BCN Calibration Simulation Results
Pause
Sample Window
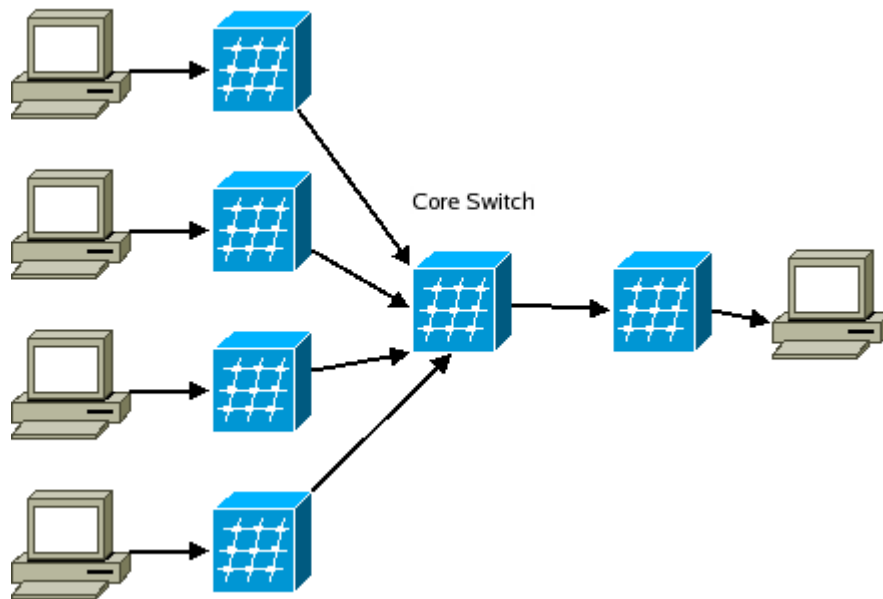
Zhi-Hern Loh
zloh@fulcrummicro.com
10/19/2006

FULCRUM microsystems

Workload

- **Traffic Type: 100% UDP (or raw Ethernet) Traffic**

- **Destination: EP0-EP3 sending to EP4**

- **Frame Size Distribution: 1500 byte fixed**

- **Arrival Distribution: Bernoulli temporal distribution**

- **Offered load at endpoint = 50%**

FULCRUM
microsystems

Topology

Core Switch

- Link capacity 10Gbps
- Core switch egress port buffer size infinite
- Rate limiter queue buffer size 150KB
- Switch latency (1 us)
- Link length (not modelled, 0 latency)
- Endpoint response time (not modelled, 0 latency)

FULCRUM
m i c r o s y s t e m s

BCN Parameters

- **Qeq 375 * 64 byte pages**

- **Frame Sampling 150KB +- 5KB (random jitter)**

- **W = 2**

- **Gi = 5.3 x 10^-1**

- **Gd = 2.6 x 10^-4**

- **Ru = 1 Mbps**

FULCRUM
microsystems

Simulations

- **Parameters**

    - No limits on max queue length at congestion point

    - Data sets

        - Pause only

        - Pause + BCN

        - Pause + BCN(0,0)

        - Sample Window Sweep for normal BCN

FULCRUM
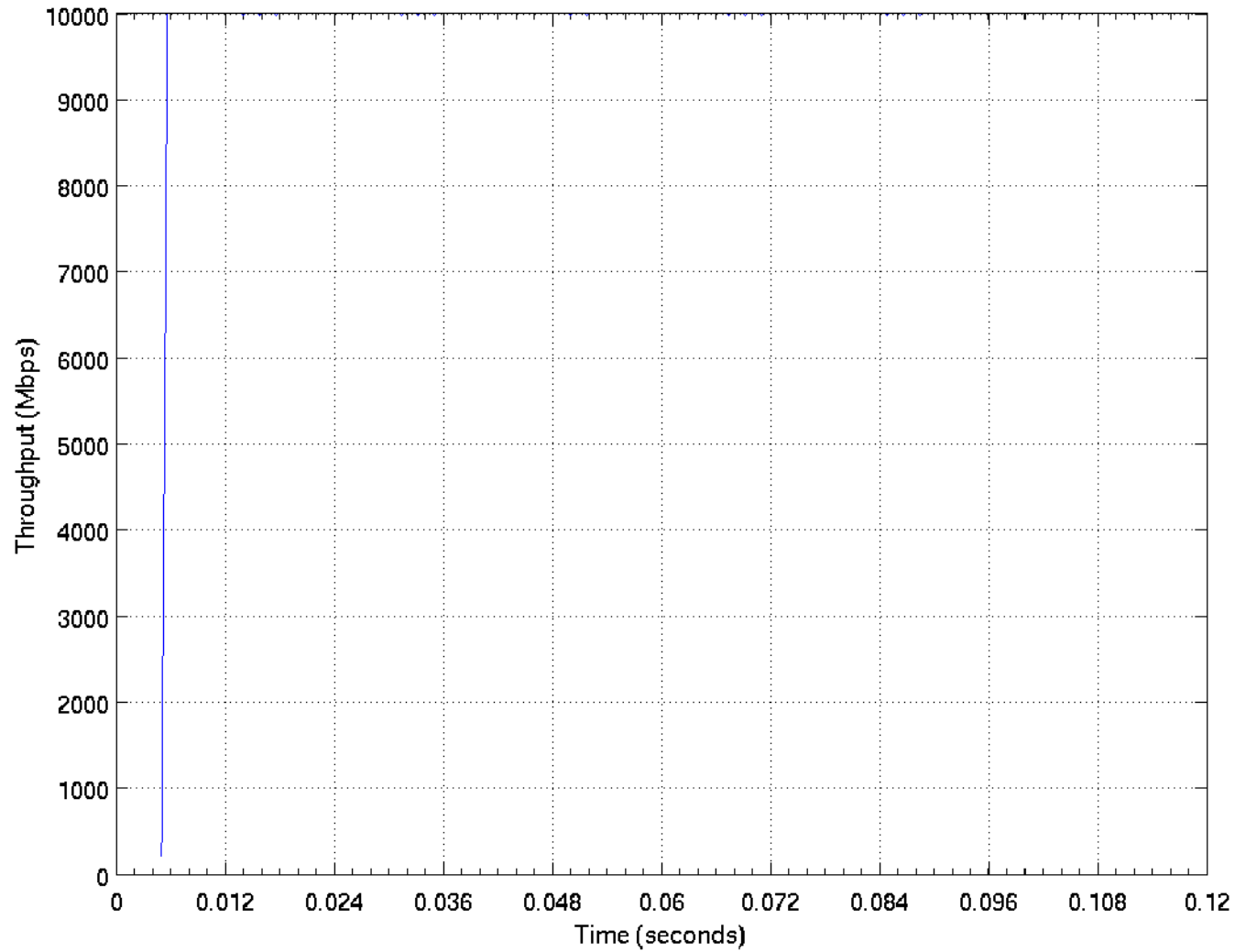microsystems

Pause

- **Global pause**

  - Pause all inputs if memory exceeds pause on watermarks

  - Unpause all inputs if memory drops below pause off watermark

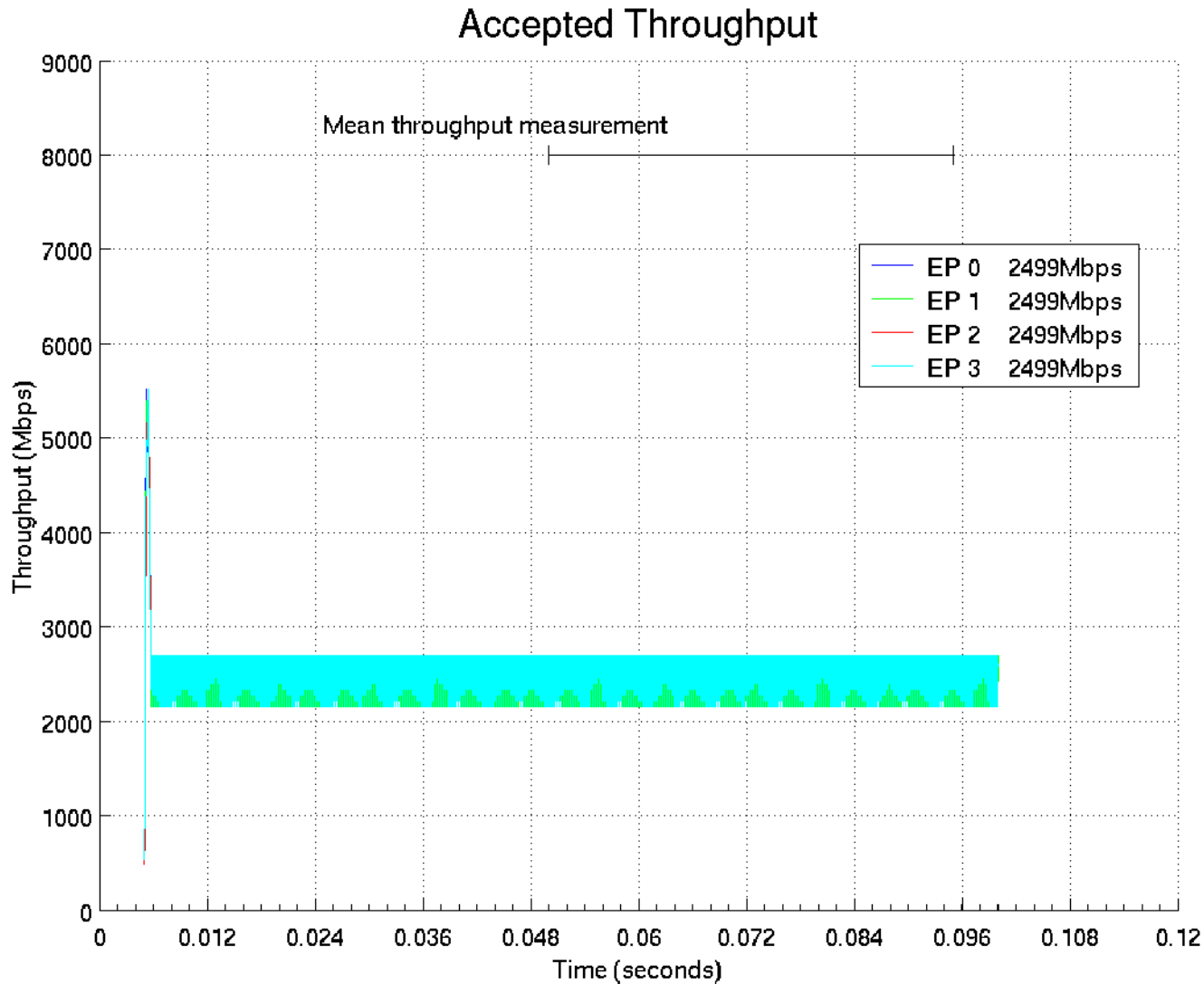 - Pause on 140 KB

  - Pause off 130 KB

Pause

Buffer Utilization at Core Switch

FULCRUM microsystems
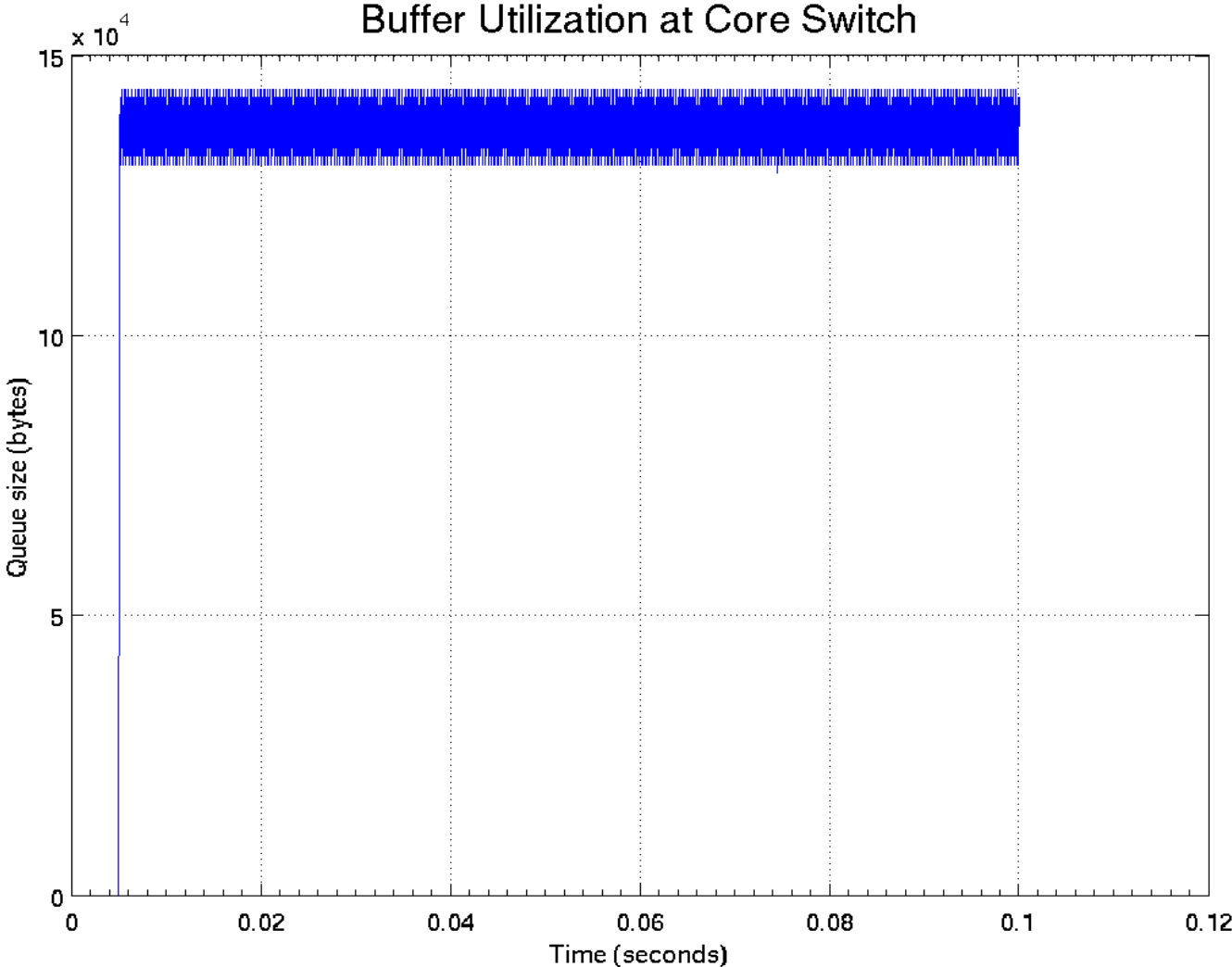
Pause Egress Throughput

