

# 一种新的逐跳 TCP 友好的主动分层多播拥塞控制方案

叶晓国<sup>1,2</sup>, 吴家皋<sup>1,2</sup>, 姜爱全<sup>1,2</sup>

(1. 东南大学计算机科学与工程系, 江苏南京 210096; 2. 东南大学教育部计算机网络和信息集成重点实验室, 江苏南京 210096)

**摘 要:** 基于 Internet 的多媒体多播应用的迅猛发展对多播拥塞控制提出了要求. 分层多播是适应网络异构性较有效的方案. 针对现有分层多播存在的问题, 将主动网技术思想引入到分层多播拥塞控制中, 提出了一种逐跳 TCP 友好的主动分层多播拥塞控制方案 (HILMA), 采用主动标记分层、逐跳探测 TCP 友好可用带宽, 以及主动速率控制机制. 仿真实验表明, HILMA 方案大大改进了分层多播拥塞控制的性能, 具有较快的拥塞响应速度、较好的稳定性和 TCP 友好特性.

**关键词:** 主动网; 分层多播; 拥塞控制; 主动分层多播

**中图分类号:** TP393 **文献标识码:** A **文章编号:** 0372-2112 (2005) 08-1432-06

## A Novel Hop-by-Hop TCP-Friendly Congestion Control Scheme for Layered Multicast Based on Active Network

YE Xiao-guo<sup>1,2</sup>, WU Jia-gao<sup>1,2</sup>, JIANG Ai-quan<sup>1,2</sup>

(1. Department of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 210096, China;

2. Key Laboratory of Computer Networks and Information Integration of Ministry of Education, Southeast University, Nanjing, Jiangsu 210096, China)

**Abstract:** Effective multicast congestion control mechanism is urgently needed with the deployment of multimedia multicast applications in Internet. Layered multicast is considered to be an efficient approach to cope with the network heterogeneity. Receiver-driven layered multicast schemes have many disadvantages, such as large congestion response delay, unstable throughput and unfriendly bandwidth sharing with TCP. A novel congestion control scheme for layered multicast, called Hop-by-Hop TCP-friendly Congestion Control for Layered Multicast based on Active Network scheme (HILMA), is proposed through introducing active label layering, hop-by-hop detecting TCP-friendly available bandwidth and active rate control mechanisms. Simulation results show that HILMA is more responsive, more stable and TCP-friendly, and can improve the performance of layered multicast congestion control effectively.

**Key words:** active network; layered multicast; congestion control; active layered multicast

## 1 引言

视频会议、网络教育、视频点播等多媒体多播应用在 Internet 上正得到越来越广泛的应用, 但是由于目前的 IP 多播基于“尽力服务”的服务模型, 缺乏有效的拥塞控制机制, 阻碍了多媒体多播应用在 Internet 上广泛部署. 多播通信中接收者数目多, 涉及的范围广, 因此多播拥塞控制对于整个网络的稳定运行具有重要的意义. 多播拥塞控制问题由于其复杂性和重要性, 一直是网络研究中的一个热点问题.

Internet 具有动态性和异构性两大特征, 多播拥塞控制必须能够适应这种异构性和动态性. 目前已经提出了多种多播拥塞控制方案, 包括单速率多播和多速率多播两种方式. 多速率多播通常又称为分层多播, 是分层编码和分层传输两种技术的结合. 分层多播能够更好地满足网络异构性的要求.

分层多播中, 发送者将数据源用分层编码的方式编码成多个数据流, 分为基层和多个增强层, 将不同层的数据流放在多个多播组中传输. 接收者根据自己的接收能力选择接收从基层到增强层的不同层数的数据, 就能接收到不同速率的数据流. 例如接收方驱动的分层多播 RLM<sup>[1]</sup>, RLC<sup>[2]</sup>等. 但是现有的分层多播单纯依靠端系统响应拥塞, 难以对拥塞做出快速有效地响应<sup>[3]</sup>.

现有的分层多播存在以下不足: (1) 需要分配和维护多个多播组, 开销较大, 频繁的组成员加入和退出容易导致系统的不稳定; (2) 对网络拥塞的响应是通过加入/退出一个多播组实现的, 延迟较大, 不适合网络动态性较强的情况; (3) 速率调节以层为单位, 粒度较粗, 速率调节不够平滑, 不适合多媒体应用的需求; (4) 网络拥塞将导致基层和增强层数据不加区分地丢弃, 引起接收者实际接收质量的剧烈抖动; (5) 接收者之

收稿日期: 2004-05-09; 修回日期: 2005-04-11

基金项目: 国家自然科学基金 (No. 90104009)

间的协调机制影响了多播应用的可缩放性和 TCP 友好特性。

现有分层多播存在的问题归根结底是由于目前的“尽力服务”的 Internet 服务模型缺乏有效的资源控制和灵活的服务定制功能。主动网是一种新型的具有高度灵活性和可扩展性的网络体系结构,主动网中的中间节点具有很强的处理能力和可编程能力,为用户提供了灵活的服务定制能力<sup>[4,5]</sup>。主动网为构建灵活性的、适应性的多播服务提供了全新的方案。文献[6]提出了主动网技术增强的单播拥塞控制协议 ACC。文献[7]研究基于主动网的多播应用问题,提出了一个主动网支持的可靠多播协议 RMANP。文献[8,9]研究主动网对网络异构性的支持机制,提出了基于主动节点过滤机制的异构视频多播流量控制方案。文献[10]提出了一个适应性的主动分层多播协议 ALMA,研究了主动节点中链路带宽分配问题。

本文针对传统分层多播存在的问题,引入主动网技术思想,提出了一种逐跳 TCP 友好的主动分层多播拥塞控制方案 HILMA。HILMA 方案结合了基于接收方驱动和网络驱动两种方案的优点,其基本思想是:(1)采用主动标记分层与优先级分层过滤机制,以避免传统的分层多播中维护多个多播组的开销,降低分层多播速率调节操作的延时,并增强分层多播吞吐率的稳定性。(2)以逐跳链路的方式探测 TCP 友好可用带宽,并利用主动节点进行逐跳速率控制,使网络可用带宽探测局部化、能快速适应网络状况的动态变化。仿真实验表明,与传统的分层多播方案 RLM 相比,我们提出的 HILMA 方案能更好地适应网络状况的异构性和动态性,达到较快的拥塞响应速度,也有较好的稳定性和 TCP 友好特性,因此能更好地满足多媒体多播应用的需求。

## 2 主动分层多播服务模型

主动分层多播构建在主动网服务模型之上。该模型假设在多播分发树的策略位置上部署了主动路由器,对多播分组进行主动转发处理。与 IP 多播服务模型不同的是,主动分组是主动分层多播传输和处理的基本单元,多播数据被封装在主动分组中传输。主动分组分层信息和拥塞控制的响应方法作为主动分组的代码部分,被封装在主动分组之中;在主动分组到达主动路由器时被执行,从而执行用户定制的拥塞控制动作。主动分层多播模型如图 1 所示。

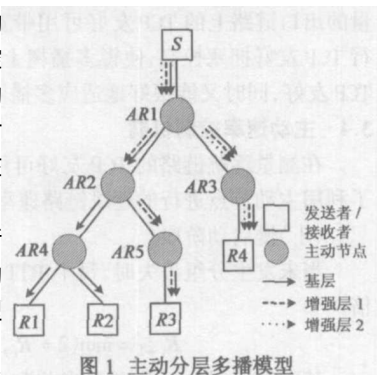


图 1 主动分层多播模型

S 为主动分层多播

发送者,R1~R4 为接收者,AR1AR5 为支持主动分层多播的主动节点,其上实现了本文提出的主动分层多播拥塞控制方案 HILMA。

## 3 主动分层多播方案(HILMA)

基于主动分层多播服务模型,我们提出了一种逐跳 TCP

友好的主动分层多播方案 HILMA,其关键实现机制包括主动分组分装、主动标记分层与优先级分层传输、逐跳链路可用带宽测量以及主动速率控制机制。

### 3.1 主动分组分装格式

HILMA 方案基于主动网报文封装协议<sup>[11]</sup>(Active Network Encapsule Protocol, ANEP),其协议报文作为 ANEP 的负载。HILMA 协议报文有两种类型,即主动控制分组(Active Control Packet, ACP)和主动数据分组(Active Data Packet, ADP)。整个协议报文头部包括两个部分,即 ANEP 头部和 HILMA 头部。HILMA 主动分组格式定义如图 2 所示。

HILMA 头部包含版本 (Version)、标志 (Flags)、类型标识 (Type-ID)、上一节点地址标识 (Previous)、生存期 (TTL)、序号 (Sequence Number)、层标记 (Layer Number)、数据/代码域

