DMZ Exercises

TCP Congestion Control, Buffer Sizing

Elie Kfoury, Jorge Crichigno
University of South Carolina

2021 NSF Campus Cyberinfrastructure (CC*) Workshop

April 15, 2021

Content

- NTP lab series
- Organization of lab manuals
- Exercise 1: TCP congestion control
- Exercise 2: Buffer sizing

NTP Lab Series

- The lab series provides learners an emulated WAN infrastructure operating at high speeds and devices running real protocol stacks
- It helps students to acquire hands-on skills on
  - Performance and measurement tools
  - Configuration of devices for high-speed networks
  - Emulate scenarios using real protocol stacks

NTP Lab Series

- The lab series can be partitioned into three parts
  - Measurement (throughput, latency, packet loss) and emulation (link bandwidth, buffer size, delay) tools
  - TCP features for high-speed transfers, router buffer size
  - Active Queue Management (AQM) algorithms

NTP Lab Series

- Lab experiments

Lab 1:	Introduction to Mininet
Lab 2:	Introduction to iPerf
Lab 3:	WANs with latency, Jitter
Lab 4:	WANs with Packet Loss, Duplication, Corruption
Lab 5:	Setting WAN Bandwidth with Token Bucket Filter (TBF)
Lab 6:	Traditional TCP Congestion Control (HTCP, Cubic, Reno)
Lab 7:	Rate-based TCP Congestion Control (BBR)
Lab 8:	Bandwidth-delay Product and TCP Buffer Size
Lab 9:	Enhancing TCP Throughput with Parallel Streams
Lab 10:	Measuring TCP Fairness
Lab 11:	Router's Buffer Size
Lab 12:	TCP Rate Control with Pacing
Lab 13:	Impact of Maximum Segment Size on Throughput
Lab 14:	Router's Bufferbloat
Lab 15:	Hardware Offloading on TCP Performance
Lab 16:	Random Early Detection
Lab 17:	Stochastic Fair Queueing
Lab 18:	Controlled Delay (CoDel) Active Queue Management
Lab 19:	Proportional Integral Controller-Enhanced (PIE)
Lab 20:	Classifying TCP traffic using Hierarchical Token Bucket (HTB)

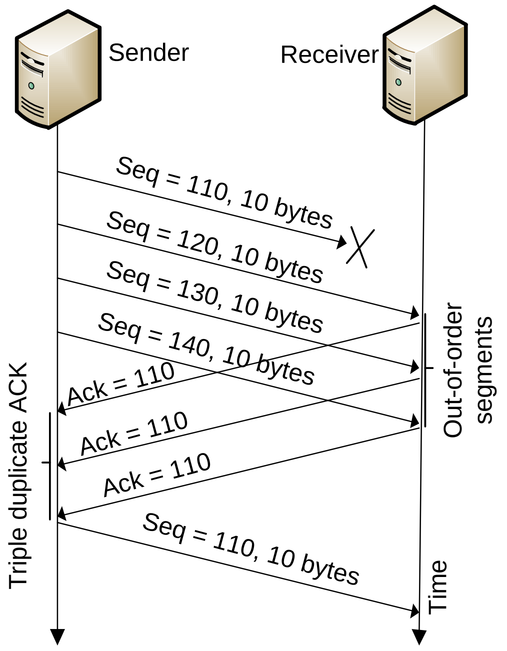
Organization of Lab Manuals

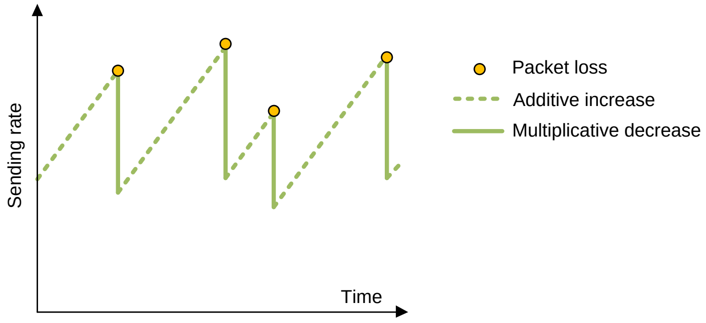
- Each lab starts with a section *Overview*
  - Objectives
  - Lab settings: passwords, device names
  - Roadmap: organization of the lab
- *Section 1*
  - Background information of the topic being covered (e.g., fundamentals of TCP congestion control)
  - Section 1 is optional (i.e., the reader can skip this section and move to lab directions)
- *Section 2… n*
  - Step-by-step directions

