Service Differentiation in ECN Networks using Weighted Window-Based Congestion Control for various Packet Marking Algorithms [1]

Vasilios A. Siris [a,*], Costas Courcoubetis [b,a], and George Margetis [a]

[a] *Institute of Computer Science (ICS), Foundation for Research and Technology - Hellas (FORTH), P.O. Box 1385, GR 711 10 Heraklion, Crete, Greece*

[b] *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 WTP (Willingness-To-Pay) and MulTCP congestion control, and the packet marking algorithms include RED, virtual queue marking, and load-based marking.

Key words: service differentiation, congestion control, active queue management, explicit congestion notification

1 Introduction

The Transmission Control Protocol (TCP) [1,13] 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

* Corresponding author. Tel.: +30 81 391726, fax: +30 81 391601.
E-mail: {vsiris, courcou, gmarget}@ics.forth.gr

1 This work was supported in part by the European Commission under IST project M3I "Market Managed Multi-service Internet" (IST-1999-11429).

telecommunications market, the ability to provide, in a flexible and efficient way, differentiated services will be extremely important. This push, along with the need to develop new algorithms better suited to the characteristics and requirements of the traffic being transferred, e.g., streaming video/audio, will continue to foster the appearance of new congestion control algorithms. Moreover, new active queue management algorithms [4] combined with Explicit Congestion Notification (ECN) [23], which has recently been approved as an IETF proposed standard, 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 network routers. 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 coherent end-to-end services. An alternative approach has emerged [8,15,17], see also overview in [9], 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 of the congestion cost their traffic is incurring. The users react to these congestion signals using some rate control algorithm. Moreover, by charging a small fixed price per mark, the network provider can give users the incentive to react to the congestion signals, hence use the network, according to their actual needs. Note that the above approach places intelligence in the end-systems, while keeping the network simple. In this sense, it adheres to the end-to-end arguments put forward in [24], which have had a profound impact on the design of Internet protocols. The reasons for following such an approach are straightforward: it is at the end-systems where information on how to respond to congestion resides, since end-systems know the requirements of applications but also the valuations of users in terms of the performance versus price tradeoff, both of which can be expressed in the form of utility functions, as discussed in Section 2. Indeed, the rate or congestion control algorithms operating at the end-systems have the role of enforcing the policies expressed by these utility functions.

In this paper we investigate the latter approach, and consider a network where congestion feedback is in the form of ECN marks. For such a network we investigate the service differentiation achieved using weighted window-based congestion control, and how its performance 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 algorithms in routers, to offer different throughput to connections with different weights or willingness-to-pay values.

The congestion control algorithms we investigate are Willingness-To-Pay (WTP), which was first proposed in [8], and MulTCP [3], which is a modification of

TCP that makes one connection approximate the behavior of multiple normal TCP flows. With WTP, the congestion window follows a *single multiplicative decrease*: the rate of ECN marks is roughly[2] proportional to the sending rate and the congestion window decreases by some *fixed* amount for each ECN mark received. On the other hand, MulTCP, as 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. WTP achieves an average rate that, depending on the marking algorithm in the routers, is proportional to a weight or willingness-to-pay, and inversely proportional to the marking probability inside the network. On the other hand, MulTCP achieves a rate that is proportional to a weight and inversely proportional to the square root of the marking and/or loss probability.

The three marking algorithms we investigate are RED (Random Early Detection) [7], virtual queue marking [8], that marks packets depending on the overflow of a virtual queue whose buffer and capacity are some percentage of the link's actual buffer and capacity, and load-based marking, where the marking probability is a linear function of the link utilization measured over some time interval.

Our simulation experiments consider both a simple network with a single congested link, and a network with multiple links. The latter configuration allows us to study service differentiation when connections encounter different round trip times, and when marking occurs on more than one links.

Our main results are summarized as follows:

- Weighted window-based congestion control is an effective way for an end-system to control its throughput. In cases where connections encounter different round trip times, these round trip times should be taken into account to appropriately set weights.
- 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 utilizations. As discussed in [15], the marking probability as a function of average utilization affects the convergence and stability behavior of the system comprising of the end-system congestion control algorithms and the marking algorithms in the routers.
- Both RED and load-based marking, where packet marking is probabilistic, exhibit smaller average queueing delay and delay variation compared to virtual queue marking, which deterministically marks packets from the time

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

the virtual queue overflows until the time it first becomes empty again.[3]

- Comparison of WTP with MulTCP, 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 some background 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 [14,16,15].

Consider a network of J resources [14,16]. 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 gain factor, 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), 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. If the network charges a fixed amount for each mark returned to the end-system, then the sending rate is adjusted so that the rate of marks multiplied by the price per mark becomes equal to w_r, which thus represents the willingness-to-pay for connection r.

Assume that the marking probability $p_j(y)$ depends on the cost $C_j(y)$ incurred

[3] Other variations of the virtual queue algorithm exist, such as the one considered in [10] where a packet is marked when it causes the virtual queue to overflow.

on resource j as follows

$$p_j(y) = \frac{dC_j(y)}{dy}\,.$$

It can be shown that, if $C_j(y)$ 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 can be generalized for the case where a connection r has a utility of a general form $U_r(x_r)$, if the willingness-to-pay changes smoothly according to

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

Up to now we have assumed instantaneous feedback. Next we summarize the approach of Kelly [15] 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

$$\frac{dx(t)}{dt} = \kappa\left(w - x(t-T)p(x(t-T))\right)\,. \tag{2}$$

The linear delay equation

$$\frac{du(t)}{dt} = -\alpha u(t-T) \tag{3}$$

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 (2), and linearizing (2) about x, we get (3) with $\alpha = \kappa(p + xp')$, where p, p' are the marking probability and its derivative at rate x. Hence, the equilibrium of (2) is stable and the convergence is non-oscillatory if

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

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 [8],

$$cwnd{+} = \bar{\kappa}\left(\frac{\bar{w}}{cwnd} - f\right) , \tag{4}$$

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 (4) 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.

3 Weighted Window-Based Congestion Control

In this section we describe a class of window-based congestion control algorithms that are based on the algorithm given by (4) [8]. The congestion window *cwnd* is updated according to

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

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

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

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 see that for connections with the same round trip time, the average rate is proportional to $\bar{w}_{\text{inc}}\bar{w}_{\text{dec}} = \bar{w}$. For $\bar{w}_{\text{inc}} = \bar{w}, \bar{w}_{\text{dec}} = 1$ we have the proportional increase algorithm given by (4). For $\bar{w}_{\text{inc}} = 1, \bar{w}_{\text{dec}} = \bar{w}$ 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 $\bar{w}_{\text{inc}}$, $\bar{w}_{\text{dec}}$ 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).

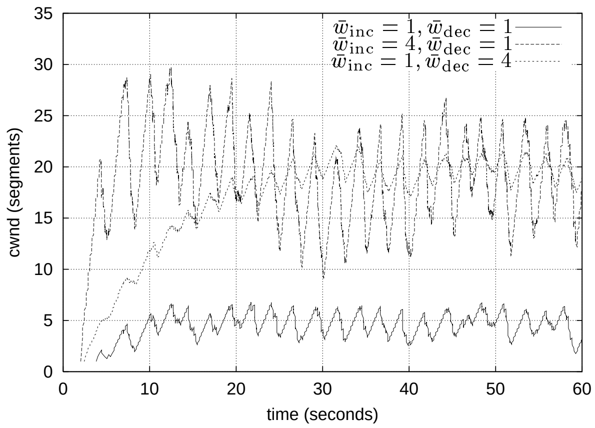


Fig. 1. For $\bar{w}_{\rm inc} = 4, \bar{w}_{\rm dec} = 1$ the congestion window reaches the equilibrium faster, but fluctuations are larger. On the other hand, for $\bar{w}_{\rm inc} = 1, \bar{w}_{\rm dec} = 4$ the congestion window reaches the equilibrium slower, but fluctuations are smaller. Both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

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

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

which, using the results of Kelly [15] that are summarized in the previous section, is stable and the convergence is non-oscillatory if

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

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

The trade-off between rate of convergence and magnitude of fluctuations in the equilibrium is depicted in Figure 1. As noted above, increasing $\bar{w}_{\rm dec}$, while keeping the product $\bar{w}_{\rm inc}\bar{w}_{\rm dec} = \bar{w}$ constant, results in smaller fluctuation around the equilibrium, but in the same average throughput. 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 larger value of p and p' results in faster convergence, but, on the other hand, will also result in larger fluctuations around the equilibrium [16]. 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 that will be investigated in the next section.

4.1 Random Early Detection (RED)

With RED [7], 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 a minimum queue length min_{th}, then the marking probability is zero. If $\bar{q}$ is between min_{th} and a 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 [4] 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 our simulation investigations we use RED with the "gentle_" variation, modified so that packets are marked rather than dropped when the average queue length is above max_{th}, and when the buffer limit is exceeded.

4.2 Virtual Queue Marking

The virtual queue marking algorithm [8] 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 [10] 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 [6], which can result in

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

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 [6,7].

4.3 Load-Based Marking

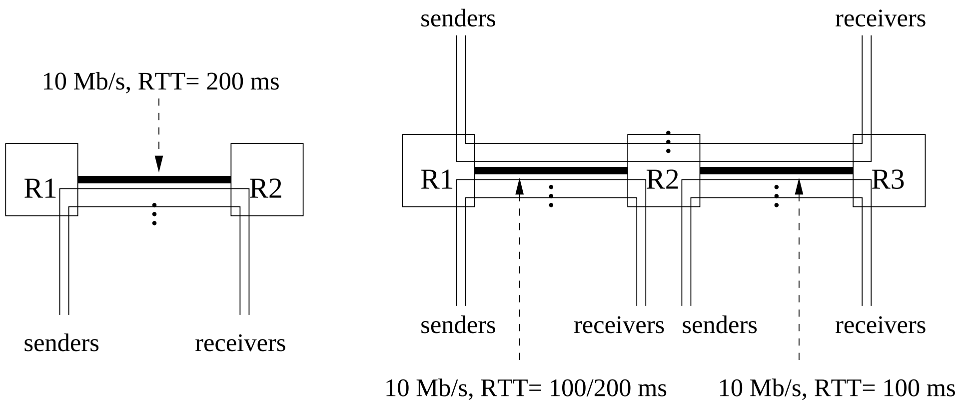
The third marking algorithm we investigate is load-based marking. According to this algorithm, the marking probability is a piecewise linear function of the average utilization measured over some time interval. In contrast, with RED the marking probability is a piecewise linear function of the average queue length. With the load-based algorithm, the marking probability is zero when the average load is less than a minimum value ρ_0. For utilization 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 utilization ρ_0, the parameter α that controls the slope of the marking probability, and the averaging interval t_{avg}; this 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 appropriate in cases where there is no buffer sharing, such as the uplink in 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 WTP and MulTCP with the three marking algorithms discussed in the previous section. 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 utilization: As discussed in Section 2, this dependence affects the stability and convergence behavior 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 remains the same.



(a) Single link (b) Multiple link

Fig. 2. Network topologies used in the simulation experiments.

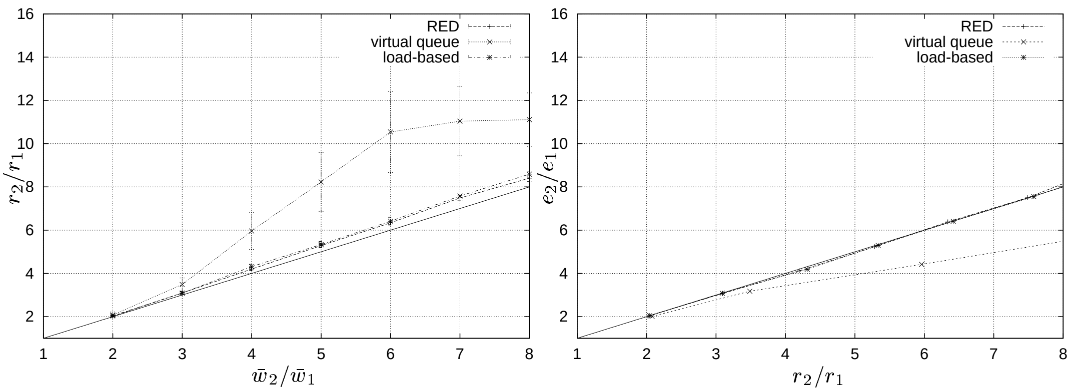
- Comparison of WTP and MulTCP : Recall that WTP follows a single multiplicative decrease, whereas MulTCP follows a double multiplicative decrease, similar to normal TCP.

The simulations were performed using the ns-2 simulator [26]. The two network topologies considered in the simulation experiments are shown in Figure 2. Since our focus is on investigating the interaction of the congestion control algorithms at the end-systems with 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.

5.1 Service Differentiation

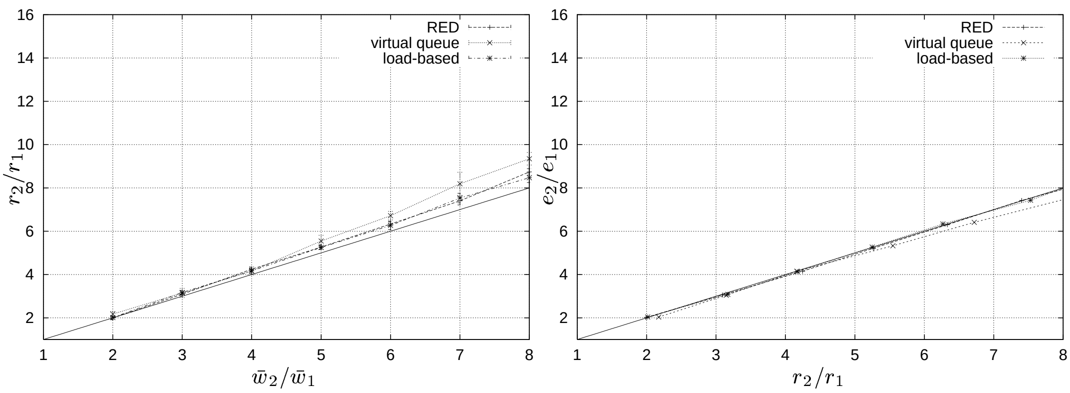
Figure 3(a) shows the average ratio of throughput for different ratios of willingness-to-pay values. Also shown is the 90% confidence interval. The results were obtained from 10 independent runs of the experiment with the same parameters. Each connection carried data from a long ftp transfer, and the 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.

Figure 3(a) indicates that a connection can control its throughput by adjusting its willingness-to-pay. Moreover, observe that the service differentiation achieved by both RED and load-based marking is roughly the same, and 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.



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

Fig. 3. Service differentiation using proportional increase WTP (single link). $\bar{\kappa} = 0.5$. $C = 10$ Mbps, $B = 30$ pkts, $RTT = 200$ ms, $N = 10$. RED: $min_{th} = 5$, $max_{th} = 15$, $max_p = 0.1$, $w_q = 0.002$. VQ: $\theta = 0.95$. 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 (single link). $\bar{\kappa} = 0.5$. $C = 10$ Mbps, $B = 30$ pkts, $RTT = 200$ ms, $N = 10$. RED: $min_{th} = 5$, $max_{th} = 15$, $max_p = 0.1$, $w_q = 0.002$. VQ: $\theta = 0.95$. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6$, $\alpha = 0.71$.

On the other hand, in the case of virtual queue marking observe in Figure 3(a) 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.

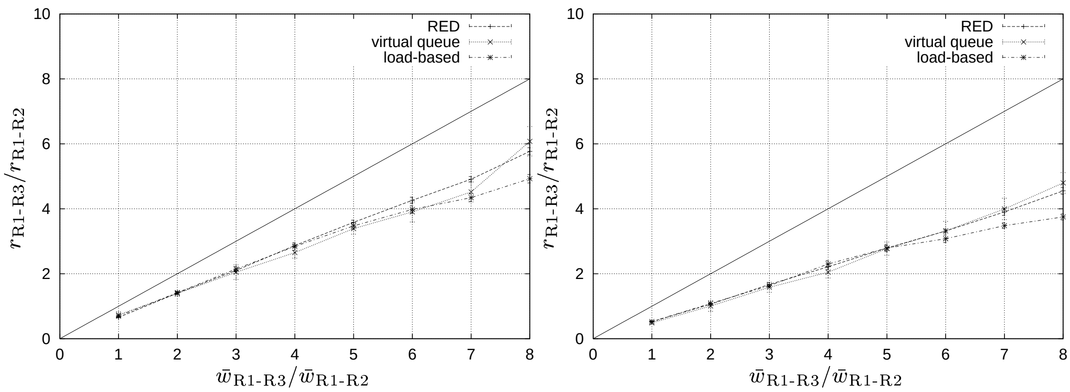
The first observation can be explained with the help of Figure 3(b), which shows the ratio of marks as a function of the ratio of average throughput. The figure shows that a smaller percentage of the packets belonging to a connection with a larger willingness-to-pay are marked, i.e., the marking probability 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; see also discussion in [8]. Hence, the algorithm marks 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 behavior of 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.

Now we explain the second observation made from Figure 3(a), namely the larger variations of service differentiation with virtual queue marking, compared to that of RED and load-based marking. The latter two algorithms mark packets probabilistically, hence marks tend to be spread out, which results in smoother changes of the congestion window. On the other hand, the virtual queue algorithm marks packets deterministically from the time the virtual queue overflows until the time it first becomes empty, hence tends to produce bursts of marks, which result in larger fluctuations of the congestion window.

Next we present simulation results for the multiple link topology in Figure 2(b). Figures 5(a) and 5(b) show the average ratio of throughput for different ratios of willingness-to-pay values for two connections: one that traverses only link R1-R2 in Figure 2(b), and one that traverses both links R1-R2 and R2-R3. Moreover, in the results of Figures 5(a) and 5(b) there are no connections that traverse only link R2-R3 (i.e., $N_{\text{R2-R3}} = 0$).

Our first observation regarding Figures 5(a) and 5(b) is that, as in the single link scenario, a connection's throughput can be controlled through its willingness-to-pay, and the performance of all three marking algorithms is roughly the same, at least for small ratios of willingness-to-pay. Indeed, the behavior of the virtual queue algorithm is similar to the other two because in the corresponding timescales of the virtual queue overflow, both connections exhibit the same burstiness.



(a) $RTT_{\text{R1-R2}} = 200$ ms (b) $RTT_{\text{R1-R2}} = 100$ ms

Fig. 5. Service differentiation for different RTT's (multiple link, proportional increase WTP). $\bar{\kappa} = 0.5$. $C_{\text{R1-R2}} = C_{\text{R2-R3}} = 10$ Mbps, $B_{\text{R1-R2}} = B_{\text{R2-R3}} = 30$ pkts, $RTT_{\text{R2-R3}} = 100$ ms, $N_{\text{R1-R2}} = N_{\text{R1-R3}} = 5$, $N_{\text{R2-R3}} = 0$. RED: $min_{th} = 5$, $max_{th} = 15$, $max_p = 0.1$, $w_q = 0.002$. VQ: $\theta = 0.95$. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6$, $\alpha = 0.71$.

The second observation regarding Figures 5(a) and 5(b) is that the line giving the ratio of throughput is lower than the diagonal. The reason is that the horizontal axis in Figures 5(a) and 5(b) depicts the ratio of willingness-to-pay values *per round trip time*, $\bar{w}_{\text{R1-R3}}/\bar{w}_{\text{R1-R2}}$. In the single link case, all connections had the same round trip time, hence the ratio $\bar{w}_{\text{R1-R3}}/\bar{w}_{\text{R1-R2}}$ was equal to the ratio of willingness-to-pay values *per unit of time*, $w_{\text{R1-R3}}/w_{\text{R1-R2}}$. On the other hand, in the multiple link case, the two connections considered experience a different round trip time: 200 msec in Figure 5(a) (100 msec in Figure 5(b)) for the connection traversing only link R1-R2, and 300 msec in Figure 5(a) (200 msec in Figure 5(b)) for the connection traversing both R1-R2 and R2-R3. Focusing on Figure 5(a), the round trip time for the connection traversing only link R1-R2 is 200 msec, hence its willingness-to-pay per round trip time is $\bar{w}_{\text{R1-R2}} = w_{\text{R1-R2}} \cdot 200$. On the other hand, the round trip time for the connection traversing both R1-R2 and R2-R3 is 300 msec, hence its willingness-to-pay per round trip time is $\bar{w}_{\text{R1-R3}} = w_{\text{R1-R3}} \cdot 300$. From the above, it follows that the ratio of willingness-to-pay per unit of time is $w_{\text{R1-R3}}/w_{\text{R1-R2}} = \frac{2}{3}\bar{w}_{\text{R1-R3}}/\bar{w}_{\text{R1-R2}}$. In the steady state, see (1), a connection's throughput is be proportional to the willingness-to-pay per unit of time, hence if the connections experience the same marking probability, we expect the ratio of throughput to be two thirds of the ratio $\bar{w}_{\text{R1-R3}}/\bar{w}_{\text{R1-R2}}$, a result that Figure 5(a) confirms. For Figure 5(b), we have $w_{\text{R1-R3}}/w_{\text{R1-R2}} = \frac{1}{2}\bar{w}_{\text{R1-R3}}/\bar{w}_{\text{R1-R2}}$, hence as the figure confirms, the ratio of throughput is one half of the ratio $\bar{w}_{\text{R1-R3}}/\bar{w}_{\text{R1-R2}}$.

One approach to take into account the round trip time is to set the willingness-to-pay per round trip time such that $\bar{w} \propto \frac{RTT}{RTT_{\text{ref}}} w$, where RTT_{ref} some reference round trip time. For example, if we consider the network scenario corresponding to Figure 5(a), we can take $RTT_{\text{ref}} = 200$ msec, and in order to achieve the same throughput on connections that traverse only link R1-R2 and connec-

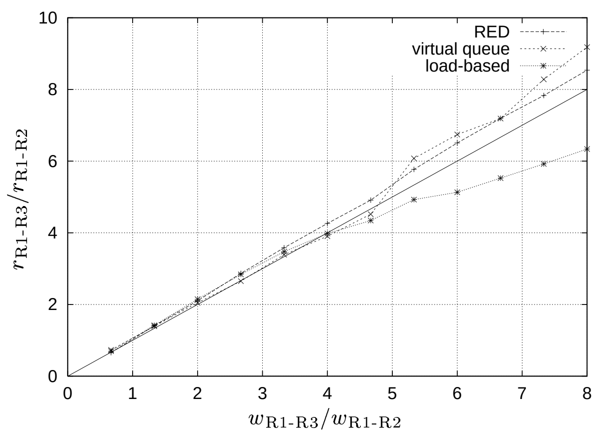
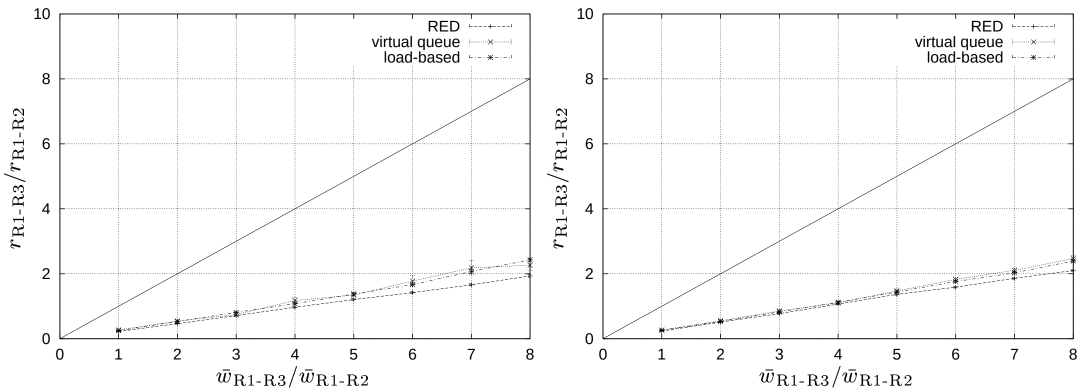


Fig. 6. Throughput ratio as a function of willingness-to-pay ratio (multiple link, proportional increase WTP). $\bar{\kappa} = 0.5$. $C_{\text{R1-R2}} = C_{\text{R2-R3}} = 10$ Mbps, $B_{\text{R1-R2}} = B_{\text{R2-R3}} = 30$ pkts, $RTT_{\text{R2-R3}} = 100$ ms, $N_{\text{R1-R2}} = N_{\text{R1-R3}} = 5$, $N_{\text{R2-R3}} = 0$. RED: $min_{th} = 5$, $max_{th} = 15$, $max_p = 0.1$, $w_q = 0.002$. VQ: $theta = 0.95$. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6$, $\alpha = 0.71$.

tions that traverse both links R1-R2 and R2-R3, we can set $\bar{w}_{\text{R1-R2}} \propto w_{\text{R1-R2}}$ and $\bar{w}_{\text{R1-R3}} \propto 1.5 \cdot w_{\text{R1-R3}}$, with the same proportionality factor in both cases. The results of such an approach are shown in Figure 6. Observe in this figure that, for large ratios of willingness-to-pay, the ratio of throughput for load-based marking tends to be smaller than that for the other two marking algorithms. This is because with load-based marking, the marking probability depends on the utilization, and when the willingness-to-pay values for the connections traversing both R1-R2 and R2-R3 becomes large, packets are also marked on link R2-R3, hence the marking probability experienced by connections that traverse both R1-R2 and R2-R3 is higher than the marking probability experienced by connections traversing only link R1-R2. (Indeed, it is for the same reason that the ratio of throughput for load-based marking in Figures 5(a) and 5(b) are lower than for the other two marking algorithms, for ratios of willingness-to-pay values larger than approximately 6).

Next, we present simulation experiments where there are three types of connections: connections that traverse only link R1-R2, only link R2-R3, and both links R1-R2 and R2-R3. The results are shown in Figure 7(a) for $RTT_{\text{R1-R2}} = 200$ msec, and Figure 7(b) for $RTT_{\text{R1-R2}} = 100$ msec. Similar to the above, these figures confirm that a connection can control its throughput through the willingness-to-pay parameter.

To illustrate the agreement between simulation results and those predicted by theory, consider a particular experiment where $\bar{w}_{\text{R1-R2}} = \bar{w}_{\text{R1-R3}} = 1$, and both links R1-R2 and R3-R4 implement RED. The ratio of the throughput of a connection traversing both link R1-R2 and R2-R3, and the throughput of a connection traversing only link R1-R2, $r_{\text{R1-R3}}/r_{\text{R1-R2}}$, and the marking probabilities at links R1-R2 and R2-R3 are shown in Table 1.



(a) $RTT_{\text{R1-R2}} = 200$ ms (b) $RTT_{\text{R1-R2}} = 100$ ms

Fig. 7. Service differentiation for different RTT's (multiple link, proportional increase WTP). $\bar{\kappa} = 0.5$. $C_{\text{R1-R2}} = C_{\text{R2-R3}} = 10$ Mbps, $B_{\text{R1-R2}} = B_{\text{R2-R3}} = 30$ pkts, $RTT_{\text{R2-R3}} = 100$ ms, $N_{\text{R1-R2}} = N_{\text{R1-R3}} = N_{\text{R2-R3}} = 5$. RED: $min_{th} = 5$, $max_{th} = 15$, $max_p = 0.1$, $w_q = 0.002$. VQ: $\theta = 0.95$. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6$, $\alpha = 0.71$.

In the equilibrium, from (1), we have that a connection's rate r will satisfy $r = \frac{w}{p}$ where w is the willingness-to-pay and p is the aggregate marking probability the connection encounters. For $RTT_{\text{R1-R2}} = 200$ msec, the connection traversing only link R1-R2 will have $w_{\text{R1-R2}} = \bar{w}_{\text{R1-R2}}/200 = 1/200$, whereas the connection traversing both links R1-R2 and R2-R3 will have $w_{\text{R1-R3}} = \bar{w}_{\text{R1-R3}}/300 = 1/300$. Accordingly, the aggregate marking probability encountered by each is $p_{\text{R1-R2}} = 0.023$ and $p_{\text{R1-R3}} = p_{\text{R1-R2}} + p_{\text{R2-R3}} = 0.065$. Hence, the ratio of throughput predicted by theory is

$$\frac{r_{\text{R1-R3}}}{r_{\text{R1-R2}}} = \frac{\frac{w_{\text{R1-R3}}}{p_{\text{R1-R3}}}}{\frac{w_{\text{R1-R2}}}{p_{\text{R1-R2}}}} = \frac{\frac{1}{300 \cdot 0.065}}{\frac{1}{200 \cdot 0.023}} \approx 0.236\,,$$

whereas the ratio of throughput for the case $RTT_{\text{R1-R2}} = 100$ msec is

$$\frac{r_{\text{R1-R3}}}{r_{\text{R1-R2}}} = \frac{\frac{w_{\text{R1-R3}}}{p_{\text{R1-R3}}}}{\frac{w_{\text{R1-R2}}}{p_{\text{R1-R2}}}} = \frac{\frac{1}{200 \cdot 0.084}}{\frac{1}{100 \cdot 0.042}} = 0.25\,,$$

both of which are in good agreement with the simulation results shown in the second column of Table 1.

Table 1
Multiple link experiment. $\bar{\kappa} = 0.5, \bar{w}_{\text{R1-R2}} = \bar{w}_{\text{R1-R3}} = \bar{w}_{\text{R2-R3}} = 1$. $C_{\text{R1-R2}} = 10 = C_{\text{R2-R3}} = 10$ Mbps, $RTT_{\text{R2-R3}} = 100$ ms, $N_{\text{R1-R2}} = N_{\text{R1-R3}} = N_{\text{R2-R3}} = 5$. RED: $min_{th} = 5$, $max_{th} = 15$, $max_p = 0.1$, $w_q = 0.002$.

$RTT_{\text{R1-R2}}$(msec)	$r_{\text{R1-R3}}/r_{\text{R1-R2}}$	$p_{\text{R1-R2}}$	$p_{\text{R2-R3}}$
200	0.231	0.023	0.042
100	0.251	0.042	0.042

Finally, we note that results from experiments with even more complex network topologies, including cases with different link capacities, agree with the results and conclusions concerning service differentiation presented in this subsection.

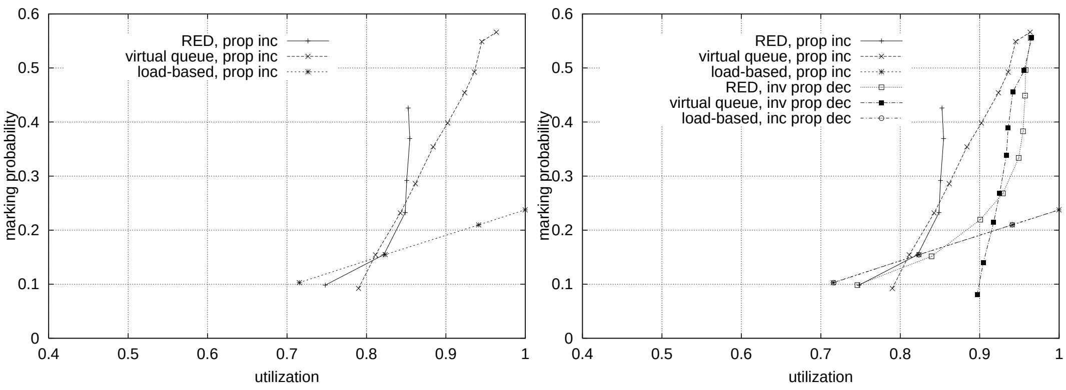
5.2 Marking Probability

Next we investigate the marking probability as a function of the average utilization, 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 [15], affects the convergence and stability behavior of the particular marking algorithm.

Figures 8(a) and 8(b) show the marking probability as a function of average utilization for the three marking algorithms, and for the proportional increase (4) and proportional decrease (6) WTP algorithms. 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 steeper as the utilization increases (we will see later that this is not always the case). Using Figure 8(a), we can assess the robustness of a marking algorithm with the load as follows: For RED configured with the specific parameters, the performance in term of convergence and utilization would be similar to that of the other two marking algorithms for utilization levels less than approximately 80%. However, for utilization levels above 85%, because the marking probability curve becomes very steep, RED would exhibit large fluctuations and stability can be compromised.

Figure 8(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 inversely proportional decrease WTP, the same marking probability is achieved for a higher utilization, compared to proportional increase WTP. This is expected, since inversely proportional decrease WTP results in smoother traffic, as shown in Figure 1, hence can achieve a higher utilization, for the same marking probability.

We have performed experiments for other round trip times, and when connections have heterogeneous round trip times. The results show that when κ, w_{inc} are kept constant, the marking probability curve is slightly higher for larger round trip times: the reason for this is that larger round trip times result in larger values for $\bar{\kappa} = \kappa T$ and $\bar{w}_{\text{inc}} = w_{inc} T$, which results in the window-based algorithm producing burstier traffic, resulting in a higher marking probability



(a) Proportional increase (b) Inversely proportional decrease

Fig. 8. Marking probability as a function of utilization 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, $B = 30$ pkts, $RTT = 200$ ms. RED: $min_{th} = 2$, $max_{th} = 6$, $max_p = 0.1$, $w_q = 0.02$. VQ: $\theta = 0.95$. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.6$, $\alpha = 0.71$. (single link)

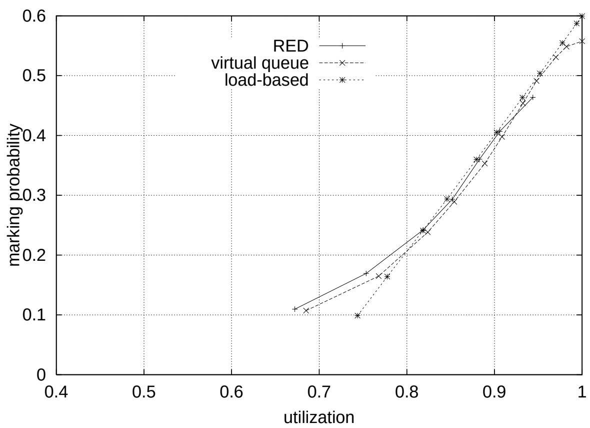


Fig. 9. Marking probability as a function of utilization (single link). $\bar{\kappa} = 0.5, \bar{w}_{\text{inc}} = 1$ and $6, \bar{w}_{\text{dec}} = 1$. $C = 10$ Mbps, $B = 30$ pkts, $RTT = 200$ ms. RED: $min_{th} = 3$, $max_{th} = 9$, $max_p = 0.95$, $w_q = 0.002$. VQ: $\theta = 0.95$. LB: $t_{avg} = 0.5$ s, $\rho_0 = 0.7$, $\alpha = 2$.

for the same utilization.

Figures 8(a) and 8(b) were obtained for some arbitrary [5] selection of parameters for each marking algorithm. Figure 9 shows that these parameters can be 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 behavior when their parameters are tuned appropriately. Note, however, that

[5] For RED we followed the guidelines suggested in the literature for max_p, and for the relative values of min_{th} and max_{th}.

Table 2
Queueing delay (in msec). Long ftp flows, $\rho \approx 0.90$.

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 ($\theta = 0.95$)	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

these curves are averages hence, as we will see next, do not give a complete picture of the queueing delay.

5.3 Queueing Delay

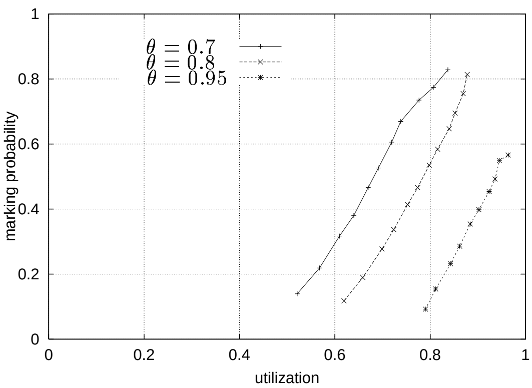
Table 2 shows the queueing delay for the three marking algorithms when the average utilization is kept the same. The bottom three rows in the table show that the queueing delay is smaller and exhibits smaller fluctuations for RED and load-based marking, compared to virtual queue marking. This is expected since, as discussed in Section 5.1, due to probabilistic marking, both RED and load-based marking smooth out marks, which results in smoother traffic; on the other hand, virtual queue marking produces bursts of marks.

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. Indeed, the first row corresponds to an operating point located at the steep region of the marking probability curve for RED, similar to the one shown in Figure 8(a).

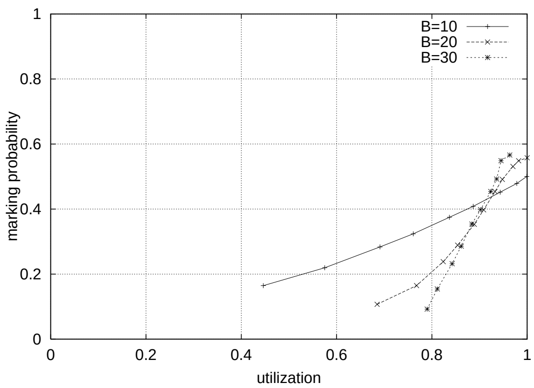
5.4 Effect of Marking Algorithm Parameters

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 of the marking algorithm affect the marking probability function.

Figure 10(a) shows how the marking probability for the virtual queue algorithm is affected by 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 probability. Also, observe that the factor has a very small effect on the



(a) factor θ



(b) buffer B

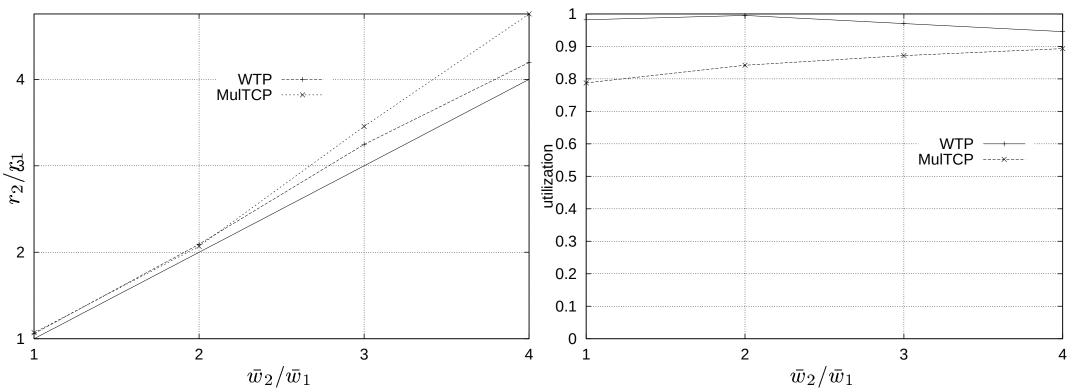
Fig. 10. Effect of the virtual queue factor θ and the buffer size (single link). $C = 10$ Mbps, $RTT = 200$ msec, $\bar{w}_{\rm inc} = 1$ and $6, \bar{w}_{\rm dec} = 1, \bar{\kappa} = 0.5$.

slope of the curve. Finally, Figure 10(b) shows that a larger buffer results in a steeper curve for the marking probability. [6] 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 utilization, and the averaging interval affects how fast the algorithm responds to changes of the load, and the timescales of traffic burstiness that affect the marking probability. Moreover, the averaging interval affects the maximum queue backlog that appears when the load changes.

Finally, for RED we have observed that when max_p is small, then the marking probability function is convex, as in Figure 8. On the other hand, when max_p approaches 1, then the marking probability function tends to be linear, as in Figure 9. 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.

[6] Varying the buffer size, while keeping the same factor for both the capacity and buffer of the virtual queue is equivalent to having a different factor for each of these two resources.



(a) Service differentiation (b) Average utilization

Fig. 11. Comparison of WTP and MulTCP (single link). WTP: $\bar{\kappa} = 0.5$. $C = 10\,\text{Mbps}, RTT = 200\,\text{msec}, N = 10$. RED Marking: $min_{th} = 5, max_{th} = 15, max_p = 0.1, w_q = 0.002$.

5.5 Comparison of WTP and MulTCP

Figure 11 compares WTP with MulTCP [3]. The latter follows TCP's double multiplicative decrease, but supports service differentiation. Figure 11(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.

Figure 11(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 than MulTCP. This is expected since MulTCP decreases its window more aggressively when congestion is detected. As the weights increase, however, the difference between the two 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 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 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 [22] 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. The loss adaptive algorithm in [22] 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 [20] investigate the convergence and steady state behavior, using both simulation and dynamical system modelling, of the rate-based congestion control algorithm given by (1), 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 [27] 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 overflow, 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 [18] 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.

Finally, the works of [5,11,19] present insightful results on the behavior of TCP and RED-like algorithms that compute the packet marking probability based on the queue length.

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 operating in the routers. Our simulation experiments included both single and multiple link topologies, where different connections had different round trip times. The congestion control algorithms [7] investigated include willingness-to-pay (WTP) congestion control, which follows a single multiplicative decrease of the congestion window upon congestion. Our results show that, for networks with no losses supporting Explicit Congestion Notification (ECN), WTP 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. Although our results where for experiments where no losses occurred, we expect that they will hold when the loss probability is much smaller than the marking probability; the latter is likely to be the case in real networks, in order to avoid the overhead due to packet retransmissions. In cases where the loss rate tends to be 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 can 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 is deterministic and occurs at the timescales of the virtual queue overflow, compared to RED and load-based marking, where marking is probabilistic. On the other hand, the virtual queue marking algorithm can differentiate flows according to their burstiness.

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. Another issue with important engineering implications, which can assist in tuning the parameters of a marking algorithm, is the estimation of the equilibrium given the congestion control and marking algorithms; a related issue, if each mark is charged by a fixed price, is how the price per mark affects the link utilization, and subsequently, given the aggregate demand, how to determine this price in order to achieve a target utilization. Results in this direction are reported in [25]. Another important direction for further work is the application of ideas and results investigated in this paper to wireless networks. This will be particularly important since there is a limited capability, compared to fixed networks, to increase the capacity of wireless networks and the emerging multimedia services and applications will increase the demand for bandwidth in wireless networks, hence the need

[7] The willingness-to-pay and packet marking algorithms have also been implemented in a real test-bed comprising of FreeBSD workstations. The implementations are available at `http://www.ics.forth.gr/netgroup/publications/wtp_vq.html`. Results from experiments conducted over the test-bed are reported in [2].

to develop efficient and robust mechanisms for controlling wireless network resources. The framework underlying the work presented in this paper is a strong candidate for constructing such mechanisms. Moreover, ECN marking for improving the performance of TCP over wireless networks has been suggested in [21], and for 3G networks in [12].

References

[1] M. Allman, V. Paxson, and W. Stevens. TCP Congestion Control. RFC 2581, April 1999.

[2] 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 Int. Symp. on Inform. Techn. 2000*, Boston, MA, 2000. Available at `http://www.ics.forth.gr/netgroup/publications/`.

[3] 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] B. Braden *et al.* Recommendations on queue management and congestion avoidance in the Internet. RFC 2309, April 1998.

[5] V. Firoiu and M Borden. A study of active queue management for congestion control. In *Proc. of IEEE INFOCOM'00*, April 2000.

[6] S. Floyd and V. Jacobson. On traffic phase effects in packet-switched gateways. *Internetworking: Research and Experience*, 3:115–156, 1992.

[7] S. Floyd and V. Jacobson. Random Early Detection gateways for congestion avoidance. *IEEE/ACM Trans. on Networking*, 1:397–413, 1993.

[8] R. J. Gibbens and F. P. Kelly. Resource pricing and congestion control. *Automatica*, 39:1969–1985, 1999.

[9] R. J. Gibbens and P. Key. Distributed control and resource marking using best-effort routers. *IEEE Network*, pages 54–59, May/June 2001.

[10] 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.

[11] C. V. Hollot, V. Misra, D. Towsley, and W.-B. Gong. On designing improved controllers for AQM routers supporting TCP flows. In *Proc. of IEEE INFOCOM'01*, April 2001.

[12] H. Inamura, G. Montenegro, M. Hata, J. James, A. Hamed, R. Garces, and F. Wills. TCP over 2.5G and 3G wireless networks. Internet Draft - draft-ietf-pilc-2.5g3g-00, February 2001.

[13] V. Jacobson and M. J. Karels. Congestion avoidance and control. In *Proc. of ACM SIGCOMM'88*, pages 314–329, 1988.

[14] F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–37, January 1997.

[15] 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/`.

[16] 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*, 49:237–252, 1998.

[17] 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.

[18] S. Kunniyur and R. Srikant. A time scale decomposition approach to adaptive ECN marking. In *Proc. of IEEE INFOCOM'01*, Anchorage, AK, April 2001.

[19] P. Kuusela, P. Lassila, and J. Virtamo. Stability of TCP-RED congestion control. In *Proc. of the 17th International Teletraffic Congress (ITC - 17)*, 2001. To appear.

[20] 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`.

[21] G. Montenegro, S. Dawkins, M. Kojo, V. Magret, and N. Vaidya. Long thin networks. RFC 2757, January 2000.

[22] 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.

[23] K. K. Ramakrishnan, S. Floyd, and D. Black. The Addition of Explicit Congestion Notification (ECN) to IP. RFC 3168, September 2001.

[24] J. Saltzer, D. Reed, and D. Clark. End-to-end arguments in system design. *ACM Trans. Comp. Sys.*, 2(4):277–288, November 1984.

[25] V. A. Siris and C. Courcoubetis. Interaction of congestion control and packet marking algorithms in ECN networks. Technical Report No. 296, ICS-FORTH, November 2001.

[26] UCB/LBNL/VINT. Network Simulator - ns (version 2). `http://www-mash.cs.berkeley.edu/ns.html`.

[27] D. Wischik. How to mark fairly. In *Proc. of Workshop on Internet Service Quality Economics*, MIT, December 1999.