A Stochastic Model for TCP with Stationary Random Losses

Eitan Altman, Kostya Avrachenkov

Chadi Barakat

INRIA Sophia Antipolis - *France*

ACM SIGCOMM

August 31, 2000

Stockholm, Sweden

Outline

- Introduction
  - Some facts about TCP and loss events in the Internet
- Our models
  - A fluid model for rate evolution of TCP
  - A general model for losses
- Performance analysis
  - Throughput calculation
  - Specification of the result to particular loss processes
  - Introduction of Timeouts and bounds for the window limitation case
- Model validation and concluding remarks

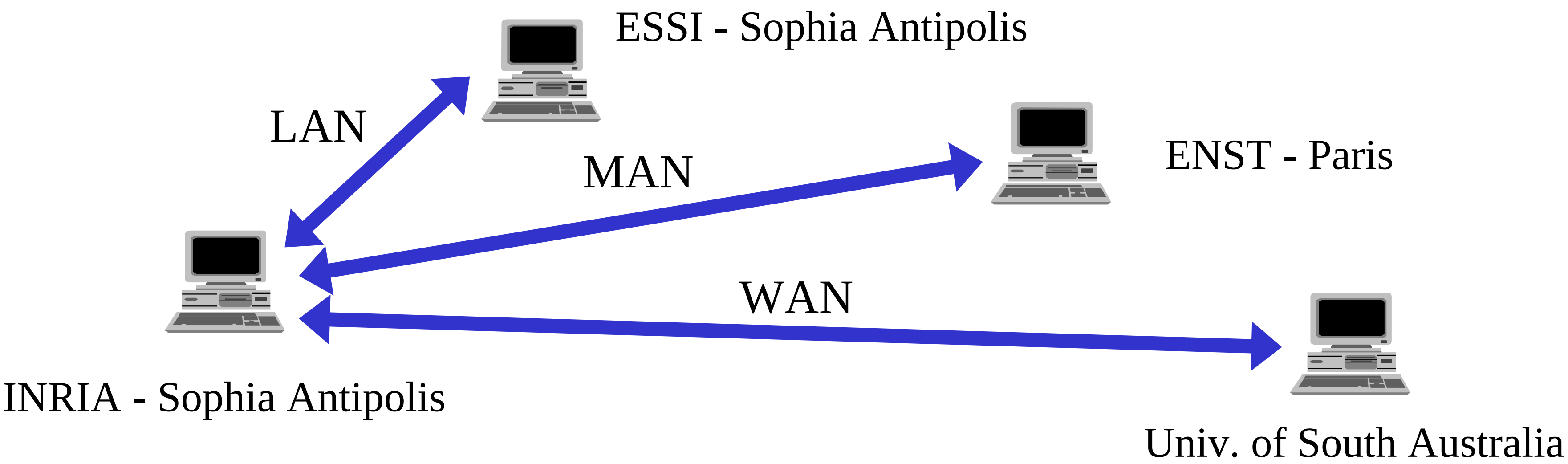
TCP and losses

- TCP congestion control
  - An additive-increase multiplicative-decrease strategy for congestion control in the Internet
  - Packet losses for congestion detection (possibly ECN)
- When looking at the rate level
  - **Loss event (or congestion event)**: An event that causes the reduction of the congestion window by a constant factor.
  - **Interpretation**: Depends on the version of TCP
    - One packet loss for Reno
    - One lossy round trip-time for Tahoe, New-Reno and SACK

Losses and TCP modeling

- TCP modeling requires a good characterization of times between loss events.
- Simple loss processes have been considered in the literature (Deterministic e.g., [Mathis, Semke, Mahdavi, 1997], Poisson e.g., [Misra, Gong, Towsly, 1999])
  - ➥ TCP throughput has been expressed only as a function of the average loss rate (e.g., p , λ)
- But, the Internet is so heterogeneous that losses may exhibit a more complicated distribution
  - ➥ What do our measurements tell us ?