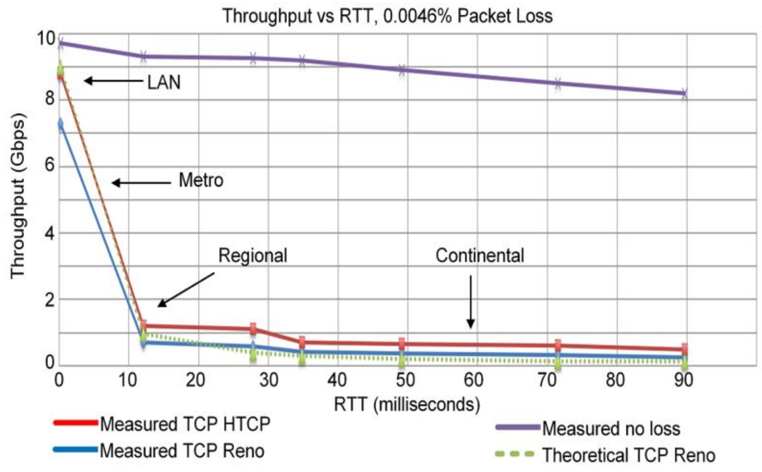
Exercise 1: TCP Congestion Control

TCP Traditional Congestion Control

- The principles of window-based CC were described in the 1980s[1]
- Traditional CC algorithms follow the additive-increase multiplicative-decrease (AIMD) form of congestion control



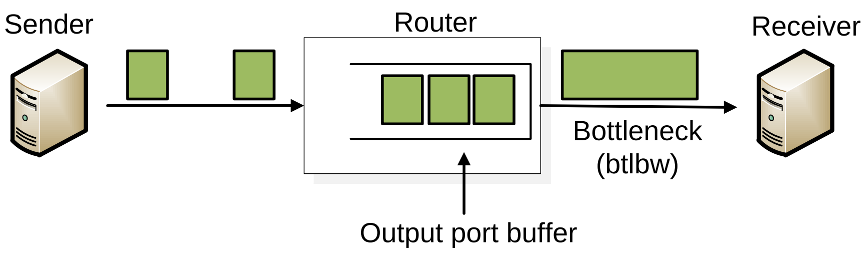


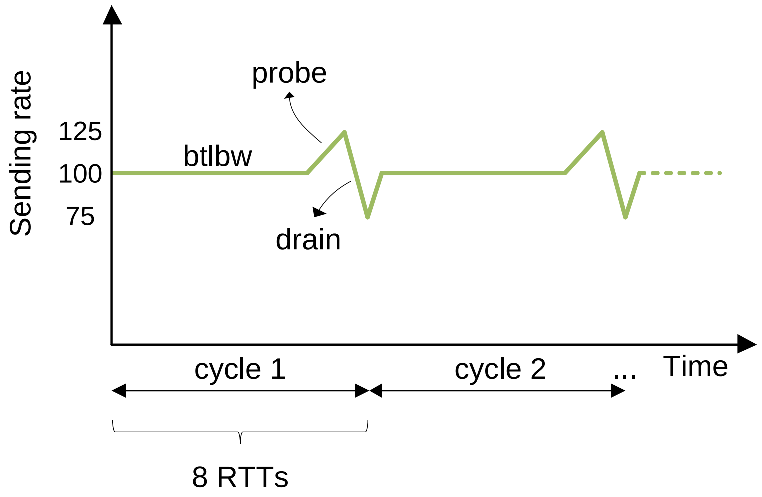


1. V. Jacobson, M. Karels, Congestion avoidance and control, ACM SIGCOMM Computer Communication Review 18 (4) (1988).

BBR: Model-based CC

- TCP Bottleneck Bandwidth and RTT (BBR) is a rate-based congestion-control algorithm[1]
- BBR represented a disruption to the traditional CC algorithms:
  - is not governed by AIMD control law
  - does not the use packet loss as a signal of congestion
- At any time, a TCP connection has one slowest link bottleneck bandwidth (btlbw)

