Analysis of Scalable TCP*

Eitan Altman, Konstantin Avrachenkov, Chadi Barakat,
Arzad Alam Kherani, and B.J. Prabhu

INRIA Sophia-Antipolis, France
{altman,k.avrachenkov,cbarakat,alam,bprabhu}@sophia.inria.fr

Abstract. Scalable TCP [2] is a proposition for a new TCP where both the increase and the decrease rate of the window size are multiplicative. It has been recently proposed in order to improve performance in high speed networks. In this paper, we present a mathematical analysis of such multiplicative increase multiplicative decrease protocols in the presence of random losses. These are typical to wireless networks but can also model losses in wireline networks with very high bandwidth delay product. Our approach is based on showing that the logarithm of the window size evolution has the same behaviour as the workload process in a standard G/G/1 queue. The Laplace-Stieltjes transform of the equivalent queue is shown to provide directly the throughput of Scalable TCP as well as the higher moments of the window size. We validate our findings using *ns-2* simulations.

1 Introduction

In very high speed networks, the congestion avoidance phase of TCP takes a very long time to increase the window size and fully utilize the available bandwidth. Floyd writes in [1]: "for a Standard TCP connection with 1500-byte packets and a 100 ms round-trip time, achieving a steady-state throughput of 10 Gbps would require an average congestion window of 83,333 segments, and a packet drop rate of at most one congestion event every 5,000,000,000 packets (or equivalently, at most one congestion event every $1\frac{2}{3}$ hours). The average packet drop rate of at most 2×10^{-10} needed for full link utilization in this environment corresponds to a bit error rate of at most 2×10^{-14}, and this is an unrealistic requirement for current networks." Thus, in the context of high speed networks, it is essential to study the effect of random packet losses on TCP, which may limit the TCP throughput more than the congestion losses do and may lead to a poor utilization of the large available capacity. The modeling of random losses is also essential in the study of wireless channels. In order to better utilize the network capacity available in high speed networks, one should use new TCP protocols that are characterized by a faster rate of increase of the window size.

* This work was partially supported by the EURO NGI network of Excellence, by the INRIA's TCP ARC collaboration project and by the *Indo–French Center for Promotion of Advanced Research (IFCPAR)* under research contract number 2900–IT.

Z. Mammeri and P. Lorenz (Eds.): HSNMC 2004, LNCS 3079, pp. 51–62, 2004.

In [2], Kelly has proposed a new TCP wherein upon each ACK it receives, the sender increases its congestion window (*cwnd*) by 0.01. When a loss event is detected, the sender decreases *cwnd* by a factor of 0.125. Hence, if the window size is $W(t)$ at some time t (meaning that there are W unacknowledged packets in the network) then, in the absence of losses, the window size after an RTT (round-trip time), $W(t+RTT)$, would be $1.01 \times W(t)$, whereas if there are losses during $(t, t+RTT)$, $W(t+RTT)$ will be around $0.875 \times W(t)$ (here we assume that, as in New Reno and SACK, the window is reduced only once during a round trip time even if there are several losses). This proposal is called Scalable TCP because, starting from a window size of some fraction of the bandwidth-delay product (BDP), the number of $RTTs$ required to reach BDP becomes independent of the link bandwidth.

Consider the class of Multiplicative Increase and Multiplicative Decrease (MIMD) congestion control algorithms where each ACK results in a window increment of $\alpha - 1 > 0$ and a loss event is responded with a reduction of window size by a fraction $1 - \beta < 1$. The Scalable TCP can then be viewed as a special instance from this class with $\alpha = 1.01$ and $\beta = 0.875$. This motivates us to study the window behaviour of MIMD congestion control algorithms. In this paper, we focus on the analytical performance study of these algorithms, and, hence, of Scalable TCP, in the presence of random losses and congestion losses.

