Marking Mechanism for Enhanced End-to-End QoS Guarantees in Multiple DiffServ Environment[1]

Woojin Park, Kyuho Han, Sinam Woo, and Sunshin An

Department of Electronic and Computer Engineering, Korea University, 1, 5 Ga, Anamdong, Sungbukku, Seoul, 136-701, Korea
{progress, garget, niceguy, sunshin}@dsys.korea.ac.kr

Abstract. Recent demands for real time applications have given rise to a need for QoS in the Internet. DiffServ is one of such efforts currently being developed by IETF. In a DiffServ network, previous researchers have proposed several marking mechanisms that are adapted to operate in intra-domain rather than in multi-domain. In this paper, we propose an enhanced marking mechanism based on a color-aware mode using the conventional tswTCM for flows crossing one or more DiffServ domains in order to guarantee effectively enhanced end-to-end QoS. The key concept of the mechanism is that the marking is carried out adaptively based on the initial priority information stored at the ingress router of the first DiffServ domain and the average queue length for promotion of packets. By statistics based on simulations, we show that the proposed marking mechanism improves an end-to-end QoS guarantee for most cases.

1 Introduction

Demands on Quality of Service (QoS) due to explosive growth of real-time traffics such as videoconferencing and video-on-demand services are persistently increasing in current Internet. Differentiated Services (DiffServ) is one of such efforts currently pursued by IETF [1], [2] to add QoS to Internet without fundamental change to the current IP networks. In DiffServ networks, service differentiation is provided based on the DiffServ Code Point (DSCP) field in the IP header and packets with the same DSCP are handled under corresponding forwarding discipline called Per-Hop Behavior (PHB) [3], [4].

DiffServ provides statistical QoS to a few predefined service classes instead of providing guarantees to individual flows. It provides service differentiation among traffic aggregates over a long time scale. A DiffServ network achieves its service goals by distinguishing between the edge and core network. It pushes all complex tasks, such as administration, traffic control, traffic classification, traffic monitoring, traffic marking etc., to the edge network where per flow based schemes may be used. Traffic passing through the edge network will be classified into different service classes and marked with different drop priorities. Core routers implement active queue management schemes such as RED with In and Out (RIO) [5], and provide service differentiation to the traffic according to preassigned service classes and drop priorities carried in the

[1] This work is supported by Korea Research Foundation under contract 2002-041-D00446.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3480, pp. 261 – 270, 2005.

packet header. RIO-like schemes achieve this objective by dropping low priority packets earlier with a much higher probability than dropping high priority packets.

In a DiffServ network, previous researchers have proposed several marking mechanisms that are adapted to operate in intra-domain rather than in multi-domain environment. In this paper, we propose an enhanced marking mechanism as an extension of tswTCM [8] to guarantee effectively enhanced end-to-end QoS for flows crossing one or more DiffServ domains. The key concept of the mechanism is that the marking is carried out adaptively based on the initial priority information stored at the ingress router of the first DiffServ domain and the average queue length for promotion of packets.

Our work is different from previous researches in that our enhanced marking mechanism based on a color-aware mode supports an end-to-end QoS guarantee in a multiple DiffServ environment. The contributions of this paper are as follows:

- We propose an enhanced tswTCM marking mechanism that can provide effectively an end-to-end QoS guarantee for flows crossing one or more DiffServ domains. In addition, our proposed marking mechanism operates in color-aware mode. The previously proposed tswTCM marker doesn't assume a specific mode.
- The proposed marking mechanism is compatible with existing markers.
- By statistics based on simulations, we show that the proposed marking mechanism can improve an end-to-end QoS guarantee in multiple DiffServ environment, in especial for middle load (20%-70% provision level).
- We compare our marking mechanism with the conventional tswTCM. Simulation results show that our marking mechanism achieves improved end-to-end QoS guarantee for most cases.

