Effects of pause and QCN on TCP sources

Berk Atikoglu, Balaji Prabhakar

Overview

- Analyze the behavior of the TCP sources
  - With and without Pause and/or QCN
  - Scenario: Baseline

TCP: Discussion

- TCP is very sensitive to drops
- When many drops occur, TCP doesn't use Fast Re-transmit; instead it waits for a timer to expire before resending packets and increasing window size

- Default TCP RTO parameters:
  - Initial RTO: 3s; Minimum RTO: 1s
  - This is too large for our setting

- Our TCP RTO parameters
  - Initial RTO: 0.6s; Minimum RTO: 01.s
  - Following: au-sim-geisler-cm-tcp-effects-1107-v1.pdf

- Maximum window size: 65536 bytes
- TCP version: TCP Reno
- Other parameters have default values

QCN and pause parameters

- W = 2.0
- Q_EQ = 26 Kbytes
- Gd = 1/128 = 0.0078125
- Base marking: once every 150kbytes
- Jitter on marking: 30%
- MIN_RATE = 10Mb/s
- BC_LIMIT = 150kbytes
- TIMER_PERIOD = 15ms
- R_AI = 5Mbps
- R_HAI = 50Mbps
- FAST_RECOVERY_TH = 5
- Quantized_Fb: 6 bits
- Jitter at RP: 30% (byte counter and timer)

- Pause threshold: 120Kbytes
- Pause interval: 96usec

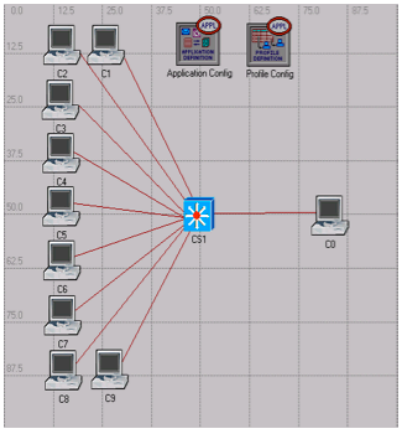
Scenario: OG Hotspot

- Baseline scenario
  - 10 nodes connected through a single switch
  - TCP connection from nodes 1--9 to node 0
  - All connections start at time 0s
  - Each link is 10Gb/s
  - Simulation duration: 0--3 seconds
  - RTT: 25us, 250us
- Switch o/p buffer size: 225Kbytes
- RL buffer size: Unlimited
- Hotspot
  - Service at one link is decreased to 2 Gb/s
  - Hotspot duration: 1sec -- 2 sec
- Simulations
  - No Pause, no QCN
  - Pause only
  - QCN only
  - Both QCN and pause