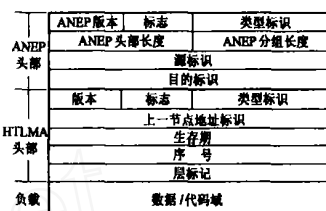


图 2 HILMA 主动分组格式

(Data/Code)等。当标志域为 0 时表示主动控制分组,此时,数据/代码域表示分层多播拥塞控制代码;当标志域为 1 时表示主动数据分组,数据/代码域表示用户数据负载。主动分标记域 Layer Number 用于标识数据分组属于哪个分层。通过标记分层策略,分层多播中基层和增强层的传输只需一个多播组,而不需要维护多个独立的多播组会话。

### 3.2 标记分层与优先级分层传输机制

在 HILMA 方案中,利用 3.1 节定义的主动分组中的层标记域 Layer Number,我们提出一种基于标记的分层机制,并将标记分层和优先级分层传输机制有效集成。

由分层编码机制可知,基层数据最重要,从基层到最高增强层,数据重要性逐层降低,相应地,传输优先级从低层到高层也依次降低。在网络发生拥塞时,优先过滤当前分层多播会话中具有最低优先级的最高层数据,从而保护高优先级低层重要数据。我们给出分层标记和分层传输优先级定义如下:

定义 1(主动分层标记与优先级) 分层多播会话 S 中第 i 层的主动分组  $AP_S^i$ ,记其分层标记为  $L(AP_S^i) = i$ ,优先级为  $Priority(AP_S^i) = N - 1 - L(AP_S^i)$ 。其中,  $i = 0, 1, \dots, N - 1$  为基层和各增强层的标记, N 为分层层数。

基于标记的分层机制将多个层放在同一个多播组中传输,具有以下优点:避免了分层多播中需要维护多个多播组的开销;避免了传统分层传输中的 IGMPP<sup>[12]</sup> leave 延时的影响,使拥塞能够得到较快缓解;易于与优先级分层传输相结合,保护较重要的低层分组。

### 3.3 逐跳链路可用带宽的测量

对链路可用带宽的检测是分层多播拥塞控制方案有效适应网络状况变化的关键。在 HILMA 方案中,我们采用逐跳链路可用带宽测量方法测量从发送者到接收者之间链路的 TCP 友好可用带宽。从发送者到接收者的链路上每对相邻的上、下游主动节点负责测量该节点对之间链路的 TCP 友好可用带宽,根据测量得到的每个出口链路的可用带宽,多播树上中

间节点可以独立地调节各出口链路上的发送速率. 与传统基于接收方和发送方的带宽探测方法相比, 这种逐跳的可用带宽测量方法分布了带宽测量, 有利于降低发送方的反馈负载.

我们利用如下的 TCP 吞吐率公式<sup>[13]</sup>计算 TCP 友好<sup>[14]</sup>的可用带宽.

$$T_{TCP} = \frac{s}{RTT \left( \frac{\sqrt{2p}}{\sqrt{3}} + \left( \frac{\sqrt{3p}}{12\sqrt{8}} \right) p(1+32p^2) \right)} \quad (1)$$

其中,  $s$  为分组大小,  $RTT$  为接收者到发送者之间的往返时间,  $p$  为丢包事件率. 对于本方案来说,  $RTT$  为相邻上、下游主动节点之间的往返时间,  $p$  为主动节点 (或接收者) 与相邻上游节点之间链路上的丢包事件率 (loss event rate). 为了估算逐跳链路 TCP 友好可用带宽, 需要估算逐跳链路上的  $RTT$  和  $p$  值.

在  $RTT$  测量过程中, 由接收者周期性地向发送者发送  $RTT$  测量请求的主动控制分组, 中间主动节点在接收到  $RTT$  测量请求分组之后, 发送一个  $RTT$  测量应答分组给接收者; 从而求出两个相邻主动节点间的  $RTT$ , 同时该请求继续向其上游节点转发, 以逐段测量上游链路的  $RTT$ .  $ACP(Measure\_RTT, Request)$  表示测量  $RTT$  的请求主动控制分组,  $ACP(Measure\_RTT, Response)$  表示测量  $RTT$  的应答主动控制分组. 主动控制分组  $ACP$  中携带有时间戳和序号. 以其中某一跳链路 (上游主动节点 AR1 到下游节点 AR2) 为例,  $RTT$  测量时序图如图 3 所示.

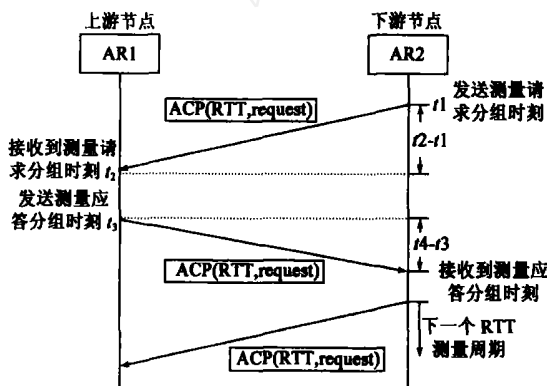


图 3 逐跳  $RTT$  测量时序图

假设接收者在  $t_1$  时刻发送  $ACP(Measure\_RTT, Request)$ ,  $ACP(Measure\_RTT, Request)$  在  $t_2$  时刻到达路径上第一个主动节点. 该主动节点在  $t_3$  时刻发送  $ACP(Measure\_RTT, Response)$ ,  $ACP(Measure\_RTT, Response)$  在  $t_4$  时刻到达接收者, 则该段链路的  $RTT$  测量值为:  $RTT = t_4 - t_3 + t_2 - t_1$ . 当这两个主动节点之间存在时钟偏差  $t$  时,  $RTT = t_4 - (t_3 - t) + (t_2 - t) - t_1 = t_4 - t_3 + t_2 - t_1$ . 因此, 逐跳  $RTT$  测量公式在相邻两节点存在时钟偏差时仍然成立.  $ACP(Measure\_RTT, Request)$  逐跳向上游节点转发, 用同样的方法可以计算出其它逐跳链路的  $RTT$  值. 为得到平滑的  $RTT$  估计值, 采用指数加权移动平均法 (Exponential Weighted Moving Average, EWMA)<sup>[13]</sup>对测量得到的  $RTT$  值进行平滑处理:

$$\hat{r}_{RTT}^i = \alpha \hat{r}_{RTT}^{i-1} + (1 - \alpha) * t_{RTT}^i \quad (2)$$

其中,  $\hat{r}_{RTT}^i$  是第  $i$  次估计值,  $t_{RTT}^i$  是第  $i$  次测量值,  $\alpha$  是指数滑动系数.

丢包事件率  $p$  由接收者或链路下游的中间主动节点负责估算. 丢包事件率  $p$  是指在一个  $RTT$  时间内的一个或多个分组丢失. 因为一个  $RTT$  时间内的多个分组丢失通常是由一次拥塞引起的, 将这些丢失看作一个丢包事件比较合理, 根据丢包事件率可得到较稳定的可用带宽估计. 相邻丢包事件之间的分组数称为丢包间隔 (Loss Interval). 依据加权平均丢包间隔法 (Weighted Average Loss Interval, WALI)<sup>[13]</sup>计算出平均丢包间隔:

$$l_{avg}(k) = \frac{\sum_{i=0}^m w_i l_{k-i}}{\sum_{i=0}^m w_i} \quad (3)$$

其中,  $l_k$  为当前测量到的丢包间隔,  $l_{k-1}, \dots, l_{k-m}$  分别为  $m$  个历史丢包间隔,  $w_0, \dots, w_m$  为对应的权值. 再根据平均丢包间隔计算丢包事件率:

$$p = 1 / l_{avg}(k) \quad (4)$$

综上所述, 在一个可用带宽测量周期内, TCP 友好可用带宽测量的基本步骤如下:

接收者发送  $RTT$  测量请求主动分组  $ACP(Measure\_RTT, Request)$ , 中间节点或发送者接收到  $ACP(Measure\_RTT, Request)$  后, 向接收者反馈响应分组  $ACP(Measure\_RTT, Response)$ , 测量该段链路的往返时间  $RTT^i$ ; 再利用 EWMA 法求  $RTT^i$  的平滑估计值  $RTT$ ;

由每跳链路的下游主动节点 (中间节点或接收者) 测量丢包间隔  $l_k$ ; 再根据当前测量的丢包间隔  $l_k$  及  $m$  个历史丢包间隔, 利用 WALI 法求平滑估计值  $l_{avg}(k)$ , 进而求丢包事件率  $p$ ;

根据测得的  $RTT$  和  $p$  计算 TCP 友好的可用带宽  $T_{TCP}$ ;

根据步骤 ~ 依次逐跳测量上游链路可用带宽.

在上述链路可用带宽测量过程中, 中间主动节点依据测量的出口链路上的 TCP 友好可用带宽相应调整发送速率, 执行 TCP 友好拥塞控制, 使得多播树上每跳链路的流量均满足 TCP 友好, 同时又能很好地适应多播网络的异构性.

