WORKSHOP ON NETWORK TOOLS AND PLATFORMS

Jorge Crichigno, Elie Kfoury, Jose Gomez
Department of Integrated Information Technology
University of South Carolina

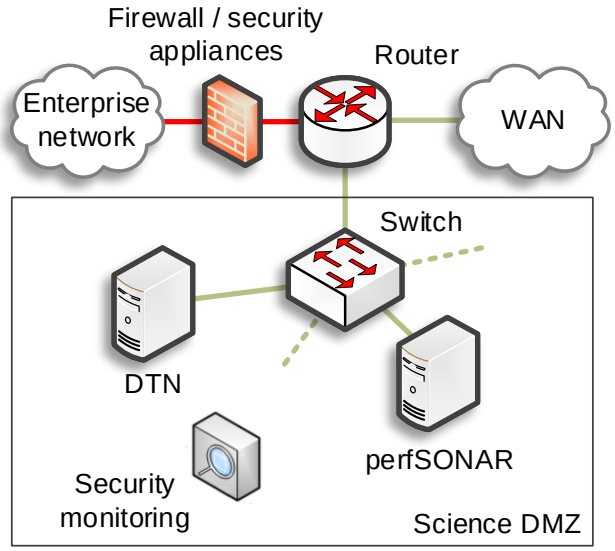
NSF Award 1829698
"CyberTraining CIP: Cyberinfrastructure Expertise on High-throughput Networks for Big Science Data Transfers"

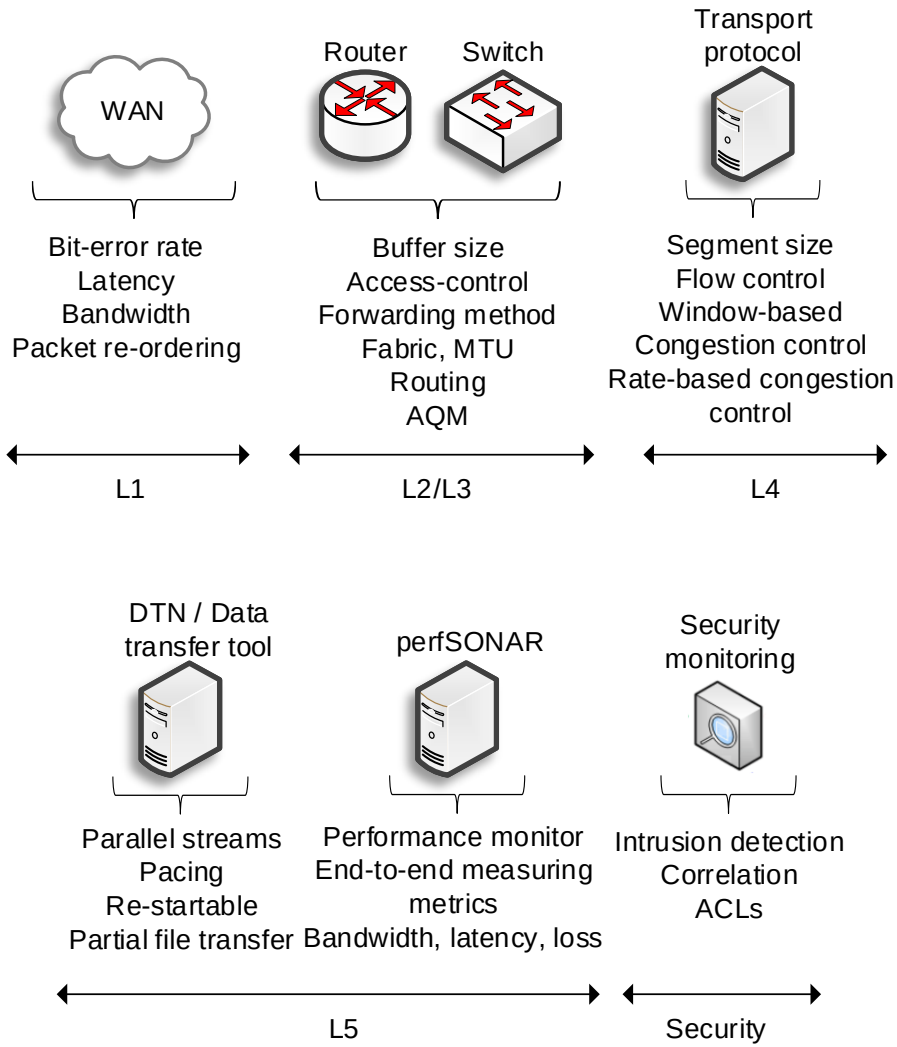
Hands-on Workshops

- The need for hands-on workshops
- Technical knowledge and skills to learn networks, high-speed protocols, security appliances, perfSONAR, etc.
- Earlier workshops: “Operating Innovative Networks” (CC*OIN): 2015 - 2017(see https://tinyurl.com/uf3fysz)

Hands-on Workshops

- Technical aspects





Hands-on Workshops

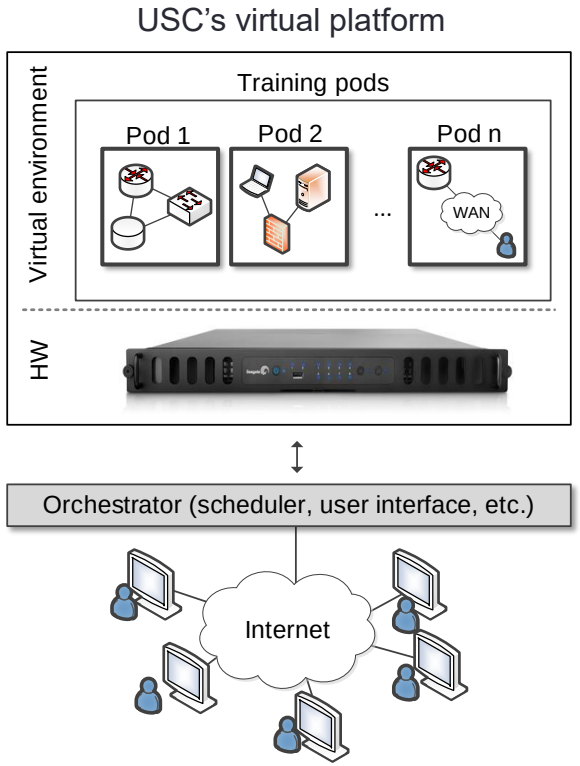
- Needs for hands-on tools at all levels
  - Industry certificate programs
  - Two-year technical colleges
  - Four-year universities
  - Graduate programs

Virtual Platform

- We aggregate content to run on a common platform
  - The Network Development Group (NDG) (netdevgroup.com)
- Multiple pod types
- A pod is a set of equipment need to complete a lab experiment
  - perfSONAR pod
  - Zeek pod
  - Next-generation firewall pod

  …

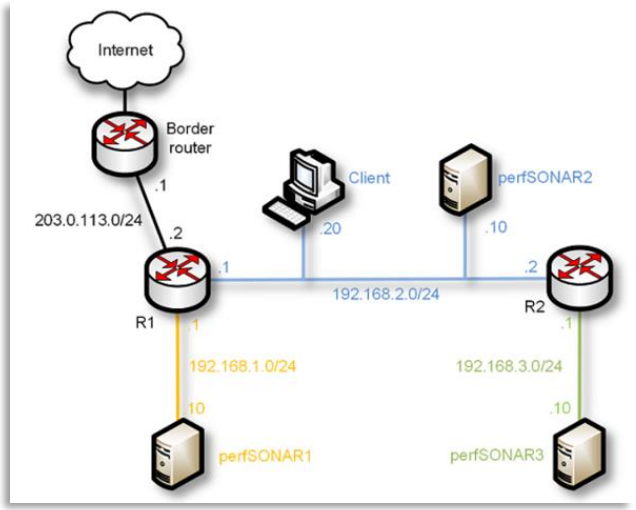
Today's platform



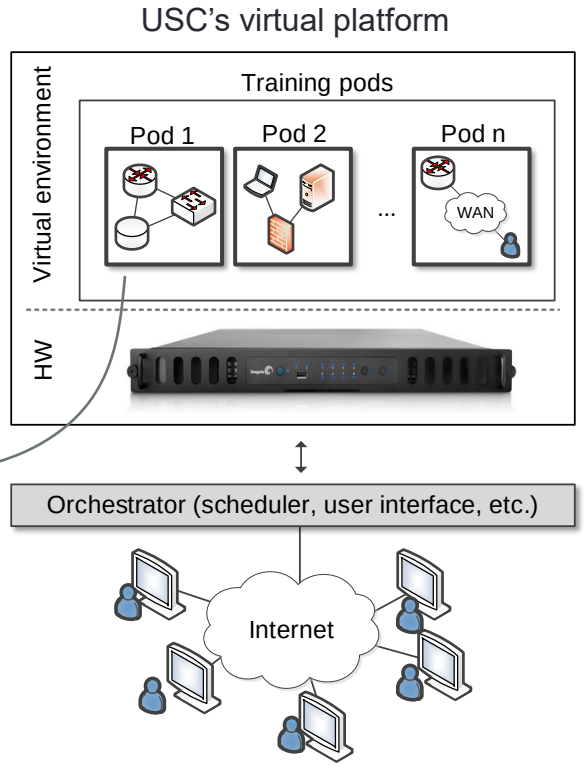
Virtual Platform

- We aggregate content to run on a common platform
  - The Network Development Group (NDG) (netdevgroup.com)
- Multiple pod types
- A pod is a set of equipment need to complete a lab experiment
  - perfSONAR pod
  - Zeek pod
  - Next-generation firewall pod

…



Today's platform



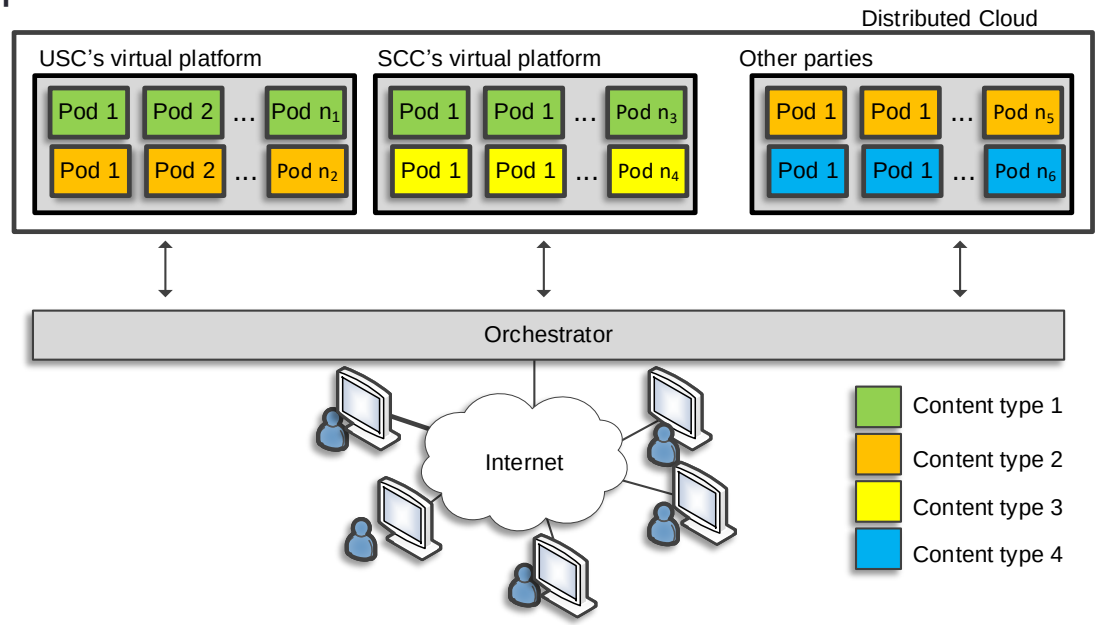
Virtual Platform

- We aggregate content to run on a common platform
  - The Network Development Group (NDG) (netdevgroup.com)
- Multiple pod types
- A pod is a set of equipment need to complete a lab experiment
  - perfSONAR pod
  - Zeek pod
  - Next-generation firewall pod

