Analysis of a TCP System under Polling-Type Reduction-Signal Procedures

Omer Czerniak
Department of Statistics and Operations Research
School of Mathematical Sciences, Tel Aviv University, Israel
omercz@gmail.com

Eitan Altman
INRIA
Sophia Antipolis, France
altman@sophia.inria.fr

Uri Yechiali
Department of Statistics and Operations Research
School of Mathematical Sciences, Tel Aviv University, Israel
uriy@post.tau.ac.il

ABSTRACT

The performance of a Transmission Control Protocol (TCP) for a system with N connections sharing a common Active Queue Management (AQM) is analyzed for both Additive-Increase Multiplicative-Decrease (AIMD) and Multiplicative-Increase Multiplicative-Decrease (MIMD) control mechanisms, where reduction signals follow either a cyclic or a probabilistic polling-type procedure. The Laplace-Stieltjes Transforms (LST) of the transmission rate of each connection at a polling instant, as well as at an arbitrary moment, are derived. *Explicit* results are calculated for the mean rate and (in contrast to most polling models) for its second moment. The analysis of the probabilistic MIMD models uses transformations yielding a system's law of motion equivalent to that of an M/G/1 queue with bulk service.

Keywords

TCP, AIMD, MIMD, Cyclic Polling, Probabilistic Polling, M/G/1 Bulk Service

1. INTRODUCTION

We analyze the performance of TCP, the widely-used transmission protocol of the Internet [9]. TCP is a reliable window-based flow control protocol where the window is increased until a packet loss is detected. TCP modeling has been studied extensively in the literature (see, e.g., [1], [2], [8] and references there). Many authors have been interested in the performance of several parallel TCP connections, see [4], [7], [12]. In some cases the parallel connections may correspond to one single transfer which is split into several connections. We assume that the connections are subject to loss events triggered by congestion caused by exogenous traffic. We assume that the losses are independent of the rate of the connections. This has been validated through measurements in [2] on long connections (i.e., having long round trip time), but turns out not to be realistic for short connections. Once a congestion event occurs, the AQM reacts to it by marking or dropping a packet of some connection. When using AQM, one can fully control which connection will be the one to loose a packet when a congestion event occurs. We shall propose and analyze two different policies for assigning the losses to TCP connections.

Consider N connections sharing a common AQM. Each of the N connections increases its transmission rate until it gets a congestion signal. A source (connection) receiving a congestion signal reduces instantaneously its rate, and then resumes increasing it. The other sources continue in increasing their transmission rates. This continues until the next congestion signaling event. Thus, each connection has two modes of operation: One during which the transmission rate grows, and one where it is reduced. Upon the receipt of a reduction signal at time t, the source that receives the signal reduces its sending rate X_t to βX_t, where $0 < \beta < 1$ is a constant. Such a reduction is termed *Multiplicative Decrease*. In absence of marking, each connection increases its sending rate. We distinguish between two methods of rate increase (i) Additive Increase, such that at time $s > t$ the transmission rate is $X_s = X_t + \alpha(s-t)$, where $\alpha > 0$ is a constant, and (ii) Multiplicative Increase, where at time $s > t$, $X_s = X_t \cdot e^{\gamma(s-t)}$, where $\gamma > 0$ is a constant. We thus have two transmission methods: (i) Additive Increase Multiplicative Decrease (AIMD), and (ii) Multiplicative Increase Multiplicative Decrease (MIMD). We assume that the marking process does not depend on the transmission rates of the sources. We introduce two signaling strategies, which determine the choice of connection that has to reduce its transmission rate: (i) The "Cyclic One Strategy" where the order of connections to which the signals are sent is cyclic, $1, 2, \ldots, N-1, N, 1, 2, \ldots$ and (ii) The "Probabilistic One Strategy" where the choice of the connection to decrease its rate is done probabilistically, where after reducing connection i, the next connection to be chosen is j with probability p_j, $\sum_{j=1}^{N} p_j = 1$. We analyze the different TCP systems using polling systems methods. Polling systems, in which a single server visits (according to some scheduling procedure) and serves (according to some service discipline) N separate queues, have been studied extensively in the literature ([10], [13], [6] and references there). In this paper the stationary behavior of the system is analyzed. In our model, TCP is not represented at packet level, but rather via direct fluid equations that describe the transmission rates for the set of

VALUETOOLS 2009, October 20-22, Pisa, Italy

DOI 10.4108/ICST.VALUETOOLS2009.7725

connections.

The paper is structured as follows. In section 2, the AIMD mechanism is analyzed for both the cyclic one (section 2.1) and the probabilistic one (section 2.2) polling strategies. The MIMD mechanism is tackled in section 3 where the probabilistic polling scheme is studied in section 3.1 and the cyclic one is examined in section 3.2. Similarities between the last two polling schemes are drawn.

Notation: For a continuous random variable X, we denote its mean by $E[X] = x$, its second moment by $E[X^2] = x^{(2)}$, and its LST by $\tilde{X}(\cdot)$.

