Inline Path Characteristic Estimation to Improve TCP Performance in High Bandwidth-Delay Networks

Cesar Marcondes
Anders Persson
Medy Sanadidi
Mario Gerla

UCLA CSD

HIDEyuki Shimonishi
Takayuki Hama
Tutomu Murase

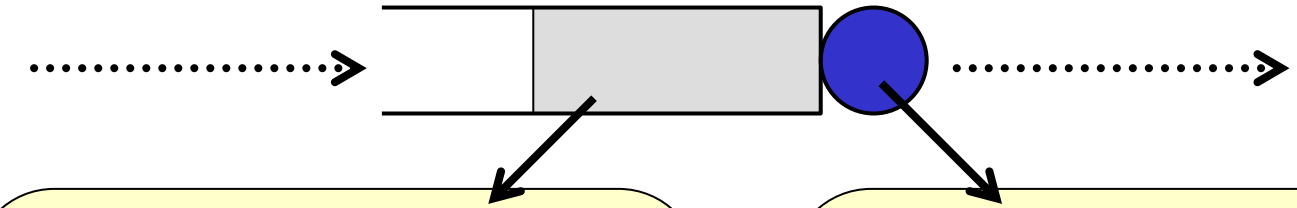
NEC

Outline

- Inferring Path Characteristics
  - Bandwidth Estimation
  - Buffer Estimation
  - Description of Testbed
- Improving TCP Performance using Inline Path Estimations
  - TCP Westwood BBE (Buffer and Bandwidth Estimation)
- Conclusions and Future Work

Motivation

- Inferring Path Characteristics -



Buffer Estimation

Loss discrimination:
Efficiency in leaky pipes

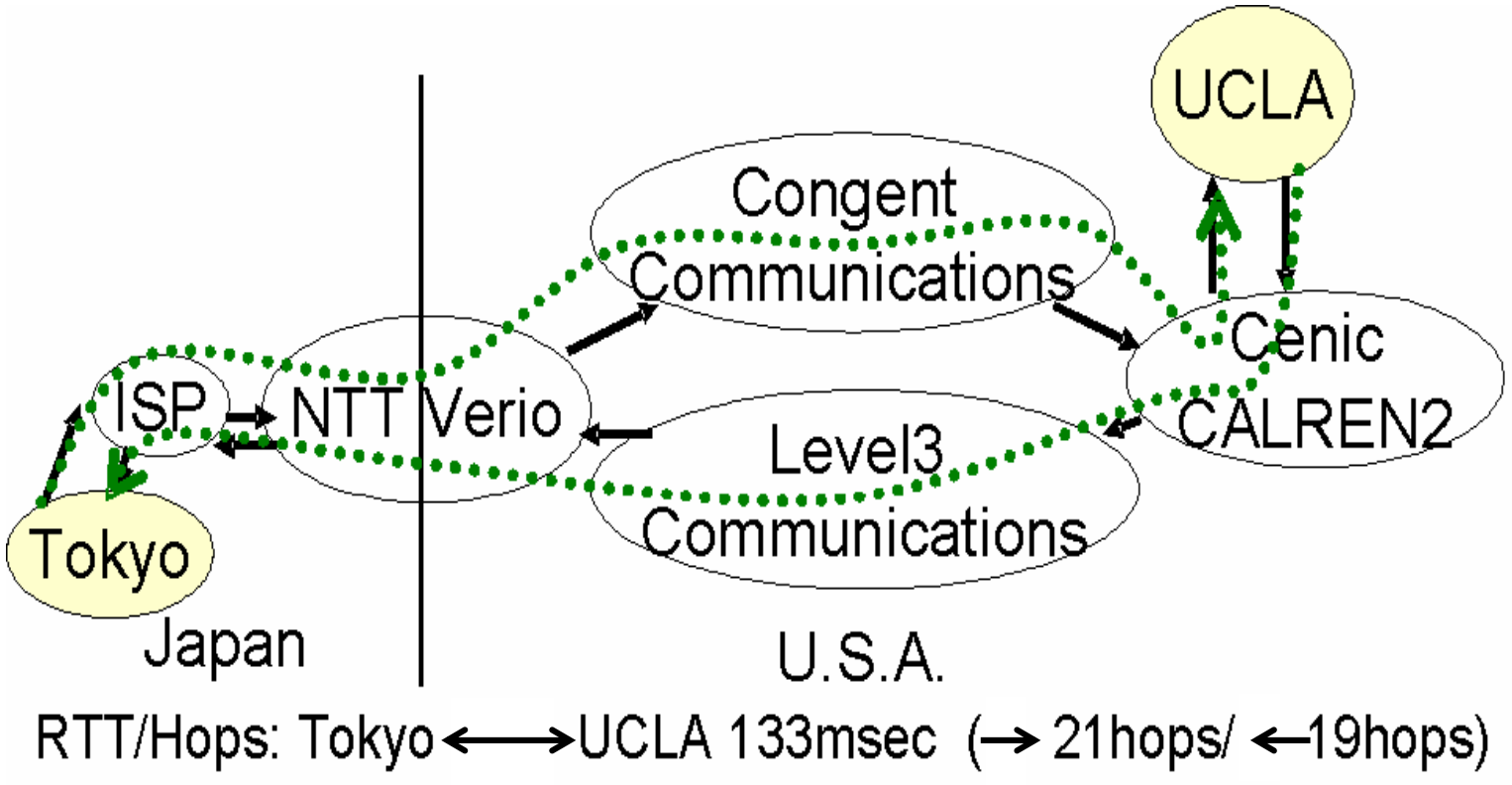
Congestion detection:
Friendliness to TCP-Reno

Bandwidth Estimation

Link capacity:
Efficiency in large pipes

Design novel protocols that make wise usage of estimations

Internet Measurement Testbed Topology
(Transpacific Link)



Capacity Estimation

- Pathrate - Active

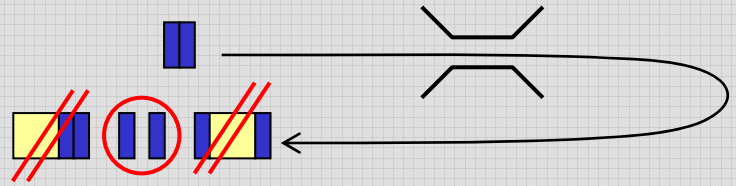
  (Infocom 2001)

  – dispersion of packet pairs and packet trains, uncover possible capacity modes, apply statistic analysis to choose

- Capprobe - Active

  (Sigcomm 2004)

  - Packet pair measurement with filtering out delayed samples



TCP Inline Measurements

- TCP-Westwood

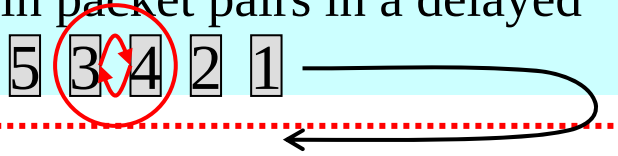
  (Global Internet 2005)

  - Packet pair like capacity estimation using TCP ACK packets

- TCPProbe

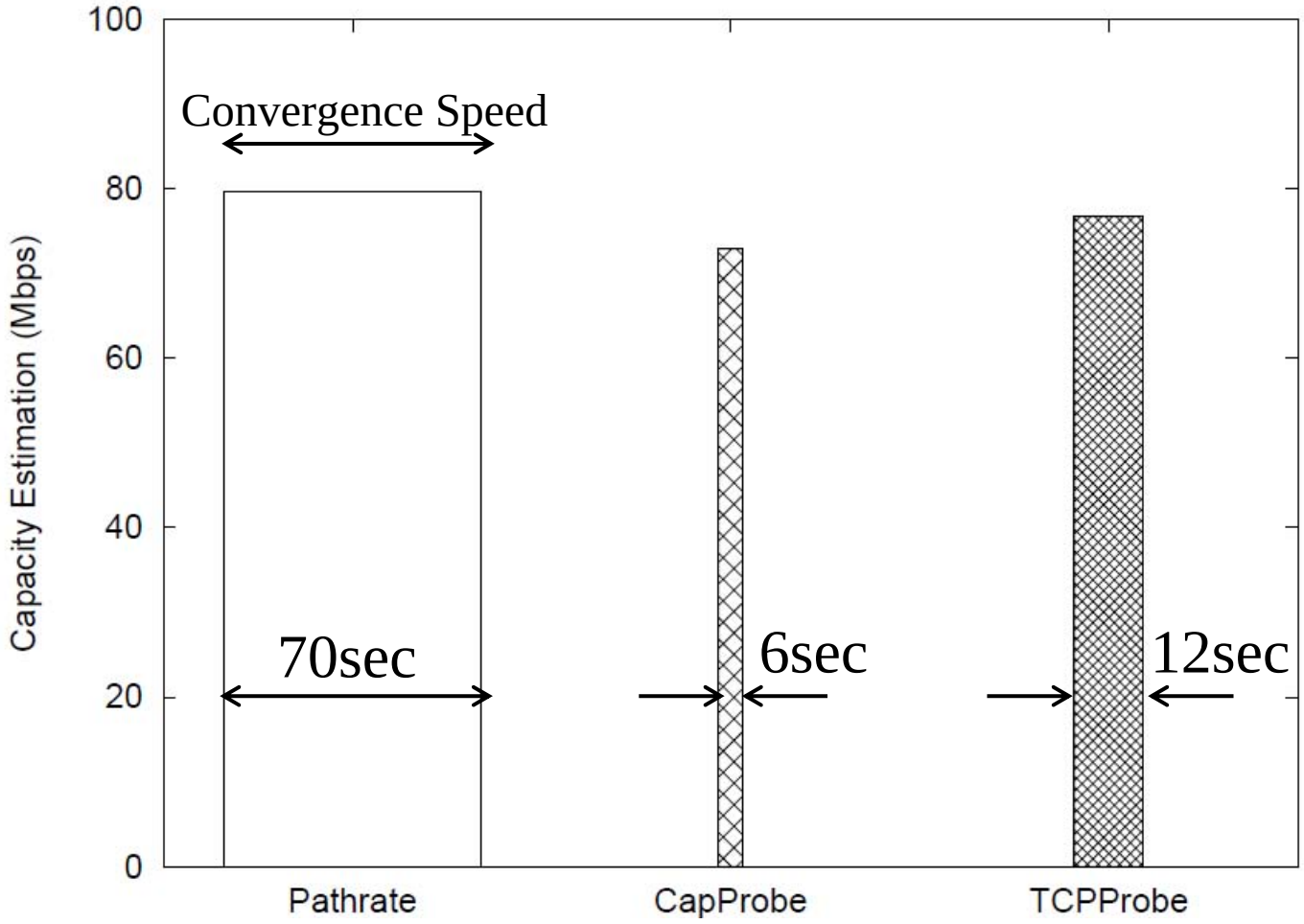
  (Global Internet 2005)