Topology & Workload

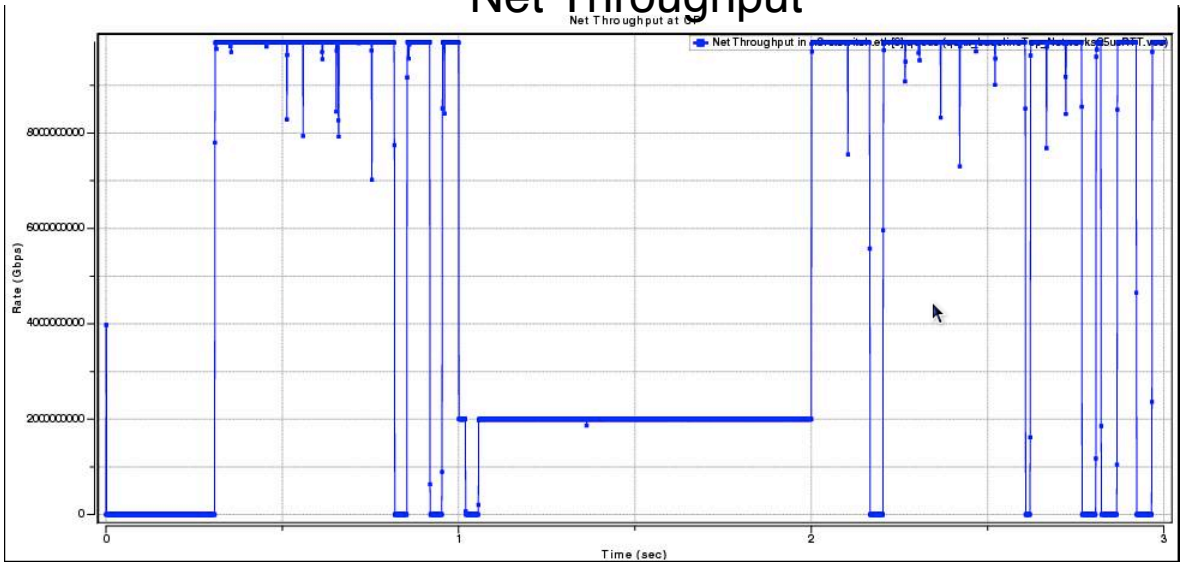
Single-Hop with output hotspot

10 Gb/s links, 2 Gb/s hotspot @ CS1→C0

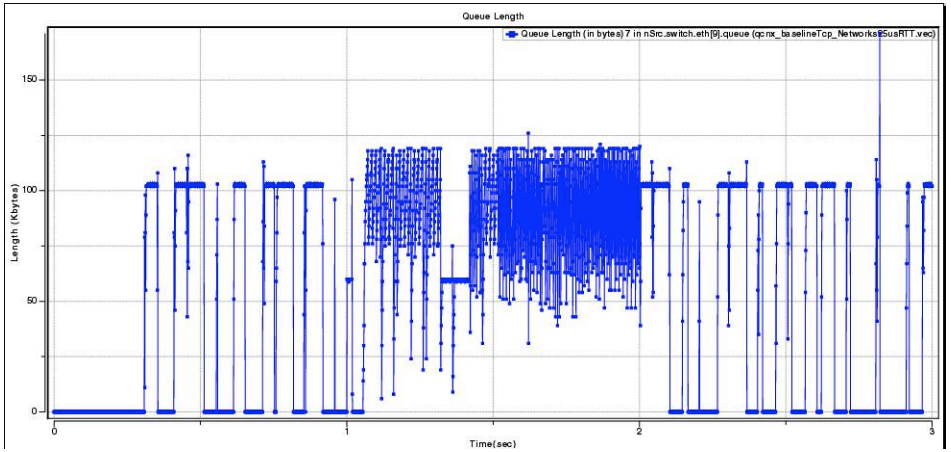


No Pause, no QCN; RTT = 25us
(Buffer overflows result in dropped packets)