The remainder of the paper is organized as follows. After the introductory part, Section 2 gives the fundamental concept of traffic markers in DiffServ networks. In Section 3, we consider the problem of previous marking mechanisms in terms of end-to-end QoS guarantee in a multiple domain environment. In Section 4, we present the proposed marking mechanism as an enhanced version of the conventional tswTCM. Section 5 describes simulation topology and configuration parameters. In Section 6, we present and discuss simulation scenarios and results. Finally, we describe conclusion and future work in Section 7.

2 Basic Concept of Traffic Markers

The marker is intended to mark packets that will be treated by the AF (Assured Forwarding) PHB in down-stream routers, which is a component that meters an IP packet stream and marks it either green, yellow, or red. The color of the packet indicates its level of priority: red has the highest priority of rejection, followed by yellow and then green. Concerning the mechanism used to check the traffic conformity to the service profile, packet marking can be further classified in two broad categories: token-bucket based marking and average rate estimator based marking. The best-known markers are a single rate Three Color Marker (srTCM) [6], a two rate Three Color Marker (trTCM) [7], and a Time Sliding Window Three Color Marker (tswTCM) [8]. The first two markers in order exploits token-bucket based marking scheme and function in two modes: the color-aware mode in which the previous color of the packets is taken into account at the entrance to the DiffServ network, and the color-blind mode in which this color is ignored. And, the third marker uses average rate estimator based marking scheme and doesn't assume a specific mode. In the average rate estimator based category, marking is performed according to the measurement of the average arrival rate of aggregated flows. In this marker, the arrival rate is calculated according to the weighted average of the arrival rate over a certain time window/interval [5].

3 Consideration of Markers in a Multiple Domain Environment

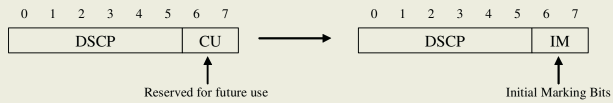
The classification and marking on each packet are carried out at the boundary routers based on conformance to the contracted throughputs for the traffic flow. There is trade-off between color-aware and color-blind mode of markers in a multiple domain environment. In color-aware mode, the pre-colored green or yellow non-conform packets can be downgraded into either yellow or red according to the level of conformity because the marking information in the previous DiffServ domain affects the current domain. In the case of a multiple domain crossing, these downgraded packets are very likely to be lost in the following domains. This is because the downgraded packets cannot restore the state of the initial marking even though the current domain has the excess bandwidth. In color-blind mode, the marking is accomplished regardless of the previous marking information. In this case, the pre-colored green or yellow packets can be downgraded by the preemption of the pre-colored red packets if red conform packets arrive faster than green or yellow packets at the current domain. This is a fatal defect in a viewpoint of an end-to-end QoS guarantee.

To guarantee effectively an enhanced end-to-end QoS in a multiple domain environment of DiffServ networks, a new marker is needed. And this new marker has to function in a color-aware mode to protect the packets with high priority from the less important packets. In this paper, we focus on the color-aware mode of the conventional tswTCM.

4 Proposed Marking Mechanism

The DSCP field in the IP header (IPv4 or IPv6) is defined to allow differentiated processing based on the value of this field. It should be noted that the value of the DSCP uses only six bits of this field. As stated in previous Section, a new traffic marker is required to operate in a color-aware mode to provide an enhanced end-to-end QoS guarantee for flows crossing one or more DiffServ domains. However, the conventional tswTCM doesn't assume a specific mode. In this paper, we propose the enhanced marking mechanism to operate in color-aware mode. For this objective, we define CU field as IM (Initial Marking) field to provide effectively an end-to-end QoS guarantee for flows crossing one or more DiffServ domains as shown in Fig. 1.

To function effectively in a color-aware mode, the initial marking information of a packet is only stored at the first boundary router of DiffServ domain and doesn't be changed at boundary routers of DiffServ domains crossing to the destination. This additional marking information is to have their initial priority level restituted to protect the Packets with high priority from the less important packets when the current domain



< Initial Marking Bits >

Initial Priority	Values of bits IM
No initialization	00
Green	01
Yellow	10
Red	11

Fig. 1. Definition and usage of Initial Marking bits

has the excess bandwidth. The usage of IM bits as follows. If a packet has "00" as the value of IM field, it means that the packet doesn't be marked at other boundary routers. Also, this IM value enables our proposed marking mechanism to support the compatibility with existing markers. That is, because the previous markers don't use the IM field, the value of IM bits is always "00". If a packet has the other values except "00" as the value of IM field, it means that the packet has been marked at other boundary router. These values are utilized to restore their initial priority or to promote the higher priority level when the boundary router of other DiffServ domains has the excess bandwidth. The case of the higher priority promotion of packets is achieved based on the average queue length of the corresponding queue. Therefore, from the new definition and usage of this IM field, our proposed marking mechanism can support the enhanced end-to-end QoS guarantee than the previous marking mechanisms for flows crossing one or more DiffServ domains.

The pseudo-code in Fig.2 through 5 respectively describes the processing of the proposed marking mechanism according to the value of IM bits. The average queue length is calculated based on exponential weighted moving average (EWMA) [9].

```
avg-rate = Estimated Average Sending Rate of Traffic Stream
avg-qlen_n = Estimated Average Queue Length (%) of Queue N
IM = Initial Marking bits

if (IM=="00")  // initial mode
     if (avg-rate <= CIR)
        the packet and IM are marked as green;
    else if (avg-rate <= PIR) AND (avg-rate > CIR)
        calculate P0 = (avg-rate - CIR) / avg-rate
        with probability P0 the packet and IM are marked as yellow;
        with probability (1-P0) the packet and IM are marked as green;
    else
        calculate P1 = (avg-rate - PIR) / avg-rate
        calculate P2 = (PIR - CIR) / avg-rate
        with probability P1 the packet and IM are marked as red;
        with probability P2 the packet and IM are marked as yellow;
        with probability (1-(P1+P2)) the packet and IM are marked as green;
```

Fig. 2. Pseudo-code of proposed marking mechanism when IM field is "00"

The processing of pseudo-code given in Fig. 2 is same as the marking strategy of the conventional tswTCM except that our proposed marking mechanism sets the value of the IM field based on conformance to the contracted throughputs for the traffic flow.

```
14 : if (IM=="01")  // initial marking is green
15 :      if (avg-rate <= CIR)
16 :         the packet is marked as green;
17 :     else if (avg-rate <= PIR) AND (avg-rate > CIR)
18 :         if (q_len_0 <= α1)
19 :             the packet is marked as green;
20 :         else
21 :             calculate P0 = (avg-rate - CIR) / avg-rate
22 :             with probability P0 the packet is marked as yellow;
23 :             with probability (1-P0) the packet is marked as green;
24 :     else
25 :         if (q_len_0 <= α2)
26 :             the packet is marked as green;
27 :         else
28 :             calculate P1 = (avg-rate - PIR) / avg-rate
29 :             calculate P2 = (PIR - CIR) / avg-rate
30 :             with probability P1 the packet is marked as red;
31 :             with probability P2 the packet is marked as yellow;
32 :             with probability (1-(P1+P2)) the packet is marked as green;
```

Fig. 3. Pseudo-code of proposed marking mechanism when IM field is "01" (α1>α2)

The processing of pseudo-code shown in Fig. 3 is same as the marking strategy of the previous tswTCM except line 18-19 and 25-26. Also, this processing part is performed in boundary routers of other DiffServ domains crossing after a packet is marked at the first boundary router. In case that line 17 is true, the existing tswTCM downgrades the corresponding packet into Yellow. However, our proposed marking mechanism enables the packet to be marked to Green when the average queue length of the Green queue is smaller than α1. Line 25-26 operates similarly with line 17-18. In addition, pseudo-codes depicted in Fig. 4 and 5 are processed likewise. The specific values of average queue lengths (α1~α6, β1~β3) are handled in Section 5.

```
33 : if (IM=="10")       // initial marking is yellow
34 :          if (avg-rate <= CIR)
35 :         if (q_len_0 <= α3)
36 :             the packet is marked as green;
37 :         else
38 :             the packet is marked as yellow;
39 :     else if (avg-rate <= PIR) AND (avg-rate > CIR)
40 :         if (q_len_0 <= α4)
41 :             the packet is marked as green;
42 :         else
43 :             calculate P0 = (avg-rate - CIR) / avg-rate
44 :             with probability P0 the packet is marked as red;
45 :             with probability (1-P0) the packet is marked as yellow;
46 :     else
47 :         if (q_len_1 <= β1)
48 :             the packet is marked as yellow;
49 :         else
50 :             calculate P1 = (avg-rate - PIR) / avg-rate
51 :             with probability P1 the packet is marked as red;
52 :             with probability (1-P1) the packet is marked as yellow;
```

Fig. 4. Pseudo-code of proposed marking mechanism when IM field is "10" (α3>α4)

In summary, the proposed mechanism is an extension of the conventional tswTCM. The key concept of our proposed mechanism is that we define unused CU field as IM field to store the initial marking information and use this information combined with the average queue length for promotion of packets. The value of IM bits is only marked at the first boundary router. This functionality is necessary for the restitution of the initial priority level of the packets to guarantee effectively an end-to-end QoS in color-aware mode and protects higher priority packets from being downgraded in other DiffServ domains. The proposed marking mechanism operates in the following manner: the boundary routers checks IM bits, and then if the bit is "00", the router marks DSCP and IM bits into its priority level according to the traffic conformity. Otherwise, it remarks DSCP bits based on the previous DSCP bits and IM bits. In addition, the proposed marking mechanism operates in two internal modes: the initial mode and remarking mode. The initial mode is performed whenever the IM bits are "00", its algorithm is equal to that of tswTCM. Otherwise, the remarking mode is operated, and the remarking of DSCP field about priority level is determined by the comparison of the previous priority level and the IM bits.

```
if (IM=="11")        // initial marking is red
    if (avg-rate <= CIR)
        if (q_len_0 <= α5)
            the packet is marked as green;
        else if (q_len_1 <= β2)
            the packet is marked as yellow;
        else
            the packet is marked as red;
    else if (avg-rate <= PIR) AND (avg-rate > CIR)
        if (q_len_0 <= α6)
            the packet is marked as green;
        else if (q_len_1 <= β3)
            the packet is marked as yellow;
    else
        the packet is marked as red;
```

Fig. 5. Pseudo-code of proposed marking mechanism when IM field is "11" ($\alpha5>\alpha6$, $\beta2>\beta3$)

5 Simulation Topology and Configuration

Simulations have been performed using the ns-2 network simulator [10] and have been conducted using the topology depicted in Fig. 6 to study the performance of the proposed marking mechanism. With this topology we test the effectiveness and ad-

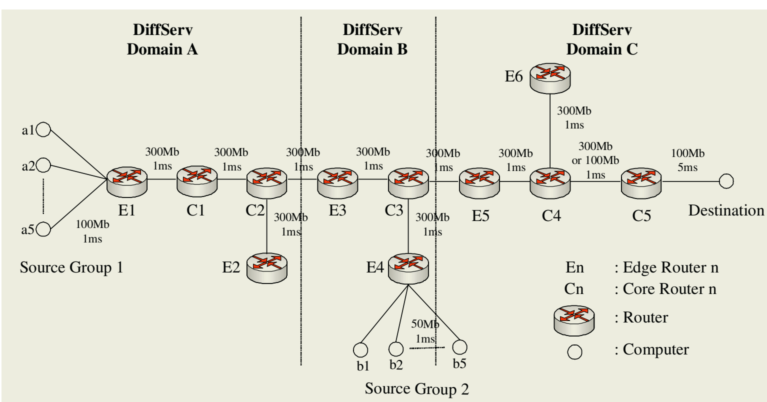


Fig. 6. Simulation network topology

vantage of using the proposed marking mechanism at the ingress router of DiffServ networks. The simulation configuration parameters are listed in Table 1 [11].

Table 1. Simulation configuration parameters

Parameter			Value
Test Time			10 sec
Packet Size			1000 bytes
Transmit Rate	a1~a5		30 or 50 Mbps
	b1~b5		15 Mbps
tswTCM	a1~a5	CIR	5 Mbps
		PIR	15 Mbps
	b1~b5	CIR	2.5 Mbps
		PIR	7.5 Mbps
MRED (RIO-D)	DP 0 min_th		100 packets
	DP 0 max_th		200 packets
	DP 0 min_Pb		0.02
	DP 1 min_th		50 packets
	DP 1 max_th		100 packets
	DP 1 min_Pb		0.1
	DP 2 min_th		25 packets
	DP 2 max_th		50 packets
	DP 2 min_Pb		0.2
Traffic (Exponential & Pareto)	Burst Time		300ms
	Idle Time		300ms
Edge & Core Router	Buffer Size		200 packets

Two UDP source groups (a1~a5, b1~b5), comprising two pareto distribution traffic sources, two exponential distribution traffic sources, and one CBR (Constant Bit Rate) source each, perform unidirectional data transmissions across links (from E1 or E4 to C5) to one corresponding destination. C1~C5 are core routers which implement RIO-D (RIO-Decoupled) [12]. And, simulations are executed using one or two UDP source groups according to test scenarios. For the conventional tswTCM and the proposed marking mechanism, PIR (Peak Information Rate) is set to CIR (Committed Information Rate) * 300%. The values that were used throughout the simulations do not necessarily correspond to an optimal configuration. However, from various simulation runs not described here, using a wide range of parameters, this configuration was found to be suitable for evaluating the proposed marking mechanism. Also, in our simulation, the specific values of average queue lengths (α1~α6, β1~β3) are shown in Table 2. In addition, these values can be adjusted from DiffServ domain administrators dynamically. The variation of parameters and their impact on the performance of the proposed marking mechanism under a number of different conditions, which will allows us to create more adaptive version of the proposed marking mechanism, is left as future work.

Table 2. Specific values of average queue lengths (α1~α6, β1~β3)

Constant	Value (%)	Constant	Value (%)	Constant	Value (%)
α1	50	α4	15	β1	5
α2	30	α5	30	β2	12.5
α3	20	α6	20	β3	7.5

6 Simulation Results and Discussion

In this section, we will compare the performance of our proposed marking mechanism, which is a color-aware mode, with the conventional tswTCM using three scenarios. The results are presented by statistics based on simulations. We list the notations used in the figures as follows.

- TR = Number of Total Received packets
- TT = Number of Total Transmitted packets
- G, Y or R = Number of transmitted Green, Yellow, or Red packets

A. Scenario 1: Source group 1 has 5 flows (each 30Mbps) and source group 2 has no flows. No bottleneck (A-1) and bottleneck (A-2) in C4-C5 link

In this scenario, as all bandwidth of each link between routers is sufficiently large, a packet doesn't be dropped at the low or middle provision level (let's say under 70%). Fig. 7 shows time versus received or transmitted packet numbers evaluated in C4-C5 link for the proposed and conventional tswTCM. An edge router (E1) meters and marks equally packets in two mechanisms. However, edge routers, E3 and E5, in other domains, marks differently as described in Section 4. From Fig. 7, we can see that many yellow packets are promoted to the higher priority level in our proposed mechanism when the boundary routers of other DiffServ domains have the excess bandwidth. Results show that our proposed mechanism is more efficient than the conventional tswTCM under medium network provision level in terms of an end-to-end QoS guarantee.

