

## 二进制流量控制算法的性能分析\*

任丰原<sup>1+</sup>, 林闯<sup>1</sup>, 任勇<sup>2</sup>, 山秀明<sup>2</sup>

<sup>1</sup>(清华大学 计算机科学与技术系,北京 100084)

<sup>2</sup>(清华大学 电子工程系,北京 100084)

### Performance Analysis of the Binary Flow Control Algorithm

REN Feng-Yuan<sup>1+</sup>, LIN Chuang<sup>1</sup>, REN Yong<sup>2</sup>, SHAN Xiu-Ming<sup>2</sup>

<sup>1</sup>(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

<sup>2</sup>(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

+ Corresponding author: Phn: 86-10-62783596, Fax: 86-10-62771138, E-mail: renfy@csnet1.cs.tsinghua.edu.cn

<http://icm.cs.tsinghua.edu.cn>

Received 2001-11-02; Accepted 2002-01-18

Ren FY, Lin C, Ren Y, Shan XM. Performance analysis of the binary flow control algorithm. *Journal of Software*, 2003,14(3):612~618.

**Abstract:** ABR (available bit rate) flow control is an effective measure in ATM network congestion control and traffic management. In large scale and high-speed network, the simplicity of the algorithm is crucial to optimize the switch performance. Though the simplicity of binary flow control is very attractive, the queue length and allowed cell rate (ACR) controlled by the standard EFCI algorithm oscillate with great amplitude, which has negative impact on the performance, so its applicability is doubted, and then relatively complex but effective explicit rate feedback algorithms are introduced and explored. In this study, based on the existed flow control model, the performance of standard EFCI algorithm is evaluated and analyzed with the describing function approach in nonlinear control theory, concluding that queue and cell rate self-oscillations are caused by the inappropriate nonlinear control law originated from intuition, but not intrinsic attribute of the binary flow control mechanism. The simulation experimental results are done to validate this analysis and conclusion. Finally, a parameter settings scheme is put forward to optimize the existed EFCI switch.

**Key words:** available bit rate traffic; flow control; nonlinear component; describing function

**摘要:** 自适应比特(available bit rate,简称 ABR)流量控制是 ATM 网络中拥塞控制和流量管理的一种有效手段.在大规模的高速网络中,简洁、有效的实现算法对优化交换机的性能是至关重要的.二进制流量控制策略以

\* Supported by the National Natural Science Foundation of China under Grant Nos.60273009, 69972040 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.G1999032707 (国家重点基础研究发展规划(973)); the National High-Tech Research and Development Plan of China under Grant No.2001AA112080 (国家高技术研究发展计划); the NSFC and RGC under Grant No.60218003 (国家自然科学基金与香港研究资助局资助项目)

第一作者简介:任丰原(1970-),男,甘肃临洮人,博士,讲师,主要研究领域为网络拥塞控制和流量控制,测控网络, Sensor network, Network calculus.

其实现上的简洁性吸引了广泛的注意,但标准 EFCI 算法作用下队列和速率的振荡却给性能造成了负面影响,使人们对二进制流量控制机制产生了疑虑,不得不放弃二进制算法的简洁性,转而研究相对复杂但有效的显式速率反馈算法.在已建立模型的基础上,用非线性控制理论中描述函数的分析方法系统地评价和分析了标准 EFCI 算法的性能.得到的结论是:依赖于直觉的启发式非线性控制算法诱发的自激振荡是造成队列和速率振荡的本质原因,而非二进制流量控制机制本身固有的属性,仿真试验的结果证实了理论分析的结论.最后,提出了一种通过修改交换机参数配置策略来优化现已有 EFCI 交换机性能的方案.

关键词: ABR 业务;流量控制;非线性环节;描述函数

中图法分类号: TP393 文献标识码: A

ABR(available bit rate)业务的流量控制是 ATM 网络拥塞控制的一种有效手段,一直是流量管理研究的一个焦点.为了适应高速网络发展的需要,ATM 网络放弃了曾在分组交换网络中广泛使用的滑窗流控机制的衍生算法——基于信誉(credit-based)的流控策略,而采用了基于速率的流量控制方案.虽然信誉方案有诸多技术优势,如零信元丢失、响应性好、对不可预测和突发性业务有更强的适应能力<sup>[1]</sup>,但扩展性和设计实现中的复杂性严重制约了它在高速大规模的 ATM 网络中的应用.速率方案设计上的简洁性和可扩展性是其最终赢得 ATM 论坛支持的关键因素<sup>[2]</sup>.在速率流控机制的实施上又出现了两种不同的具体技术:二进制速率反馈和显式速率(explicit rate,简称 ER)反馈.前者是由 DECbit<sup>[3]</sup>发展而来的,节点交换机仅需完成两个功能:(1) 探测网络拥塞的发生;(2) 向信源提供二进制反馈.最初用的是负反馈,即信源只有收到 CI 置位的 RM 信元才减小速率,而速率增加过程无须 RM 信元控制.负反馈的问题在于,如果由于网络在反向通路上过度拥塞而导致 RM 信元丢失,信源依旧持续增速,则会导致网络过载而拥塞,且很有可能导致崩溃.为了解决这一问题,后将负反馈改成了正反馈,使增速过程由 RM 信元明确控制,减速过程则无须 RM 信元指示,当然也可采用双极性反馈,使增速和减速过程都受 RM 信元的控制.为了解决二进制流量控制中的公平性问题,分别采用“智能标记”<sup>[4]</sup>和加性增加乘性减小(linear increase multiplicative decrease,简称 LIMD)<sup>[5]</sup>等技术,收到了明显效果,但有一个问题使人们对二进制流量控制算法的性能产生了质疑,这就是队列长度和信元速率在稳定状态依旧呈现显著的振荡特性.它严重地影响了网络利用率的提高.于是 MIT 的 Charny 于 1994 年在其论文工作中首次提出了称为 MIT 方案的显式速率控制策略,使交换机依据网络的负载状态公平地为信源分配它能支持的速率.这一设想得到了 R.Jain 和 Clark 等人的支持<sup>[6]</sup>,一时成为 ABR 流量控制研究的主流方向,相继产生了诸如 EPRCA<sup>[7]</sup>,ERICA<sup>[8]</sup>,DERA<sup>[9]</sup>和 Phantom<sup>[10]</sup>等典型的显式速率算法,诚然,较二进制算法而言,显式速率算法的响应性、鲁棒性和稳定性等都有较大改善,但都是以增加交换机的计算处理复杂度为代价的.ER 算法研究的一个重点就是在算法的性能和复杂度之间寻求一种合理的平衡:EPRCA 通过移动加权平均算法降低了求平均允许信元速率(mean allowed cell rate,简称 MACR)的计算复杂度,但性能不胜理想;Phantom 的“假象”虚连接(VC)思想使得在 VC 数较少时链路的利用率不高;ERICA 算法为了提高控制精度对当前信元速率(current cell rate,简称 CCR)进行实测,改进算法 ERICA+为了控制交换机队列长度,引入了两段双曲线作为控制率参考值,额外处理开销是显见的;DERCA 的计算复杂度为  $O(n)$ .在有了众多 ER 算法的背景下,众多商用 ATM 交换机,如 Cisco 的 LightStream1010<sup>[11]</sup>、IBM 的 Nways8265<sup>[12]</sup>和华为的 Radium8750A 等,依旧采用 EFCI 或其改进算法,表明目前高速交换机的设计需要明快、简洁的算法来优化和提高性能.如果能大幅度地抑制二进流控算法控制的 ACR 的振荡幅度,将队列长度限制在一定的范围之内,不会造成过多的信元丢失,那么二进制的 ABR 流量控制算法应该都是比较理想的实现算法.下面我们非线性控制系统的分析方法来揭示 ACR 和队列长度振荡的内在原因,并进行必要的修正,优化算法的性能,使二进制算法的简洁性得到应有的发挥.

## 1 算法分析

在文献[13]中,我们用流体流理论建立了二进制流量控制系统的数学模型,并用仿真试验验证了它的有效性,接着用小信号线性化方法得到了如图 1 所示的近似线性模型.

