Service Differentiation in ECN Networks using Weighted Window-Based Congestion Control for various Packet Marking Algorithms*

Vasilios A. Siris[1], Costas Courcoubetis[1,2], and George Margetis[1]

[1] Institute of Computer Science (ICS)
Foundation for Research and Technology - Hellas (FORTH)
P.O. Box 1385, GR 711 10 Heraklion, Crete, Greece
`{vsiris, courcou, gmarget}@ics.forth.gr`
[2] Dept. of Informatics, Athens University of Economics and Business, Greece

Abstract. We investigate the service differentiation, in terms of average throughput, and the performance achieved using weighted window-based congestion control in networks supporting Explicit Congestion Notification (ECN). Our results show how service differentiation, queueing delay, and average throughput are affected by the increase and decrease rules of the end-system congestion control algorithms, and how they depend on the marking algorithms operating in the routers. The end-system algorithms we investigate include algorithms that achieve an average throughput proportional to some willingness-to-pay, and the marking algorithms include RED, virtual queue marking, and load-based marking.

1 Introduction

The Transmission Control Protocol (TCP) [8] has played an important role in the Internet's growth. With TCP, however, all connections with the same round trip time receive the same average throughput in the equilibrium; hence, it can not support service differentiation. In the highly competitive telecommunications market, the ability to provide, in a flexible and efficient way, differentiated services will be extremely important. Moreover, new active queue management algorithms [3] combined with Explicit Congestion Notification (ECN) [16] open up new possibilities for alternative approaches to congestion control.

One approach for supporting service differentiation, which is followed by the differentiated services (DiffServ) architecture, is to add mechanisms inside the network. Drawbacks to this approach are the increased complexity, compared to the Internet today, and the need for co-operation among the routers in order to provide end-to-end services. An alternative approach has emerged [6, 10, 12], which suggests that service differentiation, as well as efficient and stable network operation and growth to meet increasing demand, can be achieved by a network with a simple feedback mechanism, such as ECN marking, that informs users

---

[*] This work is supported in part by the European Commission under IST project M3I "Market Managed Multiservice Internet" (IST-1999-11429).

of the congestion cost their traffic is incurring. The users then react to these congestion signals using some rate control algorithm; moreover, the network provider can give users the incentive to react to the congestion signals, hence use the network, according to their actual needs by charging a small fixed price per mark.

In this paper we consider the latter approach and assume a network where congestion feedback is in the form of ECN marks, and no losses occur. For such a network we investigate a family of weighted window-based congestion control algorithms, which we will refer to as willingness-to-pay (WTP) algorithms, that are a generalization of the algorithm first proposed in [6]. With WTP, the congestion window follows a *single multiplicative decrease*: the rate of ECN marks is roughly[1] proportional to the sending rate and the congestion window is decreased by some *fixed* amount for each ECN mark received. On the other hand, TCP follows a *double multiplicative decrease*: both the rate of ECN marks and/or losses is proportional to the sending rate and the congestion window is *halved* upon detection of congestion. The willingness-to-pay algorithms achieve an average rate that, depending on the marking algorithm in the routers, is proportional to some weight, or willingness-to-pay, and inversely proportional to the marking probability inside the network. As we discuss later, the willingness-to-pay parameter can affect the amount by which the congestion window increases when an acknowledgement without an ECN mark is received, or the amount it decreases when an acknowledgement with an ECN mark is received, or both.

Our objective is to investigate, using simulation, the service differentiation achieved by the WTP congestion control algorithms, and how the performance of these algorithms is affected by the marking algorithms operating in the routers. With service differentiation we refer to the ability of the end-system algorithms, working in conjunction with the marking algorithm in routers, to offer different throughput to connections with different weights or willingness-to-pay values. The three marking algorithms we investigate are the RED (Random Early Detection) algorithm [5], modified such that packets are only marked and never dropped, the virtual queue algorithm [6], that marks packets depending on the overflow of a virtual queue whose buffer and capacity are some percentage of the actual buffer and capacity of the link, and a load-based algorithm where the marking probability is a linear function of the link utilization measured over some time interval.

Our major results are summarized as follows:

- With the virtual queue algorithm, where the marking probability depends on the traffic burstiness, service (i.e., average throughput) differentiation depends on the characteristics of the congestion control algorithm, and in particular on the increase and decrease rules.
- With appropriate tuning, all three marking algorithms can exhibit the same marking probability for a range of average load values. As discussed in [10],

---

[1] As we discuss in more detail later, this depends on the marking algorithm operating in the router.

the marking probability as a function of the average load gives information regarding the convergence and stability behaviour of the marking algorithm.

- Both RED and load-based marking, where packet marking is probabilistic, have smaller average queueing delay and delay variation compared to virtual queue marking.
- Comparison of WTP congestion control with MulTCP [2], which follows TCP's double multiplicative decrease but supports service differentiation, shows that MulTCP can be less fair and can achieve lower link utilization compared to WTP.

The rest of this paper is structured as follows: In Section 2 we summarize the theoretical background of our work. In Section 3 we present the willingness-to-pay class of congestion control algorithms, and in Section 4 we discuss the three marking algorithms considered in our investigations. In Section 5 we present and discuss the results from our simulation experiments. Finally, in Section 6 we discuss related work and in Section 7 we present some concluding remarks.

2 Theoretical Background

In this section we summarize the theoretical background of our work as developed by Kelly and other researchers [9, 11, 10].

Consider a network of J resources [9, 11]. Each resource (link) j marks packets with some probability $p_j(y_j)$, where y_j is the aggregate arrival rate at resource j.

Let R be the set of routes, or connections, active in the network. Each connection r updates its rate x_r according to the equation

$$\frac{dx_r(t)}{dt} = \kappa_r \left(w_r - x_r(t) \sum_{j \in r} \mu_j(t) \right) , \tag{1}$$

where w_r is a weight, κ_r is a constant that controls the rate of convergence, and

$$\mu_j(t) = p_j \left(\sum_{s: j \in s} x_s(t) \right) ,$$

is the marking probability at resource j. With (1), a connection r adjusts its sending rate x_r so that the rate of marks it receives $x_r(t) \sum_{j \in r} \mu_j(t)$ becomes equal to the weight w_r. Hence, if the network charges a fixed amount for each mark returned to the end-system, then w_r represents the willingness-to-pay for connection r.

Assume that the marking probability $p(y_j(t))$ depends on the cost $C_j(y_j)$ incurred on resource j as follows

$$p_j(y_j) = \frac{dC_j(y_j)}{dy_j} .$$

It can be shown that, if $C_j(y_j(t))$ is differentiable and feedback is instantaneous, then the above system converges to a point that maximizes

$$\sum_{r \in R} w_r \log x_r - \sum_{j \in J} C_j \left(\sum_{s: j \in s} x_s(t) \right) .$$

If $U_r(x_r) = w_r \log x_r$ is taken to be the utility for a connection r, then the last expression represents the social welfare of the system (resources and connections).

The above results on social welfare maximization can be generalized for the case where a connection r has a utility with a general form $U_r(x_r)$, if the willingness-to-pay changes smoothly according to

$$w_r(t) = x_r(t) U_r'(x_r(t)) .$$

The rate-based control algorithm given by (1) can be approximated using a window-based, self-clocking algorithm, similar to TCP's congestion avoidance algorithm, if a connection's congestion window *cwnd* is updated with the reception of an acknowledgement using, see [6],

$$\mathit{cwnd}{+}= \bar{\kappa} \left(\frac{\bar{w}}{\mathit{cwnd}} - f \right) , \tag{2}$$

where $\bar{\kappa} = \kappa T$ is the gain factor per round trip time, $\bar{w} = wT$ is the willingness to pay per round trip time, and f equals 1 if the acknowledgement contains a mark or 0 if it does not. Note that (1) and its window-based version (2) follows a single multiplicative decrease, since the rate of marks is (roughly) proportional to the sending rate and the congestion window is decreased by some fixed amount for each ECN mark received. This is unlike TCP's congestion avoidance algorithm which follows a double multiplicative decrease, since both the rate of marks and the decrease of the congestion window for each mark (in particular, it is halved) is proportional to the sending rate.

The above results assumed instantaneous feedback. Next we summarize the approach of Kelly [10] for taking feedback delay into account, for the case of a single resource when all connections have the same gain parameter κ and under the assumption that the queueing delay represents a small fraction of the round trip time T (feedback delay), which is fixed and common for all connections. Let $x(t) = \sum_r x_r(t)$ and $w = \sum_r w_r$. Summing (1) for all connections and taking the time lag into account we get

$$\frac{dx(t)}{dt} = \kappa \left(w - x(t-T) p(x(t-T)) \right) . \tag{3}$$

The linear delay equation

$$\frac{du(t)}{dt} = -\alpha u(t-T) \tag{4}$$

converges to zero as t increases if $\alpha T < \pi/2$. Moreover, the convergence is non-oscillatory if $\alpha T < 1/e$. Letting $x(t) = x + u(t)$, where x is the equilibrium of (3), and linearizing equation (3) about x, we get equation (4) with $\alpha = \kappa(p + xp')$, where p, p' are the marking probability and its derivative at rate x. Hence, the equilibrium of (3) is stable and the convergence is non-oscillatory if

$$\kappa T(p + xp') < e^{-1} .$$

3 Weighted Window-Based Congestion Control

In this section we describe a class of window-based congestion control algorithms that are a generalization of the algorithm given by (2) and first presented in [6].

The congestion window *cwnd* is updated using

$$cwnd{+} = \bar{\kappa} \left(\frac{\bar{w}_{\mathrm{inc}}}{cwnd} - \frac{f}{\bar{w}_{\mathrm{dec}}} \right) ,$$

hence, the average change of the rate per unit time is

$$\frac{\bar{\kappa} \left(\frac{\bar{w}_{\mathrm{inc}}}{cwnd} - \frac{p}{\bar{w}_{\mathrm{dec}}} \right) / T}{T / cwnd} = \frac{\bar{\kappa}}{T} \left(\frac{\bar{w}_{\mathrm{inc}}}{T} - \frac{xp}{\bar{w}_{\mathrm{dec}}} \right) , \tag{5}$$

where, as before, $\bar{\kappa}$ is the gain factor per round trip time and $f = 1$ or 0 if the acknowledgement contains or does not contain an ECN mark, respectively. From the last equation we can deduce that for connections with the same round trip time, the average rate is proportional to $\bar{w}_{\mathrm{inc}} \bar{w}_{\mathrm{dec}} = \bar{w}$. For $\bar{w}_{\mathrm{inc}} = \bar{w}, \bar{w}_{\mathrm{dec}} = 1$ we have the proportional increase algorithm given by (2). For $\bar{w}_{\mathrm{inc}} = 1, \bar{w}_{\mathrm{dec}} = \bar{w}$ we have the inversely proportional decrease algorithm given by

$$cwnd{+} = \bar{\kappa} \left(\frac{1}{cwnd} - \frac{f}{\bar{w}} \right) . \tag{6}$$

The values of $\bar{w}_{\mathrm{inc}}$, $\bar{w}_{\mathrm{dec}}$ represent a trade-off between the aggressiveness in probing for available bandwidth, hence in reaching the steady state, and the stability and size of the fluctuations around the average congestion window (equivalently, around the average throughput).

Substituting $\kappa = \bar{\kappa}/T$ and $w_{\mathrm{inc}} = \bar{w}_{\mathrm{inc}}/T$ in (5), summing for all connections, and taking the time lag into account we get

$$\frac{dx(t)}{dt} = \kappa \left(w_{\mathrm{inc}} - \frac{x(t-T)p(x(t-T))}{\bar{w}_{\mathrm{dec}}} \right) ,$$

which, using the results of Kelly [10] that are summarized in the previous section, is stable and the convergence is non-oscillatory if

$$\frac{\kappa T(p + xp')}{\bar{w}_{\mathrm{dec}}} < e^{-1} . \tag{7}$$

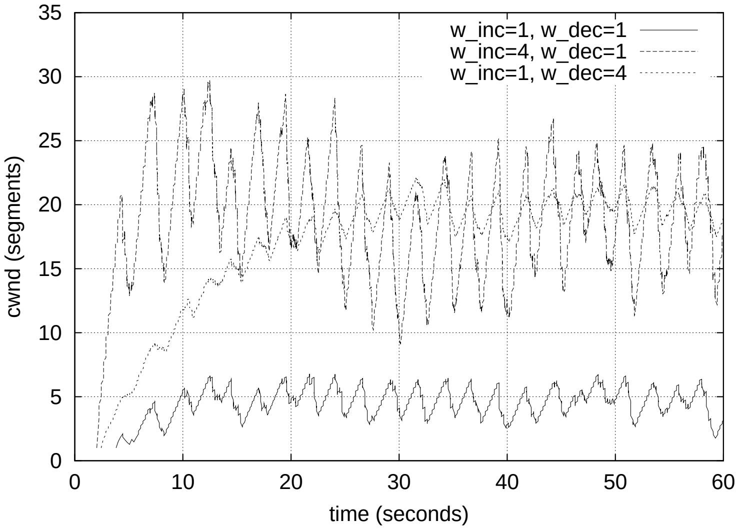


Fig. 1. For $\bar{w}_{\text{inc}} = 4, \bar{w}_{\text{dec}} = 1$ the congestion window reaches the equilibrium faster, but fluctuations are larger. On the other hand, for $\bar{w}_{\text{inc}} = 1, \bar{w}_{\text{dec}} = 4$ the congestion window reaches the equilibrium slower, but fluctuations are smaller. Both connections achieve approximately the same average congestion window, hence throughput.

The last equation is easier to satisfy for larger values of $\bar{w}_{\text{dec}}$.

The trade-off between rate of convergence and magnitude of fluctuations in the equilibrium is depicted in Figure 1. As noted above, increasing $\bar{w}_{\text{dec}}$, while keeping the product $\bar{w}_{\text{inc}}\bar{w}_{\text{dec}} = \bar{w}$ constant, results in smaller fluctuation around the equilibrium. Smaller fluctuations can be advantageous for streaming applications, which typically perform better when their sending rate does not change abruptly.

As indicated by (7), convergence also depends on the marking probability p and its derivative p': In general, a smaller value of p and p' results in faster convergence, but, on the other hand, will also result in larger fluctuations around the equilibrium [11]. As we investigate in Section 5, the marking probability depends both on the congestion control algorithm and the marking algorithm operating in the routers.

4 Packet Marking Algorithms

In this section we describe the three marking probability algorithms, namely Random Early Detection (RED), virtual queue marking, and load-based marking, that we investigate in Section 5.

4.1 Random Early Detection (RED)

With RED [5], the marking probability is a piecewise linear function of the average queue length $\bar{q}$, which is estimated using exponential averaging with weight factor w_q. Specifically, if $\bar{q}$ is smaller than some minimum queue length min_{th}, then the marking probability is zero. If $\bar{q}$ is between min_{th} and some maximum queue length max_{th}, then the marking probability ranges linearly from 0 to some maximum value max_p. Finally, if $\bar{q}$ is above max_{th} then packets are dropped. The "gentle_" modification[2] of the original RED algorithm suggests to vary the marking probability from max_p to 1, when $\bar{q}$ is between max_{th} and $2max_{th}$.

In the investigations of the next section we use the RED algorithm with the "gentle_" variation and modified so that packets are never dropped but only marked.

4.2 Virtual Queue Marking

The virtual queue marking algorithm [6] presents an early warning of congestion. The algorithm maintains a virtual buffer of size θB serviced at rate θC, where B, C are the actual buffer and capacity of the output link, respectively. Note that B is not necessarily the total buffer of the output link, but can be some value that corresponds to a maximum target delay. The algorithm marks all packets that arrive at the link from the time a loss occurs in the virtual buffer until the first time the virtual buffer becomes empty; this period is called the busy period of the virtual buffer. Note that there are other variations for when to mark packets: For example, the algorithm in [7] marks incoming packets when they cause the virtual queue to overflow.

As we will see in the next section, a property of the virtual queue algorithm is that it differentiates flows based on their burstiness. Another property, demonstrated in experiments that are not presented in this paper, is that it does not avoid cases of synchronization of phases in closed-loop congestion control algorithms, such as those observed for drop-tail routers [4], which can result in some connections achieving very large average throughput, and other connections achieving very small throughput. Indeed, one motivation for introducing RED in routers was to avoid such effects [4, 5].

4.3 Load-Based Marking

The third marking algorithm we investigate is load-based marking. According to this algorithm, the marking probability is a piecewise linear function of the average load (utilization), which is measured over some time interval. In contrast, with RED the marking probability is a piecewise linear function of the average queue length. With the load-based algorithm, the marking probability is zero

---

[2] See 'Recommendation on using the "gentle_" variant of RED', S. Floyd, 2000, `http://www.aciri.org/floyd/red/gentle.html`

when the average load is less than some minimum value ρ_0. For values of the load ρ larger than ρ_0 the marking probability is given by $\min\{\alpha(\rho - \rho_0), 1\}$. Hence, the algorithm has three parameters: the minimum load ρ_0, the parameter α that controls the slope of the marking probability, and the averaging interval t_{avg}.

The averaging interval determines how quickly the algorithm adjusts the marking probability to changes of the load. Moreover, the interval determines the timescales over which congestion is detected, hence the timescales over which traffic burstiness affects the marking probability.

