Performance of TCP/IP

E. Altman, K. Avratchenkov, C. Barakat

INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93
06902 Sophia-Antipolis Cedex

E-mail: altman@sophia.inria.fr
URL: http://www-sop.inria.fr/mistral/personnel/Eitan.Altman/me.html

July 2000

OUTLINE

Introduction

Need for simple formula for throughput

Why?

- Helps to understand fairness
- Helps to dimension the network (buffers, links etc)
- Useful for real time TCP-friendly applications.

Some related references:

T.V. Lakshman and U. Madhow, "The performance of TCP/IP for networks with high bandwidth-delay products and random loss", *IEEE/ACM Transactions on Networking*, Jun 1997.

M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm", *ACM Computer Communication Review*, Jul 1997.

V. Misra, W.-B. Gong, and D. Towsley, "Stochastic differential equation modelling and analysis of TCP-windowsize behaviour", *Performance*, Oct 1999.

J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modelling TCP throughput: A simple model and its empirical validation", *ACM SIGCOMM*, Sep 1998.

S. Floyd, M. Handley and J. Padhye, "Equation-based congestion control for unicast applications" *ACM Sigcomm*, Aug. 28 - Sept. 1, Stockholm, Sweden, 2000.

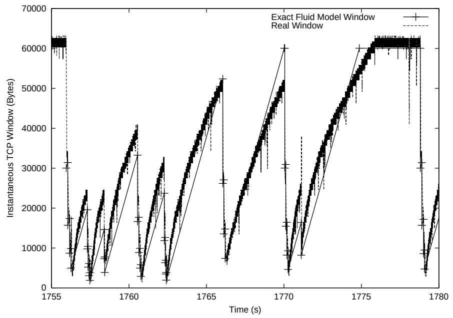

Figure 1: Measured dynamics INRIA - ESSI

E. A., K. Avratchenkov and C. Barakat, “TCP in presence of bursty losses”, *ACM SIGMETRICS*, Jun 2000. Extended version: to appear in *Performance Evaluation*.

E. A., K. Avratchenkov and C. Barakat, "A stochastic model of TCP/IP with stationary random losses", *ACM Sigcomm*, Aug. 28 - Sept. 1, Stockholm, Sweden, 2000.

Main features of TCP

- 1. The throughput variation can be approximated by a Linear Increase (rate α), multiplicative decrease upon congestion signal (typically 1/2).
  Typically $\alpha = 1/RTT^2$ or $\alpha = 1/2RTT^2$ (delayed ack version).

- 2. Congestion signals: loss of packets. Detected through duplicated acks or time-outs. Several losses in the same RTT are considered a single loss event.

- 3. Throughput = window size / RTT.

2 state Markov model

- T_n:= the time at which the nth potential loss may occur.
  Potential losses are transformed to real losses with a certain probability.
- $D_n = T_{n+1} - T_n$, i.i.d., with expectation d, second moment $d^{(2)}$ and Laplace Stieltjis Transform $D^*(s)$.
- X_n:= transmission rate just prior to the instant of the nth potential loss.
- Y_n:= the state of the channel at the nth potential loss instant.
  B:= Bad state, a potential loss is transformed to a real loss with probability p_B
  G:= Good state, a loss occurs with a smaller probability, p_G.
- We assume further that the sequences $\{Y_n\}$ and $\{D_n\}$ are independent.

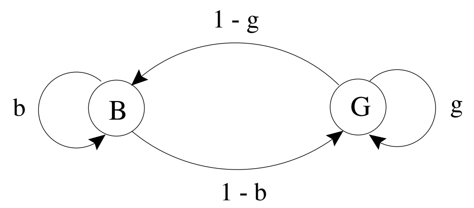

Figure 2: The Markov chain associated to the path

- $\{Y_n\}$ is a Markov chain with the following transition matrix (Figure 2)

$$P = \begin{bmatrix} g & \bar{g} \\ \bar{b} & b \end{bmatrix}.$$

- $g, b \in (0,1)$; hence $\{Y_n\}_{n=1}^{+\infty}$ is then ergodic with stationary probabilities

$$\pi_G = \frac{1-b}{2-b-g} = \frac{\bar{b}}{\bar{b}+\bar{g}}, \qquad \pi_B = \frac{1-g}{2-b-g} = \frac{\bar{g}}{\bar{b}+\bar{g}}.$$

- The average loss rate is given by,

$$R = \frac{p_G \pi_G + p_B \pi_B}{d}.$$

This is equal to the average number of times the source reduces its rate per unit of time.

- U_n and V_n:= indicators of the event at the nth instant, a potential loss occurs (at good/bad state).

A value zero however means that X_n is not affected (i.e. a real loss didn't occur). We have,

$$P(U_n = 1) = p_G, \qquad P(U_n = 0) = 1 - p_G$$
$$P(V_n = 1) = p_B, \qquad P(V_n = 0) = 1 - p_B$$

Usefulness of a two state Markov model

- Allows to obtain correlations between losses. In experimentation INRIA – ESSI we obtained a correlation coefficient $\eta = -0.2$

$$\eta := \frac{ES_{n+1}S_n - (ES_n)^2}{ES_nS_n - (ES_n)^2}.$$

(S_n: time between losses).

- Allows to model burstiness.

- Identification: often D_n is easy to choose, for example D_n =RTT.

Relevance to TCP friendly applications

- In audio over Internet, often FEC is added.
  Some applications already include the estimation of a two state Markov chain of the channel in order to optimize the amount of FEC.

- In two TCP connections with the same average loss rate, the one with higher burstiness has better throughput.
  A TCP friendly application that ignores the burstiness of the channel will be too conservative, yield bad QoS. account.

Performance Analysis of the Markov model

The evolution of the transmission rate:

$$\begin{aligned} X_{n+1} &= \left(1 - \frac{U_n}{2}\right) X_n 1\{Y_n = G\} \\ &\quad + \left(1 - \frac{V_n}{2}\right) X_n 1\{Y_n = B\} + \alpha D_n \end{aligned} \tag{1}$$

Define the column vector

$$\mathbf{X}_n = (X_n 1\{Y_n = G\}, X_n 1\{Y_n = B\})^T.$$

Define the matrix $Q_n :=$

$$\begin{pmatrix} (1 - U_n/2)1\{Y_{n+1} = G\} & (1 - V_n/2)1\{Y_{n+1} = G\} \\ (1 - U_n/2)1\{Y_{n+1} = B\} & (1 - V_n/2)1\{Y_{n+1} = B\} \end{pmatrix}.$$

Define the column vector

$$\mathbf{D}_n = (D_n 1\{Y_n = G\}, D_n 1\{Y_n = B\})^T.$$

It follows from (1): $$\mathbf{X}_{n+1} = Q_n \mathbf{X}_n + \alpha \mathbf{D}_n. \quad (2)$$

RESULTS

- **This is a linear stochastic difference equation**. X_n converges (in a strong sense) to a stationary ergodic process.
- All moments of X_n converge to the steady state moments.
- To compute moments of X_n, define the Laplace Stieltjis Transforms (LST):

$$\begin{aligned} Z_n(s,G) &= E\left[e^{-sX_n} 1\{Y_n = G\}\right], \\ Z_n(s,B) &= E\left[e^{-sX_n} 1\{Y_n = B\}\right] \end{aligned}$$

In addition, let us define

$$\mathbf{Z}_n(s) = \left[\, Z_n(s,G) \;\; Z_n(s,B) \,\right].$$

Theorem 1 *The Laplace Stieltjis Transform $\mathbf{Z}_n(s)$ satisfies the following recurrent equation,*

$$\mathbf{Z}_{n+1}(s) = D^*(\alpha s)\mathbf{Z}_n(s)P_1 + D^*(\alpha s)\mathbf{Z}_n(\frac{s}{2})P_2, \quad (3)$$

where

$$P_1 = \begin{bmatrix} g(1-p_G) & \bar{g}(1-p_G) \\ \bar{b}(1-p_B) & b(1-p_B) \end{bmatrix}$$

and

$$P_2 = \begin{bmatrix} gp_G & \bar{g}p_G \\ \bar{b}p_B & bp_B \end{bmatrix}.$$

This allows to obtain all moments of X_n as well as the time average throughput.

Model with Stationary Ergodic Random Losses

- Real loss process: $\{T_n\}_{n=-\infty}^{+\infty}$ a general stationary ergodic point process with intensity $\lambda > 0$.

$$X_{n+1} = \nu X_n + \alpha S_n, \tag{4}$$

where

- X_n is the value of $X(t)$ just prior to the arrival of the loss at T_n
- $S_n := T_{n+1} - T_n$.
- Again, stochastic linear difference equation!

Performance Analysis of the stationary-ergodic model

- Convergence to a stationary solution given by

$$X_n^* = \alpha \sum_{k=0}^{\infty} \nu^k S_{n-1-k}. \tag{5}$$

Proposition 1 *Let* $\lambda = 1/E[S_n]$ *be an intensity of the loss process and let* $R(k) = E[S_n S_{n+k}]$ *be a correlation function for the process* $\{S_n\}_{n=-\infty}^{+\infty}$. *Then,*

$$E[X_n^*] = \frac{\alpha}{\lambda(1-\nu)} \tag{6}$$

$$E[(X_n^*)^2] = \frac{\alpha^2}{1-\nu^2}[R(0) + 2\sum_{k=1}^{\infty} \nu^k R(k)] \tag{7}$$

Proof: We use $X_{n+1} = \nu X_n + \alpha S_n$. Taking expectations, we get:

$$E[X] = \nu E[X] + \alpha/\lambda,$$

which gives $E[X]$.

