# PROPERTIES OF THE VIRTUAL QUEUE MARKING ALGORITHM

R. J. Gibbens[1], P. B. Key[2] and S. R. E. Turner[3]


## Abstract

This paper explores the virtual queue marking algorithm in the context of distributed congestion control for packet networks. Queueing models and simulation experiments are used to describe the behaviour of the virtual queue marking algorithm under both slowly-varying and sudden changes in traffic demands.


## 1 Introduction

One important development goal for IP networks is the incorporation of a more sophisticated treatment of Quality of Service (QoS). This goal includes support for differential QoS for a wide range of traditional data applications together with real-time applications previously carried on entirely separate circuit-switched networks.

The current Internet relies on packet loss as the primary feedback signal, while TCP flows react to such signals by a rate reduction [5]. Explicit Congestion Notification (ECN) [1], which *marks* an IP packet if it encounters congestion, is an alternative feedback signal currently under discussion in the IETF. This approach separates the problem into two parts: adaptive user strategies on the part of increasingly intelligent end systems, and congestion feedback information relayed from the network resources to the end systems. Others have sought to go a step further, and used ECN as a key ingredient in a congestion pricing approach to QoS [4, 6].

Active queue management schemes, such as Random Early Detection [2], have been suggested as ways of determining congestion levels and hence marking behaviour. In this paper we explore a strategy based on the virtual queue marking algorithm first mentioned in [4], via simulations and an analytical approximation.

## 2 The model

### 2.1 The virtual queue marking algorithm

Figure 1 shows the features of the virtual queue marking algorithm. Packets arriving at the resource are added to the buffer unless the buffer occupancy would exceed $B$ packets. A packet not accepted by the buffer is lost. Packets are served from the buffer at the rate of 1 per unit time. Thus the choice of time unit is the service time of a packet at the resource.

The virtual queue algorithm for marking packets operates as follows. There is a virtual queue with service rate $\theta$, $0 < \theta \leq 1$, implemented as a (real-valued) counter. Packets that are carried by the real queue in the resource are also offered to the virtual queue. Such packets are accepted so long as the virtual queue is below $\theta B$. The virtual queue counter is incremented by 1 for each packet accepted, unless the counter then exceeds $\theta B$ in which case it is capped to the value $\theta B$. Any packet that would have lead to the virtual queue exceeding the value $\theta B$ is marked. The marks are returned to the user after some delay depending on the user's route. In addition, any packet lost by the real resource is treated by the user as a mark returned to the user after the same amount of delay. Notice that the case $\theta = 1$ corresponds precisely to marking according to a threshold rule on the (real) queue length.

[1]Computer Laboratory, University of Cambridge. Supported by the EPSRC under grant GR/M09551 and by a Royal Society University Research Fellowship.

[2]Microsoft Research, Cambridge.

[3]Statistical Laboratory, University of Cambridge. Supported by the EPSRC under grant GR/M09551.

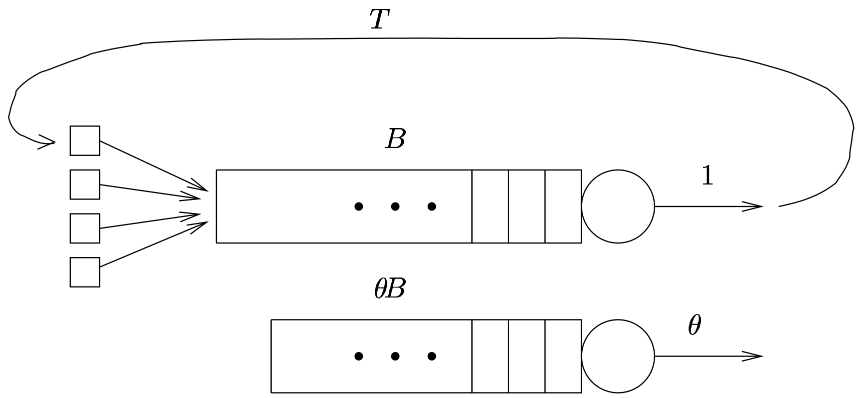


Figure 1: Virtual queue marking. The resource has the capacity to service 1 packet per unit time and can hold $B$ packets in a buffer. The virtual queue serves its packets at rate $\theta$ ($0 < \theta \leq 1$) per unit time and has a maximum length of $\theta B$. There is a delay of $T$ time units after a packet is served before the feedback is received by the user.

The intuition behind this algorithm is that it aims to give early warning of congestion by marking packets whenever the rate of offered traffic approaches $\theta$, the service rate of the virtual queue, rather than 1, the service rate of the real queue. A virtue of the algorithm is that it is parsimonious; essential trade-offs between loss and utilisation can be tuned by the single parameter $\theta$.

### 2.2 The user population

We consider two types of users, elastic and unresponsive, both of which were introduced in [4].

The elastic user rate control strategy is reminiscent of, but different from, TCP. Each elastic user has two parameters, $w$ and $\kappa$. $w$ is known as the willingness-to-pay parameter, and represents the average number of marks which the user is prepared to receive per unit time; $\kappa$ is a positive gain parameter. The user transmits $X(t) = \lfloor x(t) + z(t) \rfloor^+$ packets in the time slot $(t, t+1)$ where $x(t)$ and $z(t)$ are updated as follows:

$$z(t+1) = x(t) + z(t) - X(t), \qquad x(t+1) = x(t) + \kappa(w - f(t))$$

and $f(t)$ is the number of marks received at the end of the slot $(t, t+1)$. By adapting their sending rates in this way, and with a suitably-chosen $\kappa$, the users can ensure both stable operation and that they receive a long-run average of precisely $w$ marks per unit time.

The unresponsive users, in contrast to the elastic users, pay no attention to feedback information on the level of congestion. The unresponsive users alternate between active and idle periods, which have independent geometric distributions. When active (*on*), they send a packet at random in each time slot with probability $p$, independently for each slot. When idle (*off*) they send no packets.

## 3 The M/D/1/B approximation

As a simple analytical model of the system, we approximate the arriving traffic by a Poisson process. The rate of this Poisson process will be determined by the long-term number of marks generated, but the approximation ignores the short-term effect of the users' response to the feedback signals.

Under this approximation, the server is just an $M/D/1/B$ queue. Consider an $M/D/1/B$ queue with arrival rate $\lambda$ and a service every $m$ units of time. Let $\rho = \lambda m$. Then the number of packets in the queue immediately after each service forms a Markov process, and it is a simple exercise to calculate the stationary distribution of this process, and thus the packet loss rate, which we shall denote by $L(\rho, B)$ per service time or $L(\rho, B)/m$ per unit time, the packet loss probability $L(\rho, B)/\rho$, and the utilisation of the server $\rho - L(\rho, B)$.