Load-based marking can be particularly appropriate in cases where there is no buffering, such as Code Division Multiple Access (CDMA) wireless networks and Ethernet local area networks, since unlike the previous two marking algorithms, it does not rely on the queue length as a measure of congestion.

5 Simulation Results and Discussion

In this section we present and discuss simulation results investigating the interaction of the willingness-to-pay (WTP) congestion control algorithms discussed in Section 3, and in particular the proportional increase (2) and the inversely proportional decrease (6) algorithms, and the three marking algorithms discussed in Section 4.

The specific issues we investigate include the following:

- Service differentiation: Dependence of the average throughput achieved by different connections on their weight or willingness-to-pay, and how this is affected by the marking algorithm.
- Dependence of the marking probability on the average load: As discussed in Section 2, this dependence gives information regarding the stability and convergence behaviour of the whole system, i.e., the congestion control algorithms in the end-systems and the marking algorithms in the routers.
- Queueing delay and link utilization: In particular, we investigate the queueing delay for different marking algorithms, when the average link utilization is the same.
- Comparison of the WTP and MulTCP algorithms: Recall that WTP follows a single multiplicative decrease, whereas MulTCP follows a double multiplicative decrease, similar to normal TCP.

The simulations were performed using the ns-2 simulator [17]. The topology of the simulated network is shown in Figure 2. Since our focus is on investigating the interaction of the congestion control algorithms at the end-systems and the marking algorithms at the routers, we consider link parameters that do not result in packet loss, since this would generate retransmissions, which would affect the throughput measurements. Also, we focus on the congestion avoidance phase, hence our measurements begin after some time interval has elapsed from the start of each experiment.

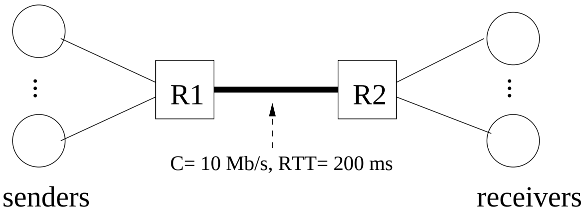


Fig. 2. Network topology used in the simulation experiments. The link connecting the two routers is the bottleneck.

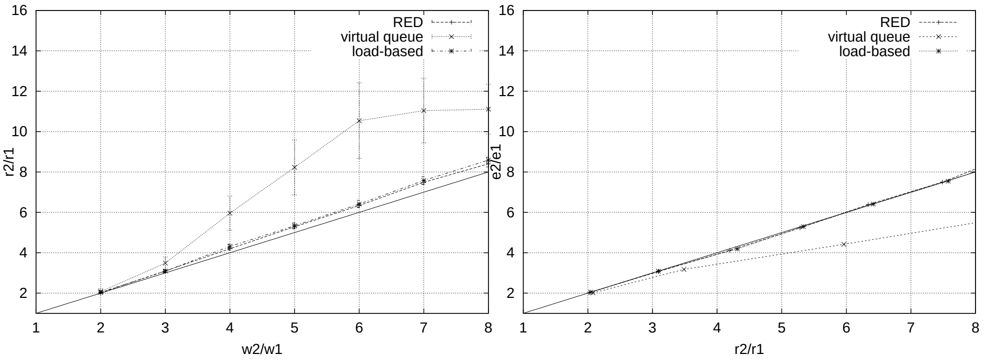
5.1 Service Differentiation

Figure 3(a) shows the average ratio of throughput for different ratios of willingness-to-pay values. Also shown is the 95% confidence interval. The results were obtained from 10 independent runs of the experiment with the same parameters. Each connection carried data from a long ftp transfer, and the start time of each connection was selected randomly from the interval [0, 5] seconds. Finally, the throughput was computed for the interval [60, 180] seconds.

Observe that the service differentiation achieved by both RED and load-based marking is roughly the same. Moreover, the ratio of the average throughput is very close to the ratio of willingness-to-pay. Note, however, that the two curves are slightly above the diagonal; this is attributed to the window-based nature of the congestion control algorithm, since in results not shown here, the ratio of the congestion windows is approximately equal to the ratio of willingness-to-pay.

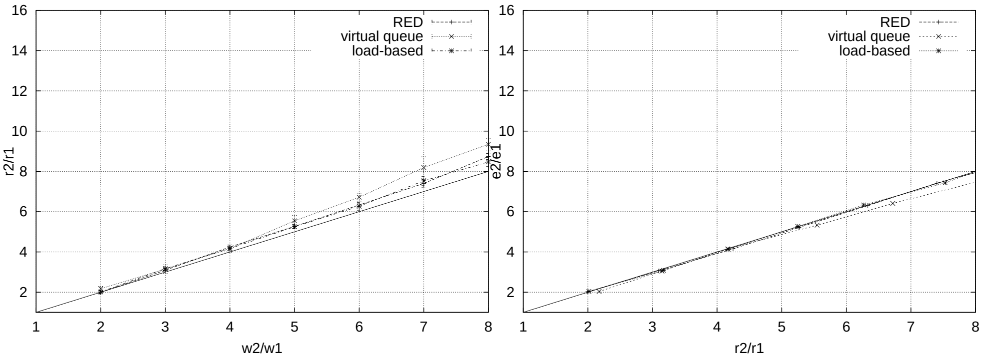
On the other hand, in the case of virtual queue marking observe that for large values of the ratio of willingness-to-pay, the ratio of throughput is larger than the ratio of willingness-to-pay. Moreover, as indicated by the confidence interval, the fluctuations in service differentiation are also much larger compared to the fluctuations for the other two marking algorithms. Next we explain both of these observations.