### 3.4 主动速率控制机制

在测量逐跳链路的 TCP 友好可用带宽基础上, 我们提出了利用主动节点进行的逐跳链路速率控制机制如下:

( ) 慢启动阶段

当未发生分组丢失时, 每个  $RTT$  时间内数据发送速率翻倍.

$$R_{i+1} = \min(2 * R_i, R_{max}) \quad (5)$$

其中,  $R_{i+1}$  与  $R_i$  分别是当前和上次发送速率,  $R_{max}$  为链路最大可用带宽.  $R_0$  设为分层多播中的基层速率.

( ) 拥塞避免阶段

当发生分组丢失时, 根据第 (3) 节逐段链路 TCP 友好可用带宽测量方法, 计算链路可用带宽  $T_{TCP}$ .

( ) 计算 TCP 友好的接收层数

根据下式计算 TCP 友好的接收层数  $L$ ,

$$L: \sum_{k=0}^{L-1} b_k R < \sum_{k=0}^L b_k \quad (6)$$

其中,  $R$  为由 (I) 和 (II) 计算出的可用带宽,  $b_k$  ( $k=0, \dots, N-1$ ) 为分层多播的基层和各增强层速率,  $N$  为分层多播的最大分层数。显然, 在由 (6) 式计算出的层数  $L$  情况下, 链路的分层累积传输速率  $\sum_{k=0}^{L-1} b_k$  是不超过 TCP 友好可用带宽  $R$  的最大值。接收者再将可接收层数  $L$  反馈给其相邻上游主动节点, 上游主动节点据此调节该链路上的分层多播转发层数。

## 4 仿真实验

### 4.1 仿真环境

我们在网络仿真器 NS-2<sup>[15]</sup> 上对本文提出的主动分层多播拥塞控制方案 HILMA 进行了仿真实验。我们扩展了 NS-2 内核中主动网仿真模块, 增加了 HILMA 功能模块。仿真实验网络拓扑如图 4 所示。

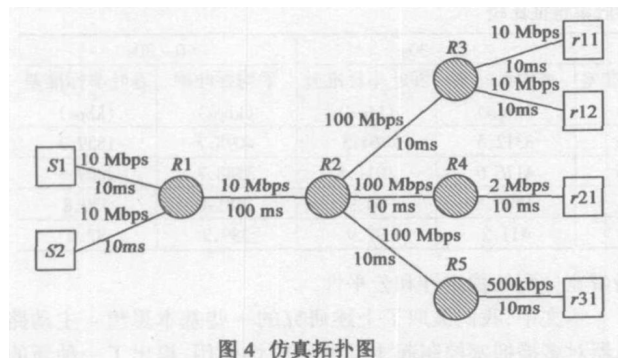


图 4 仿真拓扑图

图 4 中, 路由器 R1 ~ R5 是具有 HILMA 功能的主动路由器, R1 与 R2 之间的链路带宽为 10Mbps, 链路延时 100ms。R2 与 R3、R4、R5 之间的链路带宽均为 100Mbps, 链路延时均为 10ms。端系统 S1、S2 与 R1 之间链路带宽均为 10Mbps, 延时均为 10ms。r11、r12、r21、r31 与 R3、R4、R5 之间的链路带宽分别为 10Mbps、10Mbps、2Mbps、500kbps, 延时均为 10ms。

设定分层多播最大层数为 6 层, 基层到增强层的累积传输速率依次为 200kbps、400kbps、800kbps、1600kbps、3200kbps、6400kbps。该数据分层特性与实际多媒体流的分层传输特征和网络异构特性相一致。

r11、r21、r31 接收 S1 发送的分层多播流量, r12 接收 S2 发送的 TCP 流量。仿真时长为 30s。在 0 ~ 10s 内, S1 和 S2 均发送流量, 在 10s 时刻, S2 停止发送 TCP 流量, 在 20s 时刻, S2 开始再次发送 TCP 流量。目的是为了在 R1-R2 之间的链路拥塞状况变化的情况下, 分析 HILMA 的响应性能和稳定性。

### 4.2 仿真结果与分析

仿真实验得到 HILMA 多播会话接收者 r11、r21、r31 以及 r12 的吞吐率变化如图 5、6 所示。

从图 5 中可以看出, 由于接收者 r11、r21、r31 接入网络带宽的不同, 三个接收者有着不同的吞吐率, 分别接收到不同的接收层次, 能够较好地适应网络可用带宽的异构性。分层多播会话的起始阶段, 在网络中有 TCP 流量的情况下, r11 能够较稳定地接收 5 层数据, 吞吐率相对稳定在 3200kbps。在

