

用于缓存管理的最高速流早期丢弃策略

温蜀山, 李乐民, 孙海荣

(电子科技大学宽带光纤传输与通信系统技术国家重点实验室, 四川成都 610054)

摘 要: 为了提高 Internet 在分组转发时对 TCP 的公平性, 本文提出一种用于缓存管理的最高速流早期丢弃策略, 简称 HRFED, 并研究了它与 FCFS 和 FQ 调度算法结合时的性能. 本文用解析方法证明 HRFED 可以用 EWMA 加权系数控制突发业务成为最高速流的容限, 又用多个仿真实验, 证明 HRFED 可以在保持较高的链路利用率的同时, 获得比 LQD、RED 和 TD 更好的公平性, 并维持比 LQD 和 TD 更短的平均排队等待时间.

关键词: 最高速流早期丢弃; 队列管理; TCP; 带宽; 公平性

中图分类号: TP393.07 **文献标识码:** A **文章编号:** 0372-2112 (2002) 10-1455-04

The Scheme of Highest Rate-Flow Early Discard for Buffer Management

WEN Shu shan, LI Le min, SUN Hai rong

(National Key Lab of Broadband Optical Fiber Transmission and Communication Networks,
University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China)

Abstract: In order to enhance the fairness to TCP when packets are forwarded in the Internet, this paper presents the scheme of Highest Rate Flow Early Discard (HRFED) for buffer management, and researches its performance combined with FCFS and FQ scheduling algorithms. The EWMA weight in HRFED is proved analytically to control the margin of bursty traffic that will become highest rate flow. By simulations based on different network scenarios, HRFED is proved to get better fairness than LQD, RED and TD (Tail drop), and to maintain shorter average waiting time in the queue, as well as to keep high link utilization.

Key words: HRFED; buffer management; TCP; bandwidth; fairness

1 引言

在 Internet 中, 带宽的分配和利用不但与端到端拥塞控制机制(如 TCP)有关, 也受路由器(或交换机)缓存管理的影响, 必须将二者结合研究.

最易实现的缓存管理是 Tail Drop(简称 TD), 但它存在两个明显的缺陷, 一是带宽的分配不均衡, 二是队列长时间处于满状态, 增加了分组的端到端时延^[1].

随机早期检测(RED)是另一种应用广泛的缓存管理策略, 它对 TD 作了改进. RED 在缓存溢出之前就按某个概率随机丢弃到达分组, 队列越长, 丢弃概率越大, 这样, 在保证 TCP 总吞吐量的同时, 可以将队列长度维持在较短的限度内, 还能在一定程度上避免所有 TCP 流同时增加或减小发送速率的现象(即全局同步效应)^[2]. 然而, 研究表明, 在丢弃概率一定的情况下, TCP 的发送速率大致和 RTT 呈反比^[3,4]. 因为在特定时刻, RED 对所有的流都采用同一丢弃概率, 而在实际的 Internet 中, 不同 TCP 流的 RTT 往往具有较大差异, 因此, RED 无法对带宽作出公平的分配.

为了获得更好的公平性, 文献[5]提出了一种被称作最长队列丢弃(LQD)的缓存管理策略, 如果缓存尚有冗余空间, 就让到达分组参与排队, 否则就丢弃当前占用缓存最多的流的排头分组. 然而, 由于 Internet 中数据传输的突发性, 在某个时刻, 占用缓存最多的流并非就一定是较长时间意义上的最高速流, LQD 不利于短突发业务, 这是一个有损公平性的因素.

LQD 的另一个缺陷和 TD 类似, 即可能让队列长期处于满状态, 增加分组排队的等待时延和流的端到端时延.

为了进一步提高公平性, 并同时减小排队时延, 本文提出最高速率流早期丢弃的策略(HRFED, Highest Rate Flow Early Discard). 为考查该算法的性能, 本文将公平排队(FQ)和先到先服务(FCFS)的调度策略与 HRFED 结合研究, 文中分别记为 FQ-HRFED 和 FCFS-HRFED, 并与采用 FQ 和 FCFS 的 RED、LQD 等缓存管理策略作了比较.

2 HRFED 算法

HRFED 是一种共享缓存(但支持按流排队), 对不同流分别计算到达速率, 具有一定隔离和保护作用的算法. 算法模型见图 1. HRFED 包括速率统计和丢弃判决两个功能模块, 用于控制缓存分配. HRFED 和调度算法关系密切, 但逻辑上, 调度算法是独立于缓存管理的, 同一种缓存管理可以与不同的调度算法结合.

