计算机研究与发展
Journal of Computer Research and Development
ISSN 1000-1239/CN 11-1777/TP
48(1): 110-117, 2011

适用于流媒体传输的动态门限调度算法

汪学舜[1,2] 余少华[2] 徐 宁[1,2] 戴锦友[2]
[1](华中科技大学计算机科学与技术学院 武汉 430074)
[2](武汉邮电科学研究院 武汉 430074)
(wang_xueshun@163.com)

A Dynamic Threshold Schedule Algorithm for Media Traffic

Wang Xueshun[1,2], Yu Shaohua[2], Xu Ning[1,2], and Dai Jinyou[2]
[1](*School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan* 430074)
[2](*Wuhan Research Institute of Post and Telecommunication, Wuhan* 430074)

Abstract Real-time multimedia applications are getting more and more popular in the last several years with internet capability keeping improved, in order to guarantee the transmission quality of medium traffic, effective strategies for buffer management are applied in exchanging devices in flow medium network, such as switches and routers. Considering the characteristics of packet delay and loss sensitivity of medium data, this paper proposes a novel dynamic threshold algorithm at media traffic level for shared buffer packet switches, known as dual-dynamic threshold control based on E-model(EDTA). EDTA divides buffer management of switch into two parts: the global threshold strategy and the queue threshold strategy. The global threshold strategy controls the shared buffer by a whole threshold through judging different traffic scenarios to ensure good transmission performance in all situations. The queue threshold strategy decides and adjusts output queue threshold based on E-model traffic levels, which transforms several medium service quality effecting parameters into users' psychological factors, thus guarantees the quality of flow medium data. Experimental results show that our proposed algorithm (EDTA) can improve the traffic performance for flow media more effectively, achieve better quality of network flow medium transmission, comparing with those algorithms using traditional control. And it is suitable to be used in any network condition and can improve the utilization of the network resources.

Key words buffer management; dual-dynamic threshold; traffic level, fairness; media traffic

摘 要 为保证网络流媒体传输质量,在流媒体传输网络中需要对交换设备实施有效的缓存管理策略.结合流媒体数据对时延和丢包率敏感的特点,提出了一种基于E-Model传输级别的动态门限控制算法(EDTA),将交换设备的缓存管理分为全局门限控制和队列门限控制两部分.全局门限控制通过判断不同的传输场景进行门限控制,保证不同传输场景下的传输性能.队列门限控制则基于E-Model传输级别参数进行门限的划分和调整,保证流媒体的传输质量.实验结果表明该算法能有效改善流媒体的传输性能,对交换设备的缓存资源进行有效控制,并提高网络流媒体传输质量.

关键词 缓存管理;双动态门限;传输级别;公平性;媒体传输

中图法分类号 TP393

收稿日期:2009-10-09;**修回日期**:2010-05-20

基金项目:国家"八六三"高技术研究发展计划基金项目(2005AA121410)

随着 Internet 能力的增长,实时多媒体应用在过去几年越来越多地得到应用,但实时多媒体应用需要引入业务保证. 交换设备调度和缓存管理是两种主要的方法用来实现业务保证,调度程序控制传输资源,缓存管理主要对存储资源进行控制. 对于实时流媒体的服务质量缓存管理尤为重要[1].

交换设备中数据传输通常以令牌桶的形式处理,如文献[2]详细说明了固定传输速率,同时允许短期的速率波动,对超过速率规定的包,通过丢弃、整形或者标记进行处理. 但流媒体传输应用中,由于传输难以预测,可能出现传输速率在一段时间内超过规定速率,或者应用层爆发发送数据包并超过了规定爆发速率的情况,丢弃一些数据可能导致传输质量的下降,对流进行整形可能导致附加的时延,降低了传输质量. 因此对交换设备需要开发一种缓存管理机制,以最小的代价实现以下目标:

1) 对实时传输,满足用户的服务质量要求;

2) 对过量传输,服务质量要求范围内尽量传输;

3) 对各种网络应用场合,满足服务质量要求;

4) 对共享缓存资源,保证资源共享的公平性.

一些研究者对以上目标进行了研究,文献[3]提出的 ABE 允许实时应用优先对时延进行处理,不考虑传输规定的要求. 文献[4]基于共享存储交换结构和组播交换队列设计方案,提出一种与组播比例和平均扇出相关的缓存管理机制和队列调度算法. 但该算法并不能完全适应流媒体传输需要同时对时延和丢包率进行保证的情况. 文献[5]提出了一种基于门限选择丢弃的缓存管理机制,但该方法配置较复杂,只适用于某些特定场合. 文献[6]提出了一种根据目的地址进行分类的缓存管理策略,但该方案需要对每一个流进行检查,需要相应的硬件查找机制,实现较为复杂,且重载条件下适应性较差. 文献[7]结合进化计算和模糊控制提出一种进化管理机制,但进化算法本身较为复杂,不能适应网络中数据传输的多变性和复杂性,同时进化算法很难适应网络传输的快速变化.

1 网络交换模型

在交换设备中,输入端口收到数据后经过调度处理,进入输出端口的缓存队列,等待转发. 为了提高资源利用率,输出端口的缓存队列采用共享机制. 图 1 给出了分组交换机数据包的处理流程. 全局丢包模块在全局缓冲区占用率达到全局阈值时进行全局性的丢包;进入共享队列之前数据流首先经过一个令牌桶整形模块,对突发性流量进行整形,然后是输出队列共享缓冲区中局部丢包模块将其分为多个输出端口. 全局缓冲区和队列等缓存的管理机制与设备转发数据性能密切相关,如何设计缓存管理机制是一个极为重要的问题.

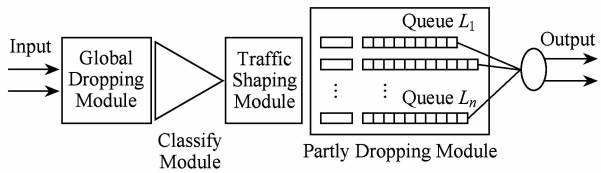

Fig. 1 Switch traffic model.

图 1 交换设备数据传输模型

缓存管理一般采用动态门限管理机制[8]. 动态门限管理机制包括两种设置和调整:第 1 种机制是全局门限设置和调整,称为全局门限设置,其基本机制是对交换设备已分配的缓存进行控制,以保证突发流量,全局门限设置需要控制各输出端口队列的总体门限;第 2 种机制是对各输出端口缓存队列门限的设置和调整,称为队列门限设置,其基本机制是对各输出端口队列长度门限进行设置. 在队列长度达到下限门限值时开始随机丢包,在队列长度达到上限门限值时,丢弃所有转发到该端口的数据包,以保证其他输出端口队列的公平性.

研究者们大多只对其中一个进行设置,实际上这两种机制都是必不可少的,如果只进行全局门限设置,容易导致某一个输出端口大量抢占缓存,设备的公平性无法保证. 如果只进行队列门限设置则不能适应网络场景的变化,对短时间内的爆发数据流无法及时响应,从而影响设备的吞吐量.

本文所提出的动态门限调度机制是一种基于 E-Model 传输级别的双动态门限控制算法,称之为 EDTA 算法,对以上两个方面同时进行调整和设置. 该算法全局门限控制通过判断不同的传输场景进行门限控制,队列门限控制则基于 E-Model 传输级别参数进行门限的划分和调整,保证流媒体的传输质量.

2 实现机制

EDTA 算法动态门限的调整包括两部分:全局门限设置和队列门限设置,以下分别说明.

2.1 全局门限设置

在全局门限设置中,门限值 $T(t)$ 为各输出端口队列分配缓存之和的上限,按照 $T(t)=k[B-Q(t)]$

进行变化，式中 B 为缓存大小，$Q(t)$ 为在时刻 t 已经分配的缓存，系数 k 表示缓存共享的程度.

在全局门限设置的研究中，如果将动态门限值与相应的 QoS 参数进行结合，系数 k 取为固定值，实际上，这种动态门限的调整只能适用于特定的场合，并不能满足实际网络传输. 文献[9]研究了动态门限系数的适应情况，根据文献[9]的分析，丢包率对参数 k 的变化较为敏感，固定 k 值很难同时使得一般传输和热点流传输都保持较高的性能. 而且热点流场景和一般传输场景优化的 k 值不存在交集.

因此，对于全局门限设置，单纯调整门限值，交换设备不能适应网络中传输场景变化，其健壮性较差. EDTA 算法除了调整门限值外，同时根据不同的场景，对参数 k 也进行调整.

为了实时检测传输场景，需要根据已分配的缓存空间 $Q(t)$ 和输出端口队列长度 $Q_i(t)$（i 表示输出端口，下同）来区分应用的场景. 用 $\delta_i=Q(t)/Q_i(t)$ 表示输出端口 i 的队列占用情况，根据文献[7]，如果存在 $\delta_i<3$，则表明当前存在热点流状况，调整 k 值为相应热点流情形，按热点流情形设置队列门限. 否则表明数据流分布比较平均，按照一般传输的 k 值设置相应的门限. 根据网络实际运行情况，一般一种场景的传输会持续数秒时间，因此实际运行中 k 值不会频繁抖动.

根据文献[7]的实验及相关曲线，对热点流，全局门限设置的最优 k 值为[0.5，2]之间，针对交换设备实际应用情况，EDTA 算法取为 1，对于一般传输，无论重载与轻载，k 的最优值大于 4，而且与端口数量无关，EDTA 算法取为 8.

为了保证最大最小公平带宽分配，在所有输出端口队列分配缓存之和超过全局门限时，控制所有输出端口缓存大小保持一致. 所有输出端口队列可获取缓存大小称之为公平共享缓存大小，用 $q(t)$ 表示时刻 t 的公平共享缓存大小. 为了能实现最大最小公平缓存分配，每一个出端口获取的缓存大小为 $\min\{Q_i(t),q(t)\}$.

如果 $Q(t)<B-T/k$，表明全局缓存不存在拥塞，不需要进行丢包处理，公平共享的缓存大小为 $q(t)=\max(Q_i(t))$. 如果 $Q(t)\geqslant B-T/k$，则表示全局缓存出现拥塞. $q(t)$ 是式(1)的唯一解：

$$\sum_{i=1}^{N}\min\{Q_i(t),q(t)\}=B-T/K, \qquad (1)$$

式(1)中，T 是全局缓存门限值，B 是全部缓存大小，N 是输出端口数量.

对于输出端口 i，如果 $Q_i(t)<q(t)$，即输出端口 i 队列长度小于交换机的公平共享缓存大小，转发到该输出端口 i 的所有包都不会丢弃. 如果 $Q_i(t)\geqslant q(t)$，则该端口收到的所有包都将被丢弃，因此其输出队列长度不可能超过 $q(t)$.

2.2 队列门限设置

队列门限设置按照 E-Model 的传输级别要求，对队列门限进行设置和调整. 对实时流媒体来说，用户的满意程度主要体现在数据包的时延和丢包率两个方面. 本节首先对时延和丢包率与队列门限之间的关系进行说明，进而根据流媒体传输，对 E-Model 模型传输级别进行简化，推导出队列门限值与传输级别之间的关系，以便根据传输级别调整队列门限.

假设输出端口队列丢弃门限为 K_i，开始按概率进行丢包的门限为 $K_i/2$. 这样设置是因为如果开始丢包门限过于接近全部丢弃门限，重载时容易造成门限变化频繁，相反，如果开始丢包门限过小，则对轻载传输不利. 因此 EDTA 算法根据实际应用情况，选取缓存占用一半为界限，超过一半为重载，否则为轻载.

因此 EDTA 算法对于各输出端口队列门限的丢包机制如下：如果当前队列长度 $Length[i]$ 小于该队列门限值 K_i 的一半，则不丢包；如果当前队列长度大于该队列的门限值，即 $Length[i]>K_i$，则直接丢弃；如果 $K_i/2<Length[i]<K_i$，则按概率 p_i 进行丢包，概率 p_i 根据当前队列长度 $Length[i]$ 动态变化，取为 $(2\times Length[i]-K_i)/K_i$. 实验表明这种设置是合理的.

2.2.1 门限与时延、丢包率关系

在交换设备中，假定分组按泊松分布到达，各输出端口分组到达率为 λ_i（$1\leqslant i\leqslant N$，N 为端口数量）. 尽管网络流量的到达并不完全遵循泊松分布，但类似的研究表明，利用泊松模型与用更复杂的模型得到的结果是比较符合的.

假设网络节点对到达分组的服务时间服从参数为 μ 的指数分布，对各输出端口进入不同的队列，队列长度门限分别是 K_i. 假设各输出端口队列开始丢包长度为队列门限长度的一半，当数据包到来时队列中等待的包数目在开始丢包长度和队列门限长度之间时，随机丢弃的概率分别是 p_i. 令 $\{N_i(t):t\geqslant 0,1\leqslant i\leqslant N\}$ 表示时刻 t 各输出端口队列中的数据包数（包括正在发送的数据包）. 容易得到，$N_i(t)$ 是齐次马氏链，并且是状态为 $(0,1,2,\cdots,K_i+1)$ 的 M/M/1/K 生灭过程.

分析输出端口 i 的分组数据. 令 $\pi_j^i=\lim_{t\to\infty} p\{\sum N_i(t)=j\}$,表示队列 i 中有 j 个分组的平稳概率,$j\in(0,1,2,\cdots,K_i+1)$. 分组到达队列缓冲区的速率分别为:

1) 如果 $N_i(t)\leqslant K_i/2$,则 $\lambda_1^i=\lambda_i$;

2) 如果 $K_i/2\leqslant N_i(t)\leqslant K_i$,则 $\lambda_2^i=(1-p_i)\lambda_i$;

3) 如果 $N_i(t)=K_i+1$,则 $\lambda_3^i=0$.

令 $\rho_1^i=\lambda_1^i/\mu$,$\rho_2^i=\lambda_2^i/\mu$,根据排队理论和上面的分析,可以得到输出端口 i 分组的丢包率函数为

$$L_i=f(K_i)=\pi_0^i(\rho_1^i)^{1+K_i/2}(\rho_2^i)^{K_i/2}, \tag{2}$$

其中,
$$\pi_0^i=\frac{1}{1+\sum_{m=1}^{1+K_i/2}(\rho_1^i)^m+\sum_{m=1}^{K_i/2}(\rho_2^i)^m}. \tag{3}$$

输出端口 i 分组的时延函数为

$$D_i=\frac{\pi_0^i}{\mu}\Big[\sum_{m=0}^{K_i/2}(m+1)(\rho_1^i)^m+(\rho_1^i)^{K_i/2}\sum_{m=1}^{K_i/2}\Big(\frac{K_i}{2}+m+1\Big)(\rho_2^i)^m\Big]. \tag{4}$$

由式(2)~(4)可以看出,不同输出端口分组的丢包率和时延都是相应的门限值参数的函数,在确定了丢弃概率的情况下,仅仅通过改变门限值参数 K_i 就能够定量地改变不同出端口的丢包率和时延. 这提供了控制分组丢包率和时延的一种手段.

2.2.2 E-Model 模型

为了进一步确定门限,需要根据丢包率和时延的合理优化来确定门限. E-Model 为分组网传送语音的测量方法,适用于对数据网络中语音质量的评估,关注数据全面的网络损伤因素,对于传输设计具有高重要度的参数组合情况,诸如延时、抖动和丢包等数据网络特有的问题,可以被确信应用. 下面根据 E-Model 模型,说明流媒体传输中时延和丢包率的确定方法.

在分组网络上传送流媒体的情况也可以根据 E-Model 模型,将影响服务质量多种传输参数转换成用户的心理因素,并量化为传输级别参数 R. 由于 E-Model 与语音的特点以及语音传送的诸多技术密切相关,直接将模型应用于一般实时信息的传输并不合适,但可以将 E-Model 的计算公式作一些变换,以适应其他流媒体的应用场合.

在传输流媒体业务时,用传送级别参数 R 描述网络传输的服务质量等级,假定其他损伤因子取缺省值的情况下,重点研究时延和丢包率与 R 的关系,这就是 EDTA 算法研究的出发点. EDTA 将语音研究中 R 的计算公式[10]简化为式(5). 在式(5)中,α_1 是不考虑时延损伤因子 I_d 和丢包率损伤因子 I_e 两个因素的传送级别参数 R 的值,β_1 和 γ_1 则是 I_d 和 I_e 的权重.

$$R=\alpha_1-\beta_1\times I_d-\gamma_1\times I_e. \tag{5}$$

通常情况下,时延损伤因子 I_d 和丢包率损伤因子 I_e 的具体计算公式较为复杂,文献[11]给出了时延损伤因子 I_d 的简化计算公式,为适合交换设备的需要,时延损伤因子 I_d 计算公式进行修正为式(6),这样,时延损伤因子 I_d 与时延 D 等价于分段线性关系.

$$I_d=0.144D+0.65(D-30)\times H(D-30), \tag{6}$$

其中,$H(x)=\begin{cases}0, & x\leqslant 0,\\ 1, & x>0.\end{cases}$

文献[11]给出了几种评价标准下的丢包率损伤因子 I_e 计算公式,通过丢包率 L 计算丢包率损伤因子 I_e 的计算方法总结为式(7). 在式(7)中,a,b 和 c 都是常数,在不同的应用场景取不同的值:

$$I_e=a+b\times\ln(1+c\times L). \tag{7}$$

式(7)中,当丢包率很小时用式(8)近似计算丢包率损伤因子 I_e 是可行的:

$$I_e\approx a+b\times c\times L=a+u\times L. \tag{8}$$

式(5)传输级别参数 R 可以用式(9)来计算:

$$R=\alpha-\beta\times[0.144D+0.65(D-30)\times H(D-30)]-\gamma\times L, \tag{9}$$

式(9)中,α,β,γ 均为常数,结合式(2)~(4)和式(9),可得到传送等级值 R 与队列门限值 K_i 之间的关系. 因此根据确定的传送等级值 R 可以获得队列门限值 K_i,根据获得的 K_i 值就可实施对队列门限大小的调整. 通过这些计算和判断可有效实施对队列门限的自适应管理.

2.3 算法实现

基于以上的分析可知,EDTA 算法通过一张队列长度历史记录表 $Length[i]$ 记录一定时间长度内输出端口 i 的平均队列长度,用 $total$ 表示一定时间长度内全局缓存占用长度,交换设备端口数为 N,输出端口队列全部丢弃门限为 K_i,开始按概率进行丢包的门限为 $K_i/2$. 交换设备收到数据帧的处理步骤如下:

步骤 1. 初始时,各端口按最小门限长度 $K_{\min}$ 分配缓冲区,全局门限值按轻载时进行设置.

步骤 2. 交换设备收到数据帧,将全局缓存占用长度 $total$ 加 1,相应出端口队列长度 $Length[i]$ 加 1.

步骤 3. 全局门限判断:

如果 $total<B-T/k$,表明全局缓存不存在拥塞,因此不需要进行丢包处理.

如果 $total \geqslant B-T/k$，则表示全局缓存出现拥塞，如果 $Length[i] > total/N$，则需丢弃该数据包，否则继续.

步骤 4. 队列门限判断：

如果 $Length[i] < K_i/2$，继续转发；

如果 $K_i/2 < Length[i] < K_i$，则按概率 p_i 进行丢包，概率 p_i 按照当前队列长度 $Length[i]$ 动态变化，按 $(2\times Length[i]-K_i)/K_i$ 计算.

步骤 5. 门限值调整：

当收到数据包个数 $total$ 达到 n 时，计算每一个端口的缓存占用比 $\delta_i = total/Length[i]$，如果存在 $\delta_i < 3$，则全局门限值 T 按存在热点流方式进行设置，否则全局门限值按一般传输进行计算.

当收到数据包个数 $total$ 达到 n 时，对每一个输出端口队列，首先计算这段时间该端口队列缓冲区的平均占有率 δ. 若 δ 小于等于平均占有率门限值下限 $\delta_{\min}$，线性减少该队列的门限值(即该队列的 K_i 值)，开始丢包长度值也作相应地减少. 对每一个输出端口，按照式(2)～(4)和式(9)，计算当前输出端口队列长度对应的平均传输级别 R_i 值，若平均传输级别 R_i 值能满足用户的要求，新的队列门限值有效，否则维持不变. 为了满足公平性的需要，队列门限值减少到下限值 $K_{\min}$ 后，即使短时间内 δ 继续低于 $\delta_{\min}$ 也不应再减低. 减少的资源可以由其他端口共享，以改善交换设备的轻载传输性能. 若 δ 大于平均占有率门限值上限 $\delta_{\max}$，线性增加该出端口队列的门限值，但不能超过全局门限设置的公平共享缓存大小.

步骤 6. 重复步骤 5 至所有的输出端口分析完毕. 清空队列长度历史表 $Length$ 中的记录，重复步骤 2 至传输结束.

3 传输分析

本节对 EDTA 算法的稳定性进行分析，假设在时刻 0($t=0$)有 $N-S$ 个输出端口处于重载，重载输出端口的队列长度等于它们的公平缓存分配长度，S 为轻载输出端口数量，为了分析方便，假定轻载输出端口的队列长度为队列门限下限值 $K_{\min}$，根据网络实际运行情况，一般一种场景的传输会持续数秒时间，因此可以假定 k 值保持为常量. 另外，假定全部缓存大小减去各端口最小分配 $K_{\min}$ 后，剩余缓存大小为 B，$Q(t)$ 表示端口缓存变化，$q(t)$ 表示时刻 t 公平缓存大小，当时刻 $t=0$ 时，M 个轻载输出端口开始获取缓存空间，根据式(1)可以得到：

$$q(0) = (B-T/k)/(N-S); \tag{10}$$

$$Q_i(0) = \begin{cases} \dfrac{B-T/k}{N-S}, & 1 \leqslant i \leqslant N-S; \\ 0, & N-S < i \leqslant N. \end{cases} \tag{11}$$

对于 $t>0$，可以得到以下微分方程：

$$\dot{q}(t) = -\left(\sum_{i=1}^{M} \dot{Q}_i(t)\right)/(N-S). \tag{12}$$

假定 $N-S+M$ 个输出端口中每一个端口以固定比率 $r(r>1)$ 接收数据包，其中 r 为输入数据包的速率与输出能力之比. 很明显，队列长度的变化比率范围为 $[-1, r-1]$. 根据负载 r 的值，有两种情况：$r \leqslant 1+(N-S)/M$，或者 $r > 1+(N-S)/M$.

首先分析 $r \leqslant 1+(N-S)/M$ 情况：对于前期 $N-S$ 个重载输出端口，其公平共享的缓存大小按式(13)递减：

$$\dot{q}(0^+) = -\frac{M\times(r-1)}{N-S}. \tag{13}$$

缓存队列长度的变化如下：

$$\dot{Q}_i(0^+) = \begin{cases} -\dfrac{M\times(r-1)}{N-S}, & 1 \leqslant i \leqslant N-S; \\ r-1, & N-S < i \leqslant N-S+M. \end{cases} \tag{14}$$

假定时刻 t_1 有 M 个端口由轻载状态变为重载状态，对于 $t \leqslant t_1$，队列长度和公平共享缓存大小可由式(15)计算：

$$q(t) = \frac{B-T/k}{N-S} - \frac{M(r-1)}{N-S}\times t, \quad t \leqslant t_1; \tag{15}$$

$$Q_i(t) = \begin{cases} \dfrac{B-T/k}{N-S} - \dfrac{M(r-1)}{N-S}\times t, & 1 \leqslant i \leqslant N-S, \\ \quad t \leqslant t_1; \\ (r-1)\times t, & N-S < i \leqslant N-S+M, \ t \leqslant t_1. \end{cases} \tag{16}$$

根据式(15)和式(16)，时刻 t_1 可由式(17)计算：

$$t_1 = \frac{B-T/k}{(N-S+M)\times(r-1)}. \tag{17}$$

对于时刻 $t=t_1$，满足式(18)：

$$Q_i(t_1) = q(t_1) = \frac{B-T/k}{N-S+M}; \quad 1 \leqslant i \leqslant N-S+M. \tag{18}$$

对于 $r > 1+(N-S)/M$ 的情况，类似的分析可以得到式(18)的结果. 为了节省篇幅不再详细说明.

因此对于时刻 $t=t_1$，共有 $N-S+M$ 个输出端

口为重载,其输出端口缓存与该状态下的公平缓存大小一致,即为稳定状态分配. 对于 $t>t_1$,每一个重载端口丢弃超过接收速率 $r-1$ 的包,并保持队列长度为常量. 所以 EDTA 的缓存分配是一种公平稳定状态的分配.

4 实验与分析

在本节用图 2 所示的拓扑图,对基于 E-Model 传输级别的动态门限控制算法(EDTA)的流媒体业务传输等级和质量进行实验. EDTA 算法与另外两种常用调度机制 RED 和 DropTail 部署在中间路由器上. S 为发送端,D 为接收端,S 和 D 由不同 PC 机和网络测试仪表端口组成. 网络中的发送和接收链路带宽为 100 Mbps. 实验通过不同端口发送不同的 TCP 流或者 UDP 流,测试接收数据流的丢包率和时延来计算传输等级. TCP 流的源端能根据接收速率自动调整发送速率,UDP 流速率固定. 所有链接的传播延时为 1 ms,TCP 最大窗口尺寸无上限,使其不会成为限制 TCP 流吞吐量的因素,TCP 的数据流由 FTP 传输大文件生成,UDP 的数据流采用 MPEG-2 的视频或音频. 所有包大小均为 1 KB.

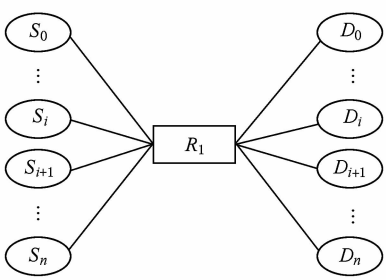

Fig. 2 Network topology.

图 2 实验拓扑图

实验中数据帧的缓存为 1 MB;UDP 的发送比特率为 60 Mbps. 测试时间为 3 min. 传输级别的计算按式(9)计算,根据文献[11],参数 α,β,γ 分别为 93.69,7.81 和 59.26. RED 算法中的参数设置为最大门限 256 KB,最小为 64 KB.

在传输中无热点流传输的情况下,图 3 给出了 3 种算法的传输级别平均值的比较. 从图 3 可以看出,EDTA 算法的传输级别值要明显高于传统的 RED 和 Drop Tail 算法. 从图 3 中还可以看出 RED 和 Drop Tail 算法有周期性的曲线趋势,这种周期性的趋势主要是因为 TCP 协议本身的流量控制机制所导致的,而 EDTA 算法呈水平趋势,说明了动态调节队列长度的效果. 实验结果表明本文提出的 EDTA 算法相比于 Drop Tail 和 RED 算法可以达到更高的传输级别.

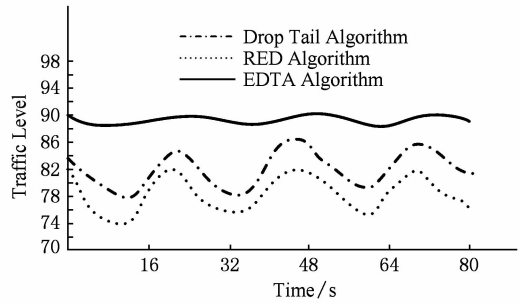

Fig. 3 Comparison of average traffic levels without hot-spot.

图 3 无热点流平均传输等级比较

在有热点流的传输环境下,Drop Tail,RED 的传输级别上下波动较大,而且每次测试结果都不一致. 而 EDTA 算法同样呈水平趋势. 因此即使存在热点流的情况下,EDTA 算法相比于 Drop Tail 和 RED 算法可以达到更高的传输级别.

为了进一步测试 EDTA 算法对视频流传输质量的改进,对评价视频传输质量广泛采用的峰值信噪比 PSNR(peak signal to noise ratio)进行测试,测试拓扑与图 2 相同,建立 4 个 TCP 流和 1 个 UDP 流,在视频编码中,周期地插入 I 帧,由 I 帧和 P 帧组成帧组的大小为 60,在验证时,只对帧组中的 P 帧进行分析. 在不同的视频帧发送速率下,随机抽取 100 帧,计算平均的 PSNR 值. 实验结果如图 4 所示,从图 4 可以看出,当视频流发送速率低于平均带宽时,EDTA 算法的 PSNR 值比 Drop Tail 和 RED 算法至少高 0.6 dB 左右,这是由于 EDTA 尽量保护 UDP 流,其他算法存在出队列缓存溢出情况,只有

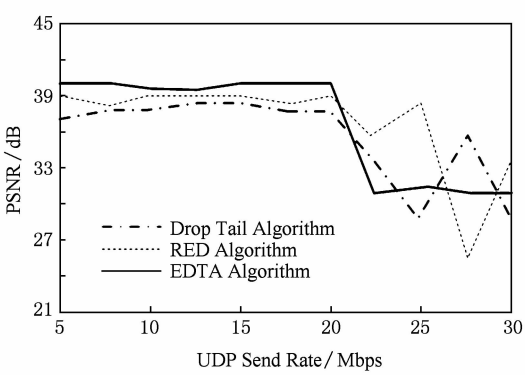

Fig. 4 Comparison of PSNR for different algorithm.

图 4 不同算法的 PSNR 比较

等待 TCP 流速度降下后才能保证视频流带宽. 当视频发送速率高于平均带宽时，Drop Tail 和 RED 算法的 PSNR 值由于跳帧而出现剧烈变化，造成图像间质量的突变，而 EDTA 算法的 PSNR 值较为稳定，较大地提高了图像的质量. 这是由于 EDTA 在传输时充分考虑流媒体数据的各项数据特征进行相应的处理.

为了测试 EDTA 算法在不同负荷条件下的吞吐量和公平性性能，按图 2 建立 9 个 TCP 流和 1 个 UDP 流，实验目的为不同 UDP 发送速率下 TCP 流可以获取的带宽. 实验结果如图 5 所示，图 5 显示了不同 UDP 流负载下 TCP 流的转发吞吐量. 从图 5 可以看出随着 UDP 流速率的增加，UDP 流速率超过平均带宽，其他 TCP 流带宽下降，这是由于 UDP 流抢占了缓存资源，致使其他流丢包率增加. 在 EDTA 算法中，随着 UDP 流速的增加 TCP 带宽下降并不明显，而其他算法则急剧下降. 这说明其他算法完全不能保证各输出端口的公平性，而 EDTA 则能保证各端口之间的公平分配.

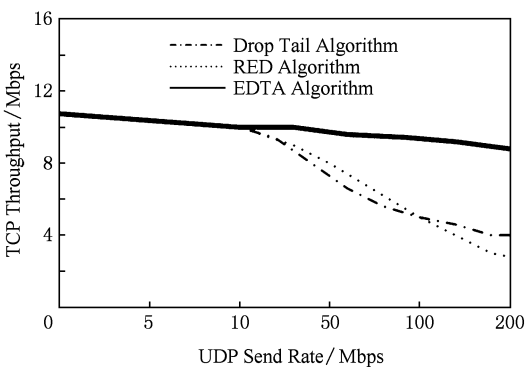

Fig. 5 TCP throughput comparison under different UDP load.

图 5 不同 UDP 流负载下 TCP 流性能比较

通过以上实验可以看出，EDTA 控制算法无论对轻载、重载或者热点流传输，其性能都是稳定的，并且能保证较高的传输等级，较好地保证流媒体传输质量和公平性，因此适合流媒体的传输.