  …

Platform under development

University of South Carolina
NDG
Stanly Community College



"HANDS-ON SESSION TCP, NETWORK TOOLS"

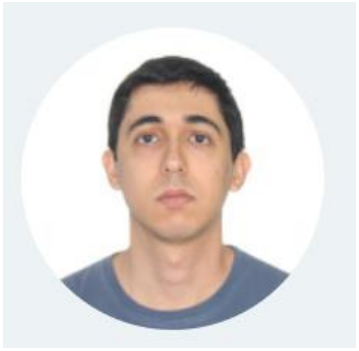
Jorge Crichigno, Elie Kfoury, Jose Gomez
Department of Integrated Information Technology
University of South Carolina

NSF Award 1829698
"CyberTraining CIP: Cyberinfrastructure Expertise on High-throughput Networks for Big Science Data Transfers"

Lab Series: Networks Tools and Protocols

- **Elie Kfoury** is a second-year PhD student in the College of Engineering and Computing at the University of South Carolina. His research focuses on programmable switches, telecommunications, network security, Blockchain, Internet of Things (IoT), and Software Defined Networking (SDN). Besides working in academia, Elie acts as the CTO of the Swedish startup “Secumobi”. The company develops security solutions backed by hardware encryption and Trusted Execution Environments (TEE).

- **Jose Gomez** is a second-year PhD student in the College of Engineering and Computing at the University of South Carolina. Prior to joining USC, he received his bachelor degree in Electrical Engineering from the Catholic University in Paraguay.



LAB SERIES: NETWORK TOOLS AND PROTOCOLS

Lab Series: Networks Tools and Protocols

- **Lab 1: Introduction to Mininet**
- **Lab 2: Introduction to iPerf**
- Lab 3: Emulating WAN with NETEM I Latency, Jitter
- Lab 4: Emulating WAN with NETEM II Packet Loss, Duplication, Reordering, and Corruption
- **Lab 5: Setting WAN Bandwidth with Token Bucket Filter (TBF)**
- Lab 6: Understanding Traditional TCP Congestion Control (HTCP, Cubic, Reno)
- Lab 7: Understanding 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 Series: Networks Tools and Protocols

- Lab 15: Analyzing the Impact of 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)

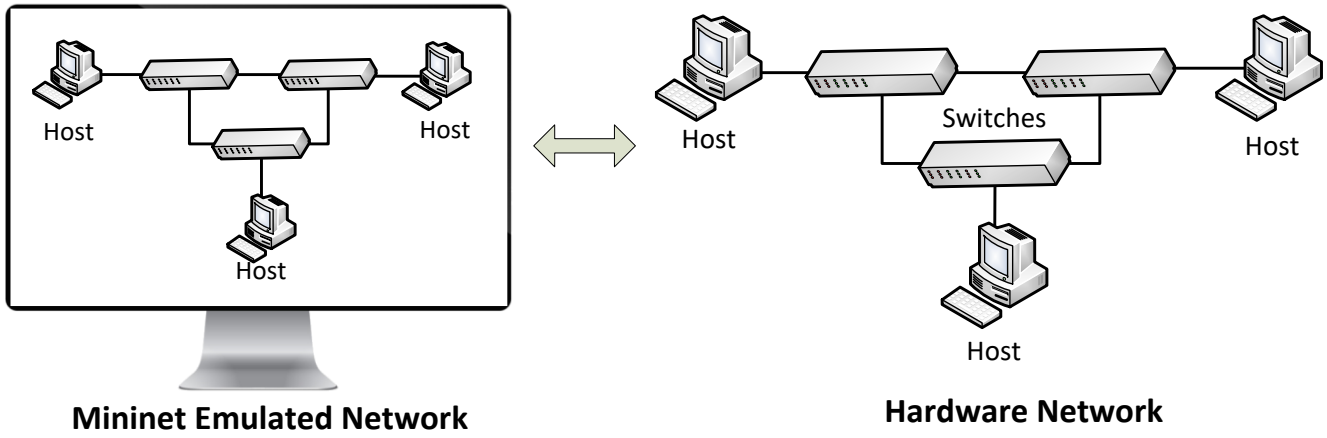
Organization of Lab Manuals

- 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

LAB 1: INTRODUCTION TO MININET

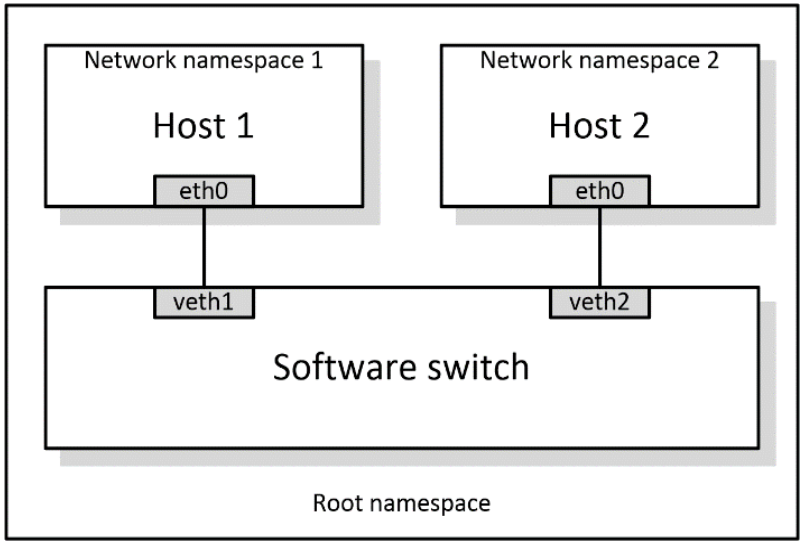
What is Mininet?

- A virtual testbed capable of recreating realistic scenarios
- It enables the development, testing of network protocols
- Inexpensive solution, real protocol stack, reasonably accurate



What is Mininet?

- Mininet nodes are network namespaces
  - Each node has different / separate virtual interfaces, routing tables
- Nodes use the underlying protocol stack of the host device
- Nodes are connected via virtual Ethernet (veth) links, which behave as Ethernet links



Where is Mininet used?

- Reproducing network research (buffers size, congestion control, long fat networks, datacenters)[1]
- Prototyping Software Defined Networking (SDN) ideas[2]
- Open Networking Foundation (ONF)[3]: SDN, NG-SDN, P4.org
- Teaching and research at undergraduate and graduate levels[4]

[1] B. Heller, “Reproducing network research with high-fidelity emulation," http://yuba.stanford.edu/~nickm/papers/brandon-thesis.pdf

[2] De Oliveira, Rogério Leão Santos, et al. "Using mininet for emulation and prototyping software-defined networks." 2014 IEEE Colombian Conference on Communications and Computing (COLCOM), 2014.

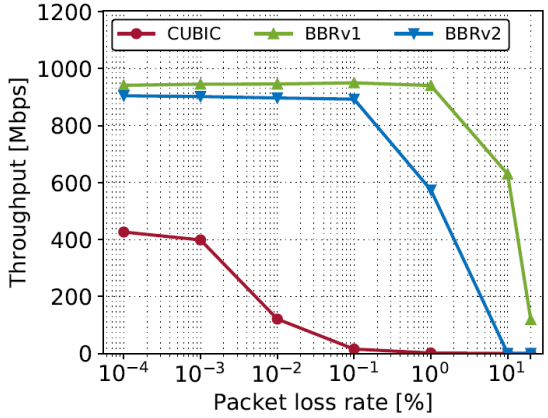
[3] Open Networking Foundation, “Mininet ONF," https://www.opennetworking.org/mininet/

[4] Lantz, Bob, Brandon Heller, and Nick McKeown. "A network in a laptop: rapid prototyping for software-defined networks." Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, 2010.

Where is Mininet used?

- Reproducing network research (buffers size, congestion control, long fat networks, datacenters)[1]
- Prototyping Software Defined Networking (SDN) ideas[2]
- Open Networking Foundation (ONF)[3]: SDN, NG-SDN, P4.org
- Teaching and research at undergraduate and graduate levels[4]

Throughput tests over a
1 Gbps 100ms network



[1] B. Heller, “Reproducing network research with high-fidelity emulation," http://yuba.stanford.edu/~nickm/papers/brandon-thesis.pdf

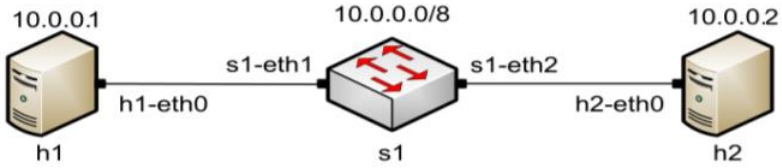
[2] De Oliveira, Rogério Leão Santos, et al. "Using mininet for emulation and prototyping software-defined networks." 2014 IEEE Colombian Conference on Communications and Computing (COLCOM), 2014.

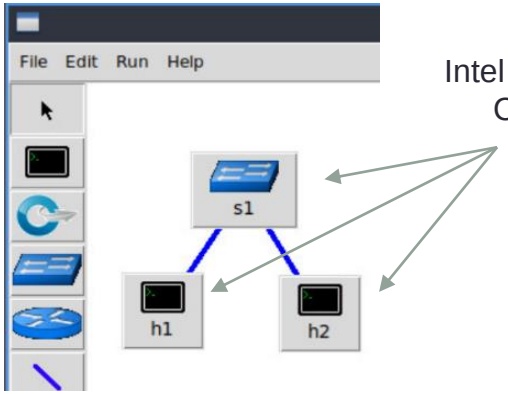
[3] Open Networking Foundation, “Mininet ONF," https://www.opennetworking.org/mininet/

[4] Lantz, Bob, Brandon Heller, and Nick McKeown. "A network in a laptop: rapid prototyping for software-defined networks." Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, 2010.

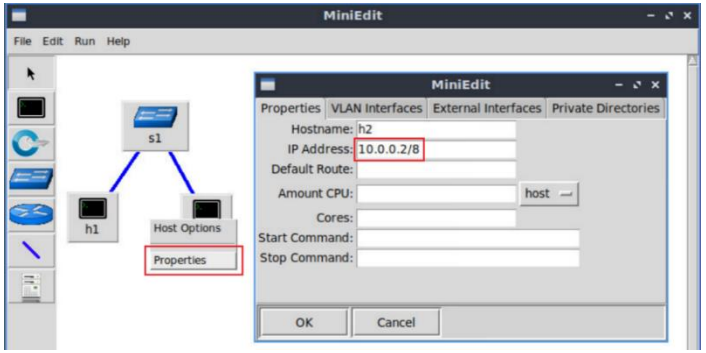
Pod Design with Mininet

- A pod is a set of virtual equipment (routers, switches, etc.)
- A pod is created every time a lab reservation is made
- For the Network Tools and Protocols Lab series, pods are embedded into Mininet





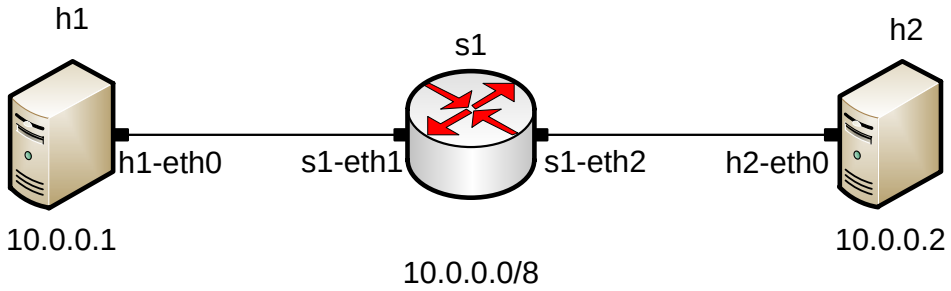
Intel Xeon Gold 6130
CPU, 2.1 GHz



LAB 2: INTRODUCTION TO IPERF

What is iPerf?

- iPerf is a real-time network throughput measurement tool
- The user can set client (sender) and server (receiver)



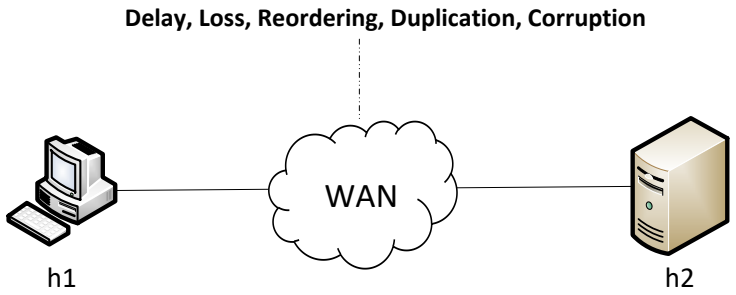

```
"Host: h1"
root@admin-pc:~# iperf3 -c 10.0.0.2
Connecting to host 10.0.0.2, port 5201
[ 19] local 10.0.0.1 port 52850 connected to 10.0.0.2 port 5201
[ ID] Interval           Transfer     Bitrate         Retr  Cwnd
[ 19]   0.00-1.00   sec  5.11 GBytes  43.9 Gbits/sec    0    730 KBytes
[ 19]   1.00-2.00   sec  5.09 GBytes  43.7 Gbits/sec    0    887 KBytes
[ 19]   2.00-3.00   sec  5.14 GBytes  44.1 Gbits/sec    0    887 KBytes
[ 19]   3.00-4.00   sec  5.15 GBytes  44.2 Gbits/sec    0    932 KBytes
[ 19]   4.00-5.00   sec  5.15 GBytes  44.2 Gbits/sec    0    932 KBytes
[ 19]   5.00-6.00   sec  5.19 GBytes  44.6 Gbits/sec    0    932 KBytes
[ 19]   6.00-7.00   sec  5.18 GBytes  44.5 Gbits/sec    0    932 KBytes
[ 19]   7.00-8.00   sec  5.05 GBytes  43.4 Gbits/sec    0   2.19 MBytes
[ 19]   8.00-9.00   sec  5.15 GBytes  44.3 Gbits/sec    0   2.19 MBytes
[ 19]   9.00-10.00  sec  5.15 GBytes  44.3 Gbits/sec    0   2.19 MBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[ 19]   0.00-10.00  sec  51.4 GBytes  44.1 Gbits/sec    0             sender
[ 19]   0.00-10.04  sec  51.4 GBytes  43.9 Gbits/sec                  receiver

iperf Done.
root@admin-pc:~#
```

LAB 5: SETTING WAN BANDWIDTH WITH TOKEN BUCKET FILTER (TBF)

Network Emulator (NETEM)

- A Linux network emulator that reproduces long-distance WANs in the lab environment




```
"Host: h1"
root@admin-pc:~# iperf3 -c 10.0.0.2
Connecting to host 10.0.0.2, port 5201
[ 13] local 10.0.0.1 port 52822 connected to 10.0.0.2 port 5201
[ ID] Interval           Transfer     Bitrate         Retr  Cwnd
[ 13]   0.00-1.00   sec  4.87 GBytes  41.9 Gbits/sec  319   233 KBytes
[ 13]   1.00-2.00   sec  4.86 GBytes  41.8 Gbits/sec  471   181 KBytes
[ 13]   2.00-3.00   sec  4.86 GBytes  41.8 Gbits/sec  454   199 KBytes
[ 13]   3.00-4.00   sec  4.92 GBytes  42.2 Gbits/sec  408   255 KBytes
[ 13]   4.00-5.00   sec  4.92 GBytes  42.2 Gbits/sec  410   375 KBytes
[ 13]   5.00-6.00   sec  4.92 GBytes  42.2 Gbits/sec  542   243 KBytes
[ 13]   6.00-7.00   sec  4.90 GBytes  42.1 Gbits/sec  454   249 KBytes
[ 13]   7.00-8.00   sec  4.84 GBytes  41.6 Gbits/sec  545   182 KBytes
[ 13]   8.00-9.00   sec  4.92 GBytes  42.3 Gbits/sec  498   202 KBytes
[ 13]   9.00-10.00  sec  4.90 GBytes  42.1 Gbits/sec  248   209 KBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[ 13]   0.00-10.00  sec  48.9 GBytes  42.0 Gbits/sec  4349            sender
.00-10.04  sec  48.9 GBytes  41.8 Gbits/sec                  receiver
e.
n-pc:~#
```

High loss-rate scenario

```
"Host: h1"
root@admin-pc:~# ping 10.0.0.2
PING 10.0.0.2 (10.0.0.2) 56(84) bytes of data.
64 bytes from 10.0.0.2: icmp_seq=1 ttl=64 time=101 ms
64 bytes from 10.0.0.2: icmp_seq=2 ttl=64 time=100 ms
64 bytes from 10.0.0.2: icmp_seq=3 ttl=64 time=100 ms
64 bytes from 10.0.0.2: icmp_seq=4 ttl=64 time=100 ms
64 bytes from 10.0.0.2: icmp_seq=5 ttl=64 time=100 ms
^C
--- 10.0.0.2 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 8ms
rtt min/avg/max/mdev = 100.069/100.253/100.961/0.354 ms
root@admin-pc:~#
```

Long-distance 100ms RTT scenario

Network Emulator (NETEM)

- A Linux network emulator that reproduces long-distance WANs in the lab environment
- Example applications:
  - ESnet[1]
  - Lawrence Berkeley National Laboratory[2]
  - Evaluate TCP enhancements[3]
  - Evaluate Active Queueing Management (AQM) algorithms[4]
  - Evaluate network behavior in various conditions (loss, latency, duplication)

[1] ESnet Energy Sciences Network, “Network Switch Buffer Testing,” Online: https://fasterdata.es.net/network-tuning/router-switch-buffer-size-issues/switch-buffer-testing/

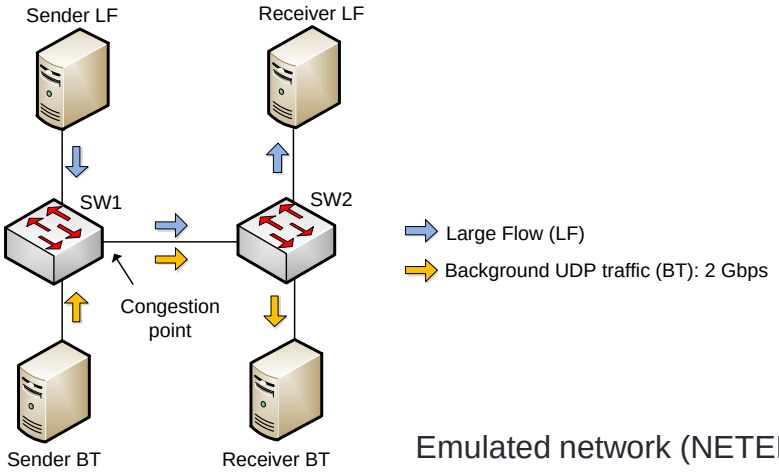
[2] Kissel, Ezra, et al. "Scalable integrated performance analysis of multi-gigabit networks." 2012 IEEE Network Operations and Management Symposium, IEEE, 2012.

[3] Hock, Mario, Roland Bless, and Martina Zitterbart. "Experimental evaluation of BBR congestion control." 2017 IEEE 25th International Conference on Network Protocols (ICNP), IEEE, 2017.

[4] Gong, Yixi, et al. "Fighting the Bufferbloat: on the coexistence of AQM and low priority congestion control." Computer Networks 65 (2014): 255-267.

Network Emulator (NETEM)

- ESnet vs emulated network, 10 Gbps links, 70 msec RTT

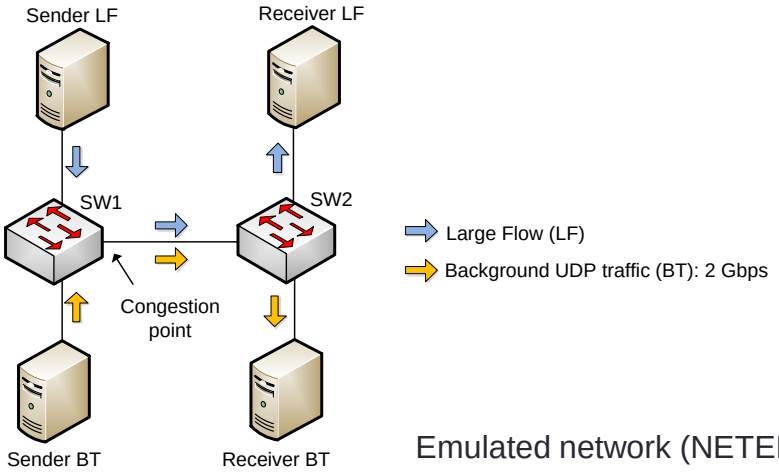


Emulated network (NETEM)

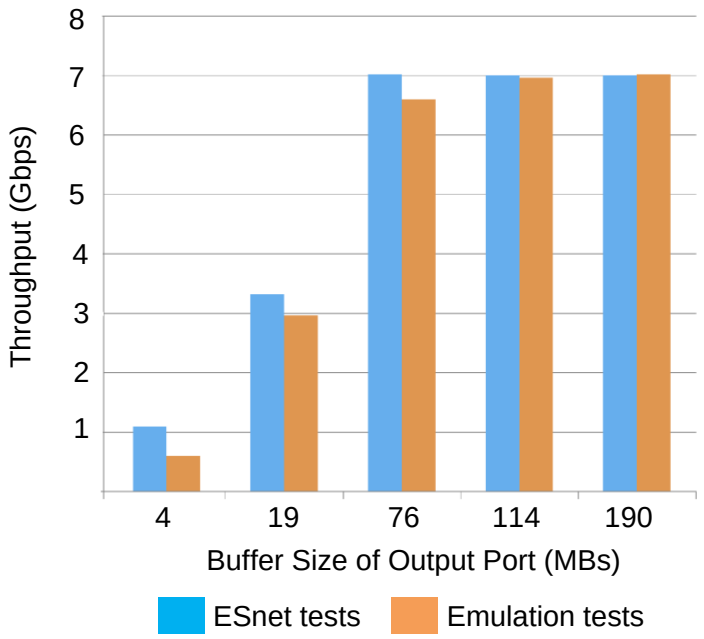
M. Smitasin and B. Tierney, “Evaluating network buffer size requirements,” Proc. Technol. Exchange Workshop, Oct. 2015. Online: https://meetings.internet2.edu/media/medialibrary/2015/10/05/20151005-smitasin-buffersize.pdf

Network Emulator (NETEM)

- ESnet vs emulated network, 10 Gbps links, 70 msec RTT



Emulated network (NETEM)



