Los Alamos Technical Report LA-UR 01-6506

On the Compatibility of TCP Reno and TCP Vegas

W. Feng and S. Vanichpun

Abstract—**Despite a plethora of research demonstrating the superiority of TCP Vegas over TCP Reno, Reno is still the most widely deployed implementation of TCP. The primary reason for this predicament is due to the alleged incompatibility of Vegas with Reno. While Vegas in isolation performs better with respect to overall network utilization, stability, fairness, throughput and packet loss, and burstiness; its performance is generally awful in any environment where Reno connections exist. Consequently, there exists no incentive for any operating system to adopt TCP Vegas.**

In this paper, we show that the accepted (default) configuration of Vegas is indeed incompatible with TCP Reno. However, with a careful analysis of how Reno and Vegas use buffer space in routers, Reno and Vegas can be compatible with one another if Vegas is configured properly. Furthermore, we show that overall network performance actually improves with the addition of properly configured Vegas flows competing head-to-head with Reno flows, thus lending credence to the incremental adoption of Vegas.

Keywords— **TCP Reno, TCP Vegas, congestion control, congestion avoidance, compatibility, fairness.**

I. Introduction

To address a series of congestion collapses starting in October 1986, Jacobson proposed a congestion-control mechanism in TCP that later became known as TCP Tahoe [1]. Since that time, many modifications have been made to TCP, resulting in two more notable implementations — TCP Reno [2] and TCP Vegas [3].

TCP Reno, like TCP Tahoe, allows congestion to occur (i.e., induces packet loss) in order to estimate the available bandwidth in the network. Once packet loss is detected, Reno recovers by cutting its window size in half. This behavior causes a periodic oscillation in the window size; an oscillation that many next-generation Internet applications (such as video teleconferencing) do not tolerate well. Furthermore, recent work shows that this oscillatory behavior induces chaotic behavior into the network [4–6], thus adversely affecting overall network performance.

W. Feng is with the Computer & Computational Sciences (CCS) Division at Los Alamos National Laboratory (LANL) and the Dept. of Computer & Information Science at the Ohio State University. S. Vanichpun is also with the CCS Division at LANL and the Dept. of Electrical & Computer Engineering at the University of Maryland.

In contrast, TCP Vegas generally performs better with respect to overall network utilization [3,7], stability [8,9], fairness [8, 9], throughput and packet loss [3, 7], and burstiness [5] when the entire network consists of Vegas-only connections. However, research to date has shown that when Reno and Vegas perform head-to-head, Reno generally steals bandwidth from Vegas [7, 8], thus eliminating the possibility of an incremental adoption of Vegas to the Internet. Consequently, while Vegas has been around for over five years, its adoption has been non-existent due to perceived incompatibilities between Reno and Vegas.

With a careful analysis of how Reno and Vegas use buffer space in routers, we will show that Reno and Vegas can be compatible with one another if Vegas is configured properly. Further, we show that overall network performance actually improves with the addition of properly configured Vegas flows competing head-to-head with Reno flows, thus lending credence to the idea that incremental adoption of Vegas is possible.

The rest of the paper is organized as follows. Section II briefly discusses the congestion detection and avoidance mechanisms of Reno and Vegas. Section III analytically shows that Reno and Vegas in its default (mis)configuration are indeed incompatible but that a properly configured Vegas *is* compatible with Reno. We confirm our analysis with simulation results in Section IV. This confirmation then leads us to more thoroughly examine fairness issues between Reno and Vegas in Section V. Finally, in the remaining sections, we critique our research, discuss future work, and conclude.

II. Congestion-Control Mechanisms

To ensure efficient use of network bandwidth, TCP controls its sending rate based on feedback from the network. In order to control the sending rate, TCP estimates the available bandwidth in the network via a *bandwidth-estimation scheme* [8]. In Tahoe and Reno, the bandwidth-estimation scheme uses packet losses (as an indication of network congestion) to estimate available bandwidth while Vegas uses the difference in the expected and actual sending rates. This difference in schemes allows Vegas to achieve higher network efficiency and fewer packet re-transmissions than Reno.

A. TCP Reno

While there are no packet losses, Reno continues to increase its window size, and hence sending rate, by one each round-trip time (RTT). With this continual increase in window size, TCP Reno allows congestion to occur; it then detects congestion via packet loss and recovers from it by halving the size of the sender window (i.e., halving the sending rate).[1] This *additive increase, multiplicative decrease (AIMD)* algorithm, based on the assumptions and results in [10], results in a fair allocation of bandwidth between connections. However, Reno fails to achieve such fairness because it is not a synchronized, rate-based control scheme as assumed in [10].

In addition, the above periodic oscillation in window size due to the continual update of the window size leads to an oscillation in the RTT of packets. This results in larger delay jitter and an inefficient use of available bandwidth, particularly for streaming applications.

B. TCP Vegas

Vegas enhances Reno by adopting a bandwidth-estimation scheme that tries to avoid rather than react to congestion. Specifically, Vegas uses the difference in the expected and actual flow rates to estimate the available bandwidth in the network. When the network is uncongested, the actual flow rate is close to the expected flow rate; otherwise, the actual rate is smaller than the expected rate, indicating that buffer space in the network is filling up and that the network is approaching a congested state. The difference in flow rates can be translated into the difference between the window size and the number of acknowledged packets during the RTT, respectively, i.e., $Diff = (Expected - Actual) \times BaseRTT$ where $Expected$ is the expected rate, $Actual$ is the actual rate, and $BaseRTT$ is the minimum-observed RTT.

To adjust the size of the congestion window ($cwnd$) appropriately, Vegas uses two threshold values, α and β (whose default values are 1 and 3, respectively), to control the adjustment of $cwnd$ at the source host as follows:

- Let $Expected = \frac{cwnd}{BaseRTT}$.
- Estimate $Actual = \frac{cwnd}{RTT}$ where RTT is the actual RTT of a packet.
- Compute the estimated backlog of packets in the network queues using $Diff = (Expected - Actual) \times BaseRTT$.
- Adjust $cwnd$ as follows:

$$cwnd \leftarrow \begin{cases} cwnd + 1 & \text{if } Diff < \alpha \\ cwnd - 1 & \text{if } Diff > \beta \\ cwnd & \text{otherwise } (\alpha \leq Diff \leq \beta) \end{cases}$$

Conceptually, Vegas tries to keep at least α packets but no more than β packets queued in the network. Thus, when there is only one Vegas connection, the window size of Vegas converges to a point that lies between $window + \alpha$ and $window + \beta$ where $window$ is the maximum window size that does not cause any queueing, i.e., uncongested bandwidth-delay product.