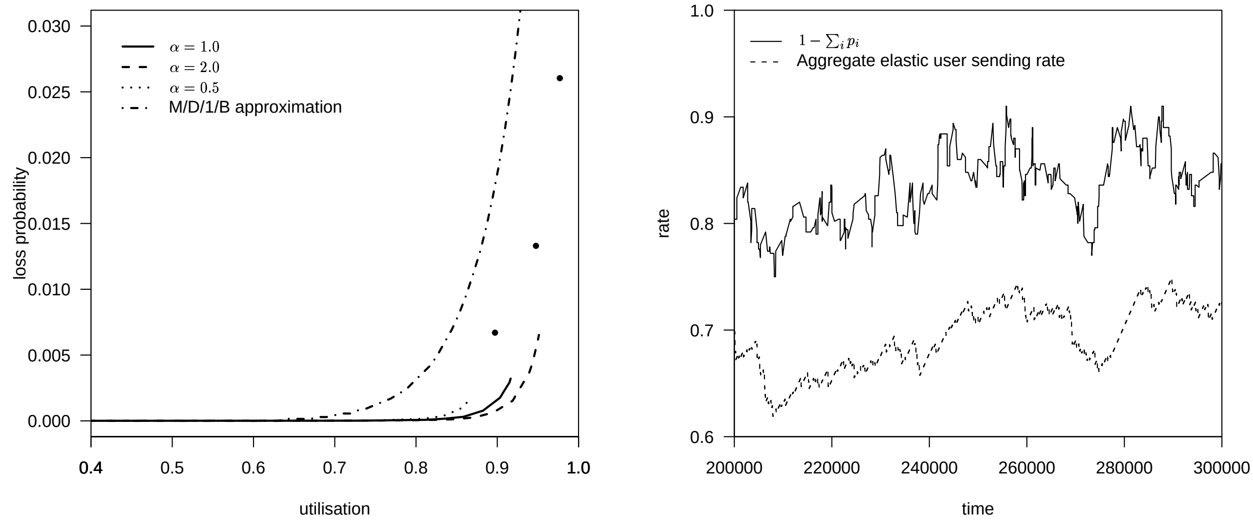


Figure 2: Experimental results with slowly varying demand. The left hand panel shows the loss-utilisation curves corresponding to the $M/D/1/B$ model and to the simulations as $\theta$ varies. The right hand panel shows a measure of fluctuating levels of capacity available to the elastic users together with their aggregate sending rate.

To calculate the arrival rate $\lambda$ in our system, we consider the marks generated by the virtual queue. We shall model the virtual queue as if it were another $M/D/1/B$ queue with capacity $\theta B$, arrival rate $\lambda$, and service time $1/\theta$. This is a further approximation because the virtual queue actually serves $\theta$ packets each time unit, whereas the $M/D/1/B$ approximation serves 1 packet every $1/\theta$ time units. The approximation thus takes only integer values, whereas the virtual queue can take real values; indeed, the capacity $\theta B$ may not even be an integer and we have to interpolate between the adjacent integer values. Nevertheless, this gives an approximation, $L(\rho, \theta B)/\rho$, to the marking probability in the virtual queue when the traffic is Poisson of rate $\lambda$, where now $\rho = \lambda/\theta$.

Now the elastic users react to marks so as to ensure that their long-term marking rate is $w_+$, where $w_+$ is the sum of all the individual users' $w$'s. Suppose that there is also a set of unresponsive users, generating an average aggregate load of $\lambda_U$. Then we choose $\lambda$ to satisfy

$$(\lambda - \lambda_U)\frac{L(\rho, \theta B)}{\rho} = w_+,$$

where $\rho = \lambda/\theta$, since the left hand side is the product of the packet arrival rate of elastic users and the packet marking probability, and this must equal the packet marking rate, $w_+$.

This is the value of $\lambda$ we use to estimate the packet loss probability and utilisation of the queue, using the formulae $L(\rho, B)/\rho$ and $\rho - L(\rho, B)$, where this time $\rho = \lambda$ because $m = 1$ in the real queue. Figure 2 shows the loss probability and utilisation given by the $M/D/1/B$ approximation.

Now note that this loss-utilisation curve is just the loss-utilisation curve of an $M/D/1/B$ queue as $\rho$ varies. In other words, the curve does not depend on the details of the marking scheme used — any marking scheme would give the same curve, under this approximation. This is because the traffic used for the model is non-adaptive, given the arrival rate; with a good marking scheme we would hope that the system could obtain better performance than this curve (that is, below and to the right), by sending more packets at times when the system is less busy. Furthermore, we expect a finite population of users to have less variability than that implied by the Poisson assumption, and this could also improve the performance of the system. A finite user population model was considered in [3].

The approximation does, however, give guidance to the network operator as to the choice of $\theta$ to achieve a chosen point on the loss-utilisation curve. As $\theta$ is increased towards 1 both the packet loss and utilisation are increased. Thus the parameter $\theta$ allows a network operator to tune the system to achieve the desired trade-off between packet loss and server utilisation.

