An Effective Remarking Scheme for Diffserv AF Service through Multiple Domains

Shoichi Motohisa*, Hiroyuki Fukuoka†, Ken-ichi Baba‡ and Shinji Shimojo‡
* Graduate School of Information Science and Technology, Osaka University
5–1 Mihogaoka, Ibaraki, Osaka 567–0047, Japan
Tel: +81–6–6879–8795, Fax:+81–6–6879–8794
Email: hisa@ist.osaka-u.ac.jp
† Telecommunications Advancement Organization of Japan
Email: fukuoka@ais.cmc.osaka-u.ac.jp
‡ Cybermedia Center, Osaka University
Email: {baba, shimojo}@cmc.osaka-u.ac.jp

Abstract— **This paper proposes a new packet re-marking scheme that can improve per-flow Quality of Service (QoS) of Assured Forwarding (AF) service traversing multiple domains of Differentiated Services (Diffserv) networks. The base concept of the scheme is to distinguish packets re-marked to out-of-profile at the domain boundaries from those already marked as out-of-profile at the time of entering the network, and to give chances to the re-marked packet to recover back to in-of-profile that can enjoy its rightful QoS within the networks. Basic performance of the proposed scheme is evaluated through simulation study, and the results show its effectiveness in preserving QoS of the inter-domain flows.**

I. Introduction

Differentiated Services (Diffserv)[1] is a type of architecture aimed at providing Quality of Service (QoS) in IP networks. In contrast to the Intserv architecture that requires frequent message exchange and per-flow state management within the core network, Diffserv architecture has the advantages of simplicity and scalability.

In Diffserv networks, service differentiation is provided based on the Diffserv Code Point (DSCP)[2] field in the IP header; packets with the same DSCP are handled under corresponding forwarding discipline called Per-Hop Behavior (PHB). The two basic PHBs currently defined in IETF are Expedited Forwarding (EF)[3] and Assured Forwarding (AF)[4] PHBs. The EF PHB is used to provide premium "Virtual Leased Line" type of services. It is suitable for real-time applications that demand higher QoS such as low loss, low delay, low jitter and assured bandwidth. On the other hand, AF PHB is used to provide more elastic type of services.

The AF service class is a framework to provide minimum bandwidth guarantee for each traffic flow by introducing drop precedence property marked in DSCP field on each packet. The classification and marking on each packet are carried out at the ingress router based on conformance to the contracted throughputs for the traffic flow. Unconformable packets are marked out-of-profile (OUT) at the ingress router while conformable packets are marked in-profile (IN). At the time of congestion in the core networks, OUT packets are more likely to be dropped than IN packets by the result of differentiated PHB, and thus the contracted throughput for IN packets are to be maintained even during congestion.

The situation is slightly different in a multiple domain environment. When AF service flow traverses more than one Diffserv domain, at the boundary router, each packet is verified if it is conformable with the agreed contract on the inter-domain traffic aggregate. At this time, an IN packet may possibly be recognized as unconformable by the boundary router and re-marked to OUT[5]. This is due to the microscopic fluctuating characteristics of the aggregated flow such as jitters that may not fit well with the re-marking algorithm adapted at the boundary router. This re-marking behavior at the boundary router may cause undesirable drop of packets within distant congested domains downstream even though they were originally marked as IN. And this may result in the failure of QoS assurance of the corresponding end-to-end traffic flow[6].

The base concept of the proposed scheme is to distinguish packets re-marked to OUT at the domain boundary from those already marked as OUT at the time of entering the network, and to give chances to recover back to the IN packets that can enjoy its rightful QoS within the networks.

By the conventional scheme, the packets re-marked to OUT and the packets marked as OUT from the beginning are treated equally in terms of packet dropping behavior. This may cause unfairness between inter-domain flows containing re-marked packets and intra-domain flows containing no re-marked packets with regard to the delivery of conformable packets on the end-to-end basis. The proposed scheme gives a solution to this problem imposing no additional cost for the packet handling.

The rest of the paper is organized as follows. Section II describes the target network model dealt in this paper. Then, the proposed scheme of packet re-marking is detailed in Section III. Performance evaluation of the proposed scheme is presented in Section IV. Simulation results shows that the proposed scheme outperforms the conventional scheme in reducing the degradation of QoS of inter-domain traffic flows. Section V concludes the paper.

0-7803-7978-0/03/$17.00

II. TARGET NETWORK

Fig. 1 shows our target network model that provides AF service under multiple domains environment. When a packet from Source arrives at the edge router of Domain A, the edge router meters the packets, and judges whether its arrival rate is within a contract rate between the Source and Domain A. If the arrival rate is within the contract rate, it marks the packet with IN, otherwise marks excess packets with OUT. Then, the packet is forwarded to the adjacent core router.

When the packets come to the boundary of Domains A and B, they are again metered by the ingress edge router of Domain B, and classified into IN and OUT this time based on the conformance with the contract rate at the aggregation level between Domains A and B. Excess IN packets are re-marked to OUT. Packets are handled similarly at the subsequent domain boundaries.

Within each domain, at the time of congestion, core routers begin to drop OUT packets first to avoid loss of IN packets. This behavior is achieved by the RIO (RED with IN and OUT)[7] mechanism, where two sets of RED[8] parameters, one for IN packets and one for OUT packets, are defined. The maximum queue length and packet drop probability parameters for OUT packets are set with more stringent values than those for IN, so that OUT packets are dropped easier when the queue length grows at the time of congestion, while IN packets are queued successfully. Note that packets are routed to the same queue regardless of its marking, IN or OUT, which will preserve the sequence of packets within a flow.

As for the metering algorithm at each ingress router, we assume use of TSW (Time Sliding Window)[9] algorithm is assumed for metering packet arrival rate. TSW calculates a packet arrival rate $rate_n$ by the following formulas.

$$rate_n = \frac{rate_{n-1} \times Itvl + size}{t_n - t_{n-1} + Itvl}$$

$rate_n$: The packet arrival rate

$size$: The packet size

t_n : The time of the current packet arrival

$Itvl$: Time window over which history is kept

$Itvl$ is a constant parameter defined in each domain. Ref. [9] recommends to set $Itvl$ with 1second.

Based on the above algorithm, packets are metered at domain boundaries. And if it detects excess rate of IN packets over contracted rate, which may due to fluctuation caused by traffic aggregation or surge of other traffic flows sharing the same inter-domain link, those excess packets are re-marked to OUT. Those re-marked packets are treated with equal priority as other OUT packets in the following domains, and may be dropped if they encounter congestion within subsequent domains (Fig. 2). And this may result in the failure of rate assurance of IN packets.

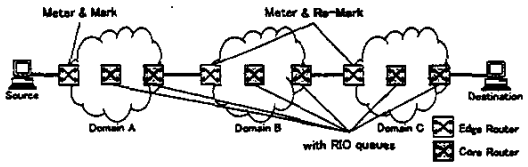

Fig. 1. Network model.

III. PROPOSED RE-MARKING SCHEME

In order to solve this problem, this paper proposes a new re-marking scheme to be used at each domain boundary. The proposed method distinguishes packets that are re-marked to OUT at domain boundaries from the packets marked as OUT from the beginning at the edge router between the user and the first domain. And it allows re-marked OUT packets recover back to IN if there is a room left within contract rate at the subsequent domain boundaries. This contributes to preserve IN packets within the flow, and to provide desired QoS on the end-to-end basis.

The proposed scheme uses all three drop precedence code points which AF service can use at maximum; Green and Red corresponds to IN and OUT, respectively, and Yellow is the newly assigned code point for packets re-marked from Green at inter-domain boundaries. The marking at the ingress edge router between the user and the first domain is similar to the conventional IN/OUT scheme; it meters the arrival of packets, and marks the packet with Green if it complies with the contracted rate, and marks the packet with Red if it exceeds the contracted rate between the user and the domain.

The ingress border router at the adjacent domain meters the packets based on the conformance with the contract rate of the aggregated link between domains, where packets of different flows from various domains may be aggregated within. Here, Green packets and Yellow packets are treated equally, i.e. the border router assumes both types of marking are the indication of in-of-profile when they arrive. The border router meters total amount of Green and Yellow packets, and marks packets within the contract rate with Green and excess packets with Yellow.

At the core routers in each domain, Yellow packets are distinguished from Green packets in terms of dropping precedence. In principle, all three colors can be assigned with different RED dropping precedence parameters, but in this paper, we assign Yellow and Red packets with the same RED parameters for the ease of comparison with the conventional IN/OUT re-marking scheme.

In the conventional IN/OUT marking scheme, the packet re-marked to OUT at domain boundaries on the path are undistinguishable from the packets marked with OUT from the beginning at the ingress edge router of the first domain. This will cause undesired drop of packets that was originally IN and accounted for the assured bandwidth of the flow, if the flow encounter substantial congestion at domain boundaries along the path.

In the proposed method, on the other hand, by the use of the third code point exclusively used to indicate occurrence of re-marking at domain boundaries, these packets are distinguishable. If a domain boundary is congested, similar amount

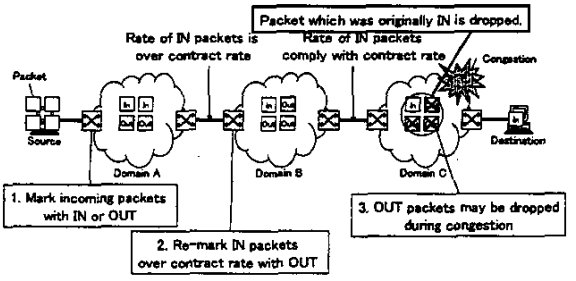

Fig. 2. Conventional Scheme.

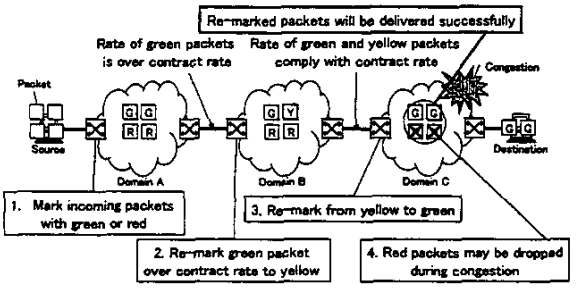

Fig. 3. Proposed Scheme.

of packets will be re-marked as in the case of conventional IN/OUT scheme, but those re-marked packets can be recovered back to in-of-profile at subsequent domain boundaries if it complies with their contract rates (Fig. 3).

IV. Evaluation of the Proposed Scheme

We performed simulations in which the traffic is UDP in order to evaluate the basic performance of a proposed scheme. We used Network Simulator version 2 (ns-2) as a simulator. First, we investigated the number of packets re-marked on the domain boundary. Next, we investigated the probability of packet loss of the flow which transits multiple domains. Finally, we investigated the fairness between an inter-domain flow and an intra-domain flow.

A. Simulation model

Fig. 4 shows the simulation model. The propagation delay of each link is 1ms and the bandwidth is 1Gbps. As the bandwidth of a link is large, a packet does not drop except for the Bottleneck Link. An edge router meters and marks packets. It adopts the TSW algorithm denoted in Sec. II for metering and marking. A core router consists of single RIO queue. The buffer size of the queue is 200packets. The parameters of a RIO queue are $(min_{in}, max_{in}, Pmax_{in}) = (100, 150, 0.02)$, $(min_{in}, max_{in}, Pmax_{in}) = (50, 100, 0.1)$ used in Ref. [5][6]. In this model, we use UDP traffic. The maximum size of packets is 1500bytes, and average size is 1000bytes. Packets arrive according to a Poisson process. The contract rate at each domain boundary is 20Mbps.

B. The number of re-marking packets

Flow a of Fig. 4 is UDP traffic with 1Mbps and consists of 20flows. Flows b, c and d, have no traffic, that is, they do

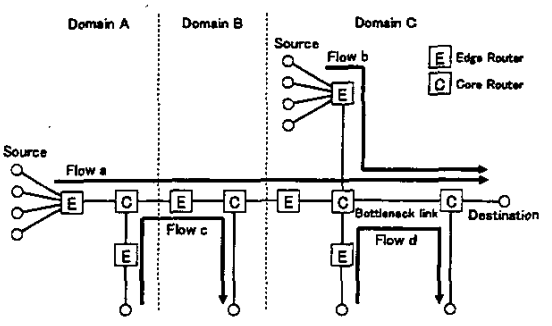

Fig. 4. Simulation Model.

not exist. As the contract rate between Source and Domain A is large, all packets that Source sends are marked with IN or Green. We investigated the number of packets marked at the boundary router of Domains B and C.

TABLE I shows the result of this simulation. By the conventional scheme, the number of OUT packets increases at the ingress routers of both Domains B and C as packets run through from Domain A to C. However, by the proposed scheme, as a Yellow packet may return to a Green packet at ingress of Domain C, the number of Yellow packets does not change so much in Domains B and C.

C. Packet loss probability

In this section, Flow a is UDP traffic with 1Mbps and consists of x flows. In Flow b, traffic does not exist. Flows c and d are UDP traffic flows with 5Mbps and 25Mbps, respectively. We assume that the contract rate of Flows a and c is fully large and the contract rate of Flow d is 0Mbps, so that all packets of Flows a and c are marked with IN or Green, and all packets of Flow d are done with OUT or Red at each ingress. And the bandwidth of Bottleneck link at Domain C is 25Mbps. We investigated the probability of packet loss of Flow a.

Fig. 5 shows the probability of packet loss when setting the number of flows to X-axis. We denoted the result when the contract rate between domains is large enough and packets are not re-marked as "Re-marking Disabled". This result shows a performance limit. In this figure, IN packets which arrive between Domains A and B exceeds domain contract rate 20Mbps when x of Flow a is 15 or more, so that IN packets which arrive between Domains A and B are re-marked to OUT at $x = 15$ or more. By the conventional scheme, as the re-marked OUT packets between Domains A and B are

TABLE I
THE NUMBER OF PACKETS AT DOMAIN INGRESS.

	DSCP	Domain B ingress		Domain C ingress	
Conventional Scheme	Total	7486073	(100%)	7486073	(100%)
	IN	7450456	(99.524%)	7436151	(99.333%)
	OUT	35617	(0.476%)	49922	(0.667%)
Proposed Scheme	Total	7494078	(100%)	7494078	(100%)
	Green	7455264	(99.482%)	7455520	(99.485%)
	Yellow	38814	(0.518%)	38558	(0.515%)
	Red	0	(0%)	0	(0%)

dropped at the Bottleneck link in Domain C, the probability of packet loss of Flow a is high. By the proposed scheme, however, the Yellow packets re-marked from Green packets between Domains A and B are again re-marked to Green from Yellow between Domains B and C. Because the contract rates between Domains B and C are 20Mbps. As most packets which arrives at the "Bottleneck link" when $x = 20$ or less are Green packets, the graph of a proposed scheme is almost the same as the graph of "Re-marking Disabled." The probability of packet loss of a proposed scheme is lower than that of the conventional scheme when x is more than 20. The proposed scheme is more effective than the conventional scheme.

D. Fairness between inter-domain flow and intra-domain flow

In this section, we investigate fairness between inter-domain flow (Flow a) and intra-domain flow (Flow b). Flows a and b are UDP traffic with 1Mbps and consist of x flows. Flows c and d are UDP traffic flows with 10Mbps and 30Mbps, respectively. As the contract rates of Flows a and c are large, all packets are marked on IN or Green. As the contract rate of Flow d is 0Mbps, all packets are marked with OUT or Red. The Bottleneck link in Domain C is 30Mbps. We investigate the probability of packet loss of Flows a and b.

Fig. 6 shows the probability of packet loss when setting the number of flows to X-axis. By the conventional scheme, an inter-domain flow does not differ from the number of intra-domain flow when x is less than 10. However, the probability of packet loss of an inter-domain flow is higher than that of intra-domain flow when $x \geq 10$. As IN packets which arrive at the boundary between Domains A and B when $x \geq 10$ exceeds a domain contract rate of 20Mbps, IN packets are re-marked to OUT packets between Domains A and B. As OUT packets are dropped at the Bottleneck link, the probability of packet loss of an inter-domain flow is higher than that of an intra-domain flow. By the proposed scheme, as the domain contract rates between Domains B and C are 20Mbps, the probability of packet loss of an inter-domain flow does not differ from that of an intra-domain flow when $x \geq 20$, and fairness between flows is kept. For example, at $x = 15$, the probability of packet loss of an inter-domain flow is 1.784×10^{-1}, and that of an intra-domain flow is 2.494×10^{-3} by conventional scheme. The difference is about 10^2 order. By the proposed scheme, that of an inter-domain flow is 1.529×10^{-2}, and that of an intra-domain flow is 1.512×10^{-2}. The probability of packet loss does not have a difference between inter-domain and intra-domain, and the proposed scheme keeps fairness. The probability of packet loss of an inter-domain flow is higher than that of an intra-domain flow while $x > 20$ by the proposed scheme. The difference of the probability of packet loss between the flows by the proposed scheme is smaller than it by the conventional scheme.

V. Conclusion

This paper proposed a new re-marking scheme for Diffserv with multiple domains environment that can preserve end-to-end QoS of the AF service flows traversing multiple domains.

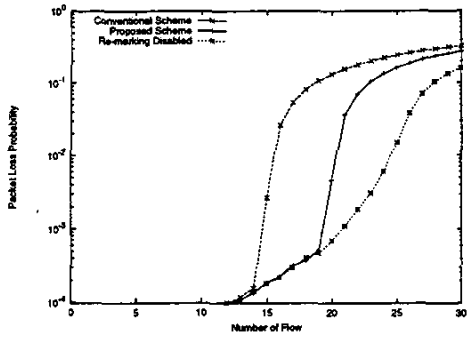

Fig. 5. Packet Loss Probability.

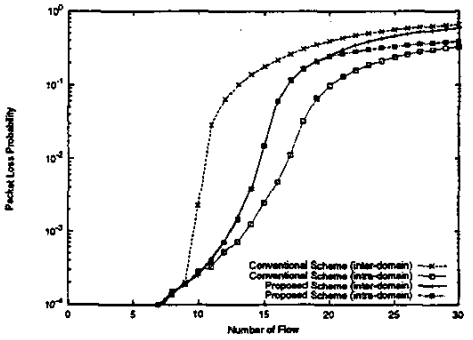

Fig. 6. Packet Loss Probability of inter-domain and intra-domain.

Basic performance of the proposed method was explored through simulation studies, and the results show that the proposed scheme is effective for suppressing QoS degradation of inter-domain AF flows under network congestion and for decreasing unfairness of inter-domain flows against intra-domain flows.

Performance of the proposed method will be examined more in detail in the further studies, which include the influence of flow control behavior of TCP traffic that the AF service class is mainly intended for.

References

[1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Services", *RFC 2475*, Dec. 1999.

[2] K. Nichols, S. Blake, F. Baker and D. Black, " Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", *RFC 2474*, Dec. 1998.

[3] V. Jacobson, K. Nichols and K. Poduri, "An Expedited Forwarding PHB", *RFC 2598*, June 1999.

[4] J. Heinanen, F. Baker, W. Weiss and J. Wroclawski, "Assured Forwarding PHB Group", *RFC 2597*, June 1999.

[5] W. Fang, "The "Expected Capacity" Framework: Simulation Results", *Princeton University Technical Report*, TR-601-99, Jan. 1998.

[6] K. Kumazoe, Y. Hori, T. Ikenaga and Y. Oie, "Quality of Assured Service through Multiple DiffServ Domains", *IEICE Transactions on Information and Systems*, Vol.E85-D, No.8, pp.1226-1232, Aug. 2002.

[7] D. D. Clark, W. Fang, "Explicit Allocation of Best Effort Packet Delivery Service", *IEEE/ACM Transactions on Networking*, Vol. 6, No. 4, pp. 362-373, Aug. 1998.

[8] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, pp. 397-413, Aug. 1993.

[9] W. Fang, N. Seddigh and B. Nandy, "A Time Sliding Window Three Colour Marker (TSWTCM)", *RFC 2859*, June 2000.