

# 一种基于 PGM 的单速率组播拥塞控制方案\*

谭连生<sup>+</sup>, 熊乃学, 杨燕

(华中师范大学 计算机科学系,湖北 武汉 430079)

## A PGM-Based Single-Rate Multicast Congestion Control Scheme

TAN Lian-Sheng<sup>+</sup>, XIONG Nai-Xue, YANG Yan

(Department of Computer Science, Central China Normal University, Wuhan 430079, China)

+ Corresponding author: Phn: +86-27-67867651, E-mail: L.tan@ccnu.edu.cn, <http://www.ccnu.edu.cn>

Received 2003-08-29; Accepted 2003-11-07

Tan LS, Xiong NX, Yang Y. A PGM-Based single-rate multicast congestion control scheme. *Journal of Software*, 2004,15(10):1538-1546.

<http://www.jos.org.cn/1000-9825/15/1538.htm>

**Abstract:** PGM (pragmatic general multicast) is a reliable multicast transport protocol, which runs on IP networks. Because PGM standard does not include any congestion control scheme in itself, it can't achieve real-time response to satisfy the requirements of network and can't timely adjust the sending rate. Based on the PGM protocol, a novel close-loop controller, which is suggested to be implemented between the sender and the CLR (current limiting receiver), is introduced in this paper to timely adjust the sending rate. The proposed rate control scheme brings the source's sending rate of the network to be stable gradually and has fast response. Simulations are carried out for the proposed control scheme, and the simulation results demonstrate that the proposed algorithm has the following properties: scalability, stability, fast convergence, and fast response under various network environments. They show good performance of the proposed multicast control schemes.

**Key words:** buffer occupancy; congestion control; multicast; PGM (pragmatic general multicast); stability

**摘要:** PGM(pragmatic general multicast)是一种在IP协议中广泛应用的可靠的组播传输协议.但PGM标准本身没有拥塞控制方案,不能实时响应网络需求,及时地调节源端发送速率.针对这个问题,在保证PGM协议可扩展性的基础上,在发送方与CLR(current limiting receiver)之间采用了一种新的闭环控制器来实时地调节源端的发送速率,使其逐渐趋于稳定,并具有较快的响应速度.而且在网络拓扑结构动态变化的情况下,对所提出的拥塞控制方案进行了仿真实验.仿真结果表明,所提出的算法具有较好的可扩展性、稳定性和较快的响应速度,控制方案使网络性能表现良好.

**关键词:** 缓冲占有量;拥塞控制;组播;PGM(pragmatic general multicast);稳定性

中图法分类号: TP393 文献标识码: A

\* Supported by the National Natural Science Foundation of China under Grant No.60174043 (国家自然科学基金); the Key Project of Natural Science Foundation of Hubei Province of China under Grant No.2002AB025 (湖北省自然科学基金重点项目)

**作者简介:** 谭连生(1965—),男,湖南宁乡人,博士,教授,博士生导师,主要研究领域为计算机通信网络;熊乃学(1978—),男,硕士生,主要研究领域为计算机通信网络;杨燕(1979—),女,硕士生,主要研究领域为计算机通信网络.

所谓拥塞(congestion),简单地讲,即在某一时刻当网络中某一资源的到达量超过了该资源在相关网络节点的承载量时,则称该资源在该节点该时间内产生了拥塞.在计算机网络中,单播(unicast)、广播(broadcast)和组播(multicast)是最基本的 3 种通信方式,网络节点之间报文交换都是以这 3 种通信模式之一来完成的.

随着因特网应用的日益增长,网络中常见的单播和广播技术的局限性表现得越来越明显.单播技术即传统的点到点的数据传输技术,由于要在双方之间建立单独的数据通道,如果大量用户同时访问同一个数据服务器,所建立的数据通道数目必然和用户数相同,导致了大量的数据冗余传输,加重了网络负担,使网络扩展和应用受到限制.而广播的机制决定了它只能应用在局域网中(路由器为其传输边界),同时,它传输数据给所有的用户,没有定制预约能力.

组播是一种为优化使用网络资源而产生的技术,依托 IP 协议完成.一方面,克服了网络的冗余使用,强制在数据传递树的分叉处(路由器)进行数据包复制,而不是由源端节点多次重复发送数据拷贝.另一方面,虽然组播技术有效地缓解了网络的拥塞情况,但是随着组播应用的发展,其本身的拥塞控制问题越来越突出,也成为组播发展的瓶颈.IETF 将组播划分为两类<sup>[1]</sup>:单点对多点模式和多点对多点模式.对于多点对多点模式,由于其在组播树的生成和组播管理等方面存在较大的难度,现阶段在 Internet 上推广的希望不大,而且它可以通过多组单点对多点的通信来实现<sup>[9-17]</sup>.

