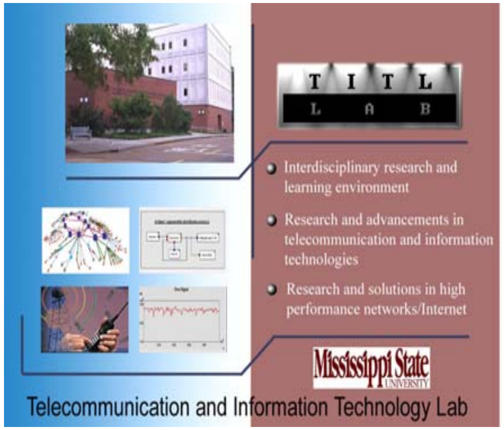
On the Modeling of TCP Latency and Throughput



TITL LAB

- Interdisciplinary research and learning environment
- Research and advancements in telecommunication and information technologies
- Research and solutions in high performance networks/Internet

Mississippi State University

Telecommunication and Information Technology Lab

Dong Zheng

MASTER'S THESIS PRESENTATION

Department of Electrical and Computer Engineering
Mississippi State University

Outline

- Introduction and Motivation
- Background Information on TCP
- Building the Stochastic Models
- Model Validation by Simulation
- Conclusion and Future Work

What Is The Problem ?
- From Practical View

- TCP's performance dominates behavior of Internet traffic inspiring tremendous research on stochastic TCP model

1. *improve TCP performance by understanding the sensitivity of TCP performance to the network conditions*
2. *help design of active queue management*
3. *aid in the design of TCP-friendly transfer multicast protocols*

— ***An accurate model of TCP performance is needed***

What Is The Problem ?

-From the Model’s View

- Most existing models doesn’t include the analysis of time-outs effects
- Models including the analysis of time-outs underestimate it
- None of the existing steady state model include the slow start phase
- Not accurate modeling of the delayed acknowledgment’s effect in the slow start phase

— ***New coupled models are needed***

What Is This Research All About?

- Develop better and tractable model for slow start
- Develop complete steady state model including the slow start phase
- Develop accurate model for short-lived TCP flows

Why Include the Slow Start?

- Slow start phase begins whenever TCP recovers from time-out phase
- Empirical studies observed that slow start phase occurs often for long-lived TCP flows
- Models that ignored slow start overestimate TCP performance

— ***Including slow start phase into steady state analysis results in accurate performance predictions***

Why Need New Models for short-lived TCP connections?

- 85% of TCP traffic are short-lived flows
- Connections ends while in slow-start phase
  - never enter congestion avoidance
  - steady-state model doesn’t apply

TCP Features

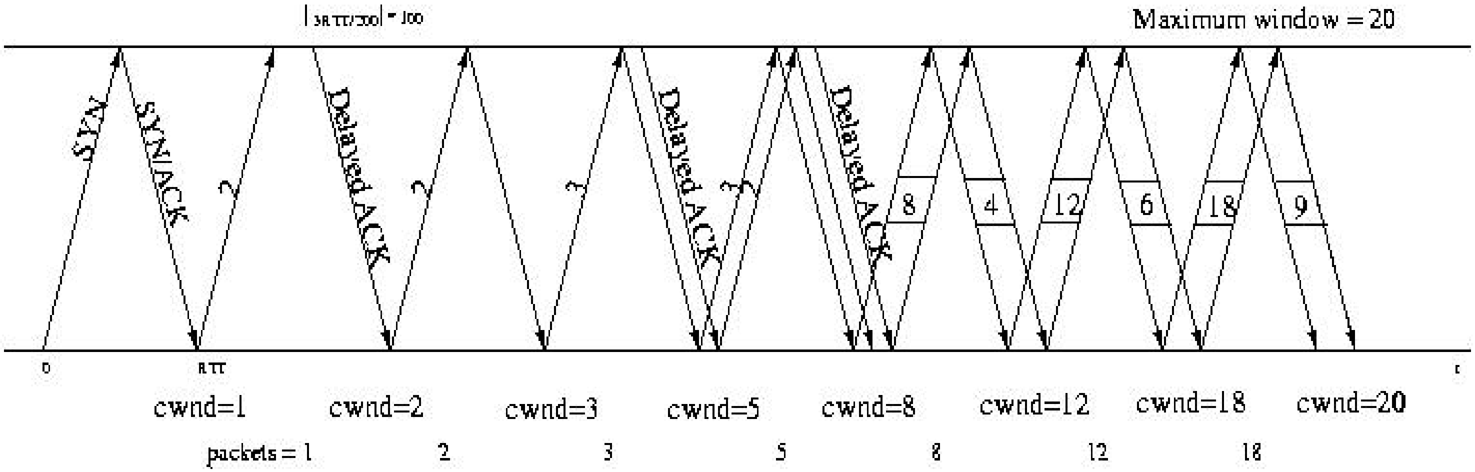
- Connection oriented
  - Explicit and acknowledged connection establishment
- Reliable stream exchange
  - every packet has sequence number
  - acknowledging the receipt of the right packet (usually delayed)
  - set retransmission timer for every packet sent
- Congestion control