HRFED 设计原则: (1) 若需丢弃分组, HRFED 首先丢弃最高速流在队列中的排头分组, 如果最高速流在队列中已没有分组可供丢弃(尽管这种情况不常见), HRFED 就丢弃次高速流在队列中的排头分组. 以此类推. (2) 每隔时间 τ , 以 EWMA 方式对流速作一次刷新, 并按流速的大小顺序对流排序. (3) 对每个到达分组, 先让其进入缓存并排队, 如果缓存已满, 就按原则(1)丢弃一个分组. 如果缓存未满, 就按当前缓存的占用量 L 计算丢弃概率 p , 并以此概率按原则(1)作随机丢弃.

下面对以上原则作一些解释. 采用原则(1)是为了让 TCP 对网络拥塞有更快的反应, 如果 FCFS-HRFED 不采用按流排队(注意, 即使 FCFS-HRFED 也可按流排队), 由于所有流的分组混排在一个队列当中, 丢弃特定流的排头分组需要某种查找算法, 但可以构造特定的数据结构加速查找过程. 如果是 FQ-HRFED, 采用按流排队, 则可以很方便地丢弃某个流的排头分组.

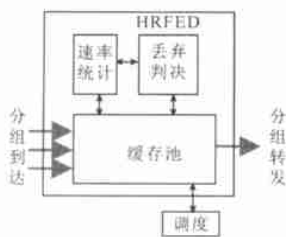


图1 HRFED 算法模型

原则(2)中, HRFED 以 EWMA 方式计算流的平均到达速率, 这有利于那些突发性强, 但长期平均吞吐量小的流. 为了计算各个流的速率, 路由器需要为每个流保留一个变量, 为流 i 保留的变量记为 B_i (初值为 0). 当分组到达时, 路由器根据分组头部的信息查看该分组所属的流^{*}, 如果该分组属于流 i , 则将分组长度值加到 B_i 上, 如果流 i 有分组被路由器丢弃, 则从 B_i 中减去该分组长度值, 因为这个分组不再会占用下游链路的带宽. 经过时间 τ 后, B_i 的值就为流 i 在时间 τ 内(将)占用下游链路带宽的分组长度之和. 设第 n 次观测后(时刻 $n\tau$, $n = 1, 2, \dots$), 流 i 的平均发送速率为 $M_i(n)$, B_i 的值为 $B_i(n)$, 令 $M_i(0) = 0$, 则按 EWMA 的规则, 有

$$M_i(n) = (1-w)M_i(n-1) + wB_i(n)/\tau \quad (1)$$

其中, w ($0 < w < 1$) 为 EWMA 加权系数. 每一次依照式(1)计算完 M_i 后, 都将 B_i 置 0, 便于下一次统计, 并按 M_i 的大小为流排序. 如果流 i 已结束发送, M_i 会逐渐递减, 当 M_i 小于某个既定的门限, 就释放为流 i 保留的状态变量. 本文仿真时将这个门限设定为 $40\text{Bytes}/\tau$, 这个速率对应于在 τ 内仅发送一个 TCP 的 ACK 分组, 同时, τ 的取值为下游链路传播时延 $\times 2$, 它对应于经过这段链路的 TCP 连接的 RTT 的下限.

在原则(3)中, p 与 L 的关系和 RED 中 p 与 EWMA 平均队长 avg 的关系相同. HRFED 采用 L 而不采用 avg , 是因为在原则(2)中已经用 EWMA 对到达速率作了平滑, 无需再采用一次 EWMA 将 L 转换成 avg . 我们曾对基于 avg 的随机丢弃和基于 L 的随机丢弃两种情形下的 HRFED 算法性能作了比较, 发现其在吞吐量、链路利用率、公平性以及分组的平均排队时延等方面的性能都很接近.

当流的数目较多时, 我们按如下方式定义一种适用于 TCP 的公平性系数: 在某个时间范围内, 若观察到流的数目为 n , 流 i 的有效吞吐量($goodput$)^{**} 为 G_i , ($i = 1, 2, \dots, n$), 则该时间范围的公平性系数 F 为

$$F = \left(\sum_{i=1}^n G_i \right)^2 / \left(n \sum_{i=1}^n G_i^2 \right) \quad (2)$$

不难证明, $F \in (0, 1]$, 越接近 1, 公平性越好.