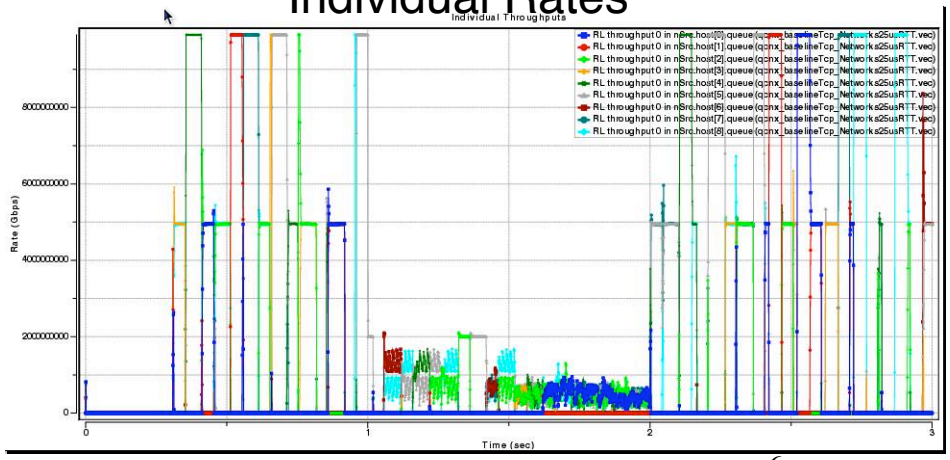
Net Throughput



Queue Length

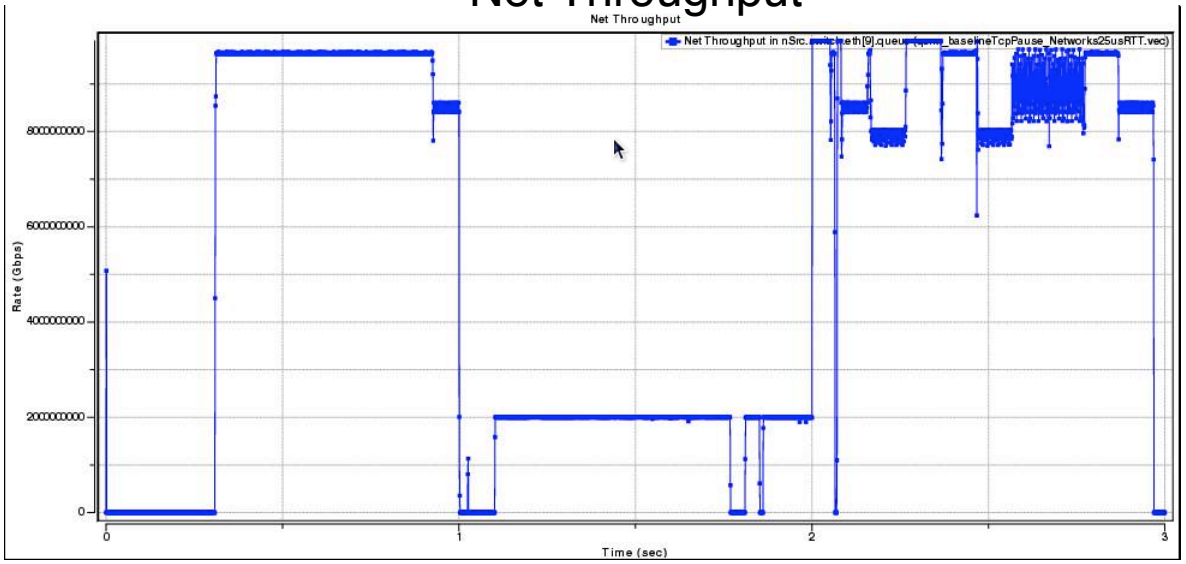


Individual Rates

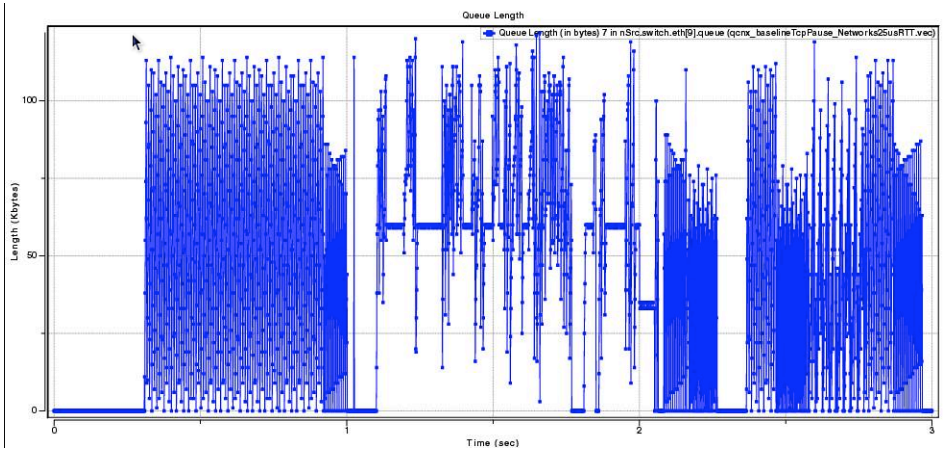


Pause Only, RTT = 25us

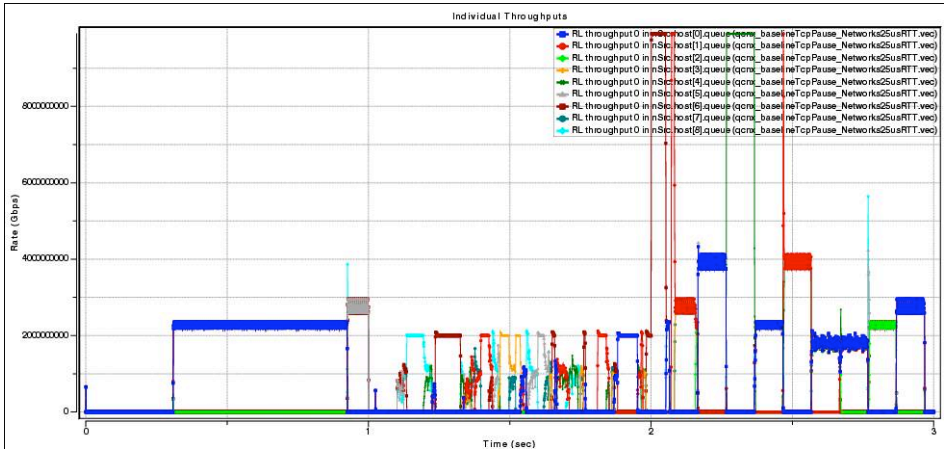
Net Throughput



Queue Length

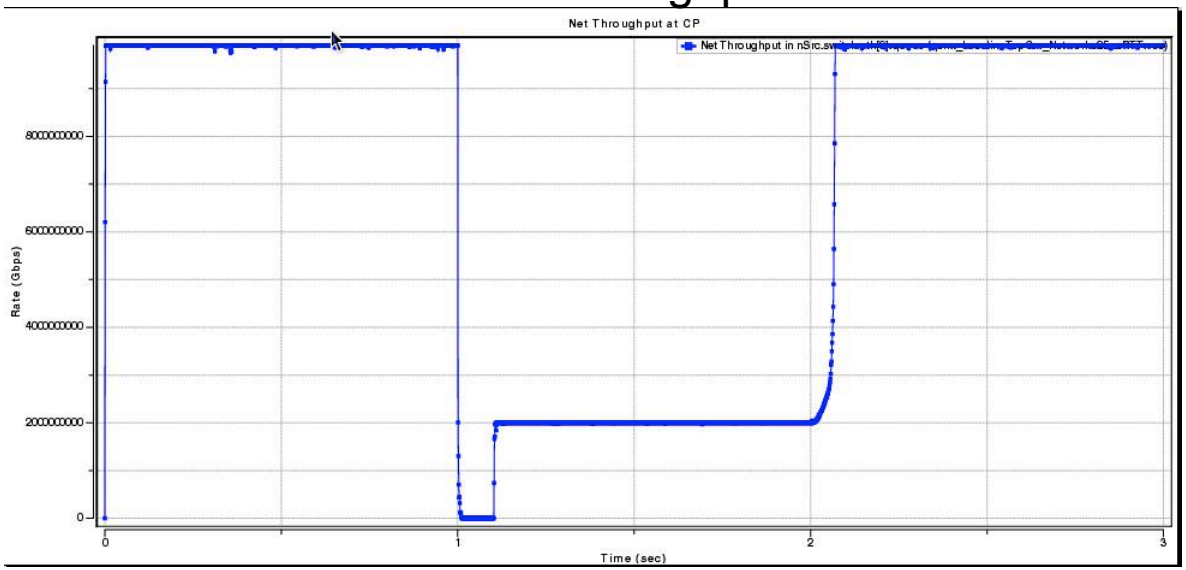


Individual Rates

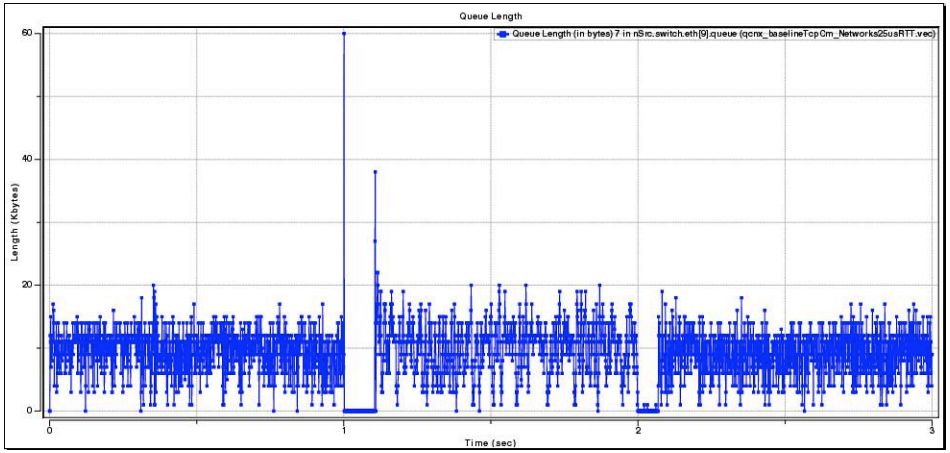


QCN Only, RTT = 25us

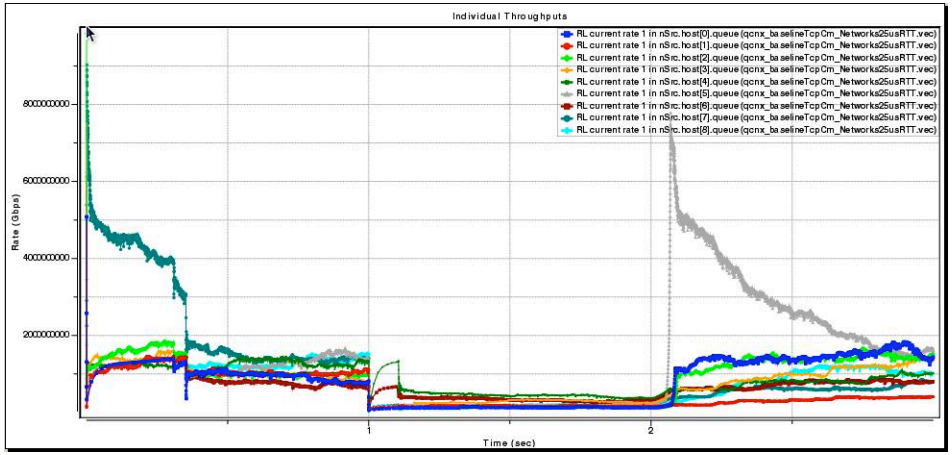
Net Throughput



Queue Length

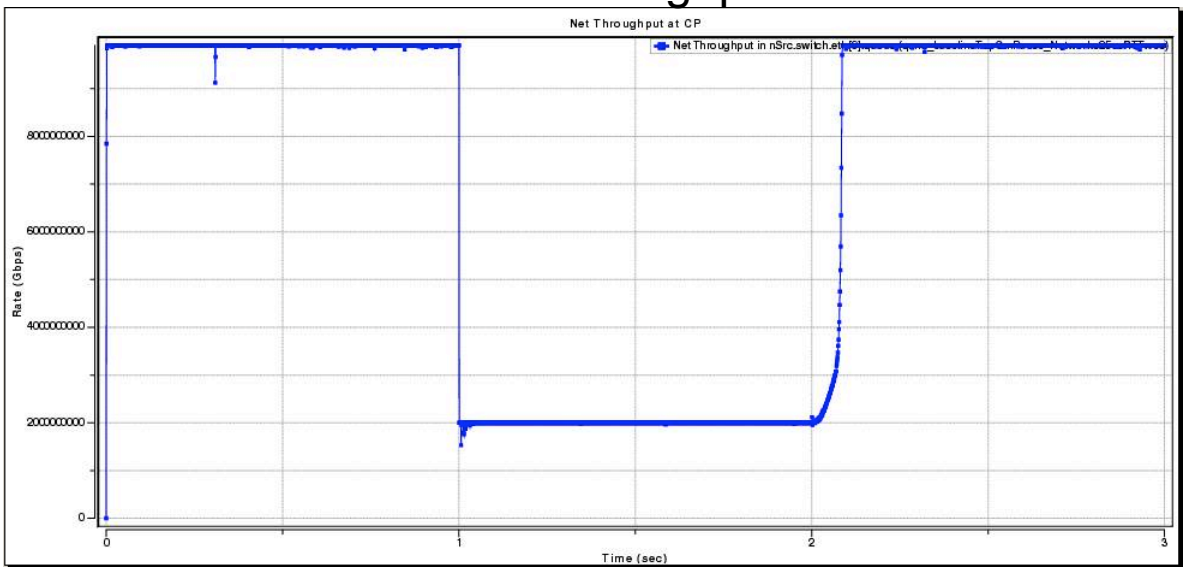


Individual Rates

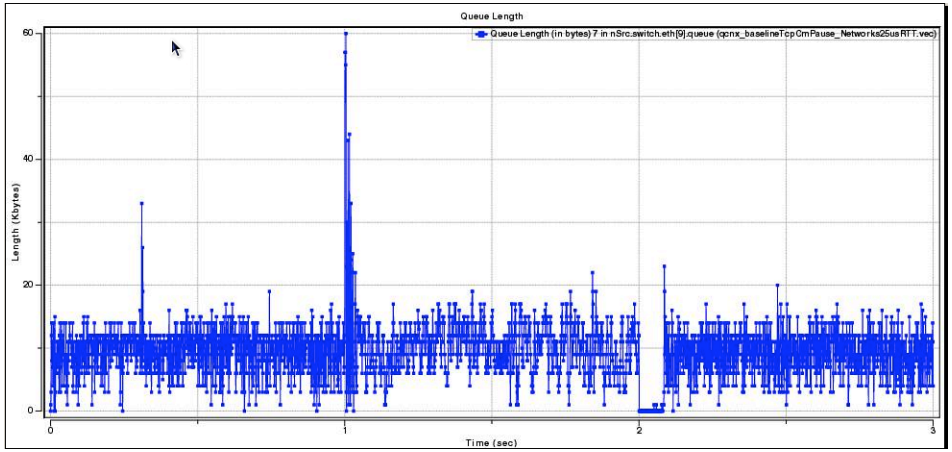


QCN and Pause, RTT = 25us

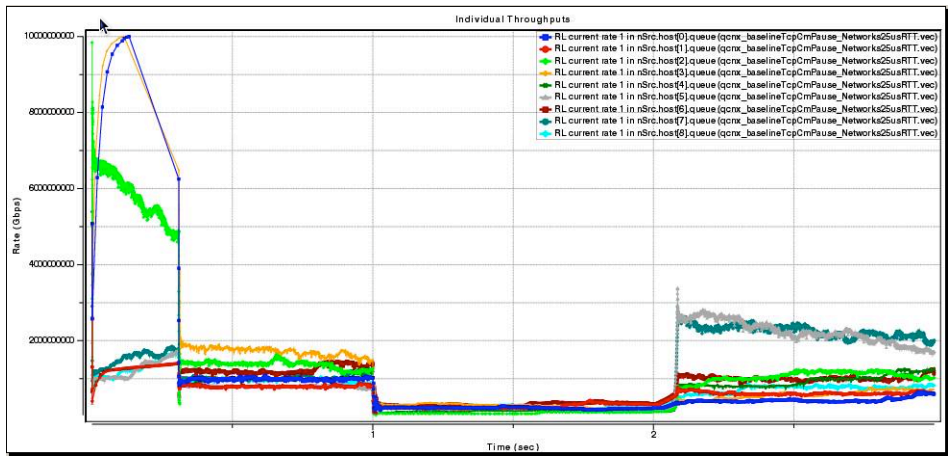
Net Throughput



Queue Length

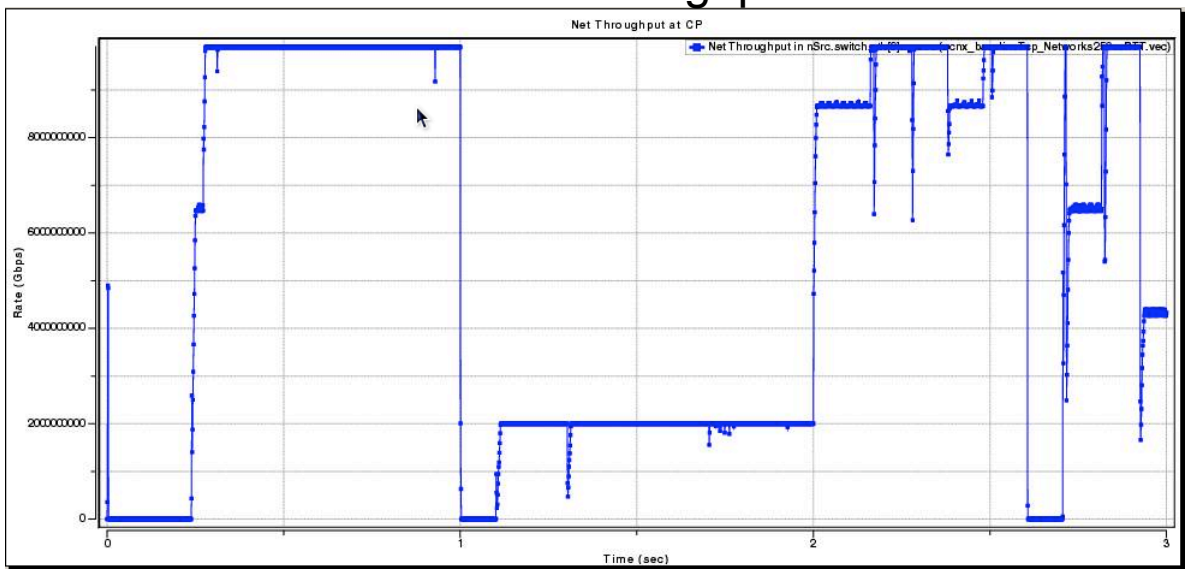


Individual Rates

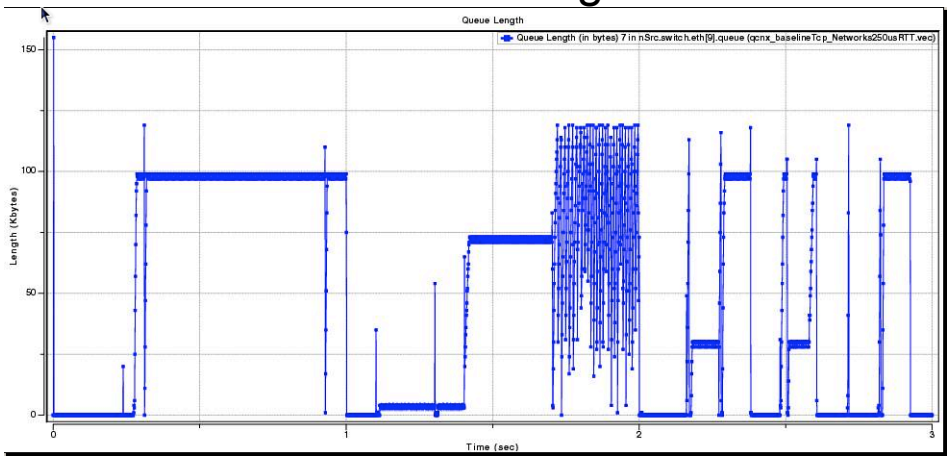


No Pause, no QCN; RTT = 250us
(Packets dropped due to buffer overflows)

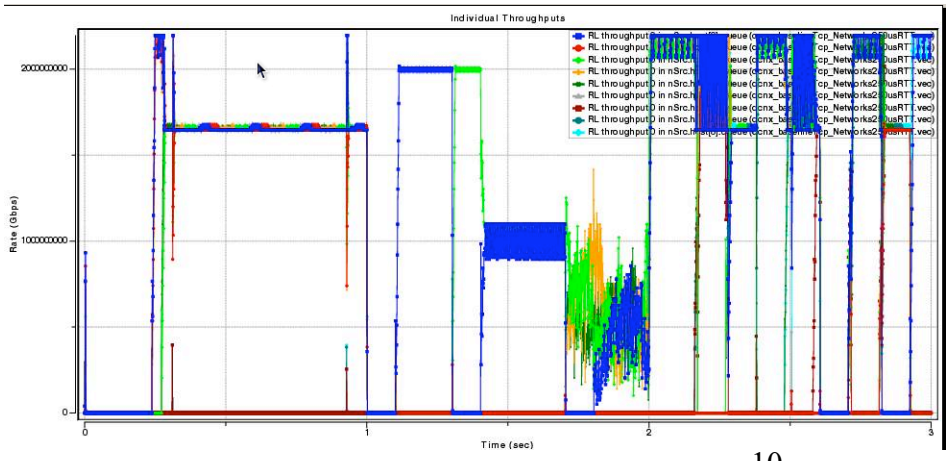
Net Throughput



Queue Length

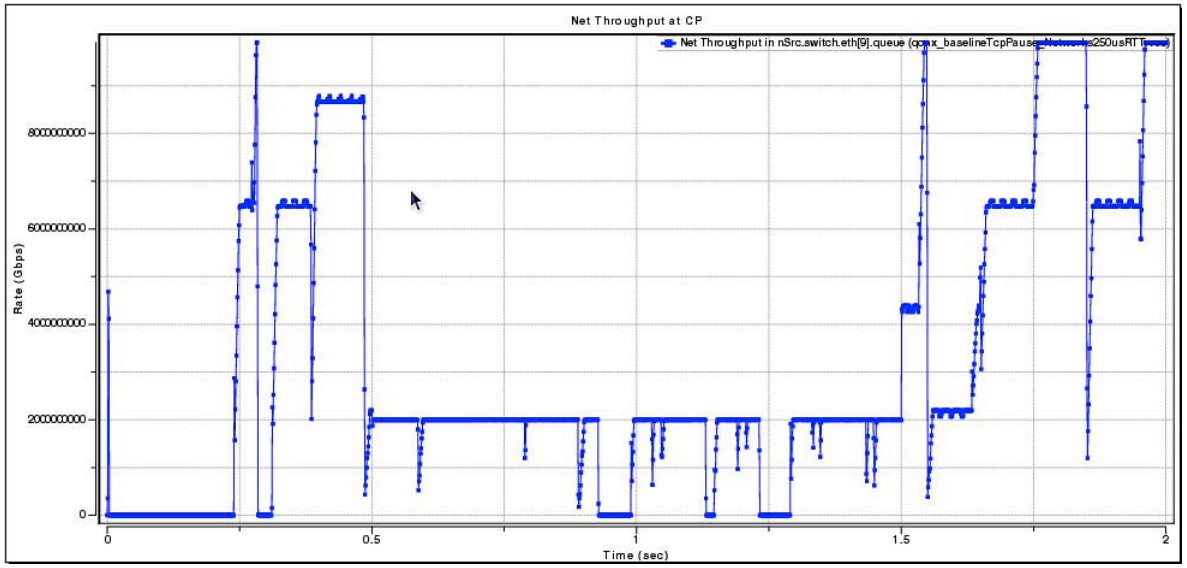


Individual Rates

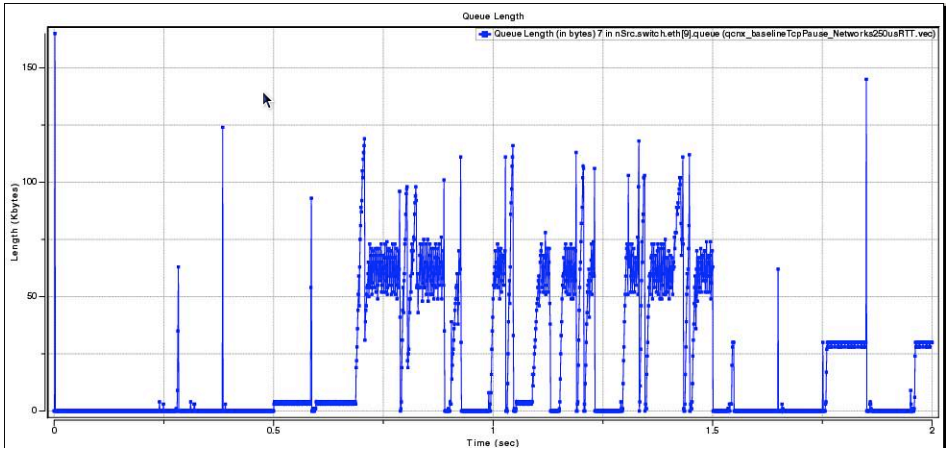


Pause Only, RTT = 250us

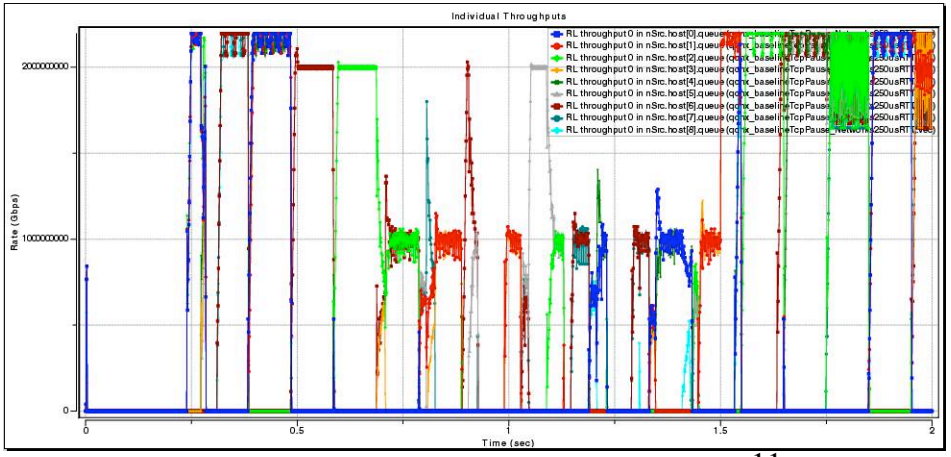
Net Throughput



Queue Length

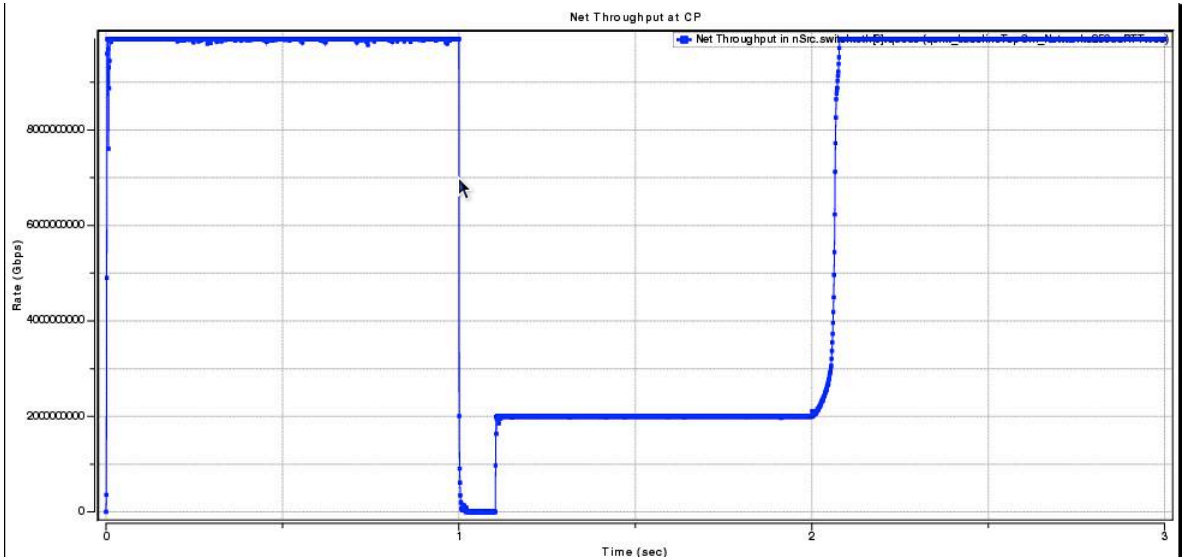


Individual Rates

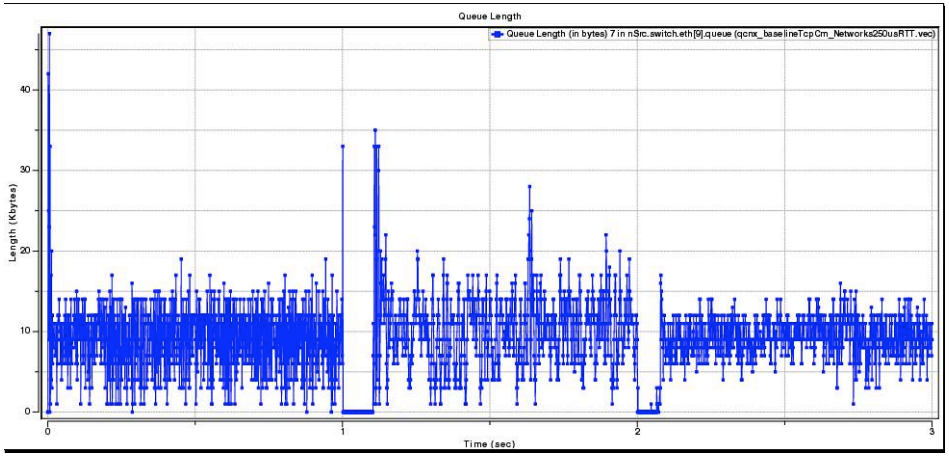


QCN Only, RTT = 250us

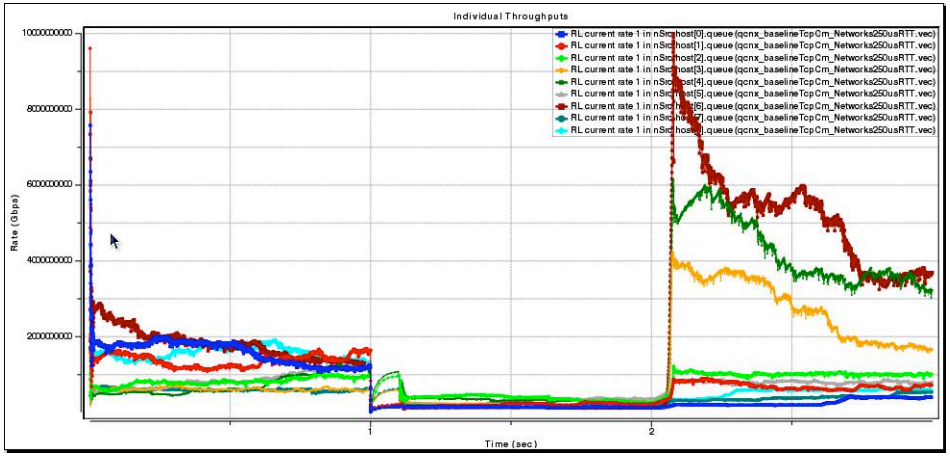
Net Throughput



Queue Length

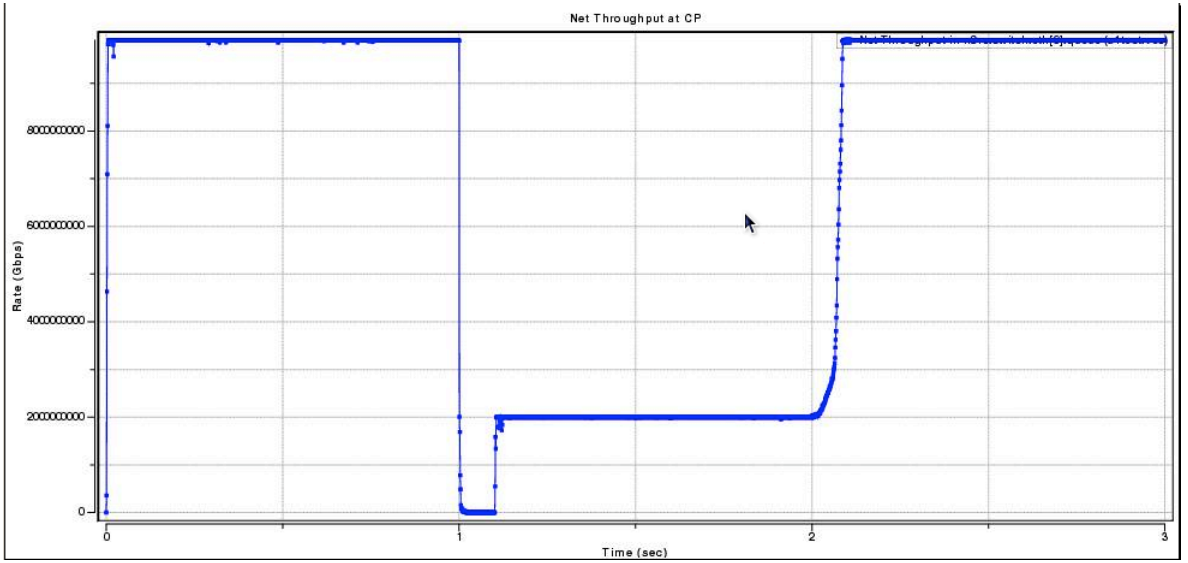


Individual Rates

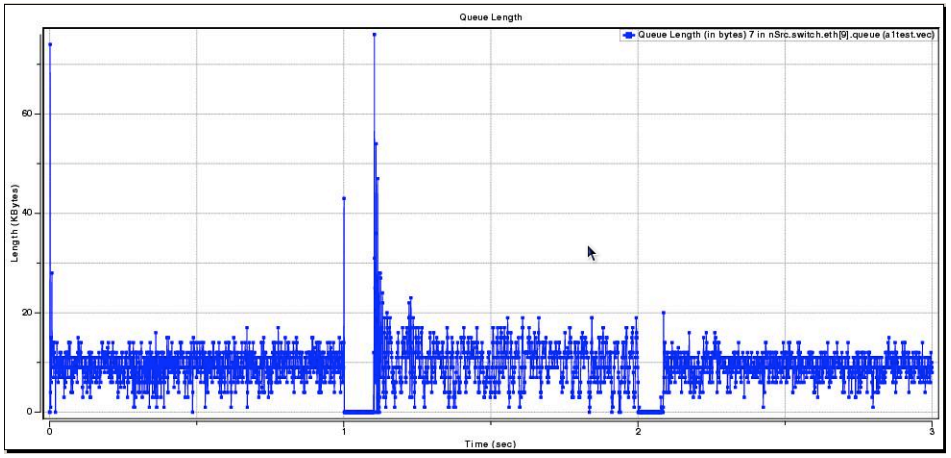


QCN and Pause, RTT = 250us

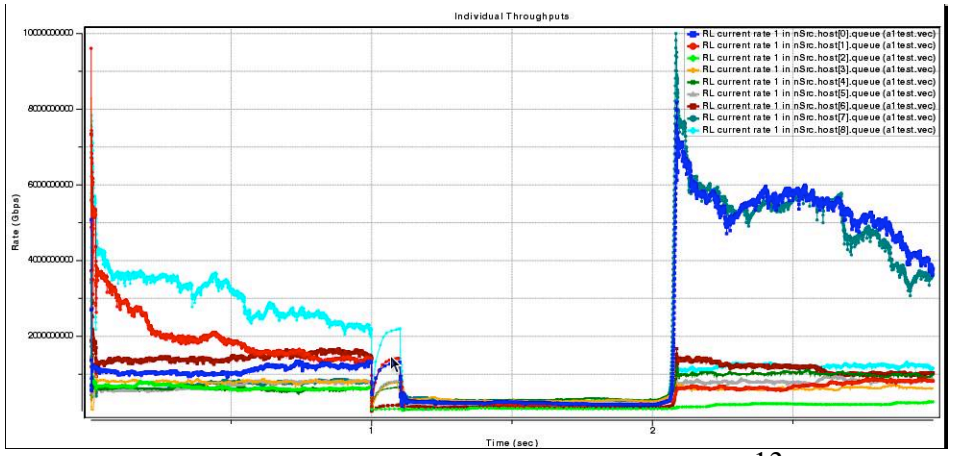
Net Throughput



Queue Length



Individual Rates



Conclusion

- It is important to have some Layer-2 CM (in this case QCN) to get good, reliable TCP performance
- Delay (RTT) makes little difference
  - Downward transience longer when RTT is longer
- Fairness looks good
  - Maximum--minimum source rate ratio is tight
- Next
  - More extensive scenarios
  - Use BIC TCP at the hosts
  - TCP flow completion times