一种区分服务网络边缘路由器动态阈值的公平性标记算法*

杨晓萍[a,b]，李　帅[b]，刘晓娇[b]，陈　虹[a,b]
（吉林大学 a. 汽车仿真与控制国家重点实验室；b. 通信工程学院，长春 130025）

摘　要： 针对区分服务分组标记算法没有考虑到流特性，从而导致不同的流在带宽竞争中出现公平性较差的问题，提出了应用在网络边缘路由器上的一种基于动态阈值的三色标记算法 DTTCM。该算法通过调整流速阈值 minth 和 maxth，按照一定的策略对数据包进行标记，从而保证数据流在带宽不足时能公平地享受带宽。理论分析和仿真结果表明，DTTCM 在 UDP 流与 TCP 流的带宽分配上，以及目标速率对于聚流间剩余带宽分配的影响上，均有着较好的公平性。

关键词： 通信技术；区分服务；边缘路由器；标记算法；公平性
中图分类号： TN915　　**文献标志码：** A　　**文章编号：** 1001-3695(2012)10-3854-03
doi:10.3969/j.issn.1001-3695.2012.10.066

Fairness marker algorithm based on dynamic thresholds for edge routers in DiffServ networks

YANG Xiao-ping[a,b], LI Shuai[b], LIU Xiao-jiao[b], CHEN Hong[a,b]
(*a. State Key Laboratory of Automotive Simulation & Control, b. College of Communication Engineering, Jilin University, Changchun* 130025, *China*)

Abstract: DiffServ packet marking algorithms have not considered traffic features. This result in that fairness is present in the bandwidth competition between different flows. In order to solve the fairness problems, this paper presented dynamic threshold three color marking (DTTCM) algorithm appling in network edge routers. It marked packets according to certain strategy by changing threshold of flow speed minth and maxth. This ensured that flows enjoy network resources fairly when bandwidth was insufficient. The simulation indicates that DTTCM has better performance in the assignment of bandwidth among UDP and TCP, as well as the assignment of excess bandwidths among the aggregates.

Key words: communication technology; DiffServ; edge router; marker algorithm; fairnes

随着 Internet 业务量的不断增长，QoS[1] 日益受到人们的重视。区分服务(DiffServ)体系结构[2] 是 IETF 为实现 Internet 服务质量控制的一种网络模型。它解决了综合服务(IntServ)[3] 的可扩展性和鲁棒性问题，且易于实现和扩展，结构层次化明显。但在 DiffServ 中，由于数据流的特性和要求不同，致使它们在带宽争用中具有不同的竞争能力。传统分组标记算法并没有考虑到这点，因而确保为不同特性的流公平地分配网络带宽成为目前研究的热点。

TCP 流、UDP 流具有不同的流特性，当网络发生拥塞时采取的策略也不相同。TCP 流利用滑动窗口机制来实现流量控制和拥塞控制[4]，当网络拥塞时，TCP 流会自动降低发送速率；UDP 流则不具有流量控制和拥塞控制机制，因此在网络拥塞时仍按原速率发送，抢占由 TCP 降低发送速率后暂时空余出来的带宽，造成带宽分配的不公平[5,6]。本文在分析分组标记算法的基础上，对 TSWTCM(time sliding window three color maker)标记算法在策略上进行改进，以改善不同特性数据流在带宽竞争中的公平性。

1　TSWTCM

TSWTCM 是一种典型的基于 TSW[7] 的三色标记器，它通常包含速率估计器和标记器两个独立的部件。通过这两个部件将每个包与一个颜色关联。其原理如图 1 所示。

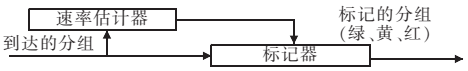

图 1　TSWTCM 原理

TSWTCM 标记器基于速率估计器在时间滑动窗口内对平均速率的估计，通过比较目标速率(CIR)和峰值速率(PIR)，依照概率将分组标记为三种颜色，每种颜色拥有不同的丢弃优先级，由低到高分别为绿(green)、黄(yellow)、红(red)。但 TSWTCM 标记算法存在较为严重的汇聚流之间带宽分配不公平的现象。Su 等人分析了该算法的缺陷，并在其基础上提出了一种改进的算法，即 ItswTCM[8]。与 TSWTCM 相比，该算法

收稿日期： 2012-03-12；**修回日期：** 2012-04-25　**基金项目：** 国家教育部长江学者和创新团队发展计划资助项目(IRT1017)；吉林省科委自然科学基金资助项目(201115030)

作者简介： 杨晓萍(1963-)，女，黑龙江鹤岗人，教授，博士，主要研究方向为网络数据传输的优化控制(yangxiaoping2004@126.com)；李帅(1988-)，男，山东烟台人，硕士研究生，主要研究方向为网络拥塞控制；刘晓娇(1990-)，女，吉林长春人，硕士研究生，主要研究方向为网络拥塞控制；陈虹(1963-)，女，浙江桐乡人，教授，博导，主要研究方向为网络优化控制和预测控制.

令带宽分配的公平性得到很大改善。但是改进的一些算法要么对 TCP 协议本身作了改动,要么需要在边界路由器上保留流的状态信息,因此有不易扩展的缺点。

2　DTTCM 标记算法

如果数据流在网络中按比例享用网络资源,那么其瞬时吞吐率与总吞吐率之比应该和其目标速率与总目标速率之比相等。如果比值不等,说明数据流之间没有公平享有带宽。基于这种思想,为了改善 TSWTCM 存在的不足,本文设计了基于动态阈值的三色标记算法(DTTCM)。

DTTCM 算法的关键是两个关于流速的阈值 minth 和 maxth,DTTCM 会根据测量到的包的到达速率动态地调整这两个阈值。根据这两个阈值,DTTCM 按照某种策略将数据包标记为 green、yellow 或 red,以确保数据流在带宽不足时能公平享用瓶颈带宽,并当带宽充足时能公平享用剩余带宽。

在 DTTCM 中,这两个阈值只有在数据流的速率超过目标速率时才起作用;而当其速率小于目标速率时,将数据包标记为 green,此时这两个阈值不起作用,主要是为了保护这些流,使其尽快达到目标速率。具体算法如下所示:

if (avgrate≤cir)

$P_{green} = 1$;

else{

fairness = (thput/thputtotal)/(cir/cirtotal);

maxth = cir × (1 + 1/fairness);

minth = cir × $\sqrt{\text{cir/cirmax/fairness}}$;

maxp = cir/[cirmax × (1 + fairness)]

P_{green} = minth/avgrate;

if(avgrate≤maxth)

$\{P_{yellow} = [1 - [\frac{avgrate - cir}{maxth - cir}]^2 \times (1 - maxp)] \times [1 - \frac{cir}{avgrate}] + [\frac{cir}{avgrate} - P_{green}];\}$

else

$\{P_{yellow} = \frac{cir}{avgrate} - P_{green};\}$

$P_{red} = 1 - P_{green} - P_{yellow};\}$

图 2 中,avgrate 为估计的数据包的平均速率;thput 为统计的瞬时吞吐率;thputtotal 为瓶颈链路带宽;fairness 为公平因子;cir 为到达包的目标速率;cirtotal 为所有流的目标速率之和;cirmax 为各个流当中的最大目标速率;P_{green}、P_{yellow}、P_{red}分别为标记绿包、黄包和红包的概率。

如果数据流享用的带宽资源超过了其在公平情况下享用的带宽资源,则 fairness 值大于 1;反之小于 1。fairness 值越大或者越小,都说明公平性越坏,DTTCM 需要减小(或增加)数据包标记为绿包和黄包的概率,增加(或减小)标记为红包的概率。利用阈值 minth 来决定标记为绿包的概率,利用 minth 和 maxth 来共同决定标记为黄包的概率。因此,fairness 值越大,minth 和 maxth 值就应该越小;反之这些值应该越大。另外,考虑到目标速率对 TCP 流公平享用网络资源的影响[9],DTTCM 在计算 minth 和 maxth 时都包含了因子 cir,并且在计算 minth 时将 cirmax 也考虑进去,以确保目标速率较大流的数据包有更多的机会被标记为绿包和黄包,从而使得该流能够更公平地享用网络资源,减小目标速率对这种公平性的影响。