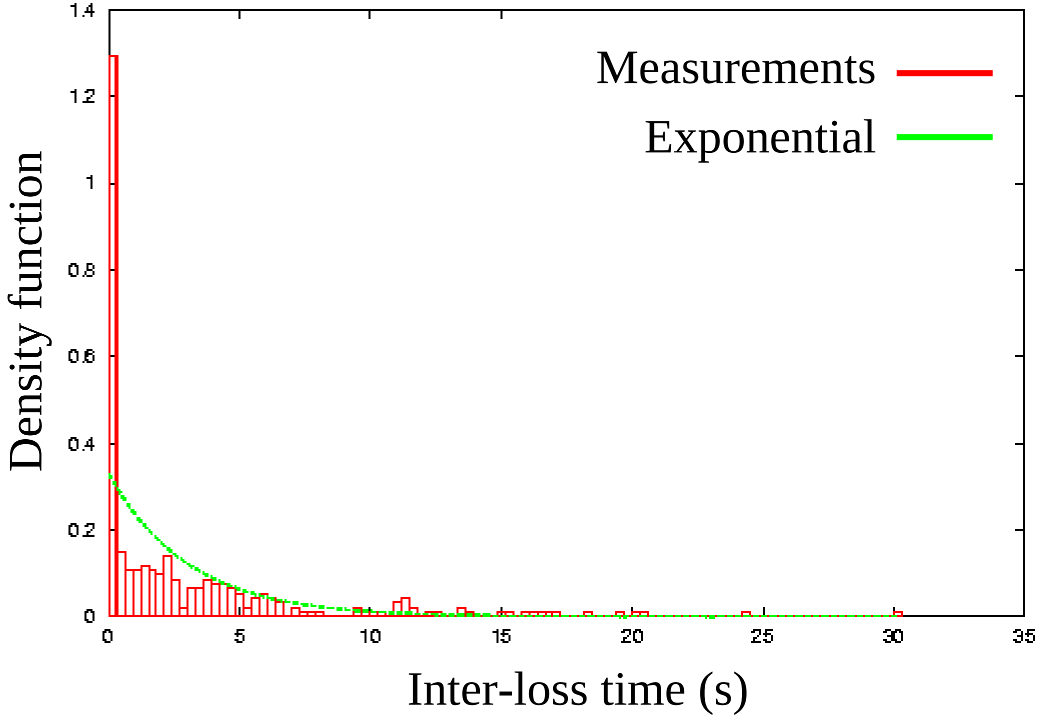
Measurement testbed



- Three long-life unlimited-data TCP transfers (New Reno version)
- Develop and run a tool at INRIA that detects loss events
- Store traces in separate files at fixed intervals (20, 40, and 60 min)

Some inter-loss time distributions

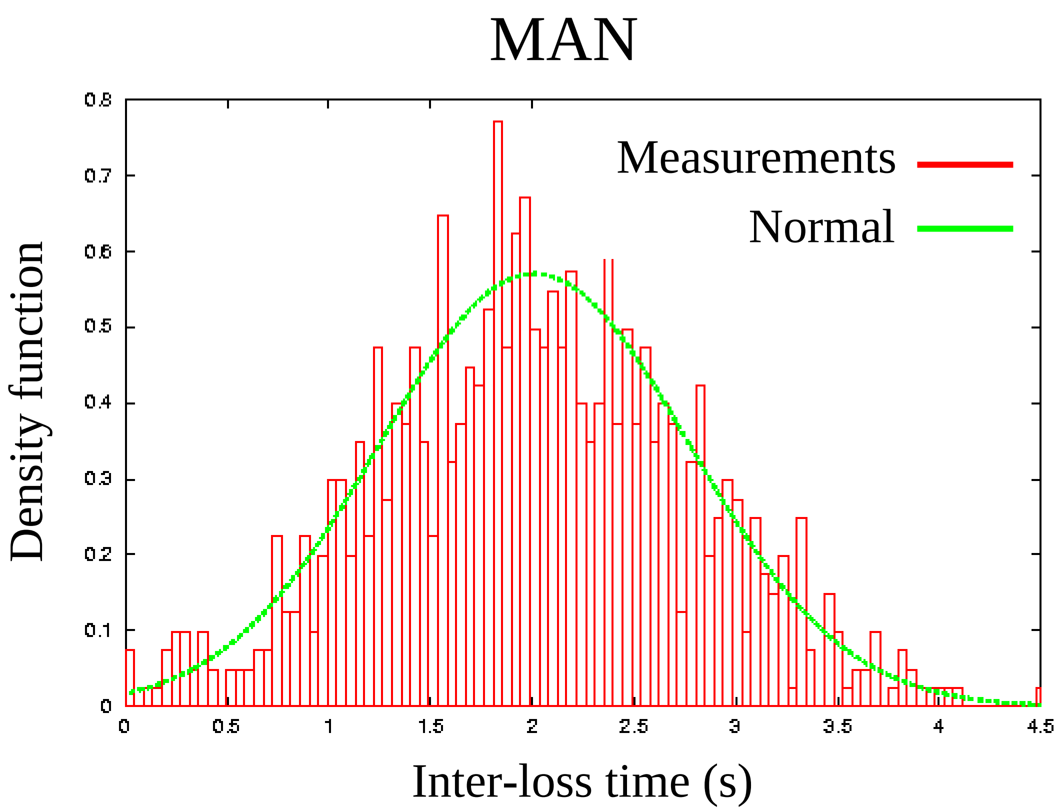
LAN



Highly bursty

INRIA

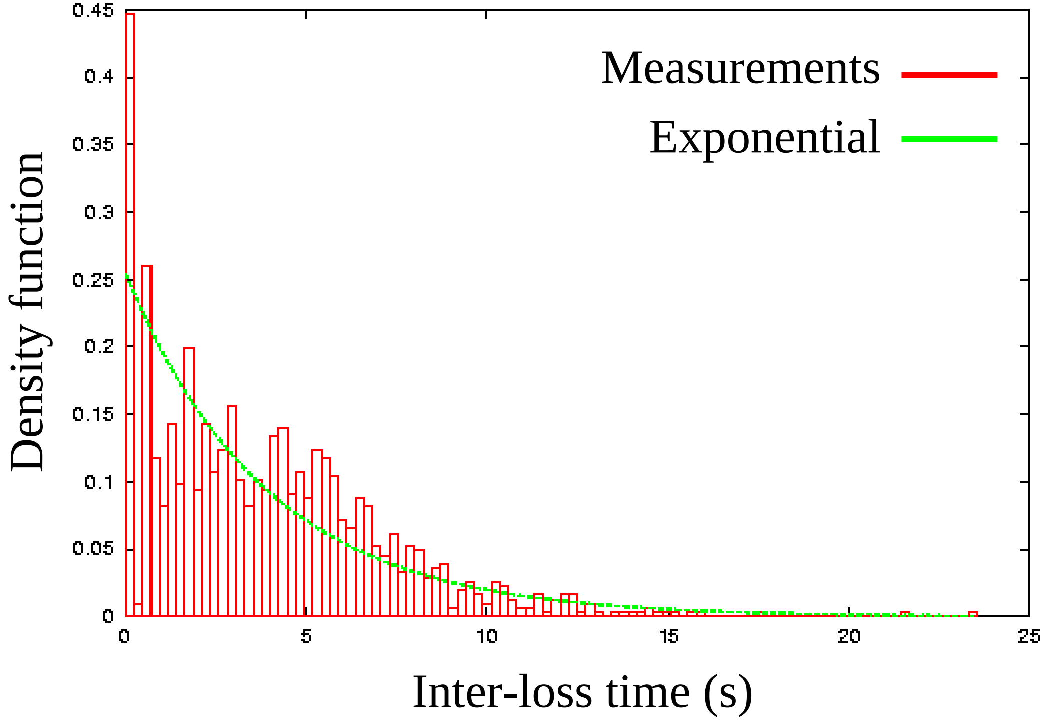
Some inter-loss time distributions



Close to Normal

Some inter-loss time distributions

WAN



Close to Poisson

Some correlation coefficients

LAN

Hour (Traces of 20 min)	Covariance coefficient $Cov(S_n, S_{n+1})/Var(S_n)$
11:00	+ 0.034
12:00	+ 0.041
12:30	+ 0.113
13:00	+ 0.001
13:30	- 0.191
14:00	- 0.078

WAN

Hour (Traces of 60 min)	Covariance coefficient $Cov(S_n, S_{n+1})/Var(S_n)$
11:00	- 0.197
12:00	- 0.001
14:00	- 0.102
16:00	- 0.107
20:00	+ 0.023
22:00	- 0.09

☞ Higher correlation is expected on paths with a significant memory (e.g., wireless links, satellite links).

Our model for losses

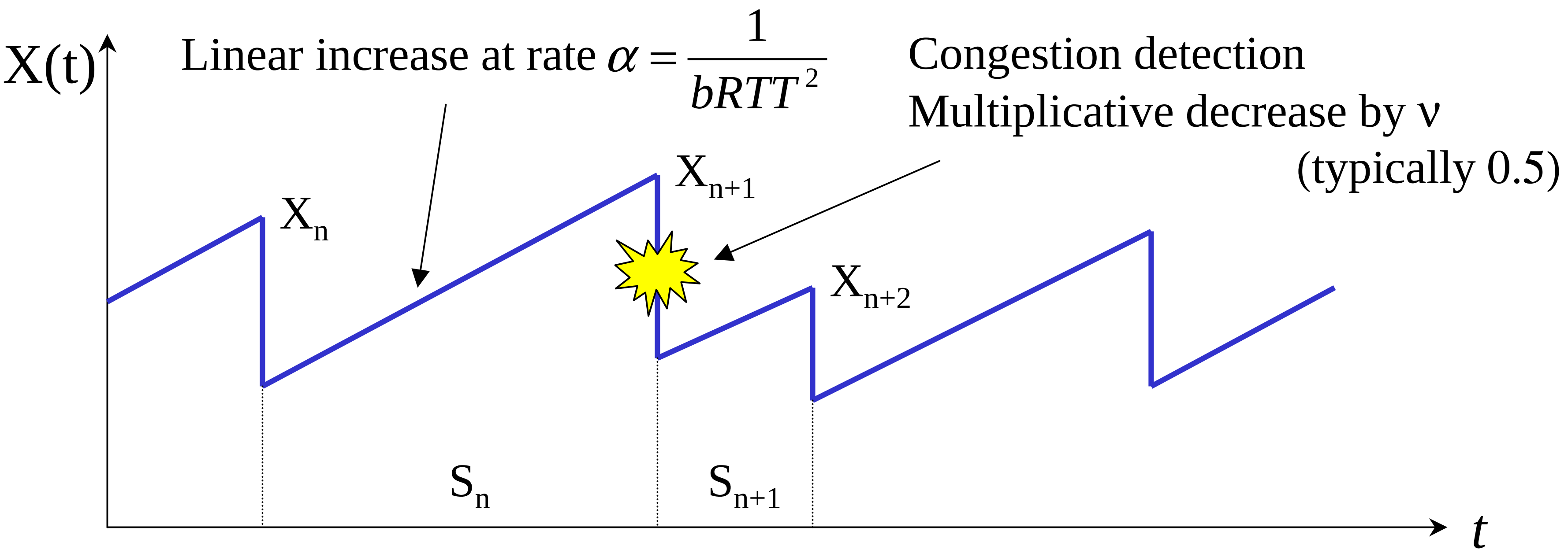
- Consider the loss events as a point process
  - Denote by $\{S_n\}$ the times between losses
- The *only* assumption we made:

$$\{S_n\} \text{ stationary and ergodic}$$

  - Notation:
    - Average inter-loss time: $d = 1/\lambda = E[S_n]$
    - Correlation functions: $R(k) = E[S_n S_{n+k}]$, $k=0,1,\ldots$
    - Covariance functions: $C(k) = R(k) - d^2$, $k=0,1,\ldots$
    - Normalized functions: $\hat{R}(k) = R(k)/d^2$, $\hat{C}(k) = C(k)/d^2$

Our fluid model for TCP

- Consider a long TCP transfer with infinite amount of data to send
- Denote by X(t) the rate of the connection at time t (=W(t)/RTT)



Performance analysis

- Input parameters
  - Parameters of the loss process
  - Parameters of the connection
- Output parameter
  - Throughput of the connection

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

Stochastic Difference Equation

Consider the rate between loss “n” and loss “n+1”

$$X_{n+1} = \nu \; X_n + \alpha \; S_n$$

Using Theorem 2A in [Glasserman and Yao,1995]

- 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

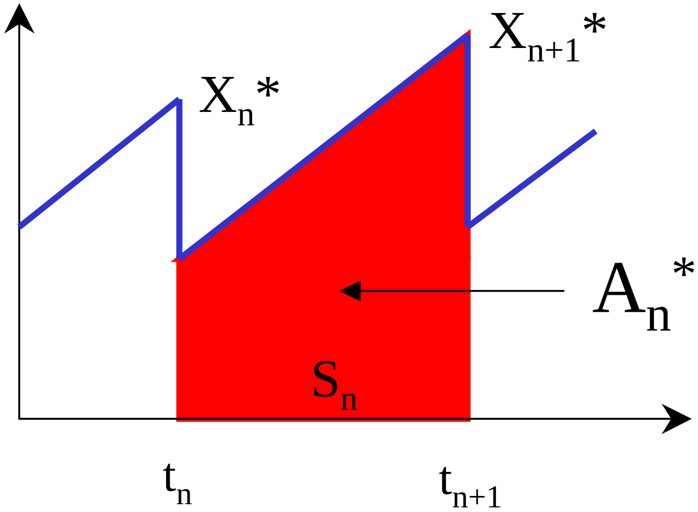
Calculation of the throughput

Using Palm theory

$$\overline{X} = E[X(t)] = \lambda E^0[A_n^*] = \lambda E^0[\nu X_n^* S_n + \tfrac{1}{2}\alpha S_n^2]$$

with

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



$$\overline{X} = \lambda\alpha[\tfrac{1}{2} R(0) + \sum_{k=1}^{\infty} \nu^k R(k)]$$

With packet drop probability

Let p denote the probability at which a packet is lost in average

$$\lambda = p\overline{X}$$

$$\overline{X} = \frac{1}{RTT\sqrt{bp}}\sqrt{\underbrace{\frac{1+\nu}{2(1-\nu)}}_{\text{Due to loss rate}}+\underbrace{\frac{1}{2}\widehat{C}(0)}_{\text{Due to variance}}+\underbrace{\sum_{k=1}^{\infty}\nu^{k}\widehat{C}(k)}_{\text{Due to correlation}}}$$

Specification of the model

- Deterministic losses: $$\bar{X} = \frac{1}{RTT}\sqrt{\frac{3}{2bp}}$$ (square root formula)

- Poisson loss process: $$\bar{X} = \frac{1}{RTT}\sqrt{\frac{2}{bp}}$$

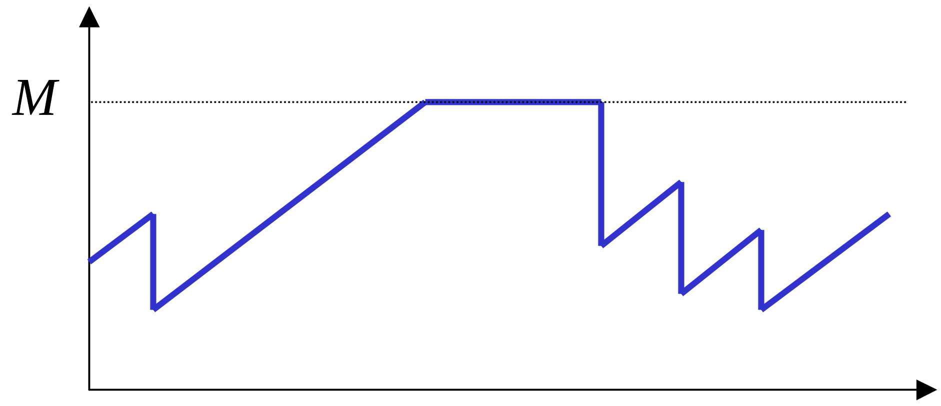
- General renewal process: $$\bar{X} = \frac{1}{RTT}\sqrt{\frac{1}{bp}\left(\frac{3}{2} + \frac{1}{2}\frac{Var(S_n)}{d^2}\right)}$$

- Markovian arrival process (see paper):

➥ The characteristics of the loss process in the connection depend on the state of the system which has a Markovian evolution. Loss events may occur between and during state transitions.

Case of rate limitation

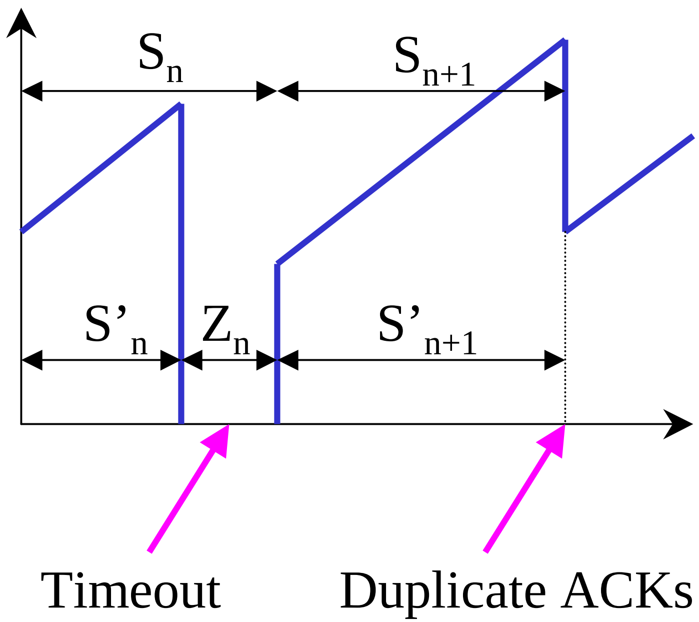
- TCP rate increase stops when the receiver window is reached.



- New SDE: $X_{n+1} = \min(\tfrac{1}{2} X_n + \alpha\, S_n, M)$
- The system converges to a stationary regime, but obtaining an explicit expression of the throughput for a general loss process seems to be impossible given the non-linearity of the model.
- We calculate bounds on the throughput which are also a good approximation of the throughput.

Introduction of Timeouts

☞ TCP may stay idle for a long time before the detection of a loss.



- $\overline{X}' =$ Average rate when excluding TO intervals

- $\overline{X} = \overline{X}'(1 - \lambda_{TO} E[Z_n \mid Z_n > 0])$

or

$$\overline{X} = \frac{\overline{X}'}{1 + p\overline{X}' Q(p) Z(p)}$$

$Q(p) = P(Z_n > 0)$

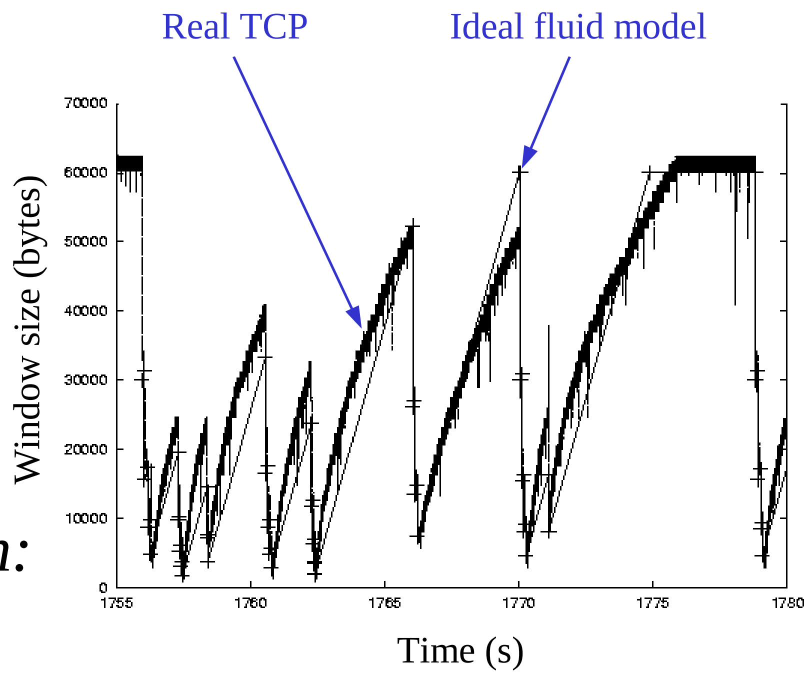
$Z(p) = E[Z_n \mid Z_n > 0]$

Validation of the model

- Instead of validating once the overall model, we chose to validate separately:

  - *The model for losses:*
    We compare the results of our model to a reference throughput, the one obtained by a fluid LIMD flow control mechanism. We construct this fluid model using the traces of a real TCP connection.

  - *The model for TCP rate evolution:*
    The main objective is to validate the assumption on the linearity of rate increase.



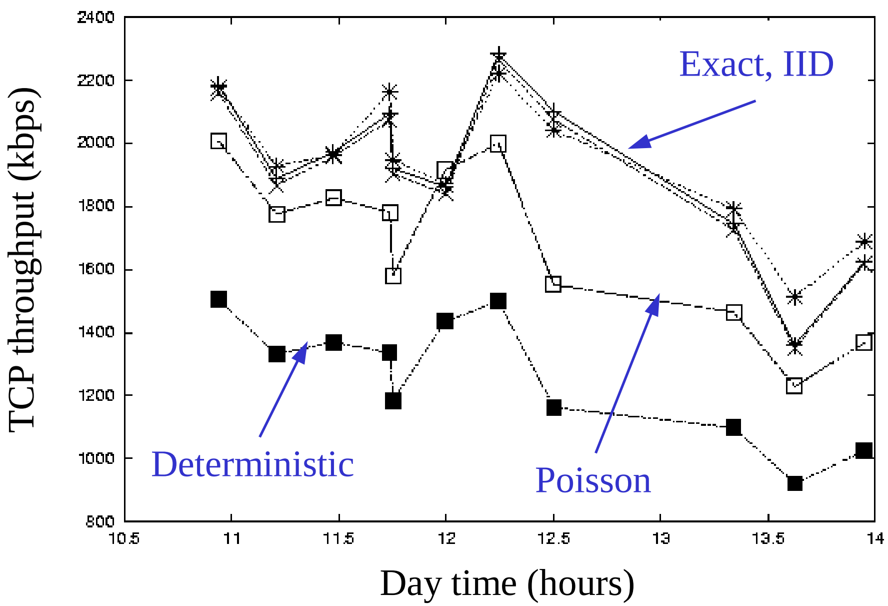
Impact of inter-loss times

- With a small number of correlation coefficients, we are able to estimate correctly the *reference* throughput.
- The correlation in our experimentation has not a great impact. A renewal model for losses is quite sufficient.
- The estimate of the throughput increases with the variance of the inter-loss times:
  - On LAN, deterministic and Poisson losses lead to an underestimation.
  - On MAN, Poisson losses lead to an overestimation.
  - On WAN, deterministic losses lead to an underestimation.

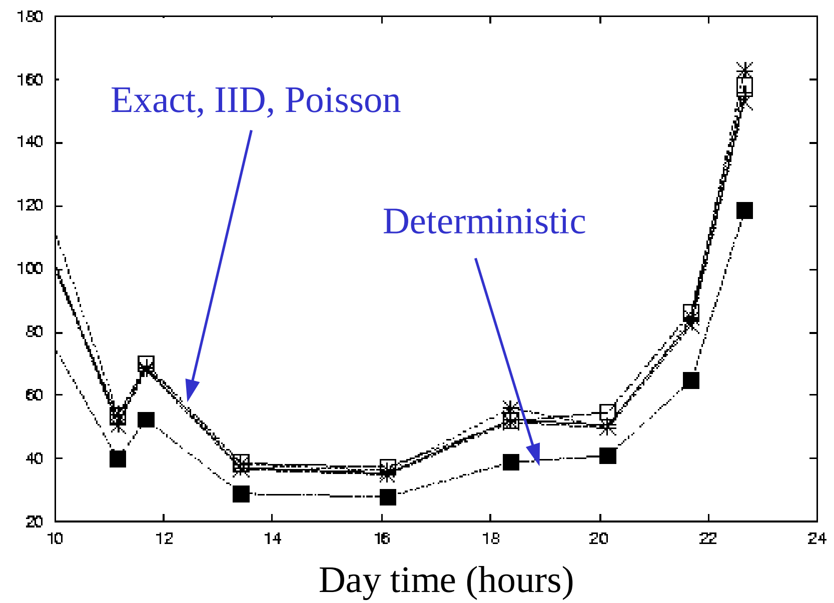
Impact of inter-loss times

An example of results: Throughput evolution during a day ...

LAN

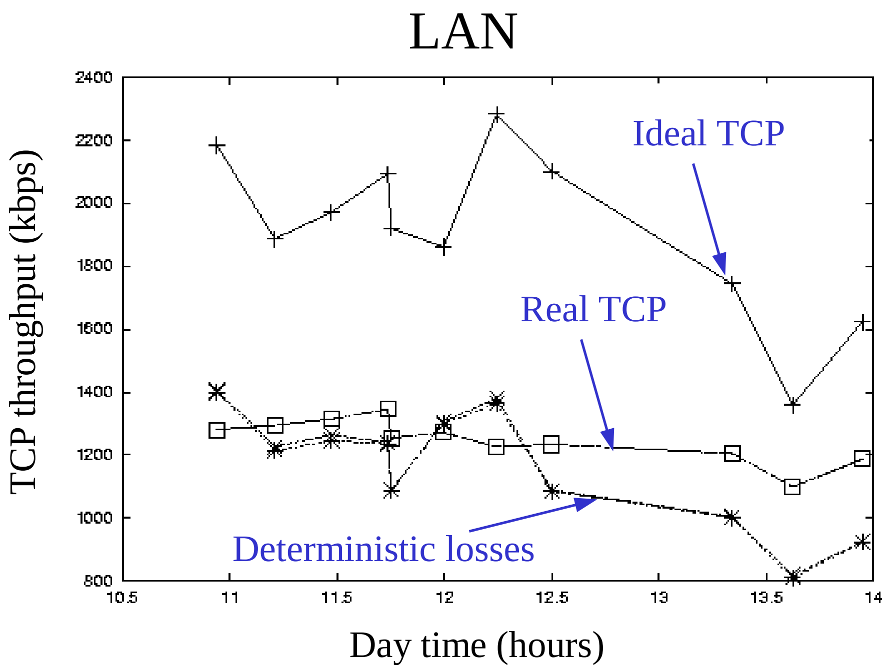


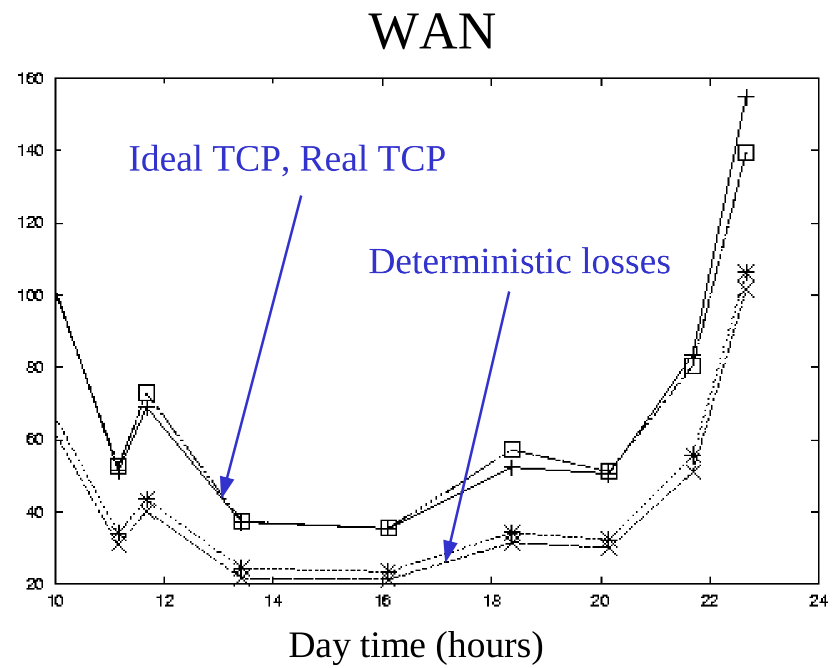
WAN



Did we model TCP correctly?

☞ We compare the reference throughput to the real one ...

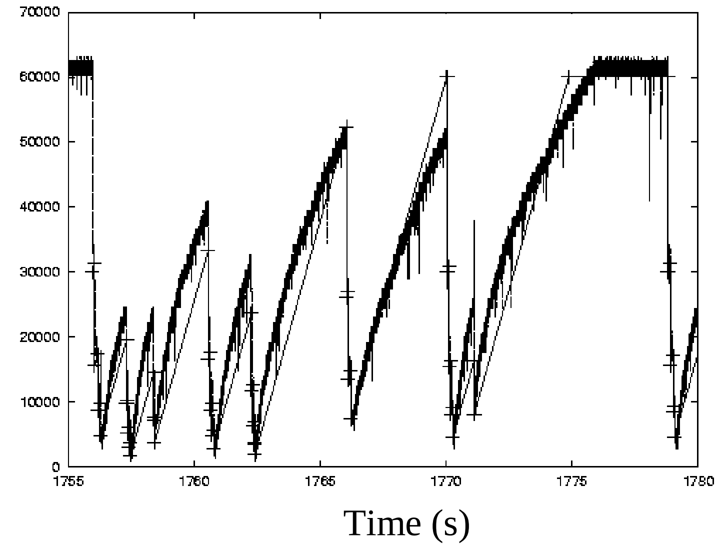




➥ Good approximation on WAN, but a significant overestimation when the window size is comparable to the bandwidth-delay product (e.g., LAN)

Sublinearity of TCP rate increase

- Main cause of the overestimation we observed in LAN:
  - TCP window increases linearly as a function of RTT number rather than time.
  - In LAN in particular, and on paths where the window size is comparable to the bandwidth-delay product in general, the increase in TCP window results in an increase in RTT (once the bottleneck link is fully utilized).
  - A correlated increase between window and RTT results in a sublinear increase in window and rate (the rate stabilizes if the connection is running alone on the path).

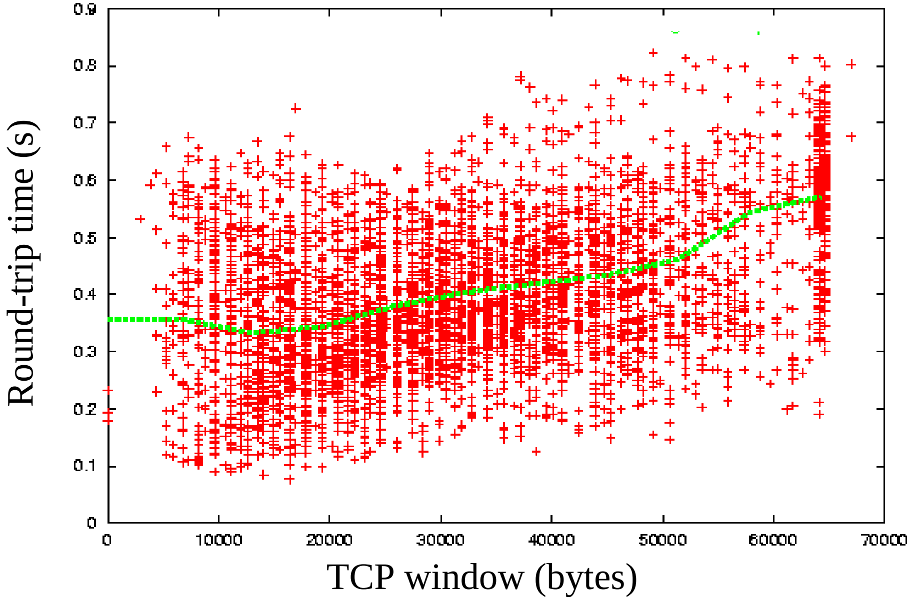


  - A sublinear rate increase leads to less throughput than a linear one.