3 用 EWMA 加权系数控制突发业务成为最高速流的容量

HRFED 中 EWMA 加权系数 w 的大小可以用来控制突

业务成为最高速流的容量. 若仅观察流 i , 设在观察之前, 流 i 的速率为 R_i 且时刻 0 的平均速率 $M_i(0) = R_i$ (即已达到平衡), 时刻 0 以后, 流 i 的速率增加到 R'_i ($R'_i > R_i$), 则时刻 τ 算得的平均速率 $M_i(1) = (1-w)R_i + wR'_i$, 时刻 2τ 算得的平均速率 $M_i(2) = (1-w)M_i(1) + wR'_i = (1-w)^2R_i + w(1-w)R'_i + wR'_i$, 以此类推, 可得时刻 $n\tau$ 的平均速率

$$M_i(n) = R_i(1-w)^n + R'_i w \sum_{k=0}^{n-1} (1-w)^k \\ = R_i(1-w)^n + R'_i [1 - (1-w)^n]$$

若令 $R'_i = R_i + \Delta R_i$, 则 $M_i(n) = R_i + \Delta R_i - \Delta R_i(1-w)^n$.

设流 i 的速率增加前, 流 j 是最高速流, 其速率为 R_j 且 $M_j(0) = R_j$, $R_i < R_j < R'_i$. 我们希望了解当 n 为多少时, $M_i(n) \geq M_j(n)$, 等价于 $R_i + \Delta R_i - \Delta R_i(1-w)^n \geq R_j$, 于是有

$$n \geq \log_{1-w} \frac{R_i + \Delta R_i - R_j}{\Delta R_i} \quad (3)$$

式(3)给出了某个流替代原有最高速流而成为新的最高速流的时间, 亦即允许某个突发业务持续到达, 而不被当作最高速流丢弃的时间. 特殊地, 如果 $R_i = 0$, $\Delta R_i = R'_i$, 则

$$n \geq \log_{1-w} (1 - R_j/R'_i)$$

当 w 取不同的值时, n 的下限随 R_j/R'_i 变化的关系示于图 2. 虽然本文主要讨论 HRFED 对提高 TCP 流的公平性的作用, 但只要灵活地配置 w ,

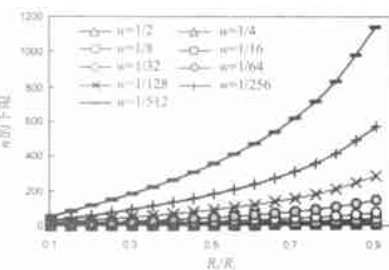


图2 n 的下限与 R_j/R'_i 的关系

HRFED 也可用于

具有不同优先级的业务. 例如, 若流 i 的优先级高于流 j , 则指定两个加权系数 w_i, w_j ($w_i < w_j$), 当流 i 的速率递增时, 采用 w_i , 递减时采用 w_j , 对流 j 可以反过来, 递增时采用 w_j , 递减时采用 w_i , 这样流 j 就比流 i 更容易成为最高速流.

4 仿真实验

我们在 NS 网络仿真器^[6]中增加了支持 FCFS/FQ 调度的 HRFED 和 LQD 的功能模块, 作了以下的仿真实验. HRFED 和 RED 的早期丢弃门限 $minth$ 和 $maxth$ 分别设为缓存容量的 $1/3$ 和 $2/3$, $maxp$ 设为 0.002, 平均分组长度为 1000Bytes, RED 算法中计算 avg 的 EWMA 加权系数为 0.002. 网络拓扑如图 3, HRFED 的 EWMA 加权系数 $w = 2^{-8} = 1/256$.

4.1 仅有两个 TCP 流的情形

两个 TCP 流分别以图 3 所示的 S_1, S_2 为源, D_1, D_2 为宿 ($i \geq 3$ 的 S_i 和 D_i 未使用) 传输 FTP 业务. R_1, R_2 之间构成

* 对于 IPv4, 可以用源地址和目的地址对区分不同的流, 对于 IPv6, 可以用与 IPv4 同样的方法, 也可以进一步结合分组头部的流标识来区分不同的流. 本文采用前一种方法.

** 所谓 $goodput$, 是指扣除重传的分组后剩余的吞吐量.

10Mb/s 瓶颈链路, 传播时延为 1ms. 为了考查 RTT 存在较大差异时的性能, S_1 到 D_1 和 S_2 到 D_2 的端到端链路传播时延(即途经的各个链路传播时延之和)分别为 10ms 和 50ms, 缓存容量为 120Packets. 我们

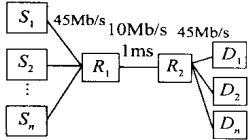


图 3 仿真网络拓扑图

对四种 TCP(Tahoe、Reno、Newreno、SACK)都作了仿真, 每次仿真运行 100 秒, goodput 见表 1, 链路利用率(总吞吐量与瓶颈带宽之比)见表 2, 分组平均等待时间(折算为具有平均分组长度的 Packets 数)见表 3.

公平性可以直接从表 1 中两个 goodput 的接近程度看出. 由于 HRFED 以分组到达速率而不是以缓存占用量来估计最高速流, 提高了对短期突发业务的宽容性, 它可以获得比 LQD 更高的公平性和较高的链路利用率. 通常, 在采用 FQ 调度时, 公平性比 FCFS 更理想, 但我们也发现, 由于 RED 原本不是为按流排队设计的, 所以会出现 FQ RED 的公平性比 FCFS RED 还差的情况.

表 1 两个 TCP 流各自的 goodput(Mb/s)

| | TD | | RED | | LQD | | HRFED | |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | FCFS | FQ | FCFS | FQ | FCFS | FQ | FCFS | FQ |
| Tahoe | 8.96/0.71 | 7.75/2.02 | 7.05/2.66 | 7.16/2.15 | 6.31/3.39 | 5.69/4.10 | 5.94/3.78 | 5.65/4.19 |
| Reno | 9.15/0.38 | 8.07/1.62 | 9.69/0.15 | 9.69/0.18 | 7.99/1.46 | 6.17/3.52 | 7.82/1.75 | 6.13/3.56 |
| Newreno | 9.15/0.74 | 7.98/1.97 | 7.10/2.86 | 7.44/2.42 | 6.43/3.29 | 5.73/4.06 | 6.46/3.50 | 5.69/4.30 |
| SACK | 9.17/0.79 | 7.28/2.68 | 7.32/2.67 | 6.67/3.23 | 6.42/3.54 | 5.92/4.06 | 5.76/4.23 | 5.15/4.83 |

表 3 两个 TCP 流各自的分组平均排队等待时间(Packets)

| | TD | | RED | | LQD | | HRFED | |
|---------|-----------|----------|-----------|----------|-----------|-----------|-----------|-----------|
| | FCFS | FQ | FCFS | FQ | FCFS | FQ | FCFS | FQ |
| Tahoe | 40.5/41.2 | 69.3/2.7 | 31.0/31.6 | 46.3/2.9 | 40.2/39.0 | 66.1/14.9 | 26.8/27.7 | 38.0/7.9 |
| Reno | 35.6/39.2 | 44.7/2.2 | 24.4/24.1 | 34.0/2.0 | 35.0/35.1 | 50.7/24.7 | 21.1/22.1 | 36.6/22.9 |
| Newreno | 42.2/42.0 | 38.3/5.1 | 27.1/27.3 | 28.9/8.8 | 40.7/40.1 | 32.5/10.7 | 21.0/19.1 | 36.9/9.2 |
| SACK | 40.9/41.4 | 59.5/9.5 | 32.5/33.0 | 45.4/2.4 | 39.2/37.8 | 55.5/34.4 | 30.3/30.4 | 43.8/24.7 |

4.2 UDP 与 TCP 竞争带宽

11 对节点 S_i 和 D_i ($i = 1, 2, \dots, 11$) 每对支持一个流, 从 1 到 10 号的流为 TCP, 第 11 号流为 UDP, 这些流经过的链路具有不同的端到端传播时延, 范围是 5 ~ 55ms. 10 个 TCP 流中, Tahoe 和 Reno 各 2 个, Newreno 和 SACK 各 3 个. TCP 传输 FTP 业务, UDP 传

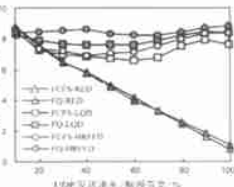


图 4 UDP 发送速率对 TCP goodput 的影响

输 CBR 业务, 缓存容量为 120Packets, 仿真运行 100 秒. 每次仿真的结果在图 4 中表示为一个点, x 轴代表 UDP 的发送速率与链路带宽之比, y 轴代表总的 TCP goodput. 从图 4 可见, FCFS RED 和 FQ-RED 所能提供给 TCP 的总 goodput 随着 UDP 发送速率的增加而呈近似线形的降低, 这是由于 UDP 属于非响应流, 即使因拥塞发生分组丢失, 也不会降低发送速率, TCP 在分组丢失后会降低发送速率, 而 RED 未对 TCP 流施加任何隔离或保护, 从而使带宽被 UDP 抢占. LQD 和 HRFED 对 TCP 都具有一定的隔离和保护作用. 另外, 在采用同样的调度