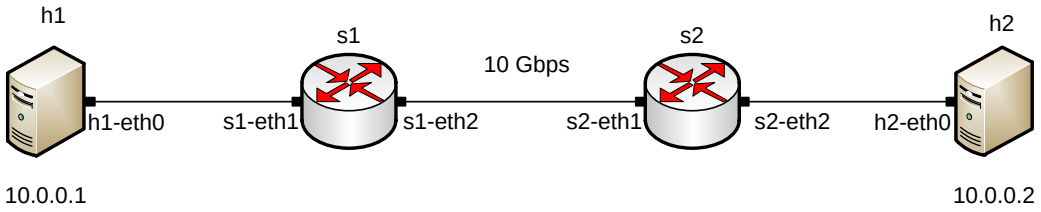
M. Smitasin and B. Tierney, “Evaluating network buffer size requirements,” Proc. Technol. Exchange Workshop, Oct. 2015. Online: https://meetings.internet2.edu/media/medialibrary/2015/10/05/20151005-smitasin-buffersize.pdf

Token Bucket Algorithm

- Algorithm used in networks to limit the bandwidth and the burstiness of the traffic
- Linux token bucket implementation is called token bucket filter (TBF)
- Parameters
  - Link rate
  - Maximum burst size
  - Buffer size

Token Bucket Algorithm

- Algorithm used in networks to limit the bandwidth and the burstiness of the traffic
- Linux token bucket implementation is called token bucket filter (TBF)
- Parameters
  - Link rate
  - Maximum burst size
  - Buffer size



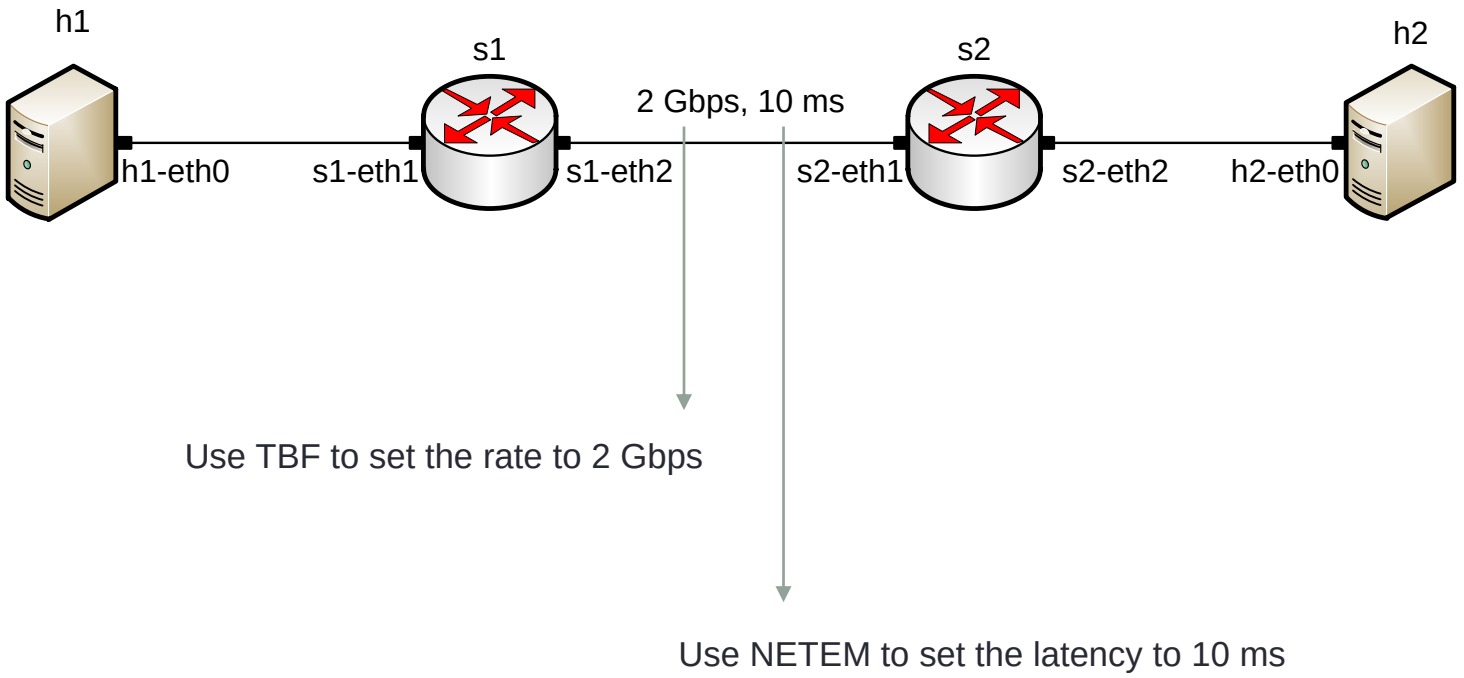

```
"Host: h1"
root@admin-pc:~# iperf3 -c 10.0.0.2
Connecting to host 10.0.0.2, port 5201
[ 15] local 10.0.0.1 port 33102 connected to 10.0.0.2 port 5201
[ ID] Interval           Transfer     Bitrate         Retr  Cwnd
[ 15]   0.00-1.00   sec  1.12 GBytes  9.62 Gbits/sec    0    540 KBytes
[ 15]   1.00-2.00   sec  1.11 GBytes  9.56 Gbits/sec    0    597 KBytes
[ 15]   2.00-3.00   sec  1.11 GBytes  9.57 Gbits/sec    0    655 KBytes
[ 15]   3.00-4.00   sec  1.11 GBytes  9.57 Gbits/sec    0    721 KBytes
[ 15]   4.00-5.00   sec  1.11 GBytes  9.56 Gbits/sec    0    757 KBytes
[ 15]   5.00-6.00   sec  1.11 GBytes  9.57 Gbits/sec    0   1.01 MBytes
[ 15]   6.00-7.00   sec  1.11 GBytes  9.56 Gbits/sec    0   1.01 MBytes
[ 15]   7.00-8.00   sec  1.11 GBytes  9.56 Gbits/sec    0   1.06 MBytes
[ 15]   8.00-9.00   sec  1.11 GBytes  9.56 Gbits/sec    0   1.06 MBytes
[ 15]   9.00-10.00  sec  1.11 GBytes  9.56 Gbits/sec    0   1.06 MBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[ 15]   0.00-10.00  sec  11.1 GBytes  9.57 Gbits/sec    0             sender
[ 15]   0.00-10.04  sec  11.1 GBytes  9.53 Gbits/sec                  receiver

iperf Done.
root@admin-pc:~#
```

Test using Linux's TBF, rate 10 Gbps

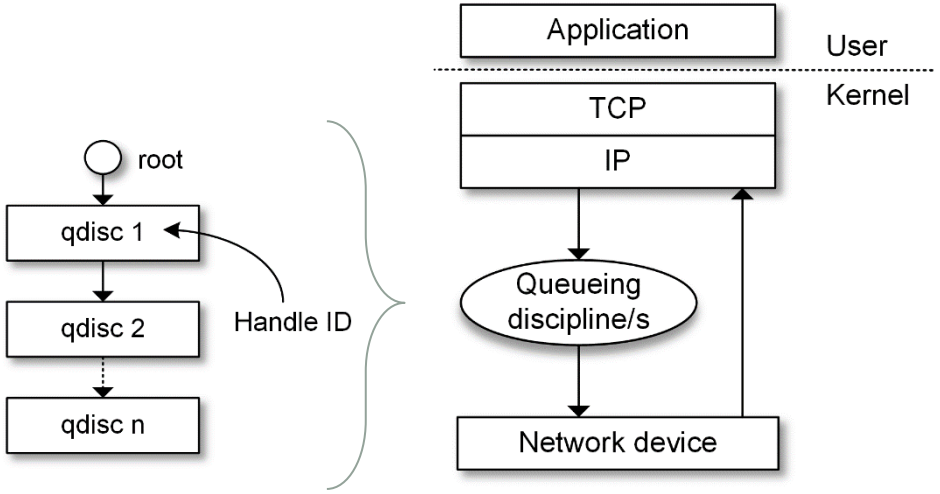
Lab 5 Topology

- Two hosts: h1 and h2; two switches: s1 and s2
- 2 Gbps, 10 ms latency from s1 to s2



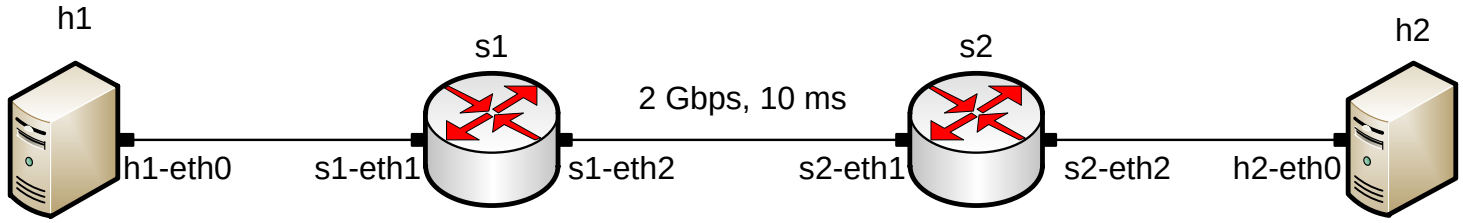
Lab 5 Topology

- NETEM is used to emulate latency
- TBF is used to limit the rate of the link s1-s2
- Then, they are combined as “queueing disciplines”
- A queuing discipline is an object with two interfaces



Lab 5 Topology

- Two hosts: h1 and h2; two switches: s1 and s2
- 2 Gbps,10 ms latency from s1 to s2




```
"Host: h1"
--- 10.0.0.2 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 7ms
rtt min/avg/max/mdev = 10.067/10.208/10.611/0.243 ms
root@admin-pc:~# iperf3 -c 10.0.0.2
Connecting to host 10.0.0.2, port 5201
[ 15] local 10.0.0.1 port 44496 connected to 10.0.0.2 port 5201
[ ID] Interval           Transfer     Bitrate         Retr  Cwnd
[ 15]   0.00-1.00   sec   222 MBytes  1.86 Gbits/sec  540   3.64 MBytes
[ 15]   1.00-2.00   sec   228 MBytes  1.91 Gbits/sec    0   3.90 MBytes
[ 15]   2.00-3.00   sec   229 MBytes  1.92 Gbits/sec  135   2.89 MBytes
[ 15]   3.00-4.00   sec   228 MBytes  1.91 Gbits/sec    0   3.03 MBytes
[ 15]   4.00-5.00   sec   229 MBytes  1.92 Gbits/sec   45   2.20 MBytes
[ 15]   5.00-6.00   sec   221 MBytes  1.86 Gbits/sec    0   2.33 MBytes
[ 15]   6.00-7.00   sec   229 MBytes  1.92 Gbits/sec    0   2.43 MBytes
[ 15]   7.00-8.00   sec   228 MBytes  1.91 Gbits/sec    0   2.52 MBytes
[ 15]   8.00-9.00   sec   229 MBytes  1.92 Gbits/sec    0   2.58 MBytes
[ 15]   9.00-10.00  sec   228 MBytes  1.91 Gbits/sec    0   2.62 MBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[ 15]   0.00-10.00  sec  2.21 GBytes  1.90 Gbits/sec  720             sender
[ 15]   0.00-10.04  sec  2.21 GBytes  1.89 Gbits/sec                  receiver

iperf Done.
root@admin-pc:~#
```

Access to Virtual Platform

Availability of Virtual Platform DURING Workshop, May 4 – May 6

- A total of 110 attendees were allocated resources to access the virtual platform
- Your credentials **have been sent** (check email used for registration)
- Please contact Assistant 1-4 for questions

Availability of Virtual Platform AFTER Workshop, May 6 – August 6

- URL: https://netlab.cec.sc.edu/
- Username: <email address used for registration>
- Password: nsf2020