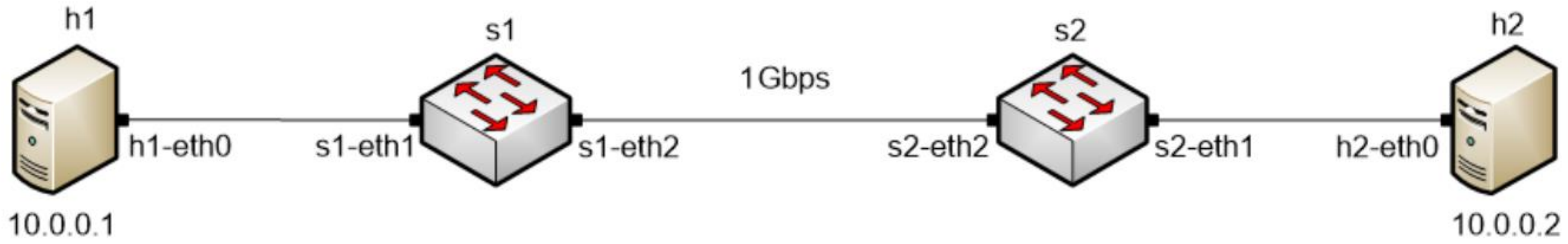




1. N. Cardwell et al. "BBR v2, A Model-based Congestion Control." IETF 104, March 2019.

Lab Goal and Topology

- Modify the TCP congestion control algorithm in Linux using sysctl tool
- Deploy emulated WANs in Mininet
- Compare the performance of TCP Reno and TCP BBR in high-throughput high-latency networks
- Lab topology:

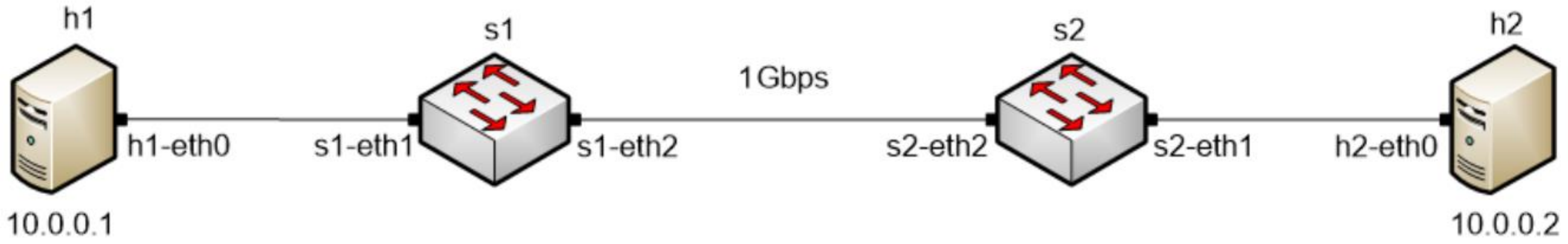


TCP Buffer Size

- In many WANs, the round-trip time (RTT) is dominated by the propagation delay
- To keep the sender busy while ACKs are received, the TCP buffer must be:

Traditional congestion controls: **TCP buffer size ≥ 2BDP**

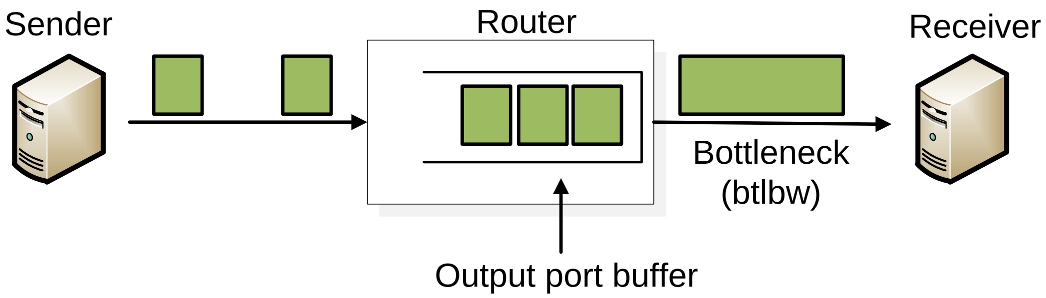
BBRv1 and BBRv2: **TCP buffer size must be considerable larger than 2BDP**



Exercise 2: Buffer Sizing

Buffer Size

- The router's buffer plays an important role in absorbing traffic fluctuations
- Buffers avoid losses by momentarily buffering packets as transitory bursts dissipate



Buffer Size

- The rule of thumb has been that the amount of buffering (in bits) in a router's port should equal the RTT (in seconds) multiplied by the capacity C (in bits per seconds) of the port[1]:

Router's buffer size = $C \cdot RTT$

1. C. Villamizar, C. Song, "High performance TCP in ansnet," ACM Computer Communications Review, vol. 24, no. 5, pp. 45-60, Oct. 1994.

