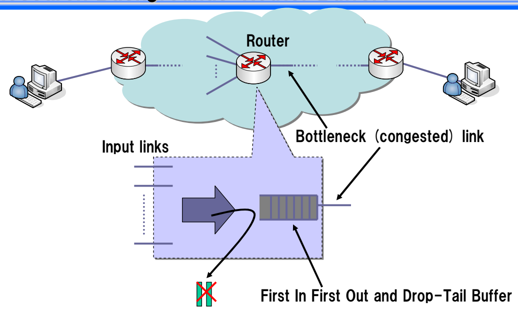
Router buffer re-sizing for short-lived TCP flows

Takeshi Tomioka, Go Hasegawa, and Masayuki Murata
Osaka University, JAPAN

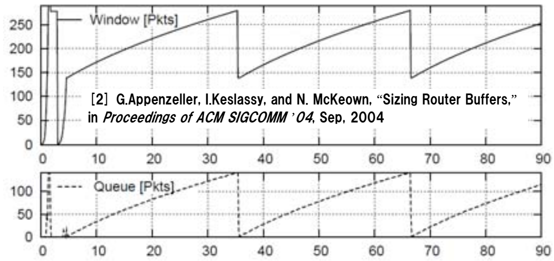
Research Backgrounds: Router Buffers



- Router buffer size affects …
  - Packet loss ratio in the network when the link is congested
  - Utilization of output link bandwidth
  - Packet transmission delay caused by queuing at the buffer

2

Two disciplines for router buffer size (1)



[2] G.Appenzeller, I.Keslassy, and N. McKeown, "Sizing Router Buffers," in *Proceedings of ACM SIGCOMM '04*, Sep. 2004

- Determined from window size control mechanism of TCP
  - It increases window size additively when no packet loss occurs, and halves it when packet loss occurs
  - This buffer size can avoid the underutilization of the output link

3

Two disciplines for router buffer size (2)

- Problem in normal discipline
  - Large cost for implementing large buffers
    - 2.5Gbits buffer is necessary for a link of 10Gbps bandwidth and 250msec delay
    - Large power consumption, board size, and monetary cost
- sqrtN discipline [2]
  - Buffer size can be decreased to bandwidth-delay product divided by the square-root of the number of TCP connections passing through the link
    - Without underutilization of output link bandwidth
    - When more than 500 TCP connections exists in the network
    - For example, buffer size decreases to 1/100 when 10,000 connections exist

[2] G.Appenzeller, I.Keslassy, and N. McKeown, "Sizing Router Buffers," in *Proceedings of ACM SIGCOMM '04*, Sep. 2004

4

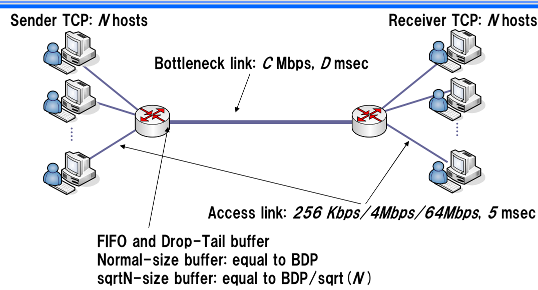
Research Objectives

- Compare two disciplines (normal and sqrtN) by simulation experiments
  - Confirm the results in [2] in terms of link utilization
  - Performance of long-lived and short-lived TCP connections passing through the link
  - Evaluations in various network environment
    - Access link bandwidth
    - bottleneck link bandwidth, bottleneck link propagation delay
    - traffic volume (number of TCP connections)

[2] G.Appenzeller, I.Keslassy, and N. McKeown, "Sizing Router Buffers," in *Proceedings of ACM SIGCOMM '04*, Sep. 2004

5

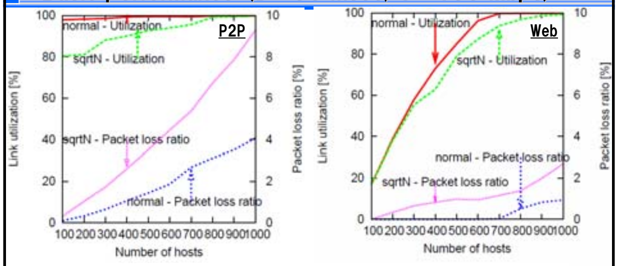
Simulation environment



- Traffic models
  - P2P … each sender TCP has infinite data to transmit
  - Web … Transmission data size and transmission interval are determined by Web traffic model

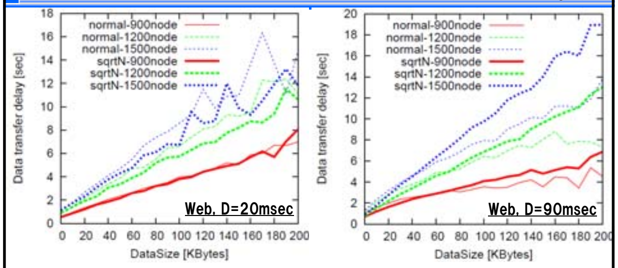
6

Basic performance（D =90 msec, C =100 Mbps)



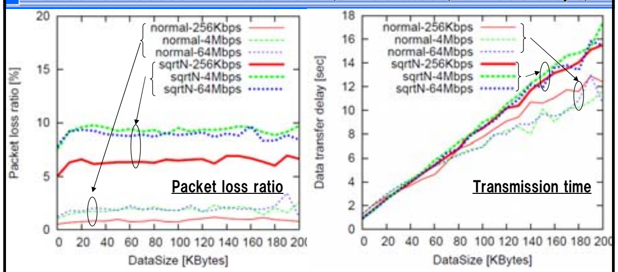
- Normal discipline provides high link utilization and small packet loss ratio
- sqrtN discipline also provides high utilization when number of flows is large enough
  - Conformance to results in [2]
  - But, packet loss ratio is significantly large

Performance of short-lived TCP connections（C =100 Mbps)



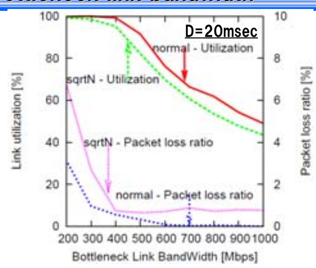
- Packet loss ratio in sqrtN discipline is always larger than that in normal discipline
- When propagation delay is small, larger packet loss ratio in sqrtN discipline does not affect the transmission time
- When propagation delay is large, however, the transmission delay is significantly deteriorated

Effect of access link bandwidth（D =90 msec, C =100 Mbps)



- Large access link bandwidth increases packet loss ratio, especially in sqrtN discipline
  - Transmission time in sqrtN discipline is up to 50% larger
  - However, small data（<100KB）transmission is not affected by the large packet loss ratio

Effect of bottleneck link bandwidth



- In sqrtN discipline, packet loss ratio never decreases to zero, although the link is underutilized
  - because bursty packet arrival can not be absorbed at the smaller buffer
- Link utilization of sqrtN discipline is smaller than that of normal discpilne
  - Using small buffer in non-congested network decrease the link utilization

Conclusion

- We tested the performance of sqrtN discipline for router buffer sizing
  - It can maintain the link utilization when there is enough traffic volume
    - But, it degrades the link utilization in non-congested network
  - It would degrade short-lived TCP performance due to large packet loss ratio
  - It may be useful only when the transmission data size is 50-100 Kbytes or when the propagation delay between the sender and the receiver hosts is significantly small
    - Otherwise, we should use large buffers
- Future work
  - The effect of pacing TCP on the buffer sizing problem