TCP in presence of bursty losses

Eitan Altman, Kostya Avrachenkov

INRIA Sophia Antipolis - France

Email : Chadi.Barakat@sophia.inria.fr
http://www.inria.fr/mistral/personnel/Chadi.Barakat

ACM SIGMETRICS
Tuesday, June 20 2000

Outline

- Introduction
  - What do we mean by bursty losses?
- Our models
  - A fluid model for rate evolution
  - A two-state Markovian model for losses
- Performance analysis
  - Moments and throughput calculation
- Model validation, conclusions and future work

TCP and losses

- TCP
  - Internet widely used transport protocol
  - An additive-increase multiplicative-decrease strategy for congestion control
  - Packet losses for congestion detection (possibly ECN)
- Loss event (or congestion event)
  - An event that causes the reduction of the congestion window
  - **Interpretation**: Depends on the version of TCP
    - One packet loss for Reno
    - One lossy round trip-time for Tahoe, New-Reno and SACK

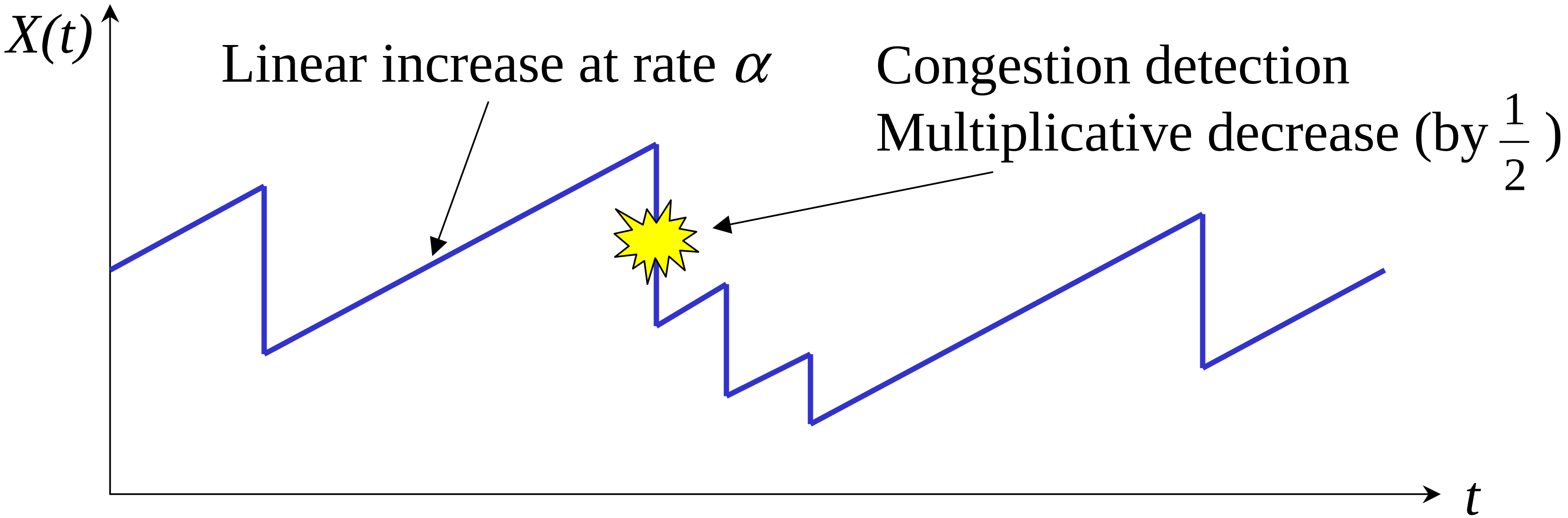
Why bursty losses?

- TCP modeling requires a characterization of inter-losses times
- Simple loss processes have been considered in the literature

  (Deterministic e.g. [Mathis et al. 1997], Poisson e.g. [Misra et al. 1999])
- TCP throughput has been only expressed as a function of the average loss rate (e.g. p , λ)
- What happens if, for the same average loss rate, loss events tend to appear in bursts?

Our fluid model for TCP

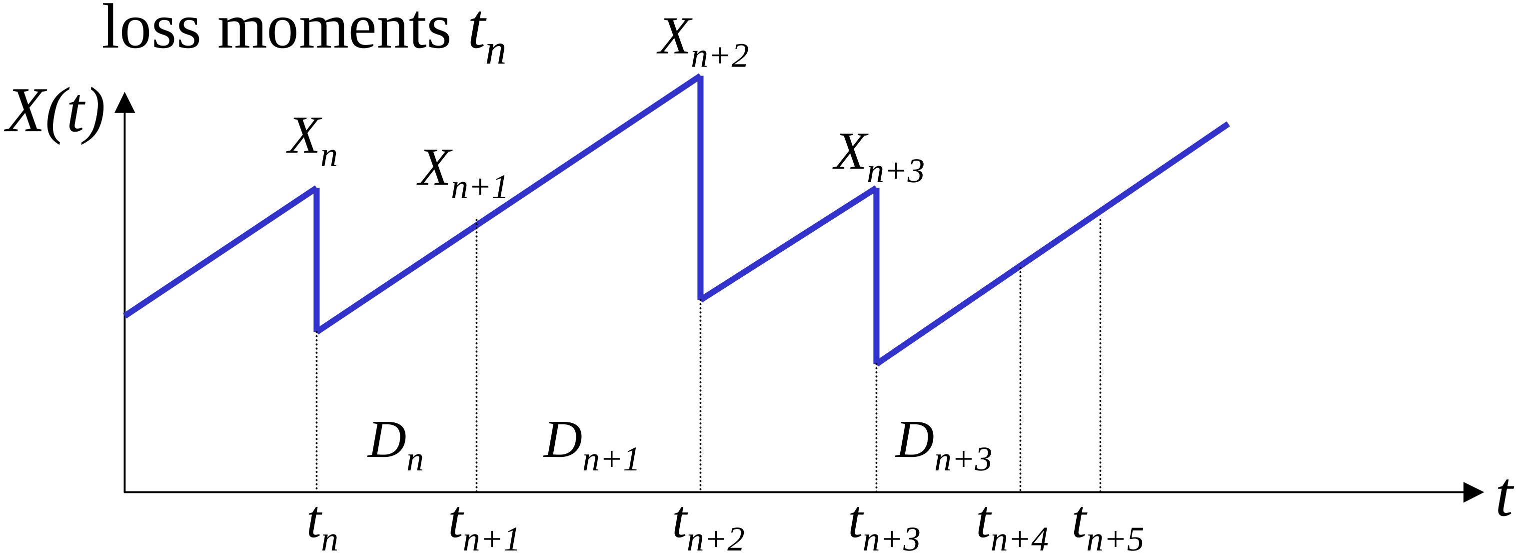
- Consider a long TCP transfer with infinite data to send
- Denote by $X(t)$ the rate of the connection at time t ($=W(t)/RTT$)
- Assume that losses are quickly detected (without long Timeout)



Our Markovian model for losses

Potential loss moments

- Reduce the rate with a certain probability at some potential loss moments t_n

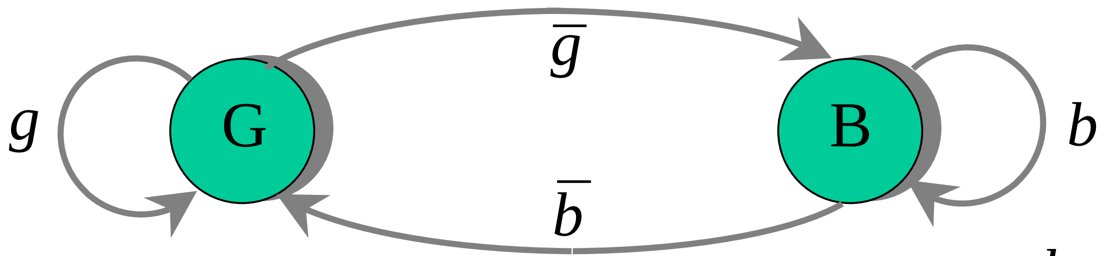


Assumptions: $\{D_n\}$ are i.i.d., $d=E[D_n] < \infty$, $d^{(2)}=E[D_n^2] < \infty$

Our Markovian model for losses

The Markov chain

- The path is in one of two states: **Bad** (B) or **Good** (G)
- Reduce the transmission rate (the window) at potential loss moments with different probabilities ($p_G < p_B$)
- Denote by *Y(t)* the state of the path
- Take $\{Y_n\}$ as a two-state Markov chain (Gilbert model)



$b, g \in (0,1)$

Performance analysis

- Input parameters
  - The average loss rate $R = \dfrac{p_G \pi_G + p_B \pi_B}{d}$
  - The burstiness of losses via *b* and *g*
- Output parameter
  - The throughput of the connection

$$\bar{x} = \lim_{t \to \infty} \frac{1}{t} \int_0^t X(\tau)\, d\tau$$

Stochastic difference equation

Let U_n (resp. V_n) $\in \{0,1\}$ indicate whether or not the rate is reduced at t_n in the **Good** (resp. **Bad**) state

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

Using Theorem 2A in [Glasserman and Yao,1995], and the fact that the reduction factor is one half, d < ∞, $\{Y_n\}$ is ergodic, we proved

- The difference equation has a unique stationary solution X_n^*
- X_n converges to X_n^* for any initial state X_0
- $\{X_n\}$ is an ergodic process

Convergence of first moments

Let $$\chi_n = \begin{bmatrix} E[X_n 1\{Y_n = G\}] \\ E[X_n 1\{Y_n = B\}] \\ P(Y_n = G) \\ P(Y_n = B) \end{bmatrix}$$

We can write $\chi_{n+1} = \Pi \chi_n$ where Π is some matrix with only one eigenvalue equal to 1 and the others less than 1. Thus, χ_n converges and we define

$$x_G = \lim_{n\to\infty} E[X_n 1\{Y_n = G\}]$$

$$x_B = \lim_{n\to\infty} E[X_n 1\{Y_n = B\}]$$

Calculation of moments

$$Z(s) = \left[Z(s,G) \quad Z(s,B) \right] = \lim_{n\to\infty} \left[E[e^{-sX_n} 1\{Y_n = G\}] \quad E[e^{-sX_n} 1\{Y_n = B\}] \right]$$

From the stochastic difference equation, conditioning on the state of the path, we get the following implicit LT equation

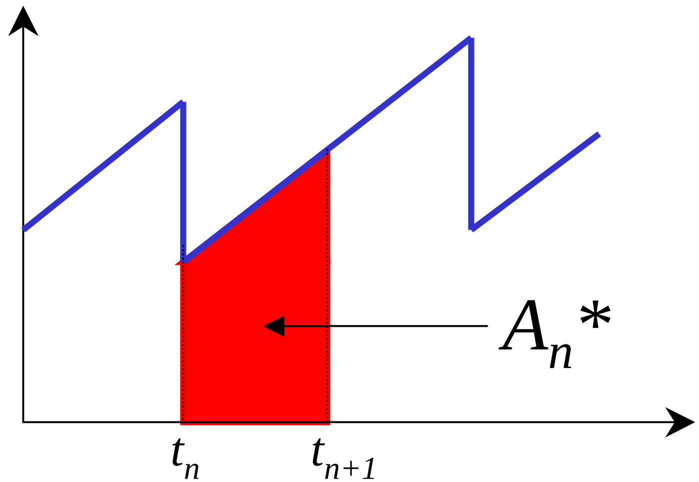
$$Z(s) = D^*(\alpha s) Z(s) P_1 + D^*(\alpha s) Z(s/2) P_2$$

$$P_1 = \begin{bmatrix} g(1-p_G) & \overline{g}(1-p_G) \\ \overline{b}(1-p_B) & b(1-p_B) \end{bmatrix} \qquad P_2 = \begin{bmatrix} gp_G & \overline{g}p_G \\ \overline{b}p_B & bp_B \end{bmatrix}$$

By taking derivatives and then setting s to 0, we can get all the moments of X_n* as a function of d, g, b, p_G and p_B

Calculation of the throughput

$$\bar{x} = \lim_{t \to \infty} \frac{1}{t} \int_0^t X(\tau)\, d\tau = \lim_{n \to \infty} \frac{\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} X(\tau)\, d\tau}{\sum_{i=0}^{n-1} D_i} = \frac{E[A_n^*]}{d}$$



The throughput

$$\bar{x} = (1 - \frac{p_G}{2}) x_G + (1 - \frac{p_B}{2}) x_B + \frac{1}{2} \alpha \frac{d^{(2)}}{d}$$

