2023 Internet2 Technology Exchange

Science DMZs and Networking for All

Hands-on session 3: Bandwidth-delay Product and TCP Buffer Sizing

Elie Kfoury
University of South Carolina
https://research.cec.sc.edu/cyberinfra/

University of South Carolina (USC)
Energy Sciences Network (ESnet)

September 18, 2023

Network Tools and Protocols (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 the labs

Each lab starts with a section *Overview*

- Objectives
- Lab topology
- Lab settings: passwords, device names
- Roadmap: organization of the lab

Section 1

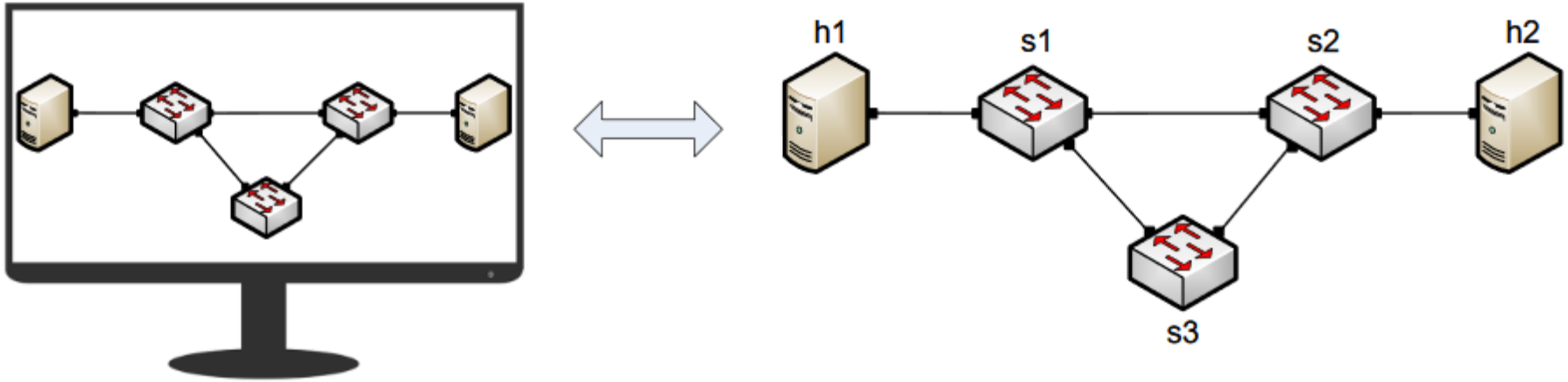
- Background information of the topic being covered (e.g., fundamentals of perfSONAR)
- Section 1 is optional (i.e., the reader can skip this section and move to lab directions)

Section 2… n

- Step-by-step directions

Mininet

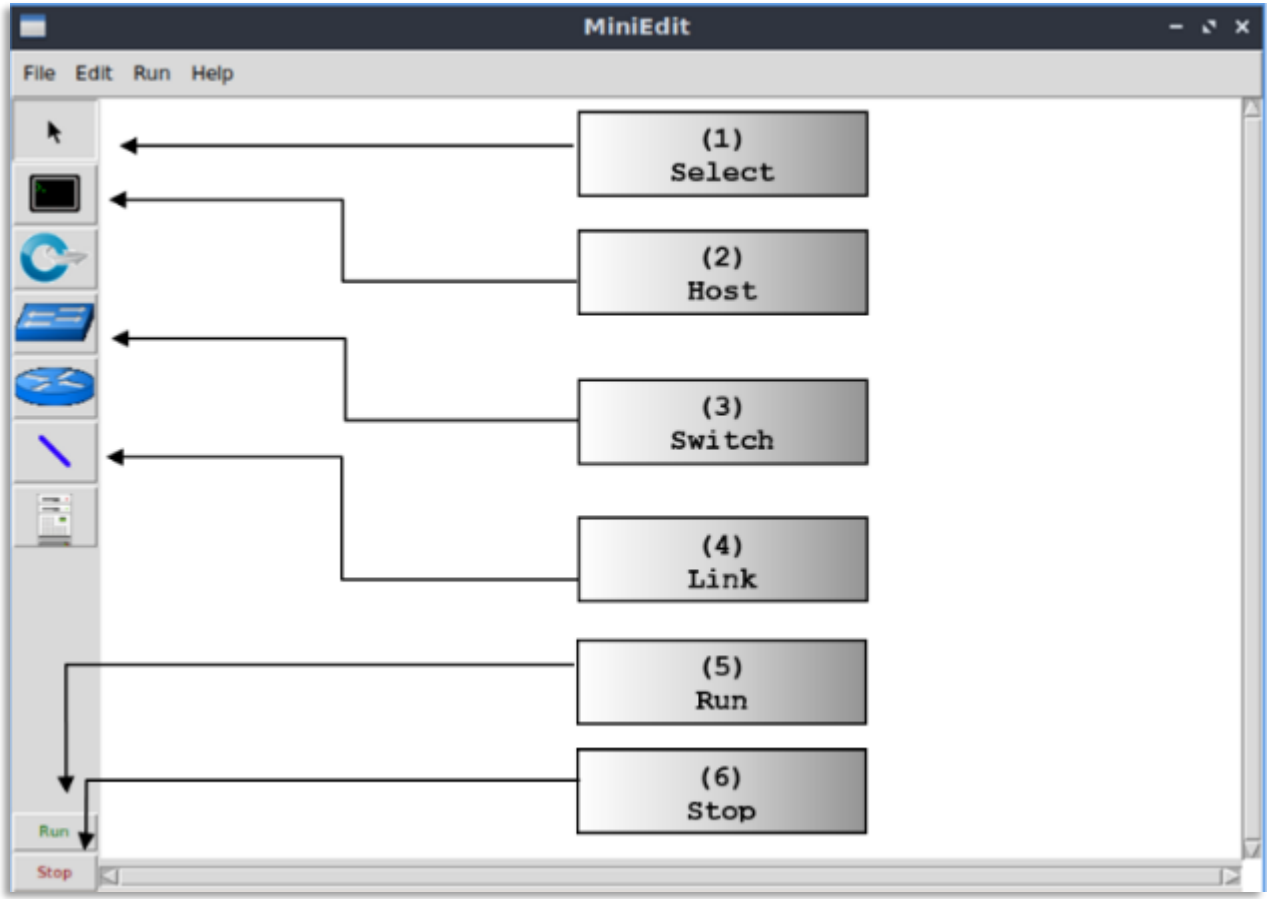
- Mininet provides network *emulation* opposed to simulation, allowing all network software at any layer to be simply run as is
- Mininet's logical nodes can be connected into networks
- Nodes are sometimes called containers, or more accurately, *network namespaces*
- Containers consume sufficiently few resources that networks of over a thousand nodes have been created, running on a single laptop



Mininet Emulated Network **Hardware Network**

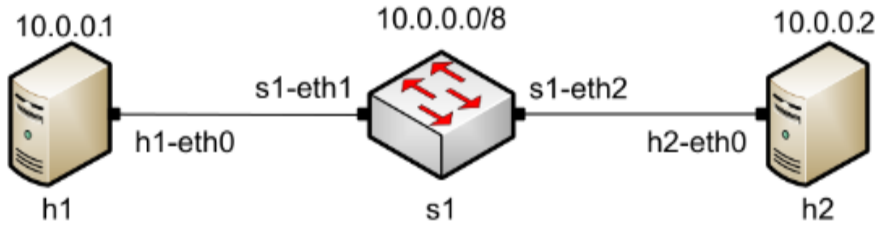
MiniEdit

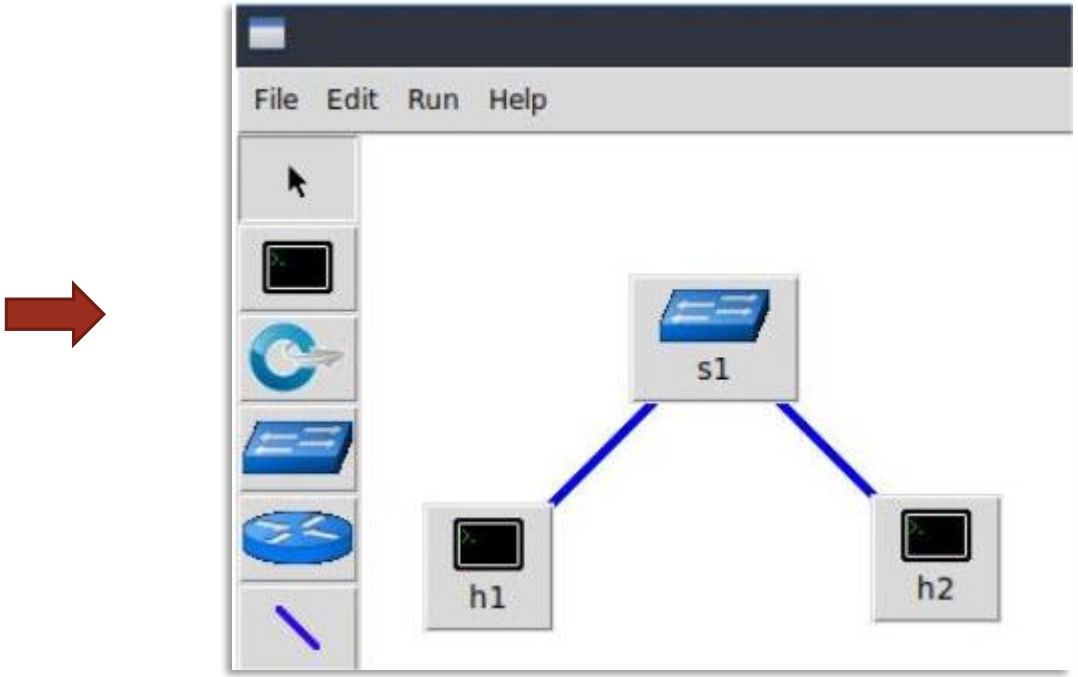
- MiniEdit is a simple GUI network editor for Mininet



MiniEdit

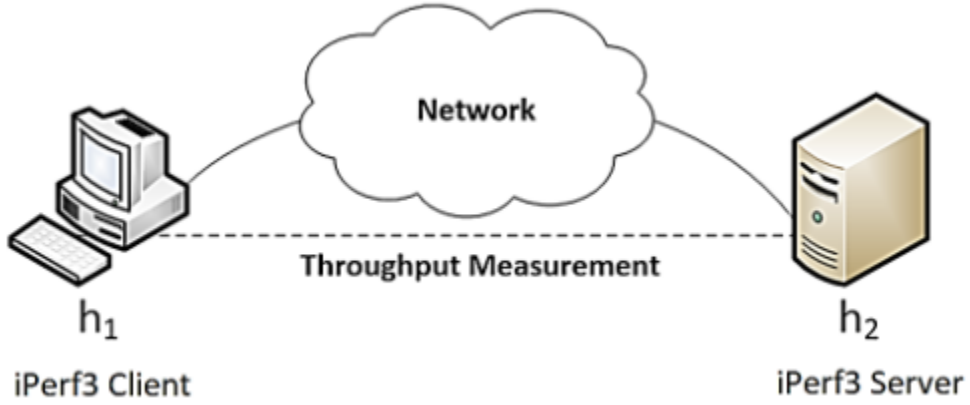
- To build Mininet's minimal topology, two hosts and one switch must be deployed





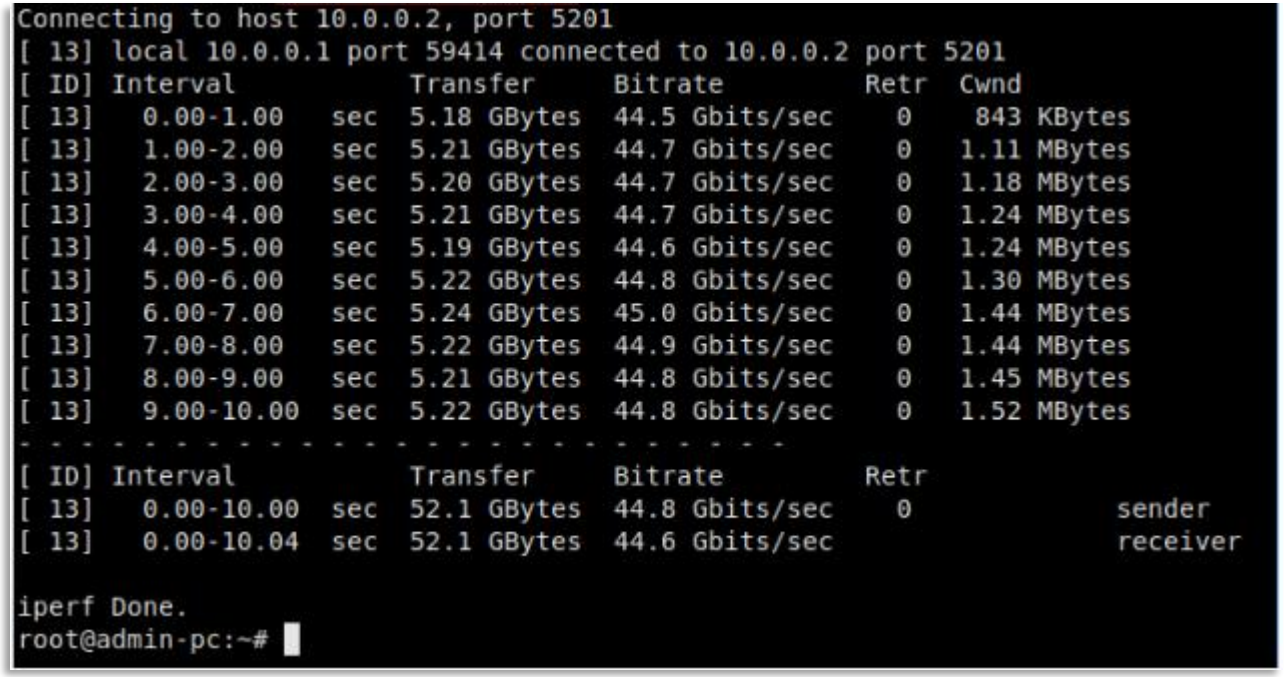
iPerf3

- iPerf3 is a real-time network throughput measurement tool
- It is an open source, cross-platform client-server application that can be used to measure the throughput between the two end devices
- Measuring throughput is particularly useful when experiencing network bandwidth issues such as delay, packet loss, etc.



iPerf3

- iPerf3 can operate on TCP, UDP, and SCTP, unidirectional or bidirectional way
- In iPerf3, the user can set *client* and *server* configurations via options and parameters
- iPerf3 outputs a timestamped report of the amount of data transferred and the throughput measured



TCP Buffer Size

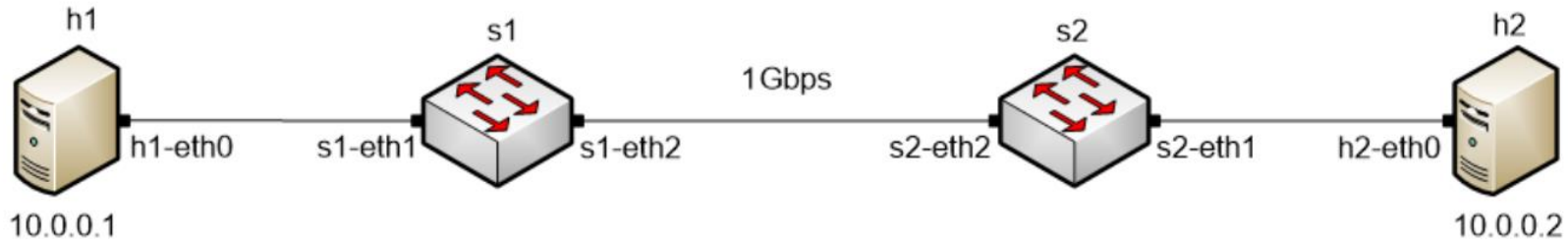
- In many WANs, the round-trip time (RTT) is dominated by the propagation delay
- To keep the sender busy while ACKs are received, the TCP buffer must be:

Traditional congestion controls:

TCP buffer size ≥ 2BDP

BBRv1 and BBRv2:

TCP buffer size must be considerable larger than 2BDP



Lab 8: Bandwidth-delay Product and TCP Buffer Size

Lab Goal and Topology

- Deploy emulated WANs in Mininet
- Modify the TCP send and receive buffers in Linux using sysctl tool
- Evaluate the performance with different buffer sizes
- Lab topology: