An Analytical Model of a new Packet Marking Algorithm for TCP flows*

Giovanni Neglia[1], Vincenzo Falletta[1] and Giuseppe Bianchi[2]

[1] Dipartimento di Ingeneria Elettrica (DIE), Università degli Studi di Palermo, Italy
Email: {giovanni.neglia, vincenzo.falletta}@tti.unipa.it
[2] Dip. di Ingegneria Elettronica, Università degli Studi di Roma "Tor Vergata", Italy
Email: bianchi@elet.polimi.it

Abstract. In Differentiated Services networks, packets may receive a different treatment according to their Differentiated Services Code Point (DSCP) label. As a consequence, packet marking schemes can be devised to differentiate packets belonging to a same TCP flow, with the goal of improving the experienced performance. This paper presents an analytical model for an adaptive packet marking scheme proposed in our previous work. The model combines three specific sub-models aimed at describing *i*) the TCP sources aggregate *ii*) the marker, and *iii*) the network status. Preliminary simulative results show quite accurate predictions for throughput and average queue occupancy. Besides, the research suggests new interesting guidelines to model queues fed by TCP traffic.

1 Introduction

Differentiated Services (DiffServ) networks provide the ability to enforce a different forwarding behavior to packets, based on their Differentiated Services Code Point (DSCP) value. A possible way to exploit the DiffServ architecture is to provide differentiated support for flows belonging to different traffic classes, distinguished on the basis of the DSCP employed. However, since it is not required that all packets belonging to a flow are marked with the same DSCP label, another possible way to exploit DiffServ is to identify marking strategies for packets belonging to the same flow.

Several packet marking algorithms have been proposed for TCP flows. The marking strategy is enforced at the ingress node of a DiffServ domain (edge router). Within the DiffServ domain, marked packets are handled in an aggregated manner, and receive a different treatment based on their marked DSCP. Generally, a two-level marking scheme is adopted, where packets labelled as IN receive better treatment (lower dropping rate) than packets marked as OUT. Within the network, dropping priority mechanisms are implemented in active queue management schemes such as RIO - Random Early Discard with IN/OUT packets [1].

* This work was funded by the TANGO project of the FIRB programme of the Italian Ministry for Education, University and Research.

The basic idea of the proposed algorithms is that a suitable marking profile (e.g. a token bucket which marks IN/OUT profile packets) may provide some form of protection in the case of congestion. A large number of papers [1–5] have thoroughly studied marking mechanisms for service differentiation, and have evaluated how the service marking parameters influence the achieved rate.

More recently, TCP marking has been proposed as a way to achieve better than best effort performance [6–8]. The idea is that packet marking can be adopted also in a scenario of homogeneous flows (i.e. all marked according to the same profile), with the goal of increasing the performance of all flows. In particular, [6, 7] consider long lived flows and adopt goodput and loss as performance metrics. Conversely, [8] focuses on WWW traffic, mostly characterized by short-lived TCP flows, and proposes a new scheme able to reduce the completion time of an http session.

In all the above mentioned marking schemes, most of the packets in the network are of type OUT. Hence, packets marked as IN will be protected against network congestion (indeed [8] relies on this property to protect flows with small window, when packet losses cannot be recovered via the fast retransmission algorithm). As shown in section 2, our marking strategy is based on a somehow opposite philosophy.

In this paper we slightly modify the mechanism proposed in [9], and we describe an analytical model to evaluate the network performance. This model can be employed to study possible variants of the algorithm. By the way, the network sub-model exhibits some novelty in comparison to previous approaches and could be useful in different network scenarios where TCP traffic is considered.

The rest of this paper is organized as follows. Section 2 describes our adaptive packet marking algorithm, focusing on some changes to the previous version. Section 3 presents the analytical model, the three submodels are detailed respectively in subsections 3.1, 3.2, 3.3, while in subsection 3.4 existence and uniqueness of a solution are proven. Section 4 deals with validation of the proposed model. Finally, conclusive remarks and further research issues are given in section 5.

2 The Packet Marking Algorithm (PMA)

In [9, 10] we proposed a new marking algorithm, able to achieve better performance in terms of average queueing delay and flow completion time versus link utilization. According to this marking scheme "long" IN-packets bursts are interleaved with a single OUT packet. The OUT packet is thence employed as a *probe* to early reveal a possible seed of congestion in the network. The algorithm dynamically updates the length of IN-packets bursts by a heuristic estimation of the experienced packet loss ratio.

The idea of marking the majority of packets as IN seems to be in contrast with some results found with other marking scheme [6–8], but the intrinsic adaptivity of our algorithm is something all these models lack.

If we think about Active Queue Management (AQM) techniques such as Random Early Detection (RED) we observe the same idea of dropping some

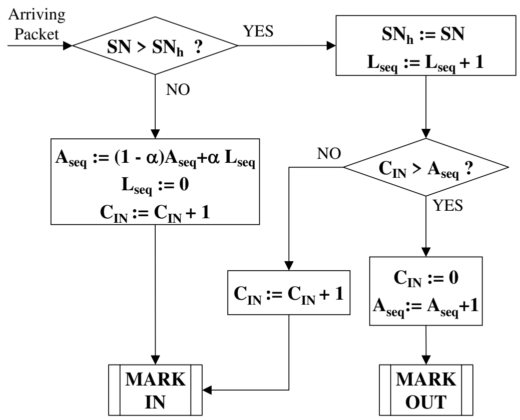

Fig. 1. PMA Flow diagram

packets when signals of an incoming congestion are received. Our algorithm moves further: it reallocates losses among the OUT packets, so it spaces them as much as possible, avoiding consecutive losses for a flow and assuring a more regular TCP adaptation behavior.

By simulative evaluation we found better performance when OUT-packets dropping probability is near 100%, while IN packets are not dropped at all.

The algorithm flowchart is shown in Fig. 1. Now we will explain how this procedure works. Each time a new SYN packet arrives at the edge router a new state vector is set, containing the following variables:

SN_h: This counter stores the highest Sequence Number (SN) encountered in the flow. It is initially set to the ISN (Initial Sequence Number) value. It is updated whenever a non-empty packet (i.e. non ACK) arrives with a higher SN.

L_{seq}: It is initially set to zero. It is increased by one unit for each new arrived packet (i.e. in-sequence packet), while is reset to zero every time an out-of-sequence packet arrives.

A_{seq}: It stores the average length of in-sequence packet burst between two consecutive losses, using an auto-regressive filter on the previous values of L_{seq}.

C_{IN}: It counts the number of IN-packets in the burst. It is reset to zero when it exceeds A_{seq} and an OUT packet is sent.

The algorithm has been slightly changed in comparison to the version presented in [9, 10]. In the previous algorithm a single variable (L_{IN}) was taking into account the number of in-sequence packets (as L_{seq} actually does) and the number of IN packets of the actual IN-packets burst (as C_{IN} actually does). This coupling required an artificial increase of the variable A_{IN} after marking an OUT packet, we chose $A_{IN} := 2A_{IN} + 1$ but its correct amount was dependant from network condition as it is discussed in [9, 10]. After the introduction of the new variable C_{IN}, a small increase of A_{IN} has been left: it assures better fairness among the flows, allowing flows with underestimated A_{IN} values to faster reach the correct estimate.

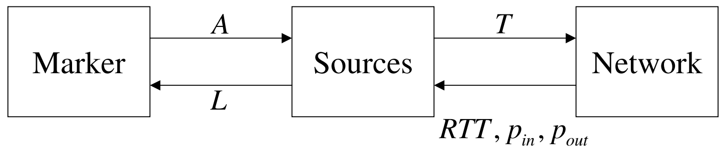

Fig. 2. The three-block model

3 The analytical model

The algorithm has shown good performance, but it essentially relies on a heuristic. In order to achieve a deeper understanding and to establish RIO setting criteria, we have developed an analytical model.

The actual model assumes n long-lived homogeneous flows sharing a common bottleneck, whose capacity is C. The model is based on three submodels, which describe respectively the TCP sources, the marker and the network status. Fig. 2 shows the relation among these elements. The number of in-sequence packets L_{seq} is employed by the marker to calculate A_{seq}. The TCP flows are marked according to A_{seq}, hence A_{seq} affects TCP throughput. The TCP traffic congests the network, and produces not-empty queues. Increase of the Round Trip Time and packet losses act as feedback signals for the TCP sources. According to the fixed point approach the influence of each block on the others is considered constant, equal to the average value. The average values for the throughput, A_{seq}, L_{seq} and the Round Trip Time are respectively indicated in Fig. 2 as T, A, L and RTT. At the same time the losses are taken into consideration only through the average dropping probabilities for IN and OUT packets, p_{in} and p_{out}.

We discuss the three submodels in the following subsections. Each of them could be replaced by a more sophisticated one.

3.1 The Sources Model

According to the previous description, we aim to obtain an expression of the average TCP throughput (T, the input to the Network block) and of the average length of the in-sequence packet burst (L, the input to the Marker block), given the marking profile (A) and the network status (RTT, p_{in}, p_{out}). We have conjectured a regenerative process for TCP congestion window ($cwnd$), thus extending the arguments in [12] to include two different service classes, with different priority levels.

We only consider loss indications due to triple duplicated acks, which turn on TCP fast retransmit mechanism. We don't consider in our analysis the fast recovery mechanism neither the time-out loss events, for the sake of simplicity. As regards time-out neglecting, this approximation appears to be not critical because PMA spaces OUT packets and hence loss events. For this reason errors are usually recovered by fast retransmission, not by time-out. Such intuition is confirmed by our simulation results, where the number of time-outs appear to be significantly reduced in comparison to a no-marker scenario.