# 4 Experimental results

The simulation environment consisted of a collection of users attached to a single resource operated in slotted time, where each time slot is equal to the time to serve one packet. The resource had a buffer capacity of $B = 10$ packets, and operated the virtual queue marking algorithm. All users had an identical delay of $T = 100$ time units before receiving the mark information.

Two simulation scenarios were considered. In the first, the demand on the resource was slowly varying over time in order to capture both busy and quiet periods. In the second, the demand remained static until a sudden shock occurred caused by a large user joining the system.

## 4.1 Response to slowly-varying demand

The first simulation scenario concerned the response of the resource to slowly-varying demands. In the first series of experiments the user population consisted of both elastic users and unresponsive users. There were 20 elastic users with willingness-to-pay parameters $w$ evenly spread in the range $5 \times 10^{-5}$ to $1 \times 10^{-3}$. The aggregate willingness-to-pay of the elastic user population was therefore 0.0105 marks per unit time. Each elastic user had a gain parameter of $\kappa = 10^{-3}$. In addition, there were 30 unresponsive users, comprising 3 copies of each of 10 $p$ values evenly spread in the range $2 \times 10^{-3}$ to $2 \times 10^{-2}$. The unresponsive users had mean on and off times of length $10^4$ time slots. Thus the mean level of demand created by the unresponsive users was given by $\lambda_U = \sum_i p_i/2 = 0.165$. Two further series of experiments were conducted with the willingness-to-pay parameters of the elastic users scaled by factors of $\alpha = 0.5$ and $\alpha = 2.0$.

The left hand panel in Figure 2 shows the experimental results. The simulations for each fixed value of $\theta$ were run for $10^6$ time slots. The simulations were repeated with different random number seeds and the estimates of loss probability and utilisation were found to be very tightly clustered around the values shown (an estimated confidence interval would be of order the width of the line). The curves show trajectories traced out as $\theta$ varies. Note that each curve is discontinuous at $\theta = 1$. Consideration of the sample paths shows that when $\theta$ is fractionally less than 1, loss occurs whenever the real queue attempts to exceed $B$, but in addition, marking occurs whenever the real queue *reaches* $B$. Thus even applying a $\theta$ which is only fractionally less than 1 produces substantially earlier feedback to the users than for $\theta = 1$, causing them to reduce their rate sooner, and so incur lower loss and lower throughput, and hence less server utilisation.

The loss-utilisation curve depends on the level of demand given by the scaling parameter $\alpha$. The curve also depends on other factors such as the relative mix of user types, the degree of departure from Poisson traffic statistics, the variability in aggregate demand caused by users turning on and off as well as on network effects such as those caused by the delays to congestion feedback information. A more refined approach than the one used in the $M/D/1/B$ approximation would be required to model these effects on the precise form of the loss-utilisation characteristic of the system. For examples of such approaches, see [3, 7].

The right hand panel of Figure 2 shows the varying demand during a part of the simulation. The solid line shows $1 - \sum_i p_i$ (summed over the active unresponsive users). The expression $\sum_i p_i$ gives the fluctuating level of demand from the unresponsive users and so $1 - \sum_i p_i$ indicates the level of resource capacity available to the elastic users. The dashed line gives the aggregate sending rate of the elastic users and can be seen to track the available capacity over time.

## 4.2 Response to sudden change in demand

The second simulation scenario concerned the response of the system to a sudden change in the demand. The experiment initially started with the same population of 20 elastic users (but no unresponsive users) until time slot 0 when a single unresponsive user with $p = 0.3$ joined the system and remained in its on state for the remaining duration of the experiment.

Figure 3 shows in the top left hand panel the response of the system to the sudden increase in demand at time slot 0. Define the shadow price to be the proportion of packets marked or lost. The figure shows that the shadow price, and the real and virtual queue lengths, increase suddenly when the unresponsive user joins the system. There is also considerably increased variability in the queue lengths. The top right hand panel of Figure 3 shows the cumulative numbers of marked and loss packets over time, gathered since time $-4000$ slots. It can be seen that the sudden shock at time 0 causes many more packets to be marked during a relatively narrow interval of approximately 1400 time slots before the system returns to the initial rate of marking. During this narrow interval around 50 packets are lost before the elastic users have had time to reduce their sending rates accordingly.

The lower panel in Figure 3 shows the effect on the aggregate source rate of the 20 elastic users. The increase in congestion causes them to back off and reduce their sending rates to a lower level consistent with maintaining their long-run willingness to pay for congestion marks.

Observe that the source rates of the elastic users appear to become synchronised with a period of around 2000 time slots. This may be due to the sudden shock causing the elastic users to back off together and thereafter react in a synchronised manner. Note, however, that no packets were actually lost during this phase. Such synchronisation effects would certainly be reduced if the elastic users also adopted random on/off behaviour.

The speed of adaptation is determined in part by the users' strategy through the value of $\kappa$ and in part by the form of the marking algorithm. For example, a marking algorithm based on a virtual queue with infinite virtual queue length would continue marking packets for much longer than the algorithm described above and hence induce the elastic users to back off much further.

## 5 Conclusion

The virtual queue is an attractive candidate for a marking strategy, since it provides early warning of congestion, and depends on only a single parameter, $\theta$. Our experiments show that the virtual queue marking algorithm, together with the response of the elastic users, produces robust system behaviour in response to both slowly-varying and sudden changes in demand. The loss-utilisation curve is a key performance measure of the system, and the network operator can choose the desired trade-off between loss and utilisation by tuning $\theta$ to attain a point on this curve.

## References


[1] S. Floyd. TCP and Explicit Congestion Notification. *ACM Computer Communications Review*, 24:10–23, 1994. http://www-nrg.ee.lbl.gov/floyd/ecn.html.

[2] S. Floyd and V. Jacobson. Random Early Detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1:397–413, 1993. ftp://ftp.ee.lbl.gov/papers/early.pdf.

[3] R. J. Gibbens and F. P. Kelly. Distributed congestion acceptance control for a connectionless network. In *Teletraffic Engineering in a Competitive World, Proceedings of the 16th International Teletraffic Congress — ITC 16*. Elsevier, June 1999.

[4] R. J. Gibbens and F. P. Kelly. Resource pricing and the evolution of congestion control. *Automatica*, 35:1969–1985, 1999.

[5] V. Jacobson. Congestion avoidance and control. In *Proc. ACM SIGCOMM '88*, pages 314–329, 1988. A revised version, joint with M.J. Karels, is available via ftp://ftp.ee.lbl.gov/papers/congavoid.ps.Z.

[6] Peter Key, Derek McAuley, Paul Barham, and Koenraad Laevens. Congestion pricing for congestion avoidance. Microsoft Research Technical Report MSR-TR-99-15, 1999. http://research.microsoft.com/research/network/publications/MSRTR99_15.pdf.

[7] Koenraad Laevens, Peter B. Key, and Derek McAuley. An ECN-based end-to-end congestion-control framework: experiments and evaluation. Microsoft Research Technical Report MSR-TR-2000-104, 2000. http://research.microsoft.com/research/network/publications/MSRTR2000_104.pdf.

