

A Quantitative Analysis and Performance Study of Fast Congestion Notification (FN) Mechanism

Mohammed M. Kadhum¹ and Suhaidi Hassan²

InterNetWorks Research Group
 UUM College of Arts and Sciences
 Universiti Utara Malaysia
 06010 UUM Sintok, MALAYSIA

kadhum@uum.edu.my¹ | suhaidi@uum.edu.my²

Abstract— Congestion in computer network happens when the number of transmission requests exceeds the transmission capacity at a certain network point (called a bottle-neck resource) at a specific time. Congestion usually causes buffers overflow and packets loss. The purpose of congestion management is to maintain a balance between the transmission requests and the transmission capacity so that the bottle-neck resources operate on an optimal level, and the sources are offered service in a way that assures fairness. Fast Congestion Notification (FN) is one of the proactive queue management mechanisms that limits the queuing delay and achieves the maximum link utilization possible with minimum packet drops. In this paper we present a detailed performance comparison of the Linear FN algorithm to RED based on the results obtained through simulations. The paper shows how FN can be tuned for different window size (Ws) and periods of time constant (T) to achieve higher link utilization; reduce the queuing delay, and lower packet drop ratio.

Keywords— Internet Congestion; Active Queue Management (AQM); Explicit Congestion Notification (ECN); Gateway buffers; Random Early Detection(RED)

I. INTRODUCTION

A queue management mechanism is the algorithm that manages the length of the packet queues by dropping packets when necessary or appropriate [1]. The efficiency of queue management mechanisms depends on how well their control decisions, on packet admission to the queue and congestion control directing, will help in satisfying their objectives regarding congestion avoidance and control. These decisions are implemented and compelled during the design of the packet mark probability and the mark activation functions.

Packet admission and congestion control directing decisions are dependent on each other. Based on the drop activation characteristic, the queue management policies can be classified into two categories. The first category is reactive (passive) queue management policies, which do not employ any preventive packet drop before the gateway buffer is flooded. The second category is proactive (active) queue management policies(AQM) which employ preventive packet drop before the gateway buffer gets full [2]. Drop-Tail, which is one of reactive queue management policies, is currently widely developed in the Internet routers. It introduces several problems, such as global synchronization, on the Internet. Active queue management policies, such as Random Early

Detection (RED), are expected to eliminate global synchronization and improve Quality of Service (QoS) of the networks. The promised advantages of AQM are increase in throughput, reduce the delay, and avoid lock-out. AQM provides preventive measures to manage the router queue to eliminate the problems associated with passive queue management. Preventive packet drop provides implicit feedback method to notify the traffic senders of the congestion onset [3]. As a reaction, senders reduce their transmission rate to moderate the congestion level. Arriving packets from the senders are dropped or marked randomly, which prevents senders from backing off at the same time and thereby eliminate global synchronization [3].

II. RANDOM EARLY DETECTION (RED)

RED [4] is the default AQM mechanism that is recommended by IETF for the Internet routers. Figures 1, 2, and 3, show the algorithm, gateway buffer, and packet drop function of RED, respectively. A router implementing RED detects the congestion early by computing the average buffer length (avg) and sets the two queue thresholds (Max_{th} and Min_{th}) for packet drop/mark. The average buffer length at time t , is defined as

$$avg(t) = (1-w) avg(t-1) + wq(t) \quad (1)$$

and it is used as a control variable to perform active packet drop/mark. The $avg(t)$ is the new value of the average buffer length at time t , $q(t)$ is instantaneous buffer length at time t , and w , which is normally less than one, is a weight parameter in calculating avg .

```

For each packet arrival
    Calculate the average queue size avg
    If Min_Threshold ≤ avg < Max_Threshold
        Calculate probability p
        With probability p:
            Mark the arriving packet
    Else if Max_Threshold < avg
        Mark the arriving packet
    
```

Fig. 1 RED algorithm

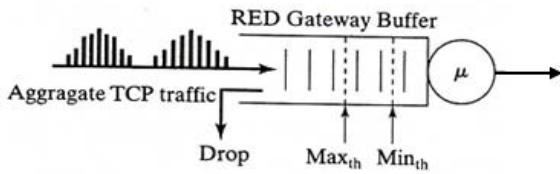


Fig. 2 RED Gateway Buffer

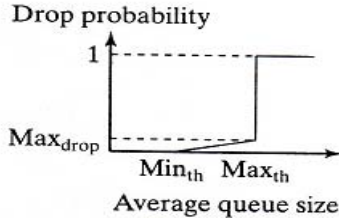


Fig. [1]3 RED packet drop function

A. Packet Drop/Mark Probability

Packet drop/mark probability function determines the probability that the packet is dropped or marked when the drop/mark activation function imposes drop/mark procedure initiation and the drop/mark position function selects the specific packet to be dropped/marked. For example in drop-tail, the chosen packet which is the packet at the tail of the queue is dropped with probability one.