2. AIMD

2.1 Cyclic One Strategy

Under a "Cyclic One Strategy" system signals occur randomly in time and are directed in a cyclic manner between the connections: $1, 2, \ldots, N$. We call an instant where a reduction signal occurs a "polling instant", we refer to a cycle of connection i as the time since connection i was polled until its next polling instant. Let X_i^j denote the transmission rate at connection j ($j = 1, 2, \ldots, N$) at the instant when the server decides to reduce the transmission rate at connection i ($i = 1, 2, \ldots, N$). $X_i = (X_i^1, X_i^2, \ldots, X_i^N)$ is the state of the system at that instant. Let the random variable U_i denote the time between the instant of the signal that causes the server to reduce the transmission rate of connection i and the one that causes the $i+1$st connection to reduce its transmission rate. All U_i's are independent, identically distributed as a generic random variable U. The transmission rate of connection i is continuously growing at a rate α_i. When the server polls connection i, the transmission rate decreases by a factor of β_i. Thus, the evolution of the stationary transmission rates of the system at a "polling instant" is given by

$$X_{i+1}^j = \begin{cases} X_i^j + \alpha_j U_i & \text{if } j \neq i \\ \beta_i X_i^i + \alpha_i U_i & \text{if } j = i. \end{cases} \tag{1}$$

That is, the transmission rate of connection j at a polling instant of connection $i+1$ ($j \neq i$) is composed of: (i) the transmission rate of connection j at a polling instant of connection i, and (ii) the growth of the transmission rate at that connection during the time between the two signals. In the case where $i = j$ the transmission rate of connection i is composed as before, except that the transmission rate of connection i just after its polling instant is $\beta_i X_i^i$.

Define the multidimensional LST $L_i(\underline{\theta})$ of the state of the system at a polling instant of connection i ($i = 1, 2, \ldots, N$). This transform is given by

$$L_i(\underline{\theta}) = L_i(\theta_1, \ldots, \theta_{i-1}, \theta_i, \theta_{i+1}, \ldots, \theta_N) = E[e^{-\sum_{j=1}^N \theta_j X_i^j}]. \tag{2}$$

Then, for $i = 1, 2, \ldots, N$, using the fact that U_i and X_i^j are independent, we obtain $L_{i+1}(\underline{\theta})$ in terms of $L_i(\cdot)$, namely, for $i = 1, 2, \ldots, N$,

$$\begin{aligned} L_{i+1}(\underline{\theta}) &= E[e^{-\sum_{j=1}^N \theta_j X_{i+1}^j}] \\ &= E[e^{-\sum_{j=1}^N \theta_j \alpha_j U_i}] E[e^{-\sum_{j=1, j\neq i}^N \theta_j X_i^j} e^{-\theta_i \beta_i X_i^i}] \\ &= L_i(\theta_1, \ldots, \theta_{i-1}, \beta_i \theta_i, \theta_{i+1}, \ldots, \theta_N) \cdot \tilde{U}(\sum_{j=1}^N \theta_j \alpha_j). \end{aligned} \tag{3}$$

Equations (3) are now used to derive moments of the variables X_i^j.

Transmission Rate at Reduction Instants: Moments

The mean transmission rate, $f_i(j) \triangleq E[X_i^j]$, at connection j when the server polls connection i is given by

$$f_i(j) \triangleq E[X_i^j] = -\frac{\partial L_i(\underline{\theta})}{\partial \theta_j}\Big|_{\underline{\theta}=\underline{0}}. \tag{4}$$

This leads to the following N^2 linear equations,

$$f_{i+1}(j) = \begin{cases} f_i(j) + \alpha_j u & \text{if } j \neq i \\ \beta_i f_i(i) + \alpha_i u & \text{if } j = i. \end{cases} \tag{5}$$

Clearly, equation (5) can also be obtained directly by taking expectation over (1).
The solution of (5) is given by

$$f_i(j) = \begin{cases} \frac{\alpha_j N u}{1-\beta_j} - \alpha_j (j-i) u & j > i \\ \frac{\alpha_i N u}{1-\beta_i} & j = i \\ \frac{\alpha_j N u}{1-\beta_j} - \alpha_j (N - (i-j)) u & j < i. \end{cases} \tag{6}$$

Denoting by C the cycle time, the explanation of (6) is as follows: since in stationary state, $f_i(i) = \beta_i f_i(i) + \alpha_i E[C]$, where clearly $E[C] = Nu$, then

$$f_i(i) = \frac{\alpha_i N u}{1 - \beta_i}. \tag{7}$$

Regarding the case where $j > i$, $f_i(j)$ equals $f_j(j) - (j-i)\alpha_j u$ since the mean time until the next polling of connection j is $(j-i)u$, and during that time connection j increases its rate by $\alpha_j (j-i) u$ to the value of $f_j(j)$ (the case where $j < i$ is explained in the same manner).
The second and mixed moments of the X_i^j are given by

$$f_i(j,k) \triangleq E[X_i^j X_i^k] = \frac{\partial^2 L_i(\underline{\theta})}{\partial \theta_j \partial \theta_k}\Big|_{\underline{\theta}=\underline{0}}. \tag{8}$$

Differentiating (3) with respect to θ_j and θ_k, we get the following N^3 linear equations,

$$f_{i+1}(j,k) = \alpha_j u f_i(k) + \alpha_j \alpha_k u^{(2)} + \alpha_k u f_i(j) + f_i(j,k) \quad k, j \neq i, \tag{9}$$

$$f_{i+1}(i,j) = \alpha_j \beta_i u f_i(i) + \alpha_j \alpha_i u^{(2)} + \alpha_i u f_i(j) + \beta_i f_i(i,j) \quad j \neq i, \tag{10}$$

$$f_{i+1}(i,i) = \alpha_i \beta_i u f_i(i) + \alpha_i^2 u^{(2)} + \alpha_i \beta_i u f_i(i) + \beta_i^2 f_i(i,i). \tag{11}$$

As opposed to most gated and the exhaustive polling regimes (see, e.g., [10] and [6]), computing the second moment of X_i^i can be done explicitly since the cycle time is independent of the transmission rate at any of these connections. Define $X_i^{i(k)}$ as the transmission rate of connection i at the kth cycle, then we have

$$X_i^{i(k+1)} = \beta_i X_i^{i(k)} + \alpha_i C, \tag{12}$$

where $C = \sum_{j=1}^N U_i$, meaning that at the beginning of a cycle, $X_i^{i(k)}$ is reduced by a factor of β_i and then it grows

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7725
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7725

linearly at a rate of α_i. Define $\tilde{X}_i^{(k)}(s) = E[e^{-sX_i^{i(k)}}]$. As $X_i^{i(k)}$ and C are independent

$$\begin{aligned}\tilde{X}_i^{(k+1)}(s) &= E[e^{-sX_i^{i(k+1)}}] = E[e^{-s(\beta_i X_i^{i(k)}+\alpha_i C)}] \\ &= \tilde{X}_i^{(k)}(\beta_i s)\tilde{C}(\alpha_i s).\end{aligned} \tag{13}$$

By iterating we have

$$\tilde{X}_i^{(k+1)}(s) = \tilde{X}_i^{(1)}(\beta_i^k s)\prod_{j=0}^{k}\tilde{C}(\alpha_i\beta_i^j s). \tag{14}$$

When $k \to \infty$ we have $\lim_{k\to\infty}\tilde{X}_i^{(1)}(\beta_i^k s) = 1$ then

$$\tilde{X}_i(s) = \prod_{j=0}^{\infty}\tilde{C}(\alpha_i\beta_i^j s). \tag{15}$$

From (15) (and the fact that $E[C^2] = Nu^{(2)} + N(N-1)u^2$) we get

$$E[(X_i^i)^2] = \frac{1}{1-\beta_i^2}\left[\frac{2\beta_i\alpha_i^2N^2u^2}{1-\beta_i} + \alpha_i^2(Nu^{(2)} + N(N-1)u^2)\right]. \tag{16}$$

Throughput of Connection i
Let L_i be the transmission rate at connection i at an arbitrary moment, and let $L_i(t)$ be the transmission rate at connection i at time t within the current cycle. The LST of L_i is calculated by dividing the expected area of the function $e^{-sL_i(t)}$ over an arbitrary cycle, by the expected cycle time. That is,

$$\tilde{L}_i(s) = E[e^{-sL_i}] = \frac{E[\int_0^C e^{-sL_i(t)}dt]}{E[C]}. \tag{17}$$

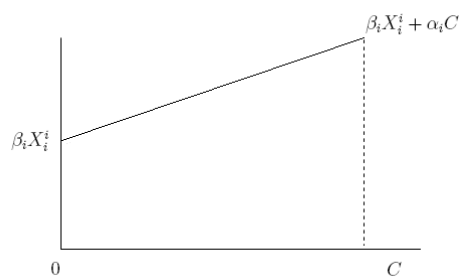

Figure 1: Transmission rate during a cycle.

Figure 1 shows the transmission rate at connection i during a full cycle, then

$$\tilde{L}_i(s) = \frac{E[\int_0^C e^{-s(\beta_i X_i^i+\alpha_i t)}dt]}{E[C]} = \frac{\tilde{X}_i(\beta_i s)(1-\tilde{C}(\alpha_i s))}{s\alpha_i E[C]}. \tag{18}$$

By taking derivative of (18) we get

$$E[L_i] = \beta_i E[X_i^i] + \alpha_i\frac{E[C^2]}{2E[C]} = \frac{\beta_i\alpha_i Nu}{1-\beta_i} + \frac{\alpha_i u^{(2)}}{2u} + \frac{\alpha_i(N-1)u}{2}. \tag{19}$$

That is, the mean transmission rate is the sum of the rate just after the polling instant ($\beta_i E[X_i^i]$) and of the accumulated rate during the mean residual time of a cycle ($\alpha_i\frac{E[C^2]}{2E[C]}$). The total throughput of the system is given by

$$\sum_{i=1}^{N} E[L_i] = Nu\sum_{i=1}^{N}\left(\frac{\beta_i\alpha_i}{1-\beta_i}\right) + \frac{u^{(2)}}{2u}\sum_{i=1}^{N}\alpha_i + \frac{(N-1)u}{2}\sum_{i=1}^{N}\alpha_i. \tag{20}$$

2.2 Probabilistic One Strategy

Under the "Probabilistic One Strategy" when the server gets a signal it decides to reduce the transmission rate to one connection, but the choice of the connection to decrease its rate is done probabilistically. Let p_i be the probability that the signal is sent to connection $i, (i = 1, \ldots, N)$, where $\sum_{i=1}^N p_i = 1$. Let $X_i^{(n)}$ denote the transmission rate at connection i just before the nth reduction (polling) instant. We assume that $X_i^{(n)}$ converges to X_i when $n \to \infty$. Let U denote the time between two successive polling instants.
The transmission rate of connection i is continuously growing at a rate α_i. When the server polls connection i with probability p_i, the transmission rate decreases by a factor of β_i. Hence the evolution of the state of the system (transmission rate) is given by

$$X_i^{(n+1)} = \begin{cases} X_i^{(n)} + \alpha_i U & \text{w.p } 1-p_i \\ \beta_i X_i^{(n)} + \alpha_i U & \text{w.p } p_i. \end{cases} \tag{21}$$

To calculate the LST of the transmission rate at polling instant, $L(\theta_1, \ldots, \theta_N)$, we express $L_{n+1}(\theta_1, \ldots, \theta_N)$ in terms of $L_n(\theta_1, \ldots, \theta_N)$. This is done by conditioning on the specific connection being chosen at the nth reduction signal,

$$\begin{aligned}&L_{n+1}(\theta_1, \ldots, \theta_{i-1}, \theta_i, \theta_{i+1}, \ldots, \theta_N|A_i) = E[e^{-\sum_{j=1}^N \theta_j X_j^{(n)}}|A_i] \\ &= E[e^{-\sum_{\substack{j=1\\ j\neq i}}^N \theta_j(X_j^{(n)}+\alpha_j U)} e^{-\theta_i\beta_i X_i^{(n)} - \theta_i\alpha_i U}] \\ &= L_n(\theta_1, \ldots, \theta_{i-1}, \beta_i\theta_i, \theta_{i+1}, \ldots, \theta_N)\cdot\tilde{U}(\sum_{j=1}^N \theta_j\alpha_j),\end{aligned} \tag{22}$$

where A_i is the event that connection i was polled at the previous (in this case, the nth) polling instant.
By unconditioning (22) and letting n approaching infinity we obtain

$$\begin{aligned}&L(\theta_1, \ldots, \theta_N) = \tilde{U}(\sum_{j=1}^N \theta_j\alpha_j) \\ &\cdot\Big(p_1L(\beta_1\theta_1, \ldots, , \theta_N) + \cdots + p_iL(\theta_1, \ldots, \theta_{i-1}, \beta_i\theta_i, \theta_{i+1}, \ldots, \theta_N) \\ &+ \cdots + p_NL(\theta_1, \ldots, \beta_N\theta_N)\Big).\end{aligned} \tag{23}$$

Transmission Rate at Reduction Instants: Moments
The moments of X_i are derived from (23) (or directly from (21)),

$$E[X_i] = -\frac{\partial L(\theta_1, \ldots, \theta_N)}{\partial\theta_i}\Big|_{\underline{\theta}=\underline{0}} = \frac{\alpha_i u}{p_i(1-\beta_i)}. \tag{24}$$

For the special case where $p_i = \frac{1}{N}$, we find that (24) is equal to the equivalent expression for $f_i(i)$ in the "Cyclic One Strategy" system (see equation (6)).

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7725
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7725

Unlike many other polling systems in this model we can derive explicit expressions for the second (mixed) moments in a non-identical connections case

$$E[X_i^2] = \frac{\alpha_i^2(u^{(2)} - 2u^2)}{p_i(1-\beta_i)(1+\beta_i)} + \frac{2\alpha_i^2 u^2}{p_i^2(1-\beta_i)^2(1+\beta_i)}, \quad (25)$$

$$\begin{aligned} E[X_iX_j] = \alpha_i\alpha_j \Big(& \frac{(u^{(2)} - 2u^2)}{p_i(1-\beta_i) + p_j(1-\beta_j)} \\ & + \frac{u^2}{p_i(1-\beta_i)p_j(1-\beta_j)} \Big) \qquad j \neq i, \end{aligned} \quad (26)$$

and

$$Cov(X_i, X_j) = \frac{\alpha_i\alpha_j(u^{(2)} - 2u^2)}{p_i(1-\beta_i) + p_j(1-\beta_j)} \qquad j \neq i. \quad (27)$$

Throughput of Connection i

Let C_i denote the time between two successive polling instants to connection i. Then

$$C_i = \sum_{j=1}^{T_i} U_j, \quad (28)$$

where all U_j's are distributed as U, and T_i is the number of polling instants between two successive polling to connection i, and is distributed geometrically with parameter p_i. Hence,

$$E[C_i] = E[T_i]u = \frac{u}{p_i}, \quad (29)$$

$$\begin{aligned} E[C_i^2] &= E[(\sum_{i=1}^{T_i} U_i)^2] = E_{T_i}[E[(\sum_{i=1}^{T_i} U_i)^2 | T_i]] \\ &= E_{T_i}[E[(U_1 + \cdots + U_{T_i})^2] \\ &= E[T_iU_1^2 + T_i(T_i - 1)U_1U_2] \\ &= E[T_i]u^{(2)} + (E[T_i^2] - E[T_i])u^2 \\ &= \frac{u^{(2)}}{p_i} + (\frac{2}{p_i^2} - \frac{2}{p_i})u^2 = \frac{u^{(2)}}{p_i} + \frac{2(1-p_i)u^2}{p_i^2}. \end{aligned} \quad (30)$$

Let L_i be a random variable denoting the transmission rate at connection i at arbitrary times. Using the same analysis as in the previous section we get

$$\tilde{L}_i(s) = \frac{E[\int_0^{C_i} e^{-s(\beta_i X_i + \alpha_i t)}dt]}{E[C_i]} = \frac{\tilde{X}_i(\beta_i s)(1 - \tilde{C}_i(\alpha_i s))}{s\alpha_i E[C_i]}, \quad (31)$$

where $\tilde{X}_i(s) = L(0, \ldots, 0, s, 0, \ldots, 0)$. Hence,

$$E[L_i] = \beta_i E[X_i] + \alpha_i \frac{E[C_i^2]}{2E[C_i]} = \frac{\alpha_i\beta_i u}{p_i(1-\beta_i)} + \alpha_i\big(\frac{u^{(2)}}{2u} + \frac{(1-p_i)u}{p_i}\big). \quad (32)$$

When $p_i = \frac{1}{N}$

$$\begin{aligned} E[L_i|\text{prob}] &= \frac{\alpha_i\beta_i N u}{1-\beta_i} + \alpha_i\big(\frac{u^{(2)}}{2u} + (N-1)u\big) \\ &= E[L_i|\text{cyclic}] + \frac{\alpha_i(N-1)u}{2}, \end{aligned} \quad (33)$$

we get that for connection i, the difference between the mean transmission rate of the "Probabilistic One Strategy" and that of the "Cycle One Strategy" is $\frac{\alpha_i(N-1)u}{2}$. This phenomenon can be better understood when looking at Figure 1: if the time intervals between rate reductions are less regular (i.e., probabilistic vs. cyclic), then the area under the graph (between two consecutive reduction instants) increases. Summing (32) for all i gives the mean total throughput of the system

$$\sum_{i=1}^{N} E[L_i] = \sum_{i=1}^{N} \left(\frac{\alpha_i\beta_i u}{p_i(1-\beta_i)} + \alpha_i\big(\frac{u^{(2)}}{2u} + \frac{(1-p_i)u}{p_i}\big) \right). \quad (34)$$

In the case where for all i $p_i = \frac{1}{N}$, the mean overall throughput under the "Probabilistic One Strategy" is larger than that of the "Cycle One Strategy" by the amount $\frac{(N-1)u}{2} \sum_{i=1}^{N} \alpha_i$.

Optimal values of p_i

By using Lagrange Multipliers we get the optimal values of p_i that maximize equation (34), denoted p_i^*, as

$$p_i^* = \frac{\sqrt{\frac{1-\beta_i}{\alpha_i}}}{\sum_{j=1}^{N} \sqrt{\frac{1-\beta_j}{\alpha_j}}}. \quad (35)$$

3. MIMD

3.1 Probabilistic One Strategy

Our approach will be based on showing that a logarithmic transformation applied to the transmission rate process results in a process that has the same evolution as the queue size in an M/G/1 bulk queue. The LST of the equivalent queueing process thus obtained provides the moments of the transmission rate of the connections. The transmission rate of connection i grows continuously, exponentially by e^{γ_i}, and when the server decides to reduce the rate of connection i, it is decreased by a factor of β_i ($0 < \beta_i < 1$). We consider a probabilistic polling strategy. As in the previous section, $X_i^{(n)}$ denotes the transmission rate at connection i just before the nth polling instant, and $U^{(n)}$ is the time between the nth and the $n+1$st polling instants (all $U^{(n)}$ are distributed as a general random variable U). Altman et al. [3] analyzed a similar model where a connection is multiplicative increased by a constant factor i.e., $X_i^{(n+1)} = \alpha_i X_i^{(n)}$ ($\alpha_i > 1$), whereas in our model $X_i^{(n+1)}$ is increased by a function of $U^{(n)}$. The evolution of the state of the system is given by

$$X_i^{(n+1)} = \begin{cases} X_i^{(n)} e^{\gamma_i U^{(n)}} & \text{w.p } 1 - p_i \\ \beta_i X_i^{(n)} e^{\gamma_i U^{(n)}} & \text{w.p } p_i. \end{cases} \quad (36)$$

We assume that the transmission rate is bounded below by a value of one. Altman et al. [3] showed the importance of the bounded value, hence (36) turns into

$$X_i^{(n+1)} = \begin{cases} X_i^{(n)} e^{\gamma_i U^{(n)}} & \text{w.p } 1 - p_i \\ \max(\beta_i X_i^{(n)}, 1) e^{\gamma_i U^{(n)}} & \text{w.p } p_i. \end{cases} \quad (37)$$

In order to evaluate the moments of X_i, we take the logarithm of equation (37);

$$\log X_i^{(n+1)} = \begin{cases} \log X_i^{(n)} + \gamma_i U^{(n)} & \text{w.p } 1 - p_i \\ \max(\log X_i^{(n)} + \log\beta_i, 0) + \gamma_i U^{(n)} & \text{w.p } p_i. \end{cases} \quad (38)$$

Dividing equation (38) by $-\log\beta_i > 0$ and using the substi-

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7725
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7725

tution $Y_i = \frac{\log X_i}{-\log \beta_i}$, we obtain

$$Y_i^{(n+1)} = \begin{cases} Y_i^{(n)} + \frac{\gamma_i}{-\log \beta_i} U^{(n)} & \text{w.p } 1-p_i \\ \max(Y_i^{(n)} - 1, 0) + \frac{\gamma_i}{-\log \beta_i} U^{(n)} & \text{w.p } p_i. \end{cases} \tag{39}$$

Define $T_i^{(n)} = \frac{\gamma_i}{-\log \beta_i} U^{(n)}$, $T_i^{(n)}$ is a non-negative random variable, $T_i^{(n)}$ and $Y_i^{(n)}$ are independent random variables. Then from equation (39) we obtain

$$Y_i^{(n+1)} = \begin{cases} Y_i^{(n)} + T_i^{(n)} & \text{w.p } 1-p_i \\ \max(Y_i^{(n)} - 1, 0) + T_i^{(n)} & \text{w.p } p_i. \end{cases} \tag{40}$$

If $T_i^{(n)}$ is an integer than equation (40) has the same form as the equation describing the number of customers in an M/G/1 queue just after the nth service (of length U), or a vacation period, where when the server finishes a service (or a vacation) period it serves the next customer with probability p_i, or takes a vacation of length U with probability $1-p_i$. $T_i^{(n)}$ is the number of new arrivals during the length of time U. Let's assume that $T_i^{(n)}$ is a fraction of an integer w. Hence $T_i^{(n)}$ can have the following values $(0, \frac{1}{w}, \frac{2}{w}, \ldots, \frac{w-1}{w}, 1, \frac{w+1}{w}, \ldots, \infty)$. Define $Q_i^{(n)} = w \cdot Y_i^{(n)}$ and $M_i^{(n)} = w \cdot T_i^{(n)}$. Then,

$$Q_i^{(n+1)} = \begin{cases} Q_i^{(n)} + M_i^{(n)} & \text{w.p } 1-p_i \\ \max(Q_i^{(n)} - w, 0) + M_i^{(n)} & \text{w.p } p_i. \end{cases} \tag{41}$$

$Q_i^{(n)}$ is an integer, thus $Q_i^{(n)}$ can be modeled as a discrete state space Markov chain. The last equation is actually the law of motion for the M/G/1 queue with bulk service of batch size w (see [5]) where upon finishing a service the server chooses whether to serve the next bulk or take a vacation. The probability generating function (PGF) of Q_i is obtained from the law of motion (41) using the following

$$\begin{aligned} E[z^{Q_i^{(n+1)}}] &= (1-p_i)E[z^{Q_i^{(n)}+M_i^{(n)}}] \\ &\cdot p_i\Big(E[z^{Q_i^{(n+1)}}|Q_i^{(n)} > 0]P(Q_i^{(n)} > 0) \\ &\quad + E[z^{Q_i^{(n+1)}}|Q_i^{(n)} = 0]P(Q_i^{(n)} = 0)\Big). \end{aligned} \tag{42}$$

Recall that $Q_i^{(n)}$ and $M_i^{(n)}$ are independent random variables. Then from (42) we obtain, when $Q_i^{(n)} \to Q_i$,

$$\hat{Q}_i(z) = \frac{p_i \hat{M}_i(z) \sum_{j=0}^{w-1} \pi_i^{(j)} (z^w - z^j)}{z^w - \hat{M}_i(z)\big((1-p_i)z^w + p_i\big)}, \tag{43}$$

where $\pi_i^{(j)}$ is the probability that $Q_i = j$. The expression for $\hat{Q}_i(z)$ contains w unknowns parameters $\pi_i^{(0)}, \pi_i^{(1)}, \ldots, \pi_i^{(w-1)}$. To determine these we use the following equality

$$\sum_{j=0}^{w-1} \pi_i^{(j)} (z^w - z^j) = (z-1) \sum_{j=0}^{w-1} v_i^{(j)} z^j, \tag{44}$$

where $v_i^{(j)} = \sum_{k=0}^{j} \pi_i^{(k)}$ (see pp. 33 in [11]). Hence, we write

$$\hat{Q}_i(z) = \frac{p_i \hat{M}_i(z)(z-1) \sum_{j=0}^{w-1} v_i^{(j)} z^j}{z^w - \hat{M}_i(z)\big((1-p_i)z^w + p_i\big)}. \tag{45}$$

Now, $\hat{Q}_i(1) = 1$ implies

$$\sum_{j=0}^{w-1} v_i^{(j)} = w - \frac{m_i}{p_i}, \tag{46}$$

which is meaningful if and only if $\frac{m_i}{p_i} < w$ (or equivalently $t_i < p_i$). That is, the mean number of arrivals between two consecutive visits to queue i, namely $\frac{m_i}{p_i}$, must be smaller than the bulk service amount w.
Assuming equation (45) to be an analytic function in the disk $z : |z| \le 1+\delta$ implies that the numerator is zero whenever the denominator vanishes in $z : |z| \le 1+\delta$. That is, the numerator and the denominator of (45) have exactly the same number of roots in the above disk. Let's state Rouché's theorem.

Theorem 1. *If $f(z)$ and $g(z)$ are analytic functions of z inside and on a closed contour C, and also if $|g(z)| < |f(z)|$ on C, then $f(z)$ and $f(z)+g(z)$ have the same number of zeros inside C.*

Define $g(z) = \hat{M}_i(z)\big((1-p_i)z^w + p_i\big)$, $f(z) = z^w$. Because $g(1) = f(1) = 1$ and $g'(1) = m_i + w(1-p_i) < w = f'(1)$, we have for sufficiently small $\delta > 0$, $g(1+\delta) < f(1+\delta)$. Consider all z with $|z| = 1+\delta$, then

$$\begin{aligned} |g(z)| &= |\hat{M}_i(z)| \cdot |(1-p_i)z^w + p_i| \\ &\le \sum_{j=0}^{\infty} P(M_i = j)|z|^j \cdot \big((1-p_i)|z|^w + p_i\big) = g(1+\delta) \\ &< f(1+\delta) = |f(z)|, \end{aligned} \tag{47}$$

where the first inequality is due to the triangle inequality. Hence $|g(z)| < |f(z)|$, and by Rouché's theorem we know that $z^w - \hat{M}_i(z)\big((1-p_i)z^w + p_i\big)$ has the same number of zeros as z^w, i.e., w roots in the disk $z : |z| \le 1+\delta$, for every sufficiently small $\delta > 0$. Let these roots be denoted by $z_1, z_2, \ldots, z_{w-1}$ and $z_w = 1$. Since the PGF $\hat{Q}_i(z)$ is analytic within the region $|z| \le 1$, the numerator of (45) should vanish at each of the roots. It follows that $\sum_{j=0}^{w-1} v_i^{(j)} z^j$ should vanish at $z_1, z_2, \ldots, z_{w-1}$. We thus have the following $w-1$ equations

$$\sum_{j=0}^{w-1} v_i^{(j)} z_k^j = 0 \qquad (k = 1, 2, \ldots, w-1). \tag{48}$$

Using the $w-1$ equations of (48) together with (46) we get

$$\sum_{j=0}^{w-1} v_i^{(j)} z^j = (w - \frac{m_i}{p_i}) \prod_{j=1}^{w-1} \frac{z - z_j}{1 - z_j}. \tag{49}$$

Hence,

$$\hat{Q}_i(z) = \frac{\hat{M}_i(z)(z-1)(p_i w - m)}{z^w - \hat{M}_i(z)\big((1-p_i)z^w + p_i\big)} \prod_{j=1}^{w-1} \frac{z - z_j}{1 - z_j}. \tag{50}$$

Finally, the moments of X_i can be obtain using $\hat{Q}_i(z)$. At steady state we have

$$X_i = \beta_i^{-Y_i}. \tag{51}$$

Therefore, the kth moment of X_i can be obtain as follows

$$E[X_i^k] = E[\beta_i^{-kY_i}] = E[\beta_i^{-\frac{k}{w} Q_i}] = \hat{Q}_i(\beta_i^{-\frac{k}{w}}). \tag{52}$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7725
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7725

The kth moment of X_i is finite as long as $\beta_i^{-\frac{k}{w}}$ is smaller than the smallest root of the denominator of (45) which is larger than 1.

3.2 Cyclic One Strategy

The analysis of the "Cyclic One Strategy" follow a direction similar to that of the "Probabilistic One Strategy". Let X_i^j and $X_i^{i(n)}$ be defined as section 2.1. Hence, the evolution of the state of the system is given by

$$X_{i+1}^j = \begin{cases} X_i^j \cdot e^{\gamma_j U_i} & \text{if } j \neq i \\ \max(\beta_i X_i^i, 1) \cdot e^{\gamma_i U_i} & \text{if } j = i, \end{cases} \quad (53)$$

or

$$X_i^{i(n+1)} = \max(\beta_i X_i^{i(n)}, 1) e^{\gamma_i C_i^{(n)}}, \quad (54)$$

where $C_i^{(n)}$ is the cycle time between the nth polling instant of channel i to the $n+1$st polling instant of that channel. Clearly $C_i^{(n)} = \sum_{j=1}^N U_j$. Using the substitution $Y_i^{(n+1)} = \frac{\log X_i^{i(n)}}{-\log \beta_i}$ as in the previous section, we get

$$Y_i^{(n+1)} = \max(Y_i^{(n)} - 1, 0) + \frac{\gamma_i}{-\log \beta_i} C_i^{(n)}. \quad (55)$$

Defining $T_i^{(n)} = \frac{\gamma_i}{-\log \beta_i} C_i^{(n)}$, and assuming that $T_i^{(n)}$ is a fraction of an integer w, we can transform the evolution equation (55) to the same evolution of an M/G/1 bulk queue, with bulk service w

$$Q_i^{(n+1)} = \begin{cases} Q_i^{(n)} - w + M_i & \text{if } Q_i^{(n)} > w \\ M_i & \text{if } Q_i^{(n)} \leq w, \end{cases} \quad (56)$$

where $Q_i^{(n)} = w \cdot Y_i^{(n)}$ and $M_i^{(n)} = w \cdot T_i^{(n)}$ are integers random variables. Then, from (56), the PGF of Q_i is

$$\hat{Q}_i(z) = \frac{\hat{M}_i(z) \sum_{j=0}^{w-1} \pi_i^{(j)} (z^w - z^j)}{z^w - \hat{M}_i(z)}, \quad (57)$$

and by using Rouché's theorem as in the previous section, we obtain (see [5])

$$\hat{Q}_i(z) = \frac{\hat{M}_i(z)(z-1)(w-m_i)}{z^w - \hat{M}_i(z)} \prod_{j=1}^{w-1} \frac{z - z_j}{1 - z_j}. \quad (58)$$

Notice the similarity between equations (50) and (58) (substituting $p_i = 1$ in (50) yields (58)) except for M_i which is defined differently in both schemes. The kth moment of X_i^i is obtained, as in equation (52), by

$$E[X_i^{i\,k}] = \hat{Q}_i(\beta_i^{-\frac{k}{w}}). \quad (59)$$

4. CONCLUSIONS

This paper analyzes a polling-type procedures of a TCP mechanism for a system with N connections sharing a common AQM. Both the AIMD and MIMD schemes are studied for the cyclic and the probabilistic polling policies. LST, mean and (explicit value for the) second moment of the transmission rate of each connection are derived and overall mean throughput is calculated. For the analysis of the MIMD scheme, an analogy to M/G/1 queue with bulk service is utilized and enables the complete analysis of the system.

5. REFERENCES

[1] E. Altman, K. Avratchenkov and C. Barakat. TCP in persence of bursty loses. ACM SIGMETRICS, June 2000.

[2] E. Altman, K. Avratchenkov and C. Barakat. A stochastic model of TCP/IP with stationary random losses. ACM SIGCOMM, October 2000.

[3] E. Altman, R. El Khoury. Analysis of scalable TCP. Proceedings of the International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks. (HET-NETS), July 2004.

[4] E. Altman, D. Barman, B. Tuffin and M Vojnovic. Parallel TCP sockets: Simple model, throughput and validation. INFOCOM 2006, 25th IEEE International Conference, 2006.

[5] M.L. Chaudhry and J.G.C Templeton. A first course in Bulk Queues. Wiley, New York 1986.

[6] O. Czerniak and U.Yechiali. Fluid polling systems. Proceeding of the Erlang Centennial-100 Years of Queueing, pp. 47-74, Copenhagen, April 1-3, 2009 (Also to appear in Queueing Systems Vol. 63, Sept.-Dec. 2009).

[7] D. Damjanovic, M. Welzl, M. Telek and W. Heiss. Extending the TCP Steady-State Throughput Equation for Parallel TCP Flows. University of Innsbruck, Institute of Computer Science, DPS NSG Technical Report 2, August 2008.

[8] S. Floyd. Connections with multiple congested gateways in packet-switched networks part 1: One-way traffic. ACM Computer Communication Review, October 1991.

[9] V. Jacobson. Congestion avoidance and control. ACM SIGCOMM, August 1988.

[10] H. Takagi. Analysis of polling systems. MIT Press, Cambridge, Mass 1986.

[11] N.U. Prabhu. Queues and inventories. John Wiley, New York, 1965.

[12] C-T. Yang, M-F. Yang and W-C. Chiang. Enhancement of anticipative recursively adjusting mechanism forredundantparallel file transfer in data grids. Journal of Network and Computer Applications 32, pp. 834-84, July 2009.

[13] U. Yechiali. Analysis and control of polling systems. Performance Evaluation of Computer and Communication System, Vol. 729, pp. 630-650, Springer-Verlag 1993.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7725
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7725