A Framework to Determine Optimal Loss Rate of RED Queue for Next Generation Internet Routers

Bing Zheng
New Focus Inc.
5215 Hellyer Ave.,
San Jose, CA 95138.
zhengbin@ieee.org

Mohammed Atiquzzaman
School of Computer Science
University of Oklahoma
Norman, OK 73019-6151
atiq@ou.edu

***Abstract*—Random Early Detection (RED) is expected to eliminate global synchronization by random active packet drop. Its packet drop probability is decided by the maximum packet drop probability in its drop function, the buffer thresholds, and the average queue length. It has been observed that for a large number of connections, a small value of maximum drop probability will not eliminate global synchronization. Further more, since RED uses its four parameter to regulate its performance, it is necessary to relate its maximum drop probability with its other parameters. The *objective* of this paper is to develop a model of maximum drop probability of RED, based on TCP channel model and traffic characteristics. The value of maximum drop probability obtained by our model will make RED queue achieve its targeted goals described in IETF documents.**

I. Introduction

The bandwidth of a TCP connection depends on the round trip time (RTT) and the packet drop probability of the connection [1], [2]. RED gateways provide bandwidth control since they use packet drop to imply congestion in the network. The packet drop probability of a RED gateway will dynamically decide the bandwidth with associated TCP connections. However, since a RED gateway regulates its performance by four parameters and one control variable, the packet drop probability depends not only on the congestion scenario, but also on the configuration parameters. As one of the key parameters in RED, the maximum packet drop probability, p_{max}, is therefore related to the traffic pattern and other configuration parameters. However, the value of p_{max} suggested in [3], [4] is independent of the traffic pattern and the RED parameters. As a result RED does not work satisfactorily in the case of RED gateways supporting large number of connections. The *objective* of this paper *is to develop a model for p_{max}, based on the TCP traffic characteristics*.

The effects of p_{max} on queue management are two folds. The first is the value of p_{max} itself, and the second is the relationship between the value of p_{max} and the other parameters (Min_{th}, Max_{th}, and w) [3] (see Fig. 1). In the first case, it has been shown that if p_{max} is too small, RED is insufficient to notify senders, and tail drop will dominate the packet drop at the RED gateway [5]; too large a value of p_{max} will lower link bandwidth. The second case is more complicated than first case, since RED queue regulates its performance by its four parameters. To illustrate the problem resulting from the second case, recall that RED uses average queue length as a control variable to calculate packet drop probability as given

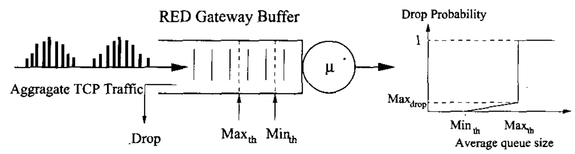

Fig. 1. Illustration of RED gateway queue.

by $p_{RED} = p_{max}\frac{avg-Min_{th}}{Max_{th}-Min_{th}}$; Therefore, the actual packet drop probability p_{RED} is decided not only by the value of p_{max} but also depends on the value of $Max_{th} - Min_{th}$. In other words, the actual packet drop probability is decided by the ratio of p_{max} and $Max_{th} - Min_{th}$. This means that p_{max} needs to be related to the link feature and the thresholds of a RED queue.

The contribution of this paper is the development of a model for p_{max} which is related to the TCP connection parameters and RED configuration parameters. The use of p_{max} suggested by our model results all packet drops being due to active drops (rather than the undesirable passive (tail) drops) while still maintaining a high link utilization and no global synchronization. The previous value of p_{max} suggested in the literature can eliminate passive drops, but *results in low utilization of the link bandwidth for TCP traffic*. Our model can be used by network engineers to determine an optimum value of p_{max} based on traffic characteristics and values of RED parameters.

In Section II, we state the assumptions in developing our proposed model for p_{max} . In section III, we develop the model based on TCP channel model and traffic characteristics. Simulation results are shown in Section IV, while conclusions are given in Section V.

II. Modeling Assumptions and Notations

To facilitate further discussion, we present the assumptions and notation used in our model.

A. Assumptions

We make the following assumptions, which will be used to develop the model for determining p_{max} of a RED queue in Section III. *Note that some of the assumptions have been used in previous work* as mentioned below.

- The RED gateway queue is initially empty (also assumed in [3]);

0-7803-7710-9/03/$17.00 © 2003 IEEE.

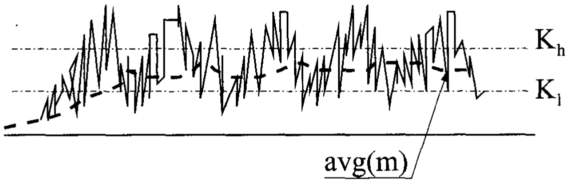

Fig. 2. Long term performance of average queue length versus instantaneous queue.

- The average queue length is initially zero (also assumed in [3]);
- When RED performs well, the average queue length will vary within a small range;
- In the long term, the active packet drop always works;
- Round trip time (RTT), τ, for a connection is constant (used in [1]);
- A TCP source's congestion window at time t is determined by the packet drops at time $t-\tau$ (as in [1], [6]).
- Long Term Congestion, which varies slowly, is described by a slow function $g(i)$, where i corresponds to the i-th calculation of the average queue length;
- Short Term Congestion, which varies fast, is described by a fast function $f(i)$ corresponding to the ith calculation of the average queue length.

$$f(i) = q_0\left(\frac{1-(-1)^i}{2} + b\frac{1-(-1)^{i-1}}{2}\right) \quad (1)$$

- The instantaneous queue length, $q(i)$, is described by the modulation of the fast function ($f(i)$) by a slow function ($g(i)$), i.e., $q(i) = g(i)f(i)$.

B. Notations

We define the following variables, which are used in our model in Section III.

- w: Weight parameter for calculation of average queue length at a RED gateway;
- $q(m)$: Instantaneous queue size of the RED gateway during the m-th calculation of the average queue length. From our assumptions in Section II-A, $q(0) = 0$;
- μ: bottleneck link rate;
- $avg(m)$: Average queue length of the RED gateway at the m-th calculation (see Fig. 2). It is defined as:

$$avg(m) = (1-w) * avg(m-1) + w * q(m) \quad (2)$$

  where $avg(0) = 0$ from our assumptions in Section II-A;
- τ: RTT in terms of calculation interval of average queue length;
- p_{max}: Maximum packet drop probability for RED queue;
- K_l: Minimum buffer threshold for RED gateway to perform active packet drop;
- K_h: Maximum buffer threshold for RED gateway to perform packet drop with probability of one;
- $W_i(m)$: TCP congestion window size for the i-th connection at time m;
- N: Total number of connections;
- $a(m)$: Normalized instantaneous queue defined by $\frac{q(m)}{K_l}$;
- $b(m)$: Normalized average queue length defined by $\frac{avg(m)}{K_l}$;

In the next section, we develop the model of p_{max} based on the above assumptions.

III. Modeling p_{max}: Lower and Upper Bounds

According to the definition of packet drop probability in RED queue, for average queue length $avg(m)$, the corresponding packet drop probability $p(m)$ is:

$$p(m) = \frac{avg(m) - K_l}{K_h - K_l} p_{max} \quad (3)$$

By defining $\alpha = \frac{p_{max}}{K_h - K_l}$, Equation (3) can be rewritten as:

$$p(m) = \alpha K_l (b(m) - 1) \quad (4)$$

On other hand, the difference equation for the instantaneous queue $q(m)$ is:

$$q(m) - q(m-1) = \sum_{i=1}^{N} \frac{W_i(m)}{\tau_i} - \mu \quad (5)$$

where, the left hand side is the net change of instantaneous queue size, the right hand side is the difference between the incoming and outgoing data. To simplify the discussion, we consider N *iid* TCP connections with the same RTT. In this case, Equation (5) becomes:

$$q(m) - q(m-1) = \frac{NW(m)}{\tau} - \mu \quad (6)$$

Expressing Equation (6) in normalized form, we have:

$$a(m) - a(m-1) = \frac{NW(m)}{\tau K_l} - \frac{\mu}{K_l} \quad (7)$$

It has been proved in [1] that $W(m)$ can be expressed as:

$$W(m) = \frac{C}{\sqrt{p(m-\tau)}} \quad (8)$$

where C is a constant. By substituting Equation (4) into (8), we have:

$$W(m) = \frac{C}{\sqrt{\alpha K_l}\sqrt{b(m-\tau)}\sqrt{1-\frac{1}{b(m-\tau)}}} \quad (9)$$

where $b(m)$ is the normalized average queue length. For active queue management to work, it must have $b(m) \geq 1$. Therefore, $\frac{1}{b(m-\tau)} \leq 1$ always holds. By using $\frac{1}{\sqrt{1-x}} \simeq (1+\frac{x}{2})$, we have:

$$W(m) \simeq \frac{C}{\sqrt{\alpha K_l}\sqrt{b(m-\tau)}}(1 + \frac{1}{2b(m-\tau)} \quad (10)$$

From the definition, the normalized average queue is expressed as:

$$b(m) = (1-w)b(m-1) + wa(m) \quad (11)$$

We have:

$$a(m) = \frac{b(m) - (1-w)b(m-1)}{w} \quad (12)$$

$$a(m) - a(m-1) = \frac{b(m)-(1+(1-w))b(m-1)+(1-w)b(m-2)}{w} \quad (13)$$

Substituting Equation (13) into (7):

$$\frac{b(m) - (1+(1-w))b(m-1) + (1-w)b(m-2)}{w} = \frac{NW(m)}{\tau K_l} - \frac{\mu}{K_l} \quad (14)$$

By substituting $W(m)$ into Equation (14):

$$\frac{b(m) - (1+(1-w))b(m-1) + (1-w)b(m-2)}{w} + \frac{\mu}{K_l} = \frac{N}{\tau K_l} \frac{C}{\sqrt{\alpha K_l}\sqrt{b(m-\tau)}} \left(1 + \frac{1}{2b(m-\tau)}\right) \quad (15)$$

Finally, from Equation (15), we have:

$$p_{max} = \frac{(CN)^2}{(\mu\tau)^2} \frac{(K_h - K_l)(1 + \frac{K_l}{2avg(m)})^2}{avg(m-\tau)} \times \frac{1}{(1 + \frac{avg(m)-(1+(1-w))avg(m-1)+(1-w)avg(m-2)}{w\mu})^2} \quad (16)$$

To simplify Equation (16), we need to evaluate the term $(1 + \frac{avg(m)-(1+(1-w))avg(m-1)+(1-w)avg(m-2)}{w\mu})$. From the definition of avg and our assumptions, we have

$$avg(m) - avg(m-1) = (1-w)^{m-1} w g(m) f(m) \quad (17)$$

$$avg(m-1) - avg(m-2) = (1-w)^{m-2} w g(m-1) f(m-1) \quad (18)$$

Therefore, we have:

$$avg(m) - (1+(1-w))avg(m-1) + (1-w) \times avg(m-2) = (1-w)^{m-1} w g(\overline{m})(f(m) - f(m-1) \quad (19)$$

From our assumption, we have:

$$f(m) - f(m-1) = q_0 \left(\frac{-(-1)^m + (-1)^{m-1}}{2} + b\frac{-(-1)^{m-1} + (-1)^{m-2}}{2} \right) \quad (20)$$

If m is even, $(-1)^m = 1$. If m is odd, $(-1)^m = -1$. Therefore, we have:

$$f(m) - f(m-1) = +q_0(b-1) \text{ if m is even} \quad (21)$$

$$= -q_0(b-1) \text{ if m is odd} \quad (22)$$

Therefore, we have:

$$(1 + \frac{avg(m) - (1+(1-w))avg(m-1) + (1-w)avg(m-2)}{w\mu}) \doteq \pm \frac{(1-w)^{m-1} g(\overline{m}) q_0 (b-1)}{\mu} \quad (23)$$

Since w is very small in a RED queue, for long term performance, m is very large, resulting in a very small value of $(1-w)^{m-1}$. Moreover, comparing with the link rate μ, $(1-w)^{m-1} g(\overline{m}) q_0 (b-1)$ is so small that the term $\frac{(1-w)^{m-1} g(\overline{m}) q_0 (b-1)}{\mu}$ can be approximated by 0. Then the expression for p_{max} can be expressed as:

$$p_{max} = \frac{(CN)^2}{(\mu\tau)^2} \frac{(K_h - K_l)(1 + \frac{K_l}{2avg(m)})^2}{avg(m-\tau)} \quad (24)$$

$avg(m-\tau)$ is called the *target average queue length* that RED tries to achieve in the long term. To have the RED gateway queue work under active queue management, the relationship $avg(m) \in [K_l, K_h]$ must be satisfied. From the drop function of the RED gateway queue, if the long term $avg(m) < K_l$, there will be no active packet drop. In this case, active queue management will not work; packet drops are due queue overflow. On other hand, if long term $avg(m) > K_h$, each arriving packet will be dropped, giving rise to the problem of *global synchronization.* From this discussion, we conclude that the target long term $avg(m)$ will control p_{max} between the *Upper Bound* and *Lower Bound* values of p^U_{max} and p^L_{max} respectively as given below.

$$p^U_{max} \leq \frac{(NC)^2}{(\mu\tau)^2} \frac{(K_h - K_l)(1 + \frac{K_l}{2K_l})^2}{K_l} \quad (25)$$

$$p^L_{max} \geq \frac{(NC)^2}{(\mu\tau)^2} \frac{(K_h - K_l)(1 + \frac{K_l}{2K_h})^2}{K_h} \quad (26)$$

IV. Performance Evaluation

To test the model developed in above section, we carried out simulations using the OPNET 5.1D network simulation tool. Before describing our results, we describe the network topology and simulation configuration.

A. Simulation Configurations

The network topology is shown in Figure 3. Three TCP sources send ftp traffic to a client via a RED gateway. To ensure a fair comparison with the value of p_{max} in the original RED, the values of the configuration parameters were the same as those suggested in [4].

- *Server0 to RED Gateway link:* Propagation delay 1 ms, link rate 100 Mbps.
- *Server1 to RED Gateway link:* Propagation delay 5 ms, link rate 100 Mbps.
- *Server2 to RED Gateway link:* propagation delay 3 ms, link rate 100 Mbps.
- *Client to RED Gateway link:* Propagation delay $5ms$, bottleneck link rate $10Mbps$. To induce congestion at the RED queue, the bottleneck link rate has been chosen to be 30 times smaller than the sum of link rates feeding the bottleneck link.
- *Gateway processing speed:* 1 ms per packet.
- *Gateway Queue Size:* 200 packets.
- w=0.07, $K^0_l = 6$, $K^0_h = 20$ *and* $K^1_l = 6$, $K^1_h = 140$
- $p^0_{max} = 0.1$ [4] and $p^1_{max} = p^0_{max} \frac{K^1_h - K^1_l}{K^0_h - K^0_l}$.

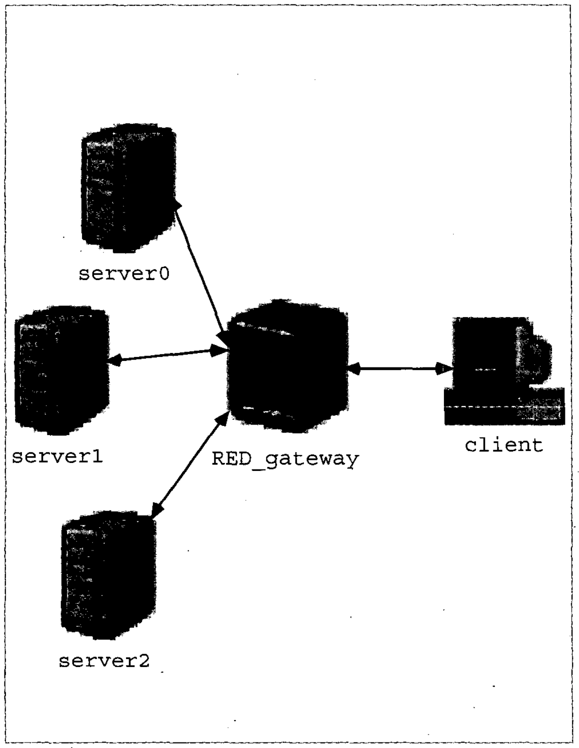

Fig. 3. Network configuration for simulation.

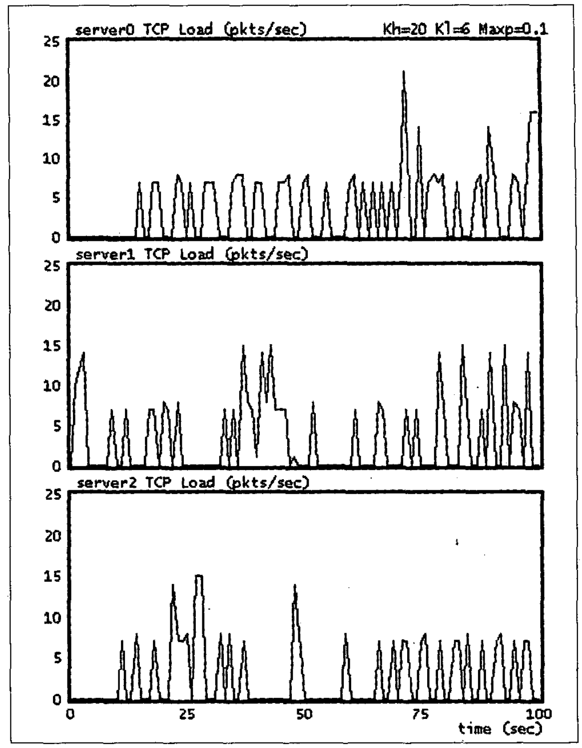

Fig. 4. TCP load for $p_{max}^0 = 0.1$ and $K_h^0 - K_l^0 = 14$.

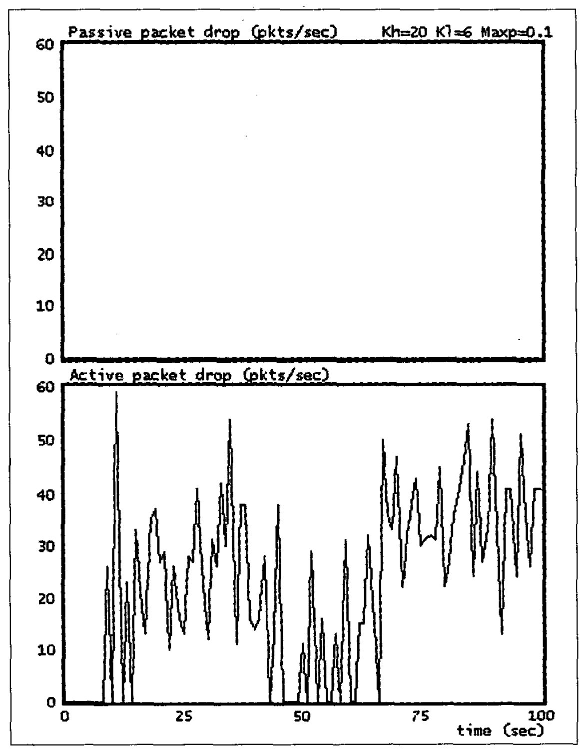

Fig. 5. Packet drop for $p_{max}^0 = 0.1$ and $K_h^0 - K_l^0 = 14$.

B. Results and Discussion

Figures 4 and 5 show the TCP traffic load at the three servers, and active and passive packet drops (tail drops) for $p_{max}^0 = 0.1$ and $K_h^0 - K_l^0 = 14$. It can be seen that all the packets are due to active drops with zero passive drops, i.e. the RED queue works in active drop. *The TCP senders did not suffer from global synchronization. The bottleneck link has a reasonable utilization as shown in Figure 6.*

Figures 7 and 8 show the simulation results for $p_{max}^0 = 0.1$ and changing $K_h^0 - K_l^0 = 134$ and 14, respectively. In the above analysis, the ratio $\frac{p_{max}^0}{K_h^1 - K_l^1}$ is so small (less than 0.0008) that the active packet drop is insufficient to eliminate passive (tail) drops. Therefore, all packet drops are due to tail drops, resulting in global synchronization of TCP senders. TCP senders stop sending after time 25 seconds. Therefore, *the bottleneck link bandwidth is wasted* after time 25 seconds, as shown in Figure 6.

Figures 9 and 10 show the simulation results for $p_{max}^1 = p_{max}^0 \frac{K_h^1 - K_l^1}{K_h^0 - K_l^0}$. As indicated by our theoretical model, the ratio $\frac{p_{max}^1}{K_h^1 - K_l^1}$ is restored to the proper value so that the *active packet drops are sufficient to eliminate passive drops. Therefore, all packet drops are active drops, resulting in elimination of global synchronization. The congestion is relieved, resulting in a higher bottleneck link utilization as shown in Figure 6.*

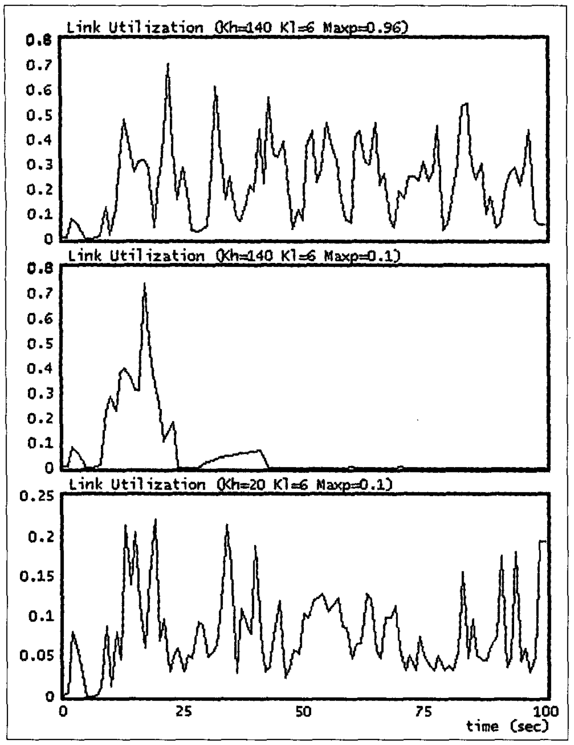

Fig. 6. Link utilization for $p^1_{max} = 0.96$ and $K^1_h - K^1_l = 134$, $p^0_{max} = 0.1$ and $K^1_h - K^1_l = 134$, $p^0_{max} = 0.1$ and $K^0_h - K^0_l = 14$.

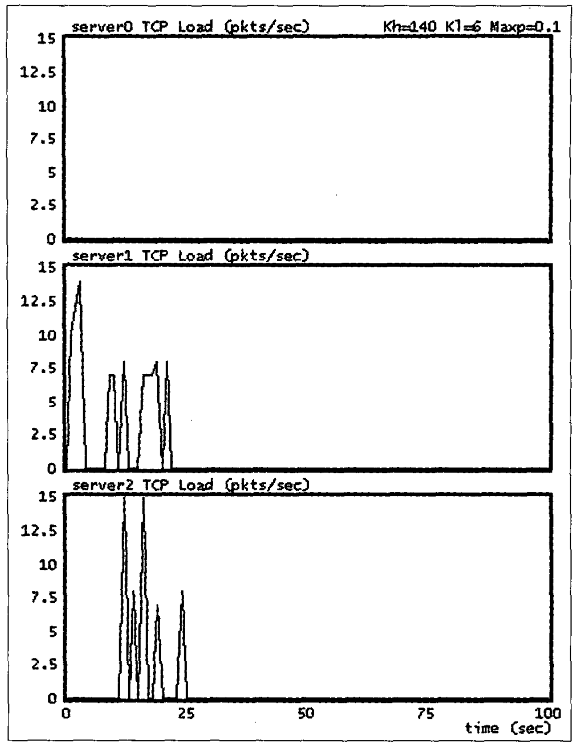

Fig. 7. TCP load for $p^0_{max} = 0.1$ and $K^1_h - K^1_l = 134$.

V. Conclusion

In this paper,a framework to determine p_{max} of RED gateways have been proposed and developed. Simulation results have shown that the model properly relates p_{max} with the buffer threshold and RED parameters. For a given TCP link, a large buffer threshold results in a large p_{max}. A fixed value of p_{max} is not universally suitable for all configurations. Inappropriate combination of p_{max} with other configuration parameters will prevent RED from achieving its desired goals. Our model can be used by network engineers to determine an optimum value of p_{max} based on traffic characteristics and values of RED parameters.

References

[1] Jitendra Padhye, Victor Firoiu, Donald F. Towsley, and James F. Kurose, "Modeling TCP Reno performance: A simple model and its empirical validation," *IEEE/ACM Transaction on Networkings*, vol. 8, no. 2, pp. 133–145, April 2000.

[2] Alhussein A. Abouzeid, Sumit Roy, and Murat Azizoglu, "Stochastic modeling of TCP over lossy links," *IEEE INFOCOM2000*, Tel-Aviv, Israel, March 26-30 2000.

[3] Sally Floyd and Van Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transanctions on Networking*, vol. 1, no. 4, pp. 397–413, August, 1993.

[4] S.Floyd, "RED: Discussions of setting parameters." www.aciri.org/floyd/REDparameters.txt, November, 1997.

[5] Wu chang Feng, Dilip D Kandlur, Debanjan Saha, and G Kang, "Self-configuring RED gateway," *IEEE INFOCOM99*, New York, pp. 1320–1328, March 1999.

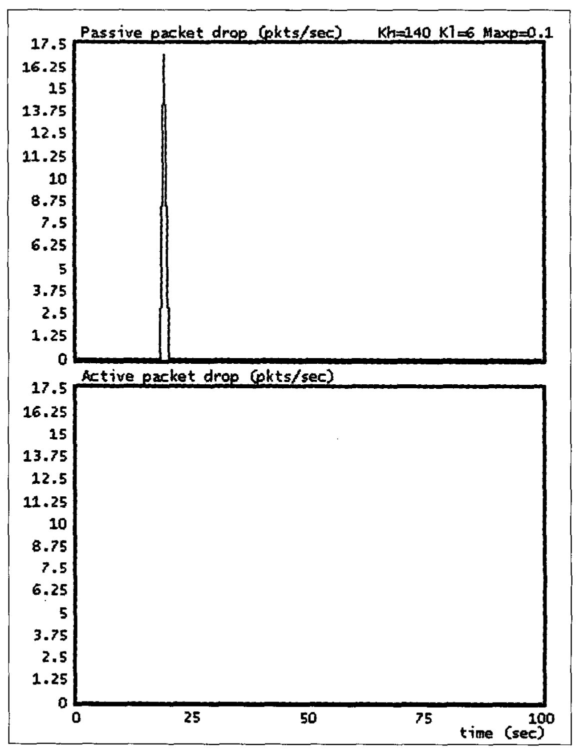

Fig. 8. Packet drop for $p^0_{max} = 0.1$ and $K^1_h - K^1_l = 134$.

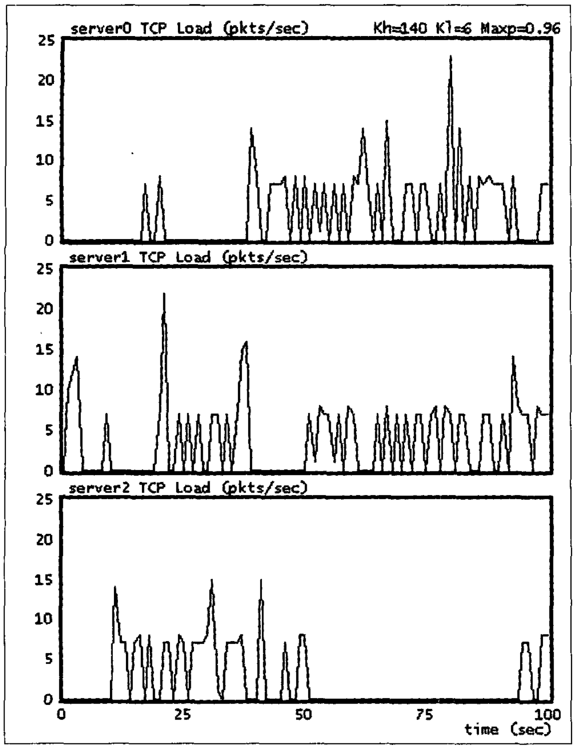

Fig. 9. TCP load for $p_{max}^1 = 0.96$ and $K_h^1 - K_l^1 = 134$.

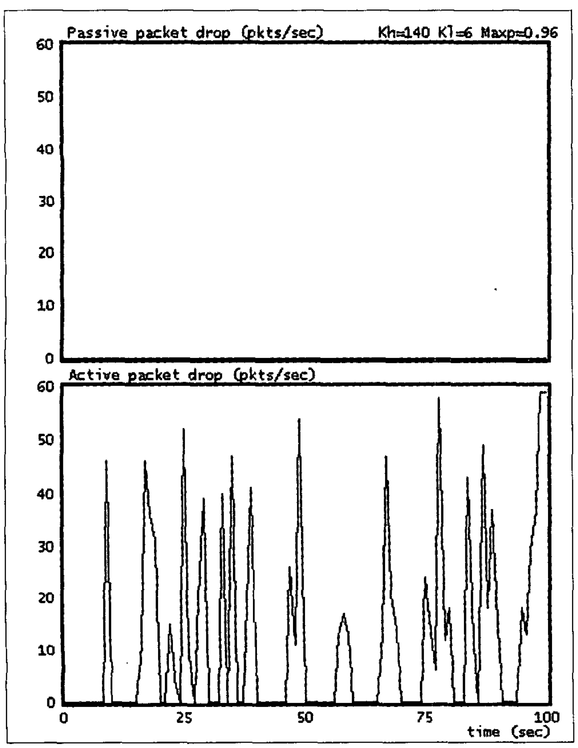

Fig. 10. Packet drop for $p_{max}^1 = 0.96$ and $K_h^1 - K_l^1 = 134$.

[6] Vishal Misra, Wei-Bo Gong, and Don Towsley, "Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED," *ACM SIGCOMM2000*, Stockholm, Sweden, 2000.