The first observation can be explained with the help of Figure 3(b), which shows the ratio of marks as a function of the ratio of average throughput. The figure shows that a smaller percentage of the packets belonging to a connection with a larger willingness-to-pay are marked, i.e., the marking probability is smaller for a connection with a larger willingness-to-pay. This is due to the combination of the following two factors: first, the connection with a smaller willingness-to-pay sends a smaller number of segments in one round trip time and second the segments are typically sent back-to-back; the latter being a property of any window-based control mechanism. As a result, the connection with a smaller willingness-to-pay produces a burstier traffic stream compared to the connection with a larger willingness-to-pay. Burstier traffic streams, however, are more difficult for a multiplexer (link) to handle, hence require more bandwidth than their average rate. The virtual queue marking algorithm has the property of differentiating streams based on their burstiness. Hence, the algorithm marks



(a) ratio of throughput (b) ratio of received marks

Fig. 3. Service differentiation using proportional increase WTP. $\bar{\kappa} = 0.5$. $C = 10$ Mbps, $RTT = 200$ msec, $N = 10$. RED: $min_{th} = 5, max_{th} = 15, max_p = 0.1, w_q = 0.002$. VQ: $vb = 0.95 \cdot 30, vc = 0.95 \cdot 10$ Mbps. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6, \alpha = 0.71$.



(a) ratio of throughput (b) ratio of received marks

Fig. 4. Service differentiation using inversely proportional decrease WTP. $\bar{\kappa} = 0.5$. $C = 10$ Mbps, $RTT = 200$ msec, $N = 10$. RED: $min_{th} = 5, max_{th} = 15, max_p = 0.1, w_q = 0.002$. VQ: $vb = 0.95 \cdot 30, vc = 0.95 \cdot 10$ Mbps. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6, \alpha = 0.71$.

a higher percentage of packets belonging to the burstier stream, which, as explained above, is the stream with smaller willingness-to-pay.

The above is not the case when the end-systems implement inversely proportional decrease WTP, as shown in Figure 4(a). In this case, and for the timescales in which packet marking occurs, the traffic from WTP connections with a different willingness-to-pay exhibit roughly the same burstiness, hence are not differentiated and the service differentiation for virtual queue marking is close to that of RED and load-based marking. This is also illustrated in Figure 4(b), which shows that in this case the ratio of marks for all three algorithms is roughly proportional to the ratio of willingness-to-pay values.

Next we explain the second observation made from Figure 3(a), namely the larger variations of service differentiation achieved with virtual queue marking compared with to that of RED and load-based marking. With the latter two algorithms packet marking is probabilistic, hence marks tend to be spread out; this results in smoother changes of the congestion window. On the other hand, the virtual queue algorithm tends to produce bursts of marks, which result in larger fluctuations of the congestion window.

5.2 Marking Probability

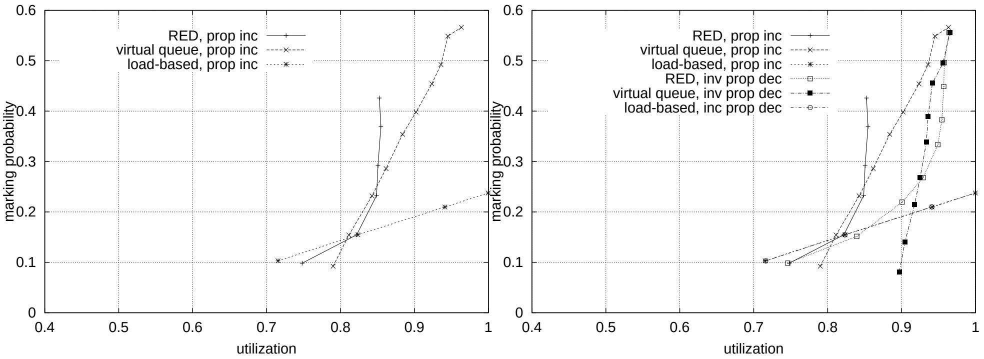
Next we investigate the marking probability as a function of the average load, for the three marking algorithms we consider. Note that this probability depends, in addition to the load, also on the burstiness of traffic, which in turn is affected by both the congestion control algorithm in the end-systems and the marking algorithm in the routers. The shape of this curve, as discussed in Section 2 and [10], gives an indication of the convergence and stability behaviour of the particular marking algorithm.

Figures 5(a) and (b) show the marking probability as a function of average load for the three marking algorithms and for the proportional increase and inversely proportional decrease algorithms running on the end-systems. Observe that for both virtual queue and load-based marking, there is a linear dependence of the marking probability on the link utilization. On the other hand, the marking probability for RED is convex and becomes steep for large values of the load, although as we will see later this is not always the case. Figure 5(b) shows that for both virtual queue marking and RED, the end-system congestion control algorithm has an affect on the marking probability. Indeed, for the inversely proportional decrease congestion control algorithm, the same marking probability is achieved for a higher utilization. This is expected, since the inversely proportional decrease algorithm results in smoother traffic, as discussed in Section 3, hence can achieve a higher utilization, for the same marking probability.

Figures 5(a) and (b) were obtained for some arbitrary[3] selection of parameters for each marking algorithm. Figure 6 shows that these parameters can be

---

[3] For RED we used parameters suggested in the literature. Our results indicate that the rules for tuning of RED in the case of WTP congestion control are not the same as in the case of TCP.



(a) Proportional increase (b) Inversely proportional decrease

Fig. 5. Marking probability as a function of load for proportional increase WTP: $\bar{\kappa} = 0.5, \bar{w}_{\text{inc}} = 1$ and $6, \bar{w}_{\text{dec}} = 1$, and inversely proportional decrease WTP: $\bar{\kappa} = 0.5, \bar{w}_{\text{inc}} = 1, \bar{w}_{\text{dec}} = 1$ and 6. $C = 10$ Mbps$, RTT = 200$ msec. RED: $min_{th} = 2, max_{th} = 6, max_p = 0.1, w_q = 0.02$. VQ: $vb = 0.95 \cdot 30, vc = 0.95 \cdot 10$ Mbps. LB: $t_{avg} = 0.5$ s$, \rho_0 = 0.6, \alpha = 0.71$.

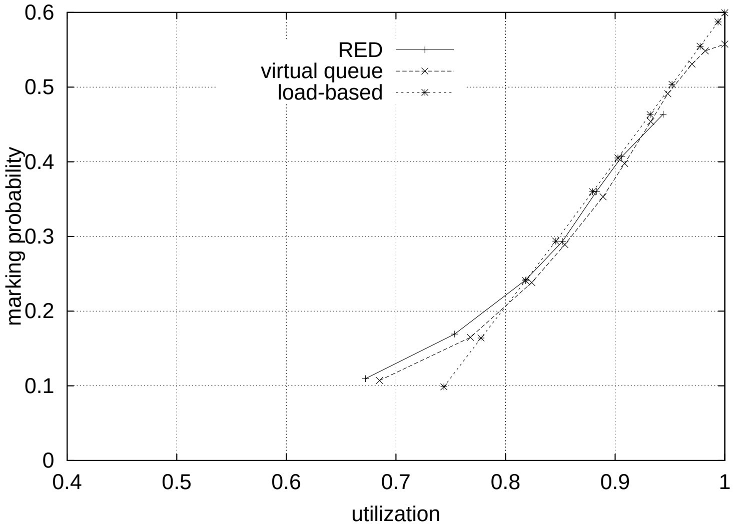


Fig. 6. Marking probability as a function of link utilization. $\bar{\kappa} = 0.5, \bar{w}_{\text{inc}} = 1$ and $6, \bar{w}_{\text{dec}} = 1$. $C = 10$ Mbps$, RTT = 200$ msec. RED: $min_{th} = 3, max_{th} = 9, max_p = 0.95, w_q = 0.002$. VQ: $vb = 0.95 \cdot 20, vc = 0.95 \cdot 10$ Mbps. LB: $t_{avg} = 0.5$ s$, \rho_0 = 0.7, \alpha = 2$.

Table 1. Queueing delay (in msec). Long ftp flows, $\rho \approx 0.90$. The bottom three rows show that the queueing delay is smaller for RED and load-based marking, compared to virtual queue marking.

marking algorithm	average	maximum	std. dev.
RED ($min_{th} = 5, max_{th} = 15,$ $max_p = 0.10, w_q = 0.002$)	26.9	89.6	26.4
RED ($min_{th} = 3, max_{th} = 9,$ $max_p = 0.95, w_q = 0.002$)	5.3	22.4	3.6
virtual queue ($vb = 0.95 \cdot 20, vc = 0.95 \cdot 10$ Mbps)	8.6	39.4	6.2
load-based ($t_{avg} = 0.5$ s, $\rho_0 = 0.7, \alpha = 2$)	6.3	38.4	4.8

chosen such that the marking probability curves for the three algorithms become very close. This indicates that, at least in terms of the macroscopic convergence and stability properties, all marking algorithms have similar behaviour when their parameters are tuned appropriately. Note, however, that these curves are averages, and as we will see in the next subsection, do not give a complete picture of the queueing delay for each marking algorithm.

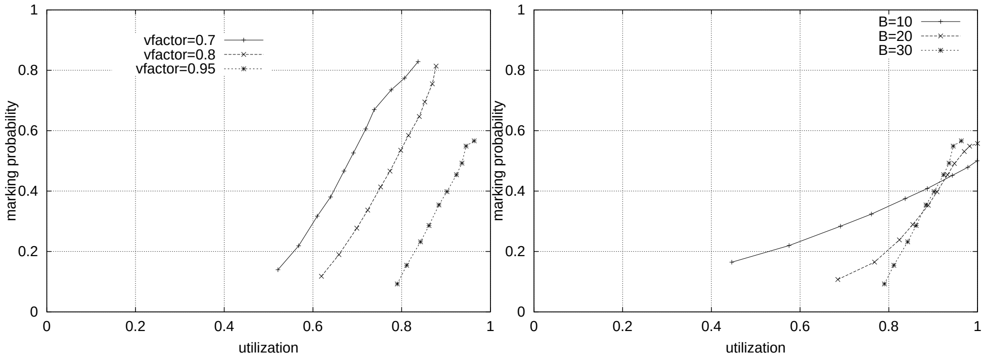
5.3 Queueing Delay

In the previous subsection we saw that the parameters of the marking algorithms we consider can be tuned so that the average marking probability is roughly the same for a range of average utilization values. Since these are average values, they do not present a complete picture of the queueing behaviour. Table 1 shows the queueing delay for the three marking algorithms when the average utilization is kept the same. The bottom three rows in the table show that the queueing delay is smaller and exhibits smaller fluctuations for RED and load-based marking, compared to virtual queue marking. Again, this is expected since, as discussed in Section 5.1, due to their probabilistic marking, both RED and load-based marking result in smoother traffic. The first two rows show that for the same marking algorithm the queueing delay depends on the parameters of the marking algorithm, even when the utilization is kept the same.

5.4 Effect of Marking Algorithm Parameters

In addition to their performance, how easy or difficult it is to tune the parameters of a marking algorithm is equally, if not more, important. Towards this end, in this subsection we provide some insight on how the parameters affect the marking probability function.

Figure 7(a) shows how the marking probability for the virtual queue marking algorithm is affected by the factor θ, which determines the buffer and capacity of the virtual queue. Observe that as θ increases, the marking probability shifts to the right, i.e., for larger θ, the same utilization corresponds to a lower marking



(a) factor θ (b) buffer

Fig. 7. Effect of the virtual queue factor θ and the buffer size. $C = 10$ Mbps, $RTT = 200$ msec, $\bar{w}_{\mathrm{inc}} = 1$ and $6, \bar{w}_{\mathrm{dec}} = 1, \bar{\kappa} = 0.5$.

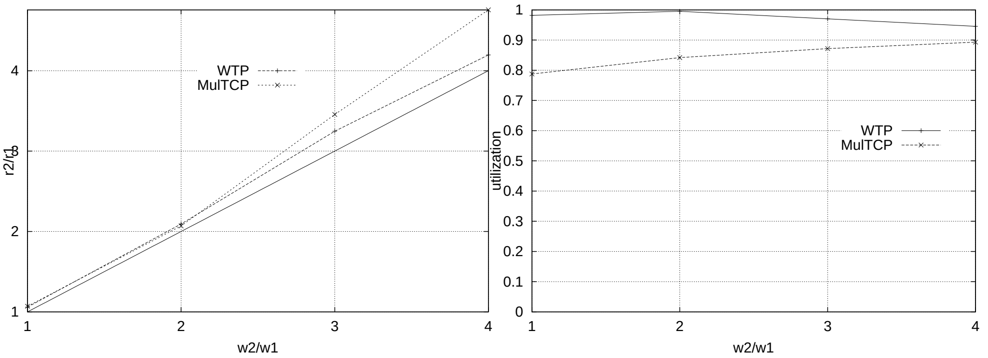
probability. Also, observe that the factor has a very small effect on the slope of the curve. Finally, Figure 7(b) shows that a larger buffer results in a steeper curve for the marking probability. Indeed, for small utilizations, the marking probability is higher for a smaller buffer, since the buffer overflows more frequently. On the other hand, for large utilizations the effect is the opposite, i.e., the marking probability is higher for a larger buffer, because now the dominating effect is the larger overflow (marking) periods, since a larger buffer needs more time to empty.

For load-based marking the effect of its parameters is straightforward: The parameter α affects the slope of the marking probability curve, ρ_0 (for constant α) affects the marking probability for a given average throughput, and the averaging interval affects how fast the algorithm responds to changes of network load and the timescales over which traffic burstiness affects the marking probability. Moreover, the averaging interval affects the maximum queue backlog that appears when the congestion level changes.

Finally, for RED we have observed that when max_p is small, then the marking probability function is convex, as in Figure 5. On the other hand, when max_p approaches 1, then the marking probability function tends to be linear, as in Figure 6. In general, we have found that the rules for tuning the parameters of RED in the case of WTP congestion control are not the same as the rules in the case of TCP.

5.5 Comparison of WTP and MulTCP

Figure 8 compares willingness-to-pay congestion control with MulTCP [2]. The latter follows TCP's double multiplicative decrease, but provides support for service differentiation. Figure 8(a) shows that the ratio of average throughput is higher than the ratio of weights for MulTCP, indicating that MulTCP favours connections with larger weights.



(a) Service differentiation for WTP and MulTCP.

(b) Average utilization for WTP and MulTCP.

Fig. 8. Comparison of WTP and MulTCP. WTP: $\bar{\kappa} = 0.5$. $C = 10$ Mbps, $RTT = 200$ msec, $N = 10$. RED Marking: $min_{th} = 5, max_{th} = 15, max_p = 0.1, w_q = 0.002$.

Figure 8(b) compares the utilization achieved with WTP and MulTCP, when the same number of connections are multiplexed in a link with RED marking. Observe that for small weights, WTP achieves a higher utilization compared to MulTCP. As the weights increase, however, the difference between the two congestion control algorithms is smaller. We conjecture that this is due to the following: For MulTCP, connections are bursty even for small weights, due to the proportional decrease of the congestion window when an ECN mark is received; moreover, as the weights increase, the resulting increase of aggressiveness helps connections achieve a higher average throughput. On the other hand, for small weights WTP connections produce smooth traffic, hence the average utilization is high; when the weights increase, however, connections become burstier, and as a result the average utilization decreases.

6 Related Work

Next we discuss some related work. Note that this is not an exhaustive survey of all the work in the area.

The authors of [15] investigate the limitations of using double proportional decrease algorithms for achieving service differentiation. Their results are in agreement with the results of Section 5.5. They also investigate an algorithm where the congestion window is adjusted based on the percentage of lost packets. In contrast, our investigations of the WTP algorithms are for networks with zero loss. The loss adaptive algorithm in [15] achieves in the steady state an average throughput inversely proportional to the loss probability. This is similar to the dependence of the average throughput on the marking probability with the WTP algorithms investigated in this paper. Comparison of the two algorithms in terms of throughput and delay is an area for further investigation.

The authors of [14] investigate the convergence and steady state behaviour, using both simulation and dynamical system modelling, of the rate-based congestion control algorithm given by (1) and when routers implement threshold-based marking, i.e., all packets are marked when the queue length exceeds a threshold. Additionally, they discuss issues regarding the timescales for a connection's rate to reach equilibrium in relation to the timescales of the inter-arrival time between new connections and the duration of each connection.

The author of [18] discusses the fairness of marking algorithms, using ideas from large deviations theory and economics. A number of marking algorithms that have been proposed in the literature are investigated, and a threshold-based marking algorithm that marks all packets in the queue when the queue length exceeds a threshold is proposed. According to the algorithm, the threshold is adjusted adaptively based on the number of packets that have contributed to the queue exceeding the threshold, hence packet marks represent the incremental cost for sending an additional packet, where cost is taken to be the number of packets exceeding the threshold level.

The authors of [13] investigate, using a fluid model, the decentralized selection of the marking rate at each router of a network in order to achieve loss-free operation. The convergence of the scheme is investigated using a timescale decomposition of the resulting system into a slow and fast system model.

7 Concluding Remarks

We have investigated, using simulation, the service differentiation and performance achieved with end-to-end weighted window-based congestion control algorithms, and how this is affected by the marking algorithms implemented in the routers. The congestion control algorithms investigated include the willingness-to-pay (WTP) algorithms that follow a single multiplicative decrease of the congestion window upon congestion. Our results show that, for networks with no losses supporting Explicit Congestion Notification (ECN), the WTP algorithms can be fairer and lead to higher utilization compared to MulTCP, which supports service differentiation and follows a double multiplicative decrease, similar to normal TCP. Note that in cases where the loss rate tends to be a large, algorithms that decrease the congestion window by a large amount, e.g., by half as in the case of TCP and MulTCP, might be more appropriate, since a large decrease of the congestion window would result in a lower loss rate, hence a smaller number of retransmissions.

The marking algorithms investigated include RED, virtual queue marking, and load-based marking. Our results show that service differentiation and queueing delay can be worse for the virtual queue algorithm, where packet marking occurs at the timescales of overflow of a virtual buffer, compared to RED and load-based marking where the marking probability depends on averages (queue length for RED and load for load-based marking). On the other hand, the virtual queue marking algorithm can differentiate flows according to their burstiness.

Our investigations were limited to the case where all connections had the same round trip time. When connections have different round trip times, which would be the case in more complex network topologies than the one considered in this paper, the appropriate selection of parameters for marking algorithms, such as RED and load-based, which determine the marking probability based on some average quantities, presents difficulties. These difficulties concern the parameters, such as the exponential weight factor of RED and the averaging interval of load-based marking, which determine the time interval over which averaging is performed, hence how fast the end-systems are informed of changes of the congestion level inside the network. Setting these parameters to achieve optimal feedback for long round trip time connections might be too slow for short round trip time connections. On the other hand, setting the parameters to achieve optimal feedback for short round trip time connections might be too fast for long round trip time connections, thus leading to larger fluctuations of the queue and instability. Moreover, the round trip time also affects the convergence and the average throughput in the equilibrium. The latter effect can be addressed, e.g., by having the increase or decrease of the congestion window upon the reception of an acknowledgement be proportional or inversely proportional, respectively, to the round trip time.

The willingness-to-pay and packet marking algorithms investigated in this paper have been implemented in a real test-bed comprising of workstations running FreeBSD[4]. Preliminary results from experiments conducted over the test-bed are reported in [1].

An issue for further investigation is how to automate the tuning of the parameters (adaptation) of the marking algorithms. Motivation for such automated adaptation is that, in agreement with the results regarding the tuning of RED parameters for TCP connections, no set of parameters are optimal for all possible traffic mixes and characteristics. Finally, another issue for further investigation is, in the case each mark is charged by a fixed price, how this price per mark affects the link utilization, and subsequently how to determine this price, given the aggregate demand, the congestion control algorithms in the end-systems, and the marking algorithms in the routers.

References

1. P. Antoniadis, C. Courcoubetis, G. Margetis, V. A. Siris, and G. D. Stamoulis. Efficient adaptation to dynamic pricing communicated by ECN marks: Scenarios for experimental assessment. In *Proc. of SPIE International Symposium on Information Technologies 2000 (Program on Internet Performance and Control)*, Boston, MA, 2000. Available at `http://www.ics.forth.gr/netgroup/publications/`.
2. J. Crowcroft and P. Oechslin. Differentiated end-to-end Internet services using a weighted proportional fair sharing TCP. *Computer Communications Review*, 28(3):53–67, July 1998.

---

[4] The implementation of the algorithms in ns-2 and in FreeBSD are available at `http://www.ics.forth.gr/netgroup/publications/wtp_vq.html`

3. B. Braden *et al.* Recommendations on queue management and congestion avoidance in the internet. RFC 2309, April 1998.
4. S. Floyd and V. Jacobson. On traffic phase effects in packet-switched gateways. *Internetworking: Research and Experience*, 3:115–156, 1992.
5. S. Floyd and V. Jacobson. Random Early Detection gateways for congestion avoidance. *IEEE/ACM Trans. on Networking*, 1:397–413, 1993.
6. R. J. Gibbens and F. P. Kelly. Resource pricing and congestion control. *Automatica*, 39:1969–1985, 1999.
7. R J. Gibbens, P. B. Key, and S. R. E. Turner. Properties of the virtual queue marking algorithm. In *Proc of 17th UK Teletraffic Symposium*, 2001.
8. V. Jacobson and M. J. Karels. Congestion avoidance and control. In *Proc. of ACM SIGCOMM'88*, pages 314–329, 1988.
9. F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–37, January 1997.
10. F. P. Kelly. Models for a self-managed Internet. *Philosophical Transactions of the Royal Society*, A358:2335–2348, 2000. Available at: `http://www.statslab.cam.ac.uk/~frank/`.
11. F. P. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49, 1998.
12. P. B. Key and D. R. McAuley. Differential QoS and pricing in network: Where flow control meets game theory. *IEE Proceedings Software*, 146(2):39–43, March 1999.
13. S. Kunniyur and R. Srikant. A time scale decomposition approach to adaptive ECN marking. In *Proc. of IEEE INFOCOM'01*, Anchorage, AK, April 2001.
14. K. Laevens, P. Key, and D. McAuley. An ECN-based end-to-end congestion control framework: experiments and evaluation. Technical Report MSR-TR-2000-104, Microsoft Research, October 2000. Available at `http://research.microsoft.com/pubs`.
15. T. Nandagopal, K.-W. Lee, J.-R. Li, and V. Bharghavan. Scalable service differentiation using purely end-to-end mechanisms: features and limitations. In *Proc. of IEEE IWQoS'00*, Pittsburgh, PA, June 2000.
16. K. K. Ramakrishnan and S. Floyd. A proposal to add Explicit Congestion Notification (ECN) to IP. RFC 2481, January 1999.
17. UCB/LBNL/VINT. Network Simulator - ns (version 2). `http://www-mash.cs.berkeley.edu/ns.html`.
18. D. Wischik. How to mark fairly. In *Proc. of Workshop on Internet Service Quality Economics*, MIT, December 1999.