An Overview and Comparison of Analytical TCP Models

Inas Khalifa and Ljiljana Trajkovic

Communication Networks Laboratory
http://www.ensc.sfu.ca/research/cnl
School of Engineering Science
Simon Fraser University

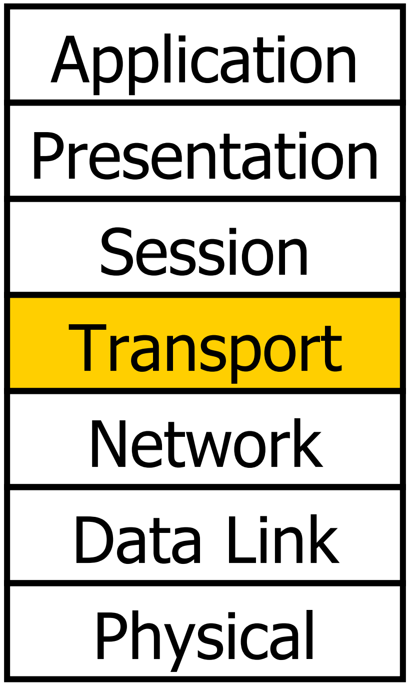
cnl
communication
networks
laboratory

Road map

- Introduction
- Motivation and objectives
- Overview of TCP mechanisms
- Model classification criteria
- Survey a set of analytical models
- Model comparison
- Summary and conclusions

Introduction

- Transmission Control Protocol (TCP) is a transport protocol
- TCP provides a reliable connection-oriented data service in packet-switched networks.
- Applications: WWW, E-mail, file transfer, remote login, database access, X-windows



OSI Model

Motivation and Objective

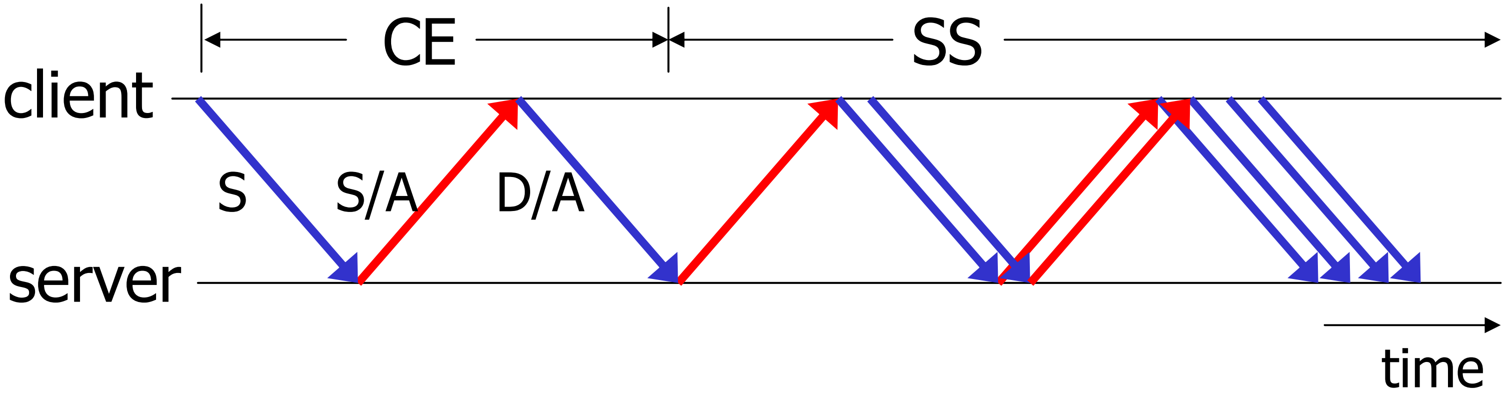
- Most network traffic is carried by TCP
- Analytical models help: evaluate TCP implementations, investigate TCP interaction with queue management algorithms, and define TCP-friendly behaviour
- Objectives:
  - examine the TCP modelling environment
  - compare several analytical TCP models
  - identify the missing features

Overview of TCP

- TCP provides reliable connection-oriented data delivery be employing ACKs, sequence numbers, and timers
- TCP Connection parameters [RFC 2581]:
    - cwnd: sender's congestion window
    - rwnd: receiver's advertised window
    - W_m: max. window size, min(cwnd, rwnd)
    - RTT: round trip time
    - RTO: retransmission timeout

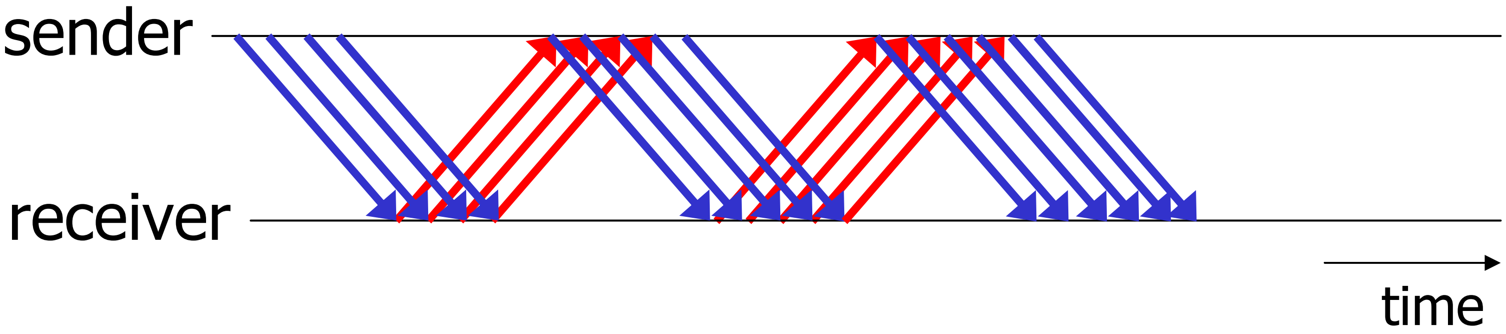
Overview of TCP (cont.)

- Connection Establishment (CE):
  - three-way handshake
- Slow Start (SS):
  - increment cwnd by 1 for each ACK



TCP Congestion Control

- Congestion Avoidance (CA):
  - switch to CA when cwnd reaches ssthresh
  - increment cwnd by 1 every RTT
  - remain in CA until:
    - TO loss: timeout ⇨ go to SS
    - TD loss: triple duplicate ACKs ⇨ go to FRT



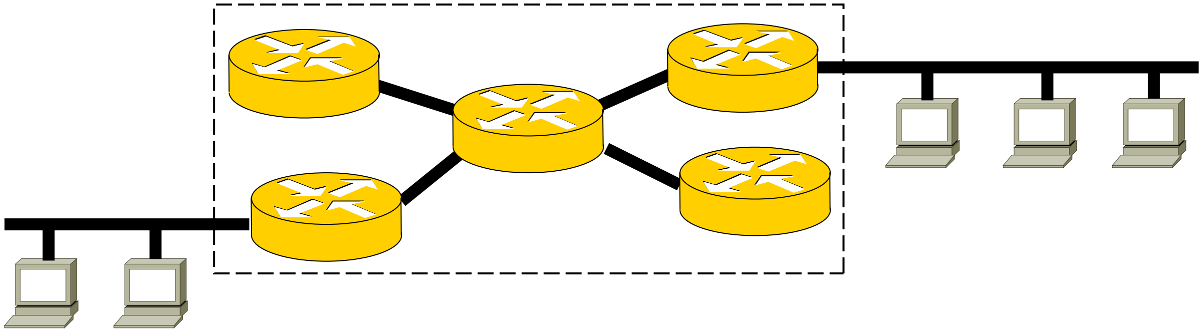
TCP Congestion Control (cont.)

- Fast Retransmit and Fast Recovery:
  - immediately retransmit the lost segment
  - set ssthresh = cwnd/2
  - set cwnd = ssthresh + 3 (inflation)
  - transmit a new segment if allowed
  - for each DUPACK: increment cwnd by 1
  - when ACK for new data arrives, set cwnd = ssthresh (deflation)

Delayed Acknowledgement

- TCP receiver may send one ACK for b segments, b≥1 [RFC 2581]
- ACKs are generated:
    - for at least every second segment
    - within 500 msec of the arrival of an unacknowledged segment
    - for out-of-order segments
    - for segments that fill gaps in the sequence number space