Impact of burstiness

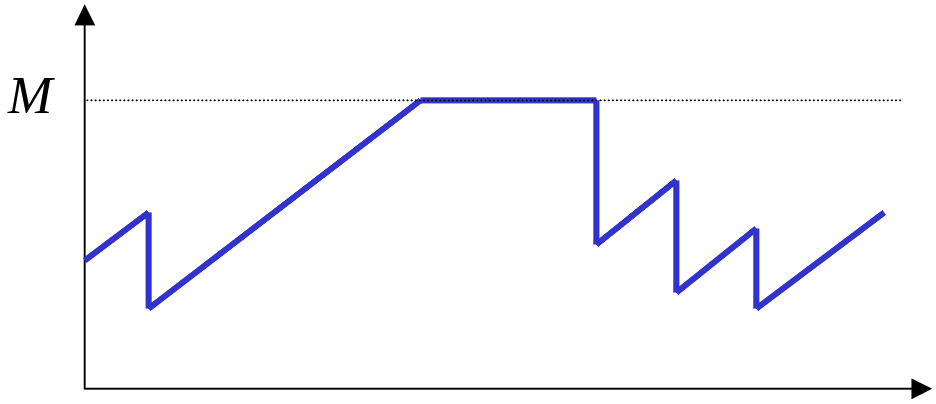
- Consider in the following the case $\boldsymbol{p_G}$**=0**, $\boldsymbol{p_B}$**=1** (no losses in the **Good** state and always losses in the **Bad** state)
- Define $\bar{x}_r$ as the throughput when losses are not bursty (a path with one state and with a loss probability π_B)

$$\bar{x} = \bar{x}_r + \alpha d \pi_G \left(\frac{1}{1-g} - \frac{1}{\pi_B} \right)$$

- The second term of the right-hand equation is always positive and **increases** with the burstiness (when g and b increase in a way that π_B and π_G remain constant)

Case of rate limitation

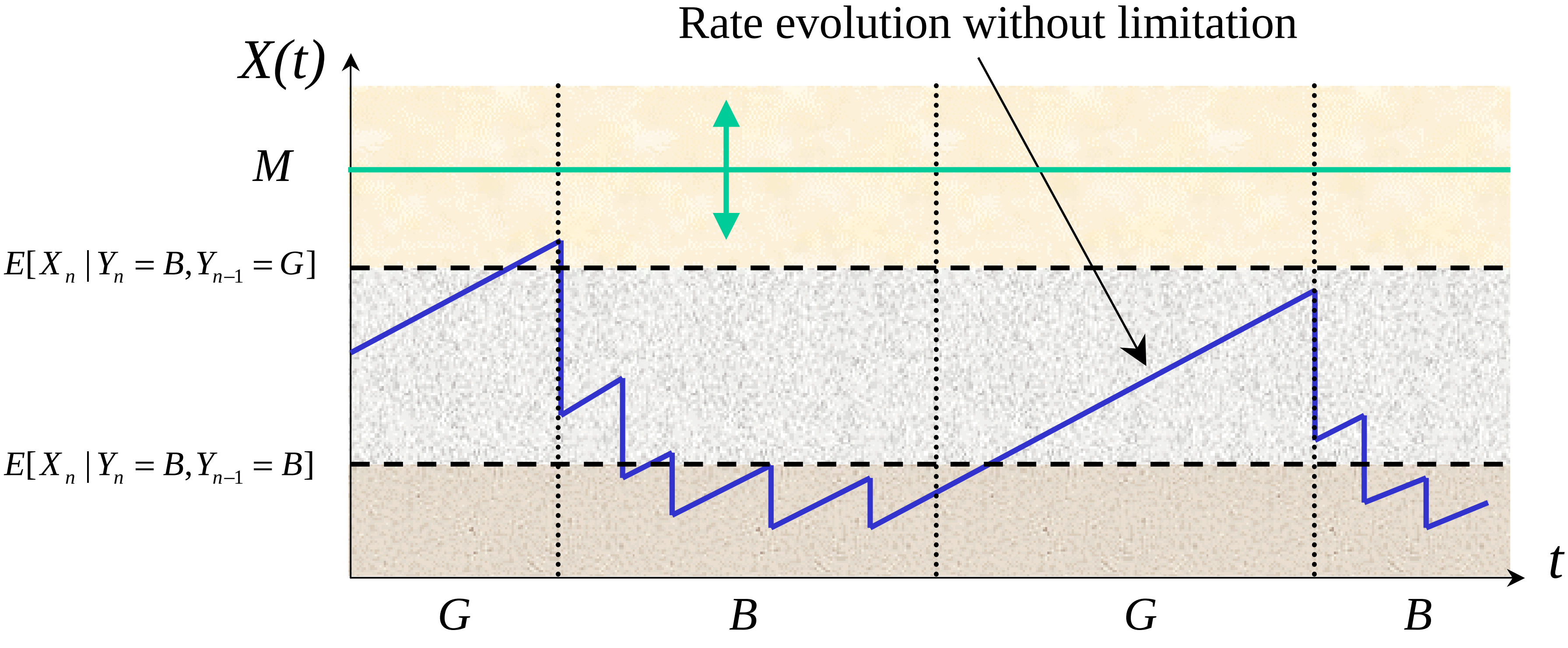
- The transmission rate may not exceed a certain limit (the receiver window in case of TCP, the PCR in case of ABR)



- We use a heuristic similar to that in [Padhye et al. 1998] to approximate the throughput

Throughput approximation

Divide the space of M into three regions



Throughput approximation

Large M

- Use the expression of $\bar{x}$ we found in case $M=\infty$

Medium M

- Assume that during the **Good** state, the rate starts at $x_0=E[X_n|Y_n=G,Y_{n-1}=B]$, reaches and stays at M until the next loss event where it drops
- Calculate the average rate during the Good state $\bar{x}_G$
- Calculate the average rate during the Bad state $\bar{x}_B$ using the expression of the throughput in case $M=\infty$

$$\bar{x} = \pi_G \bar{x}_G + \pi_B \bar{x}_B$$

Throughput approximation

Small M

- Take $\bar{x}_G = M$
- Calculate $\bar{x}_B$ using the assumption that the rate reaches M between two potential losses

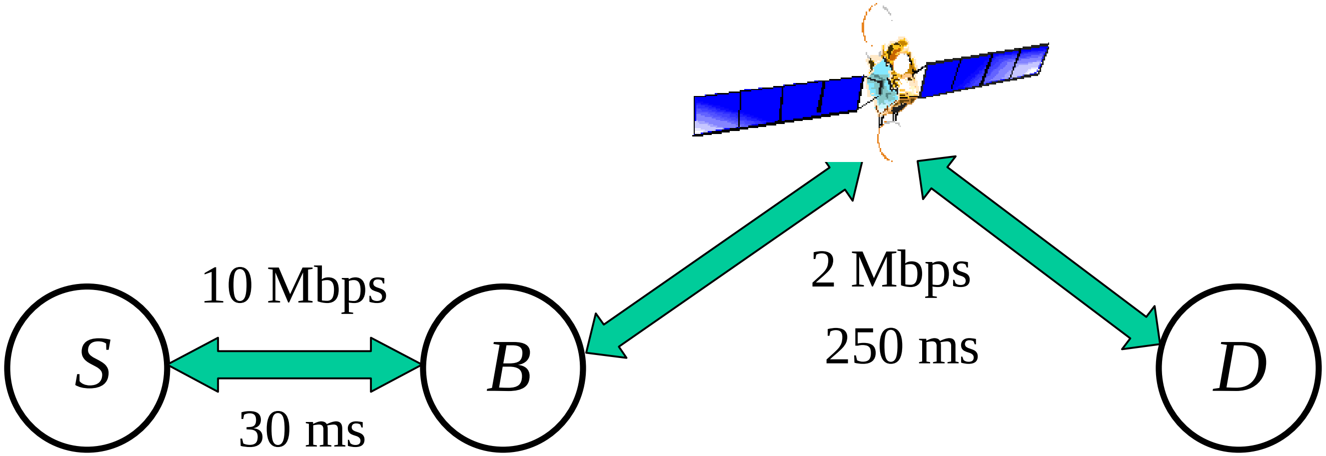
$$\bar{x} = \pi_G \bar{x}_G + \pi_B \bar{x}_B = M - \frac{M^2 \pi_B}{8\alpha d}$$

Difference from [Padhye et al. 1998]

We obtain more refined bounds by taking into account the bursty occurrence of losses

Model validation

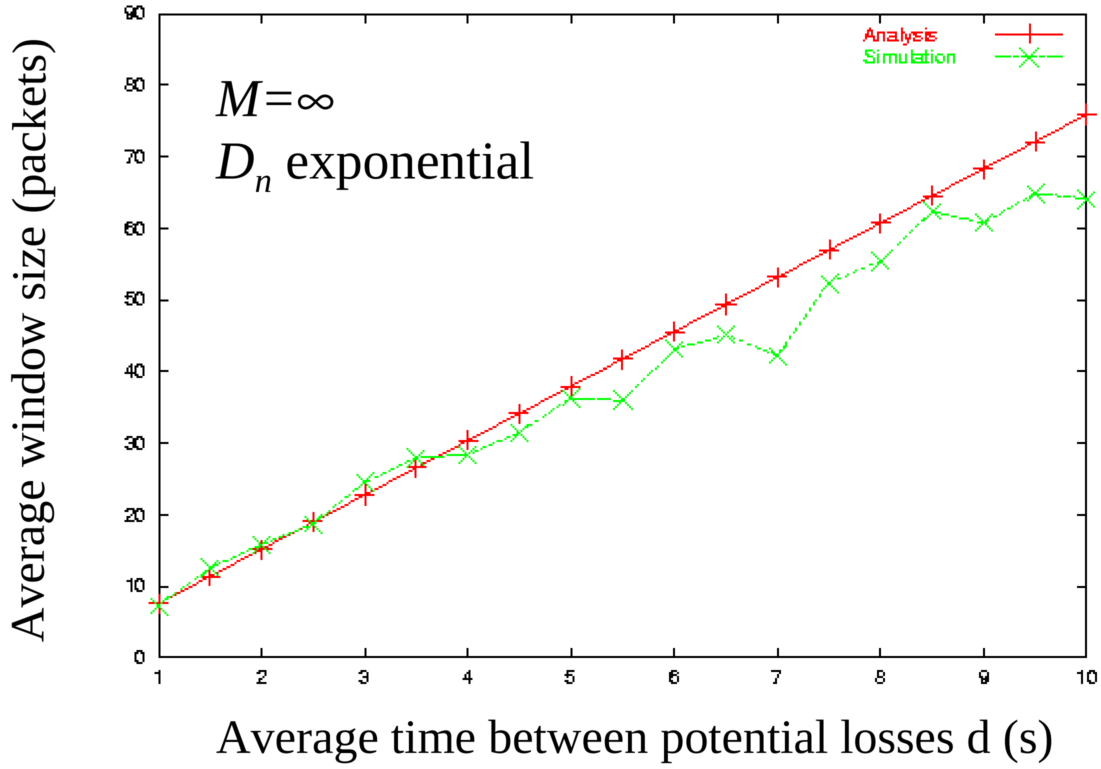
Add our model for losses to the `Error Model` object of the `ns` simulator and attach it to a satellite link



- A long ftp transfer (one hour)
- TCP version : SACK
- Packet size : 1000 bytes

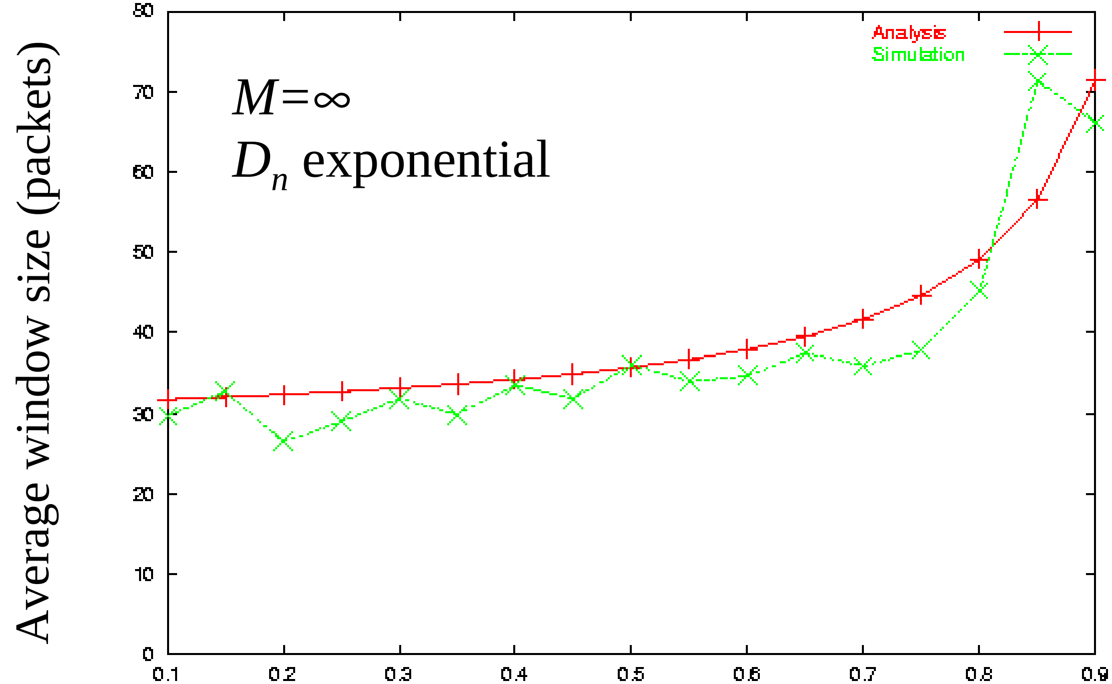
Variation of loss intensity

Fix *g* and *b* to 0.6 and change the inter-potential loss time



Variation of burstiness

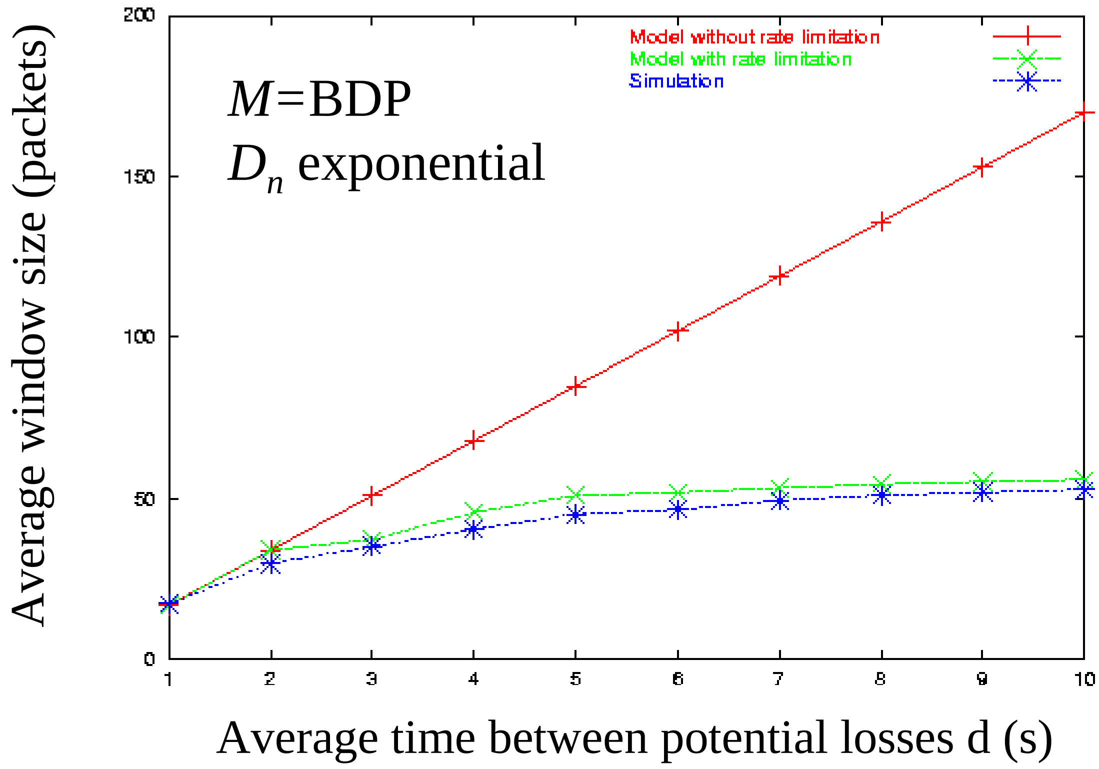
Fix d to 5 s and change b and g in a way that $\pi_G=\pi_B=0.5$



b = P[$Y_{n+1}=B$ | $Y_n=B$]

Variation of loss intensity

Fix g and b to 0.6 and change the inter-potential loss time



Conclusions

- A two-state Markovian model for TCP performance (or any other similar flow control mechanism) on bursty paths
- For the same average loss rate, the performance of TCP *increases* when loss events tend to appear in bursts

★ Current formulas are conservative in a bursty environment

Ongoing works

- Generalization of the model to multi-state paths
- Identification of the model parameters from real traces (uniformization technique)

  E. Altman, K. Avrachenkov, C. Barakat, P. Dube, "TCP over a multi-state Markovian path", under submission

- Study of TCP under a general (not only Markovian) loss process

  E. Altman, K. Avrachenkov, C. Barakat, "A stochastic model of TCP/IP with stationary random losses", to appear at ACM SIGCOMM