  - TCP inline version of CapProbe, it flips packets to obtain packet pairs in a delayed ACK scenario



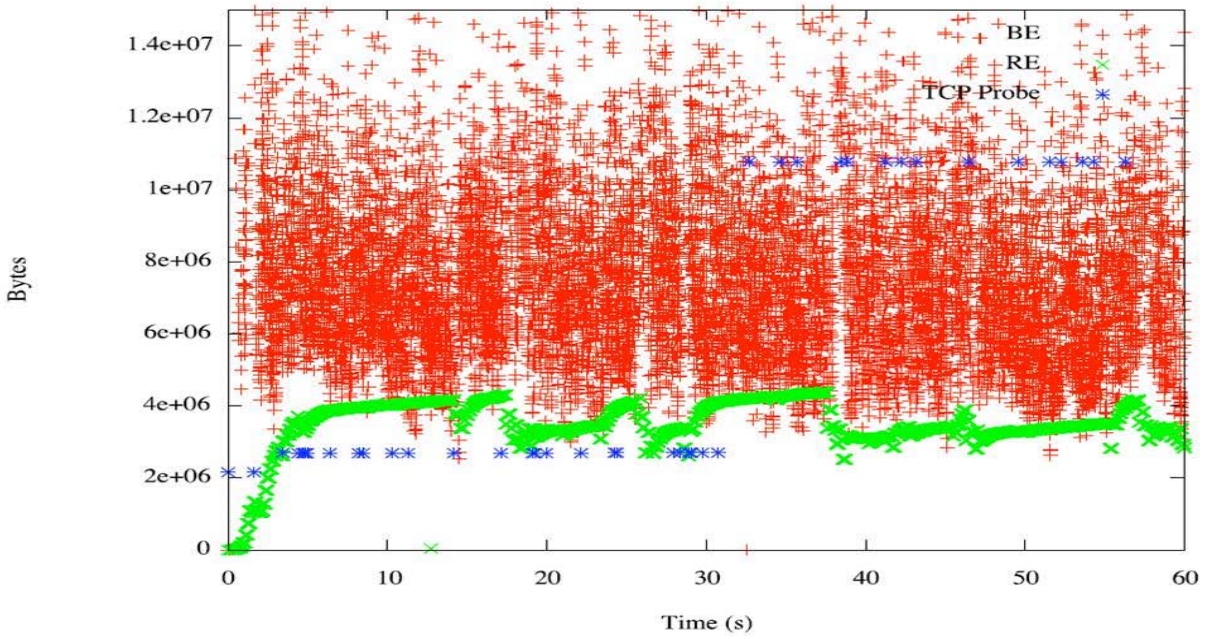
Measurement results (1)

Off-line Measurement



Measurement results (2)

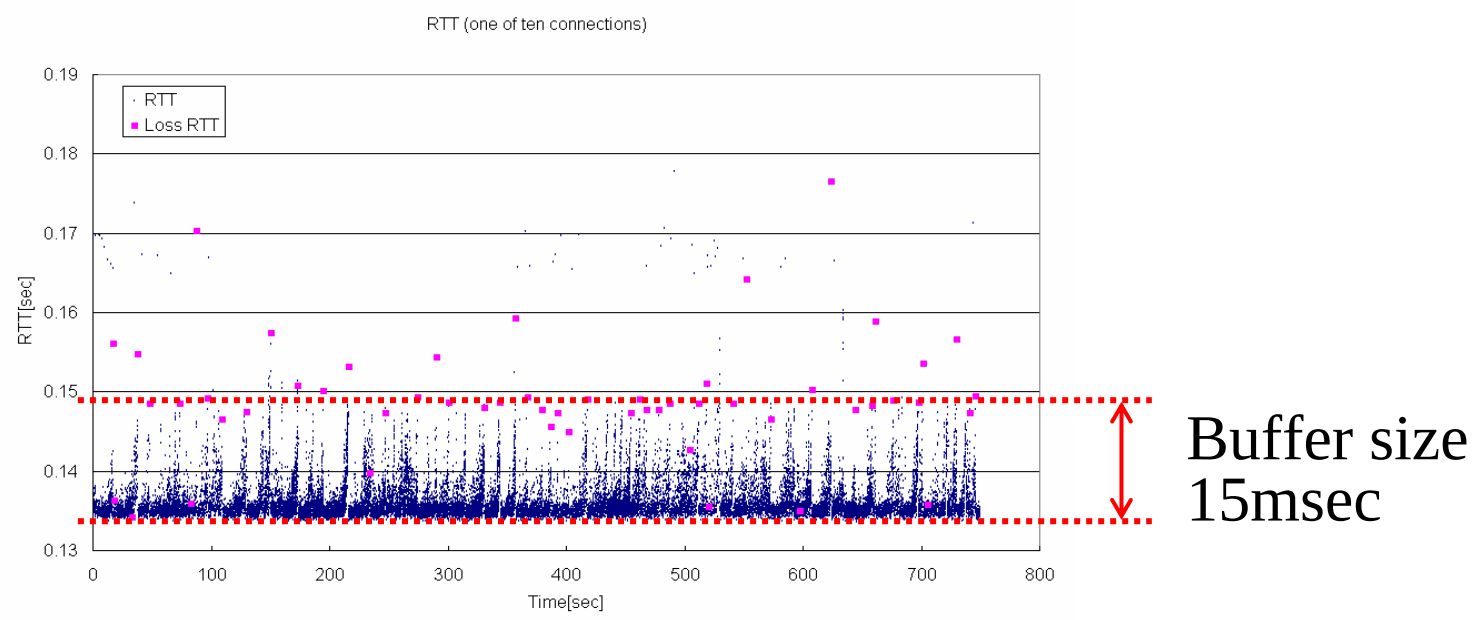
TCP Inline Measurement



	Accuracy	Convergence
Original TCPW	≈ 60Mbps	< 1sec
TCPProbe	≈ 80Mbps	32sec

Buffer Estimation

- Idea: Correlate the loss of a packet with the RTT observed just before the loss
  - Buffer size estimation (RTT value a packet loss likely to happen)
  - Congestion estimation, loss discrimination (position of current RTT)

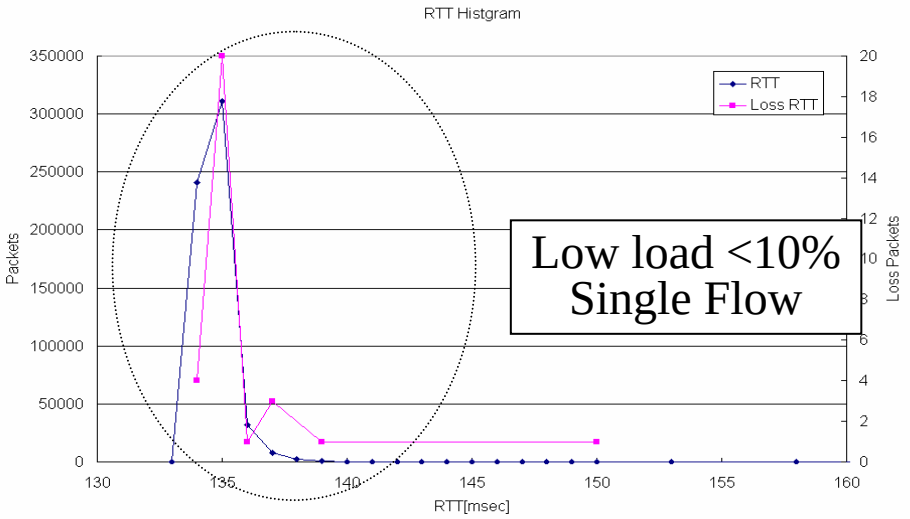


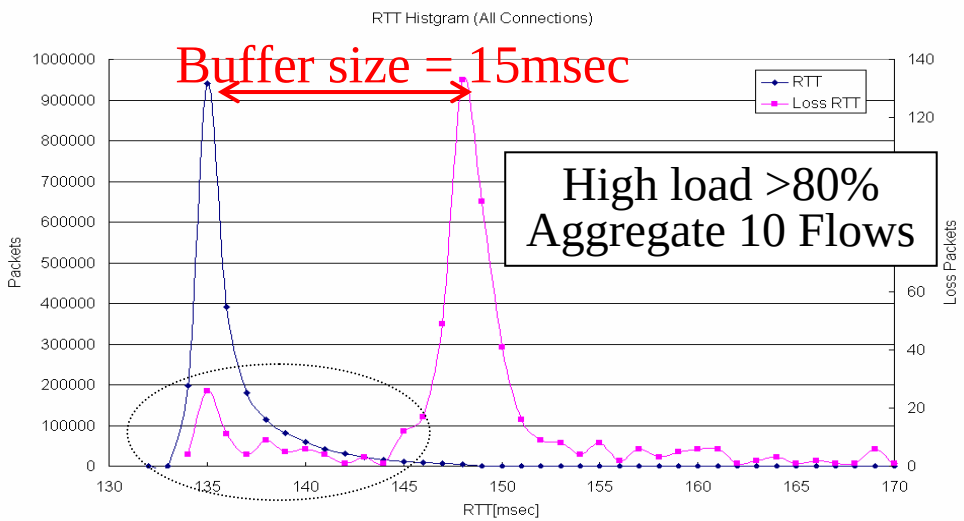
Measurement results (1)
Off-line measurement for buffer capacity

- **Traceroute + Heavy Cross-Traffic Measurements**

  track down the location of bottleneck and observed an increase of **15msec** on the average ICMP delay

- **RTT histogram of TCP flows**





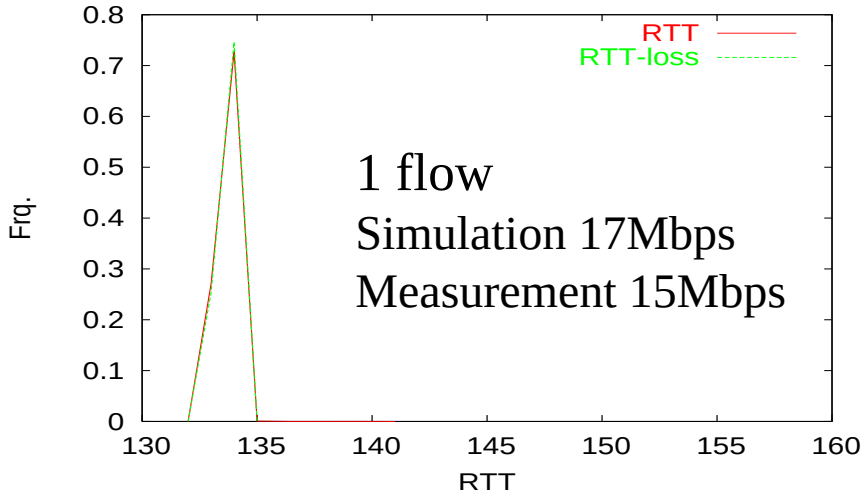
Overall loss rate = 0.005%

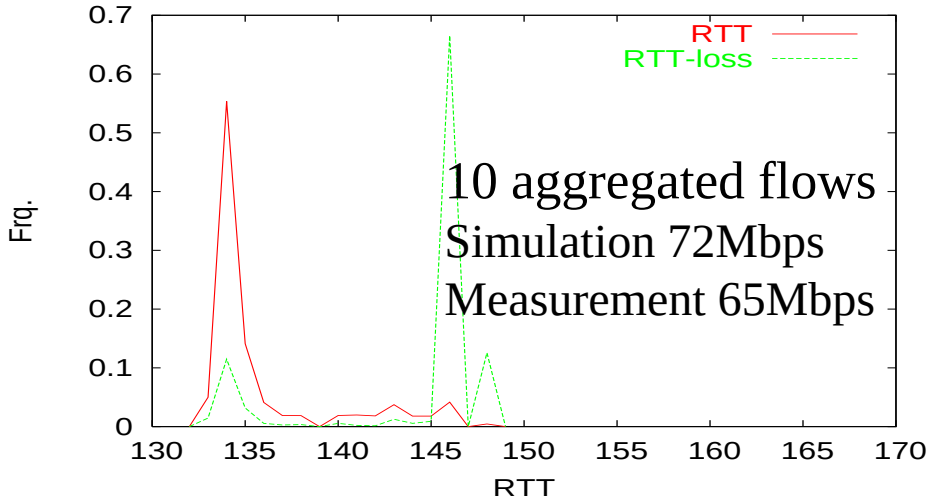
Loss rate (RTT<147msec) = 0.005%

Random loss ?

Measurement results (2)
Off-line measurement for random losses

- **Simulation Results**
  - 135msec RTT + 15msec buffer
  - 80Mbps bottleneck like
  - $4.9 * 10^{-5}$ uniform dist. random loss





- **Low Rate UDP Measurement**
  - 1-3Mbps UDP CBR traffic

➔ Loss rate = 0.001%

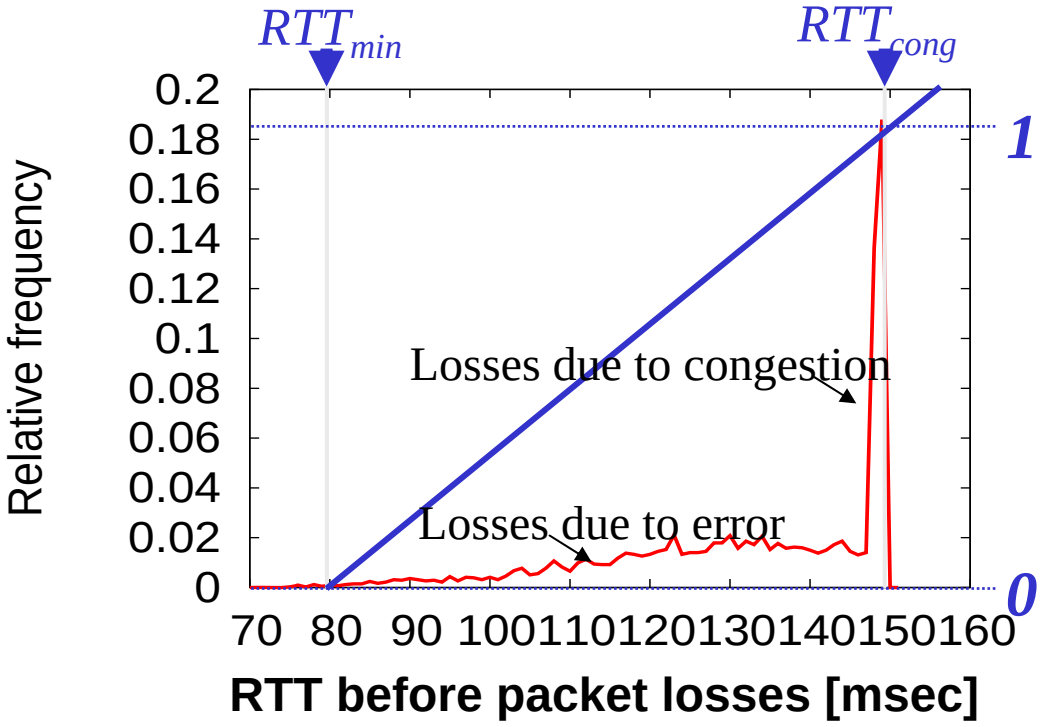
0.001-0.005% should be a good number for random loss rate

TCP Westwood BBE

- Optimize tradeoff between efficiency and friendliness to TCP-Reno
  - Robustness to buffer capacity variations and RTT variations
- Optimize window reduction upon a packet loss, using
  - In line bandwidth estimation
  - In line buffer estimation

TCPW-BBE Buffer Estimation

- RTT dynamic range
  - RTT_{min}: minimum RTT (= propagation delay)
  - RTT_{cong}: RTT value when packet losses likely to happen
- Congestion = position of current RTT in the range