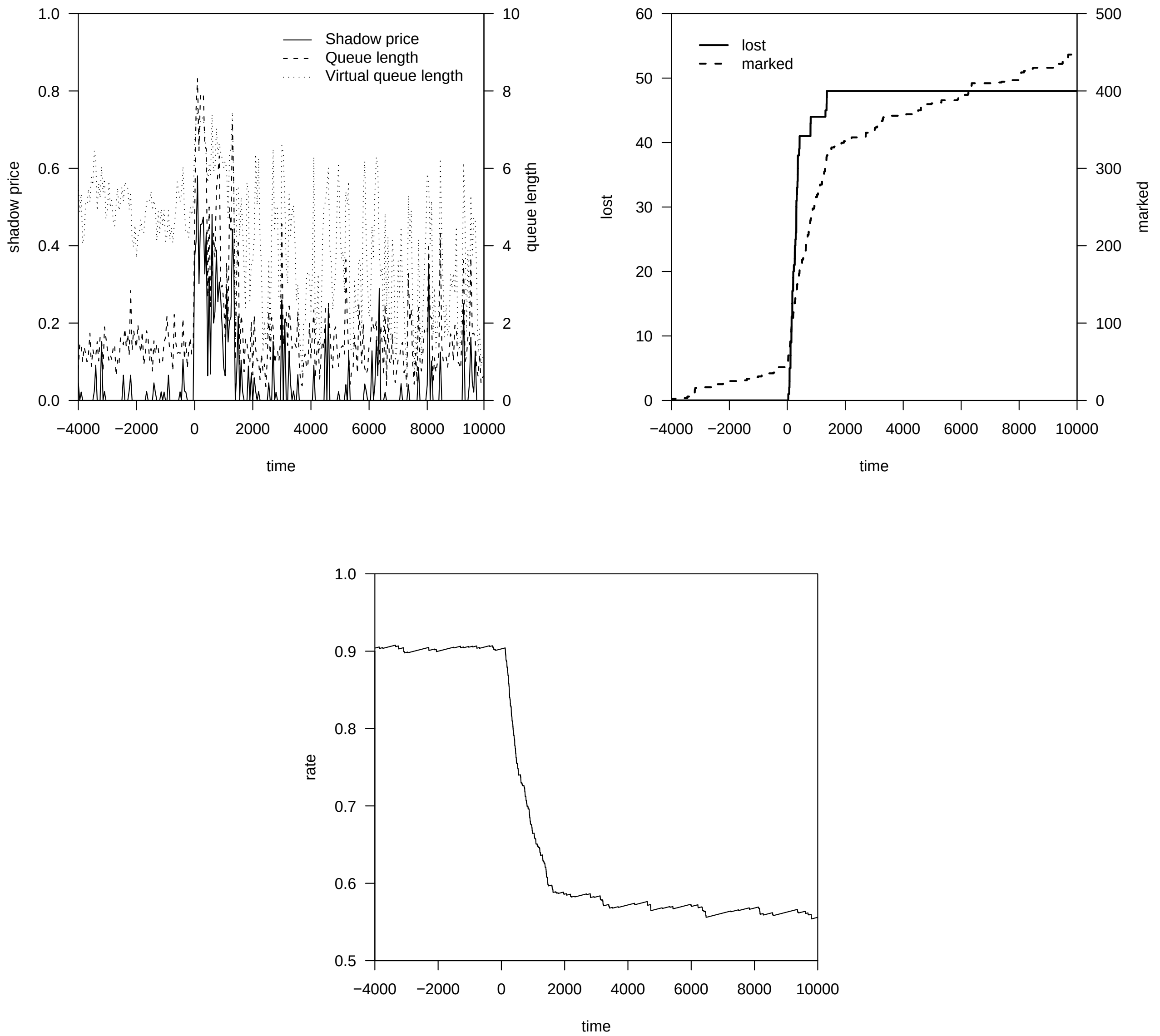


Figure 3: Experimental results with a sudden change in demand. The top left hand panel shows the effect of a single unresponsive user joining the system at time 0 on the lengths of the real and virtual queues. The shadow price for congestion is also shown. The top right hand panel shows the cumulative numbers of lost and marked packets over time. The sudden introduction of an unresponsive user causes a sharp increase in the marking rate. However, after a short interval of time, the marking rate returns to its long-run value. The lower panel shows the effect on the elastic users: their aggregate sending rate rapidly falls to a new lower value consistent with the long-run marking rate and the extra level of demand present from time 0 onwards.