Slow Start and Congestion Avoidance

If current congestion window (cwnd) is less than slow start threshold (ssthresh)	If (cwnd < ssthresh)
TCP is in slow start phase, and increase the cwnd exponentially	cwnd = cwnd + 1;
Otherwise in congestion avoidance mode, and cwnd increases linearly	Else cwnd += 1/cwnd;

A Typical TCP Connection

(No Loss Happens)



Steady State Model

- Assumptions

- Based on TCP Reno release from Berkeley
- High link speed
- Fixed packet size

– Congestion window alone determines the send rate

Steady State Model

- Assumptions (Continued)

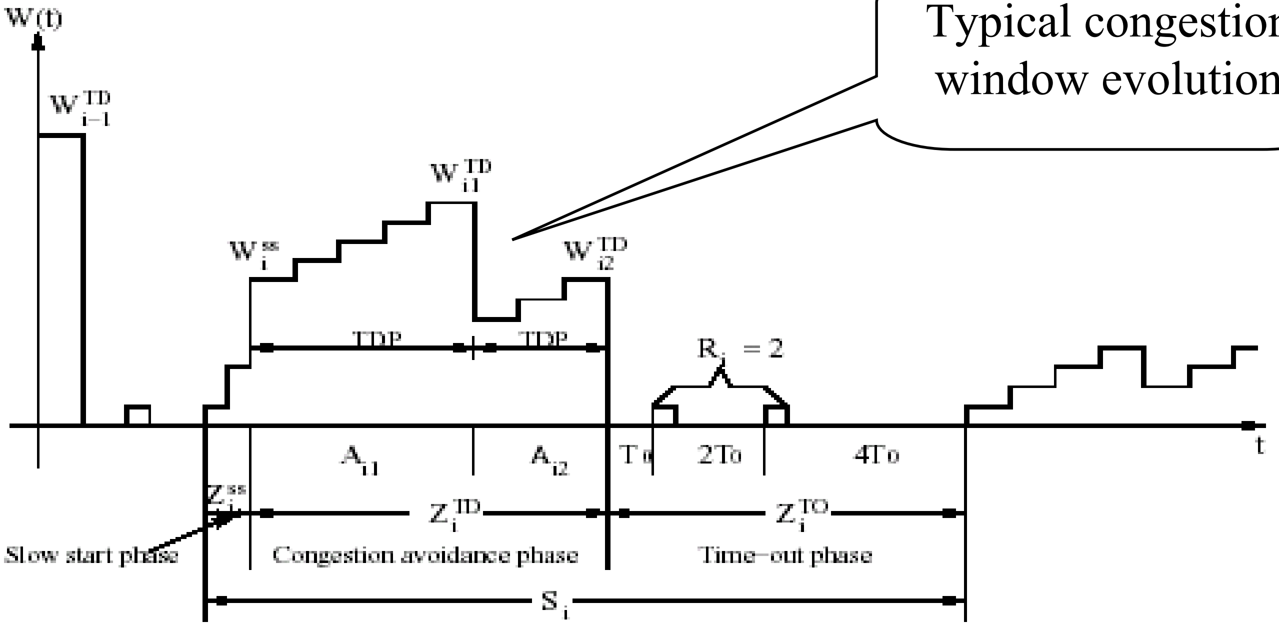
- Modeling dynamics of TCP in terms of “rounds”

— starts when a window of packets is sent and ends when one or more acknowledgments are received

- Delayed acknowledgment algorithm applied
- Packet losses in accordance with bursty loss model

— Packet losses are correlated in each round but independent between rounds

Steady State Model



Typical congestion window evolution

Steady State Model

Let M_i be the number of packets sent during the total time S_i:

$$M_i = Y_i^{SS} + \sum_{j=1}^{n_i} Y_{ij} + R_i$$

$$S_i = Z_i^{SS} + \sum_{j=1}^{n_i} A_{ij} + Z_i^{TO}$$

Assuming (M, S) to be sequences of i.i.d. random variables, the send rate is:

$$B = \frac{E[M]}{E[S]}$$

Steady State Model

Considering n_i to be i.i.d. random variables and independent of Y_{ij}, we have:

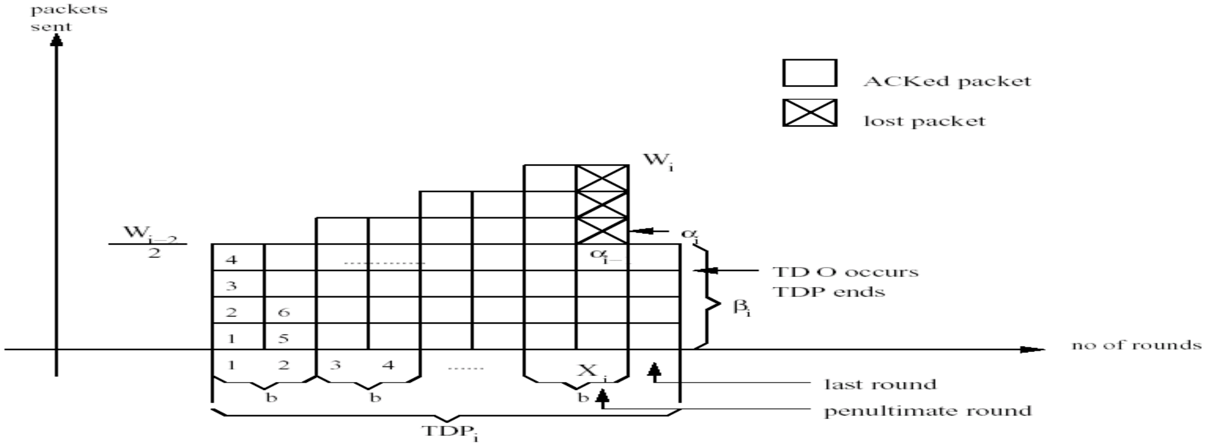
$$B = \frac{E[Y^{SS}] + E[n]E[Y] + E[R]}{E[Z^{SS}] + E[n]E[A] + E[Z^{TO}]}$$

Slow Start Phase

- Congestion window growth pattern is: $cwnd_i = [\frac{cwnd_{i-1}}{2}] + cwnd_{i-1}$
- The total number of packets sent in first n rounds : $Y_n^{SS} = \sum_{i=1}^{n} cwnd_i$

Number of packets sent	$E[Y^{SS}] = \frac{E[W^{TD}]g^2}{2} - 2$
Time duration	$E[Z^{SS}] = \log_g(\frac{E[W^{TD}]}{2C_1}) * RTT$

Congestion Avoidance Phase



Number of packets sent	$E[Y]=\frac{E[X]}{2}(\frac{E[W^{TD}]}{2}+E[W^{TD}]-1)+E[\beta]$
Time duration	$E[A]=RTT(E[X]+1)$

Congestion Avoidance Phase (continued)

- Expected congestion window size:

$$E[W^{TD}] = -\frac{2(b-2p)}{3} + \sqrt{\frac{4(bp+2(1-p^2))}{3bp} + \left(\frac{2b-4p}{3b}\right)^2}$$

- Number of packets sent in the fast retransmit:

$$E[\beta] = (E[W^{TD}]-1)(1-p)$$

- Number of rounds in TDP:

$$E[X] = b\left(\frac{E[W^{TD}]}{2}+1\right)$$

Time-outs Phase (continued)

- Padhye’s steady-state model use:

  1/E[w] = E[1/w]

- Not so good approximation:

$$E[(\frac{1}{\sqrt{W}})(\sqrt{W})]^2 \le E[(\frac{1}{\sqrt{W}})^2]E[(\sqrt{W})^2]$$

$$\Rightarrow \qquad \frac{1}{E[W]} \le E[\frac{1}{W}]$$

- Better approximation:

$$E[\frac{1}{W}] \approx \frac{1}{E[W]}(1+\frac{Var(W)}{E[W]^2})$$

Probability of Packet Loss Resulting in time-out

$$
\begin{aligned}
Q^{TD} &= E[Q^{TD}(w)] \\
&= E[\min(1, \frac{3}{w})] \\
&= \min(1, 3E[\frac{1}{W^{TD}}]) \\
&\approx \min(1, \frac{3\sqrt{3}}{E[W^{TD}]})
\end{aligned}
$$

Send Rate and Throughput

- Send Rate:

— is the number of packets sent per seconds

- Throughput:

— is the number of packets received per seconds

From Padhye's model:

Number of TDPs in a congestion avoidance phase	$E[n]=\frac{1}{Q^{TD}}$
Number of packets sent in the time-out phase	$E[R]=\frac{1}{1-p}$
Time spent in the time-out phase	$E[Z^{TO}]=T_0\frac{f(p)}{1-p}$

Send Rate

$$B = \begin{cases} \dfrac{\dfrac{E[W^{TD}]g^{2}}{2} - 2 + \dfrac{1}{Q^{TD}(E[W^{TD}])}(\dfrac{1-p}{p} + E[W^{TD}]) + \dfrac{1}{1-p}}{(\log_{g}(\dfrac{E[W^{TD}]}{2C_1}) + \dfrac{1}{Q^{TD}(E[W^{TD}])}(\dfrac{bE[W^{TD}]}{2} + b + 1))RTT + \dfrac{f(p)T_0}{1-p}} \\ \\ \qquad\qquad \mathit{When} \qquad E[W^{TD}] < W_m \\ \\ \dfrac{\dfrac{W_{mg}^{2}}{2} - 2 + \dfrac{1}{Q^{TD}(W_m)}(\dfrac{1-p}{p} + W_m) + \dfrac{1}{1-p}}{\log_{g}(\dfrac{W_m}{2C_1})RTT + \dfrac{1}{Q^{TD}(W_m)}((\dfrac{b}{8}W_m + \dfrac{1-p}{pW_m} + 2) + 1)RTT + \dfrac{f(p)T_0}{1-p}} \\ \\ \qquad\qquad \mathit{When} \qquad E[W^{TD}] \ge W_m \end{cases}$$

Throughput

To obtain throughput, changes are needed:

$E[Y] \longrightarrow E[Y'] = E[\alpha]+E[\beta]-1$

The number of packets that have been sent in a TDP → The number of packets that have been received in a TDP

$E[R] \longrightarrow E[R'] = 1$

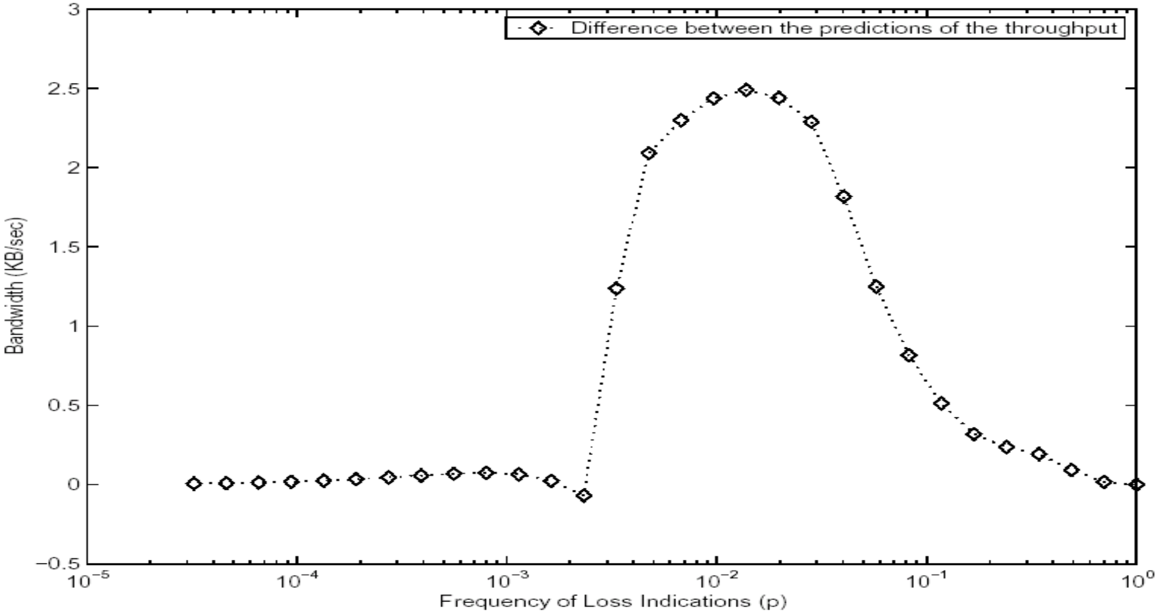
The expected number of packets sent in the time out phase → The expected number of packets received in the time out phase