本文在保证原有 PGM(pragmatic general multicast)组播协议的可扩展性的基础上,在发送方与被选的代表之间采用了一种新的闭环控制器来实时地调节源端的发送速率,使得源端的发送速率逐渐趋于稳定,这样既有效减少了源端与瓶颈节点之间的响应时间,又有效避免了网络拥塞的发生,提高了整个网络的性能.我们对所提出的拥塞控制方案在网络拓扑结构动态变化的环境下进行了仿真实验.仿真结果说明,所提出的算法具有较好的可扩展性、较快的响应速度,并且能使源端的发送速率、瓶颈节点的缓冲占有量逐渐趋于稳定,控制方案使网络性能表现良好.

本文第 1 节对系统模型进行描述.第 2 节提出一种基于 PGM 的单速率组播拥塞控制方案,并给出算法的具体实现.为了证实所提出的控制方案的性能,第 3 节给出仿真结果及相关分析.第 4 节总结并对未来工作进行讨论.

## 1 系统模型

数字通信网络通常由一系列局域分布的发送、接收节点(source/destination nodes)组成.由发送端激发的数据包通过一系列中间节点被输送到接收节点.而在网络通信过程中,发送端和接收端通过控制包 CP(control packet)来进行连接以及时了解对方的情况.通常在每个周期,发送端都会向接收端发送一个前向控制包 FCP(forward control packet),告知接收端节点有关发送端的情况.FCP 到达接收节点以后,接收节点会根据当时网络的具体情况,产生一个相应的后向控制包 BCP(backward control packets),根据 SPM(source path messages)按 FCP 的路径返回发送端,以便使发送端了解网络的情况和接收端节点的状态,及时地调节源端的发送速率以满足接收节点的需要.为了方便起见,我们考虑如图 1 所示的单点对多点组播系统模型<sup>[2]</sup>.假设该模型有如下特征:

- (1) 该模型是面向连接的,数据以固定长度的数据包来传输;
- (2) 将时间按照周期  $T$  分割成区间  $[n, n+1]$ ;
- (3) 组播树中分支节点将来自上游节点(即发送端节点)的数据包(包含控制数据包)复制后,传送给下游的各个分支节点,而且分支节点还将来自下游的 BCP 数据包进行整合后传送给其上游节点;
- (4) 发送端到最近的分支节点的链路延迟为  $\tau_0$ ,发送端到末端分支节点  $i$  的前向链路延迟记为  $\tau_i (i=1, 2, \dots, N)$ ,这里,  $\tau_0$  和  $\tau_i$  均为周期  $T$  的整数倍(如果不是,可以对周期  $T$  进行调整,使之成立);
- (5) 每个分支节点均按照先来先服务(first-come-first-service,简称 FCFS)的原则,传输接收到的数据包;
- (6) 在  $n$  时刻,第  $i$  个分支节点的缓冲占有量记为  $x_i(n)$ ,其最大值记为  $K_i$ ,理想值记为  $\bar{x}_i$  ( $\bar{x}_i < K_i$ ),若某时刻  $x_i(n) > K_i$ ,则称该节点发生拥塞;
- (7) 第  $i$  个分支节点在一个周期内可以向网络发送的数据包数为  $L_i$ ,即第  $i$  个节点按照最大发送速率<sup>[3]</sup>(PCR)在一个周期内所发送数据包数;

(8) 从发送端到各个分支节点的路径延迟从小到大排列,即当  $i \leq j$  时,  $\tau_i \leq \tau_j$ .

从图 1 所示的网络模型,分支节点  $i$  的动态缓冲占有量可以描述如下<sup>[2]</sup>:

$$x_i(n+1) = \text{Sat}_{K_i} \{x_i(n) + R(n - \tau_i) - v_i\} \tag{1}$$

其中

$$\text{Sat}_{K_i} \{x_i\} = \begin{cases} K_i, & x_i > K_i \\ x_i, & 0 \leq x_i \leq K_i \\ 0, & x_i < 0 \end{cases}$$