表 2 两个 TCP 流时的链路利用率(%)

| | TD | | RED | | LQD | | HRFED | |
|---------|------|------|------|------|------|------|-------|------|
| | FCFS | FQ | FCFS | FQ | FCFS | FQ | FCFS | FQ |
| Tahoe | 96.8 | 97.9 | 97.2 | 93.2 | 97.1 | 98.0 | 97.4 | 98.5 |
| Reno | 95.5 | 97.0 | 98.5 | 98.7 | 94.7 | 97.0 | 95.9 | 97.1 |
| Newreno | 99.0 | 99.6 | 99.6 | 98.6 | 97.3 | 98.1 | 99.7 | 99.9 |
| SACK | 99.6 | 99.7 | 99.9 | 99.0 | 99.6 | 99.8 | 99.9 | 99.8 |

在表 3 中, 仅就某种缓存管理策略而言, 如果采用 FCFS, 两个流在逻辑上共享一个队列, 所以分组平均排队等待时间相近, 如果采用 FQ, RTT 较大的 TCP 流速率较低, 它所对应的队列较短, 其分组的排队等待时间也较短. 通过比较, 可以发现 TD 和 LQD 的平均等待时间明显大于 RED 和 HRFED.

因为 TCP 主要支持对时延不敏感的数据业务, 所以将公平性和链路利用率作为衡量其性能的首要指标, 而将分组平均等待时间作为次要衡量指标. 因而, 在仅有两个 TCP 流时, 这四种缓存管理策略中, HRFED 对 TCP 是最优的.

算法时, HRFED 所能提供给 TCP 的 goodput 优于 LQD.

4.3 双向 TCP 的性能

当 ACK(确认) 分组和数据分组混合在同一个队列中, 或在数据分组中捎带 ACK 时, 会使 TCP 呈现出更大的突发性^[7]. 我们利用 NS 仿真器的 FullTCP 模块^[6] 仿真双向 TCP, 这比文献[5]中使用两个单向 TCP 连接来仿真一个双向 TCP 连接更准确, 此时, 图 3 中的 S_i 和 D_i 同时充当发送方和接收方. 每种 TCP 各占连接总数的 25%. 端到端链路传播时延分为 3、20、40、60ms 四类, 每类包括 25% 的流, 所有 TCP 都传输

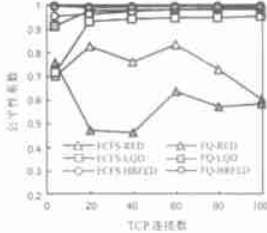


图 5 双向 TCP 连接数与公平性的关系

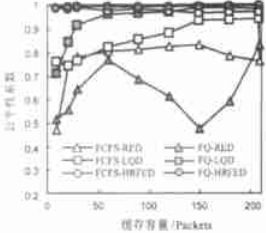


图 6 双向 TCP 缓存容量与公平性的关系

FTP, 每次仿真运行 200 秒. 当 TCP 连接总数从 4 到 100 变化时 (缓存容量为 150Packets), 公平性系数见图 5. 当缓存容量从 9 到 210Packets 变化时 (TCP 连接数为 20), 公平性系数见图 6. 可见, 在大多数情况下, HRFED 和 LQD 的公平性优于 RED, 而 FQ-HRFED 的优于 FQ-LQD, FCFS-HRFED 优于 FCFS-LQD, 另外, FQ-RED 常不如 FCFS-RED.

