

# 基于多优先级探测的准入控制实现 端到端服务质量保证的研究

吕云飞, 王行刚, 孙 巍

(中科院计算技术研究所, 北京 2704 信箱 9 分箱, 北京 100080)

摘 要: 本文针对传统的基于单优先级探测的准入控制存在的缺陷, 提出了一种可扩展性好的基于多优先级探测的准入控制方法. 它可以根据流的不同服务质量需求, 为单个流提供多种不同优先级别的端到端的服务质量保证. 本文详细论述了它的原理与实现. 为了降低探测流冲突对准入控制的性能的影响, 本文提出了在路由器中设定可变的竞争探测流数目上限的方法来降低探测流冲突, 既提高了资源利用率又保证了较高的准入成功率. 最后, 本文通过 NS 模拟验证了多优先级准入控制的优越性.

关键词: 多优先级探测; 服务质量; 准入控制; 探测流冲突

中图分类号: TP393. 02 文献标识码: A 文章编号: 03722112 (2003) 12164305

## Research on Realizing End-to-End QoS by the Admission Control Based on Multiple-Prior Probing

LV Yunfei, WANG Xinggang, SUN Wei

(The Institute of Computing and Technology of Chinese Academy of Science, P. O. Box 270429, Beijing 100080, China)

Abstract: A scalable admission control method based on multiple-prior probing, aiming at the limitation of the traditional one-prior probing admission control, is proposed to provide multiple-prior end-to-end QoS for individual flows according to their different QoS requirements and principle. The implementation of this method is explained in detail. To diminish the influence of probing collision on the performance of the proposed admission control, a method of setting the alterable maximal number of the competing probing flows in the router is put forward so as to increase the resource utilization and ensure a high access success ratio. At last, the simulation in NS validates the advantage of the proposed admission control method.

Key words: multiple-prior probing; quality of service; admission control; probing collision

### 1 引言

现在的网络用户对视频会议、远程教学、IP 电话等多媒体 Internet 应用的需求越来越大, 而这些应用需要一定服务质量要求, 譬如带宽、时延、抖动等. 但是现在的 Internet 是基于最大努力传送的原则来提供网络服务的, 根本不提供任何 QoS 质量保证. 针对此, 现在出现了很多用于提供 QoS 保证的技术, 例如 DiffServ<sup>[1]</sup>, IntServ<sup>[2]</sup> 等.

IntServ<sup>[2]</sup> 虽然可以对单个流提供很好的服务质量保证, 但是它需要大量的 RSVP<sup>[3]</sup> 信令传送网络资源消息. 网络中的路由器需要处理每个流的资源预留请求, 根据当前的可用资源对流进行准入控制, 而且要保存每个流的状态, 给路由器带来很大的负担, 从而使其可扩展性很差, 无法应用于大型网络. DiffServ 不需要准入控制和信令机制, 路由器根据数据包头中的 DS 域区分数据流的级别, 从而进行不同的处理, 不需

要保存每条流的信息, 工作简单, 可扩展性强, 但它不能针对单个流提供服务质量保证<sup>[1]</sup>. 为了能融合 DiffServ 的可扩展性的优点和 IntServ 很好的 QoS 保证的优点, 文章[4~7]提出了基于端到端探测的准入控制方法, 通过发送与申请流具有相同属性的探测流到接收端, 然后由接收端根据接收到的流的速率、丢包率、延迟等判断网络是否有足够的资源, 然后发送反馈信号通知发送端是否可以发送流. 端到端准入控制可以象 IntServ 一样提供单条流的 QoS 保障, 但与 IntServ 不同的是, 它不需要中间路由器的支持, 路由器不需要保存流的状态信息, 路由器只需要根据数据包头中的 DS 域(类似于 DiffServ) 处理数据包, 因此这种方式具有很好的可扩展性. 该方法主要适用于 soft real-time 业务, 特别适用于那些具有突发特性的 VBR 流业务, 在保证一定的服务质量的情况下, 可以获得很高的网络利用率.

虽然端到端探测法准入控制可以为每个流提供较好的服

务质量保证,可扩展性好,但它目前还存在一些问题,还无法很好的工作.

首先,目前所提出的端到端探测法准入控制对已连接流是按照 FIFO 的方式服务,即只有一种服务级别,所有已连接流获得相同的服务质量,这对于现在出现的有不同的服务质量要求的网络应用是不合适的.而用一种级别的探测流无法提供不同服务级别的准入控制,因此应有不同的探测流级别.

其次,关于探测流方式的问题.如果采用 in2band probing<sup>[4]</sup>的方式(即探测流的优先级与已连接数据流的优先级相同),则在发送探测流的过程中,可能会损害到已连接流的服务质量,而且会引起窃带宽(bandwidth stolen)的问题<sup>[4]</sup>,即高优先级探测流强占低优先级的带宽,从而使高优先级流不断准入造成低优先级流/饿死0;如果采用 ou2o2band probing<sup>[4]</sup>的方式(即探测流的优先级低于已连接数据流的优先级),虽然不会影响已连接流,但探测流必须与其他低等级的数据流竞争资源,造成得到的网络可用资源低于实际值,使网络的吞吐率与准入成功率偏低.

再次,准入控制存在探测流冲突的问题.在某一时刻,可能会有很多待准入的发送流发送探测流使探测流的总量大于网络的剩余带宽,从而即使网络的剩余带宽能满足其中任何一个流的要求,所有的流都无法准入.这种冲突随着网络规模的增大会越来越严重,导致吞吐量和带宽利用率及准入请求成功率的急剧下降.

本文将针对上述问题提出一种基于多优先级探测的端到端准入控制,它具有多个服务优先级,每个服务优先级有各自独立的数据流队列和探测流队列,采用 ou2o2band 的探测机制,保证已连接流的服务质量不受探测流的影响,并且每个优先级都分配一定的带宽,而 BE 流则利用网络的剩余带宽传送,从而既保证各级流之间互不影响(即不会出现带宽窃流现象),又保证较高的带宽利用率.

## 2 基于多优先级探测法的准入控制的原理及实现

本文所提出的这种多优先级探测法准入控制,是在 DiffServ 的基础上修改完成的.它将 DSCP 值的某一部分定义为探测信息.当路由器检查 IP 报头时,发觉其 DSCP 为本探测法所定义的某一值,便将其作为探测流包处理.

网络中的路由器分为核心路由器和边缘路由器两种.核心路由器的主要功能同 DiffServ 网的核心路由器功能类似,主要根据包头上的标记值(DSCP)区分包的优先级和类型(数据包或探测包),然后将包放入相应的队列,再者进行队列调度来完成数据包的转发.为了降低探测流的冲突,核心路由器还有一个探测包处理器,主要用于限制竞争探测流的数目,其具体功能将在第 3 节做详细介绍.边缘路由器除了要转发从用户来的数据包外,还要负责与用户进行服务协商,设定服务优先级,完成准入控制,以及对数据包进行打标记转发等.其中准入控制功能主要就是根据服务协商利用适当等级的探测流进行探测,并根据探测结果判断网络是否能满足用户数据流的服务要求,并发送信息给用户终端.同样,边缘路由器也要保存一个探测包处理器用于控制探测流的冲突

### 2.1 端到端多优先级探测准入控制过程

准入控制过程如下:

step 1 边缘路由器收到用户终端发送流请求,根据请求选择适当的服务优先级.

step 2 边缘路由器启动探测流进程,设定超时时器,在规定的时间内连续不断的按发送流特性产生探测流包,并根据优先级在报头上作标记,将发送流要求放入探测包内,发向目的端.

step 3 目的端边缘路由器在收到第一个探测包后,启动探测进程,在探测时间收集探测包信息,并在探测时间过后进行统计计算,将结果与规定的服务质量要求相比较,如不符合,则发送/失败0反馈信号,并终止探测进程,重新进入等待状态;如果符合,则发送/成功0的反馈信号给发送端边缘路由器,并进入连接成功状态,开始接收数据包,并转发给用户端.

step 4 发送端路由器若在超时之前收到反馈信息,则根据反馈信息通知用户是否准许准入.如果超时,则直接发送失败信息给用户端,并停止发送探测流;如果收到成功信息,则继续发送探测流,并给用户端发送成功信号,当用户端发送来数据包后,停止发送此探测流并转发数据包.

上述准入控制算法的主要特点是:(1)所有实现都在边缘路由器完成,对用户端透明,并且不需要核心路由器的支持,不会给其带来额外负担;(2)可以根据用户的要求提供多个优先级的服务.

### 2.2 数据包在路由器中的处理

数据包在路由器中的处理过程如图 1 所示,当一个数据包进入路由器首先经过包分类器模块,根据数据包头部的 DSCP 值判断包的优先级与类型(探测包或数据包),如果是数据包,则放入相应的数据流队列;如果是探测包,则交由探测包处理器来进行处理后送入相应的探测流队列.路由器根据 WRR(weighted round robin)队列调度算法从队列中取出数据包发送出去.队列调度算法将在 2.3 节中详述.探测包处理器主要用于控制探测流的冲突,其具体功能将在第 3 节中介绍.

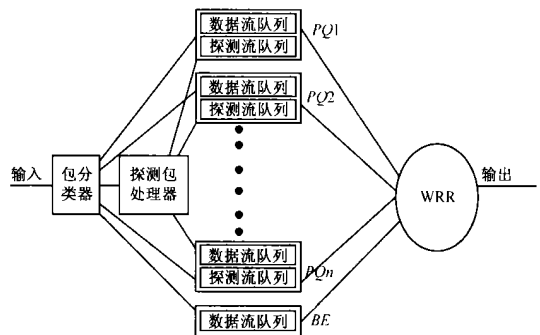


图 1 数据包的处理示意图

### 2.3 队列调度算法

每个路由器内有 n 个优先级的队列 {PQ1, PQ2, ..., PQn} 和一个 BE(best effort)队列.除 BE 队列外,每个优先级内部有两个队列:一个是数据流队列,一个是探测流队列.节点采用 WRR(weighted round robin)的队列调度机制,先从高优先级队列中取得数据包转发,然后再到低优先级队列中取数据

包转发. 为防止高优先级队列无限占用资源, 每个优先级都有固定的带宽比例, 每一轮循环, 每个优先级队列最多只能转发与其带宽比例对应数量的数据包. 在优先级内部, 数据流的优先级高于探测流的优先级, 当一个优先级队列获得发送权时, 它首先发送数据流队列中的包, 只有当此队列中没有数据流包, 且还有剩余带宽资源时, 此队列中的探测流包才可被转发. 这样做的优点是: 首先, 它保证了探测流不会影响已连接的数据流的服务质量, 因为它的优先级低于数据流; 其次, 各级探测流互不影响, 通过探测流可以正确地获得本优先级的剩余资源, 并作为准入控制的判断依据; 再次, 高优先级的探测流比低优先级探测流优先得到服务, 保证高优先级数据流可以获得优先服务和优先准入探测的服务.

具体队列调度算法如下:

```
for(;;) {
  if(PQ 1 还有足够剩余资源发送一个包) {
    if(数据流队列为空) {
      发送一个数据包
      总剩余资源减去一个发送包的大小
      PQ 1 的剩余资源减去一个发送包的大小
    }
    else if(探测流队列为空) {
      发送一个探测包
      总剩余资源减去一个发送包的大小
      PQ 1 的剩余资源减去一个发送包的大小
    }
    continue;
  }
  if(PQ 2 还有足够剩余资源发送一个包)
  ,
  if(PQ n 还有足够剩余资源发送一个包)
  ,
  if(总剩余资源足够发送一个包) {
    发送 BE 包
    总剩余资源减去一个发送包的大小
  }
  else
    分配下一轮资源并进行下一轮循环
}
```

网络中的剩余带宽由 BE 流占用, 而分配给优先级队列而未用的那部分带宽也可以被 BE 使用, 这样可以有效地利用网络带宽, 提高网络带宽利用率. 由于探测流的优先级高于 BE 流, 从而可以保证即使在 BE 流占用了所有的剩余带宽时, 即网络没有剩余资源但探测流对应优先级的数据流还没有达到其所可以获得的最大带宽资源时, 探测流就可以从 BE 流中抢回被占用的资源, 保证优先级数据流用户的准入.

### 3 对探测流冲突的解决方案

探测流冲突问题对基于探测法的准入控制的性能影响很大, 冲突严重的情况下, 即使网络有很大的剩余资源, 也无法给用户准入服务, 大大降低了网络的资源利用率和准入成功率. 探测流冲突发生的根本原因就是在同一时刻网络上

所有的探测流总量大于网络的剩余资源. 减少探测流冲突最有效的办法就是减少同时竞争的探测流的数目, 但单纯地减小竞争探测流的数目也不能完全保证准入控制的性能. 这是因为减小同时竞争的探测流数目虽然降低了探测流冲突, 但同时也限制了探测流的准入, 使网络的准入成功率降低, 特别是当网络负荷较低, 可以容纳很多探测流同时探测, 而且准入请求到达频繁时, 由于探测流数目的限制, 将会造成大量的准入请求被拒绝.

针对此, 本文提出了一种设置可变的竞争探测流数目上限的方法解决探测流冲突. 其基本思路是: 探测流数目上限与网络负荷有关. 当网络负荷较低时, 探测流数目上限较高, 可以允许数目较多的探测流同时工作; 当网络负荷较高时, 探测流数目适当降低, 减小小探测流的冲突, 提高网络利用率和准入成功率. 由于本文所提出的是多优先级探测法, 不同优先级的探测流是不会相互冲突的, 因此探测流冲突只限于同一级的探测流之间. 因此, 在路由器中每一个优先级队列对应一个探测流数目上限, 网络负荷在此是指针对特定优先级的资源利用率, 即已连接流占用的资源与分配给本优先级的总的资源的比例. 本文以如下公式计算竞争探测流数目上限:

$$n = (n_{\max} - n_{\min}) * (1 - K) + n_{\min} \quad (1)$$

其中,  $n_{\max}$ ,  $n_{\min}$  为路由器为每个优先级设定的最大和最小竞争探测流数目上限,  $K$  为特定优先级的网络负荷. 路由器定期地检测资源网络负荷, 并根据式(1)调整竞争探测流上限  $n$  的值.

探测流数目的控制是由探测包处理器来完成的, 在探测包处理器中有

流标识	定时器
-----	-----

图 2 探测流表项内容

多个探测流表, 每个探测流表对应一个优先级, 表项数为  $n_{\max}$ , 表项的格式如图 2 所示.

其中流标识能够唯一地标识一个连接, 它由 IP 包头中的各项组合而成. 定时器用来设置该流标识的在探测流表中的有效时间.

探测包处理器具体操作如下:

(1) 初始时探测流表为空;

(2) 当收到一个探测包时, 根据其优先级到相应优先级的探测流表中搜索:

如果有表项与之相同, 则将该探测包放入相应的探测队列;

如果没有, 则比较探测流表中的探测流数目  $n_p$  与探测流数目上限  $n$  的大小:

如果  $n_p < n$ , 则将该探测包的流标记加入到表中, 设置定时器, 并将探测包放入相应的探测队列中;

如果  $n_p \geq n$ , 则直接将该探测包丢弃, 并发送拒绝信息给该探测包的源节点;

(3) 当有探测流表项的定时器超时后, 释放定时器资源, 并将该表项删除.

定时器的时间设置为  $a$  倍的探测时间 ( $a > 1$ ), 保证定时器超时后, 探测过程已经结束. 对于有  $K$  个优先级的路由器,

它维护的探测流表项数最大为  $n_{max} * K$ , 与网络大小无关, 可扩展性好.

### 4 性能模拟与分析

为了验证本文所述的端到端多优先级探测准入控制的性能. 本文在 NS<sup>[8]</sup> 模拟平台上建立一个网络模型做模拟, 从已连接的数据流获得的服务质量、网络资源利用率及准入请求成功率三方面来评测本文所提出的准入控制方法的性能.

本文所用的网络模型的拓扑结构如图 3 所示:

其中 E1, E2, E3, E4 为边缘路由器, R1, R2 为核心路由器. R1, R2 之间链路带宽为 10Mb, 时延为 20ms, 边缘路由器与核心路由器之间的链路为带宽为 100Mb, 时延为 10ms.

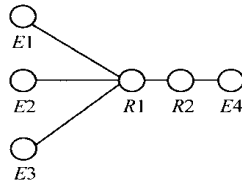


图 3 网络模型的拓扑结构

显而易见, R1, R2 之间链路为网络的瓶颈所在. 网络中的数据分为三个级别: 优先级 P1, 优先级 P2 及 BE 数据流. 链路对三种级别的数据流的带宽资源的分配比例为 P1B:P2B:BE = 3B:3B:4. 两个优先级分别有各自级别的探测流. 假设网络中的数据流都是从 E1, E2, E3 到 E4. BE 级的数据流不需要进行准入控制, 本文假定, 在从 E1, E2, E3 到 E4 之间各有 5 条 FTP 数据流作为 BE 级的数据流. P1 级和 P2 级的数据流需要由 E1, E2, E3 执行准入控制, 获得准许后才能加入的网络中. 本文假定为 P1 级和 P2 级的数据流提供的服务质量保证为: 对于优先级 P1, 流的丢包率小于 1%, 平均端到端时延小于 50ms, 对于优先级 P2, 流的丢包率小于 5%, 平均端到端时延小于 80ms. 探测时间为 1s, 超时时间为 11.5s.

由于探测法准入控制比较适合于有突发特性的 VBR 流, 因此在模拟中, 本文采用两种 ON/OFF 式的通信流, 一种是 Exponential<sup>[8]</sup>, 另一种 Pareto<sup>[8]</sup>. 其特性参数设置如表 1:

表 1 通信流特性参数

流类型	突发速率 (kbps)	On Time (ms)	Off Time (ms)	B
Exponential	Random (100, 200)	Random (50, 100)	Random (50, 100)	-
Pareto	Random (100, 200)	Random (50, 100)	Random (50, 100)	Random (1.1, 1.5)

注: Random(x, y) 表示在 x, y 之间的随机值

对于竞争探测流数目上限的设置, 在本文的模拟中  $n_{max}$  和  $n_{min}$  分别取值为 50 和 5.

#### 4.1 已连接流的服务质量

图 4, 5 是网络利用率与流的平均丢包率和平均时延的关系图. 图中 P1, P2 分别代表优先级 P1 和优先级 P2 的流的平均丢包率和平均时延.

从图 4, 5 中可以看出, 已连接流的服务质量虽然随着网络负荷的增大而有所降低, 但是始终满足规定的服务质量要求. 这是因为本文采用 cut-of-band 的探测方式, 探测流不会占用已连接流的资源, 故而若是探测流探测到有足够的可

用资源, 则准入后的流必然能获得满足要求的服务质量. 从图中还可以看出, P1 级流获得的服务质量要高于 P2 级流, 但它能获得的网络资源利用率也较低.

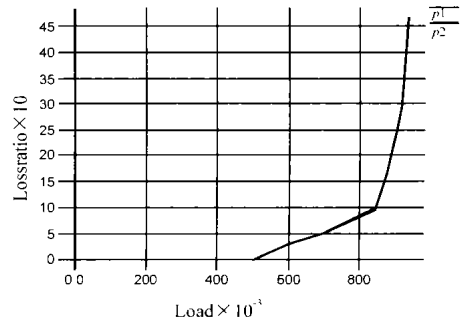


图 4 已连接流的丢包率

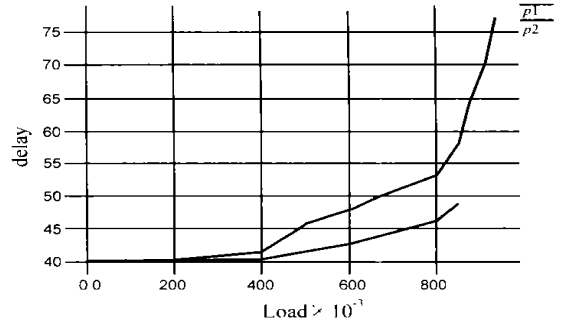


图 5 已连接流的平均时延

#### 4.2 网络资源利用率

对于本文所提出的准入控制, 探测流冲突仅会发生在同一级的探测流之间, 各优先级之间相互独立, 因此本文仅以 P1 级流为例讨论其网络资源利用率. 本文假设准入请求按泊松分布到达, 流的特性如表 1 所述设置. 为了验证本文所提出的可变式竞争探测流数目上限的方法可以很好地降低探测流冲突, 获得较高的网络资源利用率, 本文将本准入控制方法与文[4, 7]所提出的传统的探测法准入控制作比较, 在相同的服务质量要求下, 模拟在不同的准入请求到达率的情况下, 各种准入控制所能获得的最大网络资源利用率. 本文还模拟了最优情况下, 即探测流数目上限固定为 1 时, 网络能获得的最大资源利用率. 在这种情况下探测流在任何时刻始终只有一个, 不会存在探测流冲突, 只要网络剩余资源能满足该探测流的需求, 就可以准入, 因此可以获得最大的网络资源利用率. 模拟的比较结果如图 6 所示.

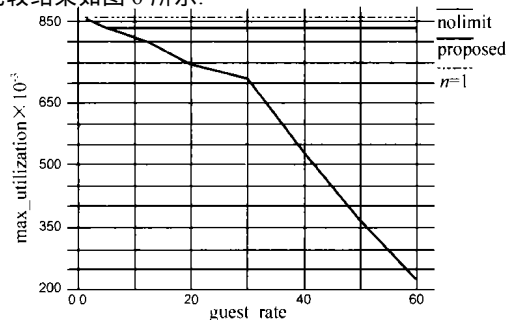


图 6 竞争探测流数目对资源利用率的影响

图 6 中横坐标  $quest. rate$  代表准入请求到达的平均速率,纵坐标为最大资源利用率; $nolimit$  曲线代表没有竞争探测流数目限制下的网络资源利用率与准入请求速度的关系; $n=1$  曲线代表竞争探测流数目上限固定为 1 的情况; $proposed$  曲线代表  $n$  可变的情况。从图 6 中,可以看出当准入请求到达速率较低时,探测流的冲突较小,各种准入控制获得的最大网络利用率相近;但是当准入请求到达速率变高时,探测流的冲突增高,传统的探测流准入控制方法的资源利用率迅速下降,而本文所提出的准入控制方法则通过适当的限制探测流数目依然可以保证获得较高的资源利用率,并且与最优情况相近。