In RED, the probability of dropping/marketing packet, P , is calculated by

$$P = Max_{drop} ((avg - Min_{th}) / (Max_{th} - Min_{th})) \quad (2)$$

The RED algorithm includes two computational parts: computation of the average buffer length and calculation of the drop/mark probability.

The RED algorithm involves four parameters to regulate its performance. Min_{th} and Max_{th} are the queue thresholds to perform packet drop/mark, Max_{drop} is the packet drop/mark probability at Max_{th} , and w is the weight parameter to calculate the average buffer length from the instantaneous buffer length. The average buffer length follows the instantaneous buffer length. However, because w is much less than one, avg changes much slower than q . Therefore, avg follows the long-term changes of q , reflecting persistent congestion in networks. By making the packet drop/mark probability a function of the level of congestion, RED gateway has a low packet-drop/mark probability during low congestion, while the drop/mark probability increases when the congestion level increases [3].

The packet drop/mark probability of RED is small in the interval Min_{th} and Max_{th} . Furthermore, the packets to be dropped (or marked if ECN-capable) are chosen randomly from the arriving packets from different hosts. As a result, packets coming from different hosts are not dropped or marked simultaneously. RED gateways, therefore, avoid global synchronization by randomly dropping/marketing packets.

The performance of RED significantly depends on the values of its four parameters [5], Max_{drop} , Min_{th} , Max_{th} , and w .

III. FAST CONGESTION NOTIFICATION (FN)

The FN [6] queue management algorithm randomly marks (if ECN) / drops (if non-ECN) the arriving packets before the buffer overflows, to effectively control the:

- Instantaneous queue length below the optimal queue length to reduce the queuing delay and avoid the buffer overflows.
- Average traffic arrival rate of the queue in the proximity of the departing link capacity to enable the congestion and queue length control.

FN integrates the instantaneous queue length and the average arrival rate of queue to compute the drop/mark probability of the packet upon each arriving packet.

The use of the instantaneous queue length in conjunction with the average queue speed (average arrival rate) can provide superior control decision criteria for an active queue management scheme [7].

A. Linear Packet Dropping/Marking Probability

The linear dropping/marketing probability function [8] is derived based on the assumption that the arrival traffic process remains unchanged over the control time constant period of length (T) seconds. In other words, it is supposed that immediately following the packet's arrival, the traffic continues to arrive at the fixed rate of (R) bits/sec, the estimated average arrival rate to the buffer computed upon the packet's arrival, for the period of the control time constant. The buffer has a capacity of (C) bits and is served by an outgoing link at a fixed rate of (μ) bits/sec. The packet drop/mark probability (P), is computed for, and applied to, every incoming packet, based on the above assumptions, with the goal of driving the instantaneous (current) queue length (Q_{cur}) to some desired optimal level (Q_{opt}) over the control time constant period (T). These are shown in Figure 4. In FN, the linear probability of dropping/marketing packet, P , is calculated by

$$P^{(i)} = \frac{((R_i - \mu) \cdot T) - (Q_{opt} - Q_{cur})}{R_i \cdot T} = \frac{\Delta Q_u - \Delta Q_d}{Q^+} \quad (3)$$

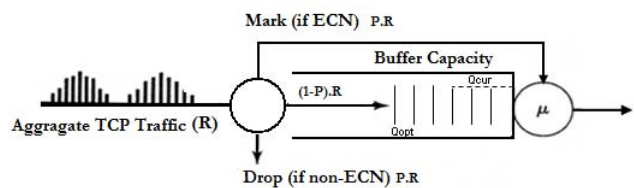


Fig. 4 FN gateway buffer

IV. EVALUATION OF FN

A. Network Simulation Topology and Parameters

For evaluating the performance of the FN mechanism, a network topology of multiple sources with one bottleneck gateway was used (Fig. 5). This network topology has become an standard topology for studying queue management algorithms (e.g., see [4])

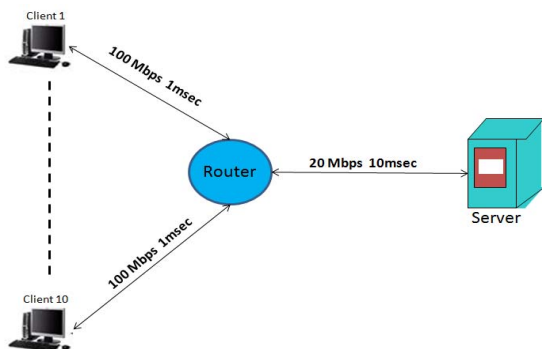


Fig. 5 Network topology

The traffic arriving on the incoming links are aggregated and multiplexed on the single outgoing link. This is the typical scenario found either in an access or core router in a network which would require the use of a queue management algorithm. In all simulations, ECN-capable TCP-NewReno was used for five flows and Non-ECN-capable TCP-NewReno for the rest. Each flow was fed with an FTP traffic generator. The access links provide full-duplex connections at a data rate of (100) Mbps with a propagation delay of (1) msec between the clients and the router. The bottle-neck link between the router and the server has a data rate of (20) Mbps and a propagation delay of (10) msec. The capacity of access links has been chosen (5) times larger than the capacity of the bottle-neck link to ensure that the link between the router and the server is the only bottle-neck in the network. The propagation delay of the bottle-neck link has been chosen (10) times larger than that of the access links to signify the large propagation delay across a network as opposed to the small delay experienced in accessing the network. The average time between the arrivals of a new session at each node (in our scenario) is 45 msec. This means that on the average, 22 new sessions arrive at each node so that the global arrival rate of sessions is 22 times 10 (number of sources), which gives 220 sessions/sec. we generate sessions with random size with mean of 10KB, with Pareto (A typical distribution which depicts the file size) with shape 1.5. The global rate of bits generation is $220 \cdot 10^4 \cdot 8$ gives 17.6 Mbps which is larger than the capacity of the bottleneck.

B. Common FN/RED Router Settings

The common FN/RED router parameters are shown in Table 1.

TABLE I
COMMON FN/RED ROUTER SETTING

Parameter	Value	Description
BufferSize	1,680,000	Physical Buffer Size [bits]
MaxPacketSize	1064	Maximum Packet Size [bytes]
ProbabilityMode1	Byte	Packet Marking Computation Mode [Byte/Packet]
ProbabilityMode2	Uniform	Packet Marking Computation Mode [Geometric/Uniform]

The size of the physical buffer was set to 1,680,000 bits which is twice the Bandwidth Delay Product (BDP). BDP is computed by multiplying the bandwidth of the bottleneck link by average round trip time (RTT). This product specifies the maximum TCP window size for each TCP session. Since a TCP sender can immediately transmit as many packets as the window size advertised by the receiver, the receiver's advertised window size is a good estimation of how many packets can arrive at the router back to back. Therefore, the router buffer sizes are set as a multiple of the TCP receiver's advertised window size to allow queuing the arriving packets at the router and subsequently forwarding them [9].

The MaxPacketSize was set to 1064 bytes. The TCP maximum segment size (MSS) is 1024 bytes plus TCP and IP headers (each of 20 bytes). This means that the buffer in our scenario can accommodate up to 197 packets. FN and RED were set to function in byte mode where the packet marking/dropping probability is computed proportional to the arrived packet size with respect to the maximum packet size (MaxPacketSize). in addition, the packet marking/dropping probability was calculated in the uniform mode where the packet marking/dropping process is uniformized using the number of packets that admitted into the queue since the last randomly marked/dropped packet.

C. RED Experiments

All simulations are run for 120 seconds of simulation time. A total of 30 experiments involving different (Q_{min} , Q_{max}) settings were executed using RED mechanism with a queue averaging weight of $W_q=0.002$ and a maximum mark/drop probability of $P_{max}=0.1$. The set of the (Q_{min} , Q_{max}) settings consists of : { (1,3), (2,6), (3,9), (4,12), (5,15), (6, 18), (7, 21), (8, 24), (9, 27), (10, 30), (11, 33), (12, 36), (13, 39), (14, 42), (15, 45), (16, 48), (17, 51), (18, 54), (19, 57), (20, 60), (21, 63), (22, 66), (23, 69), (24, 72), (25, 75), (26, 78), (27, 81), (28, 84), (29, 87), (30, 90)} where the pairs are in the units of equally sized packets. A higher link utilization was achieved by larger value of (Q_{min} , Q_{max}) pairs that allowed a larger average queue size. Among the 30 RED experiments, the (Q_{min} , Q_{max}) pairs consisted of the set { (1,3), (2,6), (3,9), (4,12), (5,15)} maintained successfully the average queue size below the required optimal queue size of $Q_{opt} = 13125$ bytes that used in FN experiments. Specifically, the highest link utilization was achieved when (Q_{min} , Q_{max}) = (5, 15). Thus, this specific RED experiment was used as the RED reference experiment for the performance studies.

D. FN Experiments

All simulations are run for 120 seconds of simulation time. A total of 192 experiments were executed for both linear and quadratic FN algorithms (96 experiments each). In all of the 196 experiments, the required optimal queue size was set to

$Q_{opt} = 13125$ bytes. The packet window size (W_s) and time constant (T) parameters were varied over 8 and 12 different values, respectively. The 96 linear and quadratic FN experiments symbolize the different (W_s, T) settings acquired from the Cartesian Product of the two sets: $W_s = \{32, 64, 128, 256, 512, 1024, 2048, 4096\}$ and $T = \{1, 8, 27, 64, 125, 216, 343, 512, 729, 1000, 1331, 1728\}$. The window size settings are in packets and the time settings are in milli-seconds. Eight experiments were chosen amongst each of the linear and quadratic sets of 96 experiments as linear or quadratic FN reference experiments for performance studies. These eight experiments, which symbolize the set of eight different window size (W_s) settings, were selected according to the average queue size noticed during the experiments. Among the twelve experiments symbolize the twelve different time (T) settings for each window size (W_s) settings, one experiments was selected as the reference experiments for specific window size (W_s) setting. The selected experiment is the one resulting in the highest link utilization for which the average queue size is smaller than the required optimal queue size ($E\{Q\} \leq Q_{opt}$) and the rate of packet drops is less than for the RED reference experiment.

E. Performance Metrics

The following performance metrics are used for the quantitative performance comparison between FN and RED:

- **Outgoing transmission Link Utilization:** Measures the ratio of time the link is utilized ($U = \frac{\text{Busy Time [seconds]}}{\text{Simulation Time [seconds]}} \cdot 100\%$)
- **Packet Loss:** Measures the ratio of the number of packets discarded at the bottleneck link to the total number of packets inserted into the bottleneck link buffer for all source $L = \frac{\text{Dropped Packets [bits]}}{\text{Total Packets [bits]}} \cdot 100\%$
- **Queue Size:** Statistical expected average queue size ($E(Q) = \frac{1}{N} \sum_{i=1}^N Q_i$)

The network efficiency, with which the network resources are utilized, can be measured based on the link utilization and packet loss values. This information is important to network owners and operators. They can maximize their profits by minimizing their cost based on this information.

The network efficiency, with which the network resources are utilized, can be measured based on the link utilization and packet loss values. This information is important to network owners and operators. They can maximize their profits by minimizing their cost based on this information.

The average queue size points out the average queuing delay experienced by the traffic passing through the router. The quality of service provided to the network users is indicated by the queuing delay and the packet loss. On the other hand, the queuing delay and the packet loss determine the reliability and response time offered by the network.

V. SIMULATION RESULTS

This section compares the performance of the RED algorithm to the Linear FN algorithm in terms of the link utilization, packet loss, and average queue size. The experiments selected for performance comparisons between RED and FN algorithms are listed in Table 2. The first line of the table shows the experiment using the RED algorithm and the subsequent eight lines show the experiments using the FN algorithm with different window size (W_s) and time (T) settings. The (Q_{min}, Q_{max})/(W_s, T) column identifies the (Q_{min}, Q_{max}) or (W_s, T) settings depending on whether the specific row in the table shows a RED or FN experiment. Q_{min}, Q_{max} and W_s are in packets units and T is in seconds units. The Utilization column identifies the ratio of utilization obtained on the outgoing transmission link over the simulation duration. The (Random, Overflow, Drop) % column shows the ratio of the traffic discarded at the router due to random packet dropping, due to buffer overflow, and the total packet drops. The [($E\{r\}, SD\{r\}$) (bps), ($E\{Q\}, SD\{Q\}$) (bits)] column identifies the mean and standard deviations of the average traffic rate and the current queue size noticed during the experiment. The $Pr(Q_{cur} \leq Q_{opt})$ column identifies the probability that the current queue size is smaller than the required optimal queue size of $Q_{opt}=105,000$ bits.

The data in the table shows that the FN provides a better performance compared to RED. All FN experiments provide higher link utilization, lower packet drop ratio, and lower average queuing delay that identified by the lower average queue size while efficiently averting buffer overflow. This means that FN has better reliability and lower queuing and end-to-end delay which are all desired by the network providers as well as the network users. In addition, by comparing the standard deviations for the RED reference experiments to that for the FN reference experiment, listed in Table 2, it is clear that FN reduces the fluctuations in the instantaneous queue size, maintains a higher link utilization, and acquires a lower ratio of packet drops compared to RED.

VI. CONCLUSIONS

In this paper we have shown how FN mechanism can be tuned with different (W_s, T) settings to efficiently control the instantaneous (current) queue length below the required optimal queue length (Q_{opt}) to reduce queuing delay, maintain a high link utilization and a low packet drop ratio. The settings that enable a tight control of the queue size below the required optimal (Q_{opt}) value, efficiently limiting the queuing delay at the cost of a relatively lower link utilization. Conversely, settings for a less strict control would improve the link utilization at the cost of a higher queuing delay. However, in either case the FN performance has been shown to be superior to RED performance in terms of maintaining a smaller average queue length with a smaller standard deviation to provide a lower queuing delay with less delay variation.

TABLE II
RED-LINEAR FN COMPARISON: UTILIZATION, PACKET LOSS, QUEUE SIZE

$(Q_{min}, Q_{max}) / (W_s, T)$	Utilization	(Random, Overflow, Drop) %	(E{r}, SD{r}) (bps) (E{Q}, SD{Q}) (bits)	$Pr(Q_{cur} \leq Q_{opt})$
(5, 15)	72.093	(4.623, 0.000, 4.623)	Q (108864.241, 107891)	0.594
(32, 0.512)	77.845	(4.175, 0.000, 4.175)	r (10019489.181, 8.162e+06) Q (37489.611, 37889.3)	0.941
(64, 0.216)	85.633	(3.582, 0.000, 3.582)	r (10001310.501, 2.138e+06) Q (56422.719, 46788.9)	0.849
(128, .0.125)	84.832	(4.510, 0.000, 4.510)	r (10060119.378, 1.891e+06) Q (97129.844, 71740.5)	0.624
(256, 0.064)	83.671	(3.787, 0.000, 3.787)	r (9678881.947, 1.584e+06) Q (93251.197, 68761)	0.632
(512, 0.064)	89.660	(4.136, 0.000, 4.136)	r (9972714.198, 1.161e+06) Q (106351.791, 70858.2)	0.548
(1024, 0.027)	88.038	(3.442, 0.000, 3.442)	r (9585846.039, 1.243e+06) Q (80289.715, 56788.7)	0.689
(2048, 0.027)	86.414	(3.583, 0.000, 3.583)	r (9519822.582, 1.352e+06) Q (83573.763, 59731)	0.655
(4096, 0.027)	86.021	(3.220, 0.000, 3.220)	r (9528891.480, 1.453e+06) Q (84688.912, 61742.1)	0.671

REFERENCES

[1] B. Braden, D. Clark, J. Crowcroft, B. Davie, S. Deering, D. Estrin, S. Floyd, V. Jacobson, G. Minshall, C. Partridge, L. Peterson, K. Ramakrishnan, S. Shenker, J. Wroclawski, and L. Zhang, *Recommendations on Queue Management and Congestion Avoidance in the Internet*: RFC Editor, 1998.

[2] S. Leonardo, P. Adriano, and M. Wagner, Jr., "Reactivity-based Scheduling Approaches For Internet Services," in *Proceedings of the Fourth Latin American Web Congress*: IEEE Computer Society, 2006.

[3] M. Hassan and R. Jain, *High Performance TCP/IP Networking: Concepts, Issues, and Solutions*: Pearson Prentice Hall, 2004.

[4] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *Networking, IEEE/ACM Transactions on*, vol. 1, pp. 397-413, 1993.

[5] F. Wu-Chang and D. K. Dilip, "Adaptive packet marking for maintaining end-to-end throughput in a differentiated-services internet," *IEEE/ACM Trans. Netw.*, vol. 7, pp. 685-697, 1999.

[6] M. M. Kadhum and S. Hassan, "Fast Congestion Notification mechanism for ECN-capable routers," in *Information Technology, 2008. ITSIM 2008. International Symposium on*, 2008, pp. 1-6.

[7] M. M. Kadhum and S. Hassan, "The Design Motivation and Objectives for Fast Congestion Notification (FN)," in *the 28th APAN Network Research Workshop*, Malaysia, 2009, pp. 25-30.

[8] M. M. Kadhum and S. Hassan, "A Linear Packet Marking Probability Function for Fast Congestion Notification (FN)," *International Journal of Computer Science and Network Security*, vol. 9, pp. 45-50, 2009.

[9] W. Richard Stevens, *TCP/IP Illustrated, Volume 1: The Protocols* Addison-Wesley Professional Computing Series, 1994.