Our approach is based on showing that an invertible transformation applied to the window size process results in a process that has the same evolution as the total workload process in a standard G/G/1 queue. The Laplace-Stieltjes transform of the equivalent queueing process thus obtained provides the throughput of the connection as well as the moments of the window size of the given MIMD protocol (Section 3). We study both the case in which there are only random losses (Section 4) as well as the case where in addition to random losses, there are either congestion losses or the window size is limitted (Section 5). We validate our findings using *ns-2* simulations (Section 6).

2 Discrete Time Models

We consider the scenario where a single FTP application transfers data using an MIMD flow control protocol with parameters α and β as mentioned in the Introduction. We assume that the file is sufficiently large to ensure the convergence to a stationary regime. In this section, we introduce different models of MIMD schemes for different network conditions.

Let τ denote the round-trip propagation delay of the high speed link (in literature this is also referred to as the fixed part of the round-trip time). Let the link capacity be c packets per second. Let $\{W(t), t \geq 0\}$ denote the window process evolving over time. Use $W_0 \stackrel{\Delta}{=} W(0)$ and let $\tau_1 \stackrel{\Delta}{=} \tau + (\frac{W_0}{c} - \tau)^+$ denote the first round-trip time. Let $W_1 \stackrel{\Delta}{=} W(\tau_1)$ and define $\tau_2 = \tau + (\frac{W_1}{c} - \tau)^+$. Proceeding in this manner, we get a sequence $\{\tau_n, n \geq 1\}$ of round-trip times and a sequence $\{W_n, n \geq 1\}$ of window sizes. Consider a sequence of time instants $\{t_n, n \geq 1\}$ where t_n is the end of n^{th} round-trip time, i.e., $t_1 = \tau_1$ and $t_n = t_{n-1} + \tau_n$. Under our definition W_n is the window size at time instant t_n. The window evolution can now be written as

$$W_{n+1} = \begin{cases} \alpha W_n, \text{ if there was no loss in interval } [t_n, t_{n+1}], \\ \beta W_n, \text{ if there were one or more losses in } [t_n, t_{n+1}]. \end{cases}$$

We shall consider the following models for the evolution of $\{W_n\}$ under random losses:

(i) There is no upper bound on the window size.
(ii) There is an upper bound B on the window size which corresponds to an explicit limitation of the window size. When this value is reached then the window stops growing.
(iii) There is an upper bound B on the window size. However, when this value is reached, the connection suffers a congestion loss (this is in addition to the random losses) and the multiplicative decrease of window is invoked.

The first model here approximates the case where the link BDP is very high and there is a significant probability of loss in a round-trip time so that the practical upper bound of BDP is reached with negligible probability. The second model above corresponds to the case where the window is bounded by the receiver's advertised window. The last model corresponds to the case where the window reaches the value of round-trip pipe size (BDP+Buffer) and suffers a loss owing to buffer overflows.

2.1 Window Evolution for the Proposed Models

Let $A_n, n \geq 1$, be a random variable such that $A_n = \alpha$ if there was no loss in the interval $[t_n, t_{n+1}]$, and $A_n = \beta$ otherwise. Throughout this section, $\{A_n, n \geq 1\}$ will be assumed to be a general stationary ergodic sequence. Now we describe the evolution of $\{W_n\}$ in terms of $\{A_n\}$ recursively for the models described above.

Model *(i)* Taking into account the fact that the window size of TCP is bounded below by a value of one packet[1], the window size evolution for this model can be written as

$$W_{n+1} = \max(A_n W_n, 1). \tag{1}$$

As mentioned earlier, this model can be expected to be useful in the presence of a very large maximum value B of the receiver advertised window size or of the bandwidth (that would result in a congestion loss if it were attained) provided that losses are sufficiently frequent so that the window level B is rarely reached. This will be made precise later.

Model *(ii)* If on the contrary, losses are infrequent and B is often reached and is sufficiently large then we can ignore the lower bound of one packet on the

[1] There is no loss of generality, as one can consider any value of the minimal window size and then rescale the model.

window size (which would rarely be attained). This is the case for the second model described before. The window evolution for this model can thus be written as

$$W_{n+1} = \min(A_n W_n, B). \tag{2}$$

Model (iii) The window evolution in this model is similar to that of model (ii). However, there is an instantaneous drop in the window upon reaching B.

In the next section we relate the window process under the different models introduced in this section to the workload evolution in a discrete time G/G/1 queue.

3 Preliminary Analysis

Taking the logarithm of equation (1) and defining $\xi = \frac{1}{\log[\alpha]}$, $Y_n = \xi \log[W_n]$, $k = -\xi \log[\beta] > 0$ and $D_n = \xi \log[A_n] + k \geq 0$, we obtain

$$Y_{n+1} = \max(D_n - k + Y_n, 0). \tag{3}$$

We now make the following important observation: The recursive equation (3) has the same form as the equation describing the workload process in a G/G/1 queue (see, for example, [5]) with D_n denoting the work arriving in n^{th} slot and k denoting the amount of service that can take place in one time slot. Since the introduced transformation $\xi \log(\cdot)$ is invertible, there is a one to one correspondence between the processes $\{Y_n, n \geq 1\}$ and $\{W_n, n \geq 1\}$. This observation allows us to study the stability of the window process $\{W_n, n \geq 1\}$ via that of $\{Y_n, n \geq 1\}$. Furthermore, the analogy with queueing theory of the process $\{Y_n, n \geq 1\}$ will allow us to obtain the steady state moments of the window size and of the throughput of TCP under some further statistical assumptions.

Theorem 1. *Assume that $E[\log A_0] < 0$. Then there exists a unique stationary ergodic process W_n^* that satisfies the recursion (1) defined on the same probability space as W_n; moreover, for any initial value $W_0 = w$, there is a random time T_w, which is finite with probability 1, such that $W_n = W_n^*$ for all $n \geq T_w$. If $E[\log A_0] > 0$ then W_n tends to infinity w.p.1 for any initial value $W_0 = w$.*

Proof. According to Theorem 2A [3], if $E[\log A_0] < 0$ then the stochastic process $\{Y_n\}$ converges to a stationary ergodic process $\{Y_n^*\}$ which is defined on the same probability space as $\{Y_n\}$ and is the unique stationary regime that satisfies (3). This implies the statement for $W_n = \exp(Y_n \log[\alpha])$. The last part of theorem similarly follows from [4, p. 36]. □

Remark 1. Due to Jensen inequality and the concavity of the logarithmic function, $E[\log A_0] \leq \log E[A_0]$. Hence $\log E[A_0] < 0$, or equivalently $E[A_0] < 1$ is a sufficient condition for the stability of the window process $\{W_n\}$ (for the existence of and convergence to a unique stationary ergodic regime). However this condition is in general not a necessary one.

Remark 2. We stress the importance of the minimum window size in model (1). Indeed, if we eliminate it and write simply $W_{n+1} = A_n W_n$ then when taking the log, we get in (3) instead $Y_{n+1} = \log[A_n]/\log[\alpha] + Y_n$. Its solution is

$$Y_n = Y_0 + \sum_{i=0}^{n-1} \log[A_i]/\log[\alpha].$$

Since $\{A_n\}$ is stationary ergodic, the strong law of large numbers implies that if $E\log[A_i] < 0$ then Y_n converges to $-\infty$ and thus W_n converges to 0 which is clearly a bad estimation for the window size process. (If $E\log[A_i] > 0$ then Y_n, and thus W_n converges to ∞ which was also predicted by the model that took the the minimum window into account.) Note that in the limiting case of $E\log[A_i] = 0$, if A_i's are independent and identically distributed (i.i.d.) then Y_n is a null recurrent Markov chain and thus unstable.

Next, we compute the moments of the window size distribution in the stationary regime. First define the Laplace-Stieltjes transform of Y_n at the stationary regime (i.e. of Y_n^*) as follows: $G(s) = E\left(e^{sY_n^*}\right)$. Then we have, for any integer $k \geq 1$,

$$E[(W_n^*)^k] = E[\exp(k\log[\alpha]Y_n^*)] = G(k\log[\alpha]).$$

Thus, since $\log[\alpha] > 0$, all moments of W_n^* are obtained from the Laplace-Stieltjes transform of Y_n^*, which we compute in the following sections. We note that the z-transform, which is defined for integer valued random variables, is a special case of the Laplace-Stieltjes transform.

