

Implementing QCN

January 2009

Nobuharu Kami, Yoichi Hidaka, Jun-ichi Higuchi, Takumi Kosawa, Takashi
Yoshikawa, Atsushi Iwata
NEC Corporation

Balaji Prabhakar
Stanford University

Zhi-Hern Loh, Oscar Ham, Joseph Chang, Uri Cummings
Fulcrum

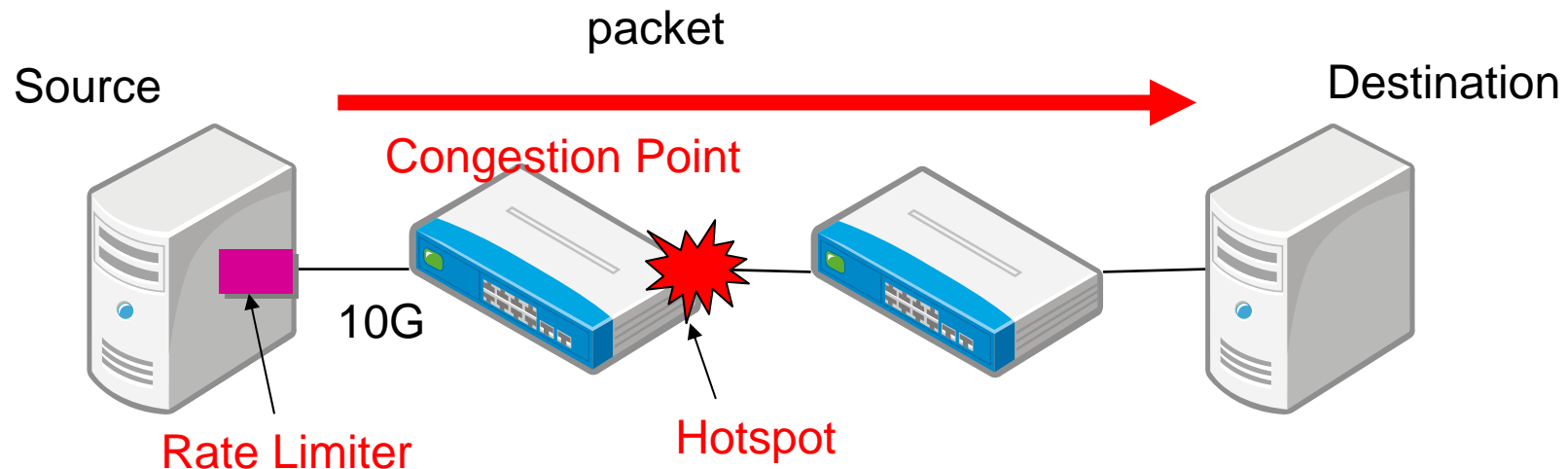
Hideaki Inagaki, Hironori Ando
Inventure

Overview

- **QCN experiments in a real network using ...**
 - QCN Switch (manufactured by Fulcrum)
 - QCN NIC (manufactured by NEC)
- **We observed how QCN goes in a real network**
 - Throughput
 - Queue length
- **Basically we confirmed that it operates as expected.**
- **Some issues in terms of implementation**
 - Burstiness of NIC's rate limiter

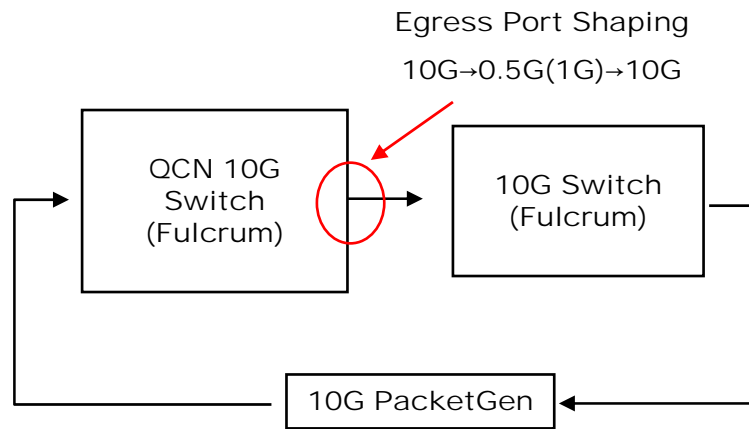
Evaluation Scenario

- **Scenario**
 - Topology: simple One source—One destination
 - Link speed: 10G
 - Traffic: CBR
 - Frame size: 312B (Fixed)
 - PAUSE disabled
- **Hotspot**
 - service at the link between switches is decreased to 0.5G (or 1G) for approx. 2 sec.

