# Comparison of ECM and QCN based on CN-SIM Benchmarks

**Davide Bergamasco**

**IEEE 802 Plenary Meeting**
**San Francisco, CA**
**July 17th, 2007**

# Objective

- Evaluate performance of ECM and QCN (both 2pt and 3pt) in some of the benchmark scenarios defined in "*Topologies & Workloads*" [1]
  - Baseline
  - Single-Hop Output-Generated Hotspot
  - Multi-Hop Output-Generated Hotspot
  - Fat-tree
- Metrics
  - Tier 1 Performance metrics from "*Discussion About Metrics*" [2]
    - Aggregate throughput
    - Packets dropped
    - % time paused
    - Signaling overhead
    - Queue length
    - Bottleneck link utilization

[1] http://www.ieee802.org/1/files/public/docs2007/au-sim-wadekar-reqd-extended-sim-list020807.pdf

[2] http://www.ieee802.org/1/files/public/docs2007/au-sim-bergamasco-on-metrics-070314.pdf

# Simulation Environment

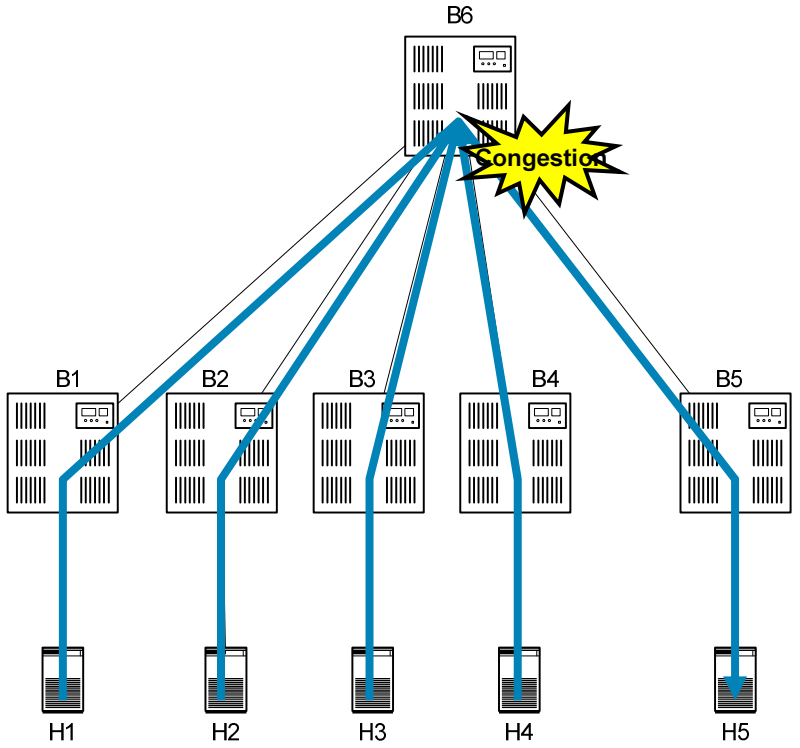
- Bridge architecture as per "*CN-SIM: A common Bridge Model*" [1]
- Bridge output buffer partitioned per input port
  - 150 KB of space for each input ➔ 2.4 MB for 16 ports
  - Pause Thresholds
    - High watermark = 130 KB
    - Low watermark = 110 KB
- NIC buffer 150 KB
- Link propagation delay = 0.5 µs
- Bridge/NIC processing delay = 1 µs
- ECM parameters
  - W = 2
  - Gi = 5.333333333 x $10^{-1}$
  - Gd = 2.666666667 x $10^{-4}$
  - Qeq = 375
  - Qmc = 1300 (ECM-Max)
  - FixedSamplingInt = 75000 B
  - RandomSamplingInt = uniform(-5000, 5000) B
  - Oversampling Enabled
  - ECM(0,0) disabled
  - Drift disabled
- QCN parameters
  - W = 2
  - Ri = 12 Mbps
  - Gd = 0.0078125
  - Qeq = 409
  - Nbits = 6
  - Pmin = 0.01
  - Pmax = 0.1
  - FRThr = 5
  - FRByteCnt = 150000
  - AIByteCnt = 37500
  - RefP = 0.04 (3pt only)
  - EFR enabled
  - HyperActive Increase disabled
  - Drift disabled



[1] http://www.ieee802.org/1/files/public/docs2006/au-sim-bergamasco-common-bridge-model-101206v2.pdf

# Baseline

- Topology & Workload



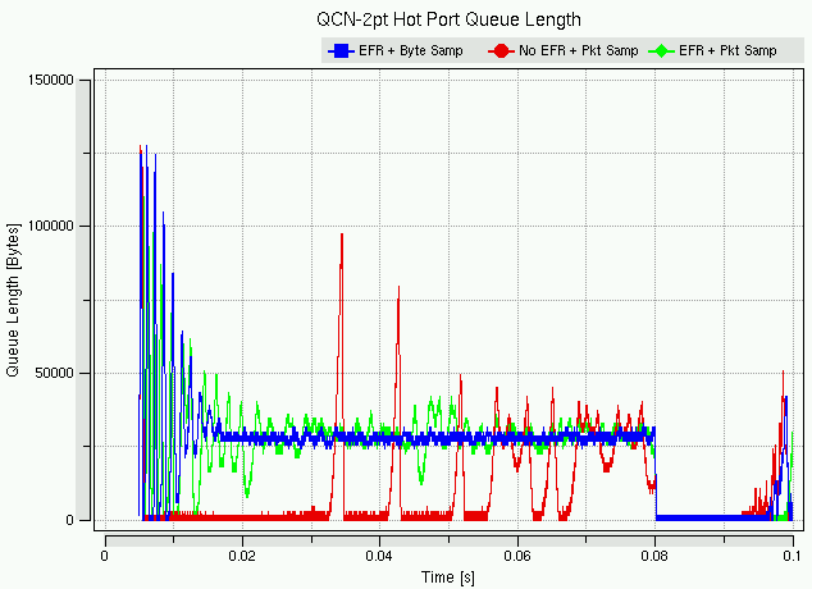

- Traffic pattern

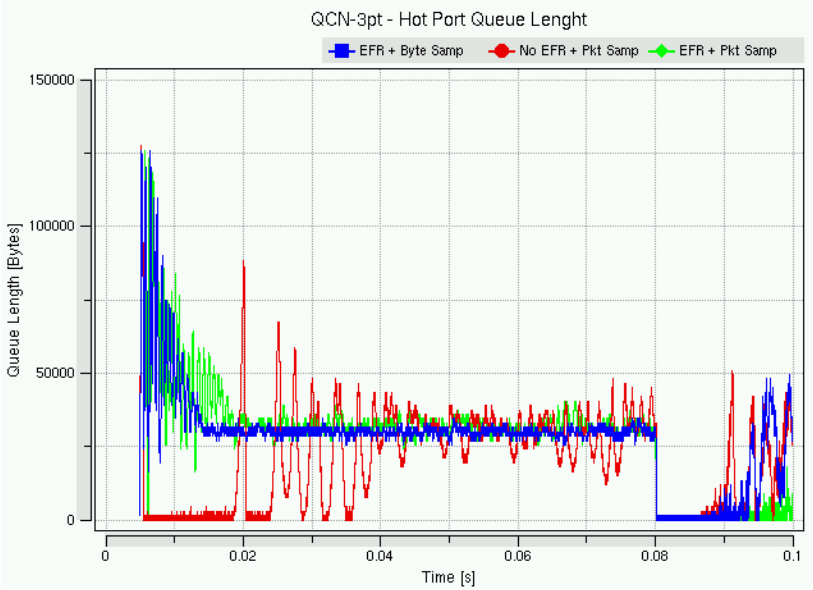
  Load: H1-H4 = 50% H5 = 0%

  Spatially: H1-H4 → H5

  Temporally: Bernoulli

# Baseline


QCN-2pt Hot Port Queue Length
EFR + Byte Samp
No EFR + Pkt Samp
EFR + Pkt Samp
Queue Length [Bytes]
150000
100000
50000
0
0
0.02
0.04
0.06
0.08
0.1
Time [s]


QCN-3pt - Hot Port Queue Lenght
EFR + Byte Samp
No EFR + Pkt Samp
EFR + Pkt Samp
Queue Length [Bytes]
150000
100000
50000
0
0
0.02
0.04
0.06
0.08
0.1
Time [s]

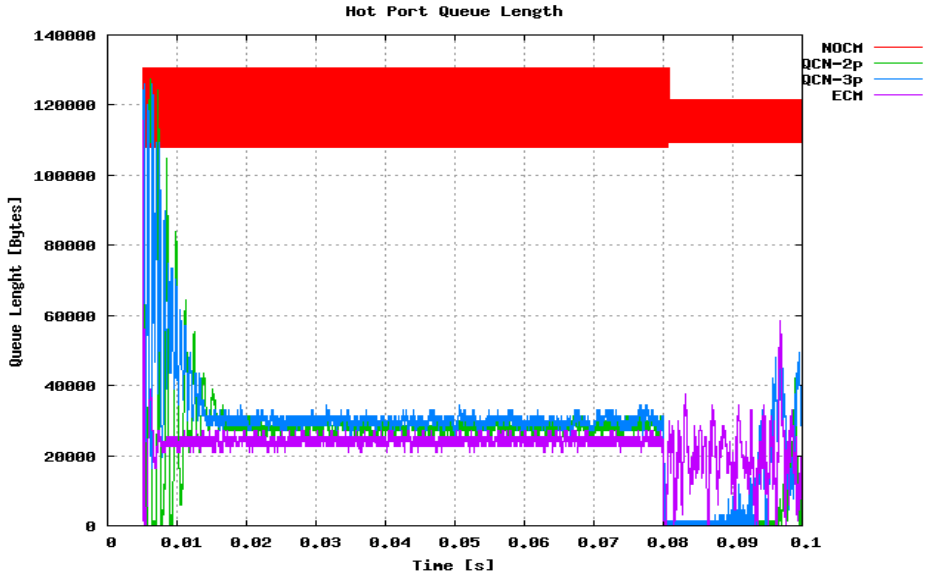

QCN-2pt - Hot Port Throughput
EFR + Byte Samp
No EFR + Pkt Samp
EFR + Pkt Samp
Throughput [Gb/s]
10
8
6
4
2
0
0
0.02
0.04
0.06
0.08
0.1
Time [s]

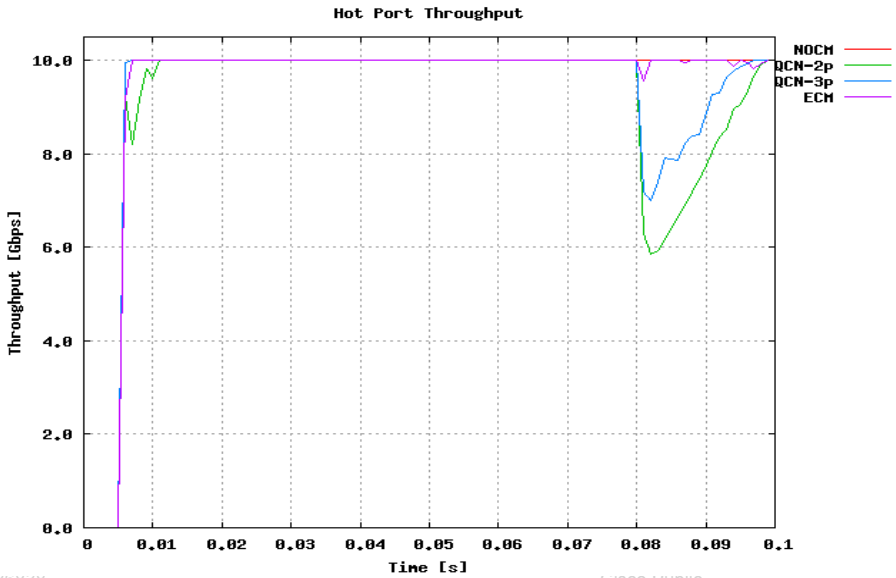

QCN-3pt - Hot Port Throughput
EFR + Byte Samp
No EFR + Pkt Samp
EFR + Pkt Samp
Throughput [Gb/s]
10
8
6
4
2
0
0
0.02
0.04
0.06
0.08
0.1
Time [s]

# Baseline

- Observations
  - Extra Fast Recovery (EFR) is essential for stability
    - Without EFR, QCN would spend too long in Active Increase (AI), where the rate grows exponentially, resulting in instability
  - Byte-based sampling is to be preferred to Packet-based sampling to reduce jitter

# Baseline



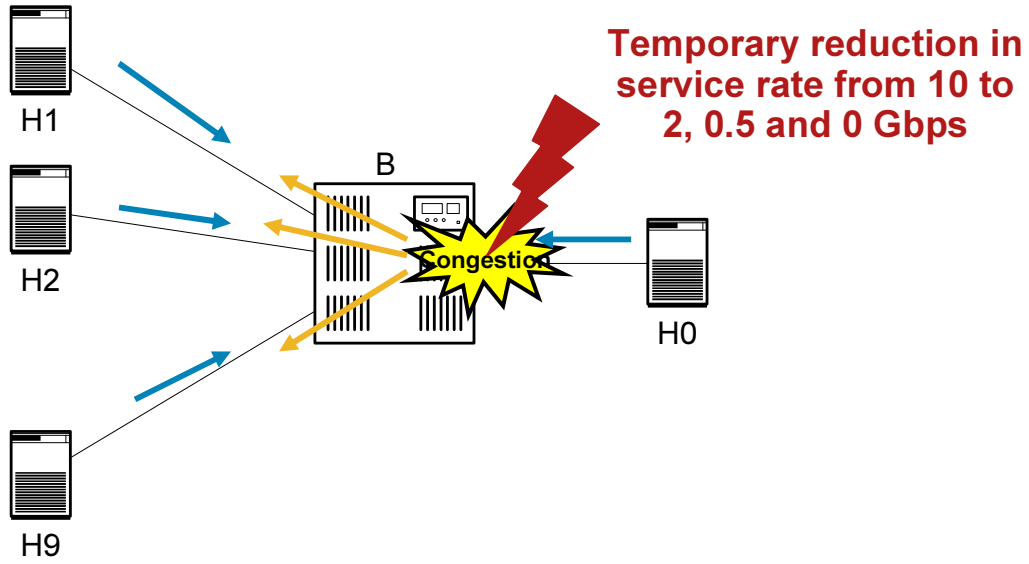




## Observations

- Both QCN-2pt and QCN-3pt seem to do their job at controlling congestion
- As expected, ECM performs somewhat better
  - QCN trades off performance for simplicity
- QCN-2pt, being the simplest algorithm, suffers more than QCN-3pt

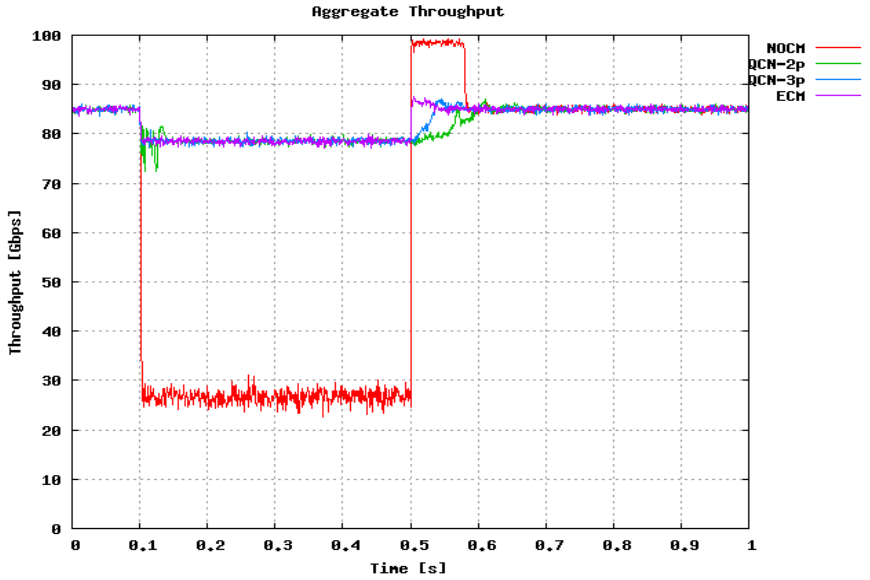
# Single-Hop Output Generated Hotspot

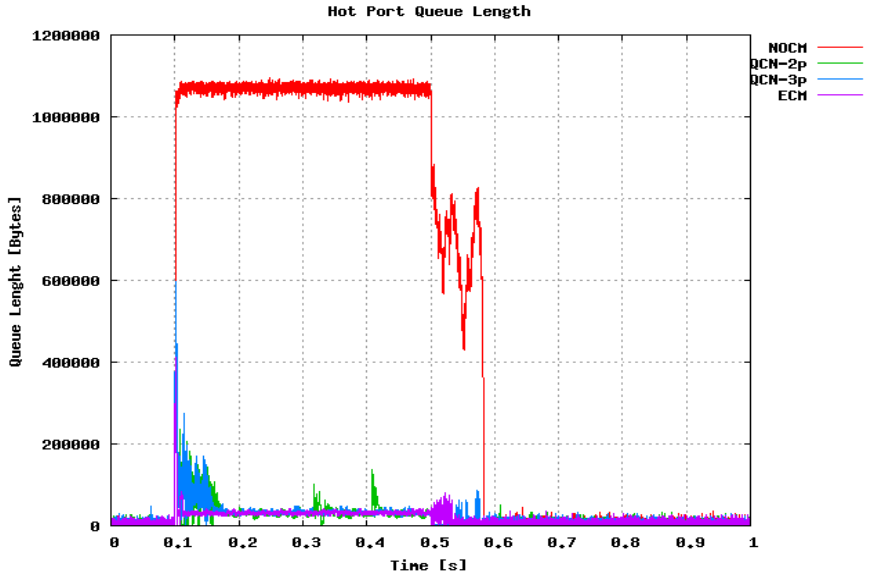
- Topology & Workload

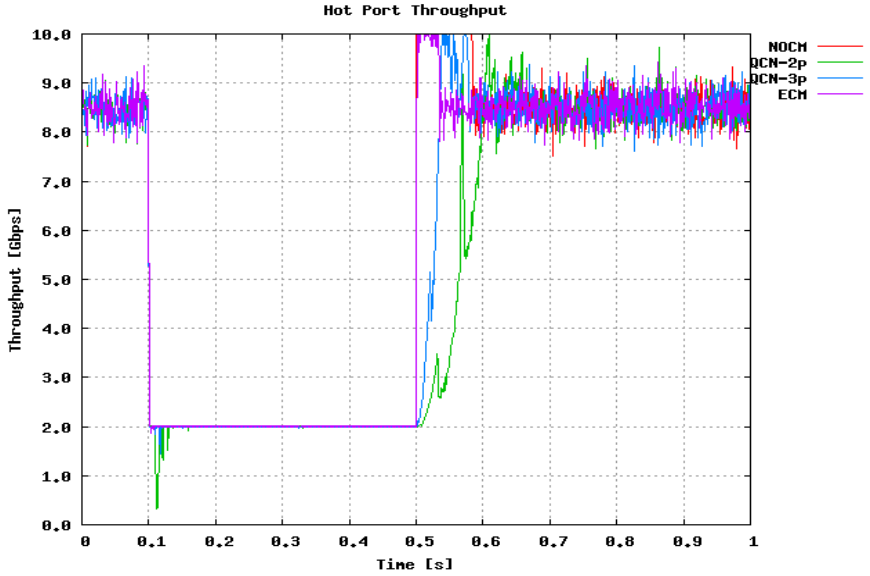



- Traffic pattern
  - Load: H0-H9 = 85%
  - Spatially: Uniform (except self)
  - Temporally: Bernoulli
- Hotspot
  - Duration: 400 ms, from $t_i$ = 100 to $t_f$ = 500 ms
  - HS degree = 9
  - HS severity = 3.25 / 15 / ∞ : 1

# 2 Gb/s Hotspot (with Pause)

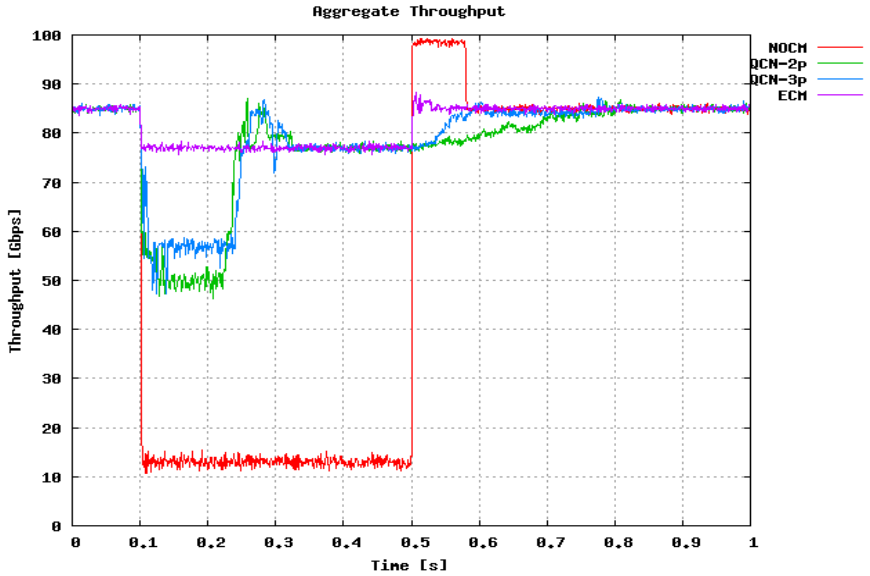


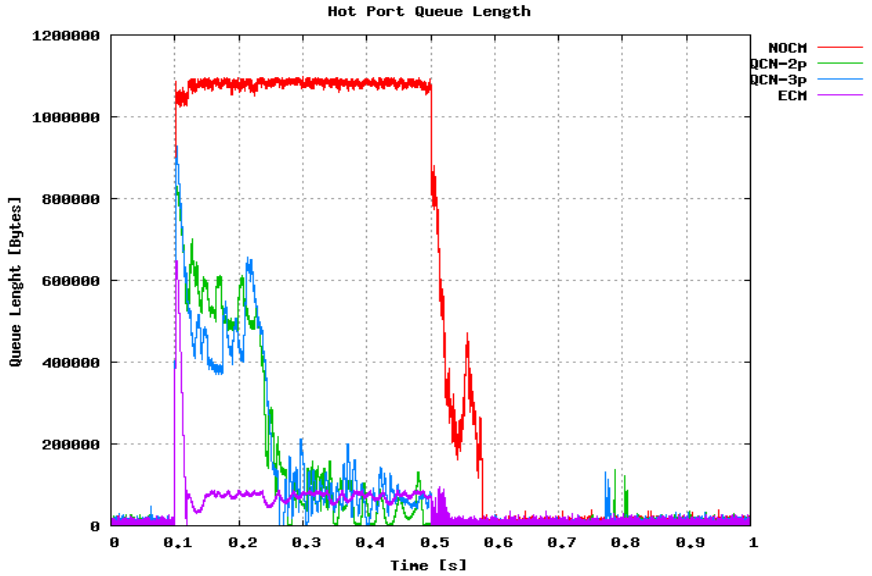



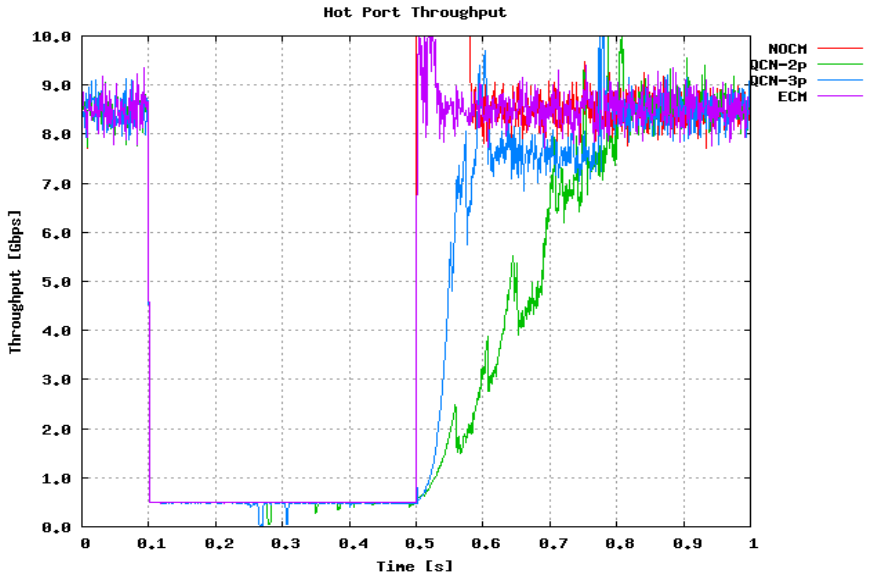




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 62.79 | 6.01 | 31.86 | 0 | 0 |
| QCN-2p | 81.99 | 5.49 | 0.15 | 0 | 0.517 |
| QCN-3p | 82.28 | 5.79 | 0.04 | 0 | 140.36 |
| ECM | 82.44 | 5.95 | 0 | 0 | 4.83 |

# 0.5 Gb/s Hotspot (with Pause)

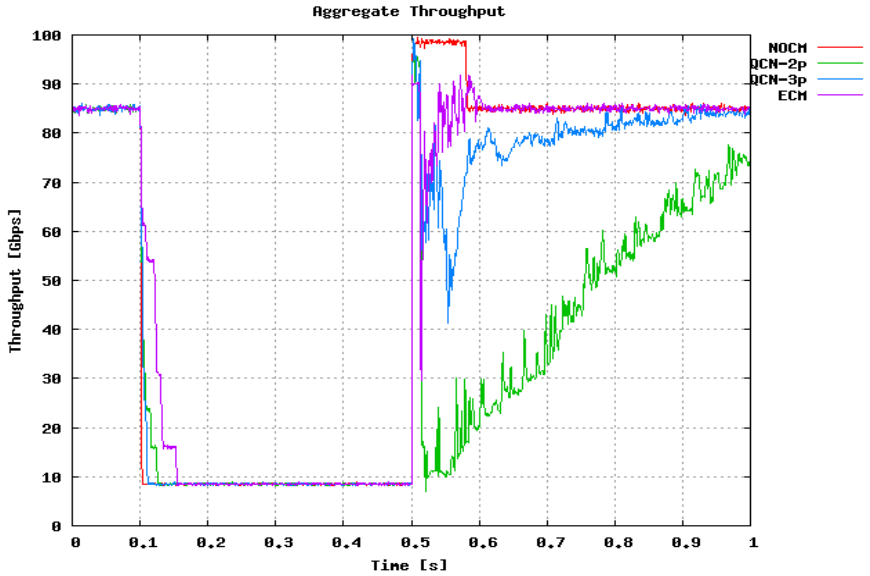


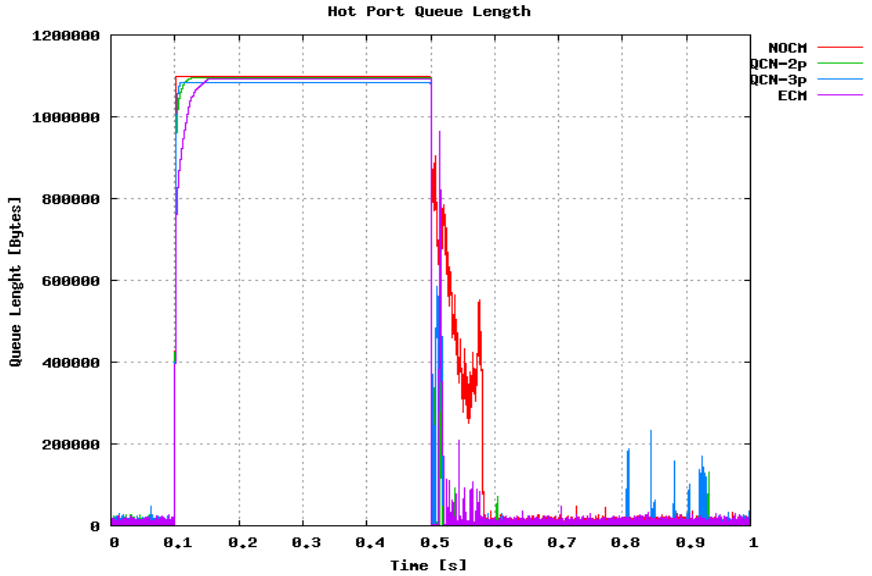



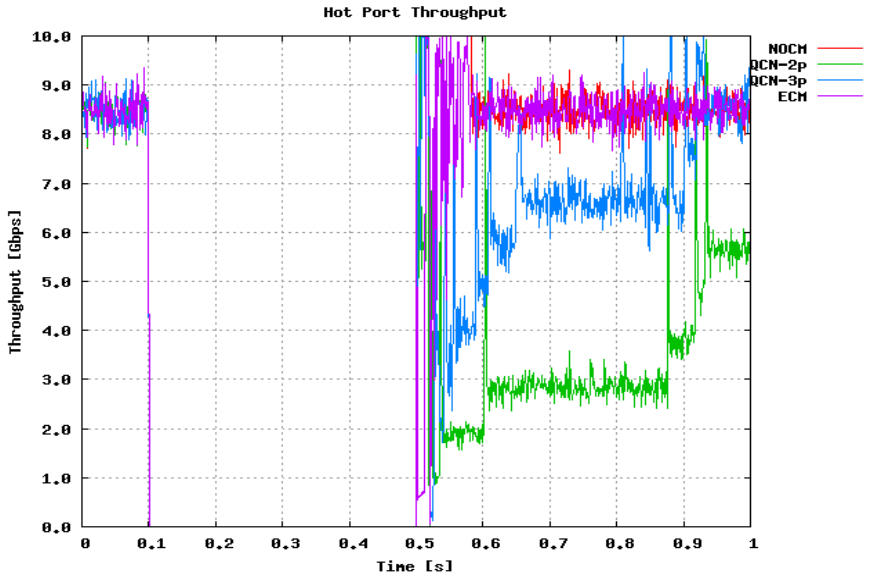




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 57.33 | 5.41 | 37.9 | 0 | 0 |
| QCN-2p | 77.40 | 4.09 | 7.07 | 0 | 0.667 |
| QCN-3p | 78.65 | 4.73 | 4.55 | 0 | 134.80 |
| ECM | 81.81 | 5.32 | 0 | 0 | 4.51 |

# 0 Gb/s Hotspot (with Pause)



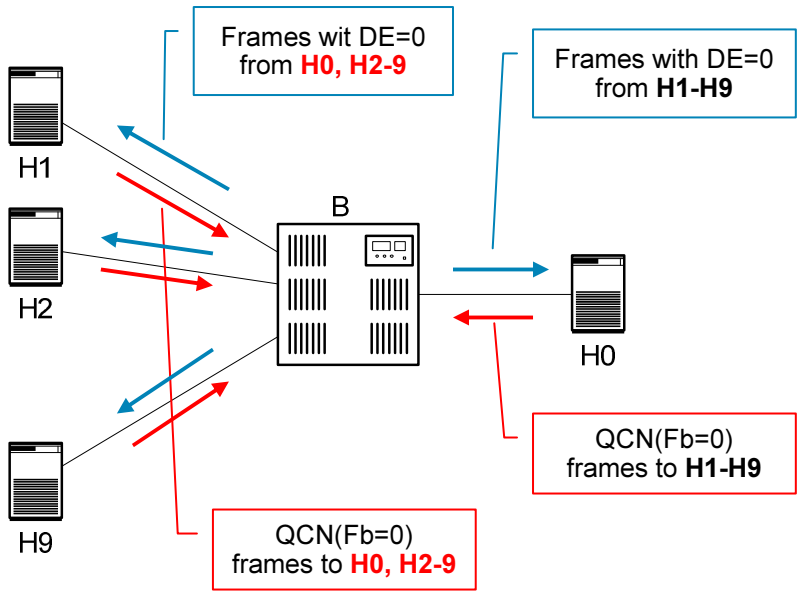




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 55.53 | 5.21 | 39.91 | 0 | 0 |
| QCN-2p | 35.55 | 2.57 | 39.40 | 0 | 0.35 |
| QCN-3p | 51.53 | 4.18 | 36.57 | 0 | 87.81 |
| ECM | 55.54 | 4.99 | 34.30 | 0 | 4.81 |

# Observations

- QCN-3pt performance is somewhere in between ECM and QCN-2pt
- However, its control traffic overhead is 1 to 2 orders of magnitude larger that the others




- Reflection Points (i.e., end nodes) generate QCN(Fb=0) messages by sampling frames with DE=0
- Unfortunately those frames are the vast majority of the traffic
- Control traffic overhead:

  QCN-3pt ~ O(N) * O(L)

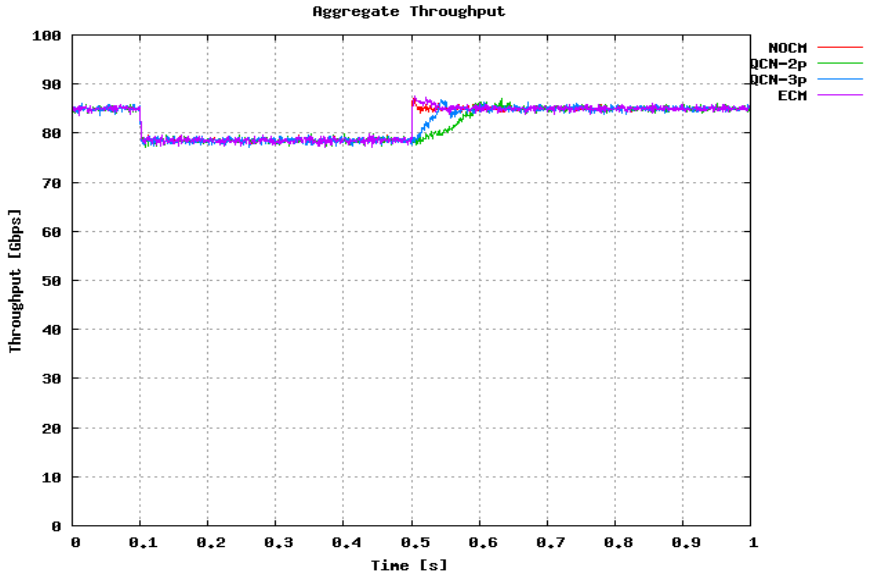
  ECM/QCN-2pt ~ O(CP) * O(B)

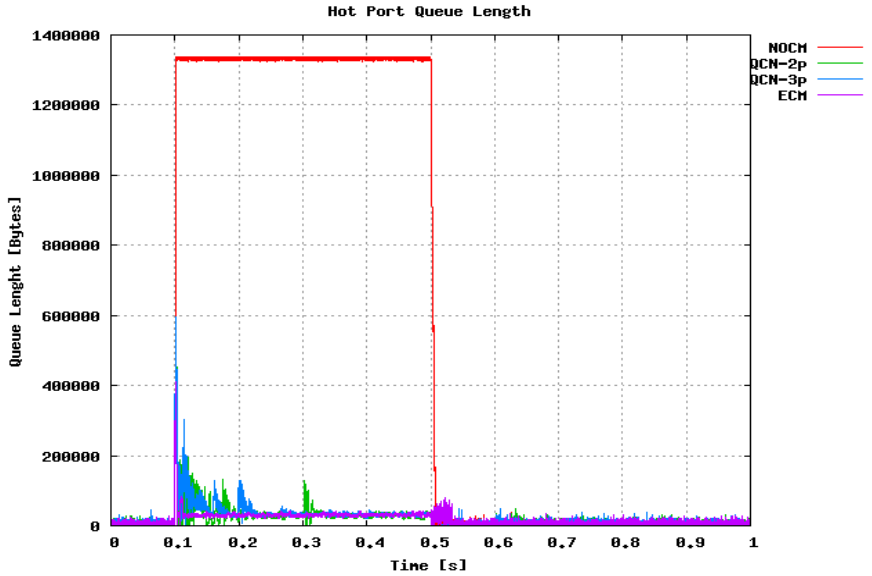
  N: number of nodes, L: offered load
  CP: number of CPs, B: bottleneck BW

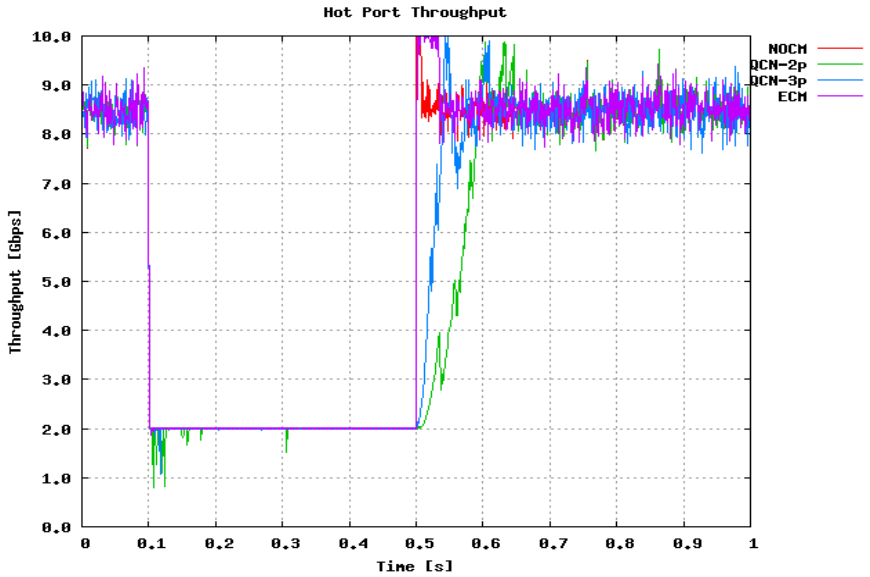
# Pause Disabled

- Same scenario and workloads re-run with pause disabled
- Results may not be representative of reality
  - There is no reliable transport layer (no retransmissions)
    - Throughput is higher
  - Topology too simple
    - Does not show blocking due to frame loss on ISLs

# 2 Gb/s Hotspot (w/o Pause)

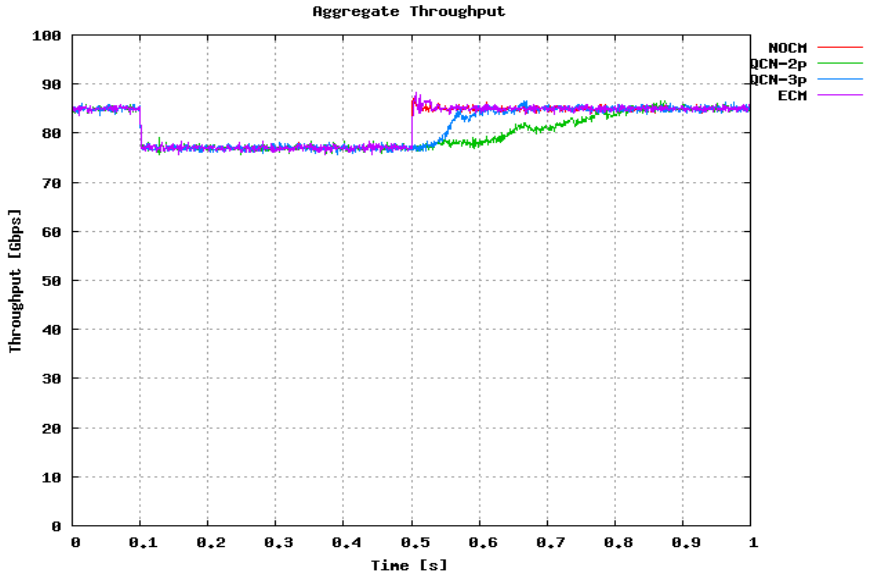


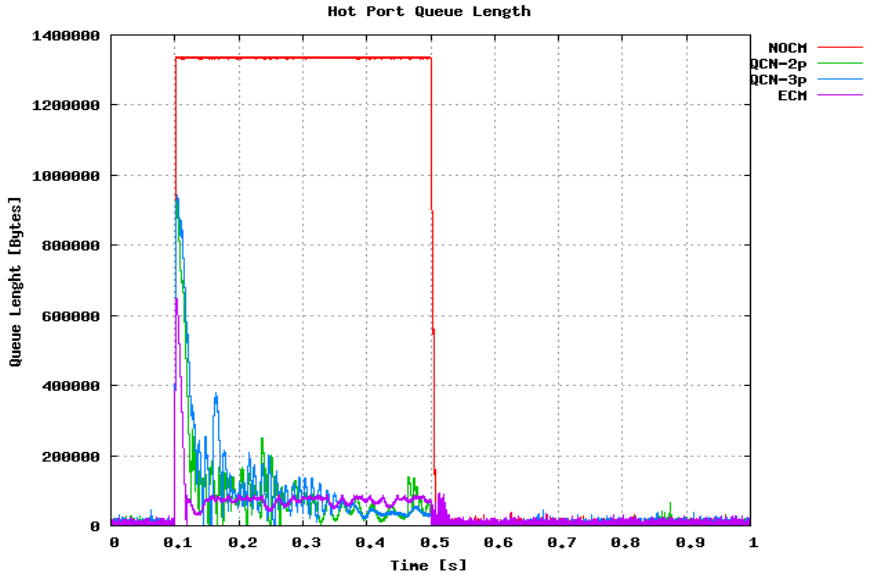



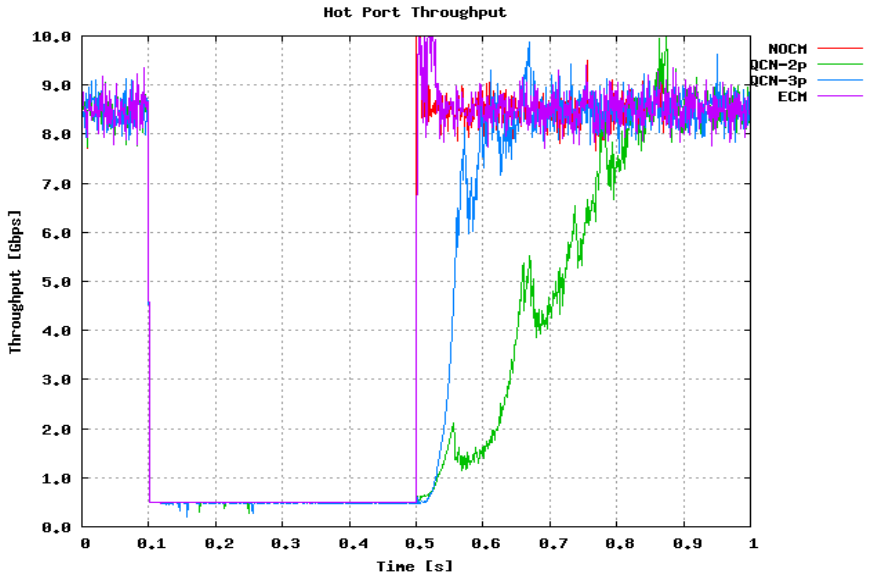




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 82.41 | 5.91 | 0 | 214569 | 0 |
| QCN-2p | 82.02 | 5.52 | 0 | 54 | 0.500 |
| QCN-3p | 82.23 | 5.73 | 0 | 37 | 140.09 |
| ECM | 82.44 | 5.95 | 0 | 0 | 4.83 |

# 0.5 Gb/s Hotspot (w/o Pause)

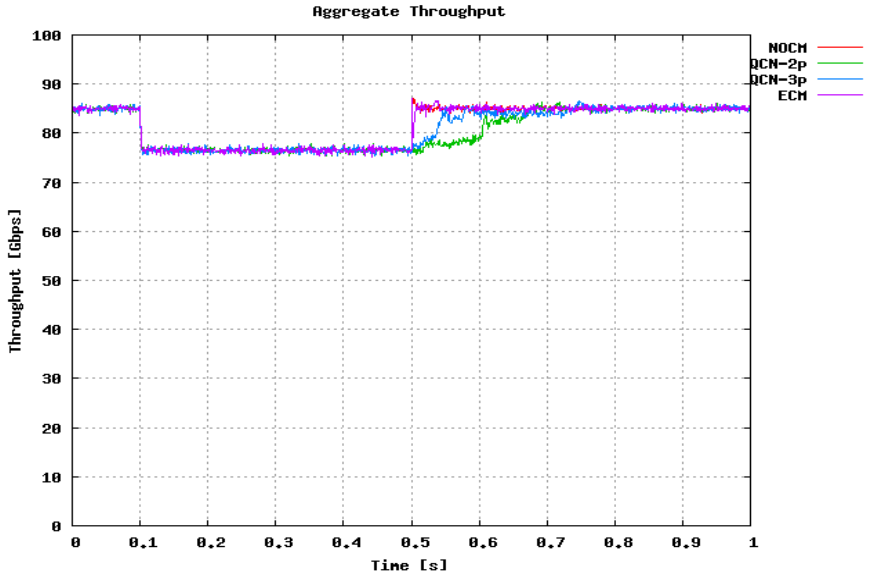


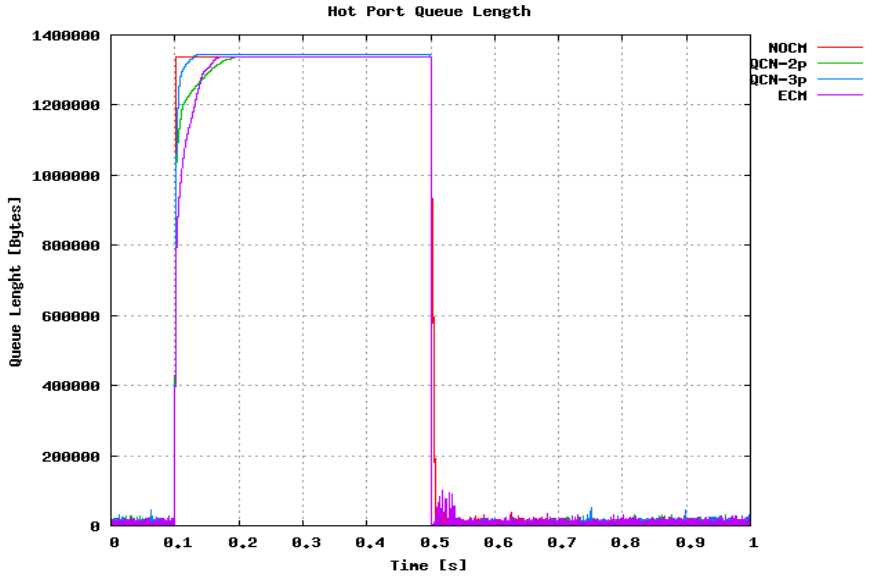



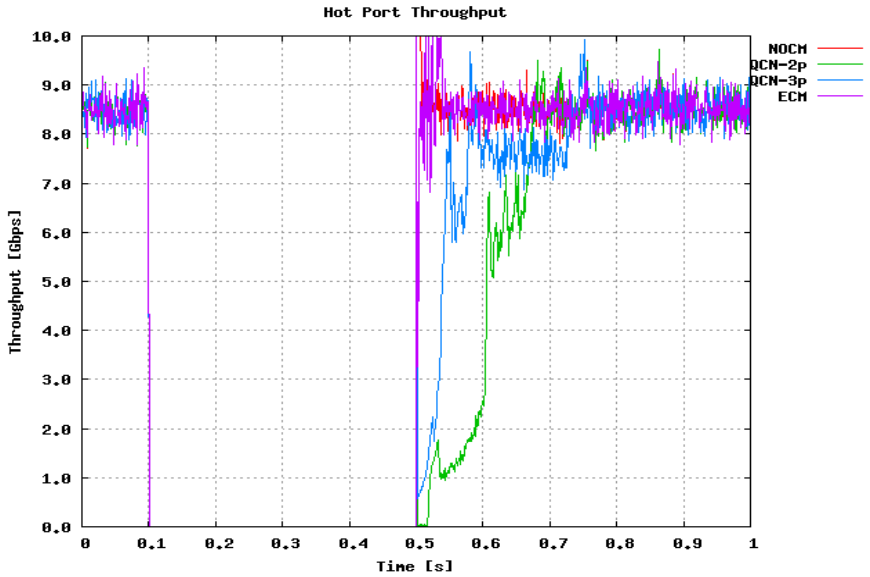




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 81.81 | 5.31 | 0 | 264575 | 0 |
| QCN-2p | 80.33 | 3.82 | 0 | 665 | 0.419 |
| QCN-3p | 81.03 | 4.81 | 0 | 204 | 139.1 |
| ECM | 81.81 | 5.32 | 0 | 0 | 4.51 |

# 0 Gb/s Hotspot (w/o Pause)



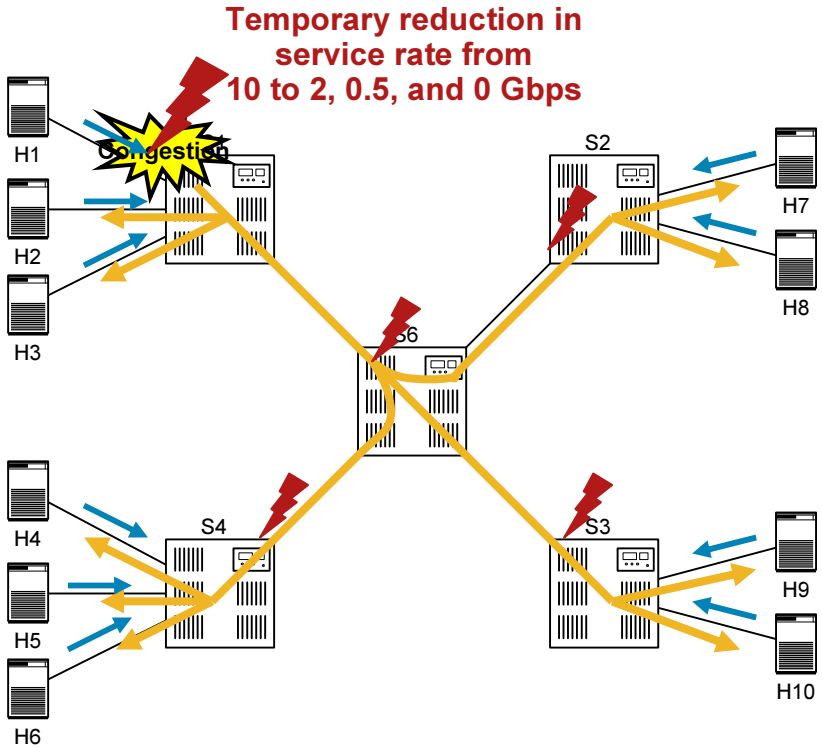

| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 81.61 | 5.11 | 0 | 281219 | 0 |
| QCN-2p | 80.71 | 4.21 | 0 | 3509 | 3.39 |
| QCN-3p | 81.15 | 4.66 | 0 | 14537 | 133.05 |
| ECM | 81.59 | 5.10 | 0 | 3143 | 27.18 |

# Observations

- Same as slide #7 and #12

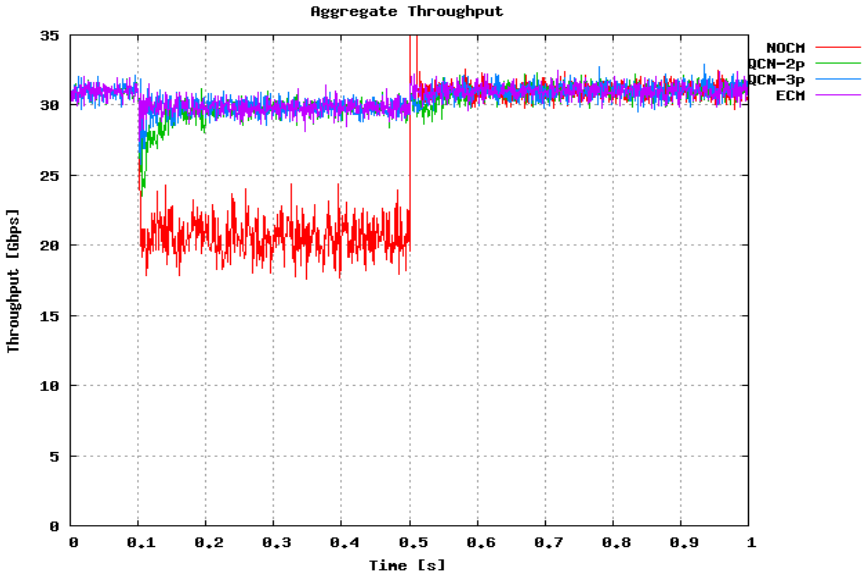
# Multi-Hop Output Generated Hotspot

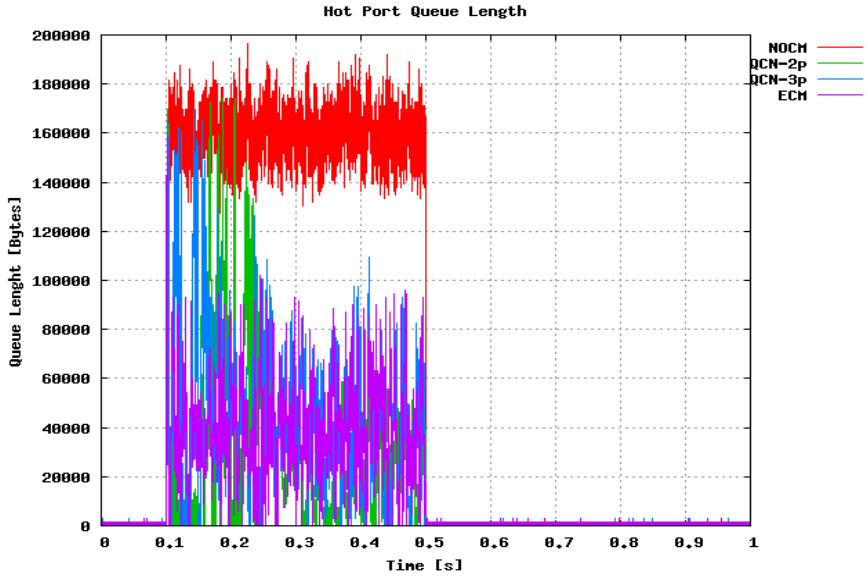
- Topology & Workload

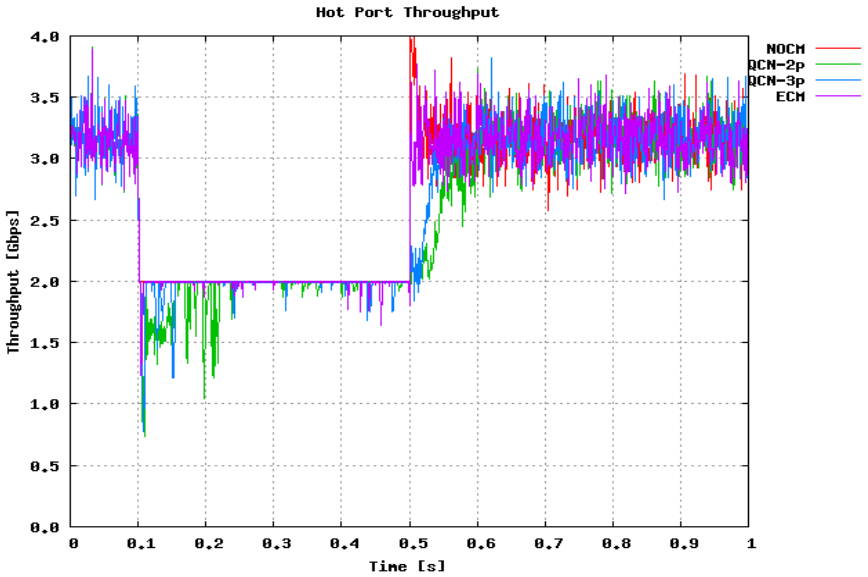



- Traffic pattern
  - Load: H1-6 = 25% H7-10 = 40%
  - Spatially: Uniform (except self)
  - Temporally: Bernoulli
- Hotspot
  - Duration: 400 ms, from $t_i$ = 100 to $t_f$ = 500 ms
  - HS degree = 9
  - HS severity = 1.58 / 6.33 / ∞ : 1

# 2 Gb/s Hotspot (with Pause)

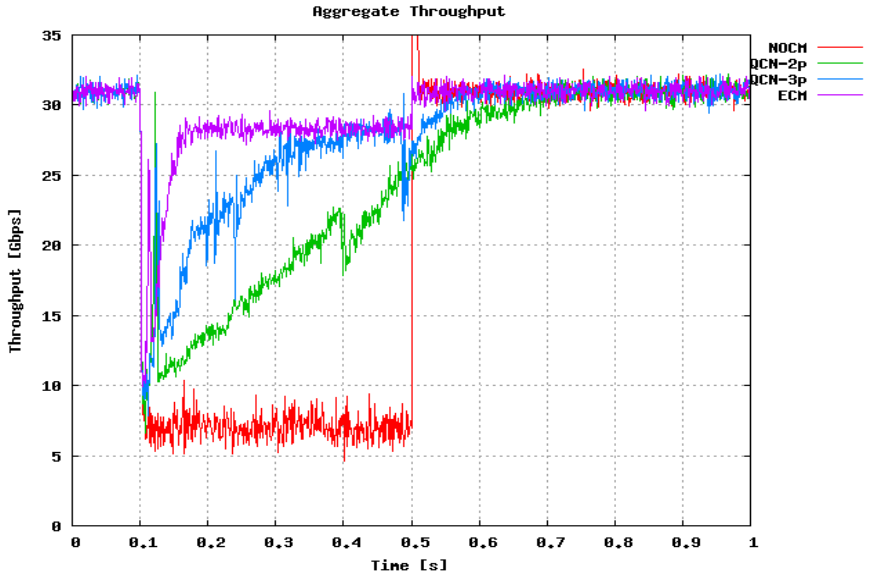


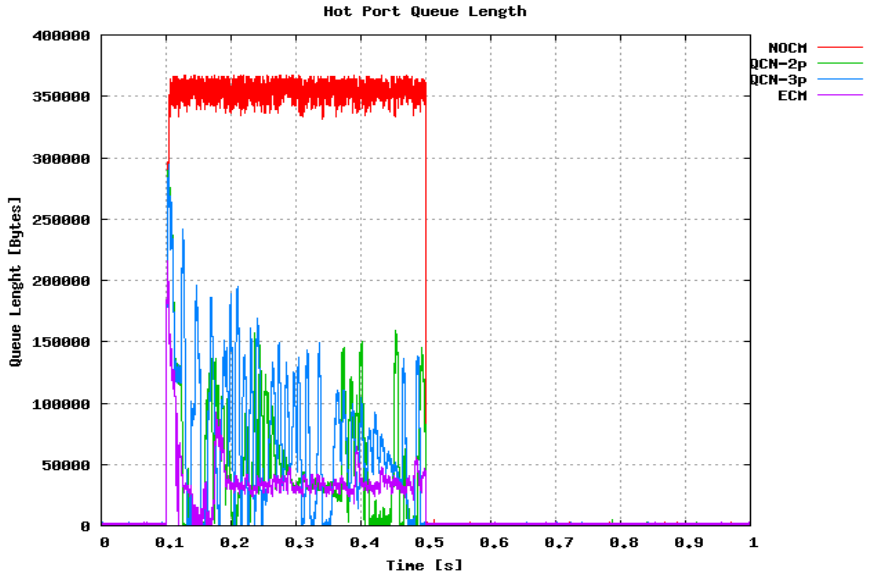



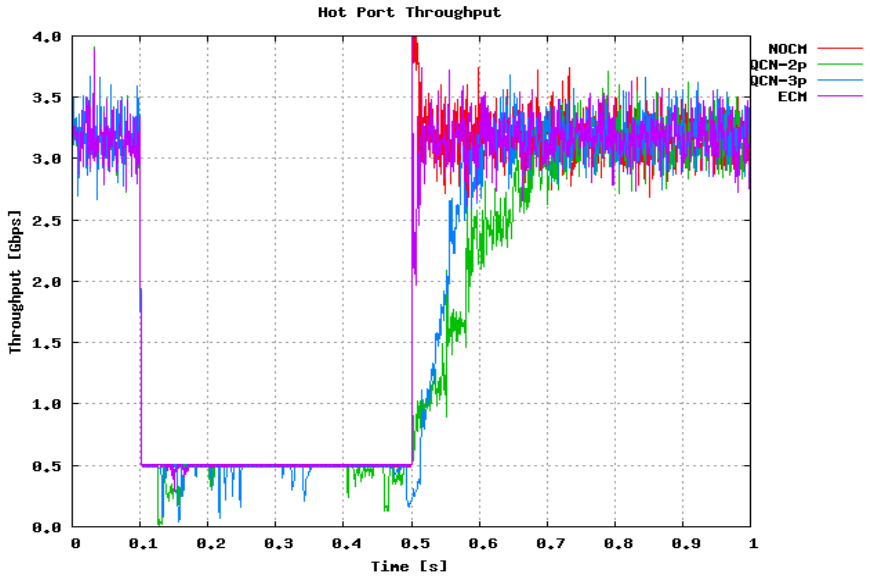




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 26.91 | 2.70 | 29.85 | 0 | 0 |
| QCN-2p | 30.32 | 2.60 | 0.10 | 0 | 0.53 |
| QCN-3p | 30.48 | 2.65 | 0.12 | 0 | 52.37 |
| ECM | 30.51 | 2.69 | 0.03 | 0 | 1.56 |

# 0.5 Gb/s Hotspot (with Pause)

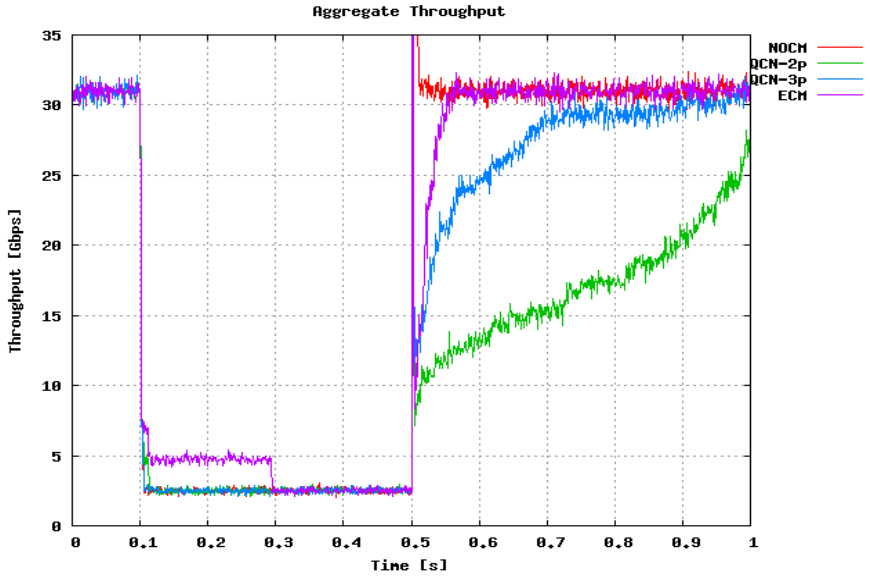


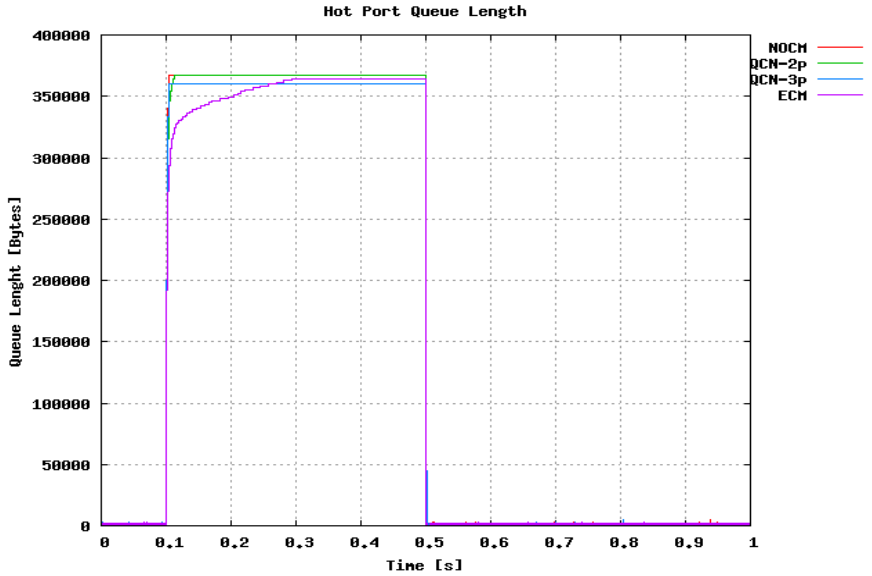



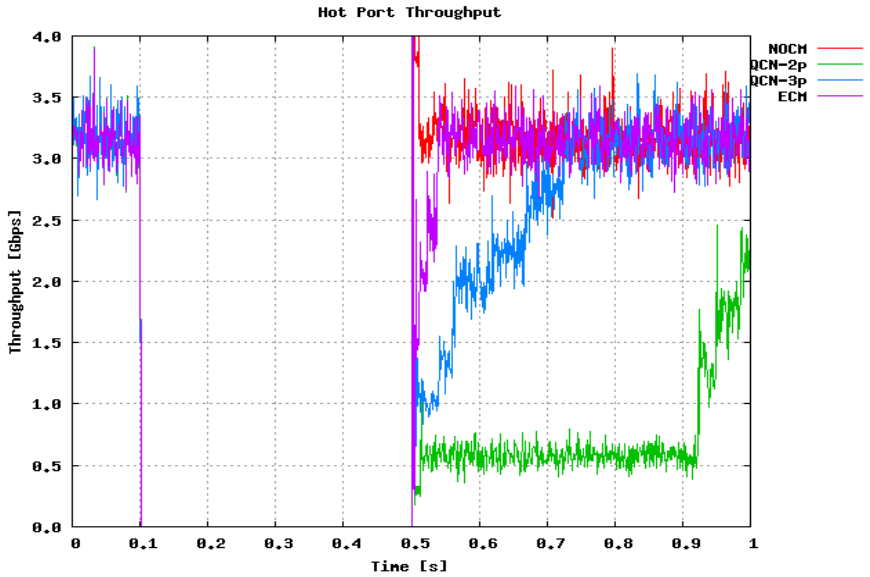




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 21.58 | 2.10 | 35.48 | 0 | 0 |
| QCN-2p | 25.19 | 1.85 | 1.96 | 0 | 0.84 |
| QCN-3p | 27.95 | 1.95 | 2.54 | 0 | 48.73 |
| ECM | 29.41 | 2.08 | 1.02 | 0 | 2.31 |

# 0 Gb/s Hotspot (with Pause)




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 19.70 | 1.91 | 40.10 | 0 | 0 |
| QCN-2p | 12.73 | 0.70 | 39.89 | 0 | 0.62 |
| QCN-3p | 17.81 | 1.64 | 37.11 | 0 | 31.06 |
| ECM | 19.62 | 1.85 | 38.43 | 0 | 2.84 |

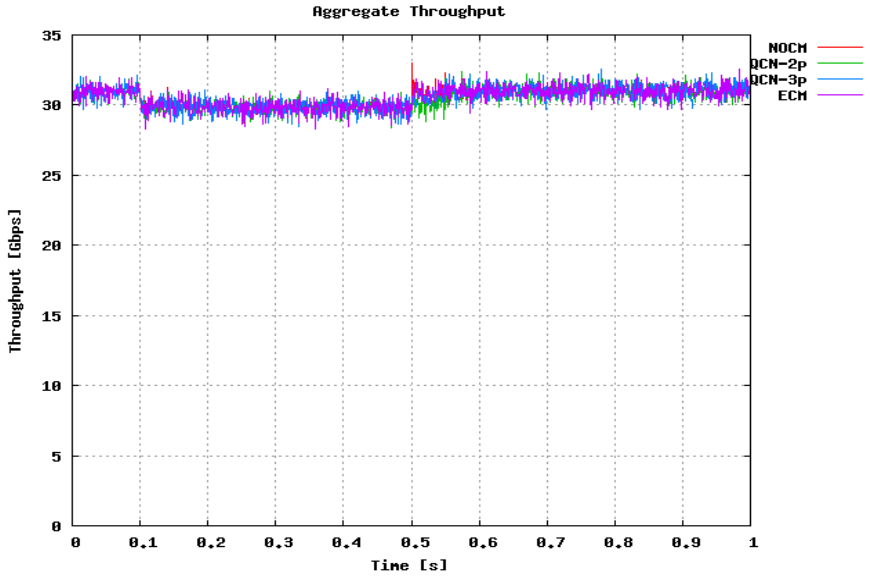
# Observations

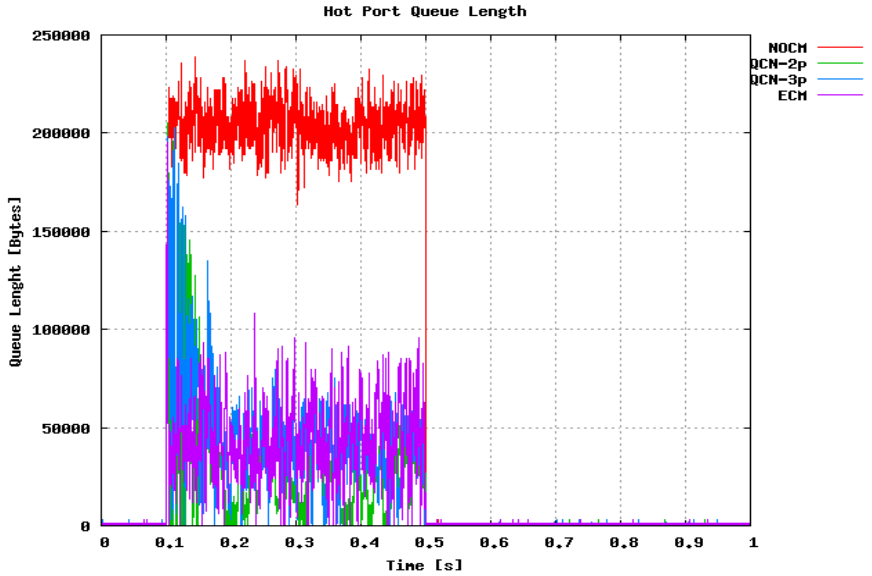
- Same as slide #7 and #12

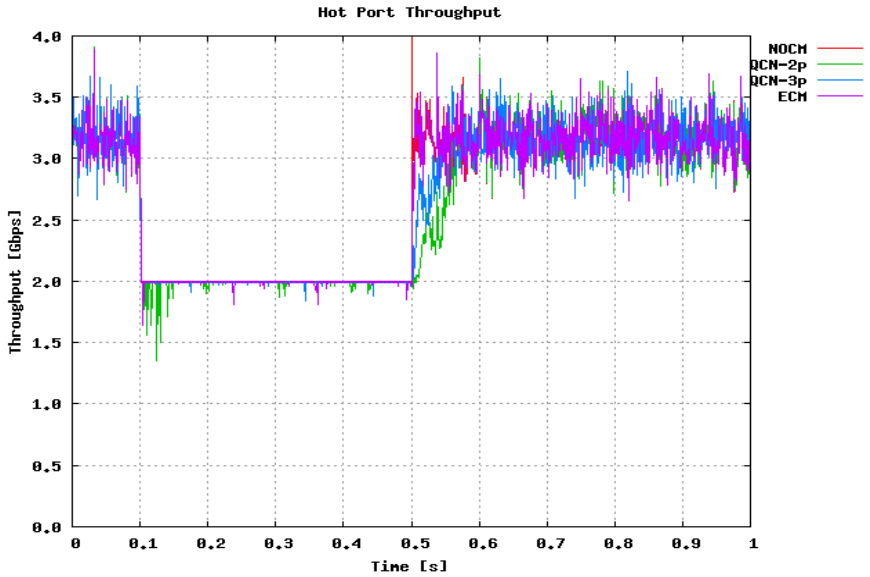
# Pause Disabled

- Same scenario and workloads re-run with pause disabled
- Results may not be representative of reality
  - There is no reliable transport layer (no retransmissions)
    - Throughput is higher
  - Topology too simple
    - Does not show blocking due to frame loss on ISLs

# 2 Gb/s Hotspot (w/o Pause)

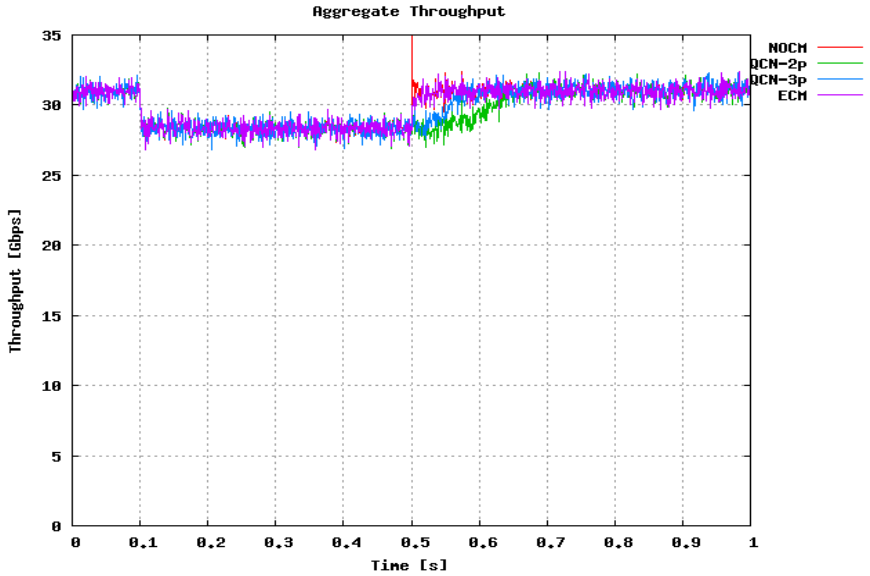


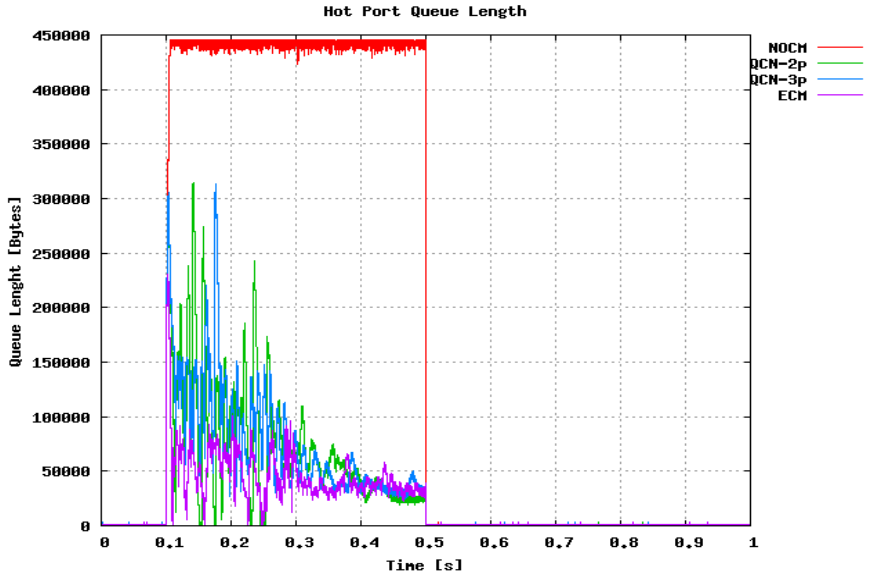



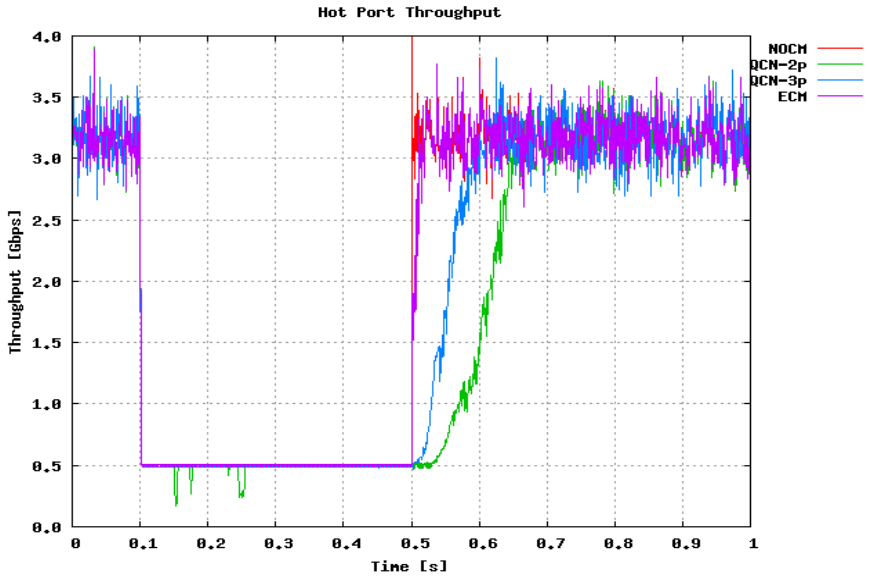




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 30.52 | 2.7 | 0 | 38945 | 0 |
| QCN-2p | 30.47 | 2.64 | 0 | 119 | 0.37 |
| QCN-3p | 30.52 | 2.67 | 0 | 115 | 51.95 |
| ECM | 30.52 | 2.69 | 0 | 30 | 0.99 |

# 0.5 Gb/s Hotspot (w/o Pause)

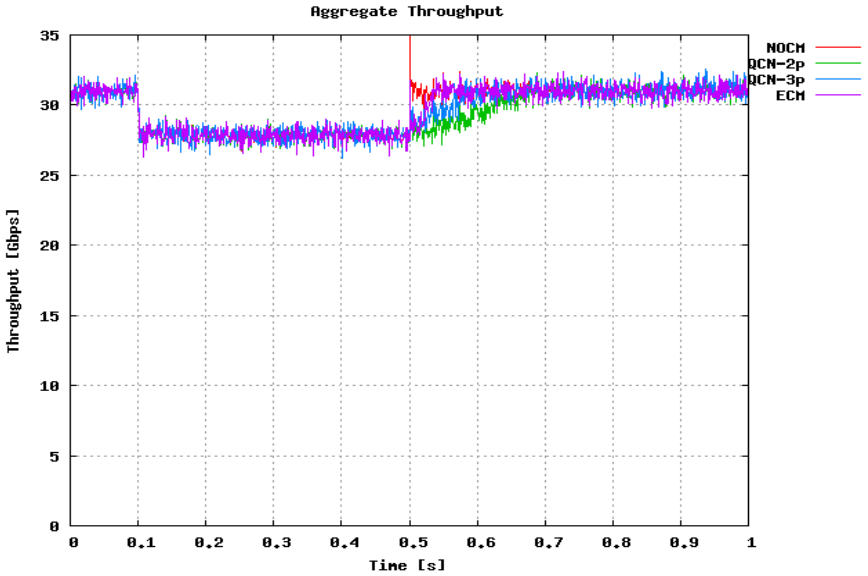


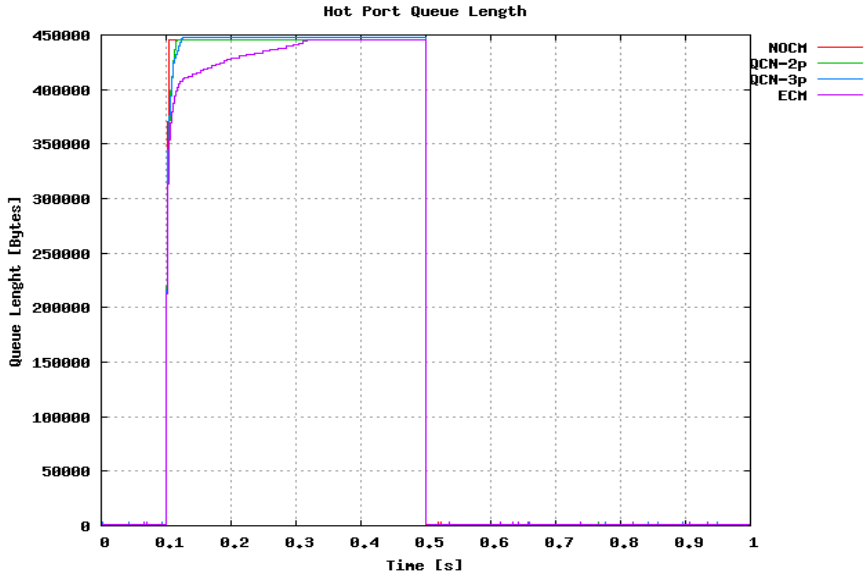



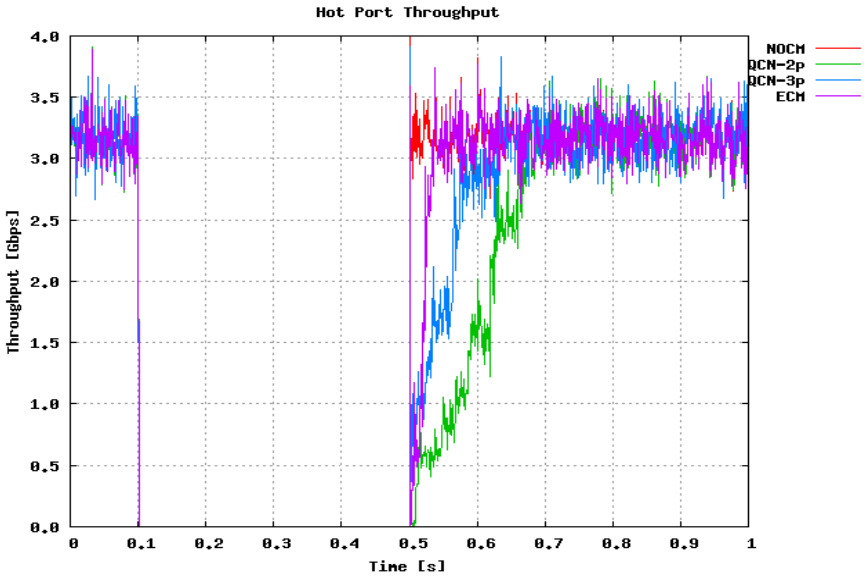




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 29.92 | 2.10 | 0 | 88785 | 0 |
| QCN-2p | 29.63 | 1.80 | 0 | 701 | 0.31 |
| QCN-3p | 29.80 | 1.96 | 0 | 888 | 51.01 |
| ECM | 29.91 | 2.08 | 0 | 256 | 0.5 |

# 0 Gb/s Hotspot (w/o Pause)



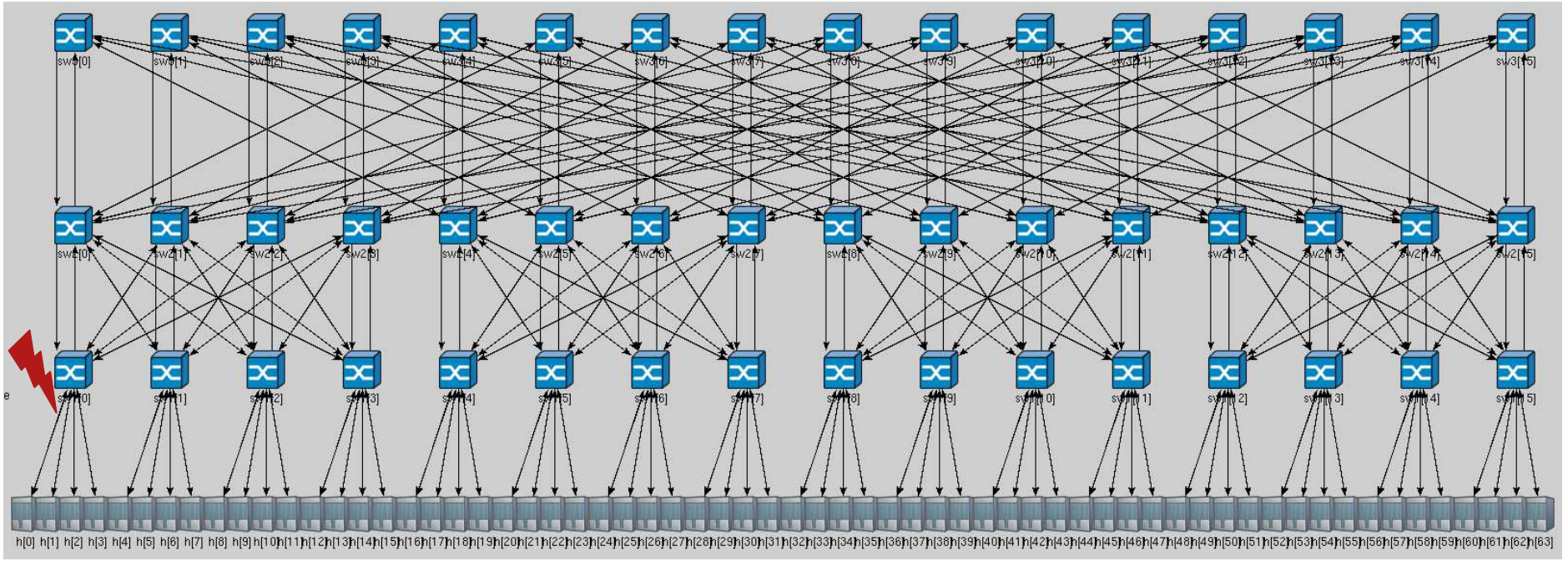

| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 29.76 | 1.90 | 0 | 210906 | 0 |
| QCN-2p | 29.42 | 1.60 | 0 | 4186 | 0.23 |
| QCN-3p | 29.62 | 1.76 | 0 | 7040 | 49.22 |
| ECM | 29.66 | 1.84 | 0 | 855 | 0.36 |

# Observations

- Same as slide #7 and #12

# Fat Tree Output Generated Hotspot

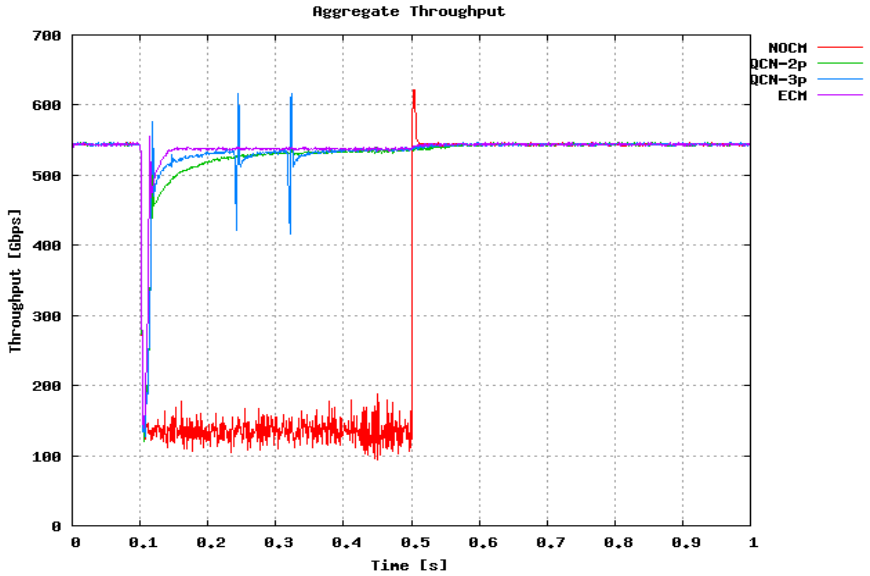
- Topology: 3-level Fat Tree (4-ary n-fly) with 48 switches and 64 hosts
  - Optimal shortest-path forwarding & optimal load balancing (perfect hashing)

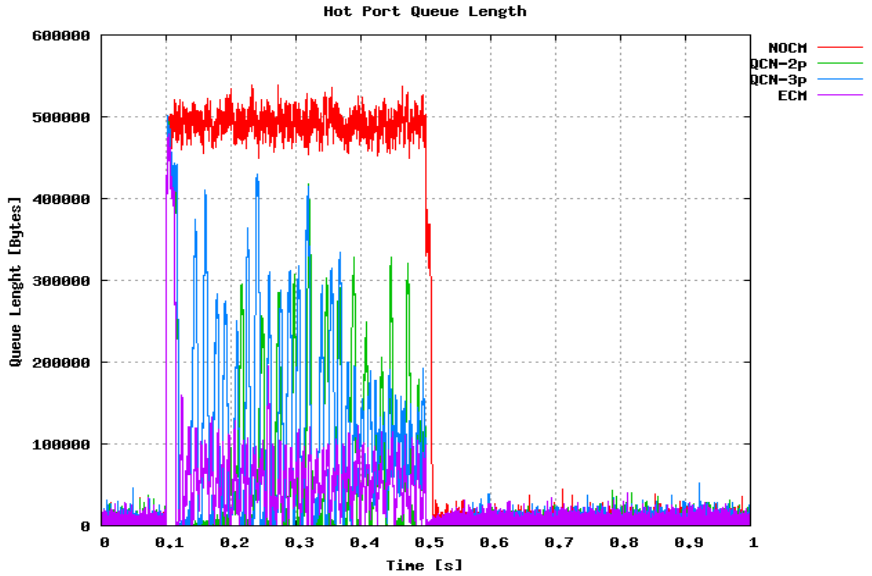


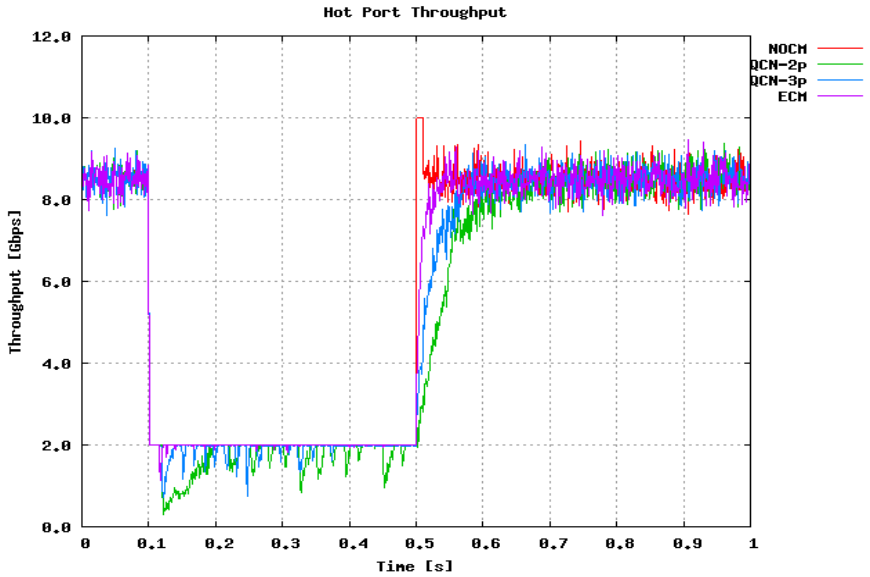
- Traffic pattern
  - Load 85%
  - Spatially: Uniform (except self)
  - Temporally: Bernoulli

- Hotspot
  - Duration: 400 ms, from $t_i$ = 100 to $t_f$ = 500 ms
  - Rate during HS: 2 and 0.5 Gbps
  - HS degree = 63
  - HS severity = 4.25 / 17 : 1

# 2 Gb/s Hotspot (with Pause)

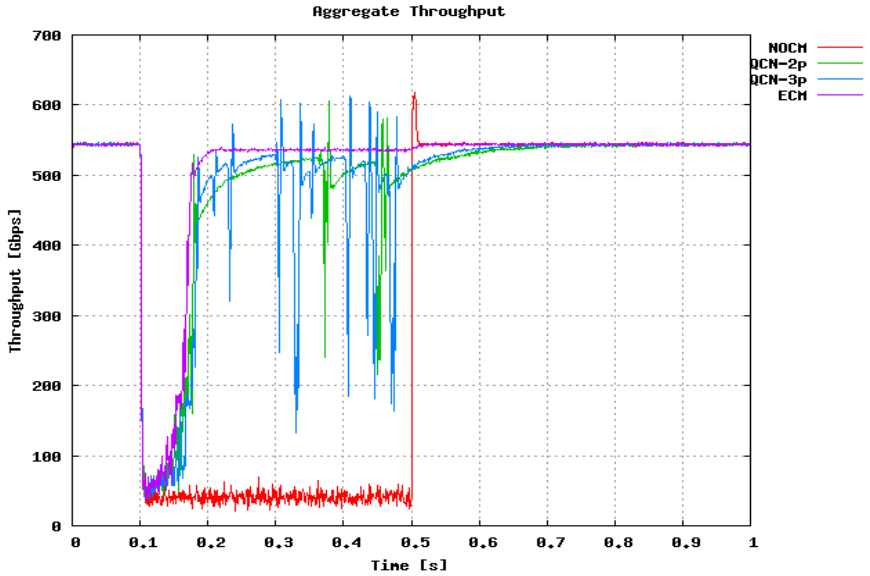


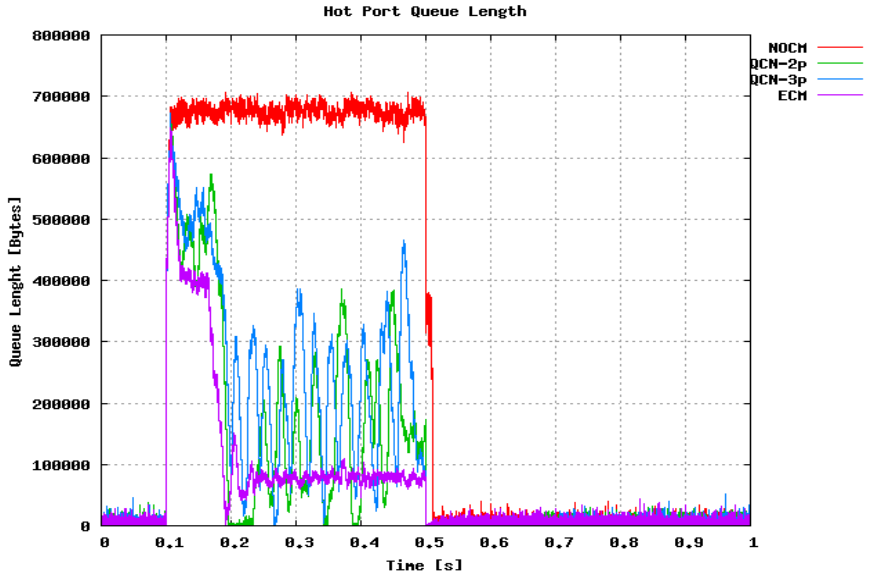



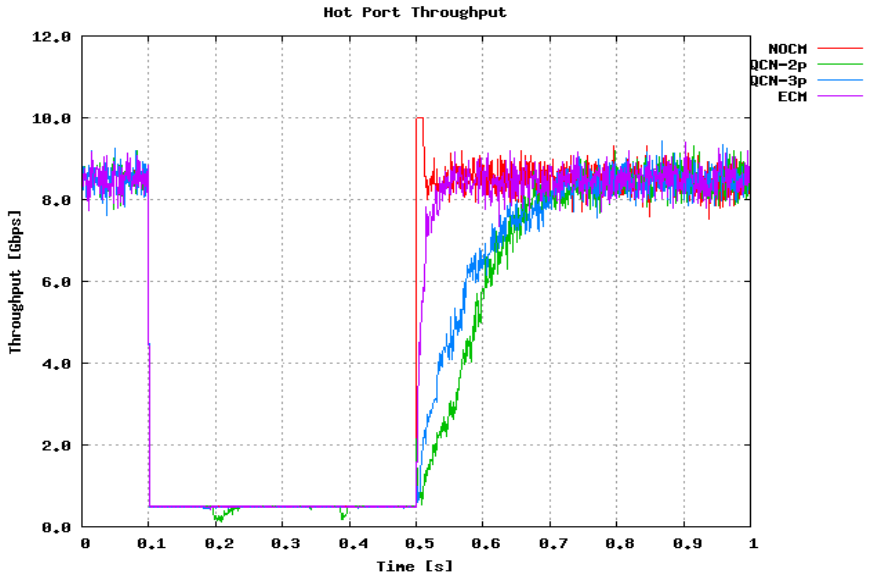




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 381.33 | 5.9 | 29.56 | 0 | 0 |
| QCN-2p | 530.91 | 5.47 | 1.04 | 0 | 12.1 |
| QCN-3p | 533.08 | 5.71 | 1.18 | 0 | 923.8 |
| ECM | 537.02 | 5.83 | 0.79 | 0 | 141.87 |

# 0.5 Gb/s Hotspot (w/o Pause)




| | Aggregate Throughput [Gb/s] | Hot Port Throughput [Gb/s] | Avg Time Paused [%] | Dropped Frames [#] | Overhead [Mb/s] |
|---|---|---|---|---|---|
| NOCM | 343.55 | 5.3 | 37.31 | 0 | 0 |
| QCN-2p | 491.50 | 4.59 | 6.54 | 0 | 25.95 |
| QCN-3p | 487.58 | 4.75 | 7.84 | 0 | 860.16 |
| ECM | 510.66 | 5.2 | 5.49 | 0 | 123.51 |

# Final Remarks

- QCN-2pt and QCN-3pt seem to do their job at controlling congestion
- However, being a simplification of ECM, their performance is somewhat lower, especially recovery from congestion events seems slower
- QCN-3pt does not seem too efficient with regards to control traffic overhead (feedback is generated also in absence of congestion)
- Recommendations: **focus on QCN-2pt and improve recovery time**

# Next steps

- Complete fat tree simulations
- Cover remaining scenarios