Buffer Size

- The rule of thumb has been that the amount of buffering (in bits) in a router's port should equal the RTT (in seconds) multiplied by the capacity C (in bits per seconds) of the port[1]:

$$\text{Router's buffer size} = C \cdot RTT$$

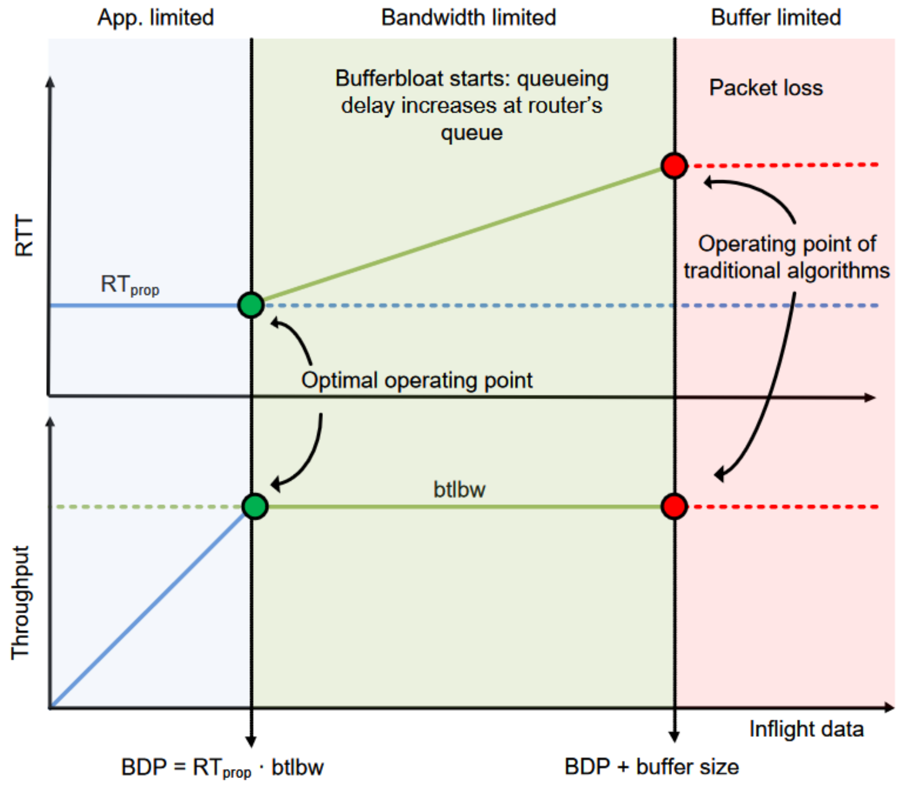
- When there is a large number of TCP flows passing through a link, say N, the amount of buffering can be reduced to[2]:

$$\text{Router's buffer size} = C \cdot RTT / \sqrt{(N)}$$

1. C. Villamizar, C. Song, "High performance TCP in ansnet," ACM Computer Communications Review, vol. 24, no. 5, pp. 45-60, Oct. 1994.
2. G. Appenzeller, I. Keslassy, N. McKeown, "Sizing router buffers," in Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications, pp. 281-292, Oct. 2004.

Bufferbloat

- Bufferbloat is a condition that occurs when the router buffers too much data, leading to excessive delays



1. N. Cardwell, Y. Cheng, C. Gunn, S. Yeganeh, V. Jacobson, "BBR: congestion-based congestion control," Communications of the ACM, vol 60, no. 2, pp. 58-66, Feb. 2017.

Lab Goal and Topology

- Understand the buffering process in a router and buffer sizing
- Explain the concept of Bufferbloat
- Modify routers’ buffer size to solve the bufferbloat problem
- Lab topology: