Analysis of Dynamic Behaviors of Many TCP Connections Sharing Tail-Drop / RED Routers

Go Hasegawa
Osaka University, Japan
hasegawa@cmc.osaka-u.ac.jp

Backgrounds

- *TCP (Transmission Control Algorithm)*
  - Majority in the current Internet, also in the future
- *Analytic investigation is necessary to understand its characteristics*

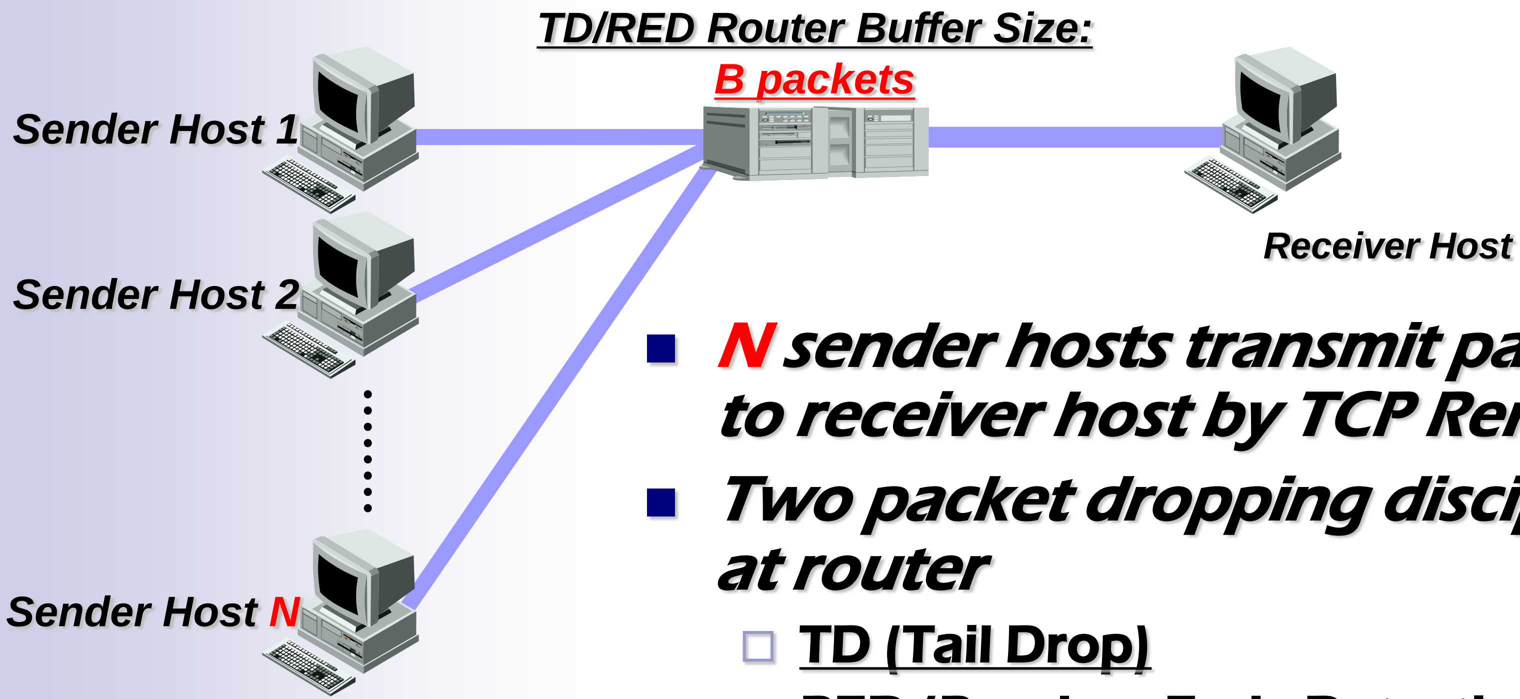
Past researches on TCP throughput analysis

- ***Long-term average throughput***
  - Short-term throughput is important for short file transfer
- ***Assume constant packet loss ratio***
  - Packet loss ratio changes dynamically due to bursty packet loss
- ***Assume RED works fine at the router***
  - Bad parameter setting degrades RED performance, causing bursty packet loss

Objectives

- *Analysis of window size distribution of TCP connections*
  - Using simple Markov modeling of TCP behavior
  - Many TCP connections accommodated
  - TD (Tail-Drop) and RED (Random Early Detection)
  - Effect of bursty packet loss
- *Evaluation of TD/RED routers in terms of ...*
  - Short-term fairness among TCP connections
  - Effect of poor parameter set of RED
  - Effect of TD buffer size

Network model



- *N* sender hosts transmit packets to receiver host by TCP Reno
- Two packet dropping disciplines at router
  - TD (Tail Drop)
  - RED (Random Early Detection)
- Focus on changes of window size, and ssthresh value of TCP connection

Markov modeling of TCP behavior

- State is a combination of window size and ssthresh values of a TCP connection
  - (w_i, t_i)
- State transition occurs at every RTT
  - *cwnd* increases when no packet loss occurs
  - *cwnd* and *ssth* decrease when packet loss occurs
- State transition probabilities are dependent on...
  - Packet loss probability at the router buffer
  - Slow start, congestion avoidance algorithms of TCP

Increasing window size

- ***When no packet loss occurs***
  - Probability: $(1 - p)^{w_i}$
- ***State transition from (w_i, t_i) to …***
  - $(2w_i, t_i)$
    - When $w_i < (1/2)t_i$ (Slow Start Phase)
  - (t_i, t_i)
    - When $(1/2)t_i < w_i < t_i$ (Phase Shift)
  - (w_i+1, t_i)
    - When $t_i < w_i$ (Congestion Avoidance Phase)

Decreasing window size

- ***When packet loss occurs***
  - Probability: $1-(1-p)^{w_i}$
- ***State transition from (w_i, t_i) to …***
  - $(1, w_i/2)$
    - When timeout occurs
  - $(w_i/2, w_i/2)$
    - When fast retransmit occurs
- ***Probability of timeout***
  - Dependent on w_i and number of lost packets in a window

Packet loss probability: p

- ***Past researches assume p is constant***
- ***Actually dependent on…***
  - **Router buffer size: B**
  - **Window size: w_i**
  - **Packet discarding discipline**
    - TD (Tail Drop), RED (Random Early Detection)

Tail-drop router

- **Bursty packet loss** occurs when the router buffer overflows
- To calculate p, we have derived ...
  - $p_{overflow}$: frequency of buffer overflow
  - $L_{overflow}$: # of lost packets in each buffer overflow
  - L_i: # of lost packets for each TCP connection in each buffer overflow
- $p = \min(1, p_{overflow} \cdot L_i / w_i)$

Tail-drop router (2)

- $p_{overflow}$: frequency of buffer overflow
  - Considering queue dynamics
  - $1/(N(W_f - Nw'))$
- $L_{overflow}$: # of lost packets in each buffer overflow
  - Each TCP connection increases its window size by 1 packet at every RTT
  - *N packets are lost in total*
- L_i: # of lost packets for each TCP connection in each buffer overflow
  - Proportional to window size of each TCP connection

RED router

- *Packet discarding probability is determined from average queue length*
- *For applying to our model, we use instantaneous queue length;*

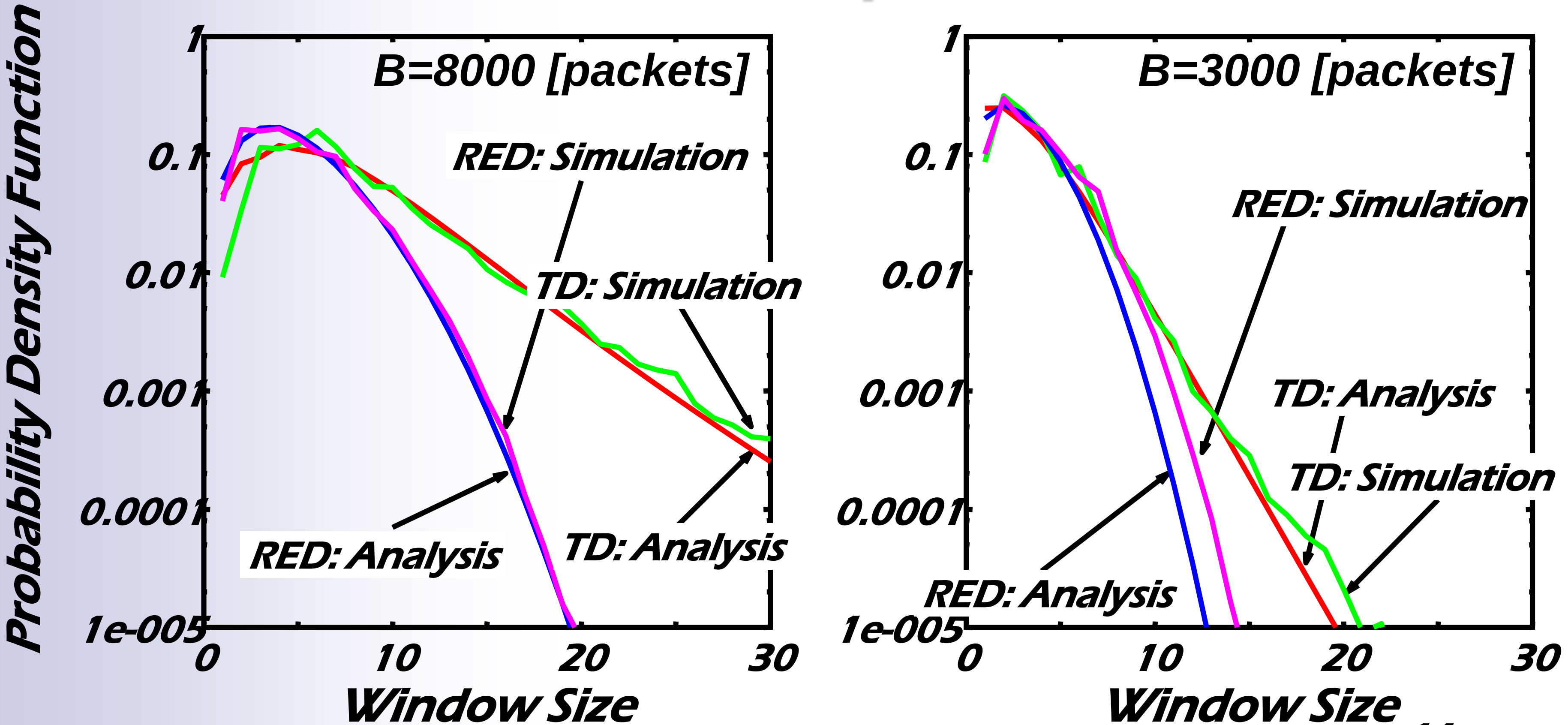
$$p_{red}(q) = \begin{cases} 0 & \text{if } q < min_{th} \\ \dfrac{q - max_{th}}{max_{th} - min_{th}} & \text{if } min_{th} \le q < max_{th} \\ \dfrac{q \cdot max_p + (q - max_{th})}{q} & \text{if } max_{th} \le q \end{cases}$$

RED router (2)

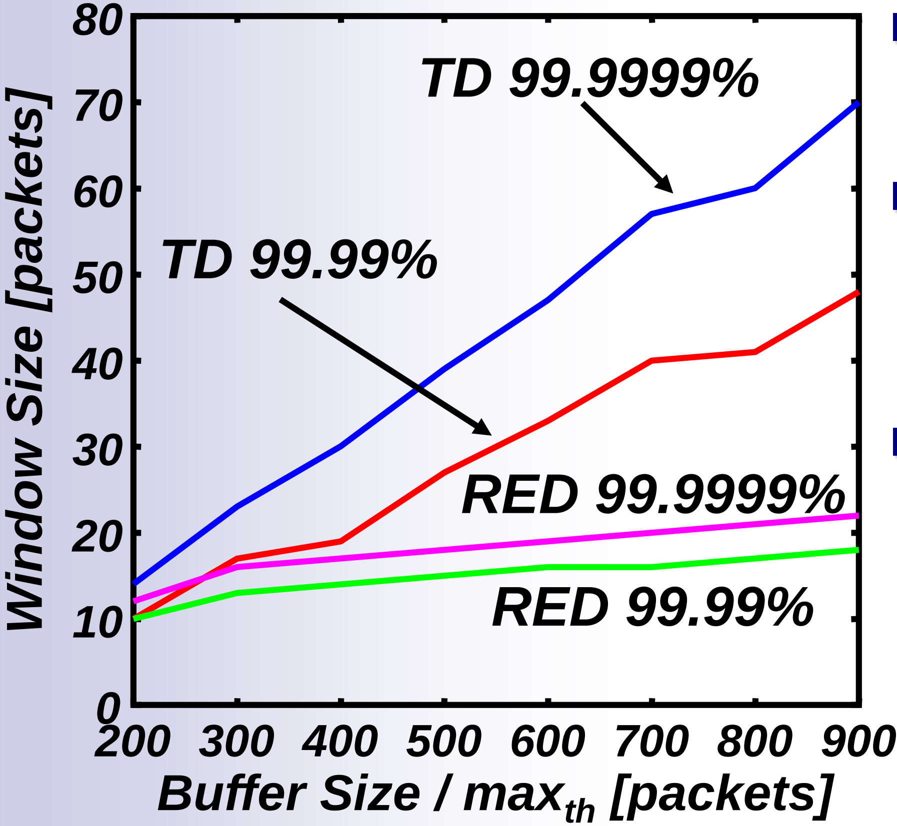
- q: queue length
  - Assume that other TCP connections are in steady state, and queue length is affected only by w_i
- $q = ((N-1)/N)w^* + w_i - 2\tau\rho$
- $p = p_{red}(q)$

Accuracy of Analysis

- N=1000, BW=1.5 Mbps, τ =2 msec



Fairness evaluation



- *Fairness of TD is much affected by buffer size*
- *Variation of window size of RED is small, <u>regardless of buffer size</u>*
- *RED can provide <u>better fairness in short-term TCP throughput</u>*

Conclusion

- ***Analysis of window size distribution of TCP connections***
  - TD/RED disciplines
  - Burst packet loss
- ***Fairness evaluation of TD/RED router***
  - RED can give short-term fairness among connections

RED router

- Probability is changed according to average queue length
- Avoid buffer overflow, keep queue length low