Selecting the α and β parameters holds an implicit tradeoff between network utilization, goodput, and fairness. By using only the default settings for these parameters, i.e., $\alpha = 1, \beta = 3$, prior research inadvertently favored Reno over Vegas [8, 11].

III. Compatibility of TCP Reno and TCP Vegas

Prior research demonstrates that Vegas (in isolation) generally performs better than other implementations of TCP [3, 5, 7–9]. Ahn et al. [7] and Mo et al. [8] also show that when a Vegas connection competes with a Reno connection, Vegas does not receive a fair share of bandwidth due to its conservative congestion-avoidance mechanism.

In this section, we illustrate that the "conservative" congestion-avoidance mechanism is *not* to blame for Vegas's inability to grab a fair share of bandwidth. Rather, the alleged incompatibility between Reno and Vegas is due to using the default (mis)configuration of Vegas parameters, i.e., $\alpha = 1$ and $\beta = 3$.

Consider one TCP Vegas connection and one TCP Reno connection over a bottleneck link. At steady state, the Vegas connection keeps (on average) approximately $\bar{q}_v \leq \beta$ packets in the queue while the Reno connection tries to gain as much bandwidth (and queue space) as it can until a packet is lost. Hence, the number of TCP Reno packets in the queue is $q_r \in [0, B - \bar{q}_v]$ where B is the buffer capacity at the bottleneck link. Assuming that the average value of q_r is $\bar{q}_r = \frac{B - \bar{q}_v}{2}$ (e.g., uniform distribution), the ratio of Vegas throughput (λ_v) to Reno throughput (λ_r) is given by

$$\frac{\lambda_v}{\lambda_r} = \frac{\bar{q}_v}{\bar{q}_r} \tag{1}$$

[1] Theoretically, without any buffering and with synchronous feedback, TCP Reno can only utilize 75% of the bandwidth.

$$= \frac{2\bar{q}_v}{B - \bar{q}_v}. \quad (2)$$

Thus, when $B = 3\bar{q}_v$, the ratio of the throughputs is one. As B increases further, Reno is favored. Hasegawa et al. [11] provide a more complete analysis of the throughputs when there are N_v Vegas connections and N_r Reno connections competing for the queue space at the same bottleneck link.

In [12], Weigle and Feng propose a heuristic algorithm for setting α and β so that Vegas can compete with Reno in both droptail and RED queues [13]. For a droptail queue, they set $\alpha = \lfloor \frac{B}{3} - 1 \rfloor$ and $\beta = \lfloor \frac{B}{3} \rfloor$ where $\lfloor x \rfloor$ denotes the floor of x. If we again assume that $\bar{q}_r = \frac{B-\bar{q}_v}{2}$ and note that the worst ratio of Vegas throughput occurs when $\bar{q}_v = \alpha$, then substituting these values into (2) gives

$$\begin{aligned} \frac{\lambda_v}{\lambda_r} &= \frac{2B-6}{2B+3} \\ &\approx 1 \quad \text{if } B \text{ is large enough.} \end{aligned}$$

In the remainder of this section, we first present a more rigorous analysis (modified from [11]) for setting α and β when there are only two connections (one Reno and one Vegas). We then generalize our analysis for multiple connections.

A. Analysis of Two Connections

Consider the case when one Vegas connection and one Reno connection share a bottleneck link using a droptail queue. Let the bottleneck link have a transmission rate of μ packets/s and a round-trip propagation delay of τ seconds with queue size B packets. Let q_v and q_r denote the number of Vegas packets and Reno packets in the queue, respectively. Then, in order to allow Vegas to compete with Reno, Vegas must set its α and β parameters so that Equation (2) is such that $\lambda_v/\lambda_r \approx 1$.

Let W, d, and D denote the window size (i.e., $cwnd$), BaseRTT delay, and actual RTT delay of Vegas, respectively. At steady state, Vegas tries to keep the difference ($Diff$) between its actual throughput ($Actual$) and expected throughput ($Expected$) between α and β where

$$Actual = \frac{W}{D} \quad \text{and} \quad Expected = \frac{W}{d}.$$

Then, as noted in Section II, we calculate $Diff$ as follows:

$$Diff = (Expected - Actual) \cdot d = W\frac{(D-d)}{D}, \quad (3)$$

and interpret $Diff$ as the number of packets in the queue, i.e., $Diff \approx \bar{q}_v$.

Let $W_{v_{max}}$ and $W_{r_{max}}$ denote the window sizes of Vegas and Reno when the queue is full, respectively. Since Vegas keeps $q_v \approx \bar{q}_v$ at all time and $Diff \approx \bar{q}_v$, we have

$$W_{v_{max}} \frac{D-d}{D} = \bar{q}_v,$$

or equivalently,

$$\begin{aligned} W_{v_{max}} &= \bar{q}_v \frac{D}{D-d} \\ &= \bar{q}_v \frac{\tau + B/\mu}{(B-1)/\mu} \\ &\simeq \bar{q}_v \frac{\tau + B/\mu}{B/\mu}, \quad \text{if } d \simeq \tau, \end{aligned} \quad (4)$$

where $d = \tau + 1/\mu$ and $D = \tau + B/\mu$ are the values of the BaseRTT delay and actual RTT delay of Vegas when the queue is full, respectively. Therefore, the window size of Reno when the queue is full is given by

$$W_{r_{max}} = B + \mu\tau - W_{v_{max}} \quad \text{if } d \simeq \tau. \quad (5)$$

After the queue is full, the probability that a Reno packet will be dropped is

$$p_r = \frac{q_{r_{max}}}{q_{r_{max}} + q_{v_{max}}} = \frac{B - \bar{q}_v}{B} \quad (6)$$

and the probability that a Vegas packet will be dropped is

$$p_v = \frac{q_{v_{max}}}{q_{r_{max}} + q_{v_{max}}} = \frac{\bar{q}_v}{B}. \quad (7)$$

We now consider two approximations for setting α and β. Approach 1: We assume that the packets dropped are only from the Reno connection and consider only the fast retransmit of Reno while ignoring its timeout mechanism. Therefore, if a Reno packet is dropped, then the Reno window size $W_{r_{max}}$ drops to $\frac{W_{r_{max}}}{2}$. Approach 2: We also consider the case of a Vegas packet being dropped. Hence, if a Reno packet is dropped, its window size evolves the same way as in Approach 1, else if a Vegas packet is dropped, then the Reno window size increases by 1.

A.1 Approach 1

Figure 1 shows a simulation of one Reno connection and one Vegas connection at steady state. Clearly, from

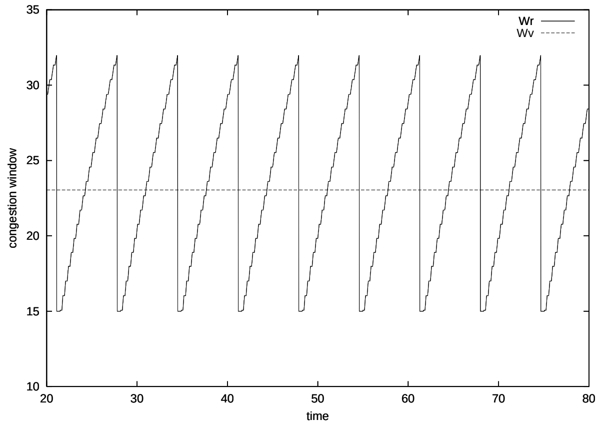

Fig. 1. Evolution of the Congestion-Control Window

this figure, the average window size of the Reno connection ($\bar{W}_r$) is given by

$$\bar{W}_r = \frac{W_{r_{max}} + \frac{W_{r_{max}}}{2}}{2} = \frac{3}{4} W_{r_{max}}. \tag{8}$$

By definition, the average queue size of the Reno connection is the fraction of $\bar{W}_r$ that is buffered in the queue, i.e.,

$$\bar{q}_r = \bar{W}_r \frac{B}{B + \mu\tau}. \tag{9}$$

Substituting (8) into (9) gives

$$\bar{q}_r = W_{r_{max}} \frac{3B}{4(B + \mu\tau)}. \tag{10}$$

Substituting (10) into (1), the ratio of the throughputs is given by

$$\begin{aligned} \frac{\lambda_v}{\lambda_r} &= \frac{4\bar{q}_v(B + \mu\tau)}{W_{r_{max}} 3B} \\ &= \frac{4\bar{q}_v}{3(B - \bar{q}_v)}, \end{aligned} \tag{11}$$

where (11) follows from (5) and (4). Now, by setting $\frac{\lambda_v}{\lambda_r} = 1$ and solving for $\bar{q}_v$, we get

$$\bar{q}_v = \frac{3}{7} B.$$

Thus, we set the TCP Vegas parameters to the following values:

$$\alpha = \lfloor \frac{3}{7} B - 1 \rfloor \quad \text{and} \quad \beta = \lfloor \frac{3}{7} B \rfloor.$$

The motivation for setting α to be one less than β is two-fold. First, setting $\alpha = \beta$ introduces stability problems [14], i.e., the congestion window oscillates around the equilibrium value. Second, setting α and β too far apart creates a larger stability region than needed, resulting in connections that can converge to opposite ends of the stability region, thus affecting fairness [9, 14].

A.2 Approach 2

If Reno or Vegas packet is dropped, the window size of Reno $\hat{W}_r$ becomes

$$\begin{aligned} \hat{W}_r &= \frac{W_{r_{max}}}{2} p_r + (W_{r_{max}} + 1) p_v \\ &\simeq \frac{W_{r_{max}}}{2} p_r + W_{r_{max}} p_v, \end{aligned} \tag{12}$$

where we assume that $W_{r_{max}} \gg 1$ and that p_r and p_v are given by (6) and (7). From (12), we then approximate the average window size of Reno ($\bar{W}_r$) by

$$\bar{W}_r = \frac{W_{r_{max}} + \hat{W}_r}{2} = W_{r_{max}} \frac{3B + \bar{q}_v}{4B}. \tag{13}$$

As in Section III-A.1 but with (9) and (13),

$$\bar{q}_r = \bar{W}_r \frac{B}{B + \mu\tau} = W_{r_{max}} \frac{B(3B + \bar{q}_v)}{4B(B + \mu\tau)}. \tag{14}$$

Hence, substituting (14) into (1) gives

$$\begin{aligned} \frac{\lambda_v}{\lambda_r} &= \frac{4\bar{q}_v B(B + \mu\tau)}{W_{r_{max}} B(3B + \bar{q}_v)} \\ &= \frac{4\bar{q}_v B}{3B^2 - 2\bar{q}_v B - \bar{q}_v^2}, \end{aligned} \tag{15}$$

where (15) follows from (5) and (4). To ensure compatibility between Reno and Vegas, we again set $\lambda_v / \lambda_r = 1$ and solve for $\bar{q}_v$, which yields

$$\bar{q}_v \simeq 0.4641B.$$

Therefore, we set TCP Vegas parameters as follows:

$$\alpha = \lfloor 0.4641B - 1 \rfloor \quad \text{and} \quad \beta = \lfloor 0.4641B \rfloor.$$

B. Analysis of Multiple Connections

In this section, we generalize our two-connection analysis from Section III-A to deal with N_v TCP Vegas connections and N_r TCP Reno connections sharing a bottleneck link of μ packet/s. Let every connection have a round-trip propagation delay of τ seconds while sharing a bottleneck buffer size of B packets. As in Section III-A, to ensure compatibility between each connection, each connection must equally share the bottleneck buffer, or equivalently, the average queue size of each connection must be identical. We also assume that the connections are synchronous, i.e., at any given instant in time, each Vegas connection i has queue size $q_{v_i} \approx \bar{q}_v \leq \beta \; \forall \; i = 1, 2, \ldots, N_v$ and each Reno connection j has the same average queue size $\bar{q}_r \; \forall \; j = 1, 2, \ldots, N_r$. Thus, TCP Vegas should set its α and β parameters so that

$$\frac{\bar{q}_v}{\bar{q}_r} \approx 1.$$

When the queue is full, the window size of Vegas $W_{v_{max}}$, derived from (3), is

$$\begin{aligned} W_{v_{max}} &= \bar{q}_v \frac{D}{D-d} \\ &\simeq \bar{q}_v \frac{\tau + B/\mu}{B/\mu}, \quad \text{if } d \simeq \tau, \end{aligned} \tag{16}$$

which is the same as (4). And similar to (5), we have

$$N_r W_{r_{max}} = B + \mu\tau - N_v W_{v_{max}}, \quad \text{if } d \simeq \tau, \tag{17}$$

where $W_{r_{max}}$ is the window size of TCP Reno when the queue is full. The dropped probabilities for Reno and Vegas are respectively

$$p_r = \frac{B - N_v \bar{q}_v}{B} \tag{18}$$

and

$$p_v = \frac{N_v \bar{q}_v}{B}. \tag{19}$$

B.1 Approach 1

Using the same approach as in Section III-A.1, the average window size ($\bar{W}_r$) and the average queue size ($\bar{q}_r$) of each Reno connection is given by

$$\bar{W} = \frac{W_{r_{max}} + \frac{W_{r_{max}}}{2}}{2} = \frac{3}{4} W_{r_{max}},$$

and

$$\bar{q}_r = \bar{W} \frac{B}{B + \mu\tau} = W_{r_{max}} \frac{3B}{4(B + \mu\tau)}, \tag{20}$$

respectively. Substituting (17) into (20), we get

$$\bar{q}_r = \frac{3(B - N_v \bar{q}_v)}{4N_r}. \tag{21}$$

Hence, the ratio of throughputs is given by

$$\frac{\lambda_v}{\lambda_r} = \frac{4N_r \bar{q}_v}{3(B - N_v \bar{q}_v)}. \tag{22}$$

And again, to ensure compatibility between Reno and Vegas, $\lambda_v / \lambda_r = 1$. Solving for $\bar{q}_v$ gives

$$\bar{q}_v = \frac{3B}{4N_r + 3N_v}.$$

Thus, we set α and β as

$$\alpha = \lfloor \frac{3B}{4N_r + 3N_v} - 1 \rfloor \quad \text{and} \quad \beta = \lfloor \frac{3B}{4N_r + 3N_v} \rfloor.$$

B.2 Approach 2

As in Section III-A.1, the window size of each Reno connection after a dropped packet is given by (12), and each average window size is given by (13). Combining (9), (13), (16), and (17) results in

$$\bar{q}_r = \frac{(B - N_v \bar{q}_v)(3B + \bar{q}_v)}{4N_r B}.$$

Substituting the above equation back into (1) gives

$$\frac{\lambda_v}{\lambda_r} = \frac{4N_r B \bar{q}_v}{(B - N_v \bar{q}_v)(3B + \bar{q}_v)}$$

Setting $\frac{\lambda_v}{\lambda_r} = 1$ helps ensure that the Reno and Vegas connection are compatible. Now, solving for $\bar{q}_v$, we have

$$\bar{q}_v = [\frac{\sqrt{(4N_r + 3N_v - 1)^2 + 12N_r} - (4N_r + 3N_v - 1)}{2N_r}]B,$$

and set the Vegas parameters as follows:

$$\alpha = \lfloor [\frac{\sqrt{(4N_r + 3N_v - 1)^2 + 12N_r} - (4N_r + 3N_v - 1)}{2N_r}]B - 1 \rfloor$$

and

$$\beta = \lfloor [\frac{\sqrt{(4N_r + 3N_v - 1)^2 + 12N_r} - (4N_r + 3N_v - 1)}{2N_r}]B \rfloor.$$

IV. Experiments

To verify our observations made through a heuristic analysis of the behavior of TCP Reno and TCP Vegas, we run two sets of simulations using the discrete-event simulation *ns*, version 2.1b8a [15].

A. Network Topologies & Parameters

We consider two networks based on the generic topology shown in Figure 2 and parametric details in Table I. The first network comes from [8] to use as a point of reference. The second network models the grid [16] between Los Alamos and Sandia National Laboratories.

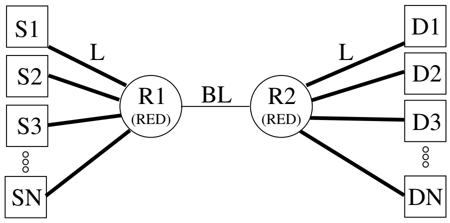

Fig. 2. Generic Topology

TABLE I
PARAMETERS OF SIMULATED NETWORKS

	Network	Reference	Grid
Node Pairs (N)		2	50
Links L	Bandwidth	10 Mb/s	100 Mb/s
	Delay	4 ms	1 ms
Link BL	Bandwidth	1.55 Mb/s	155 Mb/s
	Delay	4 ms	1.5 ms

For all our simulations, each connection starts an FTP session at time 0 and ends at 200 seconds with the packet size fixed at the standard 1500-byte Ethernet size. As in [8], we measure the number of ACKs that each connection successfully receives where ACK_R and ACK_V represent the number of ACKs from Reno and Vegas, respectively. In the general case of N_r Reno connections and N_v Vegas connections, ACK_R (resp. ACK_V) show the average number of ACKs over all N_r TCP Reno (resp. N_v TCP Vegas) connections.

B. Reference Network

We use the same network that [8] did in order to (i) confirm their results and (ii) confirm our analytic observations made in Section III. In confirming our analytic results, we show that Reno and Vegas are indeed compatible and that overall network performance improves by distributing bandwidth more evenly across *all* connections while still maintaining high overall throughput.

The remainder of this section examines the performance of our analytic heuristics in the case of two connections competing head-to-head and in the case of multiple connections competing. Within each case, we test our analytic heuristics of Approach 1 and 2.

B.1 One Reno vs. One Vegas (Head-to-Head)

The experimental set-up here is similar to [8]. The known differences in our tests are as follows: (i) Our two connections start at the same time whereas Mo et al. stagger the start of their two connections by 0.5 seconds. (ii) We use 1500-byte packets rather than 1000-byte packets. Other unknown differences could affect our results, e.g., the value of γ in Vegas. Despite the minor and potentially unknown differences between our tests, our results confirm the conclusions drawn by Mo et al. That is, as the buffer size increases, Reno uses more of the buffer to steal bandwidth from Vegas as Vegas is throttled by the "misconfigured" α and β parameters. In fact, Table II shows that in no case does Vegas get better throughput than Reno and that as the buffer size increases up to 100, Reno achieves 21 times higher throughput than Vegas.

TABLE II
REFERENCE NETWORK: HEAD-TO-HEAD WITH DEFAULT VEGAS (MIS)CONFIGURATION ($\alpha = 1$ AND $\beta = 3$)

Buffer	ACK_R	ACK_V	$\frac{ACK_R}{ACK_V}$
7	16,555	9,131	1.813
10	20,763	4,972	4.176
15	18,581	7,209	2.577
25	20,980	4,823	4.350
50	23,678	2,096	11.297
100	24,544	1,179	20.818

Tables III and IV illustrate that with the proper configuration of α and β, Vegas competes well with Reno. In fact, for the smaller buffer sizes, Vegas performs almost twice as well as Reno; this behavior occurs due to the aggressive nature of Reno's congestion-control mechanism, i.e., always trying to increase its window even though the buffer space is small. As the buffer sizes get larger and Vegas's α and β parameters adapt accordingly, we see that Reno and Vegas each get their fair share of bandwidth, i.e., the "fairness ratio" $ACK_R/ACK_V \approx 1$. Thus, these simulations confirm that our analytic heuristics from Sections III-A.1 and III-A.2 enable Reno and Vegas to be compatible with each other. In addition, the heuristics enhance overall network performance by distributing bandwidth more evenly across Reno and Vegas while keeping overall network throughput high (i.e., each row in Tables III and IV has roughly the same total throughput).

TABLE III
REFERENCE NETWORK: HEAD-TO-HEAD WITH $\alpha = \lfloor \frac{3}{7}B - 1 \rfloor$ AND $\beta = \lfloor \frac{3}{7}B \rfloor$ (APPROACH 1)

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
7	9,006	16,678	3	0.540
10	12,241	13,444	4	0.910
15	12,471	13,330	6	0.936
25	12,868	12,937	10	0.995
50	12,688	13,091	21	0.969
100	13,143	12,585	42	1.044

TABLE IV
REFERENCE NETWORK: HEAD-TO-HEAD WITH $\alpha = \lfloor 0.4641B - 1 \rfloor$ AND $\beta = \lfloor 0.4641B \rfloor$ (APPROACH 2)

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
7	9,006	16,678	3	0.540
10	12,241	13,444	4	0.910
15	12,471	13,330	6	0.936
25	12,868	12,937	10	0.995
50	11,747	14,032	23	0.837
100	12,233	13,495	46	0.906

B.2 Multiple Reno vs. Multiple Vegas (Fixed Buffer)

The set of tests performed here are identical to those in Section IV-B.1 with two exceptions: (i) There are 10 different TCP connections vying for network bandwidth. (ii) The buffer size is fixed at 250 packets.

Using the formulas for α and β from Sections III-B.1 and III-B.2, we again demonstrate that Vegas can be properly configured to be compatible with Reno. Tables V and VI show that the fairness ratio (i.e., ACK_R/ACK_V) is close to one in all cases. This is in stark contrast to Table II where the fairness ratio is 20.818, meaning that Reno gets that many times more bandwidth than Vegas.

B.3 Multiple Reno vs. Multiple Vegas (Variable Buffer)

In this set of tests, we fix the different types and number of connections while varying the size of the buffer. And as evidenced by Tables VII and VIII, we again demonstrate that with a properly configured Vegas, Vegas is compatible with Reno. Except for the first row of each table where Vegas beats Reno, Reno and Vegas each get their fair share of bandwidth as indicated by the fairness ratio of approximately one in each row of Tables VII and VIII.

TABLE V
REFERENCE NETWORK: N_r RENO AND N_v VEGAS WITH SETTING α AND β BY APPROACH 1 AND $B = 250$

N_r	N_v	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
2	8	3,153	2,407	23	1.309
4	6	2,500	2,600	22	0.962
6	4	2,645	2,427	20	1.090
8	2	2,542	2,622	19	0.969

TABLE VI
REFERENCE NETWORK: N_r RENO AND N_v VEGAS WITH SETTING α AND β BY APPROACH 2 AND $B = 250$

N_r	N_v	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
2	8	2,952	2,457	24	1.201
4	6	2,500	2,600	22	0.962
6	4	2,584	2,517	21	1.027
8	2	2,542	2,622	19	0.969

TABLE VII
REFERENCE NETWORK: 5 RENO AND 5 VEGAS WITH SETTING α AND β BY APPROACH 1

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
100	1,746	3,399	8	0.514
200	2,571	2,556	17	1.006
300	2,690	2,416	25	1.113
400	2,648	2,437	34	1.087
500	2,744	2,321	42	1.182

C. Grid Network

In this section, the experiments that were run over the reference network in Section IV-B are repeated on a model of a computational grid between Los Alamos and Sandia National Laboratories.

The experimental results for head-to-head competition on the grid network are shown in Tables IX and X. We observe that all the fairness ratios lie in the interval [0.703,1.465] and that the throughput is well distributed between both connections. These results are in stark contrast to Table II where Reno is run against the default Vegas configuration, resulting in fairness ratios that lie in the interval [1.813,20.818].

TABLE VIII
REFERENCE NETWORK: 5 RENO AND 5 VEGAS WITH SETTING α AND β BY APPROACH 2

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
100	1,746	3,399	8	0.514
200	2,571	2,556	17	1.006
300	2,621	2,486	26	1.054
400	2,648	2,437	34	1.087
500	2,719	2,346	43	1.159

For some particular cases, e.g., when the buffer size is 300 in Table IX, Reno achieves better throughput than Vegas. Why does this happen? Before steady state is reached, a Vegas packet is dropped, and a subsequent timeout occurs, thus providing Reno the opportunity to aggressively grab network bandwidth than Vegas relinquishes. The same reasoning can be used to explain Table X when the buffer size is 200.

TABLE IX
GRID NETWORK: HEAD-TO-HEAD WITH $\alpha = \lfloor \frac{3}{7}B - 1 \rfloor$ AND $\beta = \lfloor \frac{3}{7}B \rfloor$ (APPROACH 1)

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
25	1,088,323	1,279,521	10	0.851
50	1,132,451	1,377,165	21	0.822
100	1,237,816	1,337,189	42	0.926
200	1,190,797	1,310,759	85	0.908
300	1,566,684	1,004,753	128	1.459
400	1,543,070	1,322,325	171	1.167

TABLE X
GRID NETWORK: HEAD-TO-HEAD WITH $\alpha = \lfloor 0.4641B - 1 \rfloor$ AND $\beta = \lfloor 0.4641B \rfloor$ (APPROACH 2)

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
25	1,077,134	1,448,115	11	0.744
50	1,051,922	1,495,565	23	0.703
100	1,080,863	1,489,905	46	0.725
200	1,506,751	1,028,086	92	1.465
300	1,455,002	1,115,923	139	1.303
400	1,333,429	1,083,871	185	1.230

Tables XI through XIV show the experimental results for multiple TCP connections and different buffer sizes. The results in these tables closely verify the analytic heuristics that we developed in Section III. In fact, with the exception of one data point, all the fairness ratios lie within 5% of the ideal fairness ratio of one, i.e., [0.955,1.068].

TABLE XI
GRID NETWORK: N_r RENO AND N_v VEGAS WITH SETTING α AND β BY APPROACH 1 AND $B = 1,000$

N_r	N_v	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
2	8	248,824	260,565	93	0.955
4	6	258,381	257,747	88	1.002
6	4	256,206	260,656	83	0.983
8	2	259,968	244,561	78	1.063

TABLE XII
GRID NETWORK: N_r RENO AND N_v VEGAS WITH SETTING α AND β BY APPROACH 2 AND $B = 1,000$

N_r	N_v	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
2	8	223,790	266,824	96	0.839
4	6	253,251	261,166	89	0.970
6	4	255,889	261,131	84	0.980
8	2	260,206	243,617	79	1.068

TABLE XIII
GRID NETWORK: 5 RENO AND 5 VEGAS WITH SETTING α AND β BY APPROACH 1

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
500	259,967	256,245	42	1.015
750	257,939	258,098	64	0.999
1,000	257,946	258,247	85	0.999
1,250	255,283	259,859	107	0.982
1,500	255,258	259,820	128	0.982

V. FAIRNESS OF TCP RENO VS. TCP VEGAS

As shown in [13,17,18], TCP Reno favors connections with shorter delays. In contrast, Mo et al. [8] demonstrate that TCP Vegas does not suffer from this delay bias via a closed, fluid model and simulation; however, they do not consider the fairness between TCP Reno and TCP Vegas because of the demonstrated incompatibility of TCP Reno and TCP Vegas (in its default configuration, i.e,. $\alpha = 1$ and $\beta = 3$).

Because we have demonstrated the compatibility of TCP Reno and a properly configured TCP Vegas in Section III, we address the fairness issue between TCP Reno and TCP Vegas here. By following the analysis of closed,

TABLE XIV
GRID NETWORK: 5 RENO AND 5 VEGAS WITH SETTING α AND β BY APPROACH 2

Buffer	ACK_R	ACK_V	β	$\frac{ACK_R}{ACK_V}$
500	257,894	258,318	43	0.998
750	255,366	260,672	65	0.979
1,000	253,942	262,251	87	0.968
1,250	253,966	261,168	108	0.972
1,500	252,422	262,657	130	0.961

fluid-model approximation in [8], we graphically illustrate the fairness of TCP Reno and TCP Vegas at steady state in the bottleneck link.

In the steady state, let $W_r(t)$ and $W_v(t)$ be the window sizes of TCP Reno and TCP Vegas at time t, respectively. By assuming that the throughput and the queue size of each connection is relatively constant, we have

$$W_i(t) = q_i(t) + \lambda_i(t) \cdot d_i, \quad i = r, v, \tag{23}$$

where $\lambda_i(t), i = r, v$, $q_i(t), i = r, v$ and $d_i, i = r, v$ are the throughput, queue size and BaseRTT, respectively, for Reno and Vegas connections. Moreover, we also have

$$\lambda_i(t) = \frac{W_i(t)}{p(t) + d_i}, \quad i = r, v, \tag{24}$$

where $p(t)$ denotes the queueing delay of the bottleneck link at time t. Let μ and B denote the capacity and the buffer size of the bottleneck link. With large enough B, we assume that the link is fully utilized, so we have

$$\frac{W_r(t)}{p(t) + d_r} + \frac{W_v(t)}{p(t) + d_v} = \mu. \tag{25}$$

(The fluctuation of TCP Reno will not reduce the utilization if the queue is large enough.) For simplicity, we let $d_r = d_v = d$. By (25), we find $p(t)$ as

$$p(t) = \frac{W_r + W_v - \mu d}{\mu}. \tag{26}$$

Hence, from (23), (24), and (26) and given the window size of TCP Reno is W_r and the queue size of TCP Vegas is q_v; we can compute the window size of TCP Vegas (W_v) in the steady state as

$$\begin{aligned} W_v(t, q_v(t)) &= \frac{\mu d + q_v(t) - W_r}{2} \\ &+ \frac{\sqrt{(\mu d + q_v(t) - W_r)^2 + 4 q_v(t) W_r}}{2} \end{aligned} \tag{27}$$

Similarly, given W_v and q_r, we also have

$$\begin{aligned} W_r(t, q_r(t)) &= \frac{\mu d + q_r(t) - W_v}{2} \\ &+ \frac{\sqrt{(\mu d + q_r(t) - W_v)^2 + 4 q_r(t) W_v}}{2} \end{aligned} \tag{28}$$

In order to show the stability region of Reno and Vegas, we first note that Vegas keeps its q_v between α and β; hence, its window size will be stable if it lies between $W_v(t, \alpha)$ and $W_v(t, \beta)$. Furthermore, the window size of TCP Reno can be varied as a function of $q_r(t)$ where the two extreme cases are $q_r(t) = 0$ and $q_r(t) = B - q_v(t)$. Therefore, by (28) and the fact that $q_r(t) = W_r(t) \cdot \frac{B}{\mu d + B}$, $W_r(t) \in [\mu d - W_v(t), \mu d + B - W_v(t)]$ for any given $W_v(t)$. However, when TCP Reno incurs a packet loss, its window size is halved. If multiple losses do not occur, Reno resumes its linear increase up to the maximum value; therefore, we assume that the size of the TCP Reno window at the stability region lies between $\max(\frac{\mu d + B - W_v(t)}{2}, \mu d - W_v(t))$ and $\mu d + B - W_v(t)$. For convenience, we denote $Rmax = \mu d + B - W_v(t)$ and $Rmin = \mu d - W_v(t)$

As stated above, we plot the graph of the stability region for one Reno connection versus one Vegas connection over the "Reference Network" (Figure 2 and Table I) with $B = 10$. First, in Figure 3, we show the stability region of two TCP Vegas connections where the line $alpha_i$ and $beta_i, i = 1, 2$ denote $\{(W_1, W_2) | W_i - \lambda_i d = \alpha\}$ and $\{(W_1, W_2) | W_i - \lambda_i d = \beta\}$, respectively. Furthermore, the fairness line is the line such that the window sizes of both connections are the same, i.e., $\{(W_1, W_2) | \lambda_1 = \lambda_2\}$. In this case, the fairness line passes right through the stability region (also referred to as the convergence region); hence, by using only TCP Vegas in the network, all connections get fair throughput. (However, there are "throughput/stability" problems with Vegas [9,14], which we discussed earlier in the paper.) In Figure 4, we use one TCP Reno connection and one TCP Vegas connection while keeping the default parameters for TCP Vegas, i.e., $\alpha = 1$ and $\beta = 3$. For this case, the convergence region of Reno and Vegas lies between the lines $Rmax$, $Rmax/2$, $alpha$ and $beta$. It is easily seen that the fairness line hardly passes through the convergence region; hence, it introduces unfairness, and TCP Reno's throughput is higher than TCP Vegas's. However, when both α and β are set as in Section III, we improve the fairness, as shown in Figure 5 where the fairness line clearly goes through the middle of the convergence region.

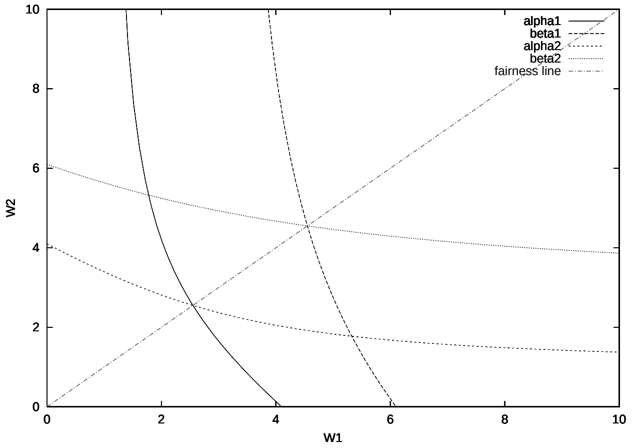

Fig. 3. Two TCP Vegas Windows in the Reference Network with $\alpha = 1$ and $\beta = 3$

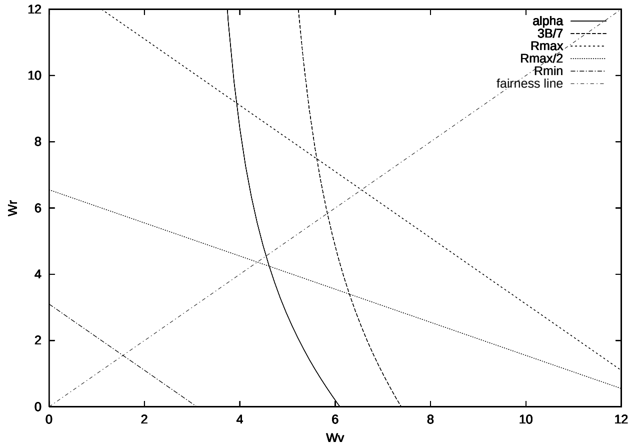

Fig. 5. TCP Reno (W_r) and TCP Vegas (W_v) in the Reference Network with $B = 10$, $\alpha = 3$, $\beta = 4$, and $3B/7 = 4.285$

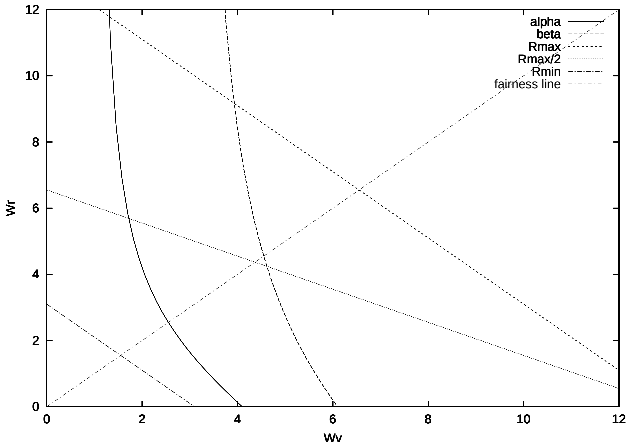

Fig. 4. TCP Reno (W_r) and TCP Vegas (W_v) Windows in the Reference Network with $B = 10$, $\alpha = 1$, and $\beta = 3$

VI. Discussion

In a Vegas-only environment, Vegas has been shown to perform better with respect to overall network utilization [3, 7], stability [8, 9], fairness [8, 9], throughput and packet loss [3, 7], and burstiness [5]. However, while Vegas is efficient in its use of available bandwidth, prior research demonstrates that it is not competitive or compatible with Reno in general. Thus, in an Internet environment dominated by TCP Reno implementations, there is no incentive for anyone to switch to Vegas. [2]

As in prior research, we confirm that the default configuration of Vegas with $\alpha = 1$ and $\beta = 3$ is *incompatible* with Reno. However, in contrast to prior research, we derive closed-form equations to properly set α and β and illustrate that such a properly configured Vegas is compatible with Reno. This insight, coupled with the ability of Vegas to maintain a relatively steady rate of throughput (e.g., see Figure 1), may provide enough benefits and incentive for the incremental adoption of Vegas. Ideally, by providing this path of incremental adoption, Vegas would enjoy wider deployment; and if it is successful in "supplanting" Reno, then Internet end-users would enjoy better overall network utilization, stability, fairness, throughput and packet loss, and smoother traffic patterns. Furthermore, with the "elimination" of Reno, the values of α and β would not have to be set so large to compete with Reno. Consequently, the queueing delay on individual connections would decrease substantially.

One perceived weakness of this paper is that TCP Vegas requires two critical pieces of information that are generally not known by end hosts: (i) the number of connections going through each router and (ii) the queue size of each router. However, the authors note that with the recent prototyping of an active queue-management (AQM) scheme called stochastic fair blue (SFB) [19] by Cisco and LVL7, the former can be estimated with SFB's Bloom filter [20], and the latter is known by the router. Thus, these two pieces of information can be transmitted back to the source (*a la* ECN [21]). Consequently, the features of these next-generation AQM routers not only can provide the above two pieces of critical information, but they can also open up new research possibilities in iterative algorithms to dynamically configure (or set) α, β, and perhaps, even γ. Furthermore, the ability of a

[2] In fact, the Linux operating system removed TCP Vegas from its 2.1 kernel and replaced it with TCP Reno for its 2.2 kernel.

properly configured Vegas to maintain a relatively steady rate of throughput greatly benefits streaming applications such as video-on-demand, video teleconferencing, voice over IP, virtual collaboratories, and remote computational steering of visualization data sets.

VII. Future Work

First, as alluded to above, we will incorporate SFB routers along with a form of ECN to provide Vegas with the necessary information to set α and β. This first step leads naturally to the next step of developing algorithms to integrate the information collected from routers. Next, we will address the issue of asymmetric networks as the two networks modeled in this paper — a reference network from [8] and a grid network that connects Los Alamos and Sandia National Laboratories — were symmetric. Finally, in parallel with the aforementioned research, we will also closely re-examine the results in the paper to better understand the subtle anomalies that appear in the results for the "fairness ratio" (i.e., ACK_R/ACK_V).

Acknowledgement

The authors would like to thank Mark Gardner for his meticulous efforts in ensuring that the theoretical aspects of this manuscript were concise and clear and for providing general feedback on the overall manuscript.

VIII. Conclusion

Prior research demonstrated the incompatibility of TCP Reno and TCP Vegas. In this paper, we showed that the incompatibility of Reno and Vegas is *not* inherent to their congestion-control algorithms but an artifact of misconfiguring Vegas's congestion-avoidance parameters.

In particular, we showed how inappropriate the default values of α and β in Vegas are (when in competition with Reno), explained the relationship of these parameters to variations in network performance, and demonstrated how to set the parameters appropriately so that Reno and Vegas are compatible. In summary, given that TCP is the ubiquitous transport protocol in the Internet, our results indicate that many could benefit from a more widespread adoption of TCP Vegas.

References

[1] V. Jacobson, "Congestion Avoidance and Control," *ACM Computer Communications Review*, vol. 18, no. 4, pp. 314–329, August 1988.

[2] V. Jacobson, "Modified TCP Congestion Avoidance Algorithm," Tech. Rep., Technical report, April 1990.

[3] L. Brakmo and L. Peterson, "TCP Vegas: End to End Congestion Avoidance on a Global Internet," *IEEE Journal on Selected Areas in Communication*, vol. 13, no. 8, pp. 1465–1480, October 1995.

[4] A. Veres and M. Boda, "The Chaotic Nature of TCP Congestion Control," in *Proc. of IEEE INFOCOM 2000*, March 2000.

[5] W. Feng and P. Tinnakornsrisuphap, "The Failure of TCP in High-Performance Computational Grids," in *Proc. of SC 2000: High-Performance Networking and Computing Conf.*, November 2000.

[6] Y. R. Yang, M. S. Kim, and S. S. Lam, "Transient Behaviors of TCP-friendly Congestion Control Protocols," ftp://ftp.cs.utexas.edu/pub/lam/transient_tech.ps.gz, July 2000.

[7] J. S. Ahn, P. B. Danzig, Z. Liu, and L. Yan, "Evaluation of TCP Vegas:Emulation and Experiment," in *Proc. of ACM SIGCOMM 1995*, August 1995, pp. 185–195.

[8] J. Mo, R. J. La, V. Anantharam, and J. Walrand, "Analysis and Comparison of TCP Reno and Vegas," in *Proc. of IEEE INFOCOM 1999*, March 1999, pp. 1556–1563.

[9] G. Hasegawa, M. Murata, and H. Miyahara, "Fairness and Stability of Congestion Control Mechanisms of TCP," in *Proc. of IEEE INFOCOM 1999*, March 1999, pp. 1329–1336.

[10] D. Chiu and R. Jain, "Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks," *Computer Networks and ISDN Systems*, vol. 17, pp. 1–14, 1989.

[11] G. Hasegawa, K. Kurata, and M. Murata, "Analysis and Improvement of Fairness between TCP Reno and Vegas for Deployment of TCP Vegas to the Internet," in *Proc. of IEEE ICNP2000*, November 2000.

[12] E. Weigle and W. Feng, "A Case for TCP Vegas in High-Performance Computational Grids," in *IEEE International Symposium on High Performance Distributed Computing (to appear)*, August 2001.

[13] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, August 1993.

[14] C. Boutremans and J. Le Boudec, "A Note on the Fairness of TCP Vegas," in *Proc. of International Zurich Seminar on Broadband Communications*, February 2000, pp. 163–170.

[15] ns, "Network Simulator, version 2.1b8a," http://www.isi.edu/nsnam/ns.

[16] I. Foster and C. Kesselman, Eds., *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers, 1999.

[17] S. Floyd and V. Jacobson, "Connection with Multiple Congested Gateways in Packet-Switched Networks, Part 1: One-Way Traffic," *ACM Computer Communications Review*, vol. 21, no. 5, pp. 30–47, August 1991.

[18] A. Mankin, "Random Drop Congestion Control," in *Proc. of ACM SIGCOMM 1990*, September 1990, pp. 1–7.

[19] W. Feng, D. Kandlur, D. Saha, and K. G. Shin, "Stochastic Fair Blue: A Queue Managment Algorithm for Enforcing Fairness," in *Proc. of IEEE INFOCOM 2001*, April 2001.

[20] B. Bloom, "Space/Time Trade-offs in Hash Coding with Allowable Errors," *Communications of the ACM*, vol. 13, no. 7, July 1970.

[21] K. Ramakrishnan and S. Floyd, "A Proposal to Add Explicit Congestion Notification (ECN) to IP," RFC 2481, January 1999.