Fairness of Responsive and Un-Responsive Aggregate for Assured Forwarding PHB in Differentiated Services IP Networks*

WU Hai-tao[1], LONG Ke-ping[1], CHENG Shi-duan[1], MA Jian[2], ZHANG Run-tong[2]

[1](*National Key Laboratory of Switching Technology and Telecommunication Networks, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*);

[2](*Nokia China R&D Center, Beijing* 100013, *China*)

E-mail: {xiaodan,chsd}@bupt.edu.cn

http://www.bupt.edu.cn

Received December 18, 2000; accepted April 17, 2001

Abstract: This paper defines the meaning of fairness more suitable for AF PHB in DS network. Besides, this paper argues that responsive and un-responsive flows should be treated differently only to their out-of-profile traffic. In addition, traffic responsiveness should be identified by its behavior rather than simply by protocol. Moreover, two markers, srTCM and trTCM, with eight mapping schemes are examined by simulations to check whether they can provide fairness in DS networks. This paper also defines fairness area as the network load range where the DS network can provide fairness to responsive and un-responsive aggregate by marking without discriminating, and parameters configuration of the area is also given.

Key words: quality of service (QoS); differentiated services; fairness; bandwidth allocation

Differentiated Services (DS)[1] Architecture, recently proposed by IETF, provides scalable means to deliver IP QoS based on handling of traffic aggregates. Traffic Classification state is conveyed by means of IP-layer packet marking using the DS field. Packets are classified and marked at the boundaries of network to receive a particular PHB (per-hop behavior) along the path. Sophisticated classification and traffic conditioning (TC), including marking, policing, and shaping operations, need only to be implemented at network boundaries or hosts. Within the DS domain, core router forwards packets according to the DSCP value in the packet header. Up to now, IETF only defines two sets of PHBs, Expedited Forwarding(EF) PHB and Assured Forwarding (AF)[2] PHBs. There are 4 independent AF classes, and 3-drop precedence level in each class. Recent studies have shown that under various conditions, existing Diffserv mechanism may have problems of unfairness and inefficient resource utilization, thereby failing to achieve the desired QoS for TCP flows running over assured services[3,4]. The interactions between

* Supported by the National Natural Science Foundation of China under Grant No. 69972008 (国家自然科学基金); Nokia-BUPT Advanced TCP/IP Networks Project

WU Hai-tao was born in 1976. He is a Ph. D. candidate at the Department of Computer Science, Beijing University of Posts and Telecommunications (BUPT). His research interests are QoS, TCP/IP and differentiated services. **LONG Ke-ping** was born in 1968. He is an associate professor at the BUPT. His research interests are ATM/IP performance analysis, TCP enhancement and IP QoS. **CHENG Shi-duan** was born in 1940. She is a professor and doctoral supervisor at the BUPT. Her research interests are ISDN, ATM, network performance, security, survivability and IP QoS. **MA Jian** was born in 1959. He is a project manager at the Nokia China R&D Center. His research interests are TCP enhancement, IPv6 and QoS. **ZHANG Run-tong** was born in 1963. He is an associate professor and an invited researcher at the Nokia China R&D Center. His research interests are AI, network performance and QoS.

responsive and un-responsive traffic in the DS network make the fairness much more difficult to obtain[5,6].

Meanwhile, Floyd promotes the use of End-to-End congestion control in future protocols for BE traffic[7]. We argue that End-to-End congestion control should be used in AF if the customer prefers to get excess bandwidth by generating out-of-profile traffic. A DS network may encourage the use of End-to-End congestion control by giving more excess bandwidth to responsive aggregate. If the network cannot provide fairness for responsive and un-responsive aggregate competing excess bandwidth, it may encourage customers to use unresponsive flows even they don't need to, which will cause unfairness and make the DS network difficult to maintain. However, the in-profile traffic of both responsive and unresponsive aggregate must not be punished. Based on the original intention of DS, i. e., providing customers with some kind of QoS, we argue that all the in-profile traffic must be protected, whether or not it is responsive.

The rest of the paper is organized as follows. Section 1 discusses some related works. Section 2 redefines the fairness in AF PHB in DS network. We present our marking scheme in Section 3. Then we evaluate and analyze those schemes by simulations in Section 4. Finally, Section 5 concludes the paper.

1 Motivation and Related Works

Fairness between responsive flows and unresponsive flows is difficult to solved. Floyd suggests prompting the use of End-to-End congestion control in the Internet[7]. However, network cannot rely on customers' corporation totally. Besides, it is very difficult to identify an unresponsive or malicious flow in traditional IP networks.

Clark gives the framework of AF by using RIO scheme[3], which is an extension of RED[8]. A single FIFO queue and two-drop precedence are used. Traffic exceeding target rate is marked out-of-profile, otherwise packets are marked in-profile. Two sets of parameters are used separately for In and Out packets to drop them distinguishingly in congestion. Ibanez[4] goes a step further by mixing AF and BE traffic. They conclude that AF cannot provide a strict allocation of bandwidth between several users. Goyal[6] concludes that three colors help clearly distinguish congestion sensitive and insensitive flows, and the reserved bandwidth should not be overbooked. For three color marker, they use two token buckets with two separate token generation rates for green and yellow color. Besides, they list four possible buffer management techniques by distinguishing accounting to be Single or Multi and Threshold to be Single or Multi: SAST, SAMT, MAST, and MAMT. They use SAMT (Single Account, Multi Threshold).

Seddigh[5] lists six possible mapping schemes for TCP In/Out profile traffic and UDP In/Out profile traffic, and they compare the results of MAST and MAMT. Besides, they use TSW (Time Sliding Window) for marker. The difference between their work and ours is that we use token bucket for marking and SAMT for queue management. However, since we get similar results by identical mapping scheme, we conclude that it is the mapping scheme for In/Out packets that determines the bandwidth allocation of different aggregate type. We go a step further by examining srTCM[9] and trTCM[10], both of which can mark traffic into three colors.

2 New Fairness Definition in AF PHB

The key issue is that the original idea of Diff-serv framework is to realize some kind of QoS to a customer. In AF, a customer has a target rate and if the aggregate from the customer does not exceed this rate, the traffic should not be punished, whether it is responsive or not. We must make the network able to deal with the customer's out of profile traffic, and a customer can never receive high performance by simply using un-responsive flows. Besides, we argue that Floyd's opinion of prompting the use of End-to-End congestion control is also fit for excess bandwidth sharing in AF PHB, especially when the customer prefers to get out-of-profile bandwidth.

Based on the original intention of DS, i. e., providing customers with some kind of QoS, as well as the opinion of prompting the use of End-to-End congestion control in the Internet, we redefine the meaning of fairness for AF PHB in DS network as following:

1) In an over-provisionec network, all in-profile traffic should be protected and target rate must be achieved if the customer has enough loads;

2) In an over-provisioned network, out-of-profile traffic can access the excess bandwidth proportional to their target rates. We strongly prompt the use of End-to-End congestion control and we argue that it should provide more bandwidth to responsive ones;

3) In an under-provisioned network, the network should keep responsive flows and un-responsive flows degrading in proportion to their profiles. Never can one aggregate be starved by the other.

All other works use UDP to generate constant bit rate traffic without explanation. We argue that it may lead to a misunderstanding opinion that TCP equals to responsive and UDP equals to unresponsive. It is unreasonable to identify an un-responsive or a malicious flows simply by whether it uses TCP or UDP. In this paper, we just use TCP to generate responsive traffic and UDP to generate un-responsive traffic. However, this absolutely never means that we can identify whether a flow is responsive or not simply by its transport protocol.

3 Marking Schemes

A DS node should try to protect packets with a lower drop precedence from being lost by preferably discarding packets with a higher drop precedence value. An AF class includes three-drop precedence. In this paper, green color means the lowest drop precedence and red color means the highest drop precedence, and yellow color is the middle one. We list all marking schemes in Table 1. Here G/Y means green for in-profile traffic and yellow for out-of-profile traffic. Note that schemes 1 to 6 are only fit for the case where the responsiveness can be identified at the edge of the DS network, while schemes 7 and 8 can be used if the responsiveness is un-known at the edge of DS network. Schemes 1 to 6 are also used by Seddigh[5]. However, unlike Seddigh who uses TSW maker, we use trTCM, which is based on token bucket and has far more flexibility, particularly useful with dual target rates. In this paper, we only consider one target rate. The simulation tool is NS-2.1b6.

Table 1 Marking schemes in AF class

	Marking scheme							
	1	2	3	4	5	6	7	8
TCP	G/Y	G/Y	G/Y	G/R	G/Y	G/Y	srTCM	trTCM
UDP	G/Y	Y/Y	Y/R	Y/R	R/R	G/R	three colors	three colors

G: Green, Y: Yellow, R: Red, In-profile/Out-profile Color

4 Simulation Results and Analysis

This section presents the simulation results for all schemes listed in Table 1. The simulation network topology is shown in Fig. 1. The link between core router (C1) and edge router (E3) is the only bottleneck link. Traffic from the source will send to the destination with corresponding number. There are 4 responsive aggregates; each one consists of 6 TCP micro-flows with the same total target rate of X Mbps. 2 UDP flows are used to generate un-responsive aggregate; each keeps sending at a constant bit rate of 1Mbps with a target rate of 0.5Mbps. IP Packet size is 1k-bytes.

4.1 Marking schemes for un-responsive aggregate can be distinguished from responsive ones

Figure 2 gives the results of marking scheme 1, i.e., marking both TCP and UDP in-profile traffic to green and out-of-profile to yellow. This scenario is the same as two-color marking scheme and using RIO. Two points need to be examined carefully. One is when X equals to 0.5Mbps, where TCP and UDP has the same target rate; the other is when X equals to 1Mps and the total target rate of all traffic equals to the bottleneck bandwidth. We can see that two-color scheme cannot provide good fairness even when lightly over-provisioned[6]. The results of TCP are slightly better when we use TCP SACK.

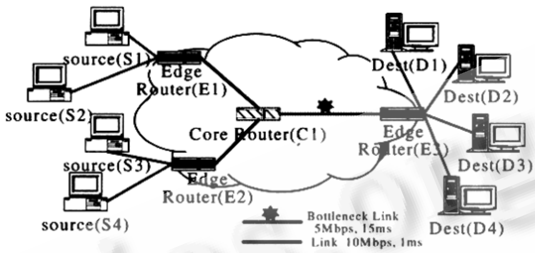

Fig. 1 Simulation network topology

Result of scheme 2 is similar to that in Ref. [5], as shown in Fig. 3. We can see that UDP is completely starved when network resource is not enough because the in-profile traffic of UDP has higher drop precedence than that of TCP. It is not true that this scheme provides fairness when the network in over-provisioned. We illustrate the scenario by Fig. 4, where we set UDP sending rate to 0.5Mbps, i.e., UDP sends traffic just at its target rate. It can be easily understood that such traffic should be protected when network is over-provisioned. But in Fig. 4 we can see UDP get poor throughput, which means high drop rate and poor QoS. Then we conclude this marking scheme cannot be used. From Fig. 5 we can see the result of marking scheme 3. Certainly UDP in this scheme is treated much unfairly than in scheme 2. In fact, if we set UDP sending rate to 0.5Mbps, we get similar results as Fig. 5. Although it can control un-responsive aggregate well, it cannot provide fairness as defined in Section 2 because target rate of UDP is not achieved even when network is over-provisioned.

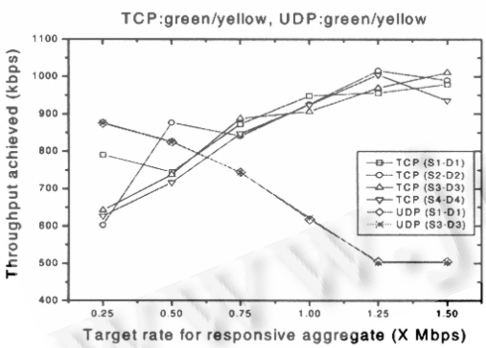

Fig. 2 Marking scheme 1

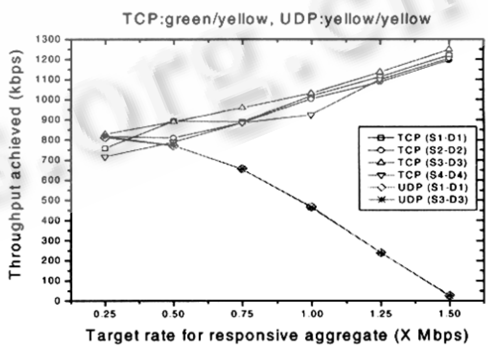

Fig. 3 Marking scheme 2

Simulation results for schemes 4 to 6 are skipped for space limitation. We also do simulations of all the schemes for different TCP versions and UDP sending rates. The results of TCP of all these six marking schemes are better than that shown in Ref. [5], however none of these schemes could bring fairness to responsive and unresponsive aggregate. Meanwhile, except for scheme 1, all the other schemes need to identify the responsiveness of an aggregate at the edge of DS domain, which is difficult in real DS network.

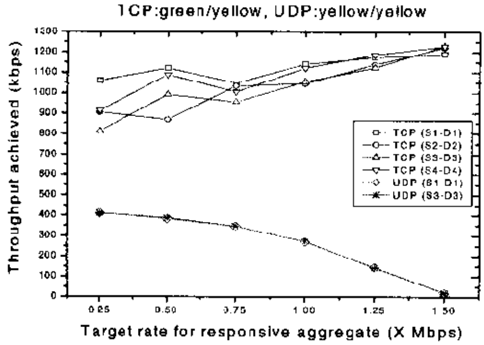

Fig. 4 Marking scheme 2 (UDP is at 0.5Mbps)

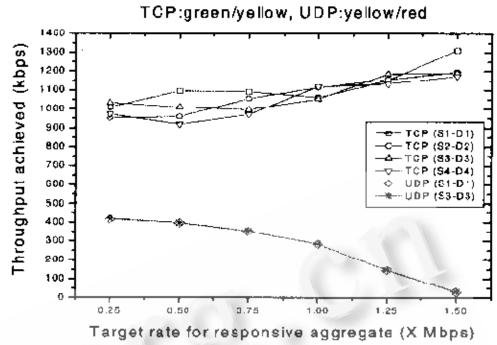

Fig. 5 Marking scheme 3

4.2 Responsiveness is un-known at edge of DS

Two token bucket based markers, srTCM[9] and trTCM[10], are examined to check whether it can provide fairness by three drop precedences to different traffic if their responsiveness are un-known at the edge or boundary of a DS domain. Here a responsive aggregate consists of 6 TCP Reno micro-flows.

For srTCM, there are three parameters, CIR is set to target rate of the aggregate. How to set CBS and EBS is an issue. We find bigger token bucket size favors bursty traffic of TCP. Besides, fairness in responsive aggregates themselves is still an issue. In our simulations all parameters are determined by target rate, irrespective it is TCP or UDP, i.e., all aggregates are treated fairly at the marker. When network is well over-provisioned (X is between 0.25Mbps and 0.75Mbps) or severely under-provisioned (X is above 1.25Mbps), the results are similar to that of scheme 1. However, when network load get closer to 100% (X is 1Mbps), it can provide good fairness for both responsive and un-responsive aggregates, satisfy all the requirements we defined in Section 2. If CBS and EBS are set to 200ms * CIR, throughput achieved by UDP is slightly less than that in Fig. 6. And the area where TCP themselves can get best fairness moves to slightly over-provisioned field. Based on these results, we conclude that when network load is near 100%, using srTCM can provide good fairness.

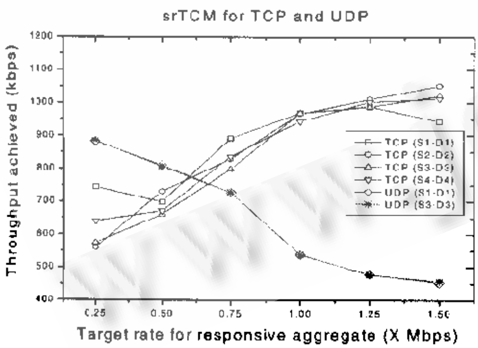

Fig. 6 srTCM with CBS and EBS at 100ms * CIR

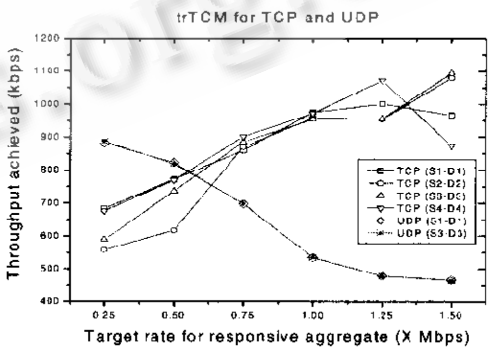

Fig. 7 trTCM with PIR at 1.2 * CIR

For trTCM, there are four parameters, CIR is set to the target rate of the aggregate. CBS and PBS are set to 100ms * CIR. Since here we have no dual target rate, how to set PIR is an issue. Like that in srTCM, we find bigger token bucket size favors bursty traffic such as TCP. We can see the results of setting PIR to 1.2 * CIR, which is shown in Fig. 7. The result is similar to those of srTCM but with some differences. To check the effect of PIR in trTCM, we also set it to 1.5 * CIR and find that un-responsive aggregate can get slightly more when network load is very near 100% or when network is much over-provisioned; while responsive aggregate can get slightly more excess bandwidth and achieve good fairness between themselves when network is over-provisioned, especially when network load is around 80%. In sum, srTCM provides better fairness when network load is near 100%, while trTCM provides better fairness when network is over-provisioned. When network is under-provisioned, those two markers give similar results.

Finally we give an explanation for the results we get from srTCM and trTCM:

1) When network is much under provisioned, two markers get similar results because all in-profile packets are marked to green, packets are marked to yellow or red only because TCP is bursty naturally and UDP may send traffic faster than target rate. Under this situation, fairness is better than that of the case where the marker marks packets into two-color, i.e., marking scheme 1 in Table 1.

2) We find that there is an area where we can provide good fairness between responsive and un-responsive aggregate. Besides, the area is not the same for srTCM and trTCM. Here we define it as fairness area. For token-based marker, the center of this field is determined by token generation rate, and the range of the field is determined by token bucket size. However, we do not mean that bigger token bucket provides greater field because bigger bucket may lead to a burst of packets in high priority, which is not friendly to RIO[3] or mRED[5] based queue management mechanism and may cause continuous packet loss.

Supposing the total target rate of all traffic is X_{all}, and the bottleneck link bandwidth is Y. Then we can use X_{all}/Y to represent the network load at bottleneck link. For srTCM, there is only one token generation rate, which equals to the target rate of the aggregate. Traffic exceeding this rate will be marked into low priority. The center of the fairness area where srTCM can provide fairness is around 100% load because in this scenario, all in profile traffic can just be supported by the bottleneck. Setting of CBS and EBS can also affect the range of this area. Bigger bucket size is fit for TCP busrty naturally traffic, while it causes bursty traffic in the same color. We find setting them to about 100ms * CIR works well in our simulations. For trTCM, there are two token rates. One is CIR, which is set to target rate; the other is PIR. If $Y>X_{all}$, setting PIR equal to $(Y/X_{all}) * \text{CIR}$ can provide better fairness. This rule can be understood by following: all in-profile traffic is marked into green; excess bandwidth is divided by out-of-profile traffic proportional to target rate by yellow color; traffic exceeding their proportional division is marked to red. Then if we set PIR as a constant $Z * \text{CIR}$, the center of the fairness area is around $1/Z$ load. For srTCM, the center of fairness is determined by CIR, so it is around 100% load; for trTCM, the center of fairness area is determined by PIR, i.e., we can move this area by setting PIR to different values. For a network load, setting PIR equal to CIR multiplying the inverse ratio of load can provide better fairness. The range of fairness area is determined by token bucket capacity. To some extent, bigger bucket size favors TCP, i.e., performance of TCP is better with bigger bucket capacity.

5 Conclusions

In this paper, we have investigated the interaction of responsive and un-responsive aggregate in an AF-based DS-capable IP network, as well as how to provide them good fairness by using marker at the edge or boundary of DS domain. Based on the original intention of DS, i.e., providing customer with some kind of QoS, we redefine the meaning of fairness more suitable for AF PHB in DS networks. This paper strongly argues the use of End-to-End congestion control for AF service in DS network, which is followed by Ref. [11]. Token bucket based markers have been evaluated by elaborate simulations. A novelty of this paper is that it introduces the idea of End-to-End congestion control to define the fairness suitable for AF PHB; another novelty is that parameter settings for srTCM and trTCM have been analyzed with simulations.

References:

[1] Blake, S., Black, D., *et al*. An architecture for differentiated services. IETF RFC 2475, Dec. 1998.

[2] Heinanen, J., *et al*. Assured forwarding PHB group. IETF RFC 2597, June 1999.

[3] Clark, D., Fang, W. Explicit allocation of best effort packet delivery service. IEEE Transcations on Networking, 1998,6(4):362~373.

[4] Ibanez, J., Nichols, K. Preliminary simulation evaluation of an assured service. IETF Internet Draft, Aug. 1998.

[5] Seddigh, N., Nancy, B. Study of TCP and UDP interaction for the AF PHB. IETF Internet Draft, Aug. 1999.

[6] Goyal, M., Durresi, A., Jain, R., Liu, C. Performance analysis of assured forwarding. Internet Draft, Feb. 2000.

[7] Floyd, S., Fall, K. Promoting the use of end-to-end congestion control in the internet. IEEE Transcations on Networking, 1999,7(4):458~472.

[8] Floyd, S., Jacobson, V. Random early detection gateways for congestion avoidance. IEEE Transcations on Networking, 1993,1(4):397~413.

[9] Heinanen, J., *et al*. A single rate three color marker. IETF RFC 2697, Sept. 1999.

[10] Heinanen, J., *et al*. A two rate three color marker. IETF RFC 2698, Sept. 1999.

[11] Wu, Hai-tao, Long, Ke-ping, *et al*. A direct congestion control scheme for Non-responsive flow control in diff-serv IP networks. IETF Internet Draft, Aug. 2000.

区分服务IP网AF PHB中响应流与非响应流的公平性

郭海涛[1], 隆克平[1], 程时端[1], 马 健[2], 张润彤[2]

[1](北京邮电大学 程控交换与通信网国家实验室,北京 100876);

[2](诺基亚中国研发中心,北京 100013)

摘要: 定义了区分服务网络中AF PHB的公平性,指出对响应流和非响应流区别的不同只应针对其超出规范的流量,并且业务量的响应性的判定应根据其行为而不是仅靠协议类型.通过仿真验证标记器srTCM和trTCM以及8种映射机制是否能提供公平性.在DS网中,存在通过不区分流的响应性的标记算法可提供公平性的负荷范围,作者将其定义为公平区间,并对其参数配置进行了分析.

关键词: 服务质量;区分服务;公平性;带宽分配

中图法分类号: TP393　　**文献标识码**: A