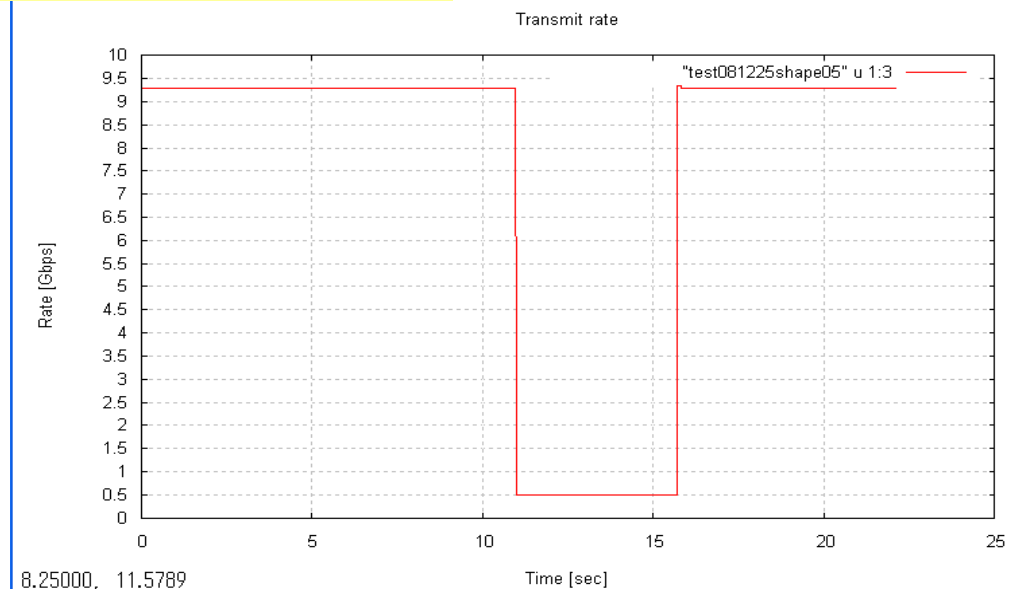


HOTSPOT (by Shaper at Egress Port)

- HOTSPOT is generated by shaping packets at the egress port of QCN switch
 - Link speed is decreased to 0.5G or 1G from 10G

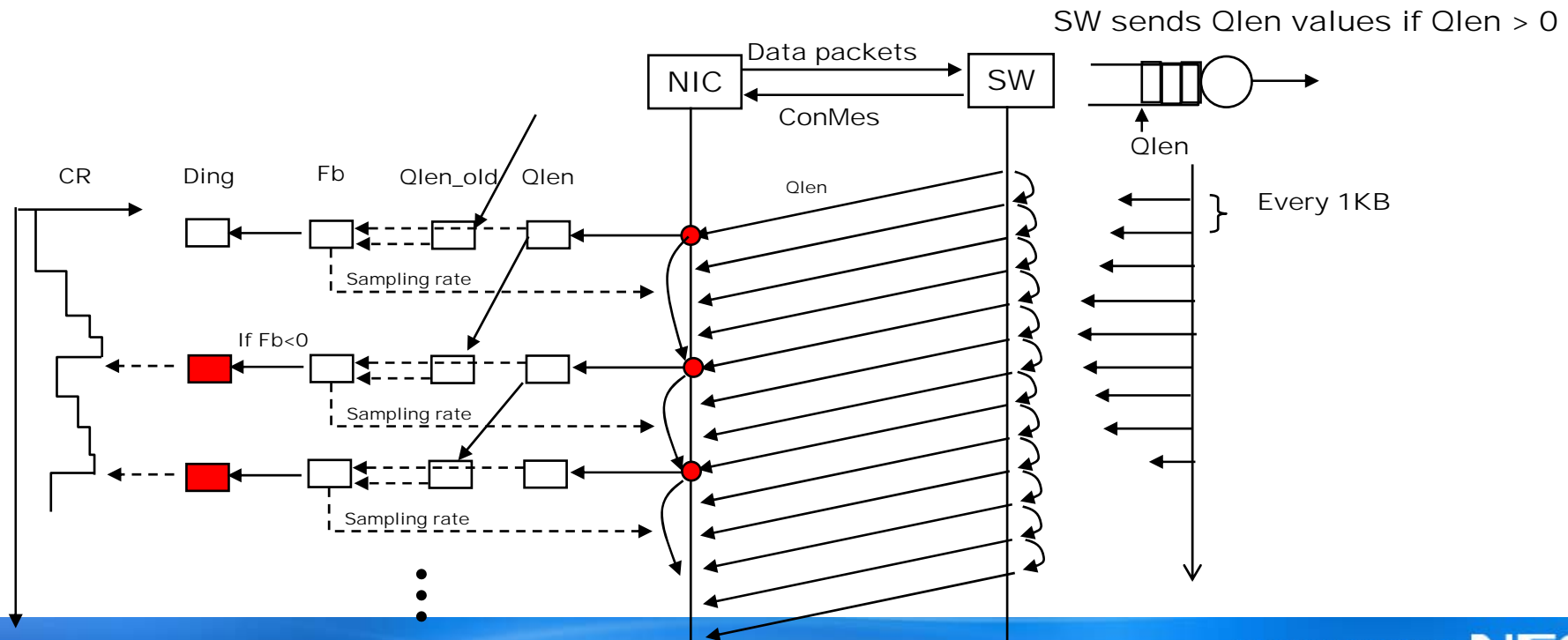


Available bandwidth



Fb Value Computation at the NIC Side

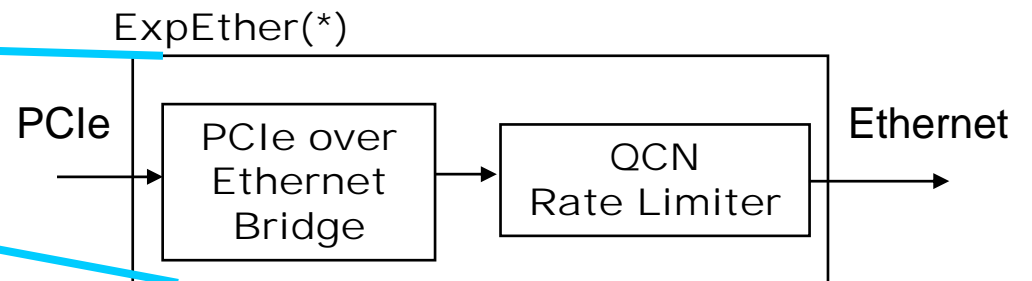
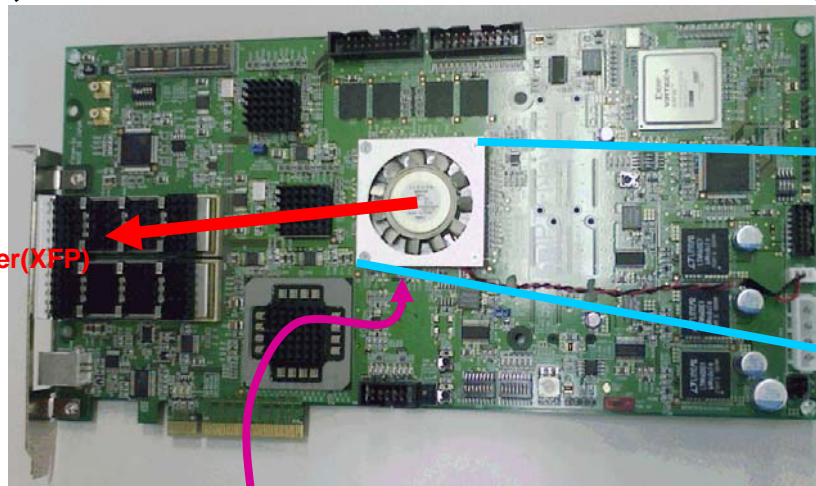
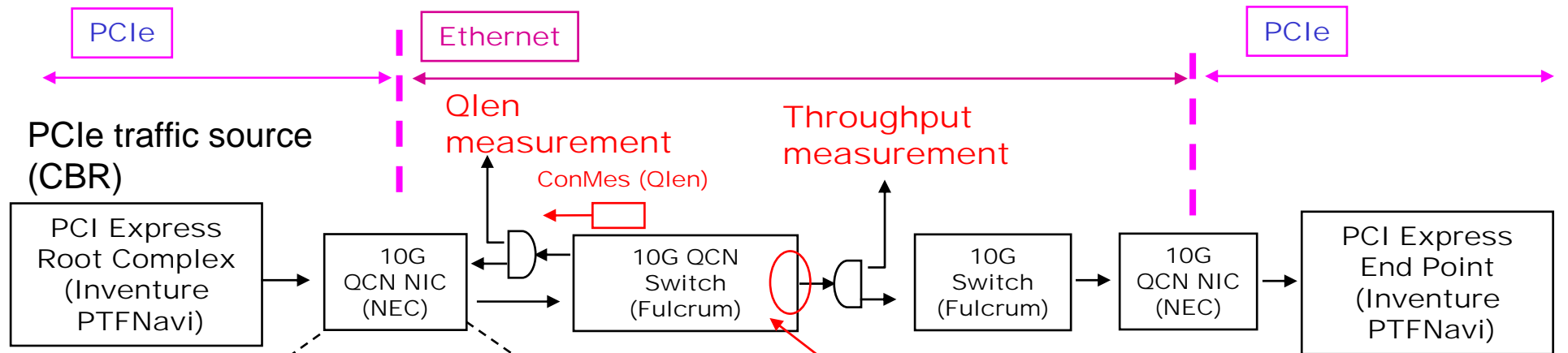
- Since we currently have no switch which can compute Fb value at high enough speed, we took the following alternative approach.
 - A switch just monitors Qlen, samples packets and send the Qlen info (say, Congestion Message: ConMes) to the NIC in every 1KB.
 - NIC “samples” the Qlen information to compute Fb values instead of the Switch and takes it as “ding” if Fb is negative
 - Next sampling interval is determined by current Fb value.



Typical Parameter Values

- **Switch**
 - Sampling rate: every 1KB
 - Switch Latency : 300ns
- **NIC**
 - FR threshold = 5
 - Fb = -1 ~ -63 (6 bits)
 - Qeq = 33000 bytes
 - w=2
 - Timer period = 10ms@FR, 5ms @AI and HAI
 - Byte counter limit = 150KB@FR, 75KB @AI and HAI
 - Gd = 0.0078125 (1/128)
 - Min rate = 10Mbps (implementation : 9.6Mbps)
 - R_{AI} =5Mbps (implementation : 3.1Mbps)
 - R_{HAI} =50Mbps (implementation : 51.2Mbps)

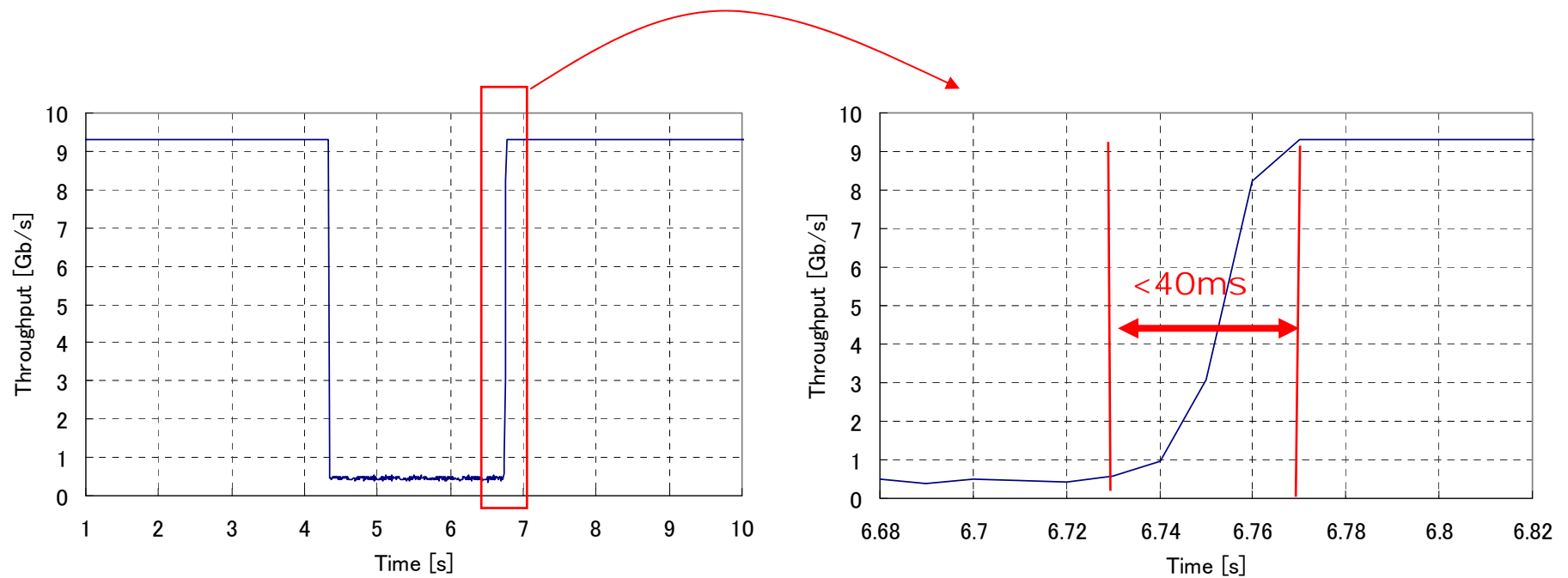
Experimental Setup: Block Diagrams



(*)ExpEther consortium
<http://www.expether.org/>

Throughput Measurements (0.5G)

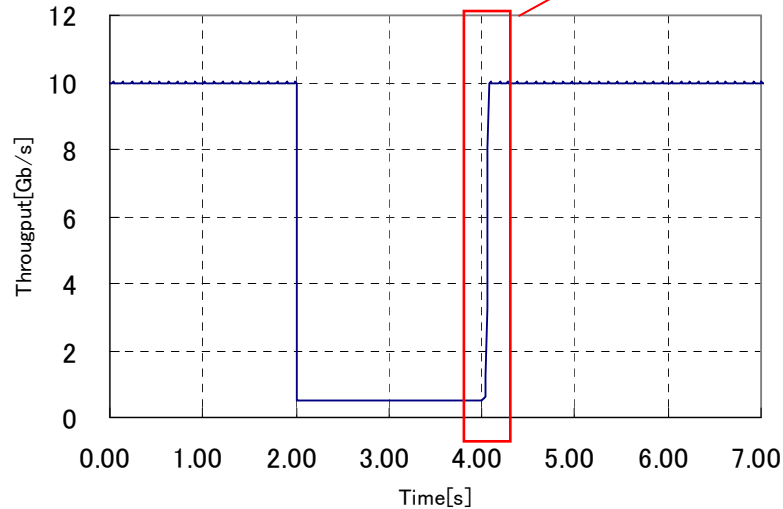
- Shaper configuration
 - 10Gbps-->0.5Gbps-->10Gbps



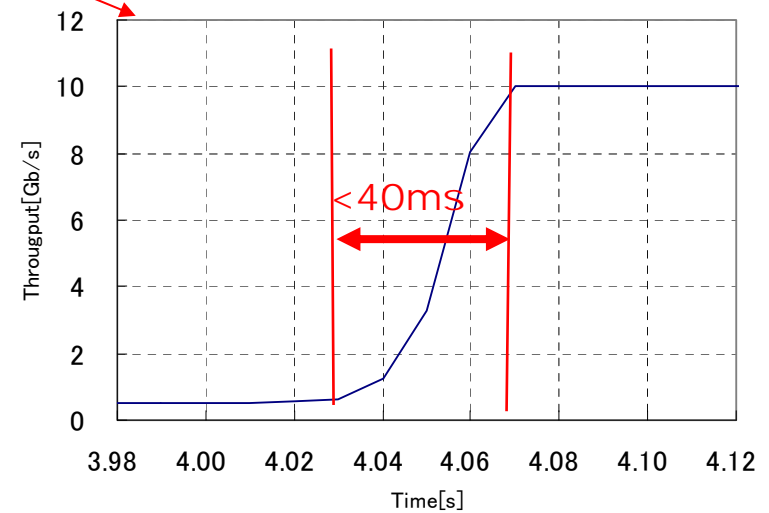
Measurement point: S/W Output Port
Resolution:10ms

NS-2 Simulation (0.5G)

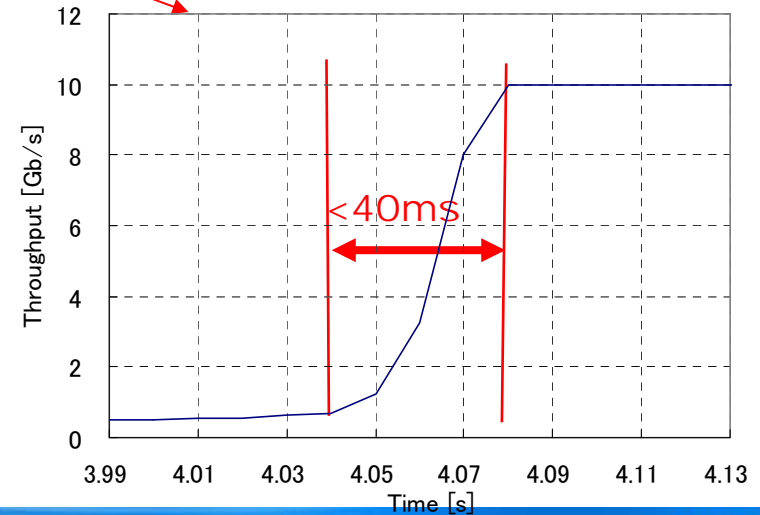
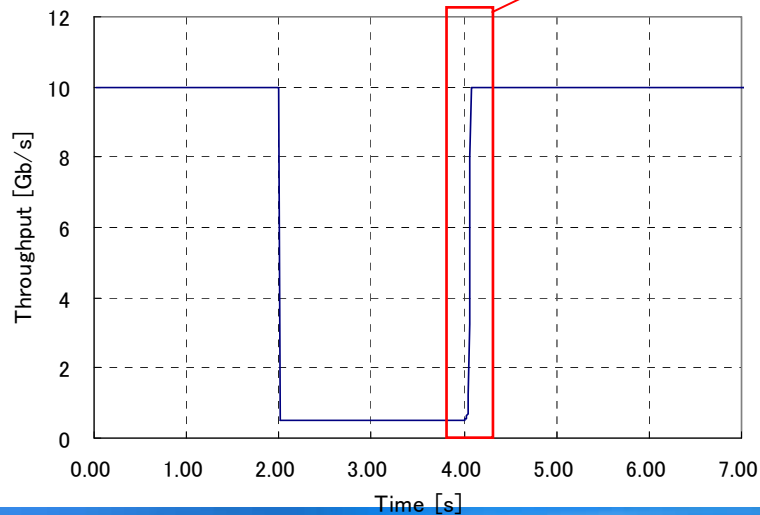
Original QCN



Measurement point: S/W Output Port
Resolution: 10ms



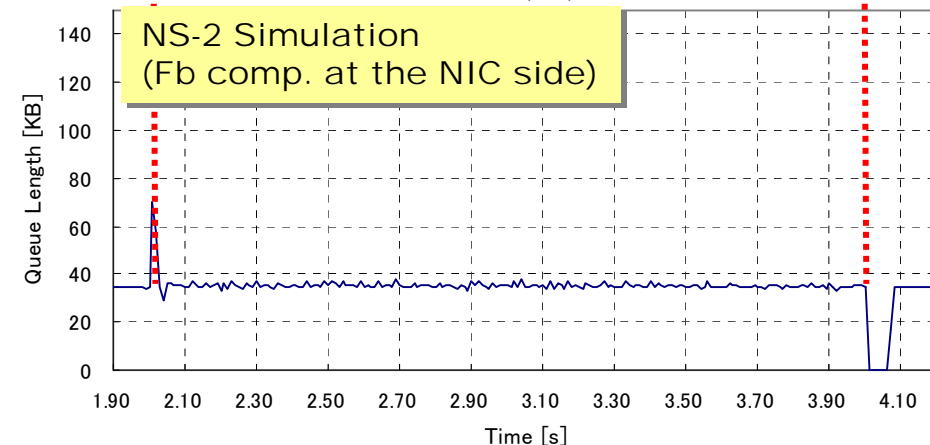
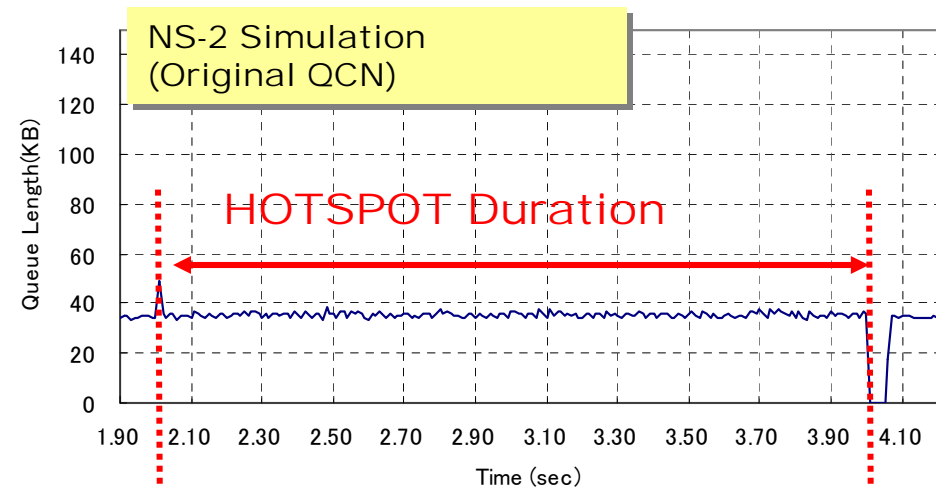
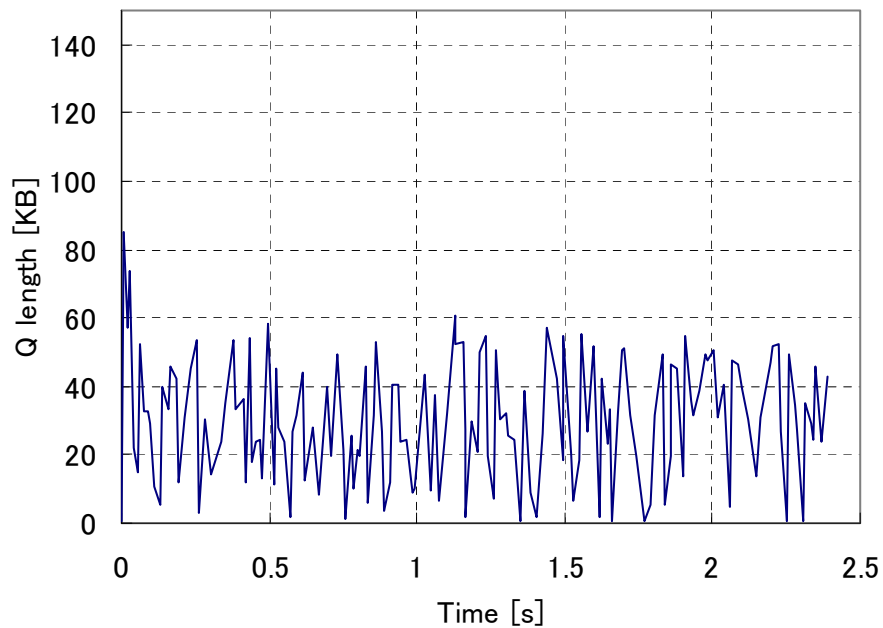
Fb computation at the NIC side



Queue Length (0.5G)

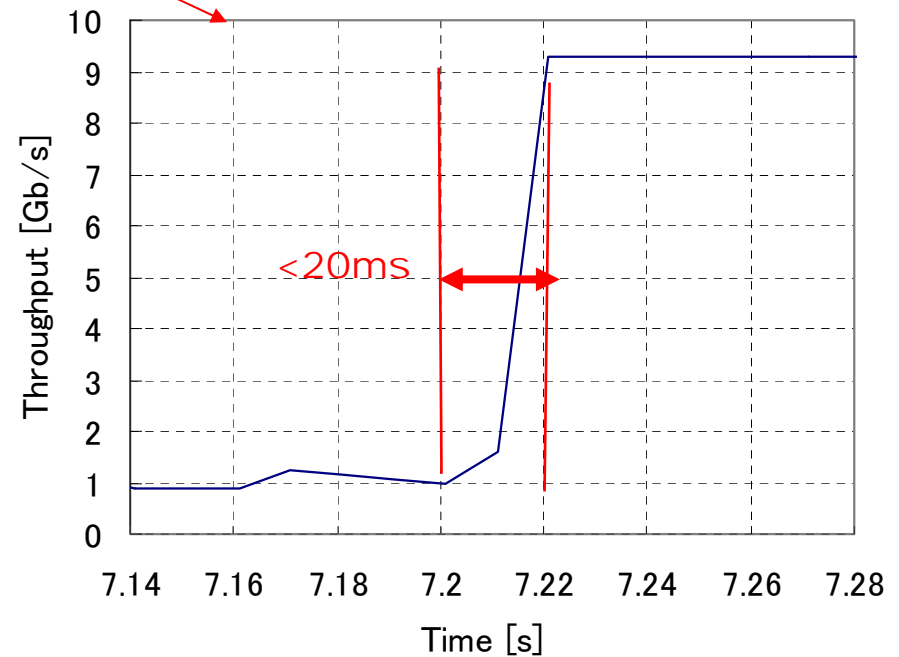
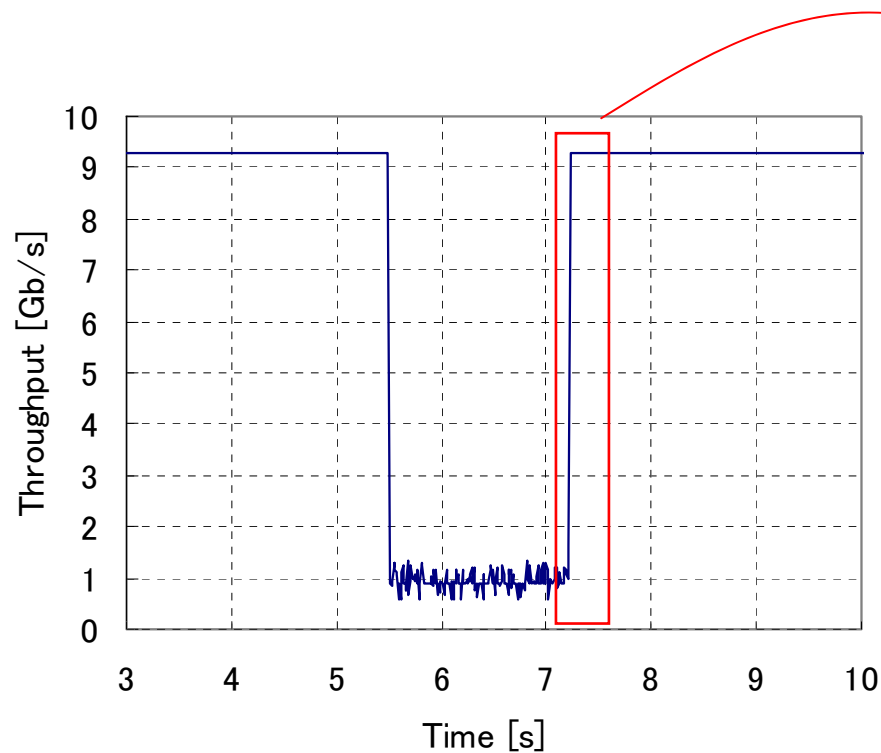
- Queue length is monitored by capturing Congestion messages from a switch
 - Note that the switch we used transmits Qlen info only when $Qlen > 0$
→ Qlen data is available only in the hotspot duration

Experiment



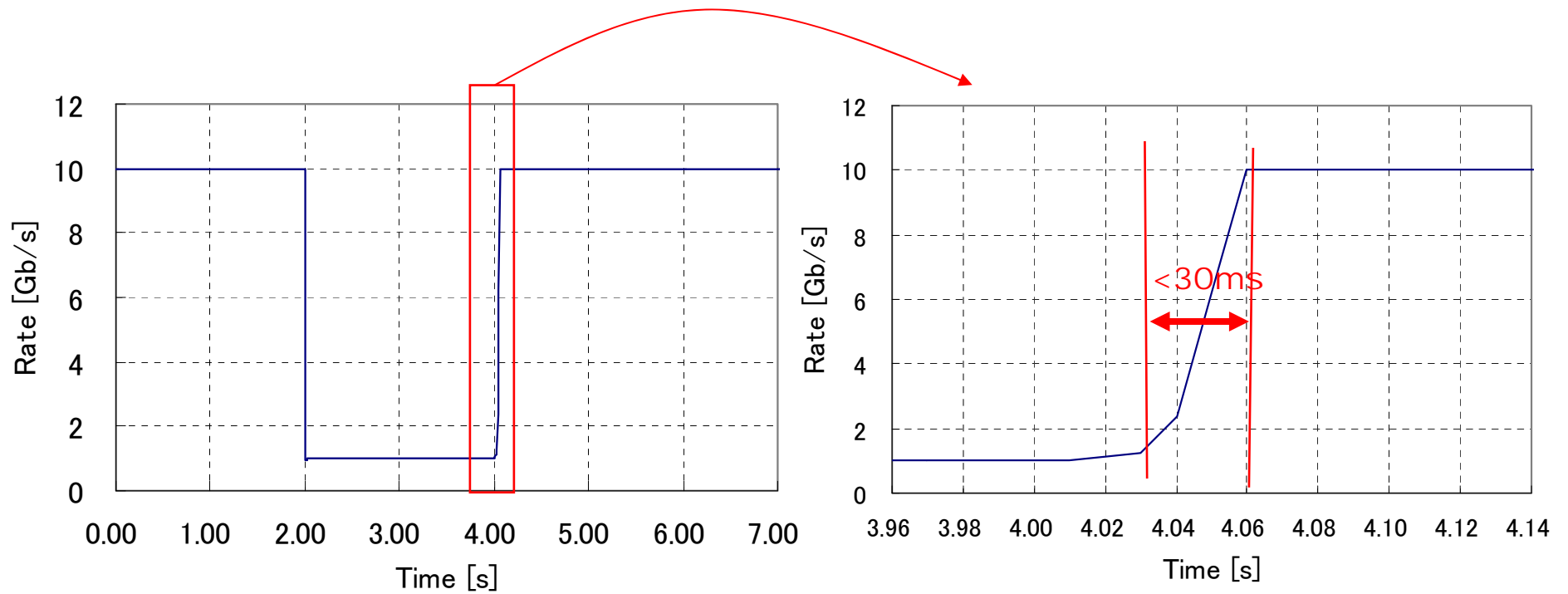
Throughput Measurements (1G)

- Shaper configuration
 - 10Gbps-->1Gbps-->10Gbps



Measurement point: S/W Output Port
Resolution: 10ms

NS-2 Simulation (1G)

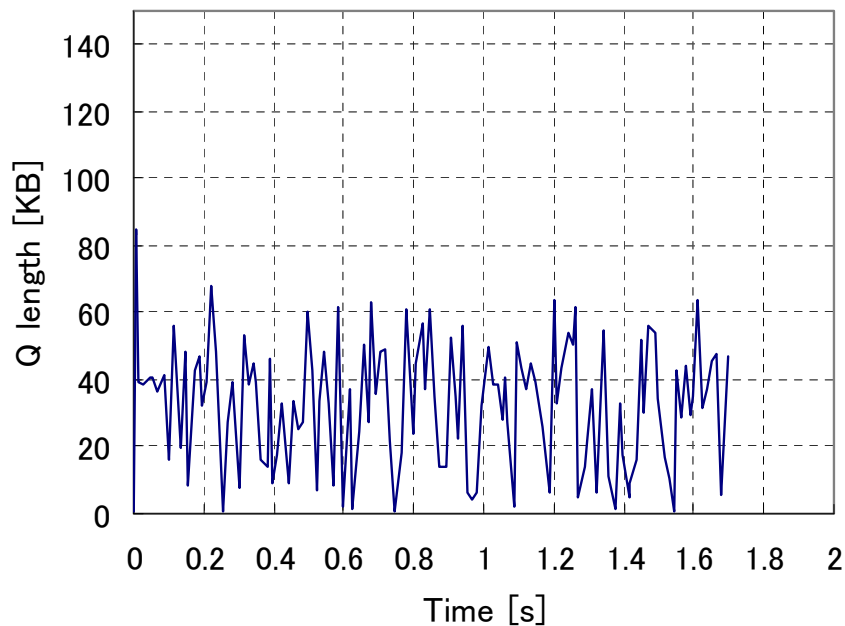


Measurement point: S/W Output Port
Resolution: 10ms

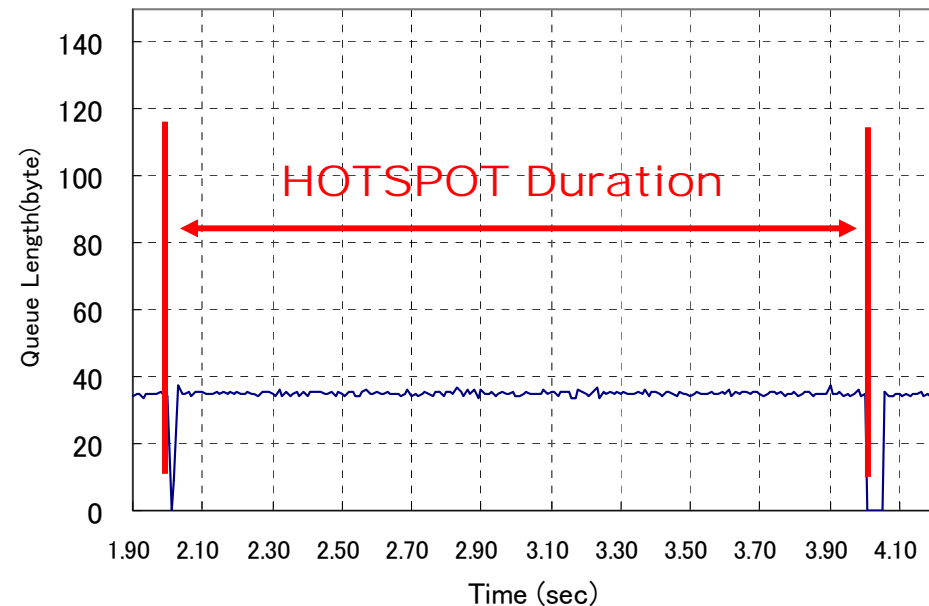
Queue Length (1G)

- Queue length is monitored by capturing congestion messages
 - Note that Fulcrum switch transmits Qlen info only when $Qlen > 0$
→ Qlen data is available only in the hotspot duration

Experiment



NS-2 Simulation



Conclusion

- **Successful QCN operation at 10G**
- **Good match of throughput data with ns-2 simulation**
 - **Approx. 40 ms (20 ms) for recovering the throughput from 0.5G (1G) HOTSPOT**
- **Queue length is wiggly**
 - **For further analysis, NIC burstiness should be examined by tuning parameters.**

Empowered by Innovation

NEC