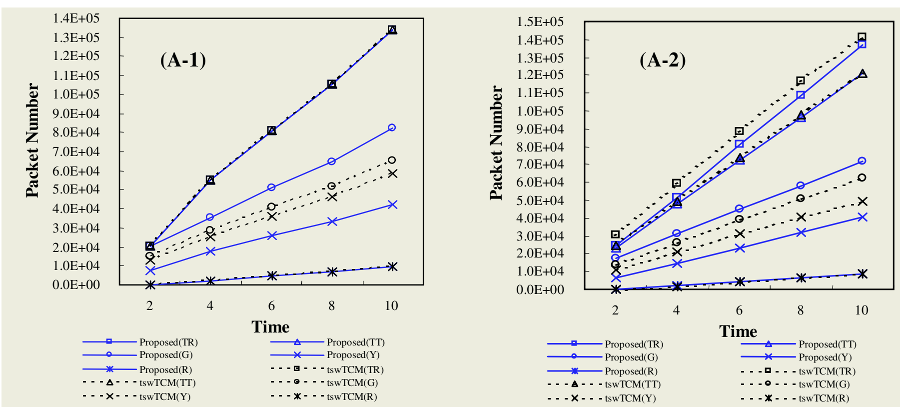


Fig. 7. Time versus received or transmitted packets. Source group 1 has 5 flows (each 30Mbps) and source group 2 has no flows. (A-1): no bottleneck in C4-C5 link (300Mbps), (A-2): bottleneck in C4-C5 link (100Mbps)

B. Scenario 2: Source group 1 has 5 flows (each 50Mbps) and source group 2 has no flows. No bottleneck (B-1) and bottleneck (B-2) in C4-C5 link

In this scenario, the throughput of two mechanisms is approximately 50% due to bottleneck link of C4-C5 (B-2). We can see that our proposed mechanism supports enhanced end-to-end QoS guarantee under the high provision level (let's say above 70%).

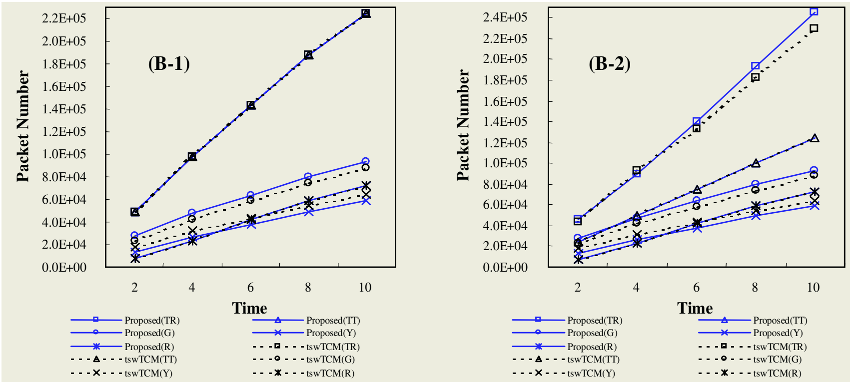


Fig. 8. Time versus received or transmitted packets. Source group 1 has 5 flows (each 50Mbps) and source group 2 has no flows. (B-1): no bottleneck in C4-C5 link (300Mbps), (B-2): bottleneck in C4-C5 link (100Mbps)

C. Scenario 3: Source group 1 has 5 flows (each 30Mbps) and source group 2 has 5 flows (each 15Mbps). No bottleneck (C-1) and bottleneck (C-2) in C4-C5 link

This scenario is same as scenario 1 except that the UDP traffics of source group 2 is generated between 3sec and 8sec. The results from this scenario show that many yellow packets is promoted to the higher priority level and prove that our proposed mechanism is operated more adaptively than the conventional tswTCM under a dynamic traffic environment.

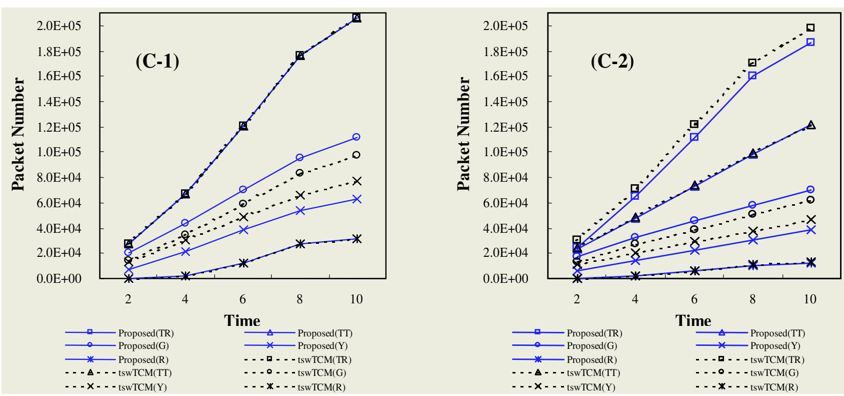


Fig. 9. Time versus received or transmitted packets. Source group 1 has 5 flows (each 30Mbps), and source group 2 has 5 flows (each 15Mbps). (C-1): no bottleneck in C4-C5 link (300Mbps), (C-2): bottleneck in C4-C5 link (100Mbps)

From extensive experiments not depicted graphically, it was found that our proposed marking mechanism performs better than the conventional tswTCM in most scenarios and the parameters for simulation are sensitive to the traffic source characteristics and link rates. In addition, the settings of RED-related threshold values and queue sizes are especially affected. The choice of CIR and PIR was also found to impact the results in our simulation. These parameters can be set adequately from DiffServ domain administrator to support more enhanced end-to-end QoS.

7 Conclusion and Future Work

We have proposed in this work an enhanced marking mechanism based on a color-aware mode as an extension of tswTCM for flows crossing one or more DiffServ domains in order to guarantee effectively enhanced end-to-end QoS. This mechanism is fully compatible with other markers and very scalable. Using simulation, we have shown that the proposed marking mechanism can improve an end-to-end QoS guarantee in multiple DiffServ environment.

As future work, we will develop more adaptive version of the proposed marking mechanism. Also, we will estimate and analyze this mechanism through the implementation on the IXP2400 network processor [13], [14] and devise improved marking mechanism to enforce fairness among different flows originated from the same subscriber network in DiffServ domain.

References

1. Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and Weiss, W.: An architecture for differentiated services. RFC 2475 (1998)
2. Nichol, K., Blake, S., Baker, F., and Black, D.: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers. RFC 2474 (1998)
3. Jacobson, V., Nichols, K., and Poduri, K.: An expedited forwarding PHB. RFC 2598 (1999)
4. Heinanen, J., Baker, F., Weiss, W., and Wroclawski, J.: Assured forwarding PHB group. RFC 2597 (1999)
5. Clark, D. and Fang, W.: Explicit Allocation of Best Effort Packet Delivery Service, IEEE/ACM Transactions on Networking, Vol. 6. No. 4 (1998) 362-373
6. Heinanen, J. and Guerin, R.: A Single Rate Three Color Marker. RFC 2697 (1999)
7. Heinanen, J. and Guerin, R.: A Two Rate Three Color Marker. RFC 2698 (1999)
8. Fang, W., Seddigh, N., and Nandy, B.: A time sliding window three color marker (tswTCM). RFC 2859 (2000)
9. Floyd, S. and Jacobson, V.: Random early detection gateways for congestion avoidance, IEEE/ACM Transactions on Networking, Vol. 1 (1993) 397-413
10. Network simulator 2 (ns-2). University of California at Berkeley, CA, Available via http://www.isi.edu/nsnam/ns/ (2003)
11. Andrikopoulos, I., Wood, L., and Pavlou, G.: A fair traffic conditioner for the assured service in a differentiated services Internet. ICC 2000, Vol. 2, IEEE International Conference (2000) 18-22
12. Seddigh, N., Nandy, B., Pieda, P., Hadi Salim, J., and Chapman, A.: An experimental study of Assured services in a Diffserv IP QoS Network, Proceedings of SPIE symposium on QoS issues related to The internet, Boston (1998)
13. IXP2400 and IXP2800Network Processor Programmar's Reference Manual, Intel Corporation (2004)
14. IXP2400 Network Processor Hardware Reference Manual, Intel Corporation (2004)