With the analogy to the queueing system, we can now recommend using (1) if $E(\log[A_n])$ is much smaller than 0, and using (2) if $E(\log[A_n])$ is much larger than 0. In the next section we study model (i).

We note, however, that models (ii) and (iii) can also be solved using a transformation to an equivalent queueing problem with infinite buffer by considering the variable $Z_n = \log[B]/\log[\alpha] - Y_n$ instead of working directly with Y_n. The throughput of the connection and the moments of the window size will then be obtained as in model (i).

In the next section, we derive explicit expression of the stationary distribution of the transformed process $\{Y_n\}$ for the case where $\{A_n\}$ is a sequence of i.i.d. random variables. As we argued in this section, such an analysis provides the stationary distribution for the window size process $\{W_n\}$ for the three models.

4 MIMD Protocols with Only Random Losses

Assume that the sequence A_n is i.i.d. with the following distribution

$$A_n = \begin{cases} \alpha \ w.p. \ 1-p, \\ \beta \ w.p. \ p, \end{cases}$$

where p is the loss rate observed by the connection (the probability that a random loss occurs in a round-trip time). Recall the notation k $(= -\frac{\log(\beta)}{\log(\alpha)})$ for the service in one time slot for the equivalent discrete time queueing model of equation (3). For this model to be stable, the necessary and sufficient condition is $(k+1)p > 1$. We assume that k is an integer. This assumption allows us to use a discrete state space, $\mathcal{S} = \{0, \log(\alpha), ..., n \cdot \log(\alpha), ..\}$. Then the recursive equation for process $\{Y_n\}$ is given by (from equation (3))

$$Y_{n+1} = \begin{cases} Y_n + 1 & w.p.\ 1-p, \\ (Y_n - k)^+ & w.p.\ p. \end{cases}$$

Let $P_n(j) \stackrel{\Delta}{=} P(Y_n = j)$. Then

$$\begin{aligned} P_{n+1}(j) &= (1-p)P_n(j-1) + pP_n(j+k), \; j \geq 1 \\ &= p\textstyle\sum_{i=0}^{k} P_n(i), \qquad\qquad\qquad\quad j = 0. \end{aligned} \tag{4}$$

Denote the z-transform of Y_n by $\mathbf{Y}_n(z)$, then

$$\begin{aligned} \mathbf{Y}_{n+1}(z) - P_{n+1}(0) &= \sum_{j=1}^{\infty} P_{n+1}(j) z^j \\ &= (1-p)\sum_{j=1}^{\infty} P_n(j-1)z^j + p\sum_{j=1}^{\infty} P_n(j+k)z^j. \end{aligned}$$

Assuming that $\mathbf{Y}_n$ converges to $\mathbf{Y}$ and that $P_n(\cdot)$ converges to $P(\cdot)$,

$$\mathbf{Y}(z) - P(0) = (1-p)z\mathbf{Y}(z) + pz^{-k}\sum_{j=0}^{\infty} P_n(j+1+k)z^{j+1+k}$$

$$\begin{aligned} \mathbf{Y}(z)(1-(1-p)z) &= P(0) + pz^{-k}\sum_{j=0}^{\infty} P_n(j+1+k)z^{j+1+k} \\ &= P(0) + pz^{-k}(\mathbf{Y}(z) - \sum_{i=0}^{k} P(i)z^i) \end{aligned}$$

and hence

$$\mathbf{Y}(z)((1-(1-p)z)z^k - p) = z^k P(0) - p\sum_{i=0}^{k} P(i)z^i.$$

Since $P(0) = p\sum_{i=0}^{k} P_n(i)$, $\mathbf{Y}(z)$ can be expressed as

$$\mathbf{Y}(z) = \frac{\sum_{i=0}^{k-1} P(i)(z^k - z^i)}{-\frac{(1-p)}{p}z^{k+1} + \frac{1}{p}z^k - 1}. \tag{5}$$