Egress Throughput vs Time

Pause Accepted Throughput

Accepted Throughput

Pause Edge Switch Occupancy

Buffer Utilization at Edge Switches
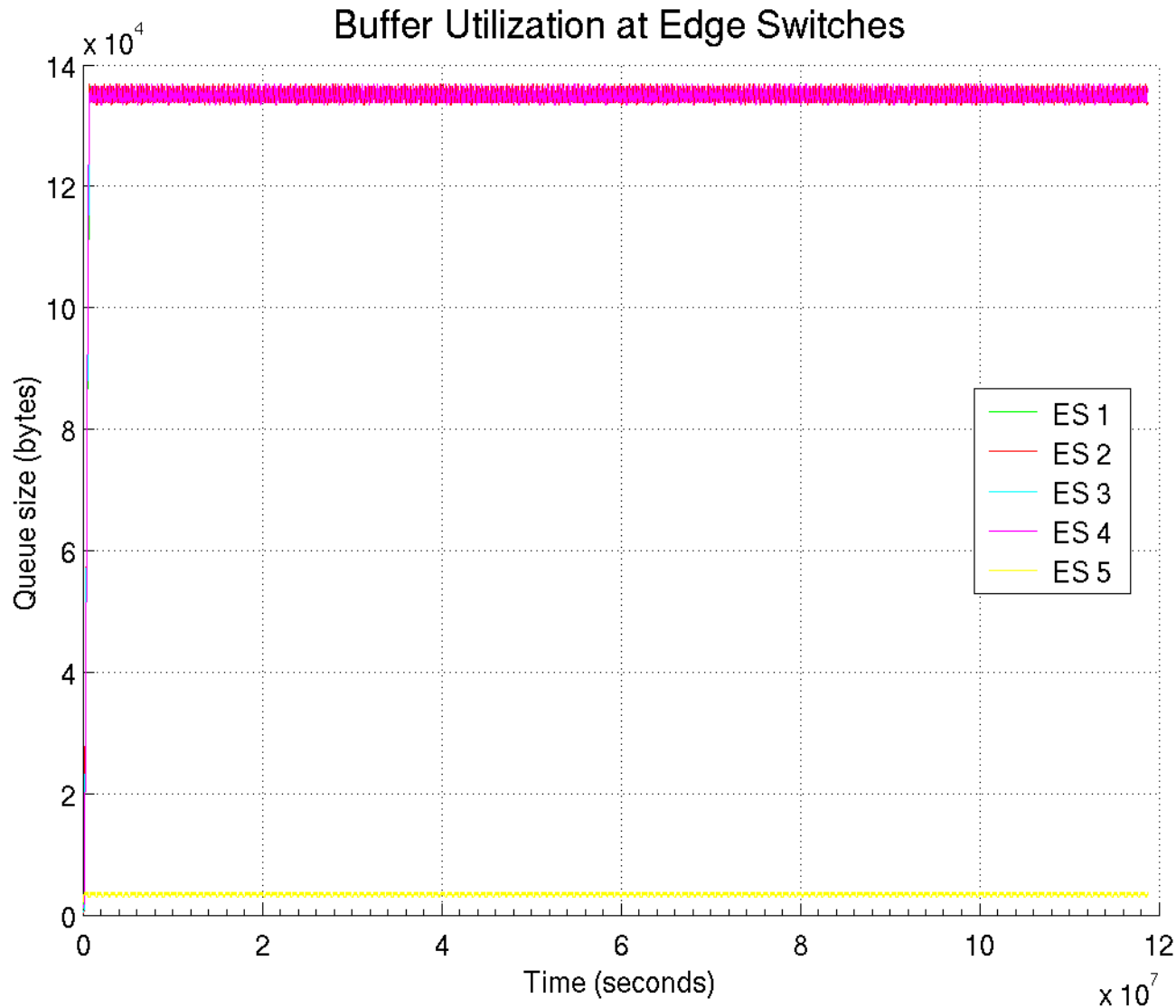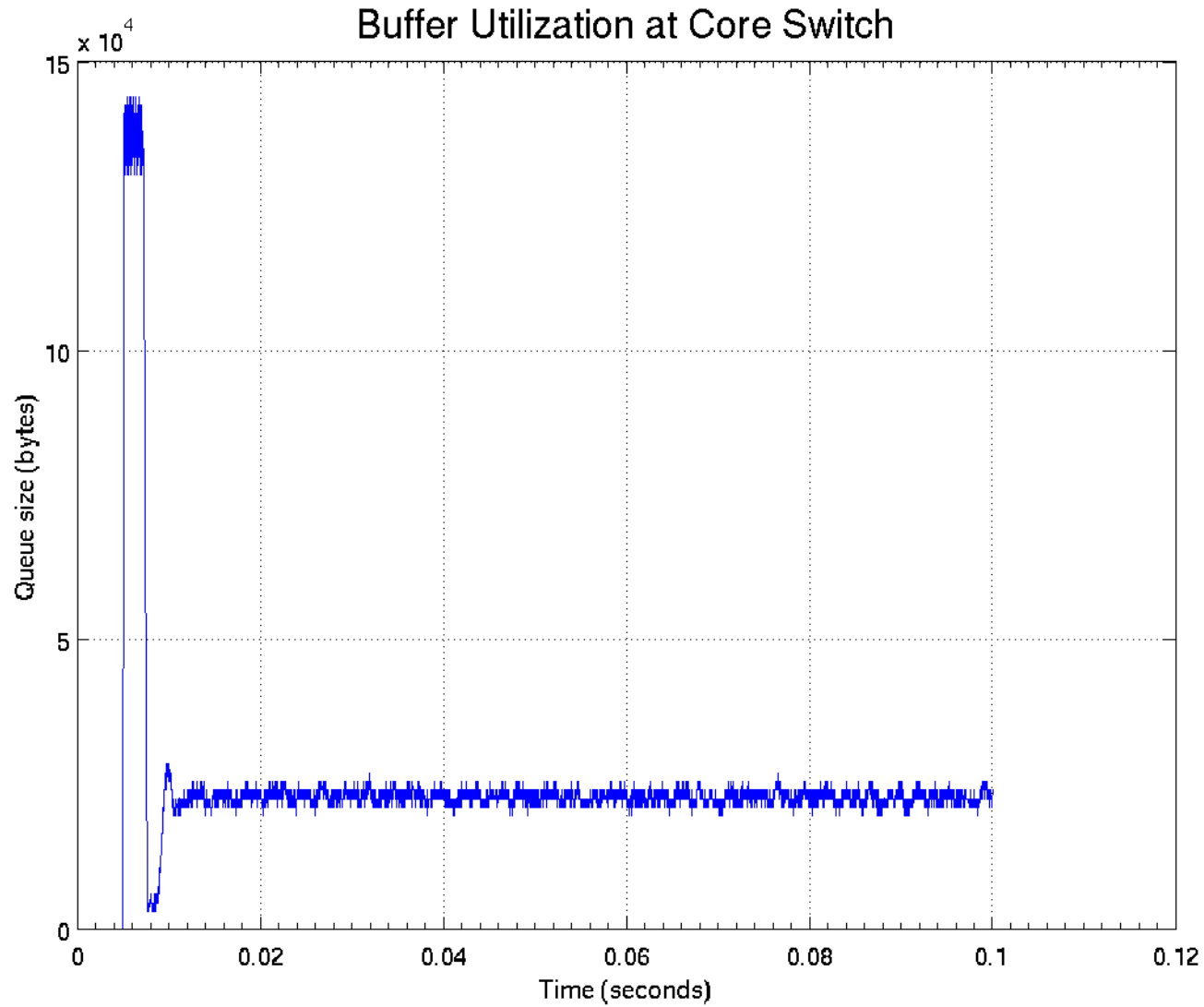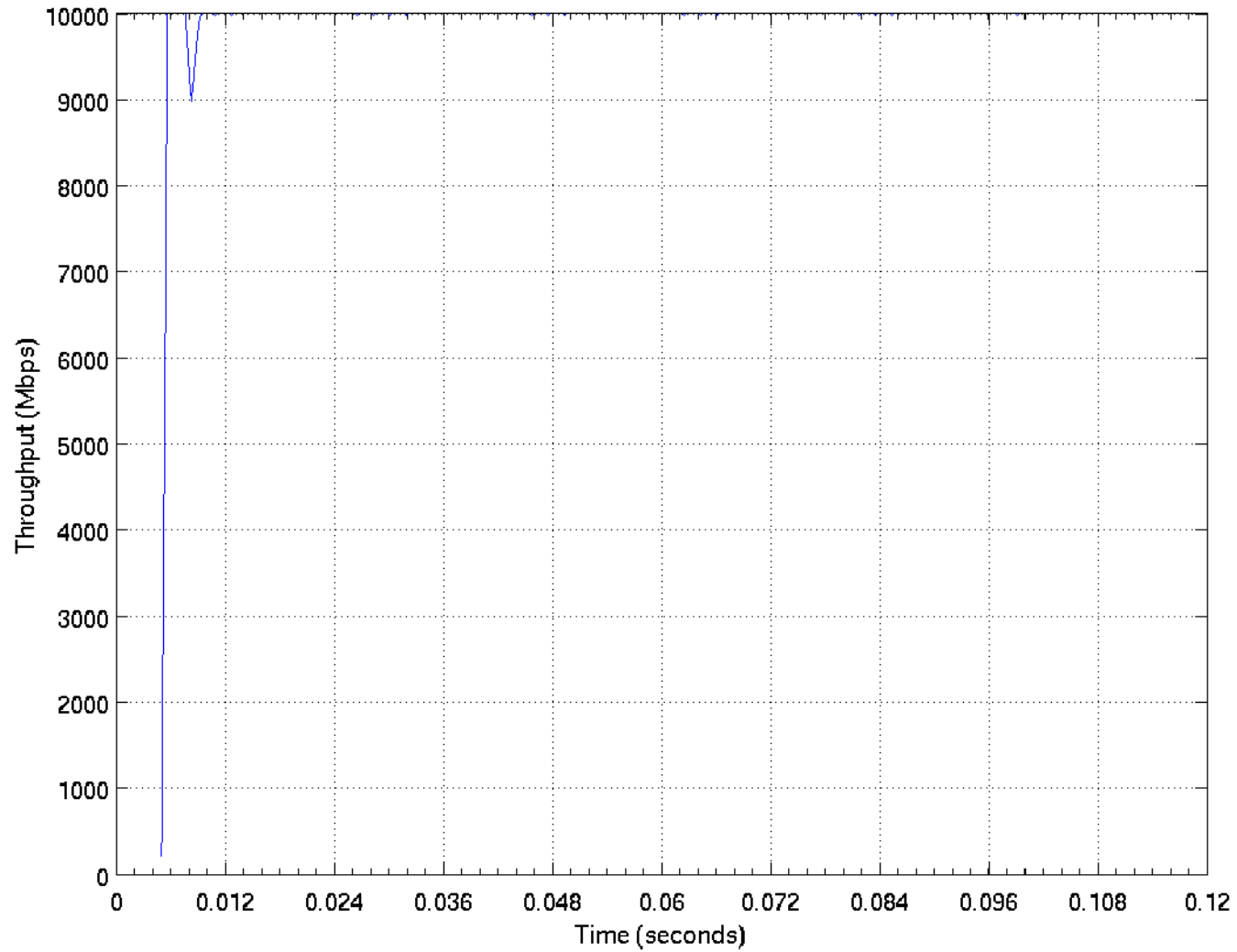
Pause + BCN

- **Rmin 1 Gbps**

- **Drift disabled**

- **BCN CP detection not disabled for edge switches ES1-4**

  - Have some BCN messages from ES1-4 during onset of

    congestion.

