Adaptive Marking for Assured Forwarding Service

Ikjun Yeom, *Member, IEEE* and A. L. Narasimha Reddy, *Senior Member, IEEE*

Abstract— **This letter looks at the problem of achieving specific QoS goals of individual flows by flexibly managing resources available to an aggregated source. We present an adaptive marker based on a TCP performance model within a Diffserv network, and show that the aggregated marker can maintain state of individual flows at the edge of the network and utilize this state effectively in adaptively marking packets of individual flows to meet their QoS goals.**

Keywords— **Differentiated service, Aggregation, Packet marking strategies**

I. INTRODUCTION

PACKET marking for Assured Forwarding (AF) Service [1][1] in the Differentiated Services (diffserv) network [2] has become an important issue since it enables customers to achieve their performance goals while providers control traffic load in each class. There have been a number of packet markers proposed to realize specific goals such as a target bandwidth of individual flows or fair bandwidth allocations among aggregated flows [3], [4], [5].

In this letter, we address a more general problem: given individual target rates for each flow, how to allocate a fixed contract rate among flows within an aggregation under dynamic network conditions? We propose an adaptive marker that relies on the TCP performance model developed by us in [6]. This letter makes the following significant contributions: (1) establishes the relationship between per-session behavior, aggregate packet marking and packet differentiation within a diffserv network. (2) presents simulations using ns-2 [7] to study the behavior of individual flows followed by an aggregated marker. (3) proposes an adaptive marking scheme for aggregated flows and evaluates it through extensive simulations.

II. BEHAVIOR OF PROPORTIONAL MARKING

We first describe the problem of packet marking when a marker does not maintain per-flow state. We call this marker a *proportional marker* since a packet is marked by aggregated sending rate, and consequently, contract rate[2] is distributed to individual flows proportional to their sending rates. Suppose that there are flows sharing aggregated contract rate, M $(= \sum_{i=1}^{n} m_i)$, and the current total sending rate of the flows is B $(= \sum_{i=1}^{n} b_i)$, where m_i and b_i are the individual marking and sending rate of i^{th} flow, respectively, and n is the number of flows. When B is less than M, every packet is marked IN. If B is greater than M, then a packet is marked IN with a probability of M/B, and we have $m_i = \frac{M}{B} b_i$.

Now we present a summary of a throughput model for TCP flows sharing a contract rate originally proposed in [6]. In this model, we assume that a TCP receiver does not employ delayed-ACK and IN-profile packet is not dropped[3].

First, we begin with a simple throughput model for an individual TCP flow in a diffserv network in [6].

$$b_i = \frac{3}{4} m_i + \frac{3k_i}{4\mathrm{RTT}} \sqrt{\frac{2}{p_i}} \qquad (1)$$

where p_i is the drop probability of OUT packet, and k_i is the packet size. Then, the sum of b_i, B is given by

$$B = \sum^{i} b_i = \frac{3}{4} M + \sum^{i} \frac{3k_i}{4\mathrm{RTT}_i} \sqrt{\frac{2}{p_i}} \qquad (2)$$

Let $c_i = \frac{k_i}{\mathrm{RTT}_i}\sqrt{\frac{2}{p_i}}$, and using (1) and $m_i = \frac{M}{B} b_i$, b_i is expressed by

$$b_i = \frac{3M}{4B} b_i + \frac{3}{4} c_i = \frac{3 c_i B}{4B - 3M} \qquad (3)$$

Applying (2), finally we have

$$b_i = \frac{3c_i}{4\sum_{i=1}^{n} c_i} M + \frac{3}{4} c_i \qquad (4)$$

So far, we have presented a simple model for aggregated flows in a diff-serv network. This model reveals that throughput of an individual TCP flow b_i is affected by other flows (more precisely $\sum c_i$). This property of the proportional marking may cause unstable state of individual TCP flows and degrade QoS consequently.

III. ADAPTIVE MARKING FOR AGGREGATED FLOWS

In this section, we discuss how to control individual throughputs with marking rates within an aggregation and propose an adaptive marking strategy. The most desirable situation is clearly to guarantee individual target rates for all the individual flows. However, it is also clear that there exist situations in which some targets cannot be reached: (i) When there is a severe congestion along the path and the current available bandwidth is less than the target. (ii) When the contract rate is not enough to achieve the target.

This work was supported in part by a Texas ATP grant and by an NSF Career Award and by a gift from EMC Corp.

I. Yeom is with the Department of Computer Science, Korea Advanced Institute of Science and Technolgy, Daejeon, South Korea (yeom@cs.kaist.ac.kr).

A. L. N. Reddy is with the Department of Electrical Engineering, Texas A & M University, College Station, TX 77843-3128 USA (reddy@ee.tamu.edu).

[1] In AF service, there are four classes defined, but we focus on a single class. Our work is easily expandable to multiple classes since packets in a class are forwarded independently from packets in another class.

[2] In this letter, we use *contract rate* to mean profile rate contracted by aggregated flows.

[3] In a diff-serv domain configured appropriately, IN-profile packet is expected to be transmitted without being discarded.

If we try to achieve the target by increasing marking rate of an individual flow observing the case (i), it makes the congestion more severe and results in resource wastage. This is an undesirable situation for both customers and service providers.

We here propose an adaptive marking scheme for aggregated flows which guarantees at least one of the following to all the aggregated flows: (1) realizes individual target rate, (2) maximizes throughput without IN packet loss when individual targets cannot be met and (3) allocates marking rate fairly when all targets can be met.

Initially, we set the marking rate of each flow proportional to its target. If every flow gets throughput more than their marking rate without IN packet loss, then the adaptive marker works as a weighted IN-fair marker. On the other hand, if the network path of a flow is oversubscribed[4] and observes IN packet losses (resource wastage), the adaptive marker adjusts marking rates of individual flows in order to avoid IN packet loss (achieving the second objective).

However, it is not easy for a marker to find whether a flow observes an oversubscribed network or not unless the marker is combined into the sender. For marking of aggregated flows, the marker cannot be combined into an individual sender. Thus, it can be just estimated from the current throughput. To estimate the current condition of a flow, we use throughput model proposed in [6]. From the model, throughput B of a TCP flow experiencing oversubscribed network is given by

$$B = \min\{\frac{3}{4}m, \frac{k}{\mathrm{RTT}}(\sqrt{\frac{1}{9} + \frac{8}{3p_{in}}} - \frac{1}{3})\} \qquad (5)$$

where m is the contract rate of the flow (or the IN-marking rate), k is the packet size, and p_{in} is the probability of IN packet loss. From (5), when throughput achieved by a flow is less than $0.75m$, the flow should observe an oversubscribed network.

Therefore we classify a flow into one of the following three states and treat these states differently. Here, t_i is the target rate of i^{th} individual flow, m_i is the marking rate, and b_i is the realized throughput. The target rate can be specified by the individual users, and $\sum m_i$ is the contract rate for the aggregation.

When $b_i \leq 0.75m_i$, the flow observes oversubscribed network, and hence to avoid wasting resources m_i is adjusted such that $b_i \geq 0.75m_i$. When $0.75m_i < b_i < t_i$, m_i is increased such that the flow may reach its target t_i. When $b_i \geq t_i$, target is reached, m_i can be reduced to meet other goals. Fig. 1 shows an algorithm for the adaptive marker.

Theorem: The adaptive marking algorithm finds m_i for which $b_i \geq \min\{t_i, b_{a,i}, b_{b,i}\}$ for $1 \leq i \leq n$ where t_i is the target rate, $b_{a,i}$ is maximum achievable rate such that $b_i \geq 0.75m_i$, and $b_{b,i}$ is rate achieved with M/n.

```
At every observation period:
1.for $i \leftarrow 1$ to $n$
2.    if $0.75m[i] < b[i] < t[i]$
3.        $m[i] = m[i] + \Delta(b[i] - t[i])$
4.    else if $b[i] \leq 0.75m[i]$
5.        $m[i] = m[i] - \Delta(0.75m[i] - b[i])$
6.    else if $b[i] > t[i]$
7.        $m[i] = m[i] - \Delta(b[i] - t[i])$
8.Do Max-Min fair with $m[i]$ and $M$
$m[i]$: Marking rate of $i^{th}$ flow
$b[i]$: current rate of $i^{th}$ flow
$t[i]$: Target rate of $i^{th}$ flow
$M$: Total marking rate = Aggregate contract rate
$n$: Number of flows
```

Fig. 1. An algorithm for adaptive marking

Assumption:
1. $\frac{\partial b_i}{\partial m_i} \geq 0$: Throughput of a flow does not decrease as the marking rate of the flow increses when other network condition is not changed.
2. $\frac{\partial^2 b_i}{\partial m_i^2} \leq 0$: There exists only one marking rate m_i at which $b_i = 0.75m_i$.
3. If $b_i < 0.75m_i$, there is IN packet drop.

Proof: For each flow,
- When $\min\{t_i, b_{a,i}, b_{b,i}\} = t_i$; If current b_i is less than t_i, b_i should be greater than $0.75m_i$ since t_i is less than $b_{a,i}$. Then, m_i increases from line 3. From the assumption that $\frac{\partial b_i}{\partial m_i} \geq 0$, b_i eventually reaches t_i. Since t_i is less than $b_{b,i}$, $m_{t,i}$, at which b_i is equal to t_i, is less than M/n, and line 8 does not change m_i.
- When $\min\{t_i, b_{a,i}, b_{b,i}\} = b_{a,i}$; If current b_i is less than $b_{a,i}$, b_i is greater than $0.75m_i$ from the definition of $b_{a,i}$ and, at the same time, less than t_i since t_i is greater than $b_{a,i}$. Then, from line 3, m_i increases until b_i reaches $b_{a,i}$. Here also, since $b_{a,i}$ is less than $b_{b,i}$, $m_{a,i}$ is less than M/n, and line 8 does not change m_i.
- When $\min\{t_i, b_{a,i}, b_{b,i}\} = b_{b,i}$; If current b_i is less than $b_{b,i}$, b_i is greater than $0.75m_i$ and less than t_i since $b_{b,i}$ is less than $b_{a,i}$ and t_i. Then, from line 3, m_i increases until M/n. Then, b_i reaches $b_{b,i}$.

Time complexity of this algorithm is $O(n \log n)$ where n is the number of flows, and this is allowable for an edge device marker. In this algorithm, we use Time Sliding Window (TSW) mechanism [8] to smooth out the individual throughput.

IV. Achieving Target Rates

In this section, we show how the proposed marking scheme realizes achievable individual target rates and finds maximized throughputs when the target rates are not achievable. We consider a multi-hop topology. There are n diffserv routers, and markers are located in between sources and diffserv routers so that every packets are marked before arriving the routers. Cross traffic is injected to this network at the i^{th} router and exits at the $(i+1)^{th}$ router.

To observe how the marking rate is adjusted and an individual flow achieves its target rate, we conducted a set of simulations. In the simulation, the link capacity is 3 Mbps

---

[4] In [6], oversubscribed network has been defined as a situation in which a flow does not transmit any OUT packets since every OUT packets are dropped or no OUT packet is sent when the sending rate is less than the contract rate. In a oversubscribed network, a flow usually experiences some number of IN packet losses.

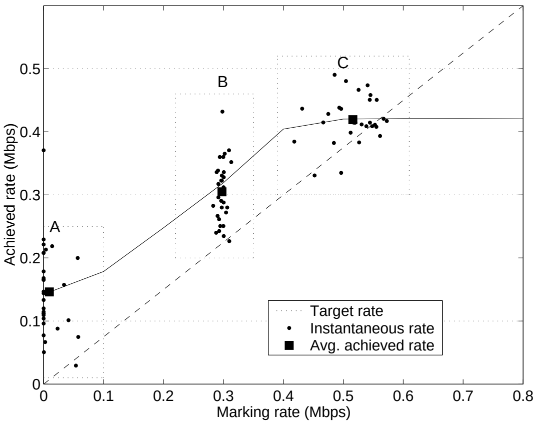

Fig. 2. Achieved and marking rates with target rates 0.1, 0.3 and 0.5 Mbps

and 10 TCP[5] flows are used for cross traffic. The contract rate for each TCP flow is randomly selected from 0 to 1 Mbps, and the total contract of cross traffic is 2.7 Mbps so that the subscription level is 90%. The number of routers (n) is five. For the tagged flow, we use single TCP flow.

First, to observe path characteristics, we use static contract rate for the tagged flow. We vary the contract rate from 0 to 0.8 Mbps. In Fig. 2, the solid line shows the achieved rate with static marking rate, and the dashed line indicates the achieved rate is equal to the 75% of marking rate. It is clear that the achieved rate increases as the marking rate increases until 0.5 Mbps. It is also observed that after 0.55 Mbps the achieved rate does not incease even if we increase the making rate upto 0.8 Mbps. This observation supports our assumption that if we increase the marking rate more than the point in which the achived rate is the 75% of the marking rate, the flow observes oversubscribed path and wastes the marking rate. In thie example, the maximum achievable rate is about 0.42 Mbps.

Now the tagged flow is an individual flow within an aggregation with aggregated contract rate. The marker for the aggregation employs the adaptive marking. We vary the target rate for the tagged flow from 0.1 to 0.5 Mbps. In Fig. 2, 'A', 'B' and 'C' show the results when the target rate is 0, 0.1 and 0.3 Mbps, respectively. In the figure, dots indicate instantaneous marking and achieved rate, and a square shows the average.

In this path, a flow gets 0.15 Mbps with zero contract rate. When the target rate is 0.1 Mbps ('A' in the figure), therefore, the marking rate stays around zero. When the target rate is achievable (less than 4.2 Mbps), it is observed that the adaptive marking scheme finds the minimum marking rate to realize the target ('B' in the figure). In 'C', it is also observed that the marking rate stays less than 0.55 Mbps to avoid resource wastage when the target is unachievable.

So far, we have looked at throughput of an individual flow within an aggregation. Now we observe aggregated flows. There are nine flows aggregated. The contract rate for the aggregation is 5 Mbps. We set the individual target rates differently and offer different cross traffic at each link to produce differently network condition.

TABLE I
Simulation results with adaptive marking

Target	Marking	Achieved	p_{in}	p_{out}	Util. by IN
0.1	0.034	0.108	0	0.051	0.321
0.3	0.301	0.302	0	0.102	0.852
0.5	0.529	0.507	0	0.154	0.947
0.7	0.468	0.339	0	0.998	1
0.9	0.591	0.408	0	0.951	0.999
1.1	0.768	0.621	0	0.973	0.999
INF	0.769	0.808	0	0.006	0.794
INF	0.769	0.801	0	0.007	0.799
INF	0.769	0.921	0	0.003	0.611

Table I shows the summary of the results. When the targets are achievable (target = 0.1, 0.3 and 0.5 Mbps), the achieved throughputs stay around their targets while the marking rates keep changing to adapt the network condition. When the target is unachievable (target = 0.7, 0.9 and 1.1 Mbps), the marking rate is managed to maintain the achieved rate to be the 75% of the marking rate. It is clear shown that the adaptive marking avoids resource wastage ($p_{in} = 0$ for all the flows) and, at the same time, maximizes throughput for the flows having unachievable targets ($p_{out} \approx 1$ and untilization by IN packet ≈ 1). The flows with infinite target rates are for simulating FTP kind of applications and consume the residual contract rates.

V. Conclusions

We have proposed a new adaptive marking algorithm based on the TCP performance model within diff-serv networks. The adaptive marker enables reaching specific QoS goals of individual flows while efficiently managing the aggregate resources. We have presented an extensive simulation study of the behavior of the adaptive marker under various parameters.

References

[1] J. Heinanen, F. Baker, W. Weiss and J. Wroclawski, "Assured Forwarding PHB Group," RFC2597, Network Working Group, June, 1999.
[2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Services," RFC2475, Network Working Group, December, 1998.
[3] W. Feng, D. Kandlur, D. Saha, and K. Shin, "Adaptive Packet Marking for Providing Differentiated Services in the Internet." *Proc. of Int. Conf. on Network Protocols*, pp. 108-117, Austin, TX, October, 1998.
[4] I. Yeom and A. L. N. Reddy. "Marking for QoS Improvement," *Computer Communications*, Vol. 14, No. 1, pp. 35-50, Jan. 2001.
[5] I. Andrikopoulos, L. Wood and G. Pavlou, "A Fair Traffic Conditioner for the Assured Service in a Differentiated Services Internet," *IEEE ICC 2000*, pp. 806-810, Sydney, Australia, June, 2000.
[6] I. Yeom and A. L. N. Reddy. "Modeling TCP Behavior in a Differentiated Services Network," *IEEE/ACM Transactions on Networking*, Vol. 9, No 1, pp. 41-46, Feb. 2001.
[7] L. Breslaau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu and H. Yu, "Advances in Network Simulation," *IEEE Computer*, Vol. 33, No. 5, pp. 59-67, May, 2000.
[8] D. Clark and W. Fang, "Explicit Allocation of Best-Effort Packet Delivery Service," *IEEE/ACM Transactions on Networking*, Vol. 6, No. 4, pp. 362-373, August, 1998.

---

[5] We use TCP Reno for the simulation.