Throughput (continued)

$$H = \begin{cases} \dfrac{\dfrac{E[W^{TD}]g^{2}}{2} - 2 + \dfrac{1}{Q^{TD}(E[W^{TD}])}\left(\dfrac{1-p}{p} + (E[W^{TD}]-1)(1-p)\right) + 1}{\left(\log_{g}\left(\dfrac{E[W^{TD}]}{2C_{1}}\right) + \dfrac{1}{Q^{TD}(E[W^{TD}])}\left(\dfrac{bE[W^{TD}]}{2} + b + 1\right)\right)RTT + \dfrac{f(p)T_{0}}{1-p}} \\ \qquad When \quad E[W^{TD}] < W_{m} \\ \\ \dfrac{\dfrac{W_{m}g^{2}}{2} - 2 + \dfrac{1}{Q^{TD}(W_{m})}\left(\dfrac{1-p}{p} + (W_{m}-1)(1-p)\right) + 1}{\log_{g}\left(\dfrac{W_{m}}{2C_{1}}\right)RTT + \dfrac{1}{Q^{TD}(W_{m})}\left(\left(\dfrac{b}{8}W_{m} + \dfrac{1-p}{pW_{m}} + 1\right) + 1\right)RTT + \dfrac{f(p)T_{0}}{1-p}} \\ \qquad When \quad E[W^{TD}] \geq W_{m} \end{cases}$$

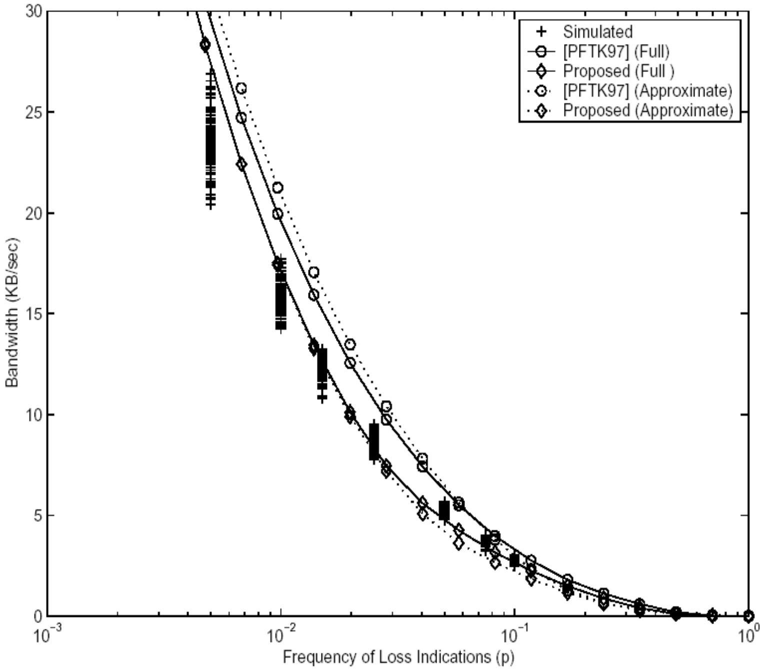
Comparison Example

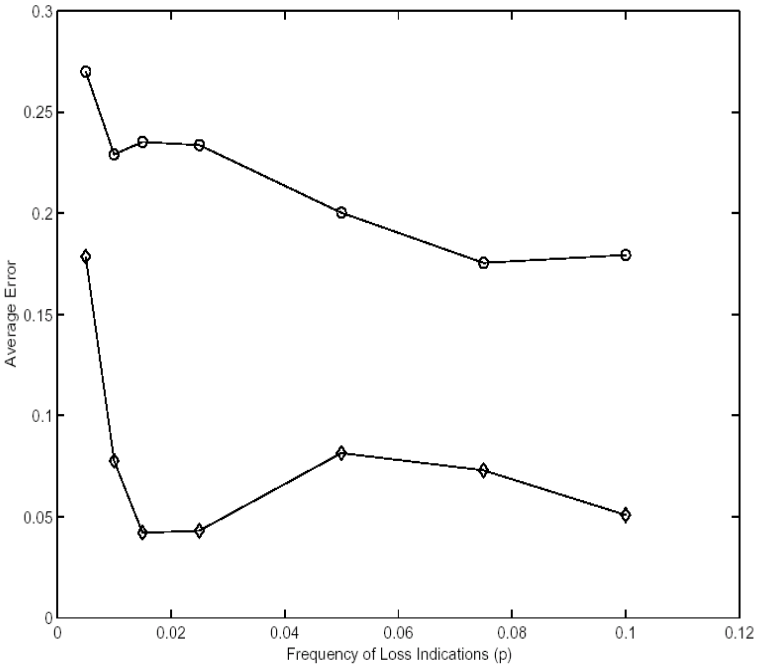
$H_{(Padhye)} - H$:



The conditions are: RTT=200ms, MSS=536Bytes, w_1=1segment, T_0=1s, W_m=20segments

Model Validation by Simulations





Short-lived TCP Connection Model

1. Initial three-way-handshake connection
   — modeled by Cardwell's paper
2. Initial slow start
   — same model used in steady state model
3. First loss
   — same analysis used for time-out phase
4. Subsequent losses
   — good approximation: Steady-state model

Short-lived TCP connection Model

- Time spent in initial slow start part:

$$E[n] = \begin{cases} [[\log_g(\frac{W_m}{C_1})] + \frac{1}{W_m}(E[Y_{init}] - g^2 W_m - 2)] \\ \qquad When \qquad E[W_{init}] > W_m \\ [\log_g(\frac{E[Y_{init}] + 2}{C_1})] - 2 \\ \qquad When \qquad E[W_{init}] \le W_m \end{cases}$$

- Time spent in the first loss part:

$$T_{loss} = (1-(1-p)^d)(Q_{init} E[Z^{TO}] + (1-Q_{init}) E[n_t])$$

- Time spent in the rest part:

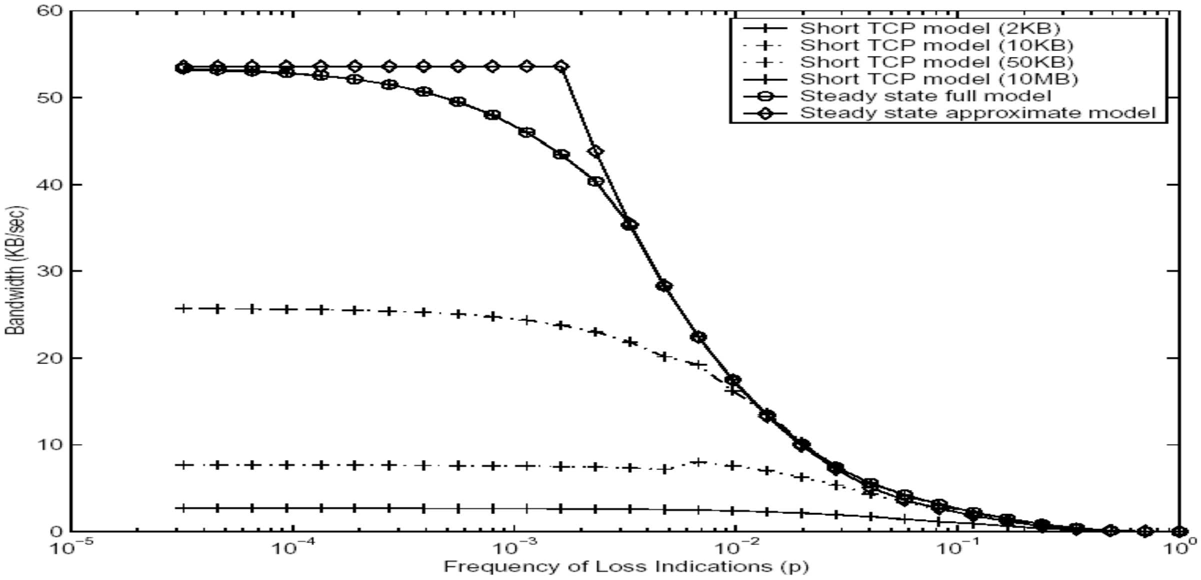
$$T_{rest} = \frac{d - E[Y_{init}]}{H} = \frac{dp - (1-(1-p)^d)(1-p)}{p * H}$$

Short-lived TCP Connection Latency

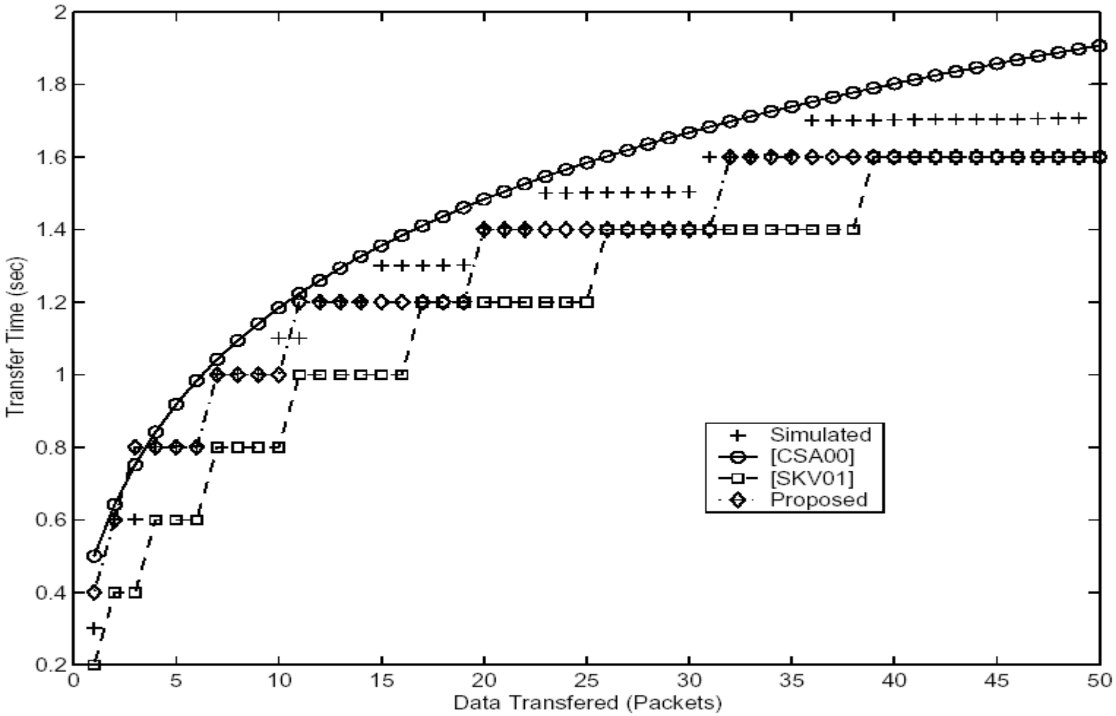
$$T_{latency} = E[T_{twhs}] + E[n]RTT + T_{loss} + T_{rest} + T_{delay} - \frac{RTT}{2}$$

- T_{delay} : caused by delayed acknowledgment for the first packet which is characterized by mean of 100ms
- Only half of a round is needed to send the last window of packets, so deduct the hlaf round trip time from the total latency.

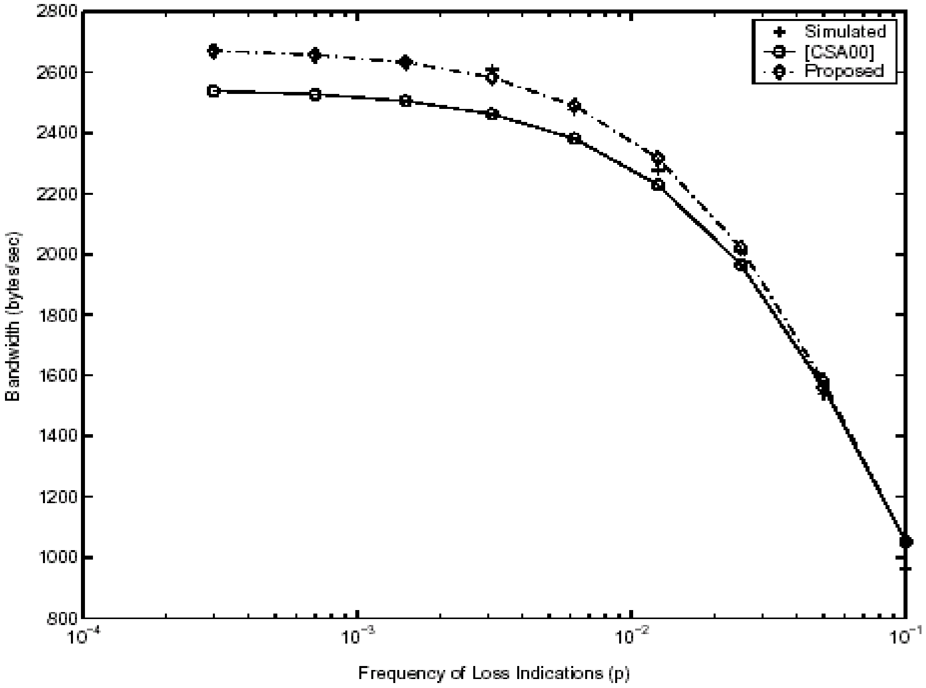
Short-lived TCP model
—> Steady state model