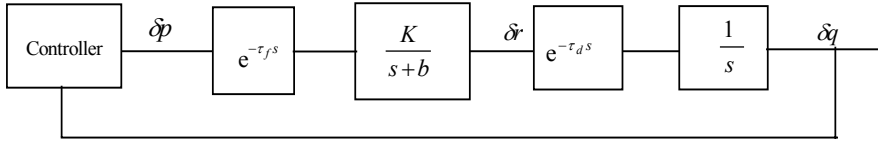


Fig.1 Model of binary ABR flow control

图 1 二进制 ABR 流量控制的模型

图中各参量的定义如下:

$$K = \frac{C}{Nrm \times S \times N^2} (\alpha N + \beta C), \quad b = \frac{\alpha \beta C}{Nrm \times S \times (\alpha N + \beta C)}. \quad (1)$$

C: 拥塞交换机的输出端口容量;

$\tau_d$ : 信源与拥塞交换机之间的广播时延;

$\tau$ : 往返时间;

S: 单信元的比特数,为 424;

N: 网络中激活的连接数;

$\tau_f$ : RM 信元自进入拥塞交换机到返回信源所经历的时间;

Nrm: 每 Nrm 个信元中包含一个 RM 信元,通常取 32;

$\alpha$ : ACR 增加过程中的单位增量,即 RIF  $\times$  PCR(峰值速率);

$\beta$ : ACR 减小过程中的倍乘因子,即 RDF;

$p(t)$ : 拥塞交换机将数据信元信头拥塞标志位置位的概率.

TM4.1 推荐的控制标准 EFCI 算法(图 1 中的控制器)可用下面的方程(2)描述:

$$p(t) = \begin{cases} 0, & q(t) > H \text{ 且 } q(t)' > 0 \\ 1, & q(t) < L \text{ 且 } q(t)' < 0 \end{cases} \quad (2)$$

从控制理论的角度分析,式(1)是一个带滞环的继电器环节,故我们可用如图 2 所示的等效系统来描述二进制的流量控制系统.它是一个典型的带非线性环节的控制系统,于是我们可以用分析该类型系统的经典方法——描述函数法<sup>[14]</sup>,来评价 EFCI 算法的性能.

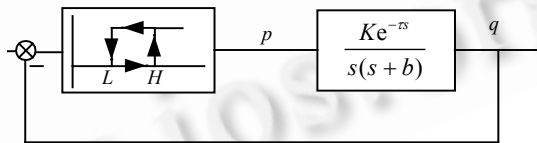


Fig.2 Equivalent system

图 2 等效系统

首先,我们需要知道图 2 中非线性环节的描述函数.对应于图 3,继电器环节的描述函数定义为

$$N(x) = \frac{A_1}{X} + j \frac{B_1}{X}, \quad X \geq H,$$

其中

$$A_1 = \frac{1}{\pi} \int_{\theta_2}^{\pi-\theta_1} y(t) \sin \omega t d\omega = (\sqrt{1-\alpha^2} + \sqrt{1-(\alpha\beta)^2})/\pi,$$

$$B_1 = \frac{1}{\pi} \int_{\theta_2}^{\pi-\theta_1} y(t) \cos \omega t d\omega = \alpha(\beta-1)/\pi,$$

所以,EFCI 算法相应的相对描述函数为

$$N_{EFCI0}(X) = H \cdot N_{EFCI}(X) = \frac{\alpha(\sqrt{1-\alpha^2} + \sqrt{1-(\alpha\beta)^2})}{\pi} + j \frac{\alpha^2(\beta-1)}{\pi}. \quad (3)$$

以如图 4 所示的网络配置为例,为方便起见,假定网络中仅有 1 个激活连接,由式(1),我们得到控制对象的传

递函数:

$$G_0(j\omega) = K_0 G(j\omega) = \frac{K}{HS} G(j\omega) = \frac{1.18 \times 10^5 e^{-j5.95 \times 10^{-3} \omega}}{j\omega(j\omega + 158.61)}$$

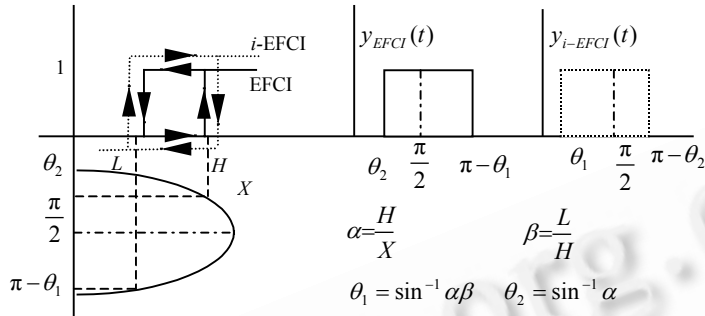


Fig.3 Describing function of nonlinear component

图 3 非线性环节的描述函数

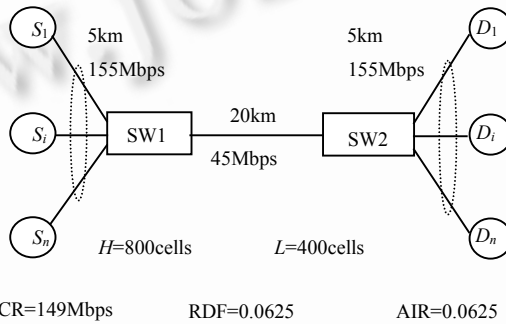


Fig.4 Simulation network configuration

图 4 仿真网络配置

需要说明的一点是,在这里,我们沿用 NIST ATM 仿真平台<sup>[15]</sup>中假定的电磁波在有线介质中的传播速度为  $2 \times 10^8 \text{m/s}$ ,因此,往返时延中固定传播时延为 0.3ms;取上下门限的均值作为队列长度的平均值,那么平均排队时延为

$$\tau_w = \frac{(H + L) \times 424}{2 \times 45 \times 10^6} = 5.65 \text{ms},$$

故

$$\tau = \tau_w + \tau_p = 5.95 \text{ms}.$$

在复平面上作控制对象的幅频特性曲线  $G_0(j\omega)$  和非线性控制器相对的负倒数描述函数曲线  $-1/N_0(X)$ , 如图 5 所示.依据稳定判据和系统产生自激振荡的条件<sup>[14]</sup>可知:系统会在两条曲线的交点处发生周期性自激振荡,振荡幅值  $X_c$  和频率  $\omega_c$  分别由曲线  $G_0(j\omega)$  和  $-1/N_0(X)$  在该点的取值确定.求解方程(4):

$$\frac{1.18 \times 10^5 e^{-j5.95 \times 10^{-3} \omega}}{j\omega(j\omega + 158.61)} = \frac{\pi}{\alpha(\sqrt{1 - \alpha^2} + \sqrt{1 - 0.25\alpha^2 - j0.5\alpha^2})}, \tag{4}$$

得到  $\omega = 191.2, \alpha = 1.215$ .

于是,  $f = \frac{\omega}{2\pi} = 30.42, X = \alpha \times H = 970$ .

为了验证上述分析方法的正确性,我们在 NIST ATM 仿真平台上用相同的配置参数进行仿真试验,并将结果表示在图 6 中,不难看到,队列长度的确在作周期性的振荡运动,幅值和频率分别近似等于 1 100 和 39,这与上面的解析解是大致吻合的,从而也就验证了用非线性控制理论的方法分析 ATM 网络中 ABR 业务的二进制流量控制是可行的.

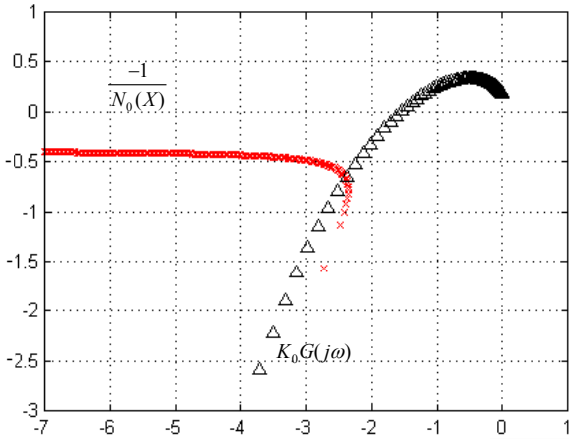


Fig.5  $G_0(j\omega)$  and  $-1/N_0(X)$  curves

图 5  $G_0(j\omega)$  与  $-1/N_0(X)$  曲线

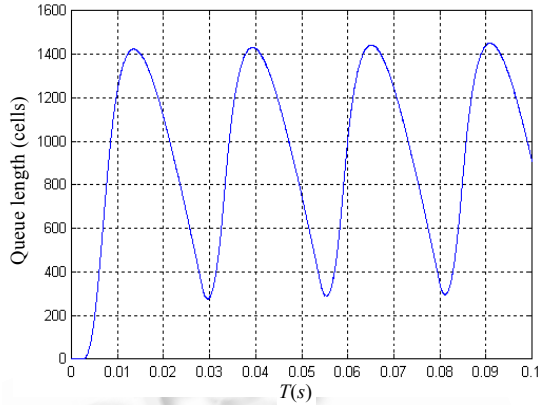


Fig.6 Simulation results ( $N=1$ )

图 6 仿真结果( $N=1$ )

## 2 算法修正

下面我们利用这一分析方法来校正一个一直被人们误解的问题,它是 EFCI 算法设计初期,依赖于直觉的启发式设计方法遗留在算法中的一个瑕疵,正是这一瑕疵使人们对二进制算法的性能和效率大打折扣,最终不得不放弃它在实现上的简洁性,转而研究复杂但有可靠性能保障的显式速率 ER 算法。

由非线性控制理论<sup>[15]</sup>可知:如果曲线  $-1/N_0(X)$  与  $G_0(j\omega)$  没有任何交点,且前者不被后者包围,则系统是稳定的.观察图 5,  $-1/N_0(X)$  的负虚部加大了两条曲线相交的可能性,如果非线性控制器的负倒数描述函数具有正实部,则会大大减小二者相交的可能.回到图 3 中,  $B_1$  为负的原因是  $y_{EFCI}(t)$  在  $\cos(\omega t)$  的负值区间  $(\pi/2, \pi)$  上的取值范围  $(\pi/2 - \theta_1)$  大于正值区间  $(0, \pi/2)$  上的取值范围  $(\pi/2 - \theta_2)$ . 如果我们将非线性环节中原来逆时针旋转的磁滞环用顺时针来替代,那么情况刚好相反,  $B_1$  也会因此而取正值,相应环节的描述函数为

$$N_{i-EFCI0}(X) = H \cdot N_{i-EFCI}(X) = \frac{\alpha(\sqrt{1-\alpha^2} + \sqrt{1-(\alpha\beta)^2})}{\pi} + j \frac{\alpha^2(1-\beta)}{\pi} \quad (5)$$

上面的改变反映在 EFCI 交换机上,只需简单地改变参数的配置,对交换机的硬件实现没有丝毫影响,也就是说,在配置现有的 EFCI 交换机时,将拥塞“高门限”设置得比“低门限”小.为了与标准的 EFCI 算法有所区别,将其称为“逆 EFCI”算法,表示为  $i$ -EFCI.我们将曲线  $-1/N_{i-EFCI0}(X)$ ,  $-1/N_{EFCI0}(X)$  和  $G_0(j\omega)$  绘制在图 7 中,这里假定网络中的激活连接数为 5,很显然,由  $i$ -EFCI 算法控制的 ACR 和队列长度是稳定的,不会产生任何自激振荡,而 EFCI 算法则不然,稳定的自激振荡影响链路利用率的提高是肯定的.为了验证从图 7 分析得到的这一结论,我们用一样的网络配置进行仿真试验,将两种算法作用下允许信元速率和队列长度的变化过程分别描述在图 8~图 12 中,仿真试验数据充分证实了我们关于算法性能的估计与预测。