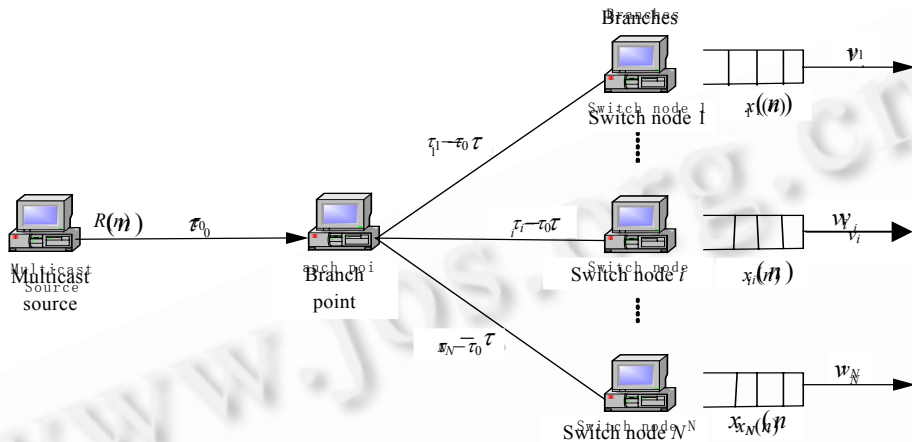


Fig.1 The multicast system model of single node to multiple nodes

图 1 单点到多点的组播系统模型

在网络发生拥塞之前,应使分支节点的动态缓冲占有量尽量稳定在阈值附近.为不失一般性,可以去掉式(1)的非线性约束<sup>[4]</sup>,代之以考虑如下线性方程

$$x_i(n+1) = x_i(n) + R(n - \tau_i) - v_i \tag{2}$$

对式(2)两边施行 Z 变换后,有

$$(z-1)X_i(z) = z^{-\tau_i}R(z) - v_iD(z) \tag{3}$$

其中,记  $X_i(z) = \sum_{n=0}^{+\infty} x_i(n)z^{-n}$ ,  $R(z) = \sum_{n=0}^{+\infty} R(n)z^{-n}$ ,  $D(z) = \frac{z}{z-1}$ .

## 2 一种基于 PGM 的单速率组播拥塞控制方案

### 2.1 PGM

PGM<sup>[5]</sup>是一种单个发送方、多个接收方的组播协议.它通过利用基于 NAK(negative acknowledgment)的重传请求,在基本 IP 组播协议基础上改进了传输的可靠性.在 PGM 中使用的基于随机延迟策略使协议具有可扩展性.而且,PGM 有效利用(不是依赖)路由器对于反馈聚集和选择性前向修复的支持.特别是,PGM 路由器将做跳式(hop-by-hop)的前向传输,聚集来自同一个子树的复制的 NAKs.为了使反馈信息通过和 FCP 相同的路径,PGM 使用一种特殊的控制信息 SPM.SPM 每经过一个路由器就被 PGM 路由器重写,并通知上游的 PGM 节点.

因为 PGM 标准本身不包括拥塞控制方案,源端按提前设定的数据速率进行传输,不能实时响应网络需求并及时地调节源端发送速率.因此,我们对 PGM 协议的修改包括对 PGM 控制包 CP 的一些选项的修改以及对发送方和接收方处理过程的修改是很有意义的.现阶段,对 PGM 的探讨多局限于基于窗口的改进,如文献 [5,8,18~20],而它们不能实时地调节源端发送速率.针对这个问题,本文提出的方案在保持 PGM 原有反馈机制的可扩展性的基础上,尽量使源端快速地响应网络的拥塞状况.通过在组播树中选择一个代表 CLR(current limiting receiver)<sup>[6]</sup>,然后在 CLR 和组播源端之间运行一个闭环的、实时反馈的控制方案.对于 CLR 的选择,与通

常的单速率组播协议相同——把具有最低吞吐量的接收者作为 CLR<sup>[18-20]</sup>.

### 2.2 接收端

对于控制包 CP 的修改,我们增加以下几个域(fields):当前瓶颈接点 CLR 的缓冲占有量(buffer occupancy)、队列阈值(threshold)和该节点的序号(identity).当接收端周期性地收到一个 FCP 时,接收端立即根据当前网络的情况负责填写 BCP 中的各域,并将填好的 BCP 返回到它的上游节点,以使组播源端能及时地了解网络拥塞状况,调节发送速率,及时地满足网络的需求.对接收端节点的分组:首先把具有相同或相近前向延迟的接收节点分成一组;然后在上面的每一组中依据接收节点的发送速率再次分组,即具有相同或相近的发送速率的接收者再分成一组.

### 2.3 发送端拥塞控制方案的设计

为简便起见,我们对所有的接收节点进行归并处理,即把具有相似吞吐量和环路延迟的接收者分为一组.在所有的分组中选取吞吐量最小的一组作为瓶颈节点 CLR.考虑在发送节点与 CLR 之间设计具有如下反馈形式的控制器:

$$R(n) = \mu + \sum_{i=1}^N a_i (x_i(n - \tau_i) - \bar{x}_i) + \sum_{j=1}^{\tau} b_j R(n - j) \quad (4)$$

其中  $R(n)$  表示源端发送速率,  $\mu$  表示源端最大发送速率,  $\bar{x}_i$  表示瓶颈节点的缓冲占有量的理想值,  $\tau$  和  $N$  分别表示瓶颈节点 CLR 的游程延迟和数目.  $a_i$  和  $b_j$  为控制参数,由系统稳定性确定.

对上式两端进行  $z$  变换,有

$$R(z) = \mu D(z) + \sum_{i=1}^N a_i (z^{-\tau_i} X_i(z) - \bar{x}_i D(z)) + \sum_{j=1}^{\tau} b_j z^{-j} R(z) \quad (5)$$

将式(3)代入式(5),整理后可得

$$\Delta(z)R(z) = \mu(z-1) \cdot D(z) - \sum_{i=1}^N a_i [(z-1)(z^{-\tau_i} v_i D(z) + (z-1)\bar{x}_i \cdot D(z))],$$

其中

$$\Delta(z) = z - 1 - \sum_{j=1}^{\tau} b_j z^{-j} (z-1) - \sum_{i=1}^N a_i z^{-2\tau_i}.$$

由于对作为瓶颈节点的一组 CLR 节点进行归并处理,则  $a_1 = a_2 = \dots = a_N$ ,  $\tau_1 = \tau_2 = \dots = \tau_N = 1/2 \cdot \tau$ . 所以上式可变为

$$\Delta(z) = z - 1 - \sum_{j=1}^{\tau} b_j z^{-j} (z-1) - N \cdot a_1 \cdot z^{-\tau},$$

$$\Delta(z) = -z^{-\tau} [-z^{\tau+1} + (1+b_1)z^{\tau} + (b_2-b_1)z^{(\tau-1)} + \dots + (b_{\tau}-b_{(\tau-1)})z + (N \cdot a_1 - b_{\tau})] \quad (6)$$

$\Delta(z)$  即为闭环控制系统(2),(4)的特征多项式(characteristic polynomial),决定系统的稳定性.

为了获得系统的稳定性,由文献[7],我们可对特征多项式(6)取特殊情况,令  $1+b_1 = b_2 - b_1 = \dots = \varepsilon$ ,  $N \cdot a_1 - b_{\tau} = 0$ . 即

$$\begin{cases} a_1 = (\tau\varepsilon - 1) / N \\ b_j = j\varepsilon - 1, (j = 1, 2, \dots, \tau) \end{cases} \quad (7)$$

把式(7)代入式(6)可以得到

$$\Delta(z) = -z^{-\tau} [-z^{\tau+1} + \varepsilon(z^{\tau} + z^{(\tau-1)} + \dots + z^2 + z)] \quad (8)$$

文献[7]中已经证明,当  $\varepsilon < \frac{1}{\tau+1}$  成立时,特征方程  $\Delta(z)$  的根均在单位圆内,此时源端的发送速率逐渐稳定.

### 2.4 中间分支节点NE(network element)的作用

这里,我们采取的方案和 PGM 协议相同,用 NE<sup>[8]</sup>执行反馈聚积策略以提高可扩展性.PGM 路由器通过两种方式节省了网络带宽:一是通过大量减少反馈 NAKs 的数量;二是通过仅仅重传数据包给有数据包丢失的节点,

从而限制了大量的数据重传.在 PGM 中,对于某个给定的数据包,仅有第 1 个 NAK 被传递到源端,其他来自接收者的 NAKs 均被 NE 过滤掉.接下来的 NAKs 都被抑制,一直等到 NE 中相应的状态被删除.表面来看,这种过滤方式可能会影响 CLR 的选择,但实际上由于具有最差接收能力的接收者通常发送更多的 NAKs,从而减少了被抑制的可能性.所以 NAK 的反馈方式具有可扩展性,而且避免了“反馈风暴”的发生.

## 2.5 基于 PGM 的单速率组播拥塞控制算法的具体描述

在未建立连接时,CLR 没有被选择出来,组播源端不能运行拥塞控制算法来调节发送速率,因此,我们先以最大发送速率进行发包.建立连接后,接收端用基于 NAKs 的机制来选择 CLR,并将 CLR 的相关信息写入 ACK 中,通过 NE 聚合后返回到源端,从而运行拥塞控制算法来调节发送速率.由于在任意时刻网络拓扑结构是确定的,因此在任意时刻,基于 PGM 的单速率组播拥塞控制算法可见如下具体描述.

### 源端算法.

在连接建立阶段,源端以最大发送速率发包,并通知所有的接收节点开始接收数据包;

Upon 收到来自其下游节点的 CLR 的 ACK 控制信息

取出 ACK 的 buffer occupancy 和 threshold 等信息;

根据本文所提出的控制方案,计算下一时刻的发送速率;

根据计算的发送速率发送数据包;

Upon 收到来自其下游节点的 NAK 控制信息

比较 NAK 和上一次会话的 ACK 的信息,由具有最差接收能力的接收者来作为新的 CLR;

### 分支节点算法.

Upon 收到一个来自其上游的包:

if 收到的是控制包(FCP) //判断包的类型

{

将当前网络的有关信息填入 FCP 的对应域;

复制并转发 FCP 到下游节点;}

else{

复制并转发该包的复本给所有下游节点;}

Upon 收到一个来自下游节点的反馈信息包

if 收到一个来自 CLR 的反馈控制包 ACK

{

继续向其上游节点转发该控制包;}

if 在一段时间内,收到是第 1 个 NAK

{

继续向其上游节点转发该控制包;}

else if 收到 NAK 包不是第 1 个 NAK

{

抑制该控制包的发送;}

### 接收端算法.

在所有的接收者中选择一个具有最差的接收能力的接收者作为 CLR;

把 CLR 的相关信息如当前的 bufer occupancy 和 queue threshold 等填入 FCP 控制包中,生成新的 ACK;

将新生成的 ACK 包反馈到其上游节点.

## 3 仿真实验与分析

在仿真实验中(如图 2 所示),有一个组播发送端(multicast source)、3 个路由器(router 1,router 2 和 router 3)

和 3 组接收者,即第 1 组接收者 Receiver 1,Receiver 2,...,Receiver 100,第 2 组接收者 receiver 101,Receiver 102,...,Receiver 200 和第 3 组接收者 receiver 201,Receiver 202,...,Receiver 300.它们之间的链路延迟分别在图中标出,其他的参数设置如下:采样时间周期  $T$  为 1m sec 源端最大发送速率为 6Mbps,第 1 组接收者发送数据包的速率为  $v_i = 2 Mbps (1 \leq i \leq 100)$ , 缓冲队列的阈值为  $\bar{x}_i = 120 Kb (1 \leq i \leq 100)$ . 第 2 组接收者发送数据包的速率为  $v_k = 4 Mbps (101 \leq k \leq 200)$ , 缓冲队列的阈值为  $\bar{x}_k = 180 Kb (101 \leq k \leq 200)$ . 第 3 组接收者发送数据包的速率为  $v_j = 1 Mbps (201 \leq j \leq 300)$ , 缓冲队列的阈值为  $\bar{x}_j = 80 Kb (201 \leq j \leq 300)$ .

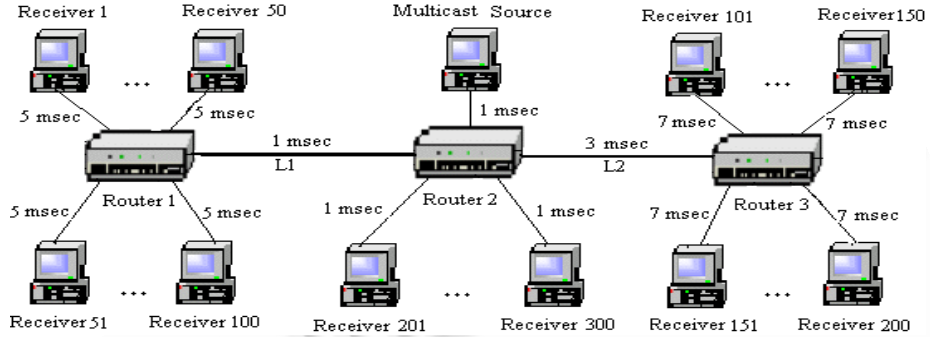


Fig.2 Simulation model

图 2 仿真模型

由于源端控制器是根据 CLR 来进行速率调节,因此需要首先在所有的接受节点分组中选取吞吐量最小的一组作为瓶颈节点 CLR,而此组 CLR 中各个节点的情况都相似,则我们只取其中 1 个节点作为代表.本文在网络拓扑结构动态变化的情况下对所提出的拥塞控制方案进行了仿真实验.

对于所提控制方案参数的优化选择问题,我们选取了使系统稳定和使系统不稳定的两组参数进行比较.在时间区间[1m sec,800m sec]内,只有第 1 组和第 2 组接收节点,此时第 3 组接收节点还未加入网络拓扑结构中,因此 CLR 节点是第 1 组接收节点.这里我们取两组  $\epsilon$  值,即  $\epsilon = 1/16$  (使系统稳定)和  $\epsilon = 5/8$  (使系统不稳定),进行仿真比较.由等式(7),通过计算可以得到,当  $\epsilon = 1/16$  时,控制参数分别为

$$a_1 = a_2 = \dots = a_{100} = -1/800,$$

$$b = [b_1, b_2, \dots, b_{14}] = [-15/16, -14/16, -13/16, -12/16, -11/16, -10/16, -9/16, -8/16, -7/16, -6/16, -5/16, -4/16, -3/16, -2/16].$$

当  $\epsilon = 5/8$  时,控制参数分别为

$$a_1 = a_2 = \dots = a_{100} = 31/400,$$

$$b = [b_1, b_2, \dots, b_{14}] = [-3/8, 2/8, 7/8, 12/8, 17/8, 22/8, 27/8, 32/8, 37/8, 42/8, 47/8, 52/8, 57/8, 62/8].$$

而在 [801msec,1600msec] 时,第 3 组接收节点加入网络拓扑结构中,此时仿真模型有如图 2 所示结构,CLR 节点转移为第 3 组节点.这里,我们取使系统稳定的  $\epsilon$  值为  $\epsilon = 1/6$ ,使系统不稳定的  $\epsilon$  值仍为  $\epsilon = 5/8$ ,进行仿真比较.由等式(7),通过计算可以得到,当  $\epsilon = 1/6$  时,控制参数分别为

$$a_1 = a_2 = \dots = a_{100} = -1/300, \quad b = [b_1, b_2, b_3, b_4] = [-5/6, -4/6, -3/6, -2/6].$$

当  $\epsilon = 5/8$  时,控制参数分别为

$$a_1 = a_2 = \dots = a_{100} = 3/200, \quad b = [b_1, b_2, b_3, b_4] = [-3/8, 2/8, 7/8, 12/8].$$

仿真图 3~图 6 中,实线表示系统稳定时的情况,而虚线表示系统不稳定时的情况.图 3 分别显示了  $\epsilon$  取不同值时组播源端发送速率随时间变化的情况.其中在时间区间 [1msec,800msec] ( $\epsilon = 1/16$ ) 内,源端逐渐调节其发送速率,到 182msec 时发送速率稳定在 2Mbps,而在时间区间 [801msec,1600msec] ( $\epsilon = 1/6$ ) 内,由于第 3 组节点加入网络,CLR 节点改变为第 3 组节点.源端发送速率迅速调整,在 844msec 时,发送速率稳定为 1Mbps.而在整个仿真期间,当  $\epsilon = 5/8$  时,系统不稳定,源端抑制数据包的发送,导致在 14msec 以后其发送速率为 0.

图 4~图 6 分别显示了  $\epsilon$  取不同值时接收端节点 1,101,201 缓冲占有量的瞬态行为的仿真.可知,在时间区间 [1msec,800msec] ( $\epsilon = 1/16$ ) 内,第 1 组的接收节点为瓶颈节点 CLR.在图 4 中,当  $\epsilon = 1/16$  时,开始由于速率没有得到

及时调节,导致缓冲占有量不稳定,然后在[1msec,185msec]期间逐渐稳定在 63.3Mbps,而当系统不稳定( $\varepsilon=5/8$ )时,缓冲占有量在 46msec 时就变为 0,这实际上是由于源端被抑制其发送速率为 0 所造成的.在图 5 中,当 $\varepsilon=1/16$ 时,仅[11msec,37msec]期间有数据堆积,而当系统不稳定( $\varepsilon=5/8$ )时,仅[11msec,31msec]期间有数据堆积.在图 6 中,因为此仿真期间第 3 组节点没有加入网络中,故缓冲占有量为 0.

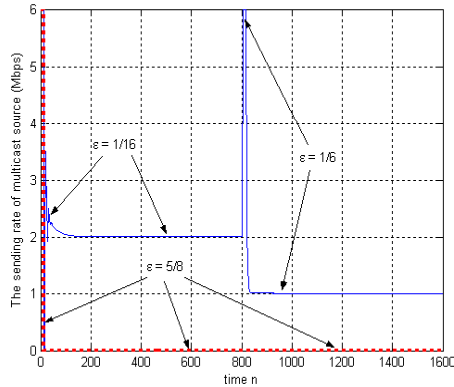


Fig.3 The sending rate of multicast source  
图 3 组播源端的发送速率

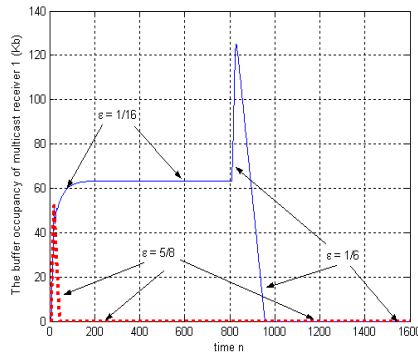


Fig.4 The buffer occupancy of multicast receiver 1 (bottleneck node)  
图 4 接收节点 1(瓶颈节点)的缓冲占有量

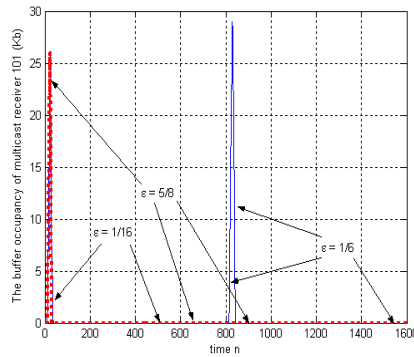


Fig.5 The buffer occupancy of multicast receiver 101  
图 5 接收节点 101 的缓冲占有量

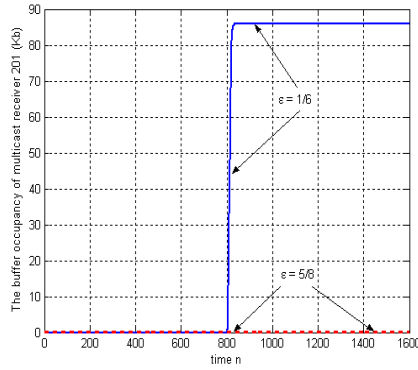


Fig.6 The buffer occupancy of multicast receiver 201

图 6 接收节点 201 的缓冲占有量

在时间区间 $[801\text{msec}, 1600\text{msec}]$ ( $\varepsilon=1/6$ )内,瓶颈节点 CLR 变为第 3 组的接收节点.在图 4 中,当 $\varepsilon=1/6$ 时,由于速率逐渐稳定在 1Mbps,导致队列中有数据包在 $[803\text{msec}, 957\text{msec}]$ 期间堆积.在图 5 中,当 $\varepsilon=1/6$ 时,仅在 $[814\text{msec}, 842\text{msec}]$ 期间队列中有数据包堆积.在图 6 中,因为此期间第 3 组节点开始加入网络,故缓冲占有量在 843msec 时稳定在 86Kb.而此仿真期间系统不稳定( $\varepsilon=5/8$ )时,缓冲占有量均仍被抑制为 0,因为不稳定的参数使控制器没有起到调节作用.从此看来,选取适当的使系统稳定的参数是必要的,而且是很有意义的.

从以上仿真结果中可知,当 $\varepsilon$ 取值使系统稳定时,在时间区间 $[0, \tau_1]$ ,组播发送端按照最大发送速率向网络发送数据包,但这些数据包并未到达分支节点.因此,各节点的缓冲占有量为 0,在时间区间 $[\tau_1, 2\tau_1]$ ,由于经过中间节点聚合后的 CP 控制包尚未到达发送端,因此,组播发送端仍然按照最大速率向中间节点发送数据包,数据包也开始在各个分支节点缓冲区中排队等待处理,各个分支节点都开始有数据包积压.经过一个 $\tau_{R1}$ 之后,CP 控制包携带终端节点的有关信息反馈到发送节点,控制器开始对发送节点的发送速率进行调节.随着时间的推移,各个分支节点的相关信息逐渐通过中间节点聚合后,反馈给发送节点.经过一段时间的调节后,发送节点的发送速率逐渐稳定,瓶颈节点的缓冲区占有量逐渐趋于稳定,而其他节点的缓冲区中积压的数据包也开始清空,缓冲区始终保持在 0Kb,当系统稳定后,源端发送速率与其发送速率相等.在网络拓扑结构发生改变时,此控制方案能重新选择 CLR,并及时调节源端发送速率,使系统重新稳定.

从源端发送速率的调节情况和缓冲占有量的稳定情况等方面综合比较,可以发现,系统稳定时( $\varepsilon = 1/16$  和  $\varepsilon = 1/6$ )比系统不稳定( $\varepsilon = 5/8$ )有更好的网络性能.而且可以做到:在接收节点缓冲占有量临近门限值时,可以尽快通知源端减少发送速率,防止数据包丢失;当系统的可用带宽增加时,可以尽快地通知源端提高发送速率,尽快地利用可用带宽,提高带宽利用率.

## 4 结 论

本文在保证 PGM 组播协议可扩展性的基础上,提出了一种新的拥塞控制方案来实时地调节源端的发送速率,使得源端的发送速率和瓶颈节点的缓冲占有量均趋于稳定.为了更好地证实所提出的拥塞控制方案具有良好的性能,我们在动态网络拓扑模型下对所提出的拥塞控制方案进行了仿真实验.仿真结果表明,所提出的算法具有较好的可扩展性、稳定性和较快的响应速度,控制方案使网络性能表现良好.对于所提出的控制方案的 TCP 友好性问题有待进一步研究.

## References:

- [1] Mankin A, Romanow A, Bradner S, Paxson V. IETF criteria for evaluating reliable multicast transport and application protocols. RFC 2357, 1998.
- [2] Lee SH, Lim JT. Multicast ABR service in ATM networks using a fuzzy-logic-based consolidation algorithm. IEE Proc. of the Communication, 2001, 148(1):8~13.



- [3] The ATM Forum Technical Committee. Traffic Management Specification Version 4.1. AF-TM-0121.000, 1999.
- [4] Benmohamed L, Meekov SM. Feedback control of congestion in packet switching networks: The case of a single congested node. *IEEE/ACM Trans. on Networking*, 1993,1(6):693~707.
- [5] Speakman T, *et al.* PGM reliable transport congestion control for layered multicast data transfer. Internet Draft. RFC 3208, <ftp://ftp.rfc-editor.org/in-notes/rfc3208.txt>
- [6] Speakman T, *et al.* RFC 3208: PGM reliable transport protocol. 2001. <http://www.networksorcery.com/enp/rfc/rfc3208.txt>
- [7] Tan LS, Pugh AC, Yin M. Rate-Based congestion control in ATM switching networks using a recursive digital filter. *Control Engineering Practice (Special Issue on Control Methods for Telecommunication Networks)*, 2003,11(10):1171~1181.
- [8] Rizzo L. pgmcc: A TCP-friendly single-rate multicast congestion control scheme. In: *Proc. of the ACM SIGCOMM 2000*. ACM Press, 2000. 17~28.
- [9] Puangpronpitag S, Boyle R. Performance comparison of explicit rate adjustment with other multi-rate multicast congestion control. In: *Proc. of the 19th UK Performance Engineering Workshop (UKPEW 2003)*. Warwick, 2003. 142~153.
- [10] Paangpronpitag S, Boyle R, Djemame K. Performance evaluation of layered multicast congestion control protocols: FLIDDL vs. PLM. In: *Proc. of Int'l Symp. on Performance Evaluation of Computer and Telecommunication Systems (SPECT 2003)*. 2003.
- [11] Li J, Kalyanaraman S. MCA: A rate-based end-to-end multicast congestion avoidance scheme. In: *Proc. of the ICC 2002*. 2002.
- [12] Shi S, Waldvogel M. A rate-based end-to-end multicast congestion control protocol. In: *Proc. of the IEEE Workshop in Enterprise Security (WETICE)*. MIT, 2001.
- [13] Widmer J, Handley M. Extending equation-based congestion control to multicast application. In: *Proc. of the ACM SIGCOMM 2001*. 2001.
- [14] Chiu DM, Kadansky M, Provino J, Wesley J, Bischof H, Zhu H. A congestion control algorithm for tree-based reliable multicast protocols. In: *Proc. of the IEEE INFOCOM 2002*. 2002.
- [15] Tan LS, Yang SH. Rate-Based congestion controllers for high-speed computer networks. In: *Proc. of the Int'l Federation of Automatic Control 15th IFAC World Congress*. 2002.
- [16] Tan LS, Liu Q, Chen L, Li J. A feedback congestion control method to ABR service in high speed ATM networks. In: Li L, Kato K, Kameda H, eds. *Proc. of the Networks, Parallel and Distributed Processing, and Applications*. 2002.
- [17] Pugh AC, Tan LS. A generalized chain-scattering representation and its algebraic system properties. *IEEE Trans. on Automatic Control*, 2000,45(5):1002~1007.
- [18] Bérard B, Bouyer P, Petit A. Analysing the PGM protocol with UPPAAL. In: *Proc. of the 2nd Workshop on Real-Time Tools (RT-TOOLS 2002)*. 2002.
- [19] Bérard B, Bouyer P, Petit A. Une analyse du protocole PGM avec UPPAAL. In: *Actes 4ième Colloque sur la Modélisation des Systèmes Réactifs (MSR 2003)*. 2003. 415~430.
- [20] Rizzo L, Iannaccone G, Vicisano L, Handley M. PGMCC single rate multicast congestion control: Protocol specification. Internet-Draft, <draft-ietf-rmt-bb-pgmcc-00.txt>, 23, 2001.