4.4 ON-OFF Pareto 源和 FTP 源共存的性能

研究表明, 以太网和广域网的业务都具有自相似性^[8,9], 而许多服从 $0.5 < \alpha < 1$ 的 Pareto 分布的 ON-OFF 业务可叠加为 Hurst 参数 $H = (3 - \alpha)/2$ 的自相似业务 ($0.5 < H < 1$)^[8]. 为了仿真 TCP 传输 Pareto 业务的性能, 我们用 $\alpha = 1.2$, ON 期平均发送 15 个长度为 80Bytes 的短分组、OFF 期平均持续时间为 1 秒的 Pareto ON-OFF 业务模拟由键盘产生的 TELNET 业务, 称第 1 组 Pareto ON-OFF; 用 $\alpha = 1.4$, ON 期平均发送 100 个长度为 1000Bytes 的长分组、OFF 期平均持续时间为 3 秒的 Pareto ON-OFF 业务模拟 WWW 业务, 称第 2 组 Pareto ON-OFF. 两组 Pareto ON-OFF 源与 FTP 源同时存在, 流的比例为: FTP: 第 1 组 Pareto ON-OFF: 第 2 组 Pareto ON-OFF = 1: 1: 2. 当保持 44 个 TCP 连接时, 缓存容量和公平性系数的关系见图 7, 该图表明 HRFED 的公平性优于 LQD. 如果与前面的仿真比较, 可以发现, 在多个 ON-OFF Pareto 源加入后, 无论对哪种缓存管理策略, 公平性会较只有 FTP 源时有所下降.

另外, 我们还对存在多条瓶颈链路的网络作了仿真, 发现随着缓存容量的增加, HRFED 和 LQD 的公平性都有所提高, 但 HRFED 仍然优于 LQD 和 RED, 当缓存容量超过 90Packets 后, FQ-LQD 较接近 FCFS-HRFED 和 FQ-HRFED. 随着连接数目的增加, 公平性呈下降趋势, 而 HRFED 的公平性仍然是这几种策略中最好的.

5 结语

作为一种缓存管理策略, HRFED 可用于路由器或交换机等分组转发设备中. 它能在保证较高链路利用率的同时, 提高 TCP 的公平性, 并维持较短的分组平均等待时间.

HRFED 的 EWMA 加权系数 w 可用于控制突发业务成为最高速流的容限. 当存在高速 UDP 流时, HRFED 对 TCP 流具有一定的保护作用. HRFED 不仅对传输 FTP 业务的单向 TCP 有效, 而且对传输 ON-OFF Pareto 业务的 TCP 和具有更大突

性的双向 TCP, 也可以提供比采用同样调度算法的 LQD 更好的公平性; 当存在多条瓶颈链路时, HRFED 的公平性仍然优于 LQD 和 RED.

另外, 在研究过程中还发现, FQ-RED 并不一定比 FCFS-RED 的公平性更好.

在具体实现时, 为了减少算法的复杂度, HRFED 所检测的流, 可以是一种广义的流, 即包含多个不同流的集合; 另外, 若 EWMA 加权系数 w 取 2^{-k} , 则乘法可简化为移位, 便于硬件实现.

参考文献:

- [1] RFC2309. Recommendations on Queue Management and Congestion Avoidance in the Internet [S].
- [2] S Floyd, V Jacobson. Random early detection gateways for congestion avoidance [J]. IEEE/ACM Trans. Networking, 1993, 1(4): 397-413.
- [3] S Floyd, K Fall. Promoting the use of end-to-end congestion control in the Internet [J]. IEEE/ACM Trans. Networking, 1999, 7(4): 458-472.
- [4] J Padhye, et al. Modeling TCP throughput: a simple model and its empirical validation [A]. Proc. of SIGCOMM 98 [C]. Canada: SIGCOMM, 1998.
- [5] B Suter, T V Lakshman, et al. Buffer management schemes for supporting TCP in gigabit routers with per flow queueing [J]. IEEE J SAC, 1999, 17(6): 1159-1169.
- [6] (network simulator) [CP]. <http://www.isi.edu/nsnam/ns/>
- [7] L Zhang, S Shenker, D Clark. Observations on the dynamics of a congestion control algorithm: the effects of two way traffic [A]. Proc. of SIGCOMM 91 [C]. Switzerland: SIGCOMM, 1991. 133-147.
- [8] W Willinger, M Taqqu, R Sheman, et al. Self similarity through high variability: statistical analysis of Ethernet LAN traffic at the source level [J]. IEEE/ACM Trans. Networking, 1997, 5(1): 71-86.
- [9] V Paxson, S Floyd. Wide area traffic: the failure of Poisson modeling [J]. IEEE/ACM Trans. Networking, 1995, 3(3): 226-244.
- [10] S Floyd. Connections with multiple congested gateways in packet switched networks, part 1: oneway traffic [J]. Computer Communications Review, 1991, 21(5): 30-47.

作者简介:



温蜀山 男, 1974 年 4 月生于云南省昆明市, 1996 年获电子科技大学学士学位, 现为电子科技大学硕博连读研究生, 研究兴趣为 TCP/IP 技术, 拥塞控制、缓存管理及其性能分析.