



Initial QCN Serial HAI Results: Effects of PAUSE and RTT Delay

Brad Matthews, Bruce Kwan & Ashvin Lakshmikantha
IEEE 802.1Qau Interim Meeting (Los Gatos, CA) – Jan 2008

Goals

- Quantify the impact of PAUSE on innocent flow throughput when operating with QCN Serial HAI
- Evaluate effects of increasing RTT delay

Parameters

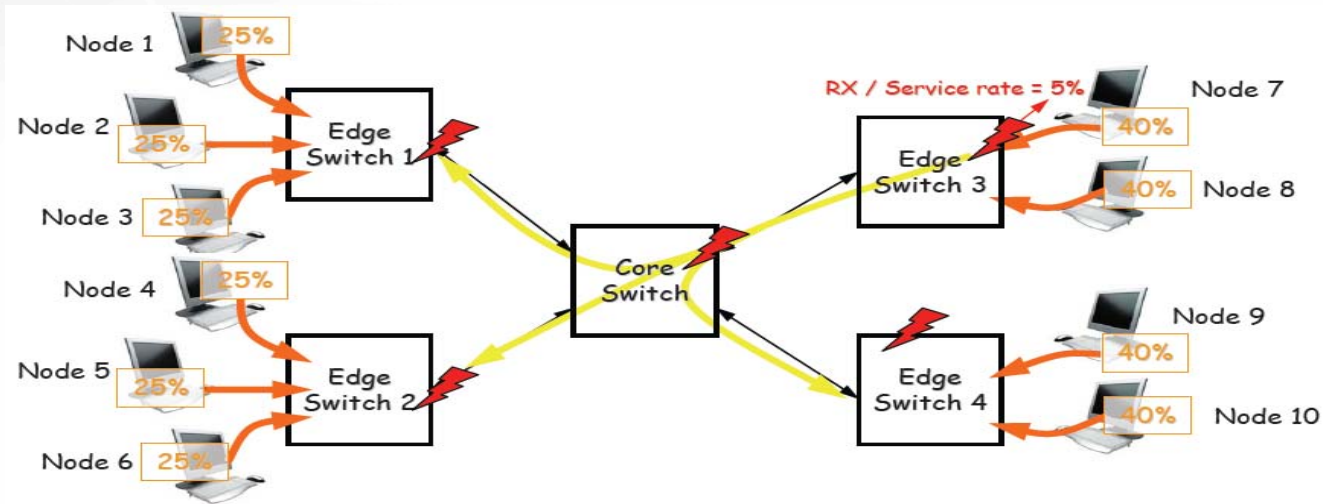
- Switch Parameters

- Shared memory switch
 - Memory Size 2.4Mbytes
- Partitioned memory per input, shared among all outputs
- PAUSE Disabled
 - Output queue limit of 150kbytes
- PAUSE Enabled
 - No output queue limit
 - Applied on a per input basis based on watermarks
 - Watermark_hi = 130kbytes
 - Watermark_lo = 110kbytes

- QCN Parameters

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

Topology and Workload: 800ms Congestion Period

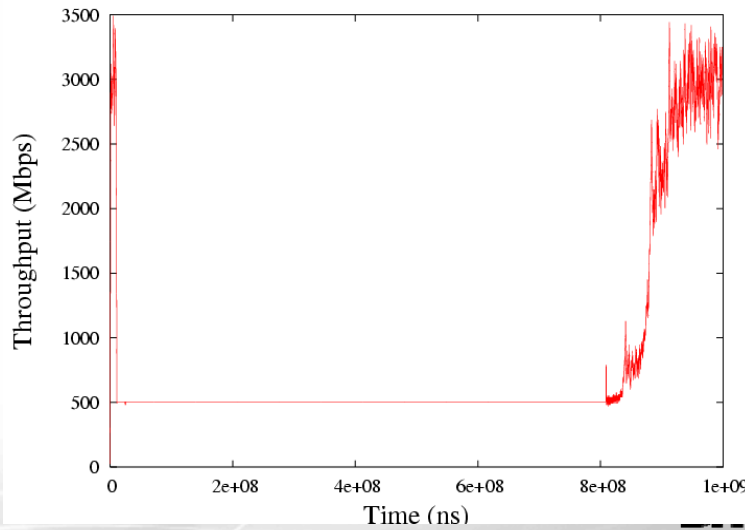
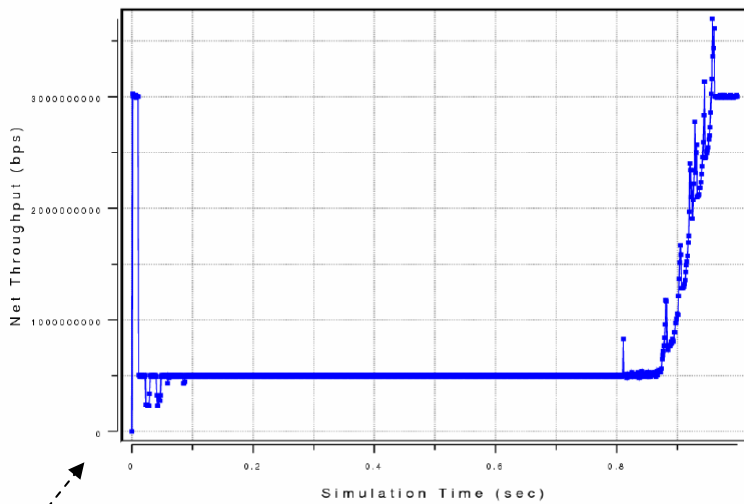
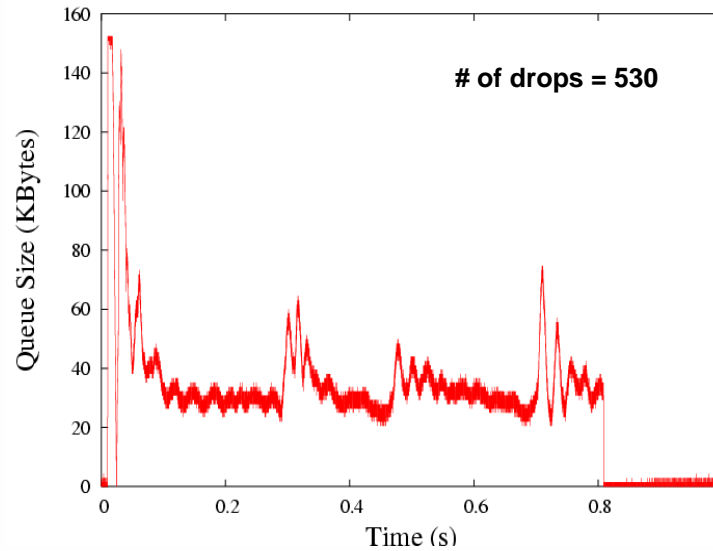
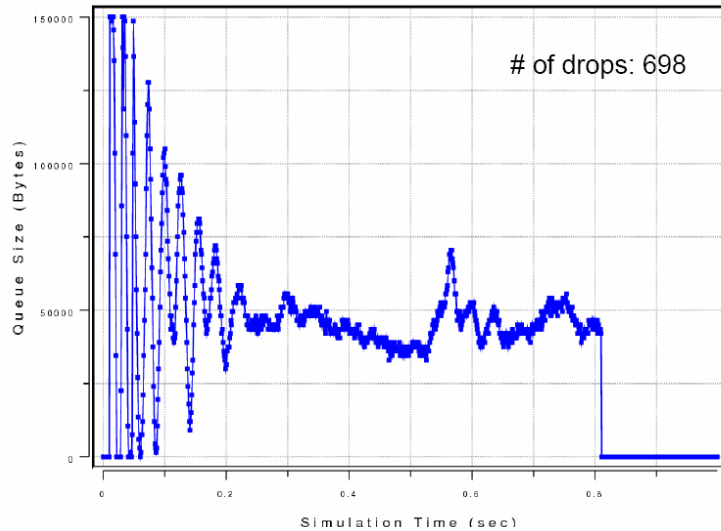


- Multi-stage Output-Generated Hotspot Scenario
 - Link Speed = 10Gbps for all links
 - Loop Latency = 16us
- Traffic Pattern
 - 100% UDP (or Raw Ethernet) Traffic
 - Destination Distribution: Uniform distribution to all nodes (except self)
 - Frame Size Distribution: Fixed length (1500bytes) frames
 - Offered Load
 - Nodes 1-6 = 25% (2.5Gbps)
 - Nodes 7-10 = 40% (4Gbps)
- Congestion Scenario
 - Node 7 temporary reduce its service rate from 10Gbps to 500Mbps between [10ms, 810ms]
- PAUSE Disabled

* Topology and Workload based on Benchmark #2: OG HS Multi-Hop. Congestion

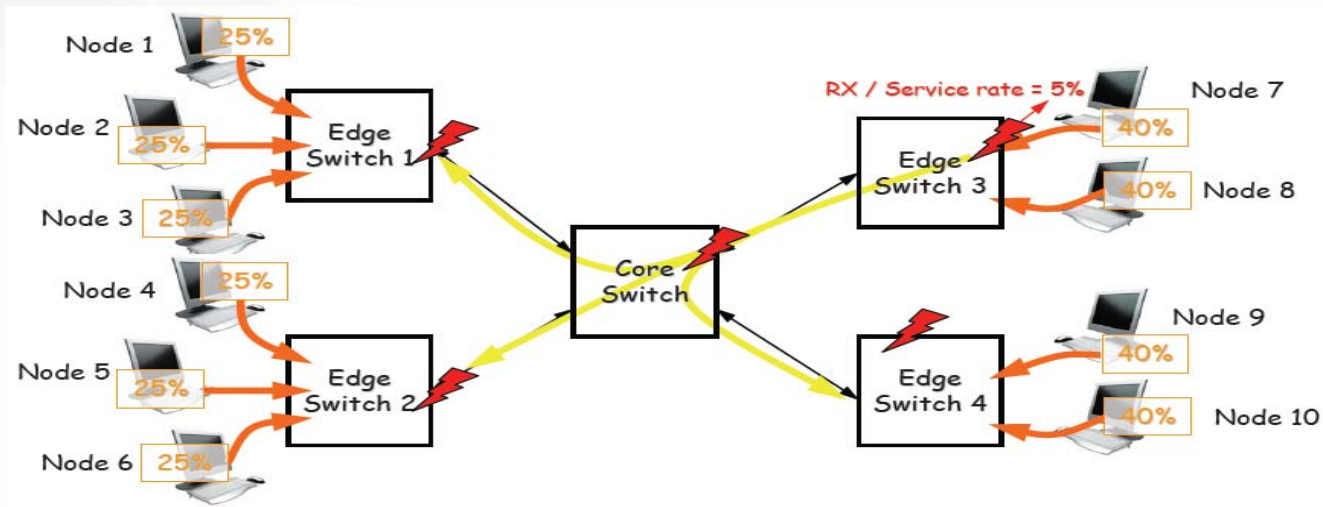
Picture is from: <http://www.ieee802.org/1/files/public/docs2006/au-sim-Zurich-Hotspot-Benchmark-OG-MS-r2.pdf>.

Queue Size & Hot Spot Throughput



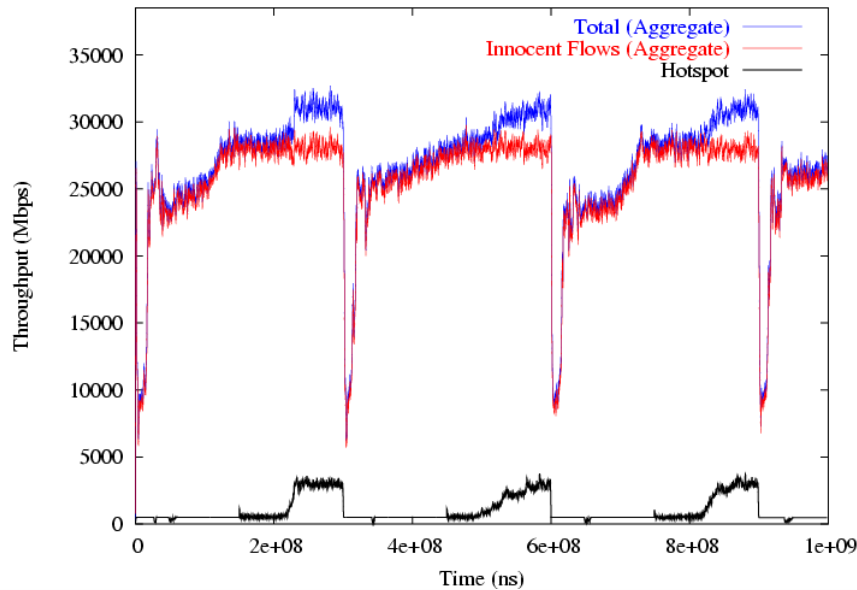
From <http://www.ieee802.org/1/files/public/docs2008/au-sim-rong-qcn-hai-updatesimu-0108-1.pdf>

Topology and Workload: Periodic Congestion Events



- Traffic Pattern
 - Same as before
- Congestion Scenario
 - Node 7 periodically reduces its service rate from 10Gbps to 500Mbps
 - Congestion Duration: 25ms – 200ms
 - Duty Cycle = 1/2
- Simulation Duration: 1 second
- Performance Metric: Aggregate Throughput
 - Ideal Aggregate Innocent Flow Throughput: 28Gbps
 - Ideal Aggregate Victim Flow Throughput: 500Mbps or 3Gbps
- PAUSE Enabled

Impact of Congestion Spread on Innocent Flow Utilization



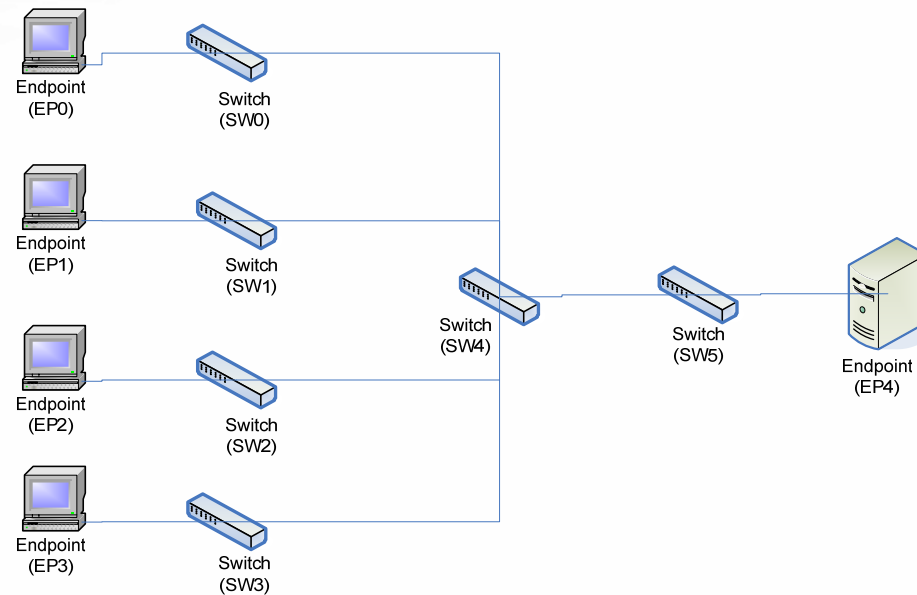
Congestion Duration: 150 ms (on/off)

QCN + PAUSE:

Innocent Flow Utilization: 91.15%
- 25.52 Gbps out of 28Gbps

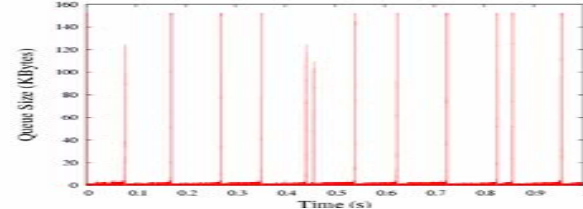
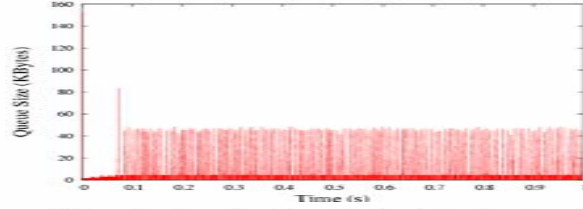
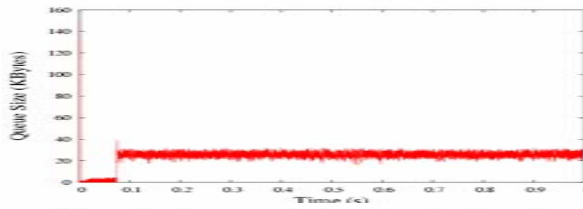
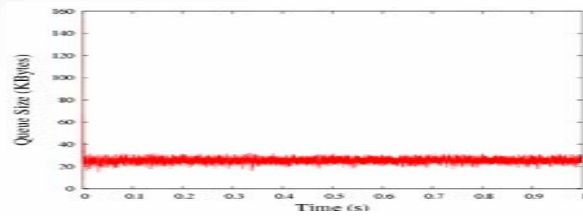
<i>Congestion Duration (ms)</i>	<i>Innocent Flow Utilization (%)</i>	
	<i>PAUSE Only</i>	<i>QCN + PAUSE</i>
25	67.29	97.40
80	63.91	95.21
125	63.82	91.31
150	61.21	91.15
200	57.48	93.77

Symmetric Topology Single HS



- Symmetric Topology Single HS
 - Link speed : 10Gbps for all links
- Traffic Pattern
 - Traffic Type: 100% UDP (or Raw Ethernet) Traffic
 - Destination Distribution: EP0-EP3 send to EP4
 - Frame Size Distribution: Fixed length (1500 bytes) frames
 - Arrival Distribution: Bernoulli temporal distribution
 - Offered Load/Endpoint = 100%
- Control Loop Delay is between the source Endpoints and the Congestion Point
- PAUSE Disabled

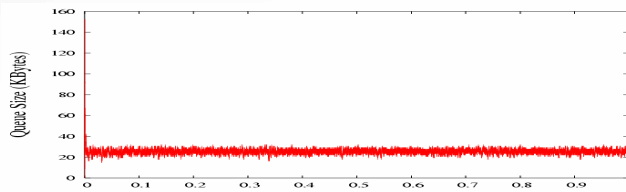
Effects of RTT Delay (4 to 1 Case)



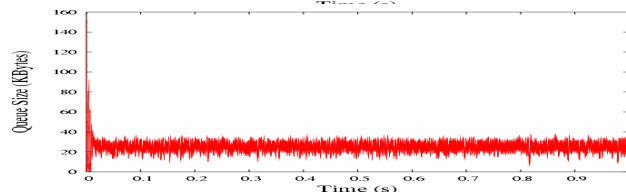
<u>Ctrl Loop Delay</u>	<u># of drops</u>	<u>Throughput</u>
8 us	6	9.99 Gbps
100 us	230	9.69 Gbps
500 us	1229	9.45 Gbps
1 ms	3746	2.23 Gbps

Varying Number of Sources

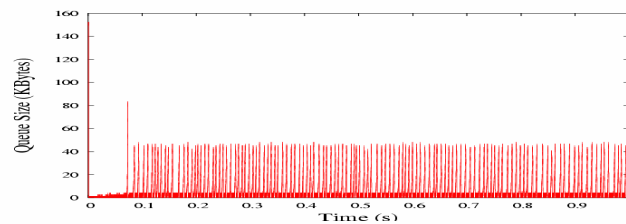
4-to-1



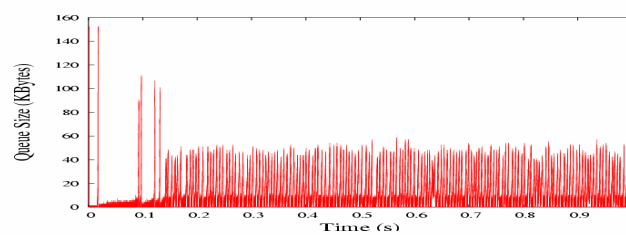
16-to-1



4-to-1



16-to-1



Ctrl Loop Delay

of drops

Throughput

8 us

6

9.999 Gbps

8 us

441

9.997 Gbps

500 us

1229

9.446 Gbps

500 us

6653

9.277 Gbps

Next Steps

- Quantify interactions with TCP traffic
- Evaluate scenarios with heterogeneous link rates
- Evaluate larger topologies