当 avgrate≤cir 时,说明该流所享用的资源未到达目标值,该流的数据包以概率 1 被标记为绿包,从而使得该流尽快达到其目标速率。

当 cir < avgrate≤maxth 时,以 minth/avgrate 概率将数据包标记为绿色,如果 fairness > 1,则 P_{green}不会大于 cir/avgrate,也就是说当竞争瓶颈链路带宽时获得的资源不会超过其目标值,从而确保瓶颈链路带宽的公平分配。标记为黄包的概率 P_{yellow} 包含两个部分,如图 2 所示(图 2、3 中各阴影部分面积表示相关的标记概率)。

$$[\frac{cir}{avgrate} - P_{green}] \tag{1}$$

$$[1 - [\frac{avgrate - cir}{maxth - cir}]^2 \times (1 - maxp)] \times [1 - \frac{cir}{avgrate}] \tag{2}$$

式(1)加上 P_{green}就等于 cir/avgrate,因而在带宽充足时可以确保该流达到目标速率。式(2)标记的数据包是用来竞争额外带宽,为确保额外带宽的公平享用,式(2)对过多享用额外资源的数据流作了“约束”。maxth 和 maxp 都与目标速率 cir 及公平因子 fairness 相关,这样公平因子偏大的流或者 cir 值较小的流的数据包被标记为黄包的概率就较小。式(2)中引入了平方项,确保了 P_{yellow}不会下降过快而影响 DTTCM 算法的稳定性。式(2)中引入了 maxp 这个因子。从图 2 中可以看到,maxp 越大时,数据包被标记为黄包的概率也就越大,并且 maxp 与 cir、cirmax 及 fairness 相关,随着 cir 的增加而增加,随着 fairness 的增加而减小。这样式(2)确保了流速越大的流其数据包标记为黄包的概率也就越小,减少了其享用额外带宽资源的机会,确保了其他流有更多机会公平享用额外网络资源。

当 avgrate > maxth 时,以 minth/avgrate 概率将数据包标记为绿色,(cir - minth)/avgrate 概率将数据包标记为黄色,其余部分都标记为红色,如图 3 所示。与图 2 相比,此时标记为红色的概率大大增加。这种情况说明该流占用的网络资源大大超过了其按比例公平应该享用的,因此要对其进行一定的“惩罚”,以维护其他流的“权益”。

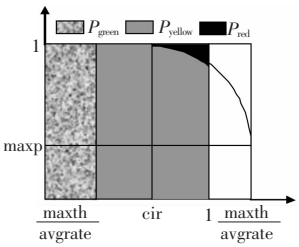

图 2　cir<avgrate≤maxth 时 DTTCM 算法标记

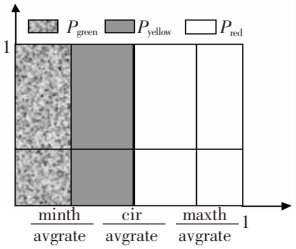

图 3　avgrate>maxth 时 DTTCM 算法标记

DTTCM 中的参数 minth、maxth、maxp 是根据瞬时吞吐率动态变化的,避免了参数敏感性问题,以适应网络流量的动态变化。

3　算法仿真

使用 NS2[10](network simulation version2)作为仿真工具,并与 TSWTCM 算法进行仿真结果比较,验证 DTTCM 算法性能。实验仿真拓扑如图 4 所示。

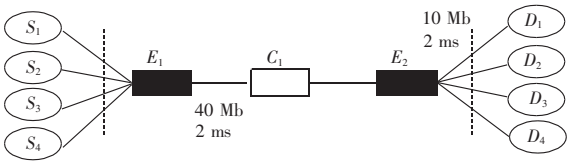

图 4　TCP 与 UDP 间的公平性仿真实验拓扑

图 4 中 $S_1 \sim S_4$ 为源端,$D_1 \sim D_4$ 为对应的目的端,E_1、E_2 为 Diffserv 的边缘路由器,C_1 为 Diffserv 的核心路由器,C_1 与 E_2 之间为瓶颈链路,带宽为 5 Mbps。各链路带宽和时延如图 5 所示。核心路由器采用 RIO-C 队列管理机制,仿真参数如表 1 所示。源端 S_1、S_2 分别发送包长度为 1 000 Byte 的连续 TCP

流,TCP 初始窗口设为 1 000。源端 S_3、S_4 分别发送包长度为 1 000 Byte、速率为 2 Mbps 的 UDP 流。

表 1　DTTCM 算法仿真参数

	AF3	AF2	AF1
minth	5	10	20
maxth	10	20	40
maxp	0.20	0.10	0.02
w_q		0.02	

设计三种情况下的仿真实验,即不足定购(under-order)、充分定购(well-order)和过量定购(over-order),其定购级别(order level)分别为 40%、80% 和 120%。在三种定购级别下,四个源端拥有相同的目标速率和峰值速率,在定购级别为 40% 时,目标速率为 0.5 Mbps,峰值速率为 0.6 Mbps;在定购级别为 80% 时,目标速率为 1Mbps,峰值速率为 1.2 Mbps;在定购级别为 120% 时,目标速率为 1.5 Mbps,峰值速率为 1.8 Mbps。仿真时间为 30 s。

仿真结果如图 5 ~ 10 所示。图中横坐标是时间,纵坐标是吞吐率。tcp1、tcp2、udp1、udp2 分别是源端 S_1、S_2、S_3、S_4 发出的数据流。

图 5 和 6 分别给出了两种算法下定购级别为 40% 时的仿真结果。由于带宽充足,两种算法的 TCP 流在前 10 s 的调整之后,都能达到目标速率。TSWTCM 算法中,UDP 流的吞吐率在 1.8 Mbps 左右,而 TCP 的吞吐率在其目标速率 0.5 Mbps 左右,剩余带宽基本上被 UDP 流占用,TCP 流几乎得不到剩余带宽;DTTCM 算法中,UDP 的吞吐率在 1.7 Mbps 左右,TCP 的吞吐率在 0.7 Mbps 左右,虽然 UDP 依然抢占了大量的剩余带宽,但 DTTCM 算法整体的公平性要好于 TSWTCM 算法。

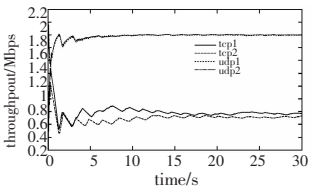

图 5　TSWTCM 算法下定购级别为 40%时 TCP 与 UDP 的吞吐率比较

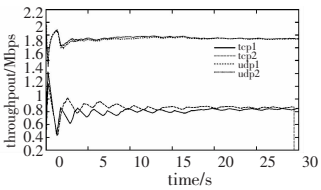

图 6　DTTCM 算法下定购级别为 40%时 TCP 与 UDP 的吞吐率比较

图 7 和 8 分别给出了两种算法在定购级别为 80% 时的仿真结果。TSWTCM 算法中,UDP 流的吞吐率在 1.6 Mbps 左右,TCP 流的吞吐率在 0.8 Mbps 左右,TCP 流无法达到其目标速率;DTTCM 算法中,UDP 流的吞吐率在经过一段时间的调整之后稳定在 1.4 Mbps 左右,TCP 流稳定在目标速率 1 Mbps 左右,从而可看出 DTTCM 公平性好于 TSWTCM 算法。

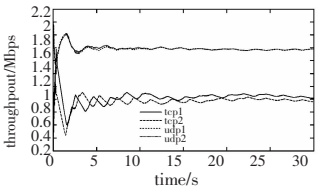

图 7　TSWTCM 算法下定购级别为 80%时 TCP 与 UDP 的吞吐率比较

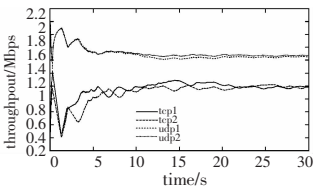

图 8　DTTCM 算法下定购级别为 80%时 TCP 与 UDP 的吞吐率比较

图 9 和 10 分别给出了两种算法下定购级别为 120% 时的仿真结果。由于处于过量定购状态,两种算法的 TCP 流都不能达到其目标速率。TSWTCM 算法中,UDP 流的吞吐率在 1.6 Mbps 左右,而两个 TCP 流的吞吐率相差较大,一个在 1 Mbps 左右,另一个在 0.7 Mbps 左右;DTTCM 算法中,UDP 流的吞吐率比 TSWTCM 中略小,在 1.5 Mbps 左右,两个 TCP 流的吞吐率都稳定在 1 Mbps 左右,从而可看出其公平性要好于 TSWTCM 算法。

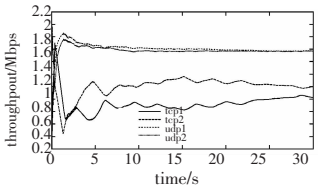

图 9　TSWTCM 算法下定购级别为 120%时 TCP 与 UDP 的吞吐率比较

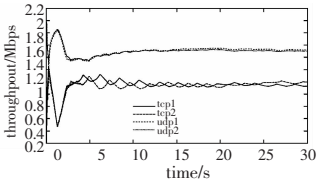

图 10　DTTCM 算法下定购级别为 120%时 TCP 与 UDP 的吞吐率比较

表 2 ~ 7 分别给出了上述仿真的数据结果。DTTCM 算法在定购级别为 40% 和 80% 时,TCP 流的丢包率要小于 TSWTCM 算法,UDP 流的丢包率大于 TSWTCM 算法;并且相对于 TSWTCM 算法,DTTCM 算法下的 TCP 流的平均吞吐率更大,UDP 流的平均吞吐率更小,DTTCM 算法下的链路利用率更高。在定购级别为 120% 时,DTTCM 算法下虽然 TCP 与 UDP 流的丢包率都有所增加,但相比 TSWTCM 算法,TCP 流的平均吞吐率更大,UDP 流的平均吞吐率更小,整体公平性更好。

表 2　TSWTCM 算法下定购级别为 40%时 TCP 与 UDP 仿真结果

源端	丢包率	throughput/Mbps	链路利用率/%
TCP1	0.044	0.545	93
TCP2	0.046	0.511	
UDP1	0.114	1.795	
UDP2	0.112	1.800	

表 3　DTTCM 算法下定购级别为 40%时 TCP 与 UDP 仿真结果

源端	丢包率	throughput/Mbps	链路利用率/%
TCP1	0.041	0.631	95
TCP2	0.039	0.637	
UDP1	0.153	1.734	
UDP2	0.146	1.746	

表 4　TSWTCM 算法下定购级别为 80%时 TCP 与 UDP 仿真结果

源端	丢包率	throughput/Mbps	链路利用率/%
TCP1	0.029	0.788	94
TCP2	0.029	0.737	
UDP1	0.266	1.580	
UDP2	0.266	1.580	

表 5　DTTCM 算法下定购级别为 80%时 TCP 与 UDP 仿真结果

源端	丢包率	throughput/Mbps	链路利用率/%
TCP1	0.027	0.945	96
TCP2	0.026	0.953	
UDP1	0.378	1.452	
UDP2	0.358	1.473	

表 6　TSWTCM 算法下定购级别为 120%时 TCP 与 UDP 仿真结果

源端	丢包率	throughput/Mbps	链路利用率/%
TCP1	0.018	0.795	98
TCP2	0.019	0.965	
UDP1	0.270	1.574	
UDP2	0.274	1.570	

表 7　DTTCM 算法下定购级别为 120%时 TCP 与 UDP 仿真结果

源端	丢包率	throughput/Mbps	链路利用率/%
TCP1	0.039	0.898	98
TCP2	0.028	0.966	
UDP1	0.313	1.523	
UDP2	0.329	1.505	

如前分析,目标速率会影响 DiffServ 聚流间剩余带宽享用的公平性,为此本文设计了图 11 所示的实验拓扑。拓扑中有两个聚流,每个聚流由三个单流组成,S_1、S_2、S_3 形成聚流 1,S_4、S_5、S_6 形成聚流 2。E_0、E_1、E_2 为 DiffServ 的边缘路由器,C_1 为核心路由器,C_1 与 E_2 间为瓶颈链路。各链路带宽和时延如图 11 所示。核心路由器采用 RIO-C 队列管理机制,仿真参数如表 1 所示。源端分别发送包长度为 1 000 Byte 的连续 TCP 流,设置峰值速率为目标速率的 1.2 倍,聚流 1 的目标速率 cir1 始终是 1 Mbps;而聚流 2 的目标速率 cir2 从 1 Mbps 增加到 8 Mbps,也就是网络定购级别从 20% 增加到 90%,TCP 初始窗口设为 1000,仿真时间为 30 s。

公平性指数的定义如下:

$$FI = \frac{(\sum_i x_i)^2}{(N\sum_i x_i^2)} \tag{3}$$

其中:$0 < FI < 1$;x_i 为流 i 所得到的剩余带宽/流 i 的目标速率 cir;N 是聚流的总数量(在本仿真中,$N = 2$)。(下转第 3860 页)

(上接第 3856 页)根据以上定义,公平性指数越接近 1,则剩余带宽在聚流中的分配越公平。

两种标记算法在不同的网络提供级别下,其公平性指数如图 12 所示。

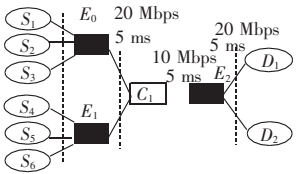

图 11　目标速率对公平性影响的仿真实验拓扑

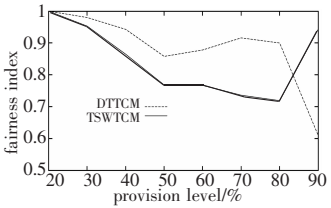

图 12　公平性指数比较

4　结束语

区分服务标记算法的公平性是保证业务流服务质量和提高网络资源利用率的一个重要环节。本文分析了 TSWTCM 算法,针对其不足提出了基于动态阈值的标记算法 DTTCM。该算法通过一个公平因子的变化来间接感知数据流参数的变化,动态调整流速阈值 minth 和 maxth,并按照一定的策略将数据包标记为 green、yellow 或 red。实验证明:相对于 TSWTCM 算法,DTTCM 在 UDP 流与 TCP 流的带宽分配上,以及目标速率对于聚流间剩余带宽分配的影响上,均有着较好的公平性和可扩展性。

参考文献:

[1] 林闯,李寅,万剑雄. 计算机网络服务质量优化方法研究综述[J]. 计算机学报,2011,**34**(1):1-14.

[2] 诸方舟,王晓东,孟祥旭. 无线网络中实时性服务区分的转发节点队列调度方法研究[J]. 计算机应用研究,2010,**27**(1):330-332,338.

[3] BRADEN R,CLARK D,SHENKER S. RFC1633,Integrated services in the Internet architecture:an overview[S]. 1994.

[4] 杨晓萍,史帅,陈虹. 一种改进的 TCP 拥塞控制算法[J]. 吉林大学学报:工学版,2006,**36**(3):433-437.

[5] ELSHAIKH M A,OTHMAN M, SHAMALA S,*et al*. A new fair marker algorithm for DiffServ networks [J]. Computer Communications,2008,**31**(14):3064-3070.

[6] 姜文刚,孙金生,王执铨. 弃头和随机结合的被动队列管理[J]. 计算机应用研究,2011,**28**(5):1798-1801.

[7] NEGLIA G, FALLETTA V,BIANCHI G. An analytical model of a new packet marking algorithm for TCP flows[J]. Computer Networks,2006,**50**(8):1176-1191.

[8] 郝俊瑞,余少华. 一种在区分服务网络中新的三色标记器[J]. 计算机科学,2008,**35**(6):55-59.

[9] 龚静,吴春明,孙维荣,等. FCA-ItswTCM:区分服务中一种公平的拥塞自适应标记算法[J]. 电子学报,2011,**20**(7):1624-1627.

[10] UCN/LBL/VINT. Network simulator-NS2 [EB/OL]. (2003-09). http://www. mash. cs. berkeley. edu/ns.

[11] SU Hong-jun,ATIQUZZAMAN M. ItswTCM:a new aggregate marker to improve fairness in Diffserv[J]. Computer Communications,2003,**26**(9):1018-1027.