If the z-transform, $\mathbf{Y}(z)$, exists, it is analytic in the open disc $\{z : |z| < 1\}$. The numerator of equation (5) has at most $k-1$ zeros inside the unit circle and one zero on the unit circle. Hence, there can be at most $k-1$ zeros of the denominator of equation (5) within the unit circle as any more zeros will make $\mathbf{Y}(z)$ non-analytic. Using Rouche's theorem [7] we can show that there are at least k zeros of the denominator inside and on the unit circle. As $z = 1$ is a zero of the denominator, there are at least $k-1$ zeros inside the unit circle. From the two previous arguments, there are exactly $k-1$ zeros of the denominator within the unit circle and they must be the same as those of the numerator for $\mathbf{Y}(z)$ to be analytic [5]. Hence, $Y(z)$ reduces to

$$\mathbf{Y}(z) = \frac{1 - 1/z_0}{1 - z/z_0}, \tag{6}$$

where z_0 is the root of

$$\frac{(1-p)}{p} z^{k+1} - \frac{1}{p} z^k + 1 = 0 \tag{7}$$

that lies outside the closed unit disc. The distribution of Y is then obtained as,

$$P(Y = j) = (1 - 1/z_0)(1/z_0)^j, \quad j \geq 0. \tag{8}$$

Since $W = \exp(Y \log[\alpha])$, it follows from the above that for $w \geq 1$,

$$P(W > w) = P(Y > \log[w]/\log[\alpha]) = (1/z_0)^{\log[w]/\log[\alpha]} = w^{-\log[z_0]/\log[\alpha]}.$$

Thus, we obtain the distribution of the stationary window size process. In order to compute the moments of W, we note that $W = \alpha^Y$. Hence,

$$E[W^n] = E[\alpha^{Yn}] = \mathbf{Y}(\alpha^n)$$

The z-transform $\mathbf{Y}(z)$ is analytic for $z < 2z_0$. Hence, the n^{th} moment of W is finite if $n < \frac{\log[z_0]}{\log[\alpha]}$. Let $a = \frac{\log[z_0]}{\log[\alpha]}$. The window size distribution can be seen to become heavy tailed for $a \leq 2$. Thus, for a given loss rate, p, either α or β can be suitably chosen in order to reduce the variance of the window size.

5 MIMD Protocols with Limit on the Window Size or with Congestion Losses

In this section we consider the discrete model where the window at the sender is limited by the receiver window size, B. We make the following transformation

$$Y_n = \frac{\log[B] - \log[W_n]}{\log[\alpha]}. \tag{9}$$

We assume that $L = \frac{\log[B]}{\log[\alpha]}$ is an integer.

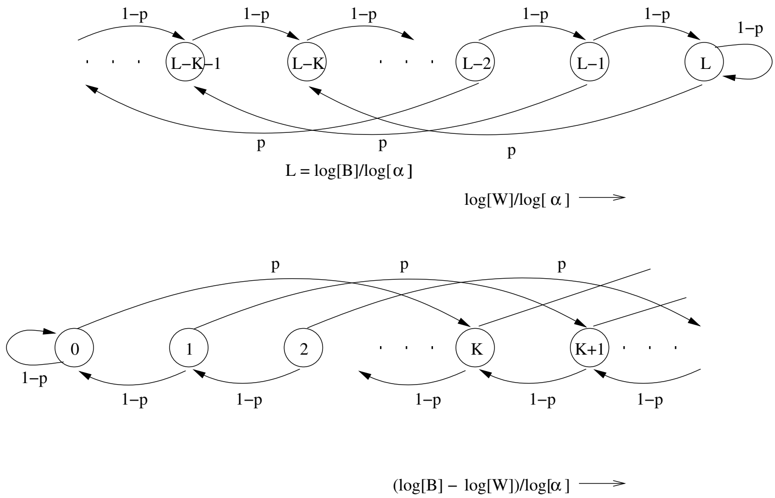

Fig. 1. Evolution, at the end n^{th} RTT, of $\frac{\log[W_n]}{\log[\alpha]}$ and Y_n.

