Workshop on P4 Programmable Switches

Hands-on Session 3: Monitoring Buffer Occupancy in P4 Switches

Elie Kfoury, Jorge Crichigno
College of Engineering and Computing, University of South Carolina
http://ce.sc.edu/cyberinfra

University of South Carolina (USC)

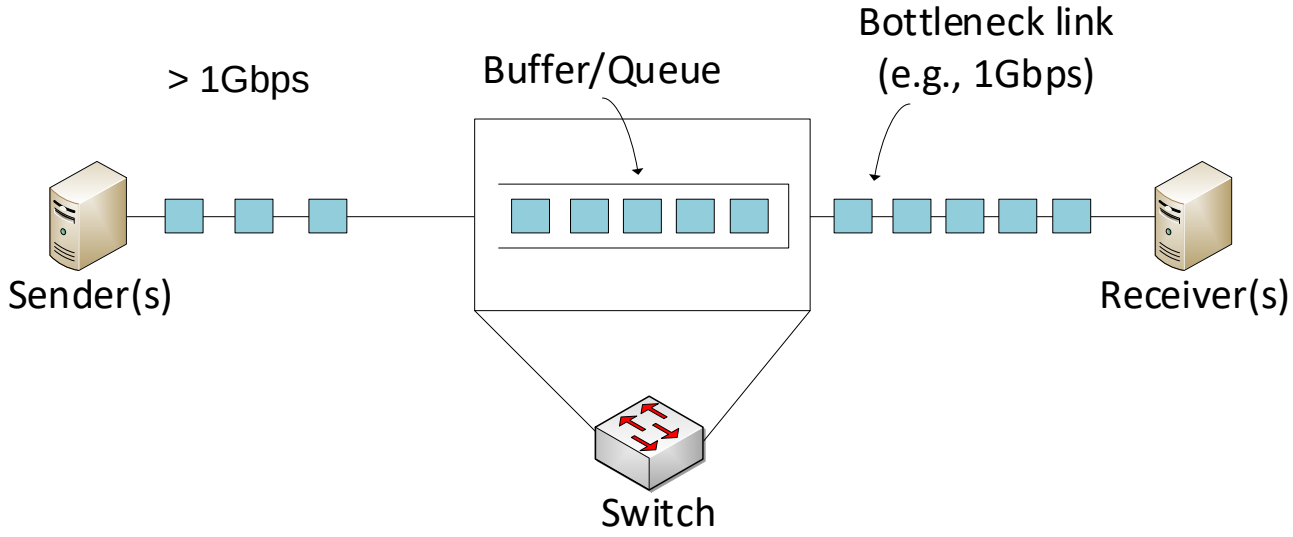
August 29, 2023

Monitoring the Switch's Queue using Standard Metadata

Lab activities are described in Lab 5, P4 Programmable Data Planes: Applications, Stateful Elements, and Custom Packet Processing lab series