10s 时刻, S2 停止发送 TCP 流量, r11 能较快地检测到可用带宽的增加, 接收层数增为 6 层, 吞吐率相对稳定在 6400kbps。在 20s 时刻, S2 开始再次发送 TCP 流量, 使 R1-R2 之间的瓶颈链路发生拥塞, 可用带宽减少, 接收者 r11 能够较快地降低接收层次以适应可用带宽的变化, 并相对稳定地接收 5 层数据。从图 5 中还可以看出, 在网络状况变化的情况下, 接收者 r21、r31 的吞吐率均维持在相对稳定的水平, 分别稳定接收 4 层和 2 层数据。这是因为 HILMA 方案引入了优先级分层传输机制, 在网络发生拥塞时, 优先丢弃最低优先级的最高层数据, 保护了较重要的基层和较低增强层的数据。从接收者吞吐率的变化可以看出, HILMA 方案能较快地响应网络拥塞, 在稳态时接收者吞吐率能维持在相对稳定的层次。

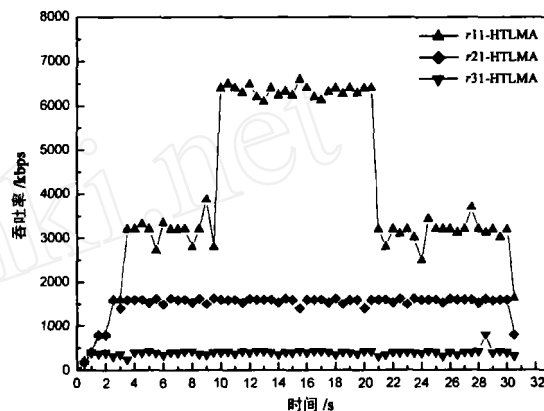


图 5 HILMA 下 r11、r21 和 r31 吞吐率变化

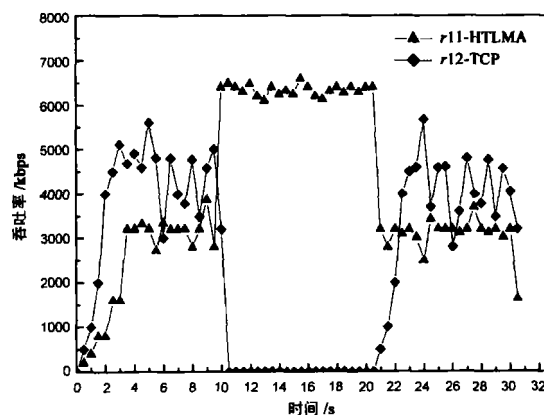


图 6 HILMA 下 r11 与 r12-TCP 吞吐率比较

从图 6 中可以看出, 在分层多播 HILMA 流量与 TCP 流量共享瓶颈链路带宽时, r11 接收到的 HILMA 流量的吞吐率均不超过 r12 接收到的 TCP 流量的吞吐率。这表明了 HILMA 流量是 TCP 友好的。这是因为本文提出的 HILMA 方案引入了逐跳链路探测 TCP 友好可用带宽和主动速率控制机制。

在相同的网络拓扑和 TCP 流量配置下, S1 改为发送 RLM 流量, 仿真实验得到 r11 接收到的 RLM 流量与 r12 接收到的 TCP 流量的吞吐率变化如图 7 所示。

从图 7 中可以看出, RLM 方案在网络状况变化时, 吞吐率存在剧烈抖动, 速率调节存在较大延时, 并且不满足 TCP 友好。这是由于 RLM 的性能受限于 IGMP 中较长的 Leave 延时和“尽力服务”模型中一致丢弃方式。

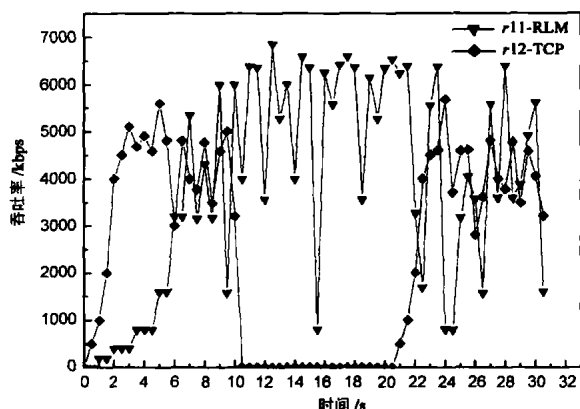


图7 RLM下r11与r12-TCP吞吐率比较

图8反映了在HILMA方案下，r11、r21、r31与r12在0~10s、10~20s以及20~30s时间内的平均吞吐率（r12-TCP表示r12接收的TCP流量）。