Model Classification

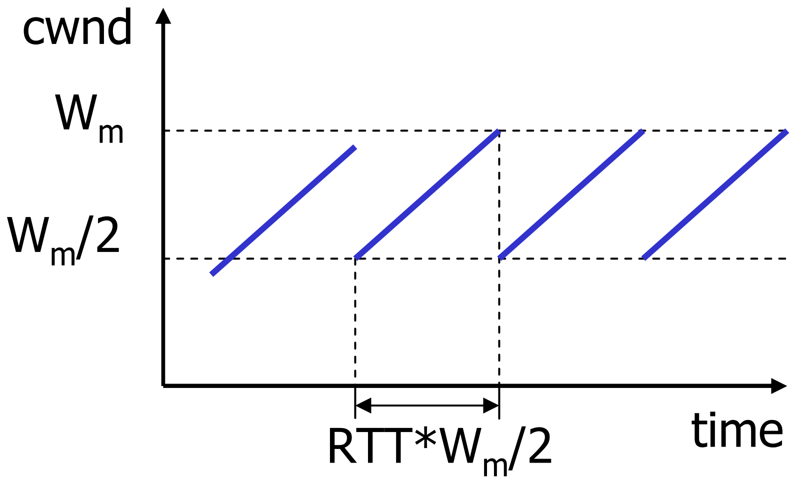


- Three perspectives:
  1. queue dynamics
  2. interaction between TCP and queue management mechanisms
  3. TCP dynamics
- TCP models can be further classified based on transfer length: long, short, or arbitrary

Models for Long-Lived Transfers

1. M. Mathis, et al. (Pittsburgh Supercomputing)

- predict steady-state throughput
- only consider CA phase and TD losses
- periodic loss p
- average throughput:



$$T = \frac{MSS}{RTT}\frac{K}{\sqrt{p}}$$

Models for Long-Lived Transfers (cont.)

2. J. Padhye, et al. (Univ. of Massachusetts)

- predict steady-state throughput
- consider TD and TO losses during CA
- use rounds (round duration=RTT)
- bursty loss model
- average throughput is:

$$T = \min\left(\frac{W_m}{RTT}, \frac{1}{RTT\sqrt{\frac{2bp}{3}} + RTO_0 \min(1,3\sqrt{\frac{3bp}{8}})p(1+32p^2)} \right)$$

Models for Transfers of Arbitrary Length

3. N. Cardwell, et al. (Univ. of Washington)

- extend model 2 to include CE and SS
- predict the expected latency L
- CE latency:

$$E[L_{CE}] = RTT + RTO_0\left(\frac{1-p_r}{1-2p_r} + \frac{1-p_f}{1-2p_f} - 2\right)$$

- expected latency:

$$E[L] = E[L_{ss}] + E[L_{loss}] + E[L_{ca}] + E[L_{delack}]$$

A Model for Short-Lived Transfers

4. M. Mellia, et al. (Turin Polytechnic)

- use CE latency from model 3
- compute average L by exhaustively enumerating all loss scenarios:

$$\begin{aligned} L_1^1 &= RTT + q\sum_{i=1}^{\infty} p^i \sum_{j=1}^{i} 2^{j-1} RTO \\ &= RTT + RTO\frac{p}{1-2p} \end{aligned}$$

- only handles transfers of a few segments, because complexity grows exponentially

Model Comparison

- Common assumptions:
  - no specific topology or queue management
  - greedy sources (Mathis)
- Common features:
  - using rounds (Padhye)
  - bursty loss model (Padhye)
  - three-way handshake latency (Cardwell)
  - closed-form solutions
  - L (T) is directly (inversely) proportional to RTT, p, RTO, and E[t_{TO}]

Model Validation

Model	Simulation	Controlled measurements	Live measurements	Compare to
Mathis	✓	✓	✗	none
Padhye	✓	✓	✗	none
Altman	✗	✓	✗	Padhye
Cardwell	✓	✓	✓	Mathis, Padhye
Sikdar	✓	✗	✓	Padhye, Cardwell
Mellia	✓	✗	✗	none

Summary and Conclusions

- Presented an overview of TCP dynamics and modelling environment
- Surveyed a number of analytical models
- Compared the models w.r.t. assumptions, approaches, and validation methods
- Missing features:
    - accurate model of delayed ACK and fast recovery
    - need for a reference set of measurements and evaluation metrics

References

- M. Allman and V. Paxson and W. Steven, TCP Congestion Control, RFC 2581, April 1999.
- M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestion avoidance algorithm," *ACM Computer Communication Review*, July 1997.
- J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Reno performance: A simple model and its imperical validation," *IEEE/ACM Transactions on Networking*, April 2000.
- E. Altman, K. Avrachenkov, and C. Barakat, "A stochastic model of TCP/IP with stationary random losses," *ACM Computer Communication Review*, October 2000.
- N. Cardwell, S. Savage, and T. Anderson, "Modeling TCP latency," *Proc. INFOCOM*, Tel Aviv, Israel, March 2000.
- B. Sikdar, S. Kalyanaraman, and K. S. Vastola, "Analytic models and comparative study of the latency and steady-state throughput of TCP Tahoe, Reno and SACK," *Proc. GLOBECOM*, San Antonio, Texas, November 2001.
- M. Mellia, I. Stoica, and H. Zhang, "TCP model for short lived flows," *IEEE Communications Letters*, February 2002.