#### 4.3 准入请求成功率

对于准入请求成功率,同前所述,各优先级之间是相互独立的,本文亦只针对某一个优先级讨论其准入请求成功率,在此,本文针对 P1 优先级。在本模拟中,本文主要研究在一定的准入请求速率的情况下,本文所提出的可变竞争探测流上限,和固定竞争探测流上限以及传统的无探测流数目控制的各种准入控制方法在不同网络负荷下的准入成功率。图 7 是当准入请求速率为 20 个每秒,  $n=1$  ( $n=1$  曲线),  $n$  可变 ( $proposed$  曲线) 及  $n$  无限制 ( $nolimit$  曲线) 时网络负荷与准入成功率的关系。可以看出  $n=1$  时,即使在网络负荷较低时,网络的准入成功率也很低,不到 10%。这是因为网络中只允许一个探测流存在,从而使大多数准入请求都被拒绝,造成准入率过低。当  $n$  无限制时,网络负荷低时,准入成功率很高,而随着网络负荷的增大,探测流冲突增大,准入成功率迅速降低。而本文所提出的可变探测流数目上限方法在网络负荷较低时,  $n$  较大;而在网络负荷变高时降低  $n$  值,减小探测流冲突,从而使准入成功率始终保持一个较高的值。

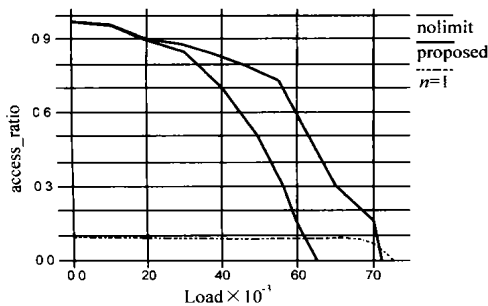


图 7 竞争探测流数目对准入成功率的影响

## 5 结论

本文针对传统的探测法准入控制存在的一些问题提出了一种基于多优先级探测法的准入控制,它可以为单个流提供

不同优先级的端到端服务质量保证。针对探测冲突问题,本文还提出了设定可变的竞争探测流数目上限的方法来降低探测冲突率,提高资源利用率和保证一定的准入成功率。最后,本文通过 NS 的模拟验证了以上方法的正确性和优化性。

#### 参考文献:

- [1] S Blake, D Black, M Carlson, E Davies, Z Wang, W Weiss. An Architecture for Differentiated Services[S]. Internet RFC 2475 December 1998.
- [2] Braden R, Clark D, Shenker S. Integrated services in the Internet Architecture: an overview[S]. Internet RFC 1633, June 1994.
- [3] Braden R, Zhang L, Berson S, Herzog S, Jamin S. Resource reservation protocol(RSVP) Version 1 Functional Specification[S]. RFC 2205, Proposed Standard, September 1997.
- [4] Breslau L, Knightly E, Shenker S, Stoica I, Zhang H. Endpoint admission control: Architectural issues and performance[A]. Proc. of SIGCOMM 2000[C]. Stockholm Swden, Association for Computer Machinery, 2000. 57- 69.
- [5] V Elek, G Karlsson, R Ronngren. Admission control based on end-to-end measurements[A]. Proceedings of IEEE INFOCOM 2000[C]. Tel Aviv, Israel: Institute of Electrical and Electronics Engineers, Mar, 2000. 623- 630.
- [6] S Floyd, K Fall. Promoting the use of end-to-end congestion control in the internet[J]. IEEE/ACM Transactions on Networking, 1999, 7(4): 458- 472.
- [7] Giuseppe Bianchi, Antonio Capone, Chiara Petrioli. Throughput analysis of end-to-end measurement based admission control in IP[A]. Proc of INFOCOM 2000[C]. Tel Aviv, Israel: Institute of Electrical and Electronics Engineers, March, 2000. 1461- 1470.
- [8] Network Simulator ns[CP]. <http://www.isi.edu/nsnam/ns>.

#### 作者简介:



吕云飞 男, 1977 年生于山东淄博, 博士研究生, 研究方向是信息网络系统设计, 目前感兴趣的方向是网络服务质量, QoS 路由和流媒体的传输技术。

王行刚 男, 1937 年生于四川自贡, 研究员, 博士生导师, 研究方向是信息网络系统设计。