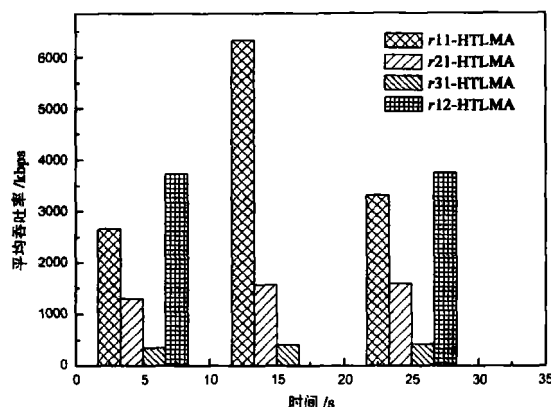


图8 HILMA与TCP流量平均吞吐率比较

其中，在10~20s时间内，s2停止发送TCP流量，这段时间内，网络中只有HILMA流量。在0~10s、20~30s内，HILMA方案下r11的平均吞吐率不高于TCP流量的平均吞吐率。这

表1 HILMA和RLM吞吐率特征比较

|           | 0~10s           |                  | 10~20s          |                  | 20~30s          |                  | 0~30s           |                  |
|-----------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|
|           | 平均吞吐率<br>(kbps) | 吞吐率标准差<br>(kbps) | 平均吞吐率<br>(kbps) | 吞吐率标准差<br>(kbps) | 平均吞吐率<br>(kbps) | 吞吐率标准差<br>(kbps) | 平均吞吐率<br>(kbps) | 吞吐率标准差<br>(kbps) |
| r11-HILMA | 2653.2          | 1428.7           | 6330.5          | 126.3            | 3312.5          | 764.3            | 4098.7          | 1859.3           |
| r11-RLM   | 2055.6          | 2013.1           | 5434.1          | 1523.5           | 4176.6          | 1911.5           | 3588.7          | 2829.5           |
| r21-HILMA | 1301.2          | 520.7            | 1565.3          | 64.7             | 1585.5          | 34.9             | 1481.0          | 330.8            |
| r31-HILMA | 352.7           | 103.9            | 405.7           | 26.2             | 411.2           | 97.9             | 389.2           | 87.0             |

进一步说明了HILMA方案下的分层多播流量具有TCP友好特性。

HILMA与RLM方案在0~10s、10~20s以及20~30s时间内的吞吐率特征的对比关系如表1所示。

在0~10s时间内，也就是分层多播会话的起始阶段，HILMA与RLM吞吐率均有一个不稳定的阶段，但在后续阶段HILMA的吞吐率稳定性显著优于RLM。表1中的实验结果表明，在0~30s时间内，HILMA方案比RLM的平均吞吐率高14.2%，而吞吐率标准差是RLM的65.7%。因此，HILMA在有较高吞吐率的同时，具有较优的稳定性。

## 5 与相关研究工作的比较

文献[7]研究了主动网支持的多播协议及其基本服务元素，按照可靠性要求的不同，将多播分为有时限的可靠多播、无时限的可靠多播和不可靠多播三种方式，提出了主动可靠多播协议RMANP，该文献着重研究多播可靠性问题，没有提出相应的多播拥塞控制方案。文献[8]提出了一个主动网上异构视频多播的应用框架，给出了多播分发树构造和主动节点位置选择的一个启发式算法，利用主动节点对视频流的过滤作用，满足异构接收者不同的质量需求。文献[9]提出了一个基于逐跳信用量机制的主动多播拥塞控制方案，满足接收者不同的可靠性和端到端延时需求。文献[10]从资源供求关系的角度研究了主动节点中链路带宽分配问题，提出了一个适应性的主动分层多播协议ALMA。文献[16]提出了利用优先级丢弃和预丢弃机制的分层多播网络带宽分配方案，提高多播带宽分配的稳定性及公平性。

本文中，我们汲取了上述研究的一些基本思想——主动路由对多播拥塞控制起主动的适应性作用，提出了一种新的逐跳TCP友好的主动分层多播拥塞控制方案HILMA，该方案基于主动网服务模型，采用主动标记分层、优先级分层传输、逐跳链路探测TCP友好可用带宽，以及利用主动节点进行逐跳速率控制。与上述同类方案相比，本文为主动分层多播拥塞控制问题提供了新型的更加完整的方案，并给出了方案HILMA中关键实现机制的描述，整个方案易于实现。此外，我们还通过仿真实验，从拥塞控制响应速度、稳定性和TCP友好等方面对本方案进行了研究分析，验证了本方案的有效性。