以上分析表明:标准 EFCI 算法使 ACR 和队列振荡导致的性能衰减并非二进制流量控制机制所固有的缺陷,而是在设计具体算法时过分依赖不完全可靠的直觉造成的.从直观上分析,为了提高链路的利用率,忽略短暂拥塞频繁启动拥塞控制而对性能造成的影响,将用以探测拥塞的“高门限”设置得相对高一些;处于保守性的考虑,将用以解除拥塞的“低门限”设置得相对低一些,似乎是合理的,但正是这种似是而非的启发式算法给发挥 EFCI 算法在实现上的简洁性带来困惑.在算法的设计中,借助一定的理论对算法性能作一个较为透彻的分析是很有必要的。

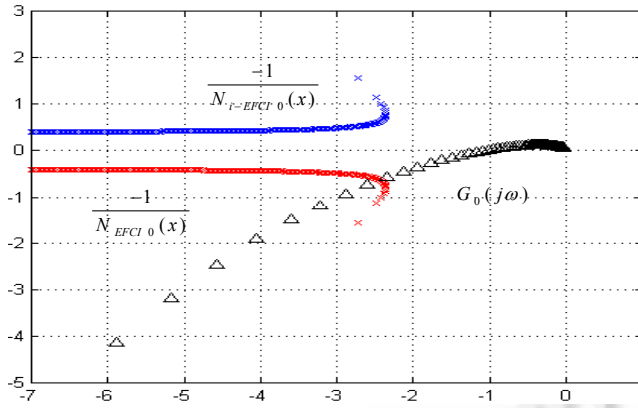


Fig.7  $G_0(j\omega)$ ,  $-1/N_{i-EFCI_0}(X)$  and  $-1/N_{EFCI_0}(X)$  curves  
 图 7  $G_0(j\omega)$ ,  $-1/N_{i-EFCI_0}(X)$  和  $-1/N_{EFCI_0}(X)$  曲线