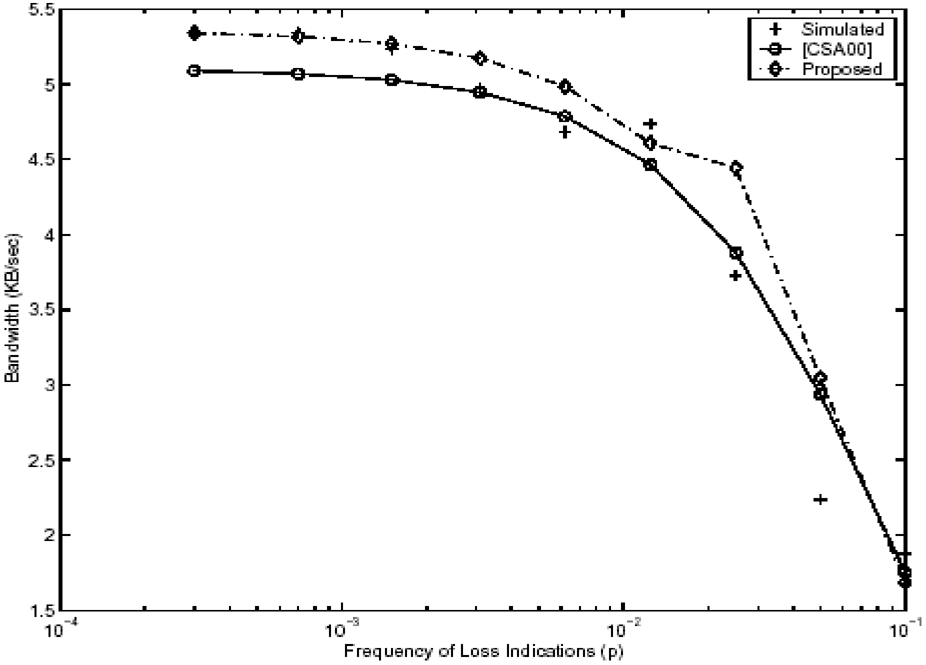
Latency of Short-lived TCP Connection



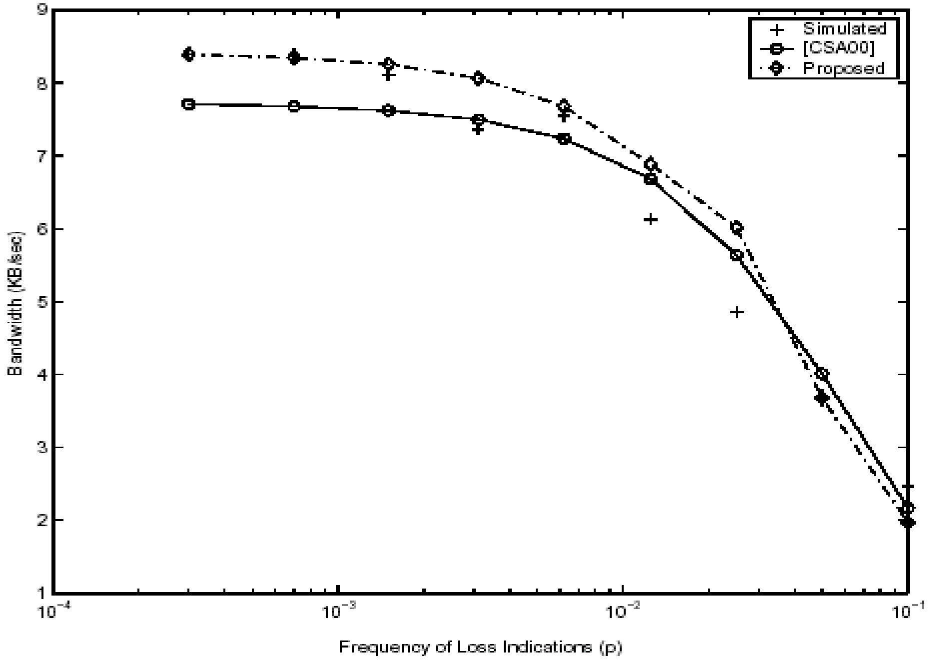
Throughput vs. Loss Rate (File Size = 2KB)



Throughput vs. Loss Rate (File Size = 6KB)



Throughput vs. Loss Rate (File Size = 11KB)



TITL

Comparison of the Average Error

Loss Rate	P = 0	$3\times10^{-3} \sim 10^{-1}$		
File Size	0.5~26KB	2KB	6KB	11KB
[CSA00]	9.40%	4.08%	6.43%	8.38%
Proposed	5.83%	0.59%	7.54%	7.64%

Conclusions

- Propose new model for the slow start phase
  - Based on discrete equation
  - Using results from Fibonacci sequence
- Develop complete steady state model
  - Integrate slow start phase
  - Accurate time-out analysis
- Develop accurate short-lived TCP model
  - Using same analysis of slow start
  - New estimate time-out analysis

1. Future Work

- Considering effect of fast recovery

  — will help building a more accurate model

- Analyze effects of different loss models to TCP’s performance

  — help design different queuing methods

- Find probability distribution of latency

  — better than the expected value