## 6 结论和进一步的工作

在分析传统的分层多播拥塞控制方案存在问题的基础上，提出了一种新的逐跳TCP友好的主动分层多播拥塞控制方案HILMA，给出了方案中关键实现机制的描述，通过仿真实验对本文提出的HILMA方案进行了深入分析，并与接收方驱动的分层多播方案RLM进行了比较。仿真实验表明，HILMA在拥塞控制响应速度、稳定性、TCP友好等方面具有较优的性能。为了尽可能减少中间节点的计算处理开销，HILMA方案只需根据分层标记进行分层传输和优先级过滤，并采用优化节点处理开销的轻量化处理模型。当中间节点存在非主动节点时，对分层多播流量提供透明的“尽力而为”服务。这就使得基于主动网服务模型的HILMA方案可以在现有的网络中得到渐进式的部署。进一步的研究工作将是在实际主动网环境中实现HILMA方案，并对HILMA方案的服务

可定制性,以及自适应动态分层策略等问题做进一步的研究.

最后,感谢导师顾冠群教授对本文的审阅与修改建议.

#### 参考文献:

- [1] McCanne S, Jacobson V, Vetterli M. Receiver-driven layered multicast [A]. Proc. of ACM SIGCOMM 1996 [C]. New York: ACM Press, 1996. 117 - 130.
- [2] Viciaro L, Rizzo L, Crowcroft J. TCP-like congestion control for layered multicast data transfer [A]. Proc. of IEEE INFOCOM 1998 [C]. New York: IEEE Computer Society Press, 1998. 996 - 1003.
- [3] Nakauchi K, Morikawa H, Aoyama T. A network-supported approach to layered multicast [A]. Proc. of IEEE ICC 2001 [C]. New York: IEEE press, 2001. 1227 - 1231.
- [4] Tennenhouse D L, Smith J M, Sincoske W D, et al. A survey of active network research [J]. IEEE Communications Magazine, 1997, 35 (1): 80 - 85.
- [5] Hicks M, Moore J T, et al. Experiences with capsule-based active networking [A]. Proc. of the DARPA Active Networks Conference and Exposition [C], New York: IEEE Computer Society Press, 2002. 16 - 24.
- [6] Faber T. Experience with active congestion control [A]. Proc. of the DARPA Active Networks Conference and Exposition [C]. New York: IEEE Computer Society Press, 2002. 132 - 142.
- [7] Calderon M, Sedano M, et al. Active network support for multicast applications [J]. IEEE Network Magazine, May/June 1998. 46 - 52.
- [8] Akamine H, Wakamiya N, Miyahara H. Heterogeneous video multicast in an active network [J]. IEICE Trans. on Comm. January 2002. E85-B (1): 284 - 292.
- [9] Lee J K, Kim T G. Active multicast congestion control with hop-by-hop credit-based mechanism [J]. IEICE Trans. on Comm. March 2002. E85-B (3): 614 - 622.
- [10] Yamamoto L, Leduc G. An active layered multicast adaptation protocol [A]. Proc. of IWAN 2000 [C]. Berlin: Springer Press, October 2000. 179 - 194.
- [11] Alexander D, Braden B, et al. Active network encapsulation protocol [DB/OL]. <http://www.cis.upenn.edu/switchware/ANEP/>. July 1997.
- [12] Fenner W. Internet group management protocol, Version 2 [S]. IETF RFC 2236, November 1997.
- [13] Widmer J, Handley M. Extending equation-based congestion control to multicast applications [A]. Proc. of ACM SIGCOMM 2001 [C]. New York: ACM Press, 2001. 275 - 285.
- [14] Floyd S, Fall K. Promoting the use of end-to-end congestion control in the Internet [J]. IEEE/ACM Trans on Networking, August, 1999. 7 (4): 458 - 472.
- [15] McCanne S, Floyd S, et al. LBNL network simulator, ns-2 [CP/OL]. <http://www.isi.edu/nsnam/ns>. 1997 - 10 - 01/2004 - 05 - 18.
- [16] 翟明玉. 组播网络带宽分配研究 [D]. 南京: 东南大学计算机科学与工程系, 2002.

#### 作者简介:



**叶晓国** 男, 1975 年生于安徽桐城, 东南大学博士研究生, 主要研究方向为高性能网络体系结构及协议、主动网络技术、网络拥塞控制等。  
E-mail: xgye@seu.edu.cn.



**吴家军** 男, 1969 年生于江苏南京, 东南大学博士研究生, 主要研究方向为高性能网络体系结构及协议、Overlay 网络等。