5 结束语

本文结合流媒体数据对时延和丢包率敏感的特点，根据网络交换设备数据转发的流程，提出了一种基于 E-Model 传输级别的动态门限控制算法，该算法依据网络传输环境，将交换设备的缓存管理分为全局门限控制和队列门限控制两部分. 全局门限控制通过判断不同的传输场景，对门限值进行不同的预留，以适应不同环境下的爆发流量. 而队列门限控制则基于 E-Model 传输级别参数，进行门限的划分和调整，保证流媒体的传输质量. 实验结果表明，该算法能有效改善流媒体的传输性能，并提高流媒体传输质量.

参 考 文 献

[1] Guerin R, Kamat S, Peris V, et al. Scalable QoS provision through buffer management [C] //Proc ACM SIGCOMM 1998. New York: ACM, 1998: 29-40

[2] Heinanen J, Guerin R. A Two Rate Three Color Marker [S]. New York: Institute of Electrical and Electronics Engineers, 1999

[3] Hurley P, Kara M, et al. ABE: Providing a low-delay service within best effort [J]. IEEE Network Mag, 2001, 15 (3): 60-69

[4] Song Meina, Song Junde, Zhan Xiaosu, et al. An efficient queuing and switching scheme for multicast in high-speed routers [J]. Journal of Computer Research and Development, 2005, 42(5): 810-817 (in Chinese)
(宋美娜，宋俊德，战晓苏，等. 高速路由器中一种有效的组播交换排队机制[J]. 计算机研究与发展，2005，42(5)：810-817)

[5] Yang Juipin, Chu Yuansun, Liang Mingcheng. Threshold-based selective drop for shared buffer packet switches [J]. IEEE Communications Letters, 2003, 7(4): 183-185

[6] Nachiondo T, Flich J, Duato J. Buffer management strategies to reduce HOL blocking [J]. IEEE Trans on Parallel and Distributed Systems, 2009, 6(1): 1-15

[7] Ascia G, Catania V, Panno D. An evolutionary management scheme in high-performance packet switches [J]. IEEE/ACM Trans on Networking, 2005, 13(2): 262-275

[8] Chu Y S, Yang J P, Wu C S, et al. Partial sharing and partial partitioning buffer management scheme for shared buffer packet switches [J]. IEEE Trans on Communications, 2002, E85-B(1): 79-88

[9] Choudhury A K, Hahne E L. Dynamic queue length threshold for shared-memory packet switches [J]. IEEE/ACM Trans on Network, 1998, 6(2): 130-140

[10] Choudhurya G, Cole R. Design and analysis of optimal adaptive de-jitter buffers [EB/OL]. (2006-04-12) [2009-03-18]. http://wotan. liu. edu/docis/dbl/cocomm /2004_27_6_529_DAAOOA. htm

[11] Sun Lingfen, Emmanuel Ifeachor. Voice quality prediction models and their application in VoIP networks [J]. IEEE Trans on Multimedia, 2006, 8(4): 809-820

Wang Xueshun, born in 1970. PhD candidate. His current research interests include quality of service in multimedia traffic and congestion control of switches.

Xu Ning, born in 1976. PhD candidate. His current research interests include computer network.

Yu Shaohua, born in 1962. PhD. Senior engineer and PhD supervisor. His main research interests include IP network and MAN.

Dai Jinyou, born in 1968. PhD and senior professor. His current research interests include multi-service Ethernet and IP network.

《计算机光盘软件与应用》杂志简介

《计算机光盘软件与应用》杂志是由中国科学院主管、大恒电子出版社主办的国内外公开发行的综合性国家级学术期刊.中国学术期刊(光盘版)全文收录期刊、中国核心期刊(遴选)数据库全文收录期刊、中文科技期刊数据库全文收录期刊.

国际标准刊号:ISSN 1007-9599

国内统一刊号:CN 11-3907/TP

邮发代号:82-271

栏目设置

1.软件设计开发:软件工程,程序设计,基于计算机、电子、自动化各领域的理论与应用研究;

2.信息技术应用研究:

3.工程技术:网络与通信技术、信息安全技术、开发研究与设计技术、人工智能及识别技术;

4.多媒体技术及应用:远程教育、多媒体教学、网络教研;

5.计算机教学与教育信息化:计算机化教学,计算机教学应用研究;

6.科技创新.

征稿对象

从事计算机、电子、通讯、教育方面学习或工作的人员均可向本刊投稿.

编辑部地址:北京市丰台区方庄太阳天地商务会馆6层 **邮编**:100078

联 系 人:彭西宁 **电话**:010-67217131

邮 箱:JSJGPRJYYYZZS@126.COM

(JSJGPRJYYYZZS为“计算机光盘软件与应用杂志社”首字母).