 - 1279.
- [2] MAO J, MOH W M, WEI B. PQWRR scheduling algorithm in supporting of DiffServ[C]//IEEE International Conference on Communications. Helsinki, Finland: IEEE Press, 2001, 3: 679 - 684.
- [3] MAY M, BOLOT J C, JEAN-MARIE A. Simple performance mod-

- els for differentiated services schemes for the Internet [C]// IEEE INFOCOM'99. New York: IEEE Press, 1999: 1385 - 1394.
- [4] SAHU S, TOWSLEY D, KUROSE J. A quantitative study of differentiated services for the Internet [C]// IEEE Globecom'99. [S. l.]: IEEE Press, 1999.
- [5] 应辉,唐小我. 排队论: 基础与应用[M]. 成都: 电子科技大学出版社, 2000: 55 - 60.
- [6] 王婷. 大规模接入汇聚路由器复用单元中的流量管理技术及实现[D]. 郑州: 中国人民解放军信息工程大学, 2007.

(上接第 1830 页)

- [2] HEIDEMANN J, LI Y, SYED A, *et al.* Research challenges and applications for underwater sensor networking [C]// Processing of the IEEE Wireless Communications and Networking Conference. Las Vegas, Nevada, USA: IEEE Press, 2006: 228 - 235.
- [3] MIN K P, RODOPLU V. UWAN-MAC: An energy-efficient MAC protocol for underwater acoustic wireless sensor networks [J]. IEEE Journal of Oceanic Engineering, 2007, 32(3): 710 - 720.
- [4] AITSAADI N, ACHIR N, BOUSSETTA K, *et al.* Differentiated underwater sensor network deployment [C]// Europe OCEANS. Washington, DC, IEEE Press, 2007: 1 - 6.
- [5] JURDAK R, RUZZELLI A G, O' HARE G M P, *et al.* Design considerations for deploying underwater sensor networks [C]// International Conference on Sensor Technologies and Applications. Valencia Spain: IEEE Press, 2007: 227 - 232.

- [6] 洪申. 水声学原理[M]. 哈尔滨: 哈尔滨船舶工程学院出版社, 1990.
- [7] 周延. 水声通信信道射线模型的研究[D]. 西安: 西北工业大学, 2003.
- [8] DESAI M, ANJUNATH M D. On the connectivity in finite Ad Hoc networks [J]. IEEE Communication Letters, 2002, 6(10): 437 - 439.
- [9] NIKOLAY K. Mobile radio channel modeling in Matlab [J]. Journal of Radio Engineering, 2003, 12(4): 12 - 16.
- [10] YANG W B, YANG T C. High frequency channel characterization for Mary frequency shift keying underwater acoustic communication [J]. Journal of Acoustics Science of America, 2006, 120(5): 2615 - 2626.