FULCRUM
microsystems

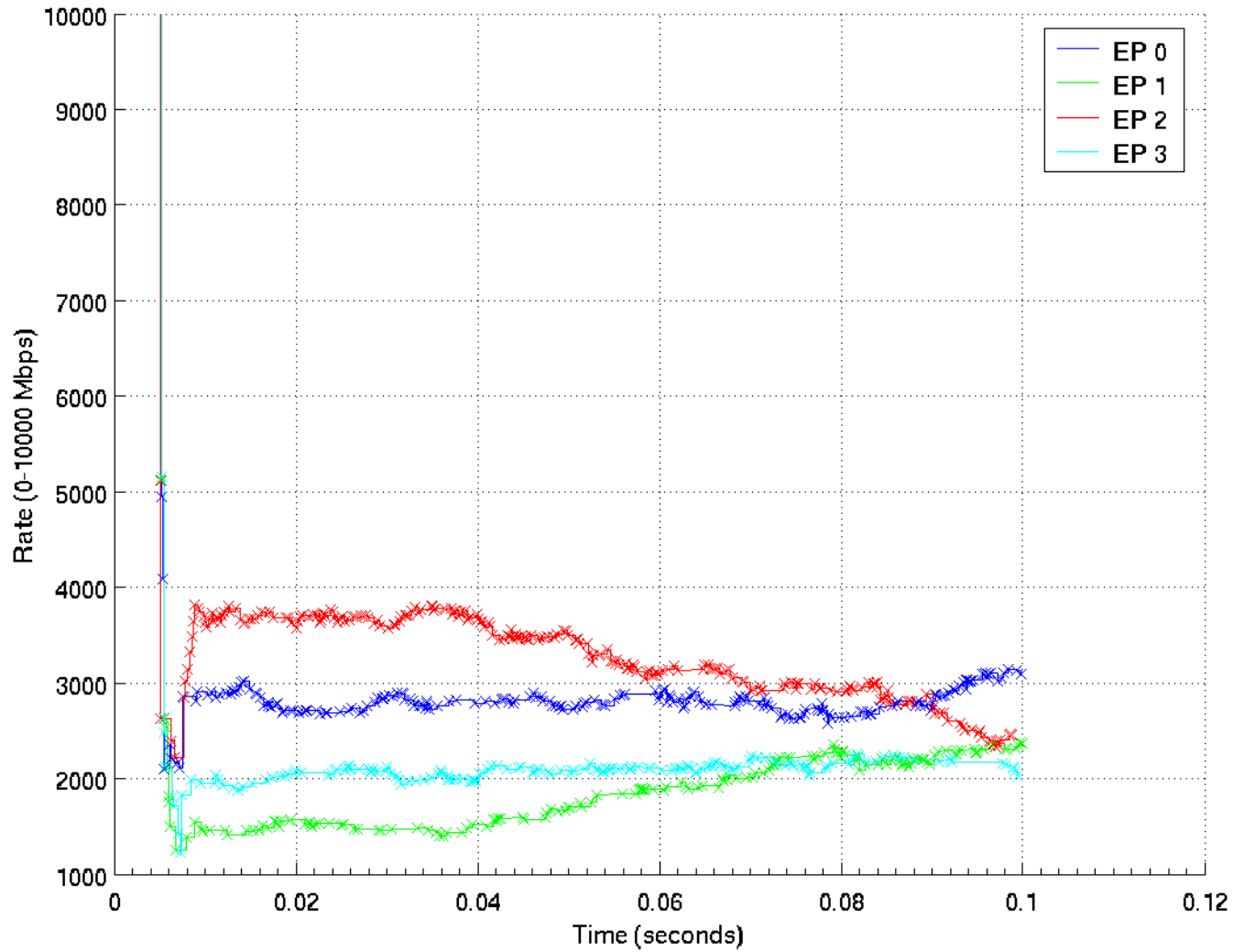# Pause + BCN

Buffer Utilization at Core Switch

Pause+BCN

Egress Throughput vs Time

Pause+BCN

RLQ Rate vs Time

Pause+BCN

Accepted Throughput

Pause+BCN

Buffer Utilization at Edge Switches
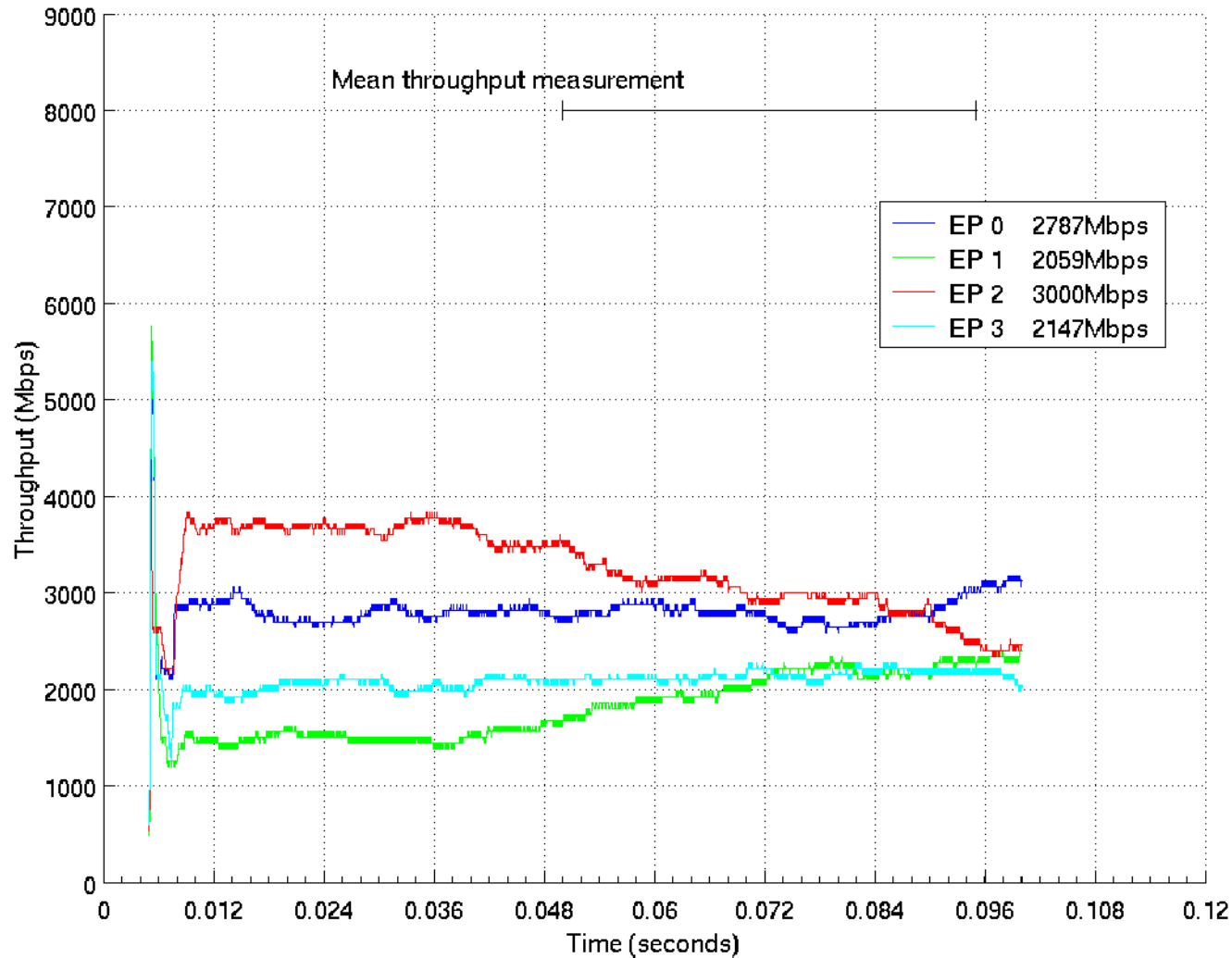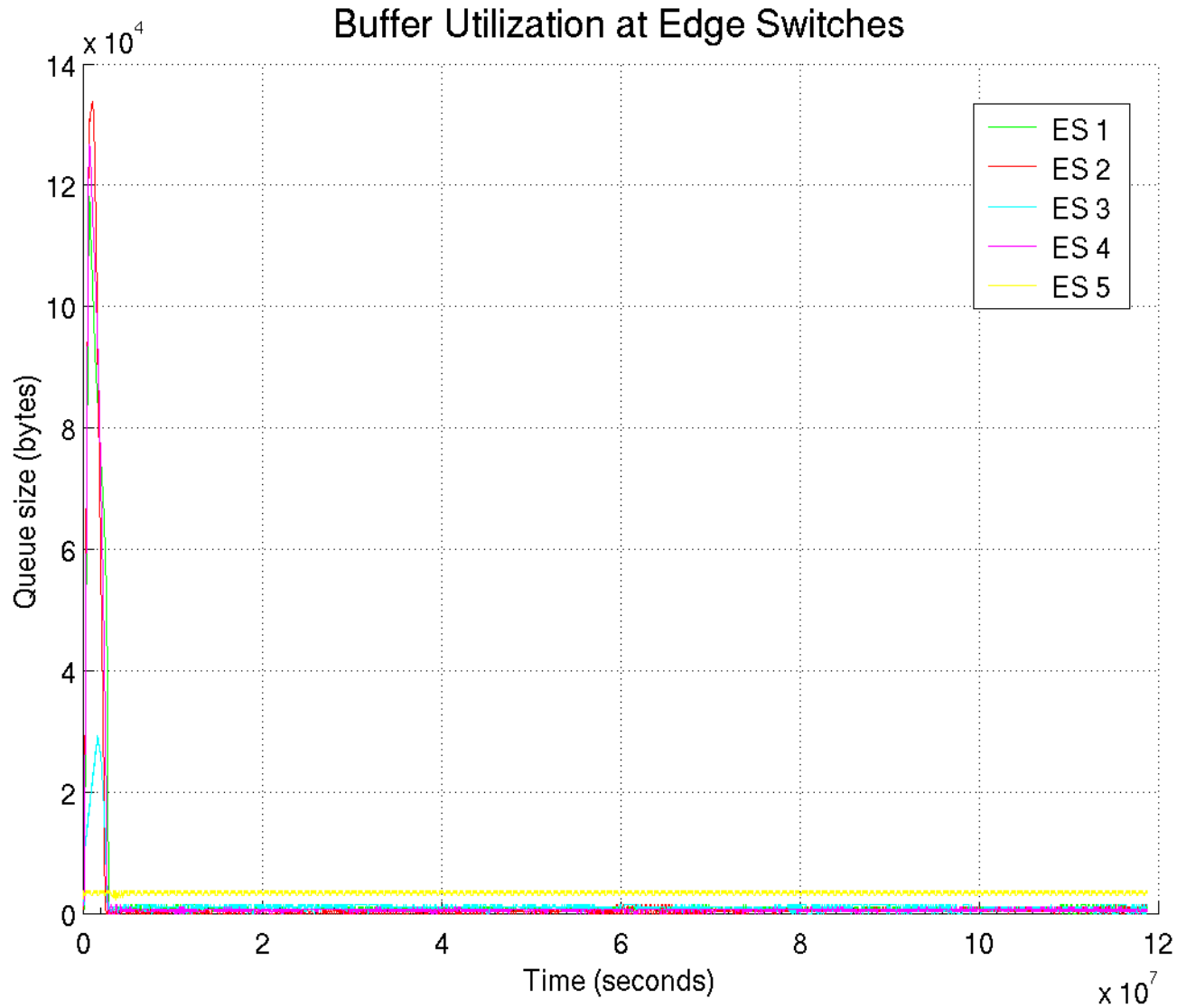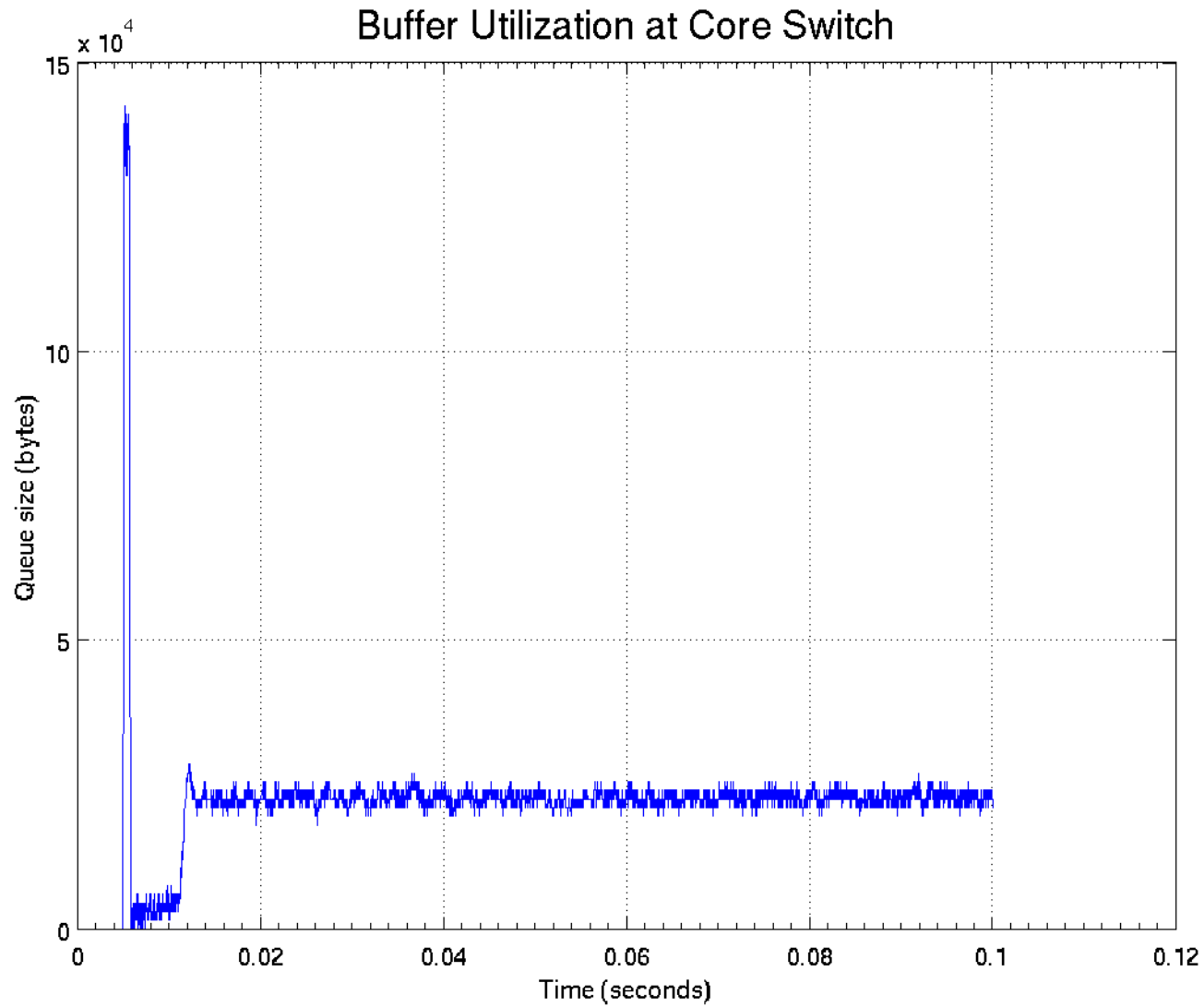
Pause + BCN

- **Differences from Bergamasco's results**

  - Edge Switches have BCN turned on

    - Congestion spike in congestion point queue has shorter

      duration due to stronger feedback.

    - Duration of spike

      – Approx 0.003 with BCN on at edge switches

      – Approx 0.007 with BCN off at edge switches

FULCRUM
microsystems

Pause + BCN(0,0)

- Qsc = 112.5KB

FULCRUM
microsystems

Pause + BCN(0,0)

Buffer Utilization at Core Switch

FULCRUM microsystems

Pause + BCN(0,0)

Egress Throughput vs Time

Pause + BCN(0,0)
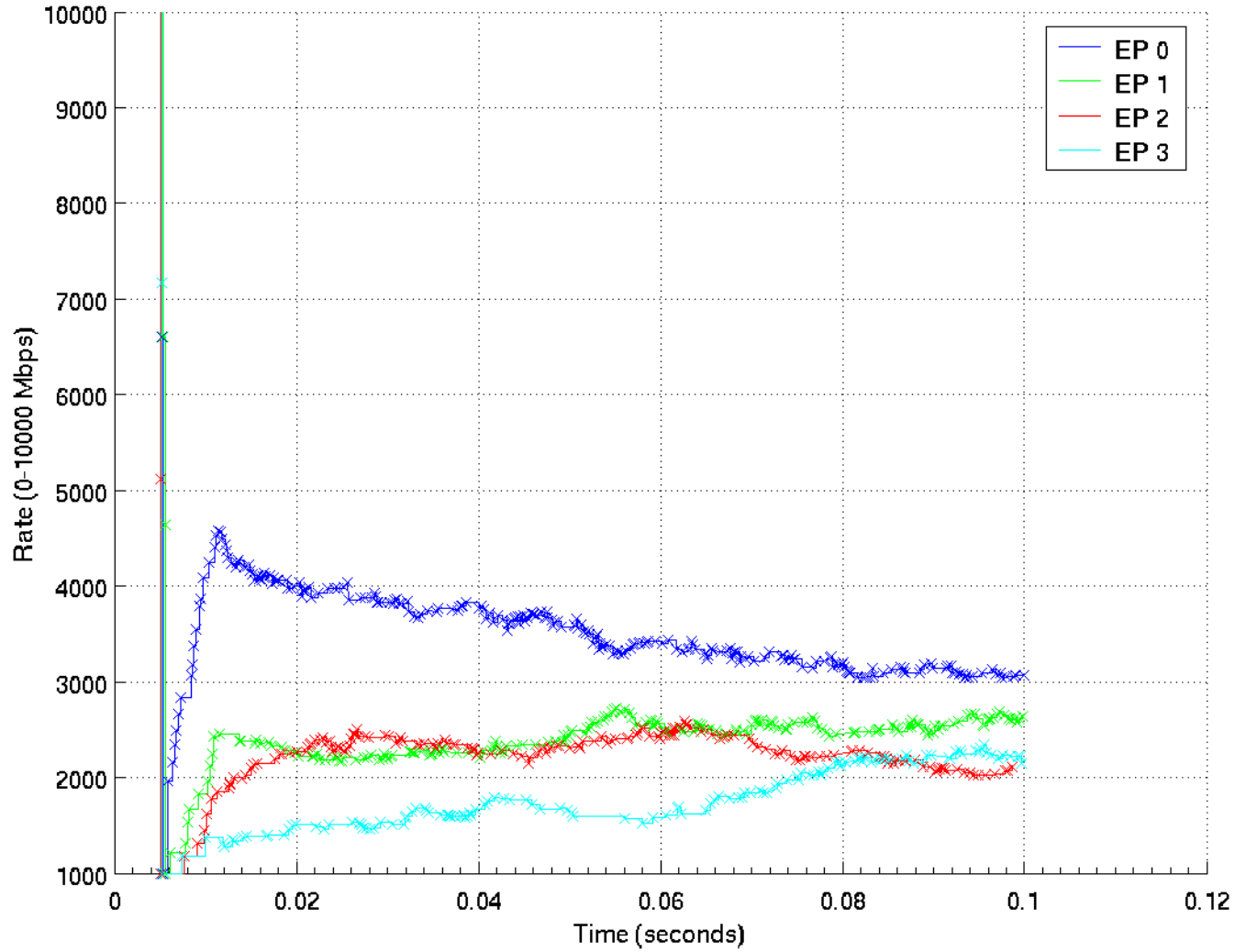
Pause + BCN(0,0)