A period of our regenerative process starts when the sender congestion window is halved due to a loss indication. Figure 3 shows *cwnd* trend as rounds succeed. W_{i-1} is the *cwnd* value at the end of the $(i-1)$-th period, hence in the i-th period *cwnd* starts from $W_{i-1}/2$ and it is incremented by one every b rounds (b is equal to 2 or 1, respectively if the receiver supports or not the delayed ack algorithm). Notice that, due to neglecting fast recovery and timeouts, each period starts with an IN retransmitted packet, hence the number of packets sent in the period (Y_i) is equal to $L_{seq}+1$, according to the marker description in section 2.

In the i-th period we define also the following random variables: I_i is the length of the period; β_i is the number of packets transmitted in the last round; α_i is the number of the first lost packet since the beginning of the period, while γ_i is the number of packets transmitted between the two losses occurred in the $(i-1)$-th and in the i-th period. We get $Y_i = \alpha_i + W_i - 1$ and $\alpha_i = \gamma_i - (W_{i-1} - 1)$.

Due to the renewal-reward theorem we can obtain the expression for the average throughput of n sources sharing the same path:

$$T(A, RTT, p_{in}, p_{out}) = n\frac{E[Y_i]}{E[I_i]}$$

We first compute $E[Y_i]$. The relation between α_i and γ_i allows us to obtain $E[Y_i]$ as a function of the marking profile (A) and the network status (in particular p_{in}, p_{out}). In general $Y_i \neq \gamma_i$, however if we consider their mean values, it holds:

$$E[Y_i] = E[\alpha_i] + E[W_i] - 1 = E[\gamma_i] - (E[W_{i-1}] - 1) + E[W_i] - 1 = E[\gamma_i]$$

Let us denote by N the expected value $E[\gamma_i]$. We compute N as:

$$N = \sum_{n=0}^{\infty} np(n) = \sum_{n=0}^{\infty} (1 - P(n)) = \sum_{n=0}^{\infty} Q(n)$$

where $p(n)$ is the probability of losing the n-th packet after $(n-1)$-th successful transmission, $P(n) = \sum_{l=0}^{n} p(l)$ is cumulative distribution function, and so $Q(n) = 1 - P(n)$ represents the probability of *not* losing any packet among these n. If we put n as $n = k(A+1) + h$, with $0 \leq h < (A+1)$ we can write $Q(n)$ as

$$Q(n) = s_{in}^{kA+h} s_{out}^{k}$$

where $s_{in} = 1 - p_{in}$, $s_{out} = 1 - p_{out}$. Then N is equal to:

$$N = \sum_{k=0}^{\infty} \sum_{h=0}^{A} s_{in}^{kA+h} s_{out}^{k} = \frac{s_{in}^{A+1} - 1}{s_{in} - 1} \frac{1}{1 - s_{in}^{A} s_{out}} \tag{1}$$

Now we compute $E[I_i]$. Denoting with X_i the round in the i-th period when a packet is lost, we obtain the period length as $I_i = \sum_{j=1}^{X_i+1} r_{ij}$, where $r_{i,j}$ is the

j-th round trip time length. Supposing r_{ij} independent of the round number j (i.e. independent of *cwnd* size), taking expectation we find

$$E[I_i] = (E[X] + 1)E[r]$$

where $E[r] = RTT$ is average round trip time.

In the i-th period *cwnd* size grows from $W_{i-1}/2$ to W_i with linear slope $1/b$, so[3]

$$W_i = \frac{W_{i-1}}{2} + \frac{X_i}{b} - 1$$

and taking expectation we get

$$E[W] = \frac{2}{b}\left(E[X] - b\right)$$

To simplify our computations we assume $W_{i-1}/2$ and X_i/b to be integers. Now let us count up all the packets:

$$\begin{aligned} Y_i &= \sum_{k=0}^{X_i/b-1} \left(\frac{W_{i-1}}{2} + k \right) b + \beta_i = \frac{X_i W_{i-1}}{2} + \frac{X_i}{2}\left(\frac{X_i}{b} - 1 \right) + \beta_i \\ &= \frac{X_i}{2}\left(W_{i-1} + \frac{X_i}{b} - 1 \right) + \beta_i = \frac{X_i}{2}\left(W_i + \frac{W_{i-1}}{2} \right) + \beta_i \end{aligned}$$

and taking again expectation it follows

$$N = \frac{E[X]}{2}\left(E[W] + \frac{E[W]}{2} \right) + E[\beta]$$

Assuming β identically distributed between 1 and $W_i - 1$ we can write $E[\beta] = E[W]/2$; therefore, solving for $E[X]$:

$$E[X] = \frac{3b-2}{6} + \sqrt{\frac{2bN}{3} + \left(\frac{2+3b}{6} \right)^2}$$

then it follows

$$E[I_i] = RTT\left(\frac{3b-2}{6} + \sqrt{\frac{2bN}{3} + \left(\frac{2+3b}{6} \right)^2} + 1 \right)$$

Now we can write down the throughput formula:

$$T(N, RTT) = n\frac{N}{RTT(E[X]+1)} = n\frac{N}{RTT}\,\frac{1}{\frac{3b-2}{6} + \sqrt{\frac{2bN}{3} + \left(\frac{2+3b}{6}\right)^2} + 1} \tag{2}$$

[3] There are actually different ways to represent *cwnd* linear growth above the i-th period in the continuous; period bounds are chosen respectively at the beginning of the first round and at the end of the last round, but while in [12] *cwnd* starts from $W_{i-1}/2$ at the beginning of the period, in our analysis it reaches $W_{i-1}/2$ only after $b/2$ rounds.

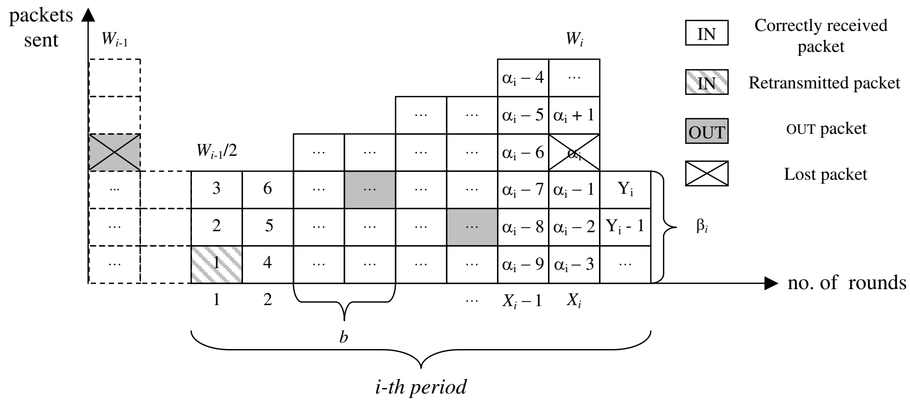

Fig. 3. Timeline and transmitted packets

Throughput dependance from A, p_{in} and p_{out} is included in N through eq.(1).

Note that if $A_{seq} = A = 0$ (i.e. there is only one class of packets) and $p_{out} = p \rightarrow 0$ we get the well-known formula [12]:

$$T(p, RTT) \simeq \frac{n}{RTT}\sqrt{\frac{3}{2bp}}$$

.

Finally, as regards the average length of the in-sequence packet burst (L), from previous remarks it simply follows:

$$L = E[Y_i] - 1 = N - 1 \tag{3}$$

3.2 The Marker Model

We have discussed before about PMA in this paper, and we have seen how the procedure acts marking one packet OUT every A_{seq} IN, where A_{seq} is obtained filtering L_{seq} with an autoregressive unitary-gain filter. Hence, given A and L respectively the average values of A_{seq} and L_{seq}, they are tied by the relation $A = L$[4]. The relation between A_{seq} and L_{seq} has been chosen according to the rationale discussed in section 2. Anyway the relation between A and L can be considered a project choice:

$$A = a(L) \tag{4}$$

A change of the $a()$ law leads to a different marking algorithm, for example pursuing a different target.

As regards the fixed-point approach approximation, we observe that the previous relation looks more suitable as long as the system reaches the state where

[4] A closer look to the algorithm reveals that this is an approximation due to the update $A := A + 1$ after each OUT-packet transmission.

$p_{in} \simeq 0$ and $p_{out} \simeq 1$. In fact, in the case of $p_{in} = 0, p_{out} = 1$ we would have $A_{seq} = L_{seq}$, not simply $A = L$. In [9] and [10] we have shown that the algorithm exhibits optimal performance under *hard differentiation* setting, which leads to $p_{in} \simeq 0$ and $p_{out} \simeq 1$. Hence fixed-point approximation appears justified for PMA.

3.3 The Network Model

In [11] we have proposed a network submodel, extending the approach proposed in [13] for a best-effort scenario to a DiffServ one. A limit of that approach was that TCP sources are intrinsically assumed to achieve full bottleneck utilization, hence the model is able to predict average queue occupation, not link utilization. Besides the model in [11] predicts a range of solutions when $max_{out} < min_{in}$. These problems could be overcome introducing in the model queue variability.

Anyway in this paper the approach is radically different, we consider that the queue can be modelled as a M/M/1/K queueing system. This allows us to evaluate the stationary distribution of the queue for a given offered load T, and then the average values we are interested in, i.e. RTT, p_{in} and p_{out}.

As regards the assumption of Markovian arrivals, it seems to be justified when the TCP connection rate increases [14]. Anyway M/M/1/K models have been widely employed in literature and have shown good performance [15–19]. In particular our framework is similar to those of [17] and [19], which model respectively Token Bucket and Single Rate Three Color Marker, but it differentiates because it assumes state dependent arrivals, rather than uniform ones.

These models take into account the presence of different class of traffic and the effect of AQM mechanism like RIO, but they assume that dropping probability depends only on the instantaneous queue size, disregarding the effect of filtering.

According to [17], the stationary distribution of the queue can be evaluated as:

$$\pi(i) = \pi(0) \left(\frac{T}{C}\right)^i \prod_{j=0}^{i-1} (1 - p(j)), i = 1, 2, ...max_{in}$$

where C is the bottleneck capacity, $\pi(0)$ is given by the normalization equation

$$\pi(0) = \left(1 + \sum_{i=1}^{max_{in}} \left(\frac{T}{C}\right)^i \prod_{j=0}^{i-1} (1 - p(j))\right)^{-1}$$

and

$$p(i) = \frac{T_{in} p_{in}(i) + T_{out} p_{out}(i)}{T_{in} + T_{out}} = \frac{A p_{in}(i) + p_{out}(i)}{A + 1}$$

Note that we assumed $max_{out} < max_{in}$, and that it is useless considering queue values greater than max_{in} because RIO drops all the incoming packets when the instantaneous queue is equal to max_{in}.

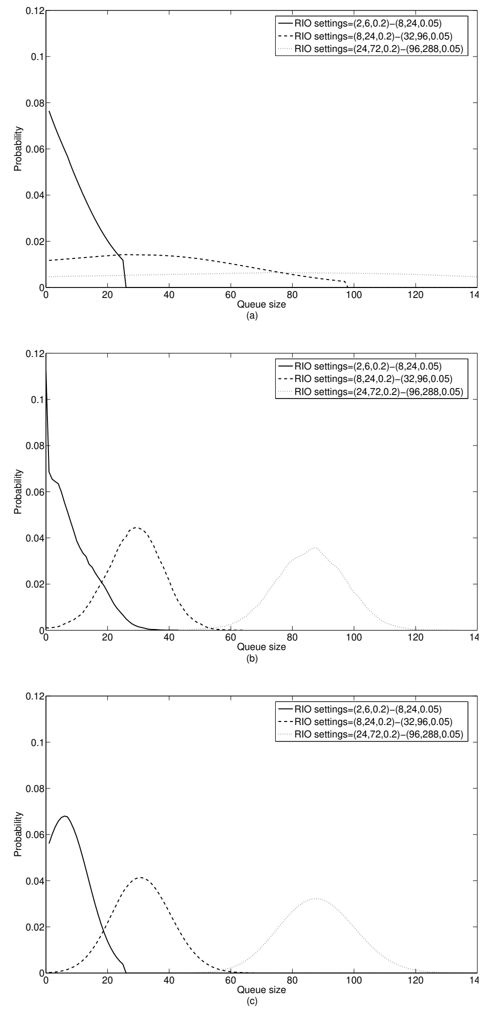

Fig. 4. (a) Queue distribution predicted by the model with uniform arrivals. (b) Queue distribution obtained by simulations. (c) Queue distribution predicted by the model with state dependant arrivals

Once $\pi(i)$ has been obtained RTT, p_{in} and p_{out} can be evaluated as

$$RTT = R_0 + q/C = R_0 + \frac{1}{C} \sum_{i=0}^{max_{in}} i\pi(i) \tag{5}$$

$$p_{in} = \sum_{i=0}^{max_{in}} p_{in}(i)\pi(i) \tag{6}$$

$$p_{out} = \sum_{i=0}^{max_{out}} p_{out}(i)\pi(i) \tag{7}$$

where R_0 is the propagation and trasmission delay.

We have followed such approach, but results are unsatisfactory. The physical explanation appears from figures 4(b) and 4(a), which show the empirical distribution coming from simulations and the queue distribution predicted by the model given the same average load, for three different configurations. The RIO settings in the legend are given in the form $(min_{out}, max_{out}, Pmax_{out}) - (min_{in}, max_{in}, Pmax_{in})$. According to the model the queue should exhibit a spread distribution, with high probability for low queue values (in particular the queue distribution decreases if $T < C$), while the empirical distribution looks like a gaussian one: the dynamic adaptive throughput of the TCP sources, which increase their throughput when RTT decreases and vice versa, appear to be able to create a sort of "constant bias".

In order to capture this behavior, we have modified the model in [17], by introducing arrival dependence from the network status. The input to the submodel is

$$F(N) = T * RTT = \frac{N}{\frac{3b-2}{6} + \sqrt{\frac{2bN}{3} + \left(\frac{2+3b}{6}\right)^2} + 1} \tag{8}$$

and the arrival rate when the there are i packets in the queue is:

$$T(i) = \frac{F}{R_0 + \frac{q}{C}}$$

Now the stationary distribution can be evaluated as:

$$\pi(i) = \pi(0) \prod_{j=0}^{i-1} \frac{T(j)}{C}(1 - p(j)), i = 1, 2, ...max_{in}$$

Figure 4(c) shows the queue distribution evaluated by the new model. The similarity with figure 4(c) is impressive, the only difference is for the first configuration $((2, 6, 0.2) - (8, 24, 0.05))$, as regards low queue occupancy. The peak for $q = 0$ is probably due to timeouts, which are more common with low RIO settings, and make TCP throughput less uniform and hence the markovian arrival assumption less accurate.

3.4 About the solutions of the system

Summarizing, our model has 8 variables (N,A,T,L,RTT,p_{in},p_{out},F) and 8 equations (1), (2), (3), (4), (5) (6), (7) and (8). In this section we afford existence and uniqueness of solutions for this system.

Firstly, let us note that F can be expressed as a function of A by equations (1) and (8), hence, given A, the variables $q(A)$, $p_{in}(A)$ and $p_{out}(A)$ are determined

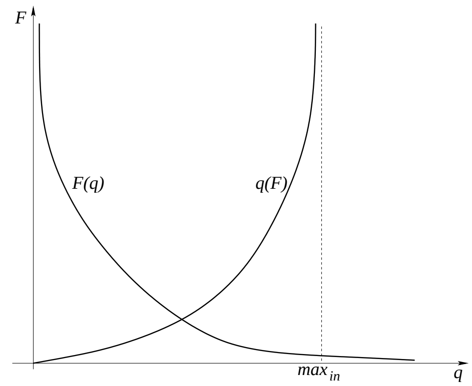

Fig. 5. Existence and uniqueness of the solution

in a univocal manner. Besides it can be proven that $\pi(i+1)/\pi(i)$ increases with A, and therefore q, p_{in} and p_{out} are continuous increasing function of A. Hence they are invertible and one can express p_{in} and p_{out} as (increasing) functions of q.

Besides, the following results hold:

$$q(F=0)=0, \qquad \lim_{F\to+\infty} q(F) = max_{in}$$

As regards A, from equations (1) and (4) ($A = L$), we obtain that A is the solution of the following equation

$$A+1 = \frac{s_{in}^{A+1}-1}{s_{in}-1}\frac{1}{1-s_{in}^{A}s_{out}} \tag{9}$$

Being $p_{in} < p_{out}$ A is a decreasing function of p_{in} and p_{out}, hence a decreasing function of q. Besides, being:

$$\lim_{q\to 0} N(q) = +\infty,$$

it holds

$$\lim_{q\to 0} A(q) = +\infty$$

Let us focus on the expression of F (8). It appears that it is a decreasing function of the queues values, because it is an increasing function of A.

The following results hold:

$$\lim_{q\to 0} F(q) = +\infty, \qquad \lim_{q\to+\infty} F(q) = 0$$

From the previous considerations and hypotheses it follows that the simplified system in F and q admits only one solution, as it is qualitatively shown in figure 5. Being all the function monotone, the original system admits only one solution.

It is possible to set up an iterative procedure to find numerically this solution, and this is just what we did using MATLAB.

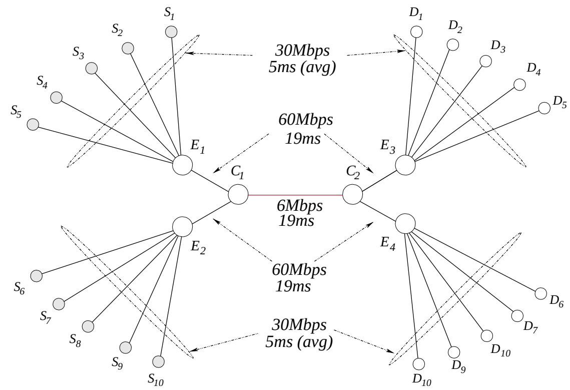

Fig. 6. Network topology

Table 1. Model vs Simulation with 10 flows

RIO	T (pkt/s)		G (pkt/s)		q (pkt)		$Pdrop$ (%)		$Pdrop_{in}$ (%)		$Pdrop_{out}$ (%)		A (pkt)	
	mod	sim	mod	sim	mod	sim	mod	sim	mod	sim	mod	sim	mod	sim
(2,6)(8,24)	486.12	476.19	476.04	469.14	8.79	8.51	2.073	1.457	0.844	0.190	68.605	76.498	54.15	58.12
(4,12)(16,48)	488.25	500.60	480.12	494.84	16.72	15.94	1.666	1.178	0.580	0.085	74.590	85.898	67.17	76.60
(6,18)(24,72)	501.29	503.64	494.25	499.19	24.52	21.71	1.405	0.935	0.408	0.050	79.988	88.121	78.83	98.06
(8,24)(32,96)	501.10	504.38	495.03	499.89	32.04	29.29	1.211	0.937	0.300	0.039	83.490	89.652	90.33	98.23
(12,36)(48,144)	500.39	503.76	495.68	499.98	46.63	44.12	0.943	0.777	0.184	0.015	87.848	90.962	114.60	118.21
(16,48)(64,192)	499.93	503.32	495.47	499.99	60.89	58.77	0.892	0.687	0.120	0.012	90.450	88.296	115.94	128.66
(24,72)(96,288)	499.44	502.39	496.84	500.00	89.00	86.31	0.521	0.517	0.059	0.014	93.627	81.998	201.19	160.50
mean error (%)	-0.49		-0.84		6.14		30.76		659.78		-3.77		-4.91	
max error (%)	-2.47		-2.97		12.92		50.29		1128.05		14.18		25.35	

4 Model validation

To validate our model we considered the network topology showed in Fig.6, consisting of a single bottleneck link with capacity equal to 6Mbps. Considering both the transmission and the propagation delay of packets and acks in the network, the average Round Trip Time is $R_0 \cong 138$ms. The IP packet size is chosen to be 1500 Bytes, for a bottleneck link capacity of $c = 500$packets/s.

As regards RIO configurations we considered *non overlapping* the ones in which $max_{out} < min_{in}$, more precisely we choose $max_{out} = 3min_{out}$, $max_{in} = 3min_{in}$ and $min_{in} = 4max_{out}$. In previous performance evaluation this kind of settings showed better results in comparison with a *overlapping* RIO configuration in which, $max_{out} \geq min_{in}$. We tested seven different configurations, varying min_{out} from 2 up to 24, and for each configuration we gathered statistics from 10 trials of 1000 seconds each. We ran our simulations using ns v2.1b9a, with the Reno version of TCP.

Table 1 compares model predictions with simulation results when the number of flows is equal to $n = 10$, as regards throughput (T), goodput[5](G), queue occupancy (q), the dropping probability for the generic packet, for IN packets and for OUT packets (respectively $Pdrop$,$Pdrop_{in}$, $Pdrop_{out}$), and the average

[5] The goodput is estimated as $G = T(1 - Pdrop)$.

length of IN packets bursts. The average mean error over the different settings and the maximum error are shown in the last two rows. The model appears to be able to predict with significant accuracy throughput, goodput and queue occupancy, which are the most relevant performance indexes when we consider TCP long lived performance flows. On the contrary dropping probability estimates are very inaccurate, in particular as regards $Pdrop_{in}$. We think the reason is that the model neglects the effect of filtering on dropping probability calculation from RIO routers. In fact some preliminary results which take into account filtering seem to suggest that filtering: i) can be neglected in order to evaluate the dynamic of the instantaneous queue, ii) it is significant for the evaluation of the dropping probabilities. In particular probabilities estimates look better. At the moment we have introduced the effect of filtering by considering a two dimensional Markov chain where the status is the pair of instantaneous queue and filtered queue (whose values have been quantized). This approach is particularly heavy from the computational point of view, for this reason, at the moment, we have not adopted it.

We evaluated also the model with the same network topology with a different number of flows ($n = 6$,$n = 20$). The differences between model predictions and simulation results are similar to those observed for $n = 10$ flows. The relative errors for these two scenarios are shown in table 2.

Table 2. Model vs Simulation with 6 and 20 flows

$n = 6$	T (pkt/s)	G (pkt/s)	q (pkt)	$Pdrop$ (%)	$Pdrop_{in}$ (%)	$Pdrop_{out}$ (%)	A (pkt)
mean error (%)	-0.87	-0.90	-2.28	3.71	760.33	-2.87	37.33
max error (%)	-1.49	-1.53	-25.96	9.47	1227.52	28.15	61.31
$n = 20$	T (pkt/s)	G (pkt/s)	q (pkt)	$Pdrop$ (%)	$Pdrop_{in}$ (%)	$Pdrop_{out}$ (%)	A (pkt)
mean error (%)	1.75	-0.24	2.76	110.99	607.65	-1.37	-45.67
max error (%)	5.93	2.15	19.37	157.58	798.09	-10.19	-54.12

5 Conclusions and further research issues

In this paper we have presented an analytical model for our adaptive packet marking scheme proposed in previous works. From preliminary simulative results, model predictions about throughput and average queue occupancy appear to be quite accurate. We are going to extend simulative evaluation and to employ such model to study possible variants of the marking algorithm and to establish optimal RIO settings.

Besides our network sub-model exhibits some novelty and seems to be more suited than traditional M/M/1/K proposals to capture the behavior of long lived TCP flows. We are going to study it deeply and to evaluate it in a simpler best effort scenario. We want to evaluate the effect of filtering, which is usually neglected in M/M/1/K models, but it appears to have a deep impact on the performance.

References

1. D. D.Clark and W. Fang, “Explicit Allocation of Best Effort Packet Delivery Service”, IEEE Transactions on Networking, Vol. 6, No. 4, pp. 362-373, Aug. 1998.
2. N. Seddigh, B. Nandy, P. Piedu, “Bandwith Assurance Issues for TCP flows in a Differentiated Services Network”, IEEE Globecom, Rio de Janeiro, December 1999
3. S. Sahu, D. Towsley, J. Kurose, “Quantitative Study of Differentiated Services for the Internet”,IEEE Globecom, Rio de Janeiro, December 1999
4. S. Sahu, P. Nain, D. Towsley, C. Diot, V. Firoiu, “On Achievable Service Differentiation with Token Bucket Marking for TCP”, Proc. ACM SIGMETRICS’00, Santa Clara, CA, June 2000
5. W. Feng, D. Kandlur, D. Saha, K. Shin, “Adaptive Packet Marking for Mantaining End-to-End Throughput in a Differentiated Services Internet”, IEEE/ACM Transactions on Networking, Vol. 7, NO:5, pp. 685-697, April 1999
6. F. Azeem, A. Rao, S. Kalyanaraman “A TCP-Friendly Traffic Marker for IP Differentiated Services” IwQoS’2000, Pittsburg, PA, June 2000.
7. G. Lo Monaco, F. Azeem, S. Kalyanaraman, Y.Xia, “TCP-Friendly Marking for Scalable Best-Effort Services on the Internet”, Computer Communication Review (CCR), Volume 31, Number 5, October 2001.
8. M. Mellia, I. Stoica, H. Zhang, “Packet Marking for Web traffic in Networks with RIO Routers”, Globecom 2001, San Antonio, Texas, November 25-29, 2001
9. G. Neglia, G. Bianchi, F. Saitta, D. Lombardo, “Adaptive Low Priority Packet Marking for better TCP performance”, Net-con 2002, Paris, October 2002
10. G. Neglia, G. Bianchi, M. Sottile, “Performance Evaluation of a new Adaptive Packet Marking Scheme for TCP over DiffServ Networks” Globecom 2003, San Francisco, USA, December 2003
11. G. Neglia, G. Bianchi, V. Falletta, “An Analytical Model of a new Packet Marking Algorithm for TCP flows: preliminary insights”, ISCCSP 2004, Hammamet, Tunisia, 21-24 March 2004
12. J. Padhye, V. Firoiu, D. Towsley, J. Kurose. “Modeling TCP Throughput: A Simple Model and its Empirical Validation”, ACM Sigcomm ’98, Vancouver, Canada, September 2-4, 1998
13. V. Firoiu, M. Borden, “A Study of Active Queue Management for Congestion Control”, IEEE Infocom 2000, Tel-Aviv, Israel, March 26-30, 2000
14. J. Cao, W. S. Cleveland, D. Lin, and D. X. Sun, “Internet traffic tends toward poisson and independent as the load increases,” in Nonlinear Estimation and Classification, Springer, 2002
15. M. May, J-C. Bolot, A. Jean-Marie, C. Diot, “Simple performance models of tagging schemes for service differentiation in the Internet”, Proc. Infocom ’99, New York, NY, March 1999
16. R.J. Gibbens, S.K. Sargood, C. Van Eijl, F.P. Kelly, H. Azmoodeh, R.N. Macfadyen, N.W. Macfadyen, “Fixed-Point Models for the End-to-End Performance Analysis of IP Networks”, 13th ITC Specialist Seminar: IP Traffic Measurement, Modeling and Management, Monterey, California, September 18-20, 2000
17. N. M. Malouch, Z. Liu, “Performance Analysis of TCP with RIO Routers”, IEEE GlobeCom 2002, Taipei, Taiwan, November 17-21, 2002
18. M. Garetto, R. Lo Cigno, M. Meo, M. Ajmone Marsan, “Closed Queueing Network Models of Interacting Long-Lived TCP Flows”, IEEE/ACM Transactions on Networking, Vol. 12, NO. 2, p. 300-311, April 2004
19. R. Stankiewicz, A. Jajszczyk, “Modeling of TCP Behavior in a DiffServ Network Supporting Assured Forwarding PHB”, ICC 2004, Paris, France, 20-24 June 2004