$E[X^2]$ is obtained from:

$$\begin{aligned} EX_{n+1}^2 &= E(\nu X_n + \alpha S_n)^2 \\ &= \nu^2 EX_n^2 + \alpha^2 ES_n^2 + 2\nu\alpha E[X_n S_n] \end{aligned}$$

$$E[X_n S_n] = \alpha E\left[\left(\sum_{k=0}^{\infty} \nu^k S_{n-1-k}\right) S_n\right]$$

- The throughput:

$$\bar{X} = \lim_{T\to\infty} \frac{1}{T} \int_0^T X(t)dt.$$

Proposition 2 *The throughput is given by*

$$\bar{X} = \lambda\alpha[\frac{1}{2}R(0) + \sum_{k=1}^{\infty} \nu^k R(k)] \tag{8}$$

Other representations

- Define the covariance function

$$C(k) = R(k) - E[S_n]^2$$

Then,

$$\bar{X} = \lambda\alpha[\frac{1}{2}C(0) + \sum_{k=1}^{\infty} \nu^k C(k)] + \frac{\alpha(1+\nu)}{2\lambda(1-\nu)}.$$

- Define $A(t)$:= number of packets transmitted on the TCP connection until time t,
  $L(t)$:= number of loss events until time t.

- The probability p of losing a packet is then simply

$$p = \lim_{t\to\infty} \frac{L(t)}{A(t)} = \lim_{t\to\infty} \frac{\lambda t}{\int_0^t X(\tau)d\tau} = \frac{\lambda}{\bar{X}} = \frac{1}{d\bar{X}}. \qquad (9)$$

- Define the normalized corelation function: $\hat{R}(n) = R(n)/d^2$. Then using (9) and $\alpha = 1/(bRTT^2)$, we get

$$\overline{X} = \frac{1}{RTT\sqrt{pb}}\sqrt{\frac{1}{2}\hat{R}(0) + \sum_{k=1}^{\infty} \nu^k \hat{R}(k)}.$$

- Define the normalized covariance as $\hat{C}(k) = C(k)/d^2$ then

$$\overline{X} = \frac{1}{RTT\sqrt{pb}}\sqrt{\frac{1+\nu}{2(1-\nu)} + \frac{1}{2}\hat{C}(0) + \sum_{k=1}^{\infty} \nu^k \hat{C}(k)}. \quad (10)$$

We conclude that for arbitrary stationary ergodic loss process, the throughput of TCP is inversely proportional to RTT and to the square root of the average packet loss probability.

The i.i.d. case

Proposition. Let $\{S_n\}_{n=-\infty}^{+\infty}$ be i.i.d. with $d := E[S_n]$ and $d^{(2)} := E[S_n^2]$. Then,

$$E[X_n^*] = \frac{\alpha d}{1-\nu},$$

$$E[(X_n^*)^2] = \frac{\alpha^2}{1-\nu^2}[d^{(2)} + \frac{2\nu d^2}{1-\nu}],$$

$$\bar{X} = E[X(t)] = \frac{\alpha}{d}[\frac{1}{2}d^{(2)} + \frac{\nu d^2}{1-\nu}]. \quad (11)$$

- The throughput is a constant, that depends only on the average inter-loss times, plus a term which grows *linearly* with the variance of the inter-loss times.
- For large burstiness in the losses we get a large under-estimate of the throughput if we use the deterministic or exponentially distributed assumption on inter-loss times.

- In particular, if the inter-loss times are exponentially distributed, we have

$$\overline{X} = \frac{\alpha d}{1-\nu}. \tag{12}$$

If the inter-loss times are deterministic, we get

$$\overline{X} = \frac{1+\nu}{2(1-\nu)}\alpha d = \overline{X} = \frac{1}{RTT}\sqrt{\frac{3}{2bp}}, \tag{13}$$

See:

T.V. Lakshman and U. Madhow, 1997,

M. Mathis, J. Semke, J. Mahdavi, and T. Ott, 1997

J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, 1998.

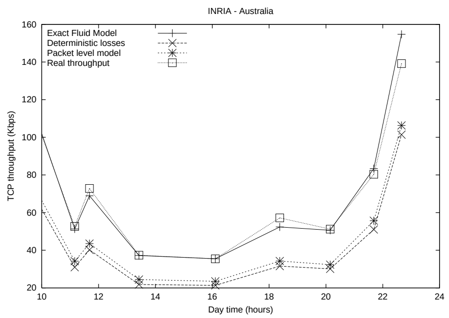

Figure 3: Comparison of throughput

Example of Experimentation

- Experimentation on January 2000, from clope.inria.fr (INRIA, Sophia-Antipolis) which uses New Reno.
- Destination: Univ. of South Australia (linus.levels.unisa.edu.au).
- TCP packets are of size 1460 Bytes (not including TCP and IP headers).
- Advertized window: 22 packets.