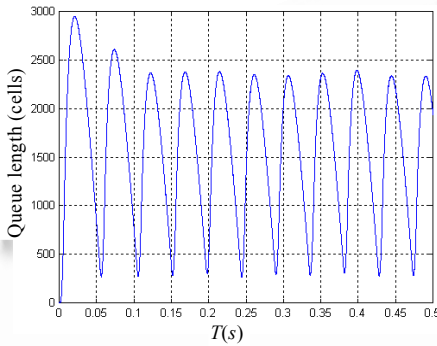


Fig.8  $Q(t)(EFCI)$

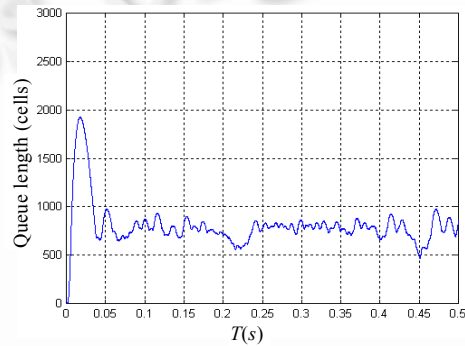


Fig.9  $Q(t)(i-EFCI)$

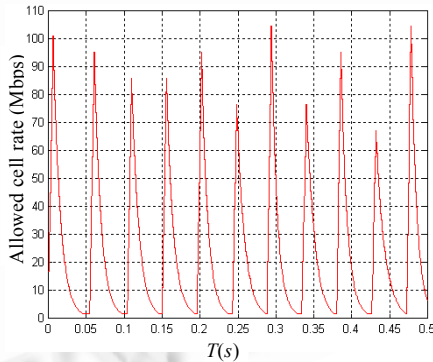


Fig.10  $ACR(t)(EFCI)$

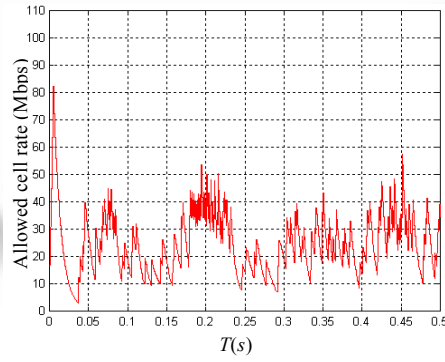


Fig.11  $ACR(t)(i-EFCI)$

### 3 结 论

ABR 的流量控制是 ATM 网络中平衡统计复用增益和服务质量保证这一相互对立的技术目标的最有力的手段。大规模的高速网络需要扩展性强、性能可靠、实现简单方便、处理开销小的流量控制算法来支持。在基于速率的流控策略中,二进制反馈因其实现上的简洁性而普遍受到交换机生产厂商的青睐。进一步的研究表明:标准 EFCI 算法控制下的信元速率和交换机队列存在大幅振荡的弊端,严重影响链路利用率的提高。对二进制流控算法有效性的质疑导致了显式速率反馈机制的产生与发展,但随之引入的计算复杂性又使人们割舍不下二进制反馈算法的简洁性。在本文中,我们利用非线性控制理论的分析方法深入评价和分析了标准的 EFCI 算法,

认为依赖于直觉的不恰当的启发式非线性控制算法诱发的自激振荡是造成队列和速率振荡的本质原因,而非二进制流量控制机制固有的属性.仿真试验结果与理论分析结果相吻合的事实表明了我们所采用的分析方法的可行性.最后,在理论分析的基础上,提出了与标准算法正好相反的队列门限参数的配置建议,从仿真结果来看,收到的效果是显著的.

#### References:

- [1] Kung HT, Robert M. Credit-Based flow control for ATM networks. *IEEE Networks Magazine*, 1995,9(2):40~48.
- [2] Kalampoukas L. Congestion management in high speed networks [Ph.D. Thesis]. University of California Santa Cruz, 1997.
- [3] Raj J. Myths about congestion management in high-speed networks. *Internetworking: Research and Experience*, 1992,3(2): 101~113.
- [4] Hluchyi M. Closed-Loop rate-based traffic management. *ATM Forum Contribution 94-0211R3*, 1994.
- [5] Kim T, Lu S, Bharghavan V. Improving congestion control performance through loss differentiation. In: *Proceedings of the IEEE International Conference on Computers and Communication (INFOCOM)*. New York: IEEE Computer Society, 1999. 140~145.
- [6] Charny A, David DC, Raj J. Congestion control with explicit rate indication. In: *Proceedings of the IEEE International Communication Conference (ICC)*. Tokyo: IEEE Computer Society, 1996. 1954~1963.
- [7] Hiroyuki O, Masayuki M. Rate-Based congestion control for ATM networks. *ACM Computer Communication Review*, 1995,25(2): 60~72.
- [8] Kalyanaraman S, Jain R, Fahmy S, Goyal R, Vandalore B. The ERICA switch algorithm for ABR traffic management in ATM networks. Part I : Description. *IEEE/ACM Transactions on Networking*, 2000,8(1):87~98.
- [9] Charny A, Ramakrishnan K, Lauck K. Time scale analysis scalability issues for explicit rate allocation in ATM networks. *IEEE/ACM Transactions on Networking*, 1996,4(4):569~581.
- [10] Afek Y, Mansour Y, Ostfeld Z. Phantom: a simple and effective flow control scheme. *Computer Networks*, 2000,32(3):277~305.
- [11] Lin A. LightStream 1010 switch architecture and traffic management. 1998. <http://www.cisco.com/warp/public/cc/so/neso/vvda>.
- [12] Haissam A. IBM 8265 ATM backbone switch hardware architecture. *ELSEVIER Computer Network*, 1999,31(6):527~539.
- [13] Ren FY, Ren Y, Shan XM. Fluid-Based analysis of ABR flow control in ATM networks. In: *Proceedings of the Asia-Pacific Conference on Communications*. 2001.
- [14] Hu SS. *Automatic Control Principle*. Beijing: Military Industry Press, 1984 (in Chinese).
- [15] Nada G. The NIST ATM/HFC network simulator operation and programming guide version4.0. 1999. <http://hsnt.nist.gov/>.

#### 附中文参考文献:

- [14] 胡松寿. *自动控制原理*. 北京:国防工业出版社,1984.