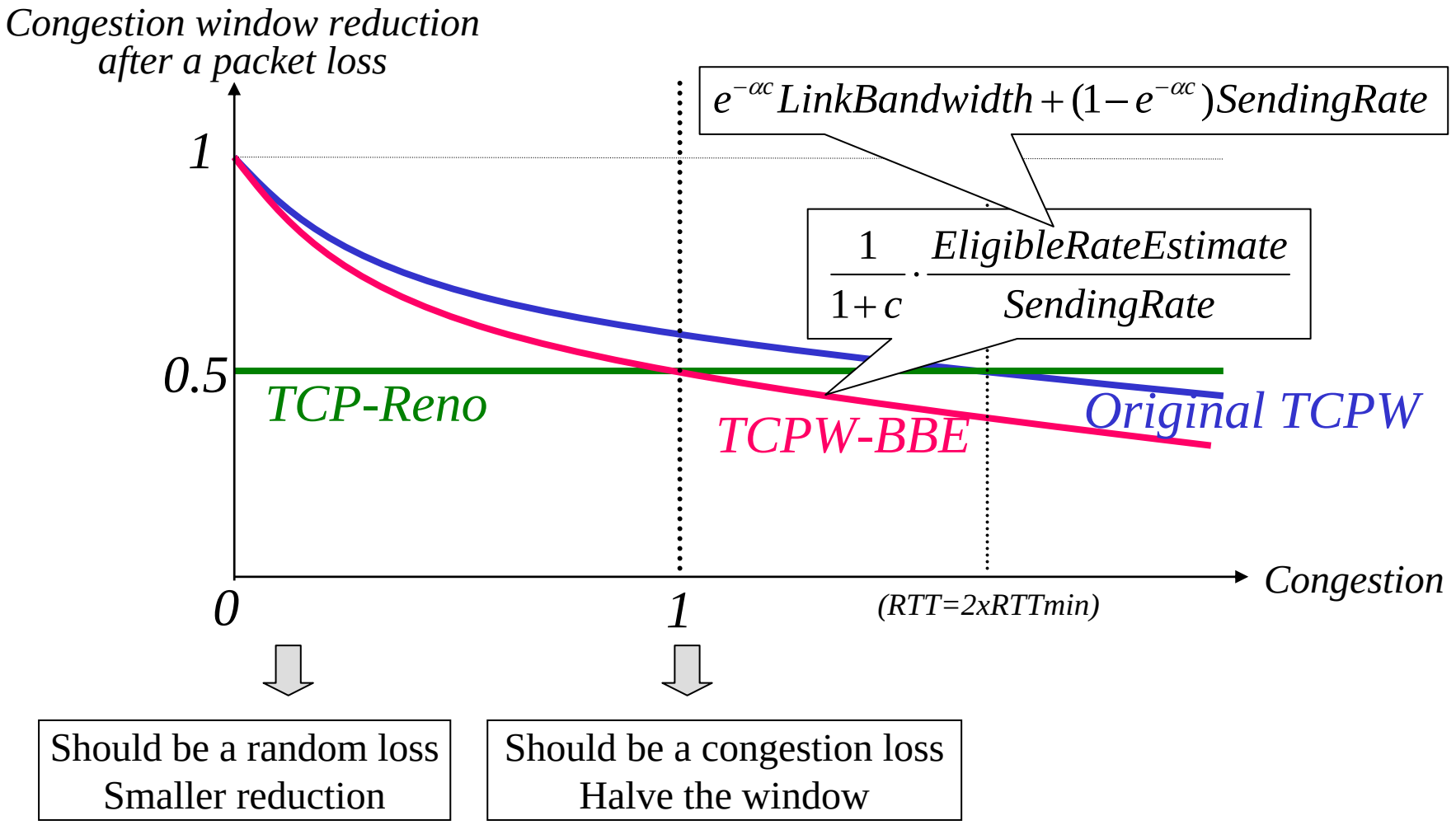


RTT_{cong} is an exponential average of RTT right before losses

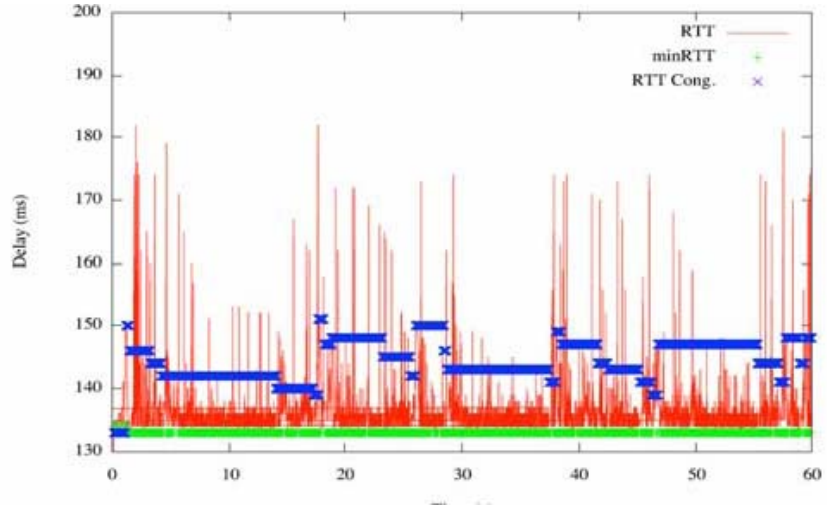
Congestion level

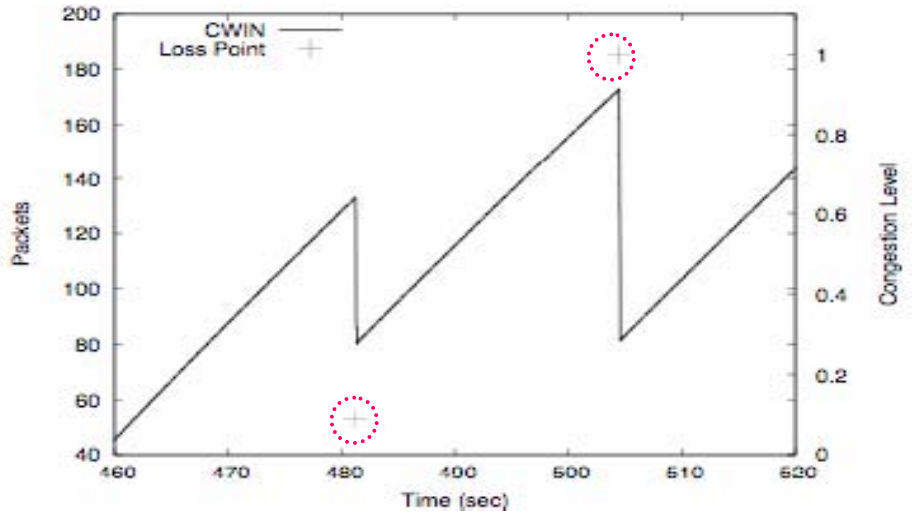
$$c = \frac{RTT - RTT_{min}}{RTT_{cong} - RTT_{min}}$$

Congestion Window Reduction



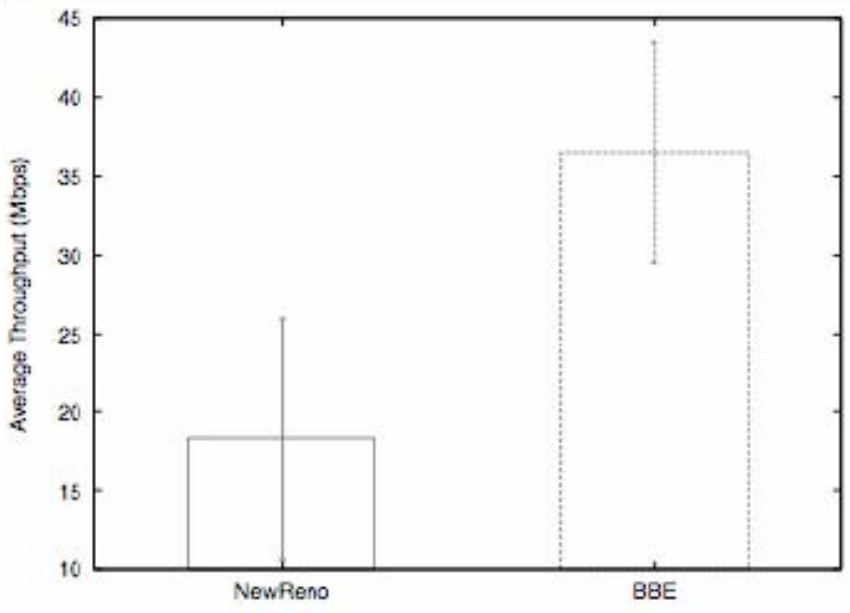
Measurement Result (1)
Buffer estimation and loss discrimination





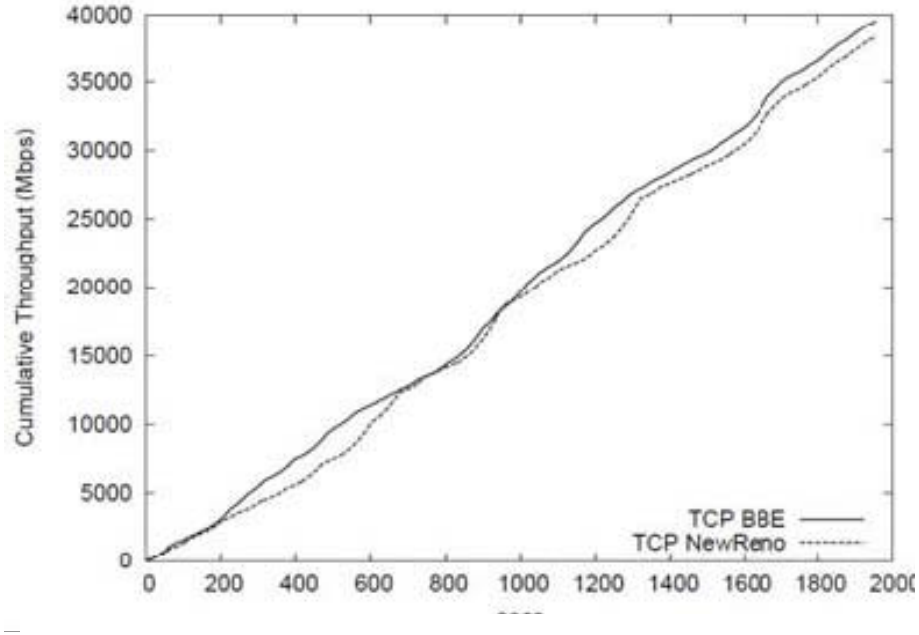
- Buffer Estimation
  - Inline estimating of buffer around 15 msec
  - * consistent with previous results
- Inline Loss Discrimination
  - Window reduction depending on the congestion level estimated
  - * more robust to non-congestion errors

Measurement Result (2)
Efficiency and Friendliness



Average Throughput of Multiple Trials

Pathload Estimation: 61-74Mbps Avail.
Conjectured Non-Congestion Error: 10^{-5}



Cumulative Throughput during 2000 Secs

Pathload Estimation : 30Mbps Avail.
More Congestion losses

Conclusions and Future Work

- Path Characteristic Estimations as presented in this paper are reliable/fast/non-intrusive and present a major advantage for future Intelligent Network Stacks
- In our measurement study, we presented, in addition to extensive Internet results, cross-validation based on more than one source and we observed that the inline capacity and buffer estimations are reliable.
- In the specific scenario tested, TCP BBE was able to improve the performance when comparing to NewReno on a non-congested scenario
- In the future, a combination of BE (Bandwidth Estimation from TCP Westwood) and TCPProbe can lead to faster capacity estimation. Additionally, new methods of using such estimates to improve start-up, cross-traffic identification, changes in route characteristics and ameliorate burstiness issues using buffer-awareness ought to be studied