TCP Scenarios with QCN / ECM

Eric Geisler

Manoj Wadekar

13 November 2007

Goals

- Study effects of CM on TCP in a congested environment.
- Ensure that L2 CM doesn't negatively interfere with TCP
- Compare CM proposals for a TCP workload

Methodology

- Topologies
  - Baseline
  - Single-hop output generated hotspot
- Scenarios
  - No CM
  - ECM
  - ECM & Pause
  - QCN
  - QCN & Pause
- Metrics
  - Aggregate Application throughput
  - Hotspot throughput
  - Queue length
  - Packet drops

Simulation Parameters

QCN

W = 2.0

Qeq = 409

minDecFactor = 0.5

Drift Factor = 1.0005, period = 200µs

Gd = 0.0078125 (1/128)

Rmin = 1000000

FR threshold = 5

EFR Max = 1000000

A = 12000000

TO Threshold = 150000

Sampling: Fb probability = 1% to 10%

Quantized Fb Nbits = 6

Fb Hat enabled, EFR enabled, HAI enabled, Drift timer enabled

ECM

W = 2.0

Qeq = 375

Gi = 0.53333

Gd = 0.0078125

Rmin = 1000000

Ru = 1000000

Rd = 1000000

Td = 1ms

Sampling Interval = 75000

Over-sampling disabled, ECM(0,0) disabled, ECM-Max disabled

TCP

Windows XP default settings

Reno Fast Recovery

To simulate within a short period (0.5 sec):

- Init RTO = 0.6 sec, Min RTO = 0.1 sec
- Max ACK delay = 0.001 sec

Switch

Link propagation delay: 500ns

TX buffers:

- 150KB max size per port
- 130KB pause high watermark
- 110KB pause low watermark

Note: parameters may be varied per simulation and will be noted on that slide.

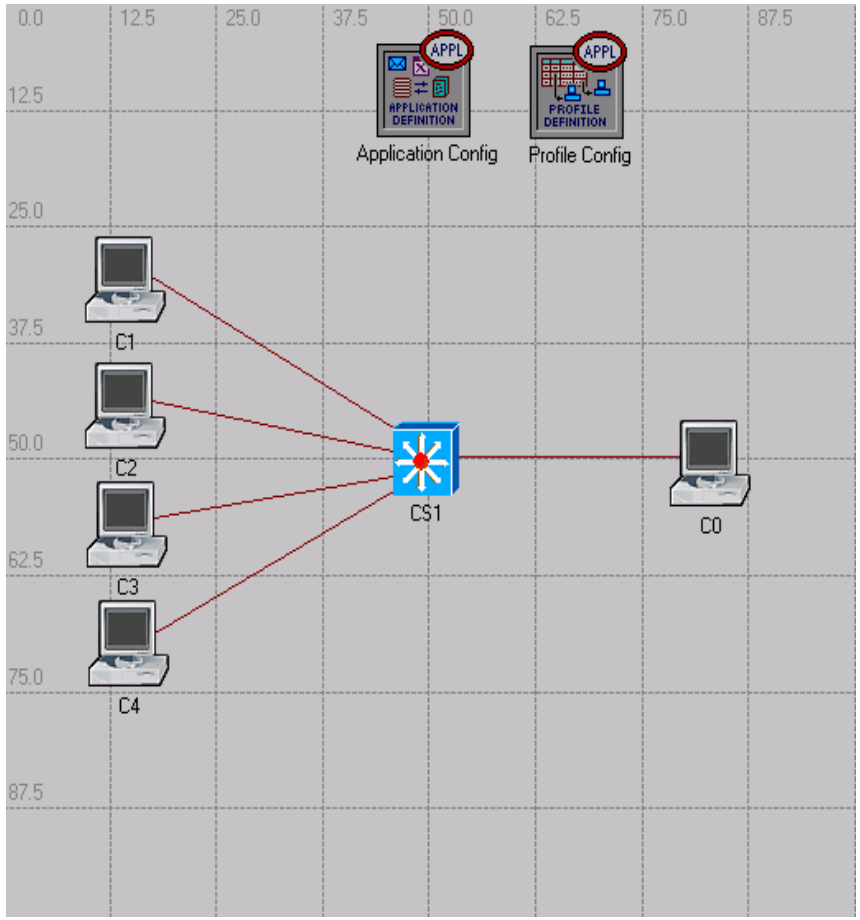
Topology & Workload

Baseline with output hotspot

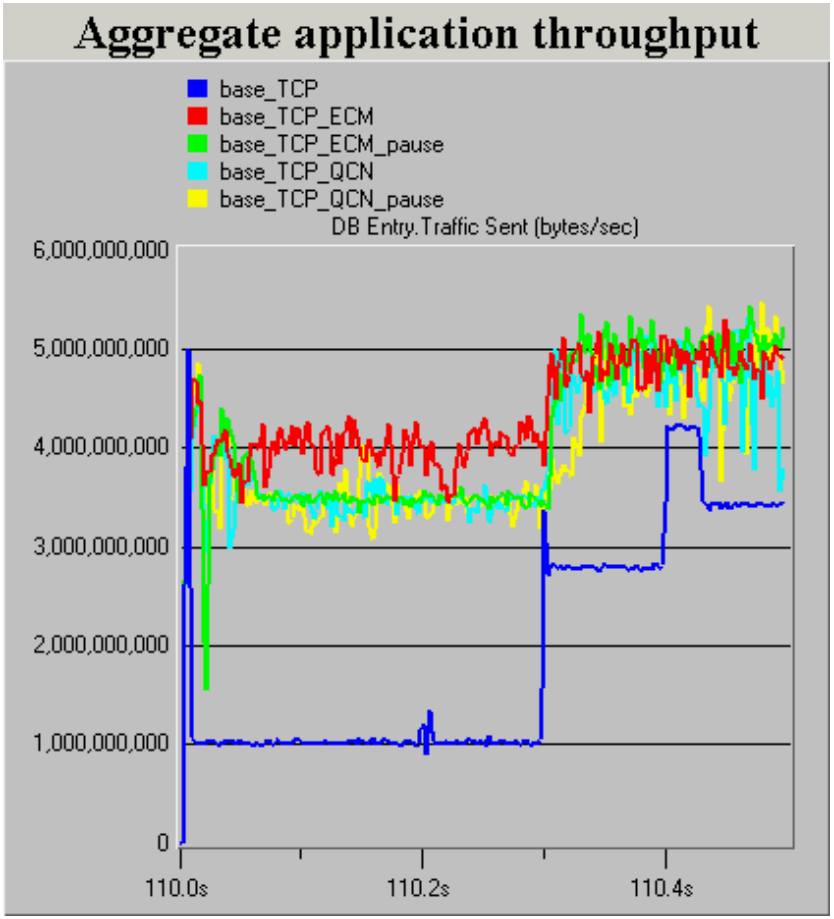
10 Gb/s links, 2 Gb/s hotspot @ CS1→C0

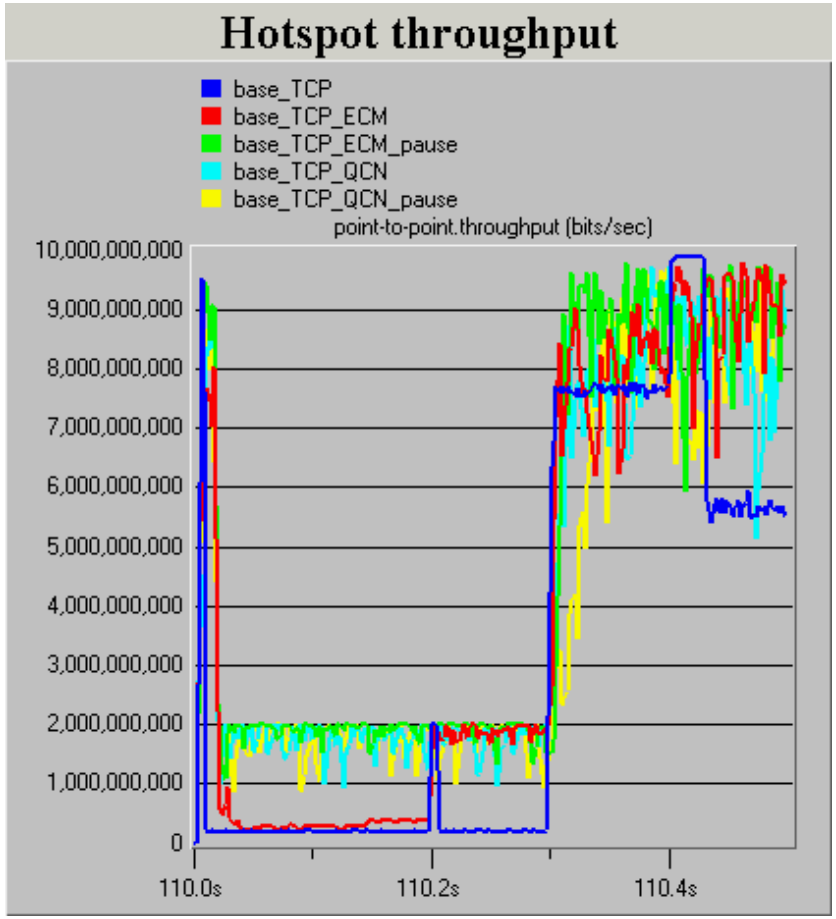
Traffic pattern:

- 5 hosts C0-C4 @ 85% loading
- Hotspot duration: ~300ms
- Spatially uniform (except self)
- Temporally Bernoulli



Results
Throughput

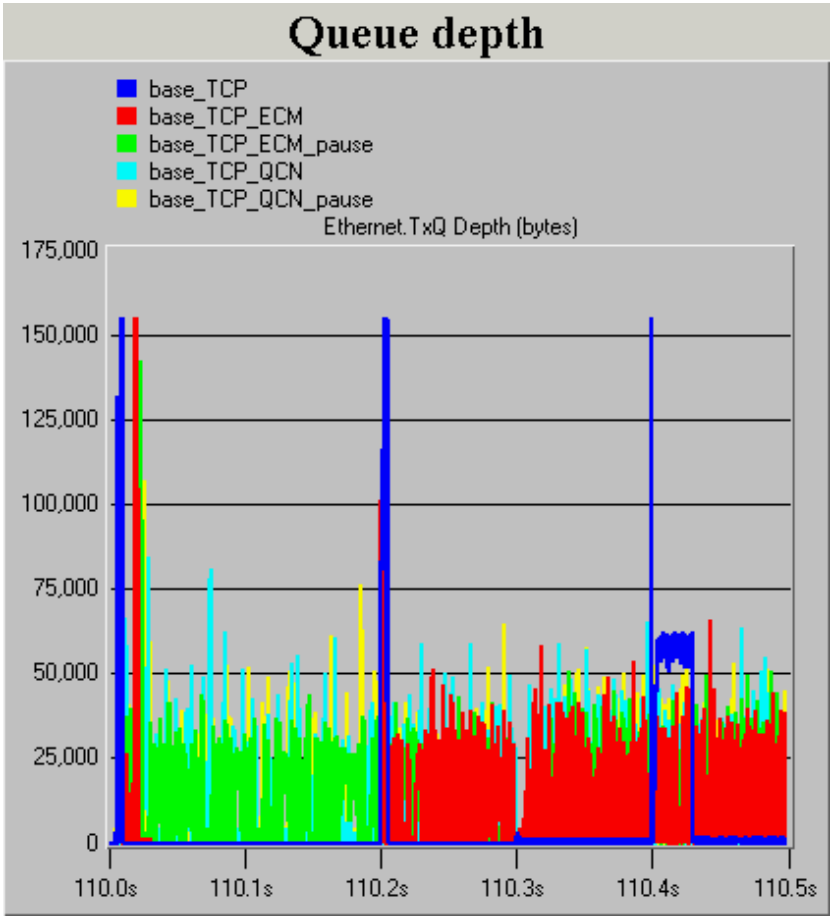


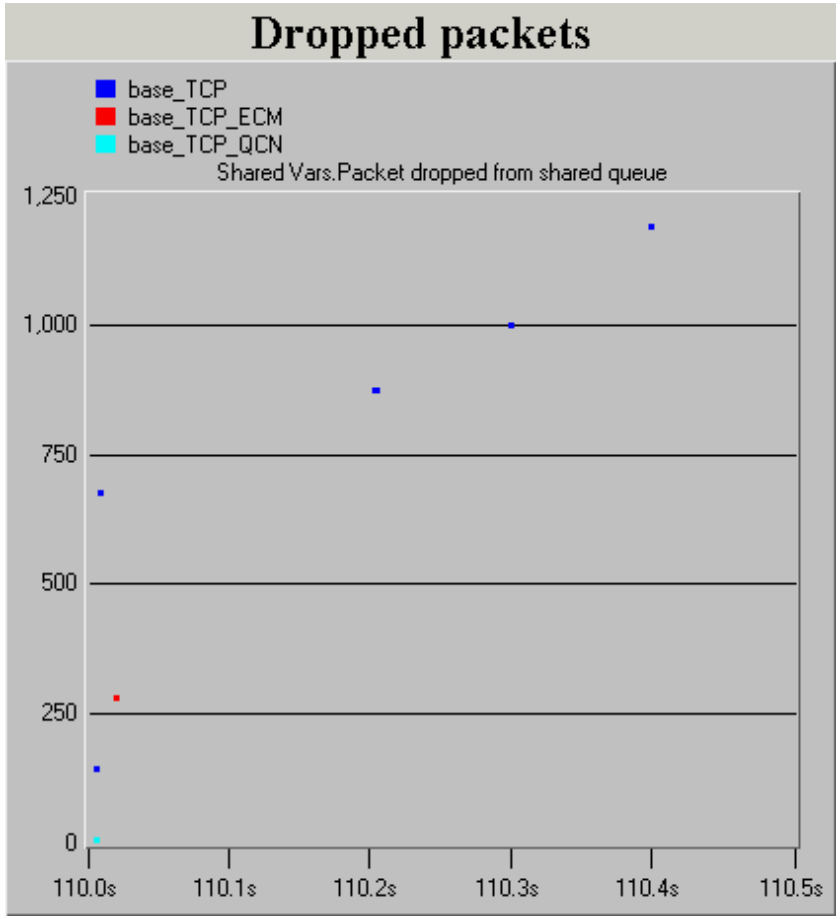


Lost connections can help few flows to achieve better throughput. E.g. base_TCM_ECM in chart above

Results

Queue depth, packet drop





Throughput lost due to timed-out TCP connections for ECM, NO_CM.

No dropped packets with PAUSE.

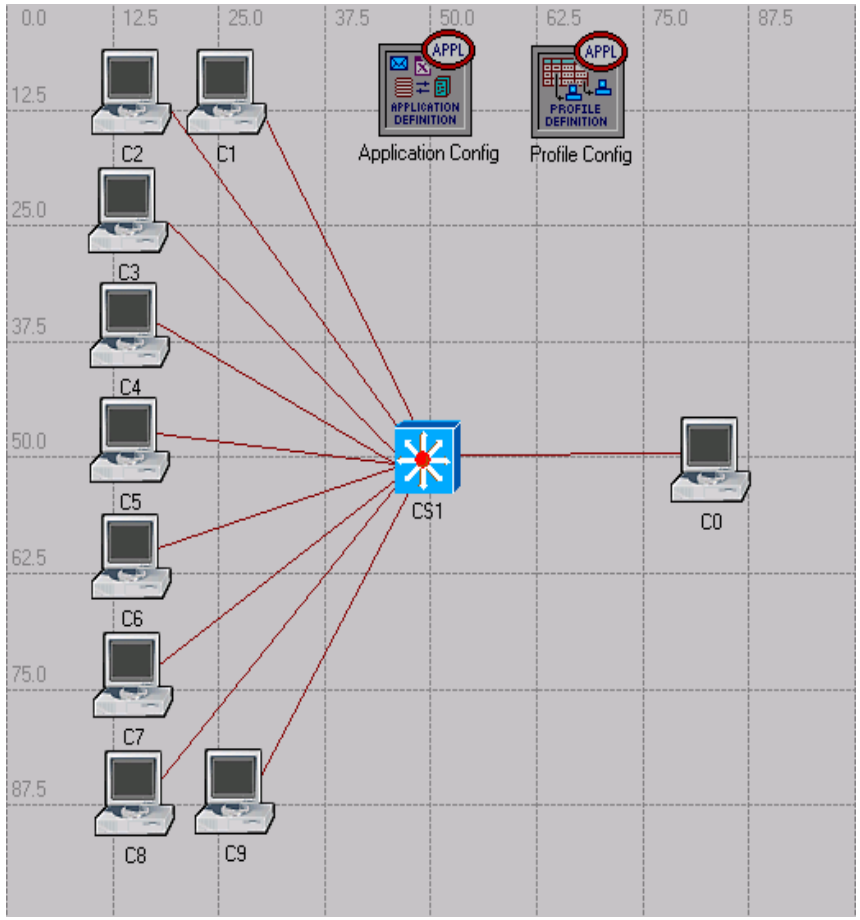
Topology & Workload

Single-Hop with output hotspot

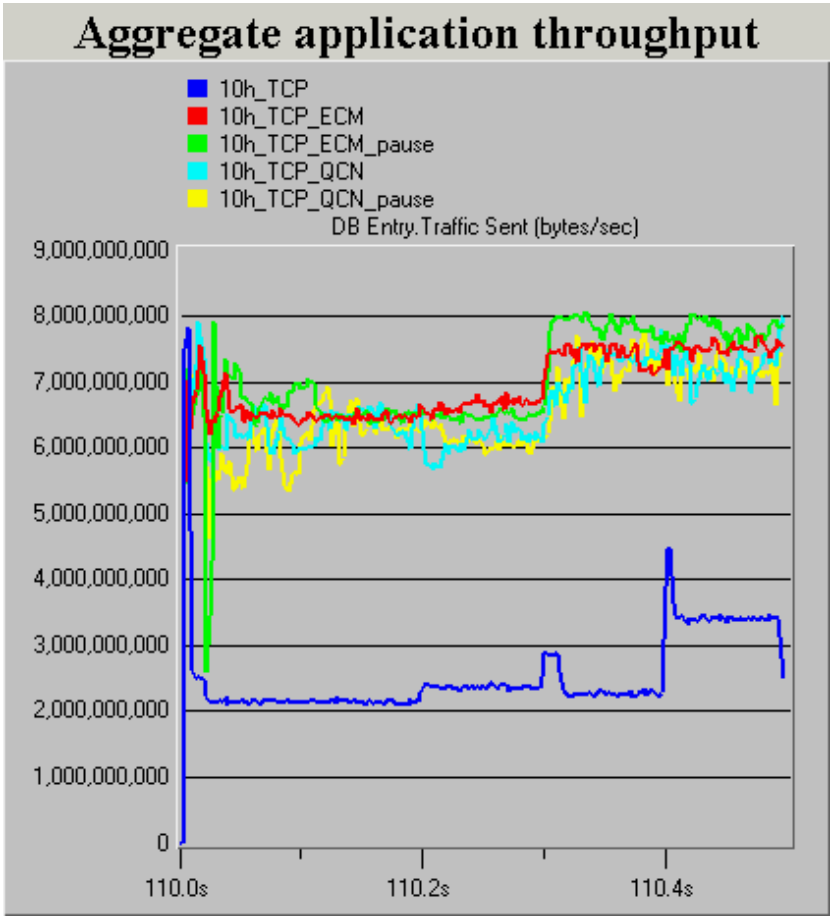
10 Gb/s links, 2 Gb/s hotspot @ CS1→C0

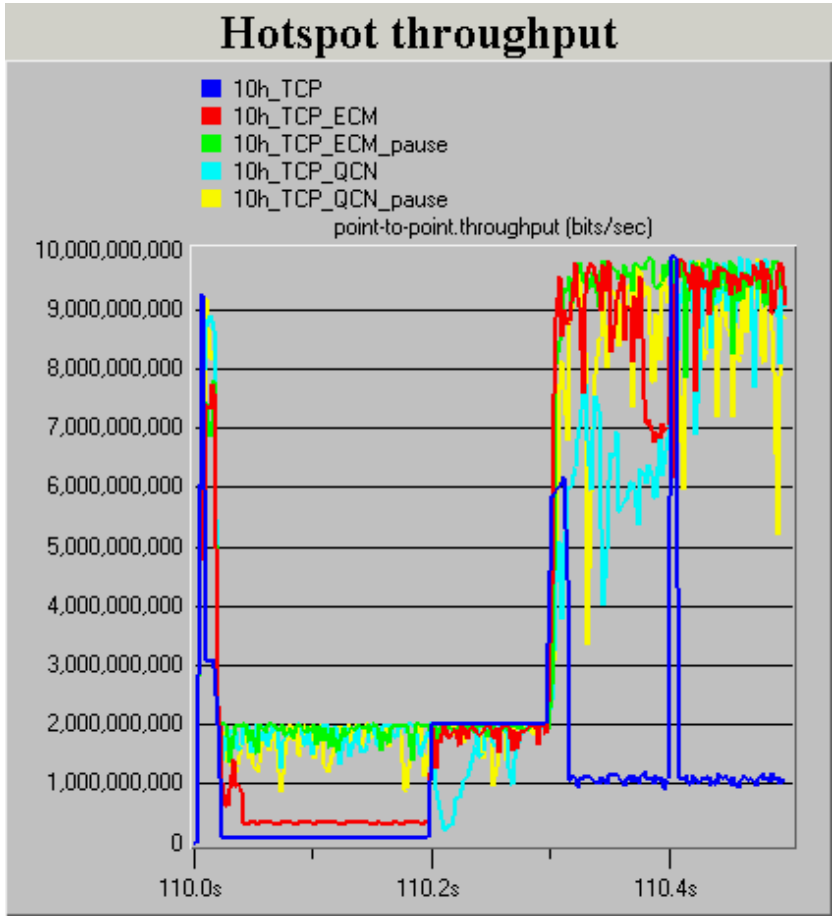
Traffic pattern:

- 10 hosts C0-C9 @ 85% loading
- Hotspot duration: ~300ms
- Spatially uniform (except self)
- Temporally Bernoulli



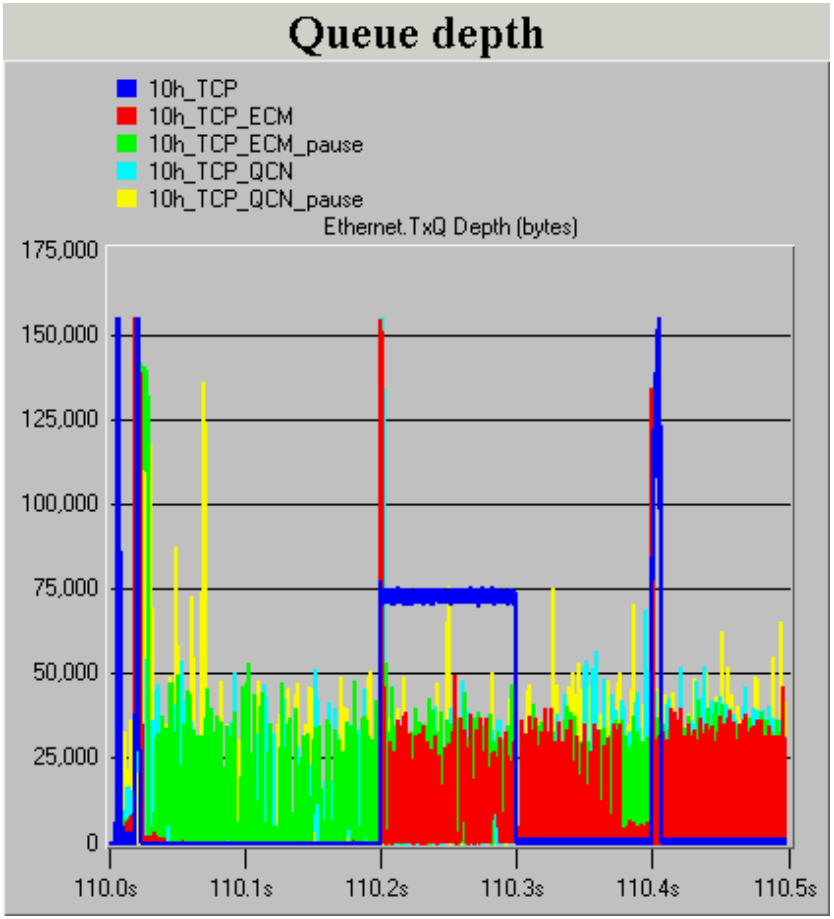
Results
Throughput

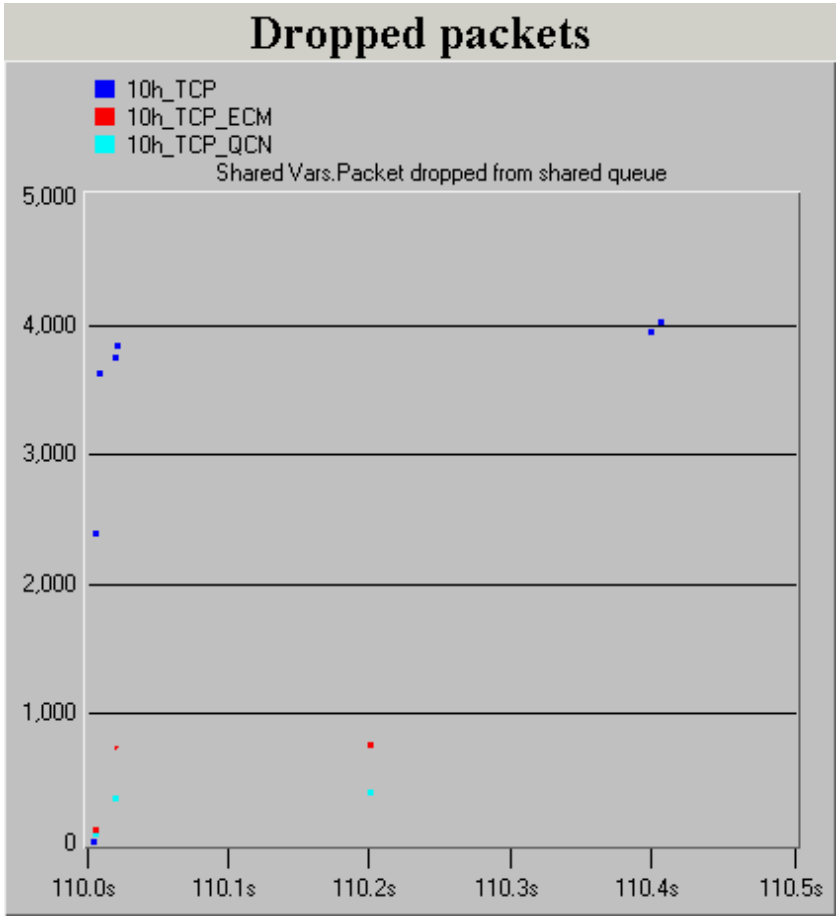




Results

Queue depth, packet drop





Summary

- CM protocols don't harm TCP
- Without CM: In DC environment – TCP burst causes packet drops and retransmission
  - CM improves throughput for all flows
  - Other alternative is to tune TCP for flows within DC
- ECM and QCN comparison in line with L2 workloads
  - ECM did not have over-sampling, ECM(MAX), ECM(0, 0) enabled
- No TCP congestion avoidance when CM avoids packet drops

Next Steps

- Long link latency effects
- FCT & dynamic flows
  - Pareto distribution of small and large flows