5.1 Model (ii)

The evolution of $\frac{\log[W_n]}{\log[\alpha]}$ and of Y_n is shown in Fig. 1. The balance equations for Y_n can be written as

$$P(0) = \frac{(1-p)}{p} P(1),$$
$$P(i) = (1-p)P(i+1), \quad i = 1, ..., k-1.$$
$$P(i) = pP(i-k) + (1-p)P(i+1), \quad i \geq k.$$

This is similar to a bulk arrival queue. Hence, following similar arguments from Kleinrock [5], we can write the z-transform as,

$$\sum_{i=1}^{\infty} P(i)z^i = \sum_{i=1}^{\infty} pP(i-k)z^i + \sum_{i=1}^{\infty}(1-p)P(i+1)z^i,$$

where $P(i-k) = 0$ for $i < k$.

$$\begin{aligned}\mathbf{Y}(z) - P(0) &= pz^k \sum_{i=1}^{\infty} P(i-k)z^{i-k} + z^{-1}\sum_{i=1}^{\infty}(1-p)P(i+1)z^{i+1} \\ &= pz^k \mathbf{Y}(z) + z^{-1}(1-p)(\mathbf{Y}(z) - zP(1) - P(0)),\end{aligned}$$

which gives

$$\mathbf{Y}(z)(z - pz^{k+1} - (1-p)) = P(0)(z - (1-p) - zp)).$$

Using $\mathbf{Y}(1) = 1$ and the L'Hôpital's rule we get $P(0) = [1 - (k+1)p]/[1-p]$. Hence, we obtain $\mathbf{Y}(z)$ as,

$$\mathbf{Y}(z) = (1 - (k+1)p)\frac{1-z}{pz^{k+1} - z + (1-p)}. \tag{10}$$

For this model to be stable, the necessary and sufficient condition is $(k+1)p < 1$. The distribution of Y can be found by inverting the $\mathbf{Y}(z)$ using partial fraction expansion. The distribution can be seen to be a weighted sum of geometric distributions. The moments of $W = B\alpha^{-Y}$ can be found as

$$E[W^n] = E[(B\alpha^{-Y})^n)] = B^n E[\alpha^{-nY}] = B^n \mathbf{Y}(\alpha^{-n}).$$

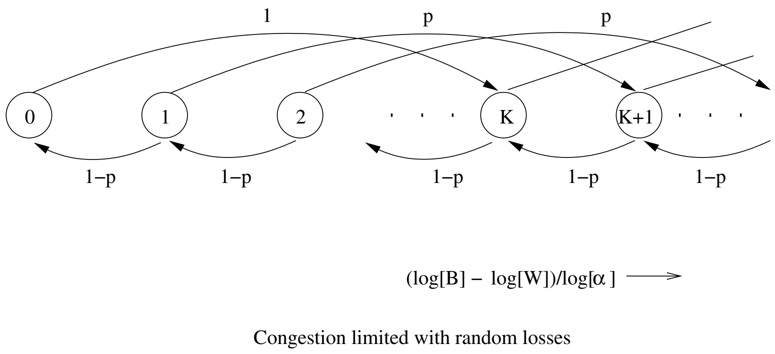

Fig. 2. Evolution, at the end n^{th} RTT, of Y_n.

5.2 Model (iii)

In this model along with random losses, a loss is detected when the window size reaches B. The evolution of Y_n is shown in Fig. 2. The evolution at $Y = 0$ is different from the previous model. Here, there is a jump with probability 1 to state k. The modified balance equations can be written as,

$$P(i) = (1-p)P(i+1), \quad i = 0, ..., k-1.$$
$$P(i) = pP(i-k) + (1-p)P(i+1) + (1-p)P(0)\delta_{i-k}, \quad i \geq k.$$

As before we can write the z-transform, $\mathbf{Y}(z)$, as

$$\sum_{i=1}^{\infty} P(i)z^i = pz^k \sum_{i=1}^{\infty} P(i-k)z^{i-k} + \frac{1}{z}\sum_{i=1}^{\infty}(1-p)P(i+1)z^{i+1} + z^k(1-p)P(0),$$

which implies

$$\mathbf{Y}(z) - P(0) = pz^k\mathbf{Y}(z) + z^{-1}(1-p)(\mathbf{Y}(z) - zP(1) - P(0)) + z^k(1-p)P(0),$$

and gives the relation

$$\mathbf{Y}(z)(z - pz^{k+1} - (1-p)) = P(0)(1-p)(z^{k+1} - 1).$$

Using $\mathbf{Y}(1) = 1$ and L'Hôpital's rule we get $P(0) = [1 - (k+1)p]/[(k+1)(1 - p)]$, and hence, we obtain $\mathbf{Y}(z)$ as

$$\mathbf{Y}(z) = \left(\frac{1-(k+1)p}{(k+1)}\right) \frac{1 - z^{k+1}}{pz^{k+1} - z + (1-p)}. \tag{11}$$

As before, we can obtain the of distribution of Y, and hence W, by inverting the z-transform. We can also obtain the moments of W directly from $\mathbf{Y}(z)$.

6 Simulation Results

We perform simulations using *ns-2*[8] to validate our model. The simulation setup has a source and a destination node. The source node has infinite amount of data to send and uses Scalable TCP with New Reno flavor. The link bandwidth is 150Mbps and the RTT is 120ms. The window at the source is limited to 500 packets to emulate the receiver advertised window. The BDP for this system is approximately 2250 packets (packet size is 1040 bytes). In the Scalable TCP we have implemented, the following assumptions are made:

- The minimum window size, MIN_W, is 8. The growth rate of Scalable TCP is very small for small window sizes. It has been recommended in [2] to use the Scalable algorithm after a certain threshold.
- There is no separate slow start phase since slow start can be viewed as a multiplicative increase algorithm with $\alpha = 2$.
- For each positive ACK received, the window is increased by $\alpha - 1$ packets. When a loss is detected, the window is reduced by a factor of β. α is taken as 1.01 and β is taken as 0.86. This value of β is taken so that $k = -\frac{\log[\beta]]}{\log[\alpha]} \approx 15$. These values are chosen so as to be close to the values recommended in [2] ($\alpha = 1.01$, $\beta = 0.875$).

The expression for the density function of W, $f(W)$, modified for the minimum window at 8 is given by

$$f(w) = \frac{a}{8}\left(\frac{w}{8}\right)^{-(a+1)}, \tag{12}$$

where $a = \frac{\log[z_0]}{\log[\alpha]}$. In the simulations, the density function of W is obtained by sampling the window at an interval of $RTT = 0.12s$. We would like to note that in the present setting RTT is very close to the propagation delay, and hence, does not vary much.

Figures 3 and 4 show the PDF of W for two different values of loss rate, p. Simulation results are observed to match well with the analysis (Eqn.(12)). Depending on the value of the root, z_0, of Eqn.(7), the distribution can be seen to become heavy tailed. For example, for $p = 0.07$, the tail decreases at rate 1.55 indicating the heavy tailed nature of the window size. In the models which

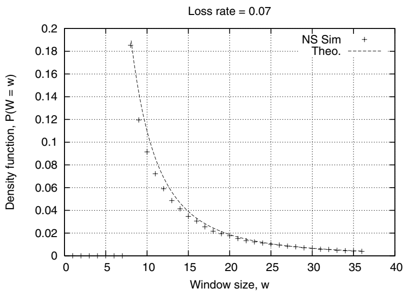

Fig. 3. Density Function of the Window size, W. $a = 1.55$.

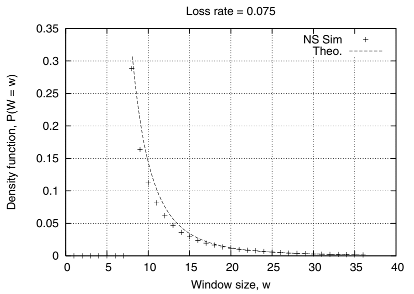

Fig. 4. Density Function of the Window size, W. $a = 2.53$.

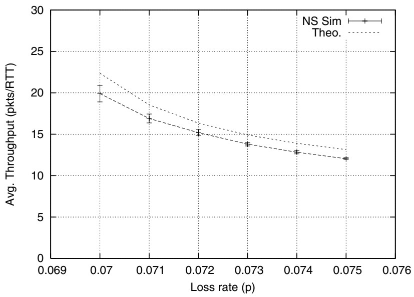

Fig. 5. Throughput (pkts/RTT) versus Loss rate, p.

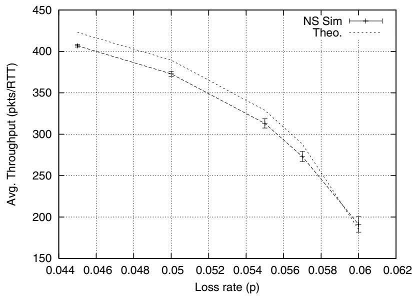

Fig. 6. Throughput (pkts/RTT) versus Loss rate, p.

we considered, the window size was assumed to take rational values. In practice, however, the window size takes only integer values. For example, when the window size is 8.5, the sender sends 8 packets. The density for the window size through simulations is, therefore, defined only at integer values whereas the theoretical plot is shown for real values. This results in a small discrepancy between the simulations and the theoretical function. Figure 5 shows the throughput in (TCP packets)/RTT as a function of the loss rate, p. The error bars are the 99% confidence intervals.

Figure 6 shows the throughput in (TCP packets)/RTT as a function of the loss rate, p, for the model in which the maximum window at the sender is limited by the receiver's advertised window. The receiver buffer is assumed to be limited to 500 packets. The error bars are the 99% confidence intervals. A good match is observed between the simulations and the analysis.

7 Conclusions

We presented a mathematical model and analysis for computing the moments of the window size and, in particular, the throughput of a single connection using

Scalable TCP in the presence of random losses. In the first model, we analyzed the scenario where the loss rate is high so that the window size would return to minimum window infinitely often. In the second model, we considered the scenario wherein the sender's window was bounded by the receiver buffer and the connection was subject to random losses with a low rate. In the third model, congestion losses were considered in addition to the random losses. The moments of the window size were shown to be equivalent to evaluating the Laplace-Stieltjes transform of the log of the window evolution process. The log of the window size was observed to be equivalent to the number of customers in a discrete time queue. The simulations were seen to match well with the analysis.

References

1. S. Floyd, "HighSpeed TCP for Large Congestion Windows", RFC 3649, Experimental, December 2003. Available at www.icir.org/floyd/hstcp.html
2. T. Kelly, "Scalable TCP: Improving Performance in Highspeed Wide Area Networks," Submitted for publication, December 2002. Available at `http://www-lce.eng.cam.ac.uk/~ctk21/scalable/`
3. P. Glasserman and D.D. Yao, "Stochastic vector difference equations with stationary coefficients", *J. Appl. Prob.*, Vol 32, pp 851-866, 1995.
4. F. Baccelli and P. Bremaud, *Palm probabilities and stationary queues*, Springer, 1987.
5. L. Kleinrock, *Queueing Systems Volume I: Theory*, Wiley & sons, 1975.
6. F. Dufour and O.L.V. Costa: "Stability of Piecewise Deterministic Markov Processes", *SIAM Journal on Control and Optimization*, vol 37 (5), 1483-1502, 1999.
7. S. G. Krantz, "Rouche's Theorem", in *Handbook of Complex Variables*, Birkhauser, 1999.
8. S. McCanne and S. Floyd. *ns Network Simulator*. `http://www.isi.edu/nsnam/ns/`