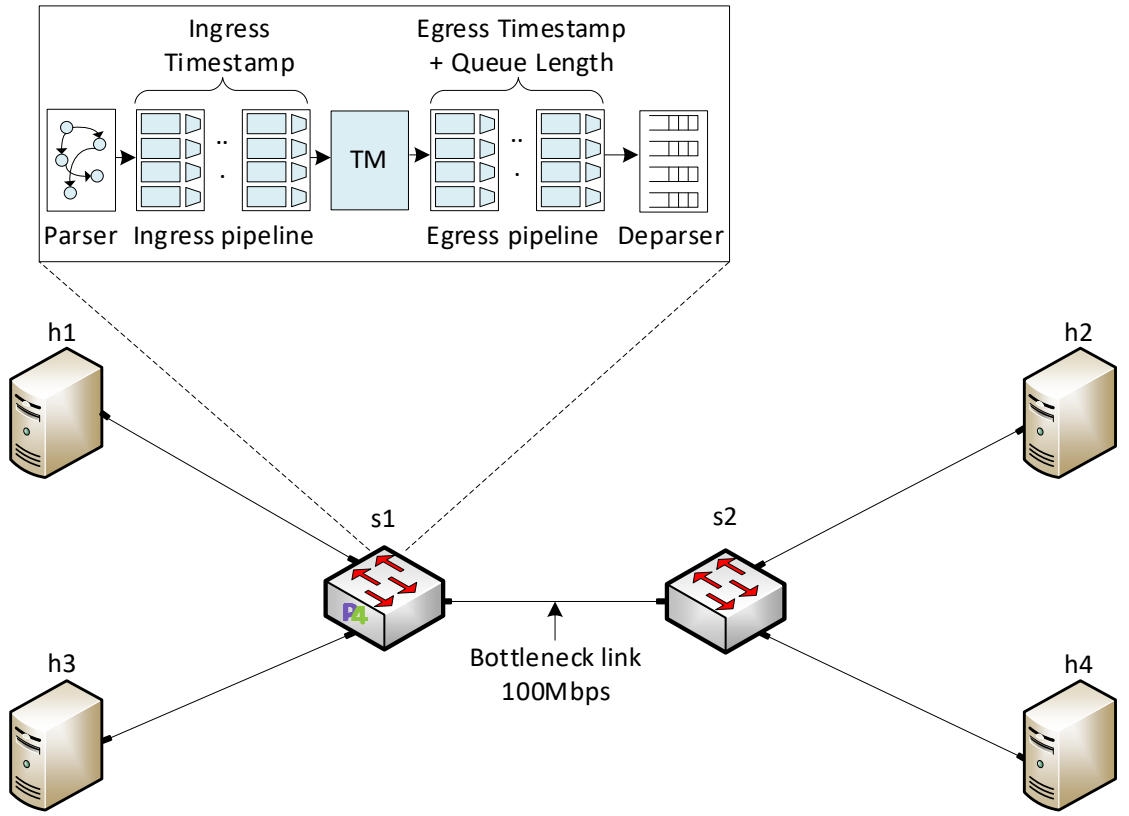
Queueing Delay

- As a packet travels from the sender to the receiver, it experiences several types of delays at each node (router/switch) along the path
- Queueing delay: the time a packet waits for transmission onto the link, in the order of microseconds to milliseconds
- Queue builds up when the output link is fully utilized (i.e., link becomes the bottleneck)



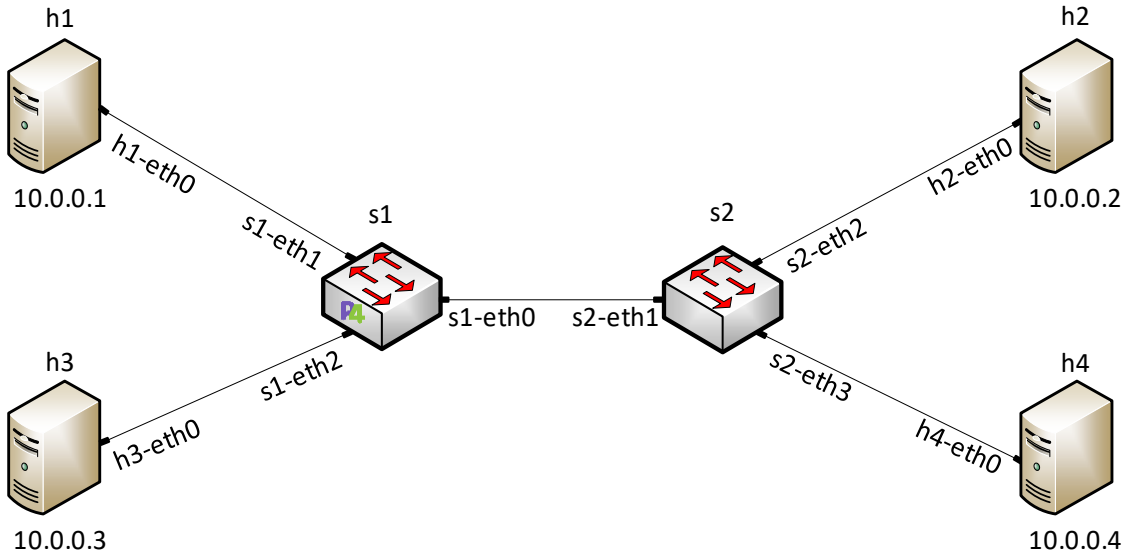
Computing Queueing Delay in P4

- The standard metadata in the switch contains:
  - The enqueue timestamp (`enq_timestamp`), which indicates when the packet is enqueued
  - The egress timestamp (`egress_global_timestamp`), which indicates when the packet enters the egress pipeline
- The difference between the two timestamps approximates the queueing delay
- The standard metadata in the switch also contains:
  - The queue depth (`q_depth`), which indicates the number of packets in the queue



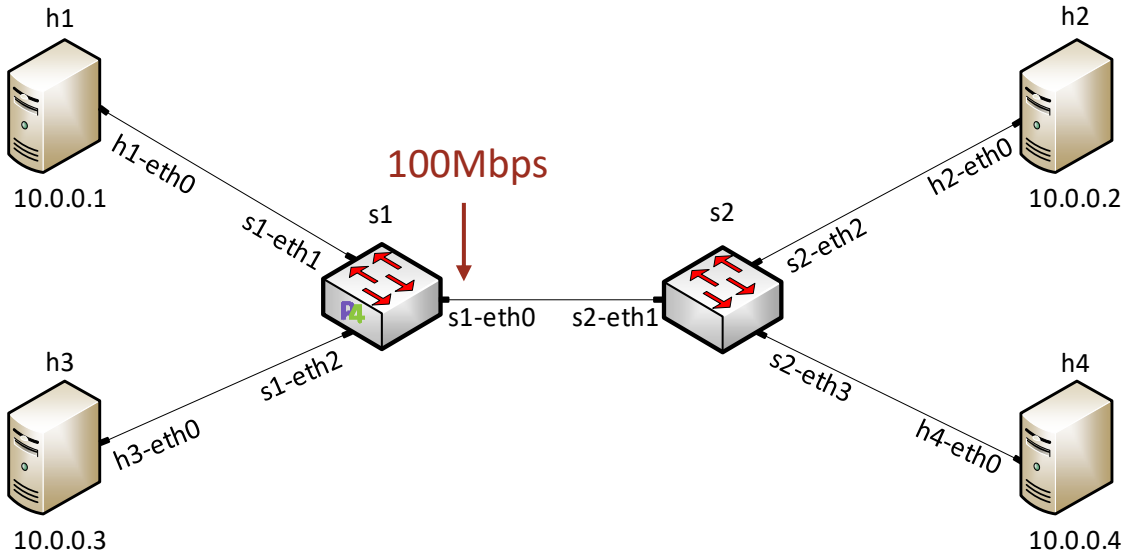
Lab Topology and Objectives

- The topology consists of four hosts: h1, h2, h3, and h4; one P4 switch: s1; and one legacy switch: s2
- The goal is to observe the queue occupancy on the switch s1



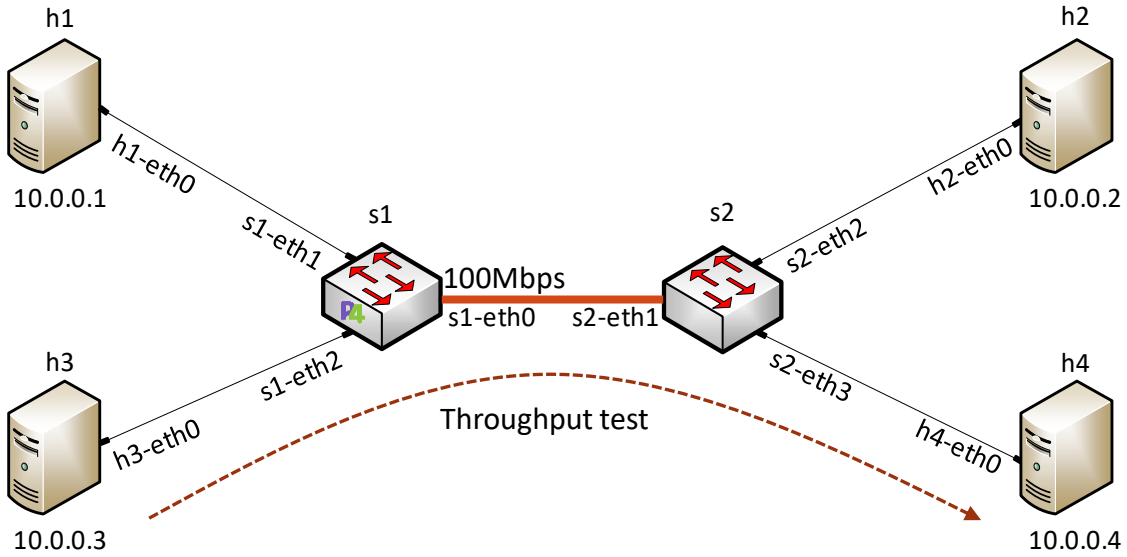
Lab Topology and Objectives

- The topology consists of four hosts: h1, h2, h3, and h4; one P4 switch: s1; and one legacy switch: s2
- The goal is to observe the queue occupancy on the switch s1
- Lab objectives:
  - Limit the rate of the output port of the switch and set the queue size



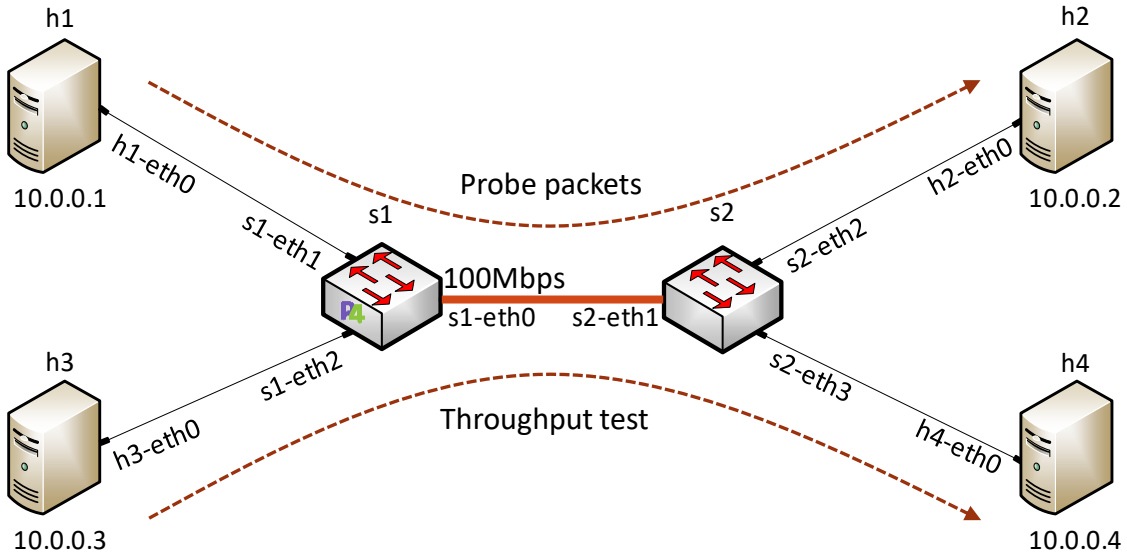
Lab Topology and Objectives

- The topology consists of four hosts: h1, h2, h3, and h4; one P4 switch: s1; and one legacy switch: s2
- The goal is to observe the queue occupancy on the switch s1
- Lab objectives:
  - Limit the rate of the output port of the switch and set the queue size
  - Run a throughput test from h3 to h4 using the iPerf3 tool



Lab Topology and Objectives

- The topology consists of four hosts: h1, h2, h3, and h4; one P4 switch: s1; and one legacy switch: s2
- The goal is to observe the queue occupancy on the switch s1
- Lab objectives:
  - Limit the rate of the output port of the switch and set the queue size
  - Run a throughput test from h3 to h4 using the iPerf3 tool
  - Observe the queue occupancy by sending probe packets from h1 to h2 using a custom protocol

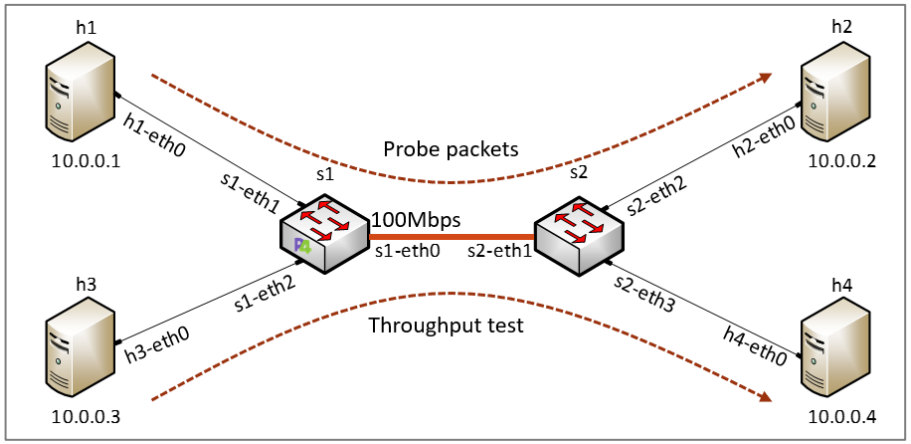


Custom Probing Protocol

- The custom protocol will be added by the sender (the device probing) on top of the IPv4 header (when `ipv4.protocol = 0xFD`)
- The fields are initialized to 0
- The P4 switch parses the custom protocol header and overwrites its fields

Field name	Size [bits]	Description
switch_ID	8	Stores the switch identifier
enq_timestamp	32	Stores the timestamp set when the packet is enqueued
deq_timestamp	48	Stores the timestamp set when the packet shows up on egress
q_delay	48	Stores deq_timestamp - enq_timestamp
q_depth	24	Stores the current number of packets in the queue

Header fields of the custom protocol



Topology