Accepted Throughput

Legend:
- EP 0 — 3251Mbps
- EP 1 — 2534Mbps
- EP 2 — 2305Mbps
- EP 3 — 1904Mbps

Mean throughput measurement

X-axis: Time (seconds)
Y-axis: Throughput (Mbps)

22

Pause + BCN(0,0)

Buffer Utilization at Edge Switches
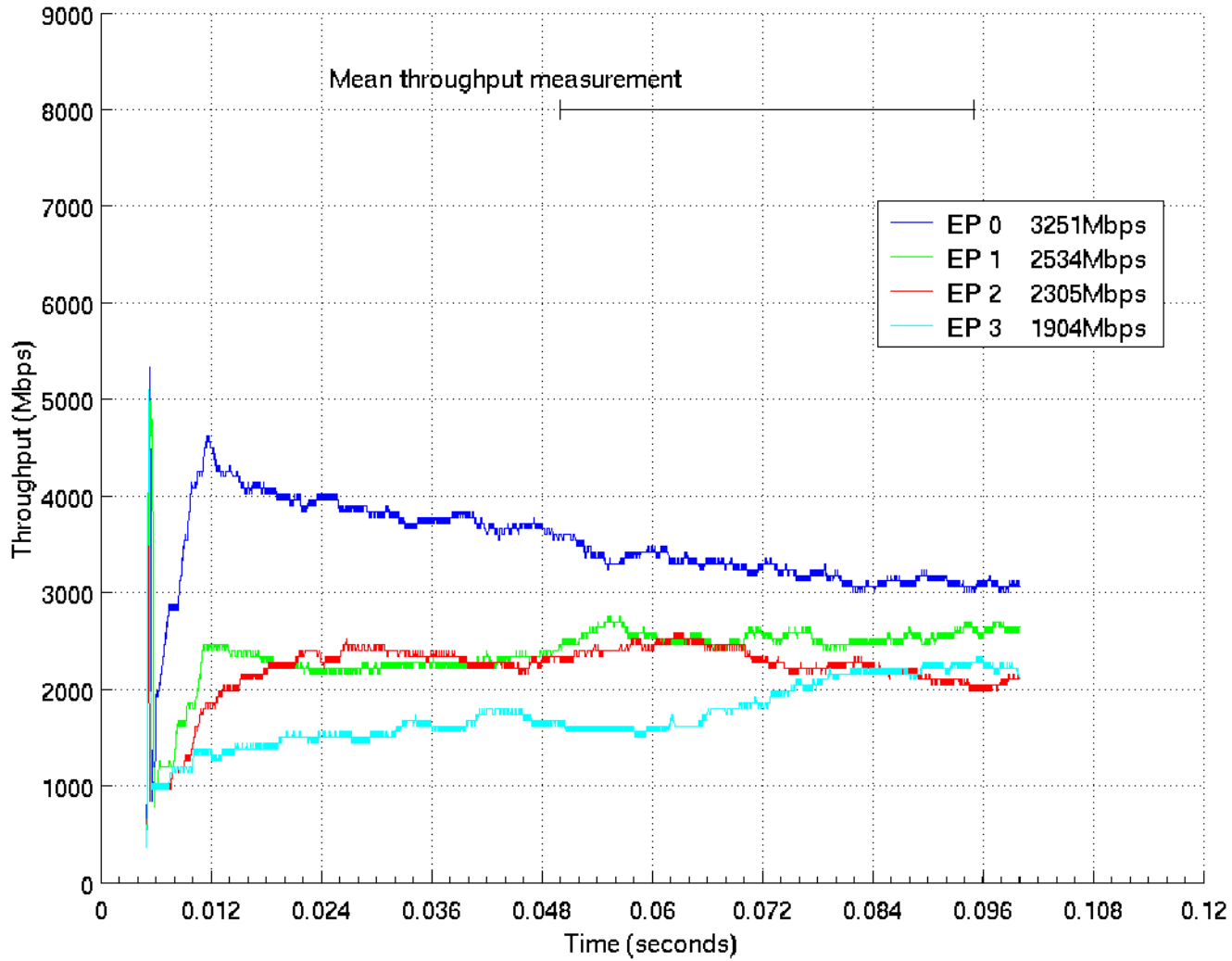
Sweep Over Sample Window Length

- **What happens when sample window is decreased?**

- **Simulation datasets**

  - Window sizes = 10, 20, 40, 60, 80, 100 KB

  - Random jitter +- 5KB for all window sizes

  - 25 runs per datapoint

- **Results**

  - Plot peak queue size vs sample window

  - Plot fairness index vs sample window

    - Fairness measured over $t = 20 - 80$ ms

FULCRUM
microsystems

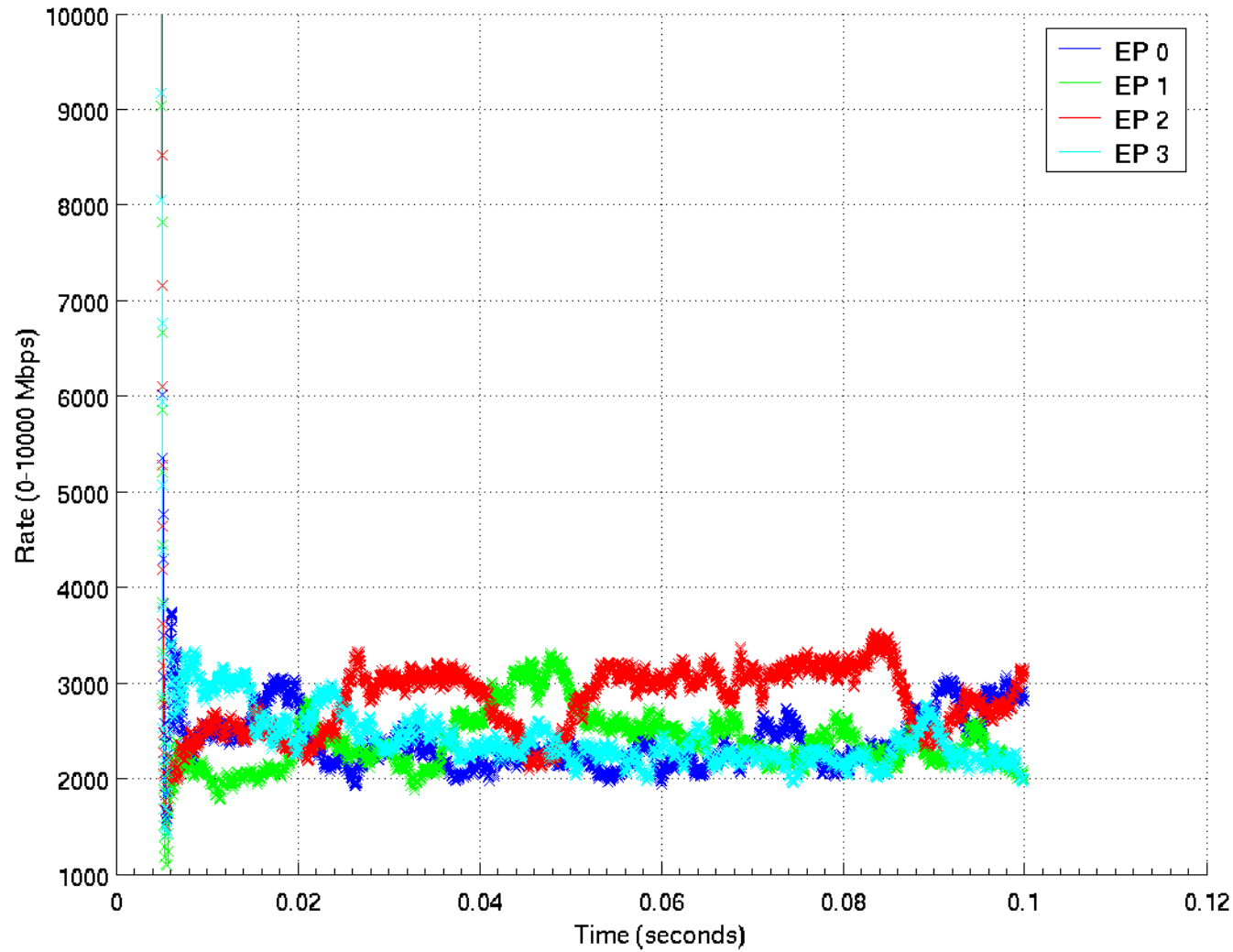# Window size 10KB

Buffer Utilization at Core Switch
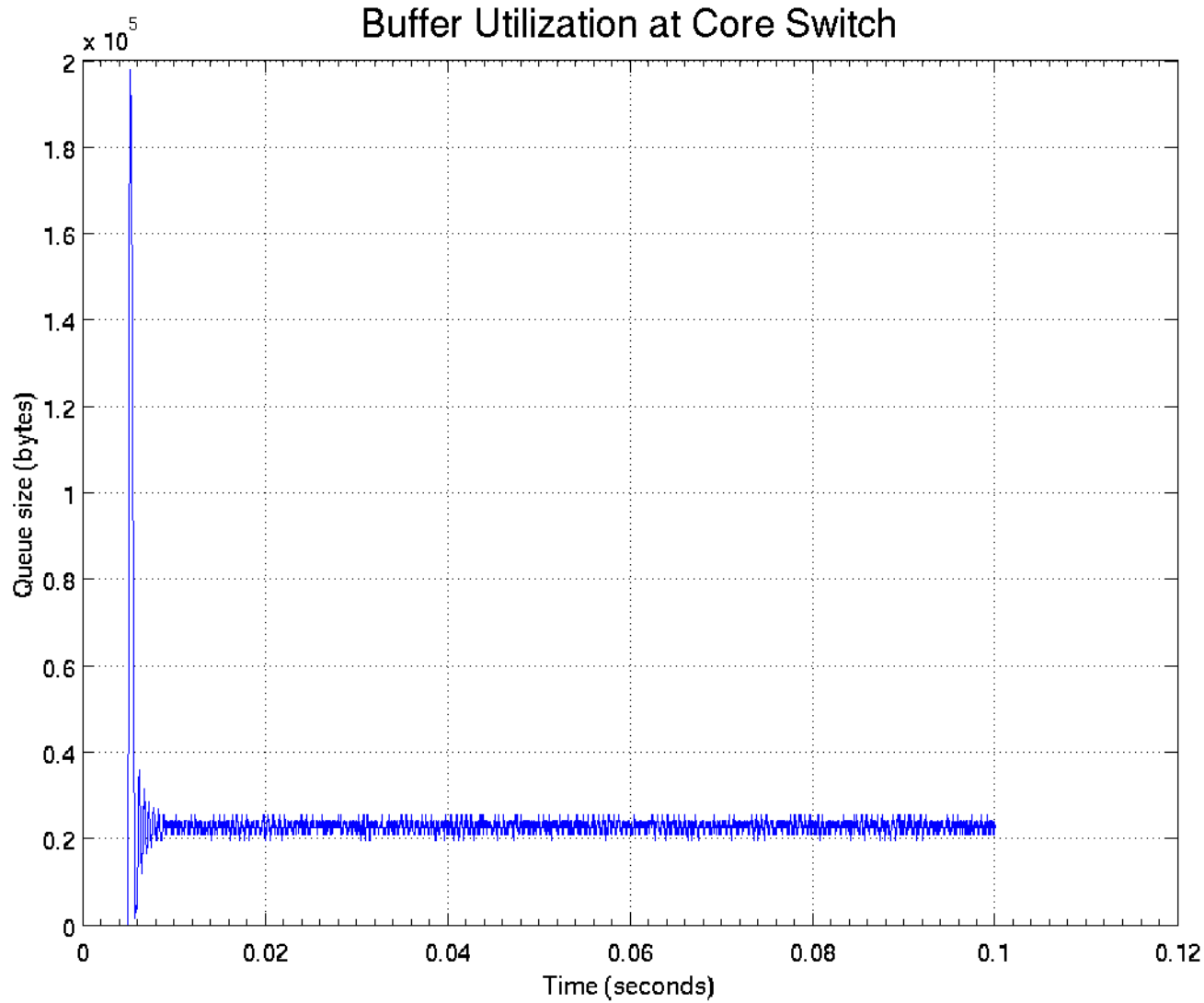
Window size 10KB

Egress Throughput vs Time
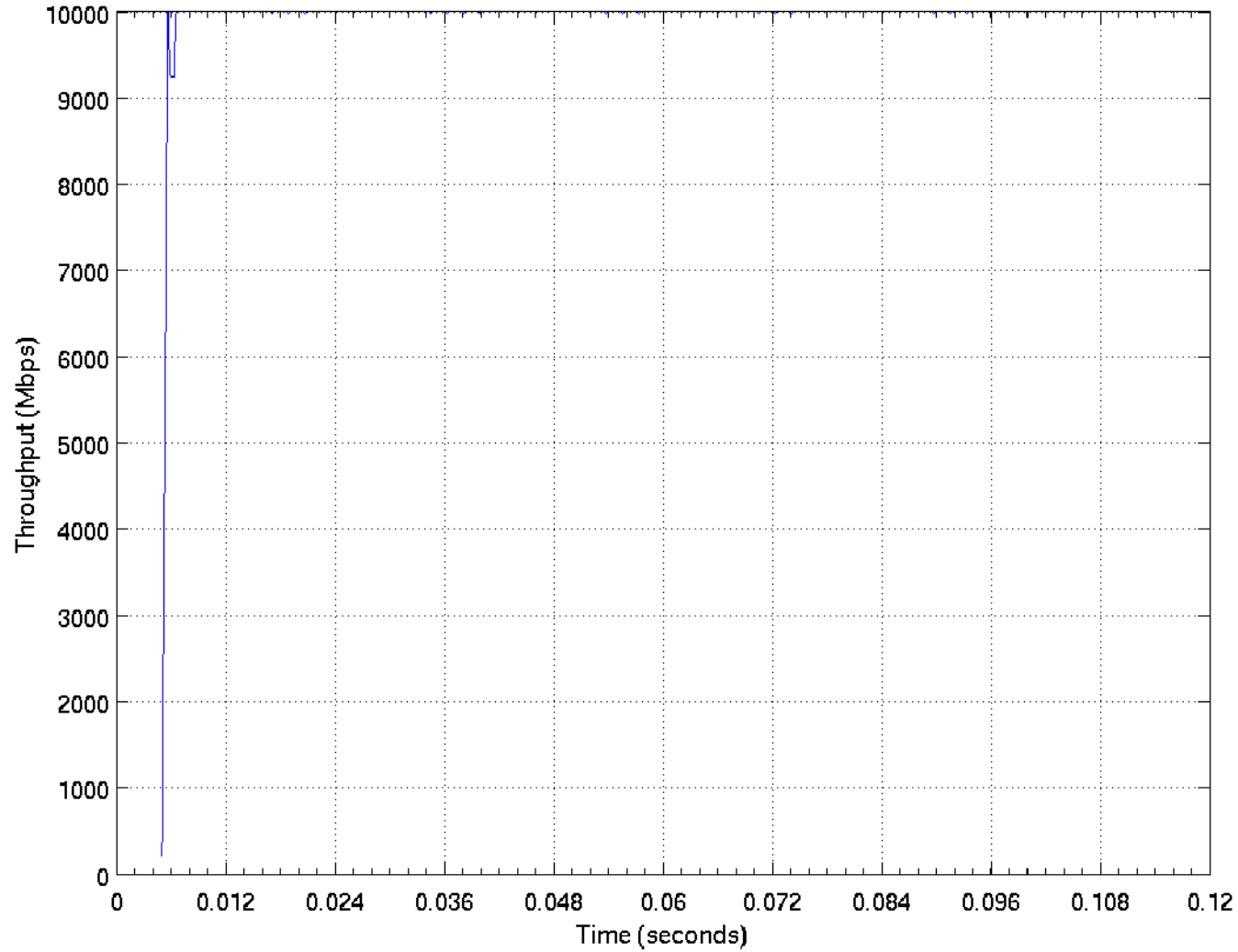
Window size 10KB

RLQ Rate vs Time

Window size 10KB
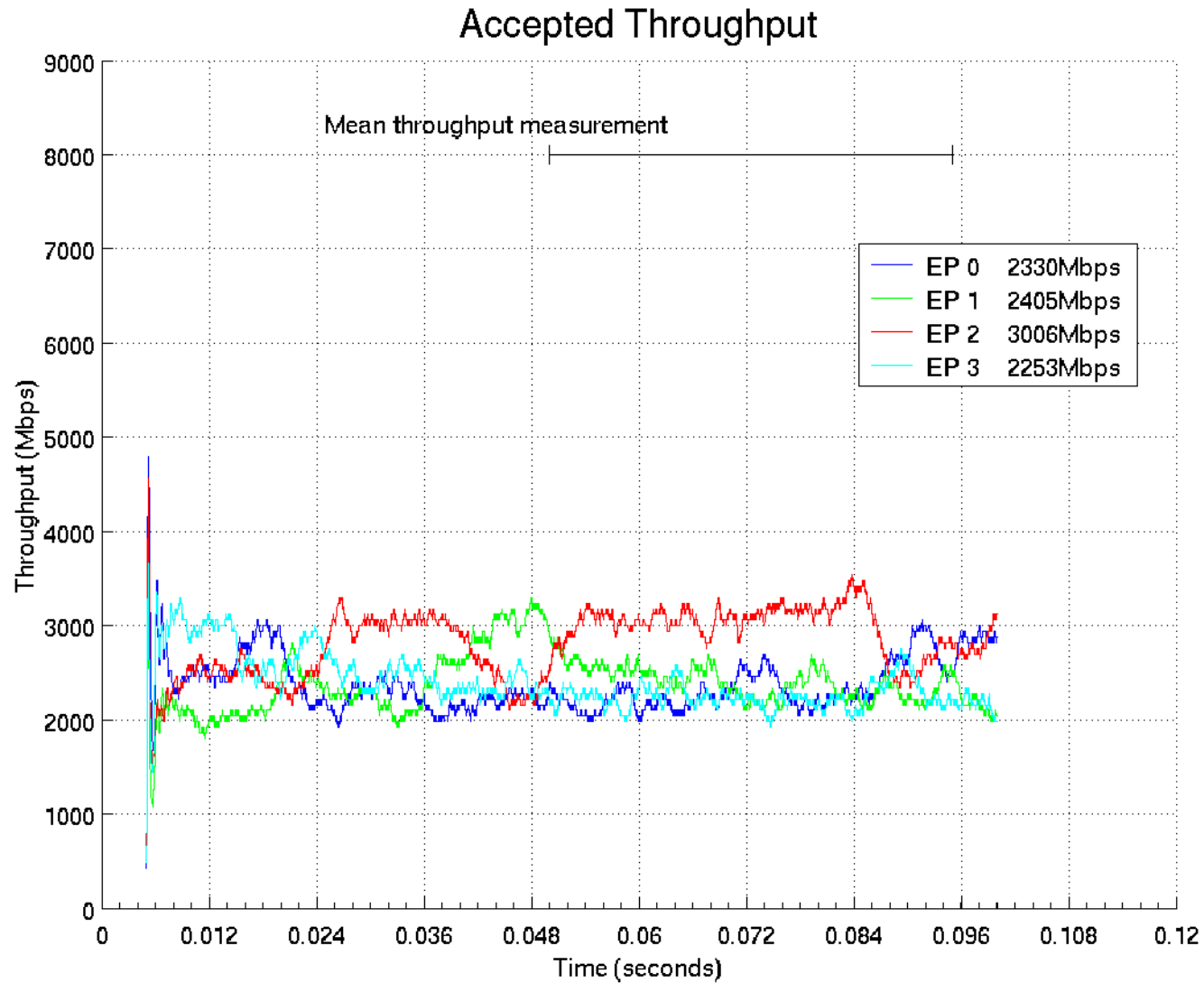
Averaged Peak Queue Size
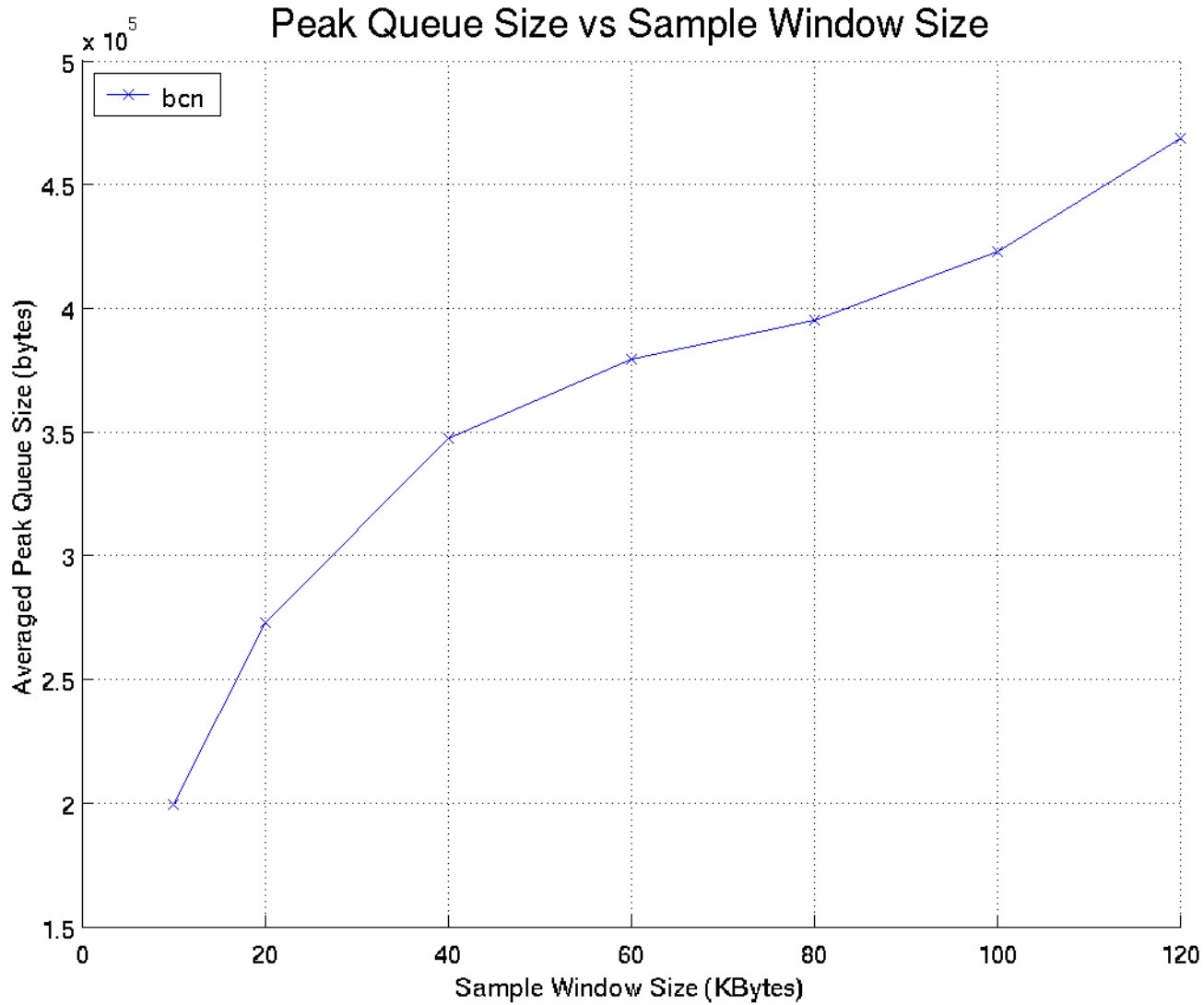

Peak Queue Size vs Sample Window Size
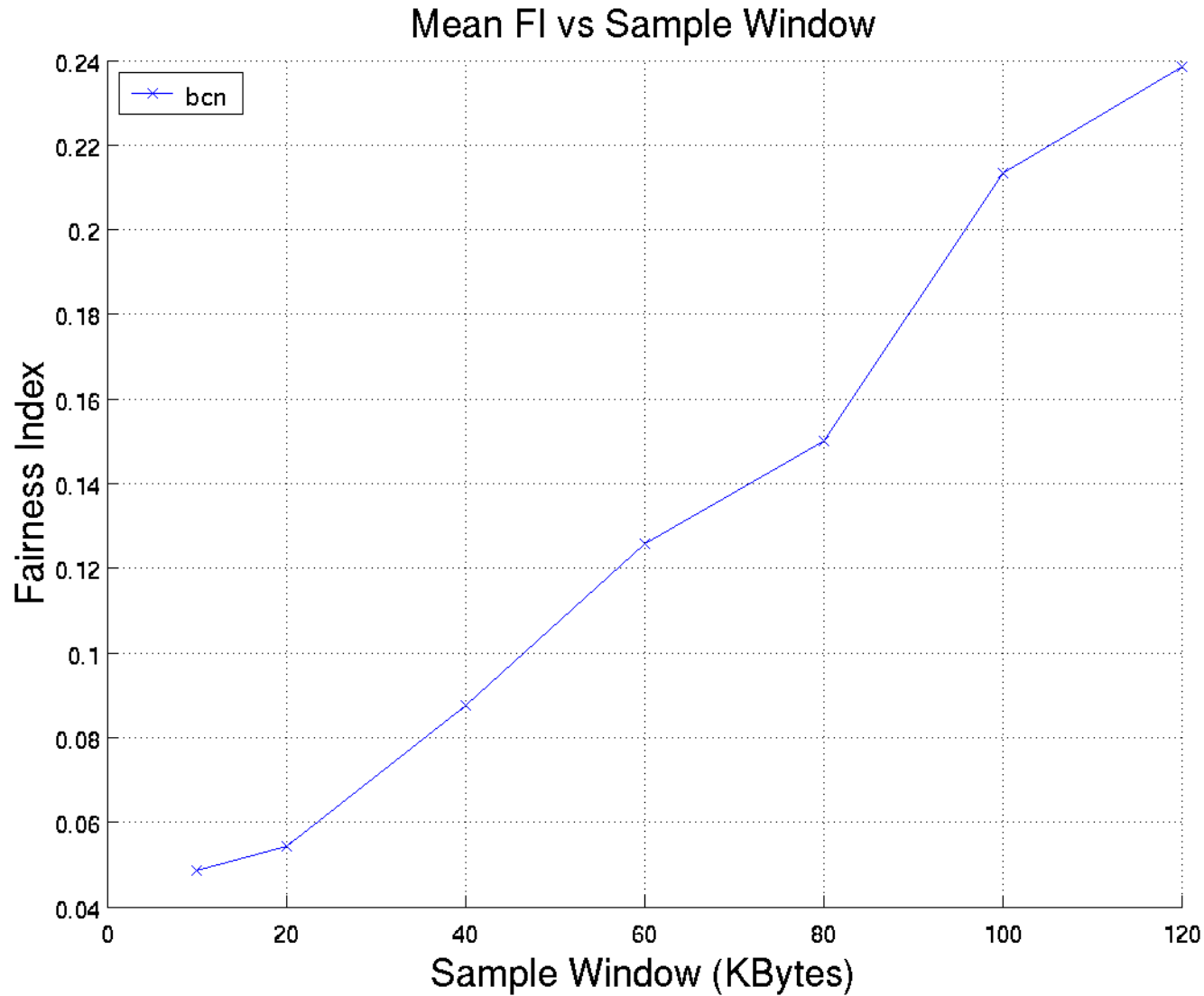
Fairness Index

Standard Deviation in Fairness Index