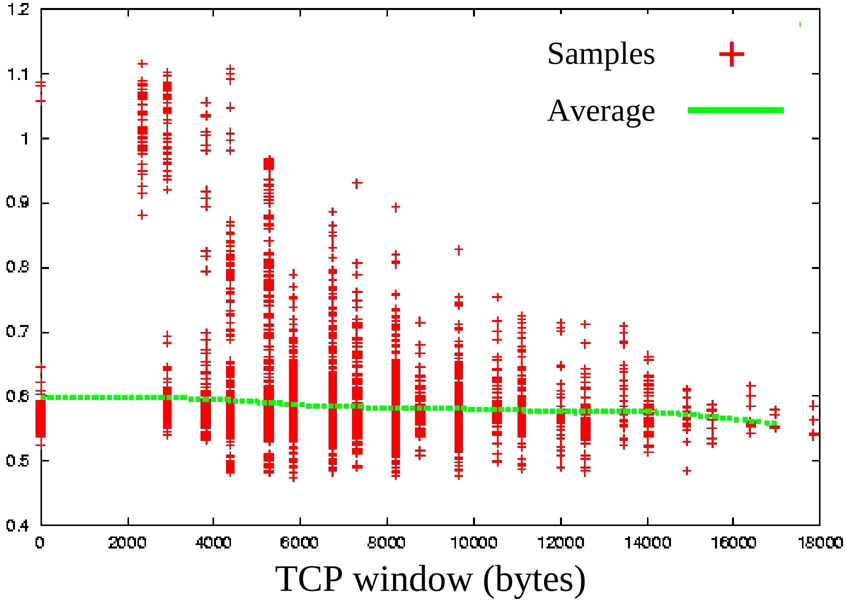
Window and RTT

For a single trace file

LAN

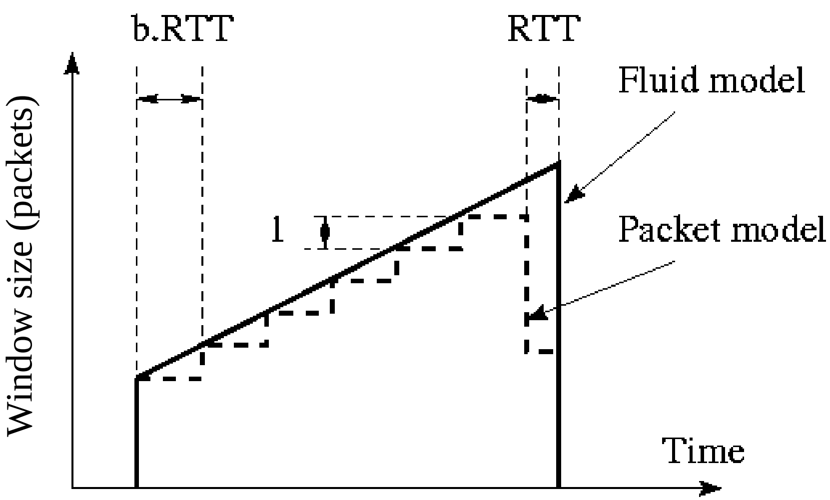


WAN

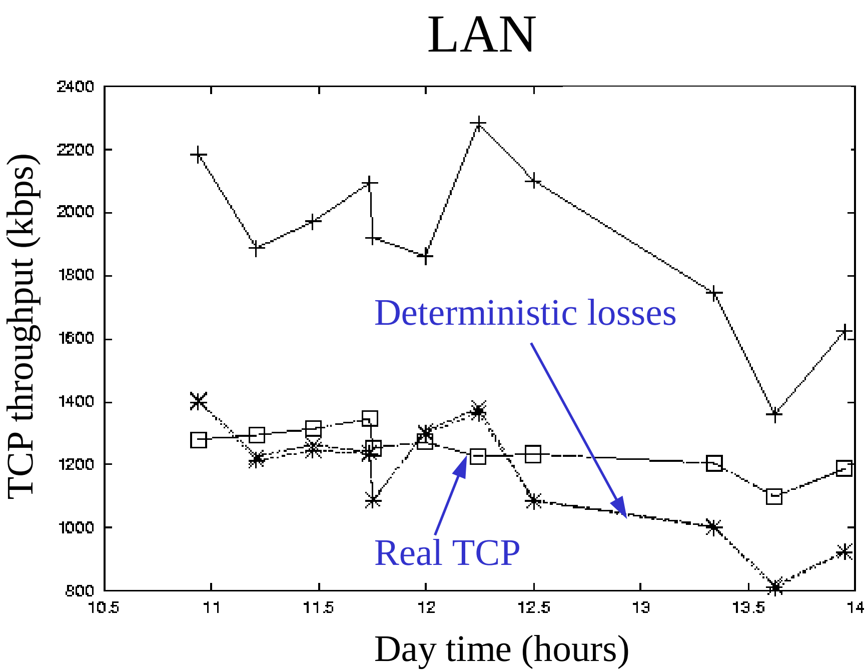


- Linear window increase models often used in the literature work well when the window and the RTT are *independent* from each other.
- Existing models need to be extended to the case when they are dependent.

Correction for the fluid assumption



$$\overline{X}_{dis} = \overline{X} - \frac{1}{2}RTT - \frac{E[X_n^*].RTT}{2E[S_n]}$$



- With this correction, our model with *deterministic* inter-loss times gives the same performance as the detailed packet level model in [Padhye,Firoiu,Towsly,Kurose,1998]

INRIA

Summary

- A good modeling of TCP requires:
  - A good characterization of inter-loss times. The second moment and some correlation coefficients need to be found.
    - These coefficients can be measured or calculated from another model for the network (e.g., as a function of the packet drop probability p).
  - A good characterization of TCP window evolution. The dependency between window and RTT needs to be considered.
- The errors introduced by these two characterizations may be in opposite directions and *cancel* each other, resulting in an overall throughput close to the real one (e.g., LAN, assuming that inter-loss times are deterministic).