



# CN-SIM: A Baseline Simulation Scenario

**Davide Bergamasco ([davide@cisco.com](mailto:davide@cisco.com))**

**IEEE 802.1 Interim Meeting**

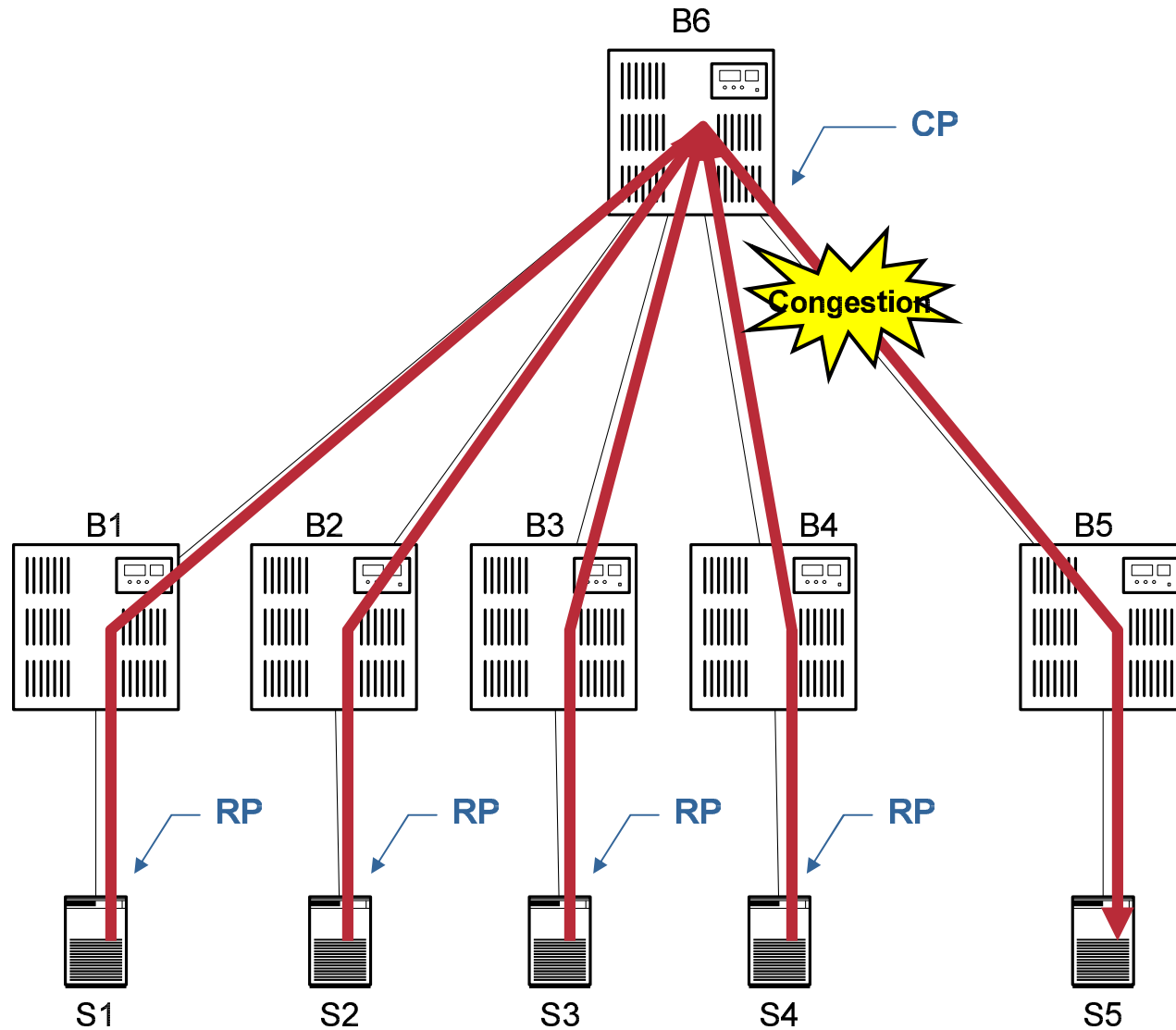
**York, UK**

**September 28<sup>st</sup>, 2006**

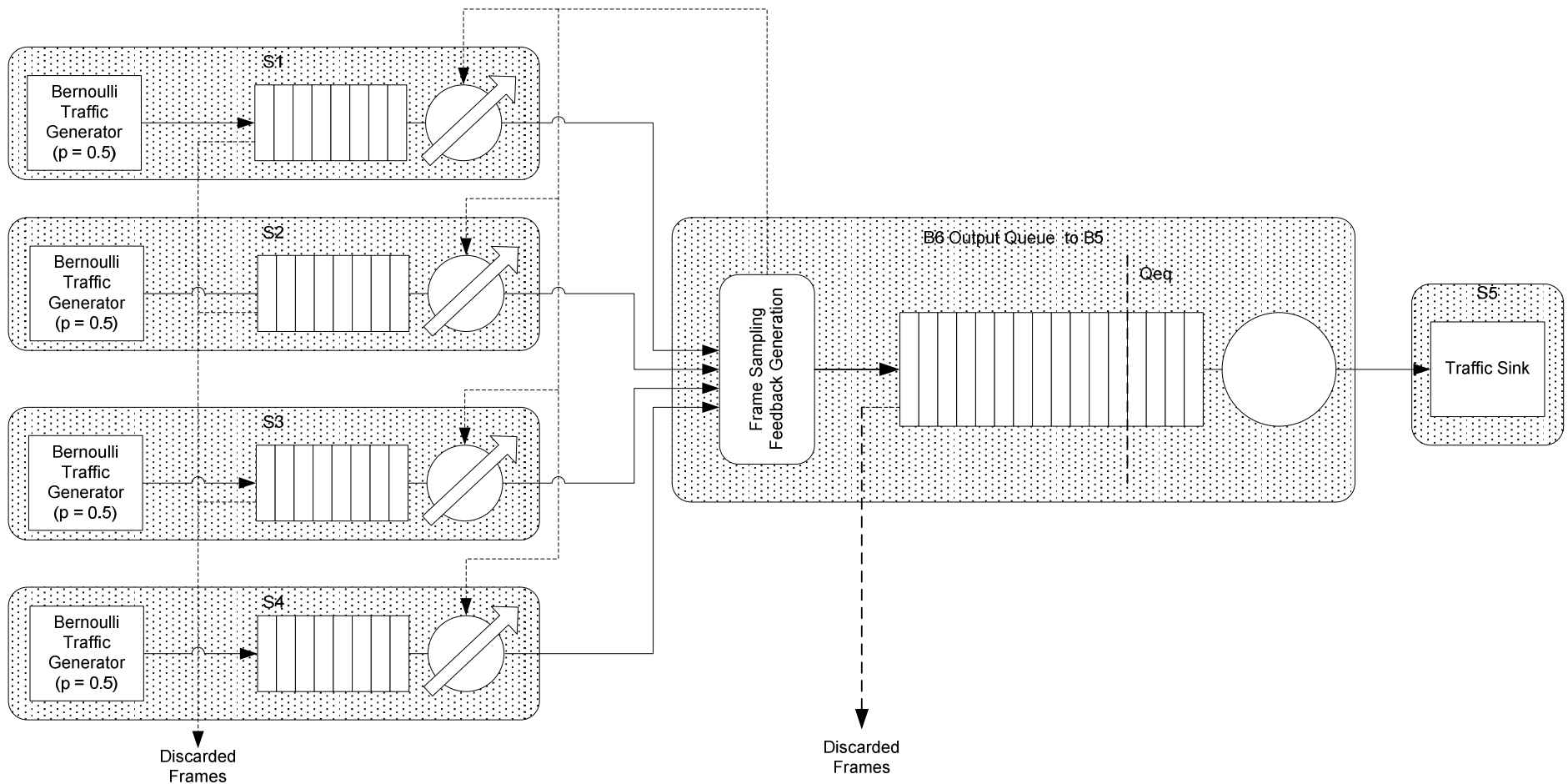
# Motivation

- **So far we have defined a set of common**
  - **Topologies**
  - **Traffic Patterns**
  - **Metrics**
  - **Bridge Model**
- **To ensure comparability of results, we also need to make sure our models and simulation tools are properly calibrated**
- **The **baseline simulation scenario** should allow us to achieve a reasonable alignment quickly and easily**

# Topology & Traffic Pattern



# Topology & Traffic Pattern



# Configuration, Parameters & Workload

- **Short Range, High-Speed Datacenter-like Network**

Link Capacity (C) = 10 Gbps

Buffer Size (B) = 150 KB (both CP and RP)

Switch latency = 1  $\mu$ s

Link Length = 100 m (.5  $\mu$ s propagation delay)

Station processing time = 2  $\mu$ s

Loop Latency = 8  $\mu$ s

- **BCN Control Loop Parameters**

Qeq = 375 64-byte pages (or 16 1500-byte frames or approx 24 KB )

S = 150 KB (frames are sampled on average every 150 KB received)

W = 2

Gi =  $5.3 \times 10^{-1}$  (Max rate increase: C/10 when Max Fb<sup>+</sup> = ( 1 + 2 \* W ) \* Qeq is received)

Gd =  $2.6 \times 10^{-4}$  (Max rate decrease: 1/2 when Max Fb<sup>-</sup> = ( 1 + 2 \* W ) \* Qeq is received)

Ru = 1 Mbps

- **Workload: 100% UDP (or Raw Ethernet) Traffic**

S1-S4: fixed-length (1500 bytes) frames, Bernoulli temporal distribution with parameter p = 0.49 (i.e., offered load = 49%)

# Simulation Run & Measurements

- **Simulation**

- Runs: 25 (different seeds)
- Duration: 100 ms
- Initial Transient: t = 5 ms (all sources start)
- Final Transient: t = 80 ms (2 sources stop)

- **Measurements**

- Throughput (Slide 7 & 8)

On congested downlink: 10 Gbps (100%, from 5 to 100 ms)

On uplinks: 2.5 Gbps (25%, from 20 to 80 ms)

- Buffer utilization (Slide 9 & 10)

Congested link

Rate limiter queues

- Fairness (Slide 11)

Maximum Deviation 
$$Dev_{\max} = \max_i \left( \frac{|R_i - T|}{T} \right) \quad i = 1 \dots N$$

Jain's Index

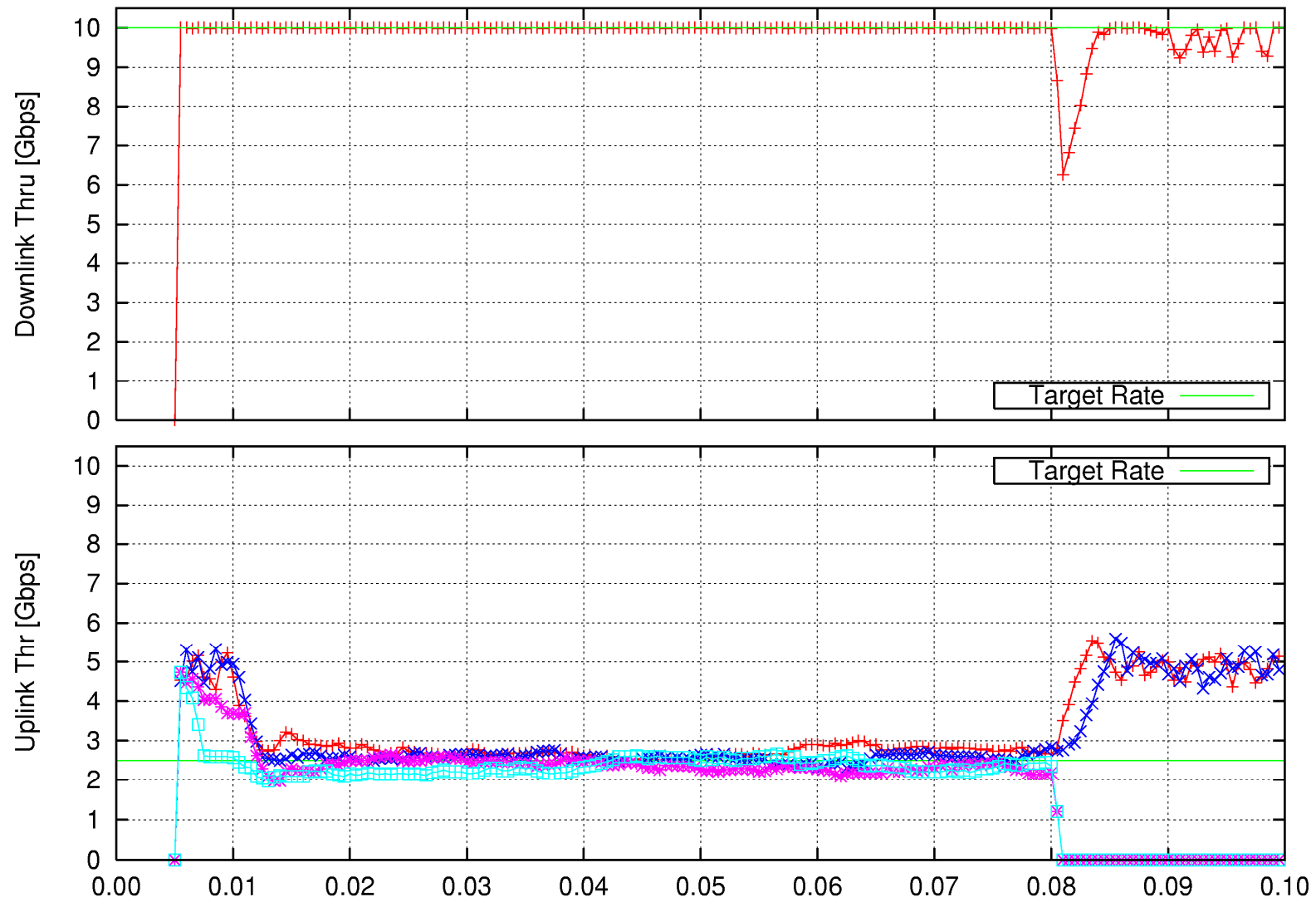
$$RJFI = \frac{\left( \sum_{i=1}^N R_i / T \right)^2}{N \sum_{i=1}^N (R_i / T)^2}$$

Alternate Index: Deviation Root Mean Square

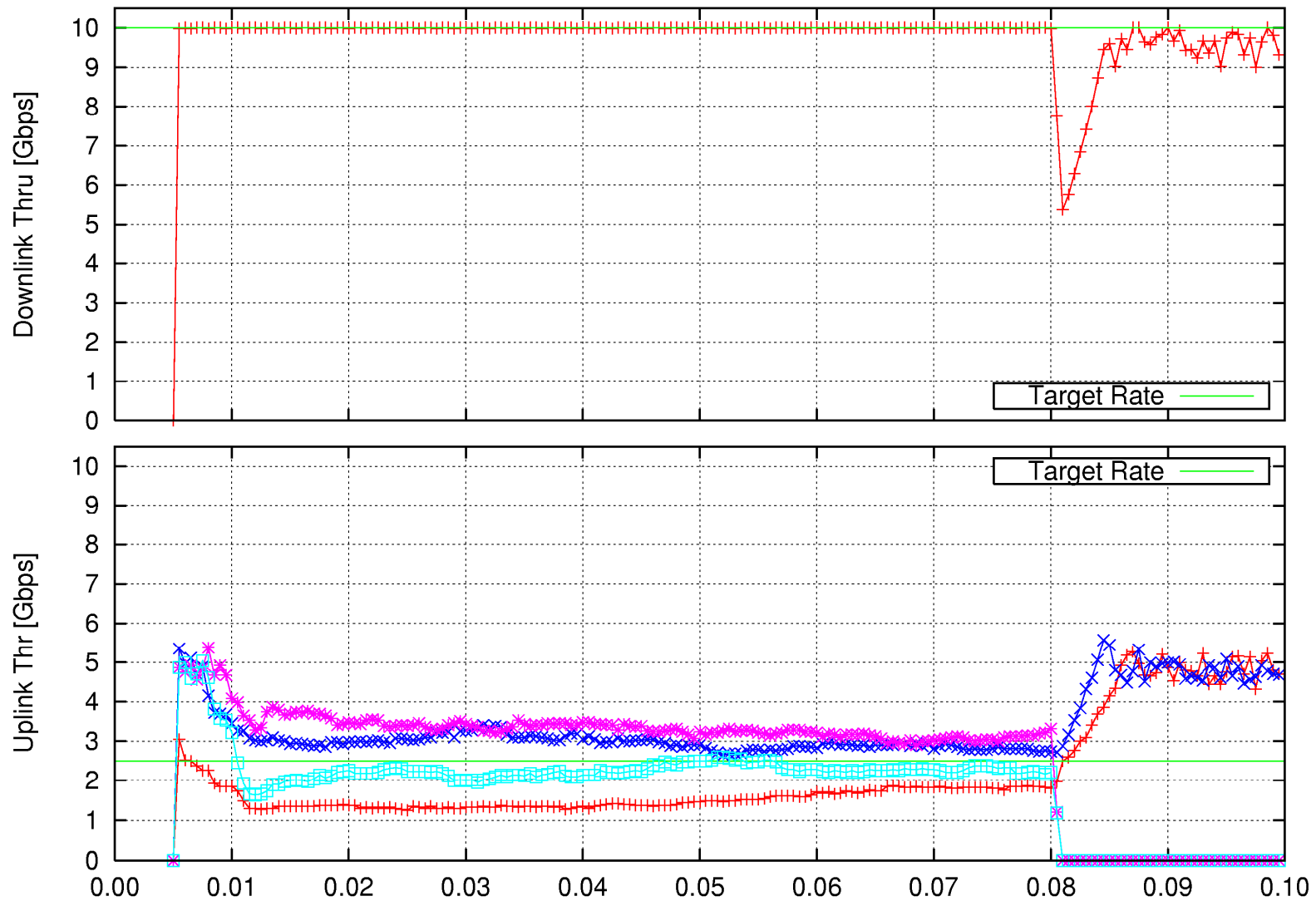
$$Dev_{rms} = \frac{1}{T} \cdot \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (R_i - T)^2}$$

Where: N = 4 number of flows  
R<sub>i</sub> = rate of each flow measured between 20 and 80 ms  
T = 2.5 Gbps Target rate

# Throughput (best)

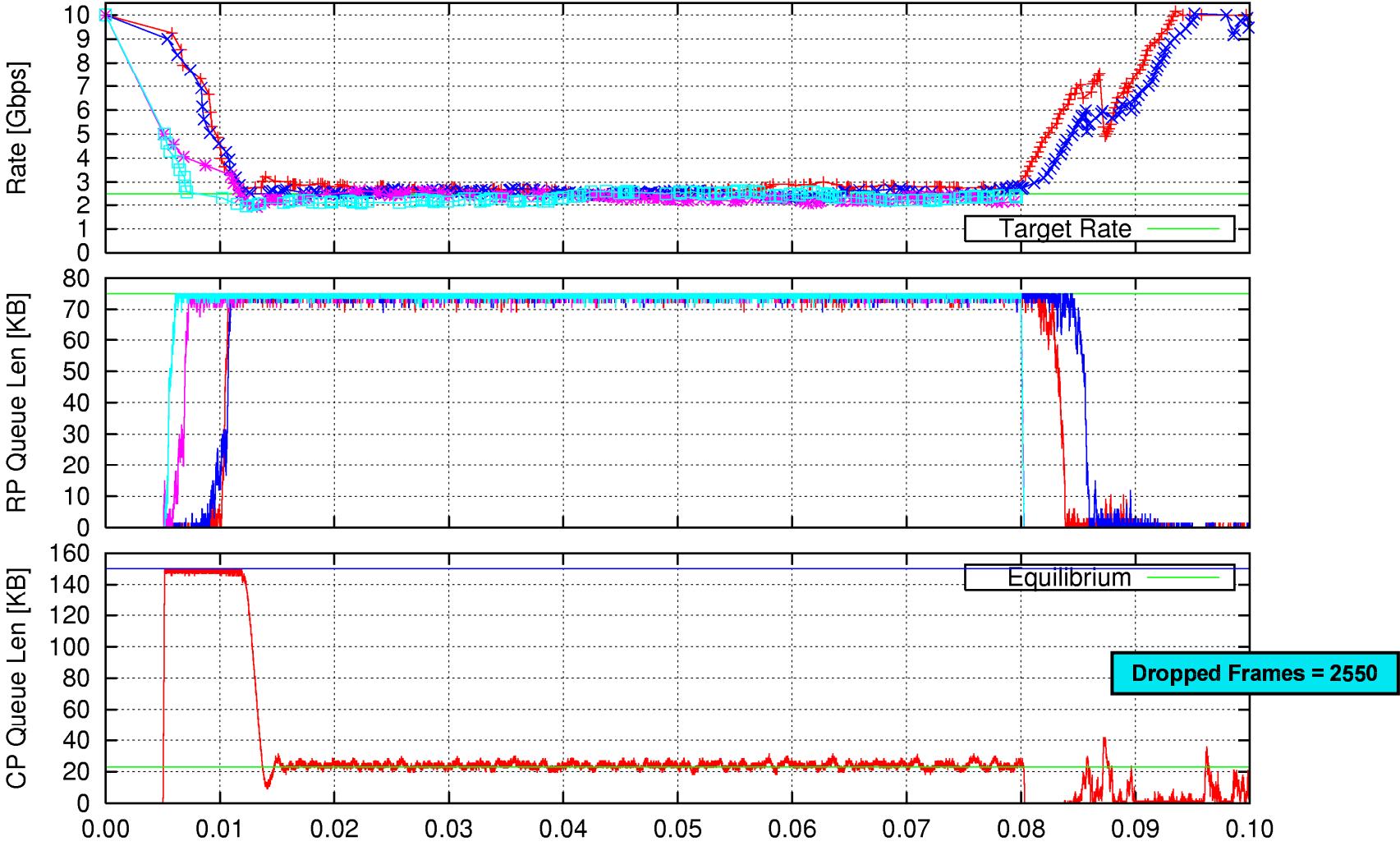


# Throughput (worst)

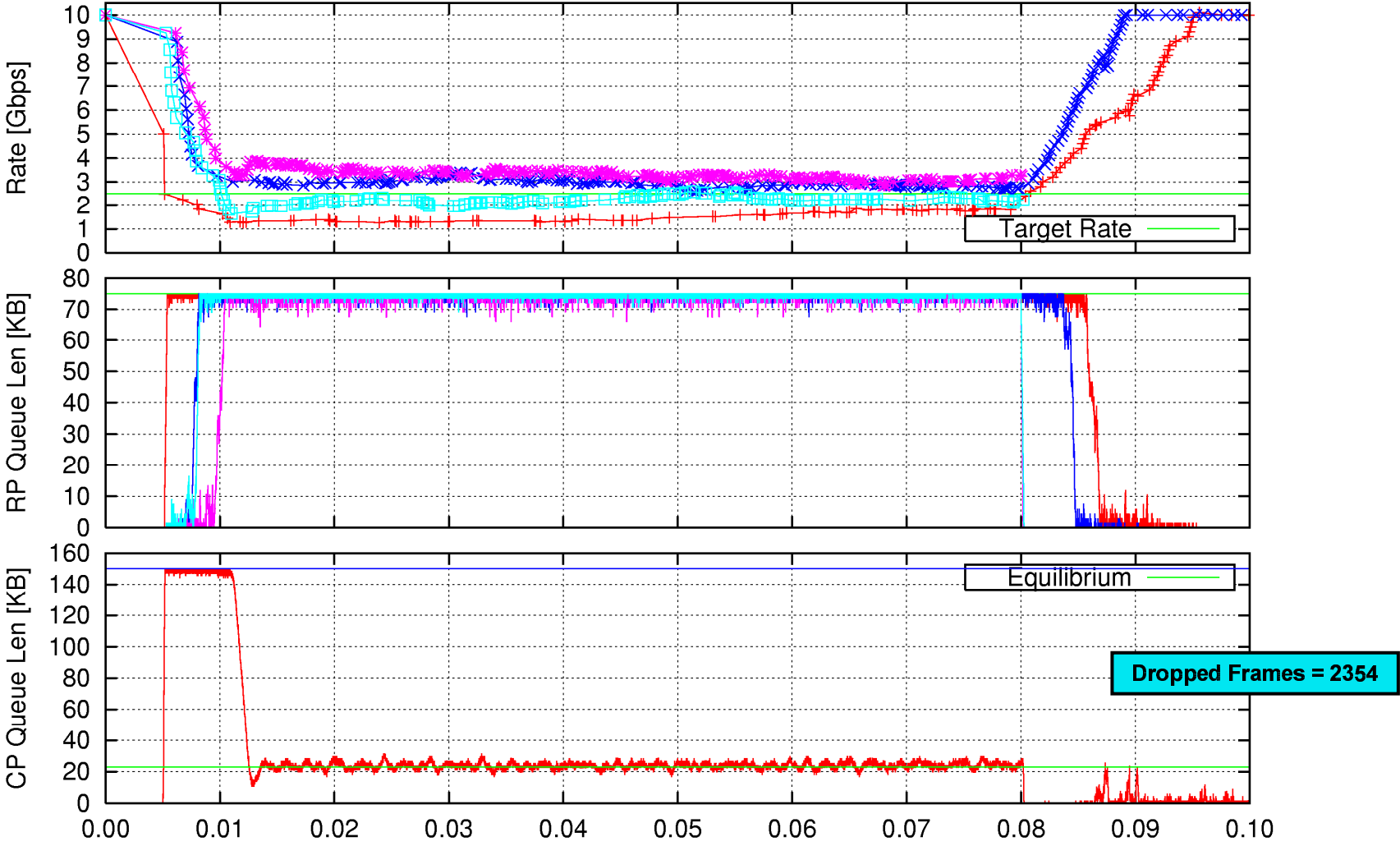




# Buffer Utilization (best)



# Buffer Utilization (worst)



# Fairness: Normalized Deviation & FI

Run	Dev <sub>max</sub>	Dev <sub>min</sub>	Dev <sub>mean</sub>	Dev <sub>rms</sub>	RJ-FI
19812	0.383	0.048	0.193	0.229	0.950
19639	0.252	0.001	0.126	0.177	0.970
19460	0.150	0.009	0.072	0.093	0.992
19285	0.143	0.035	0.078	0.088	0.992
19105	0.353	0.044	0.178	0.212	0.957
18924	0.215	0.017	0.124	0.144	0.980
18741	0.247	0.006	0.129	0.157	0.976
18578	0.103	0.032	0.064	0.069	0.995
18411	0.181	0.022	0.087	0.106	0.989
18240	0.189	0.019	0.096	0.113	0.987
18068	0.100	0.050	0.066	0.069	0.995
17790	0.138	0.041	0.089	0.096	0.991
17619	0.271	0.024	0.148	0.172	0.971
17452	0.216	0.059	0.135	0.147	0.979
17281	0.306	0.044	0.155	0.188	0.966
17112	0.083	0.027	0.057	0.061	0.996
16945	0.311	0.103	0.196	0.211	0.958
16800	0.372	0.093	0.237	0.261	0.937
16629	0.287	0.033	0.147	0.180	0.969
16462	0.249	0.021	0.128	0.163	0.974
16291	0.183	0.071	0.127	0.134	0.982
16120	0.200	0.135	0.168	0.170	0.972
15949	0.220	0.202	0.213	0.213	0.957
15794	0.089	0.009	0.048	0.057	0.997
15625	0.152	0.041	0.095	0.103	0.990

Dev<sub>max</sub> stats:  
 min = 0.083  
 mean = 0.216  
 max = 0.383

Dev<sub>rms</sub> stats:  
 min = 0.057  
 mean = 0.144  
 max = 0.261

RJ-FI stats:  
 min = 0.937  
 mean = 0.977  
 max = 0.997

# What happens with BCN(0,0)?

- **Settings**

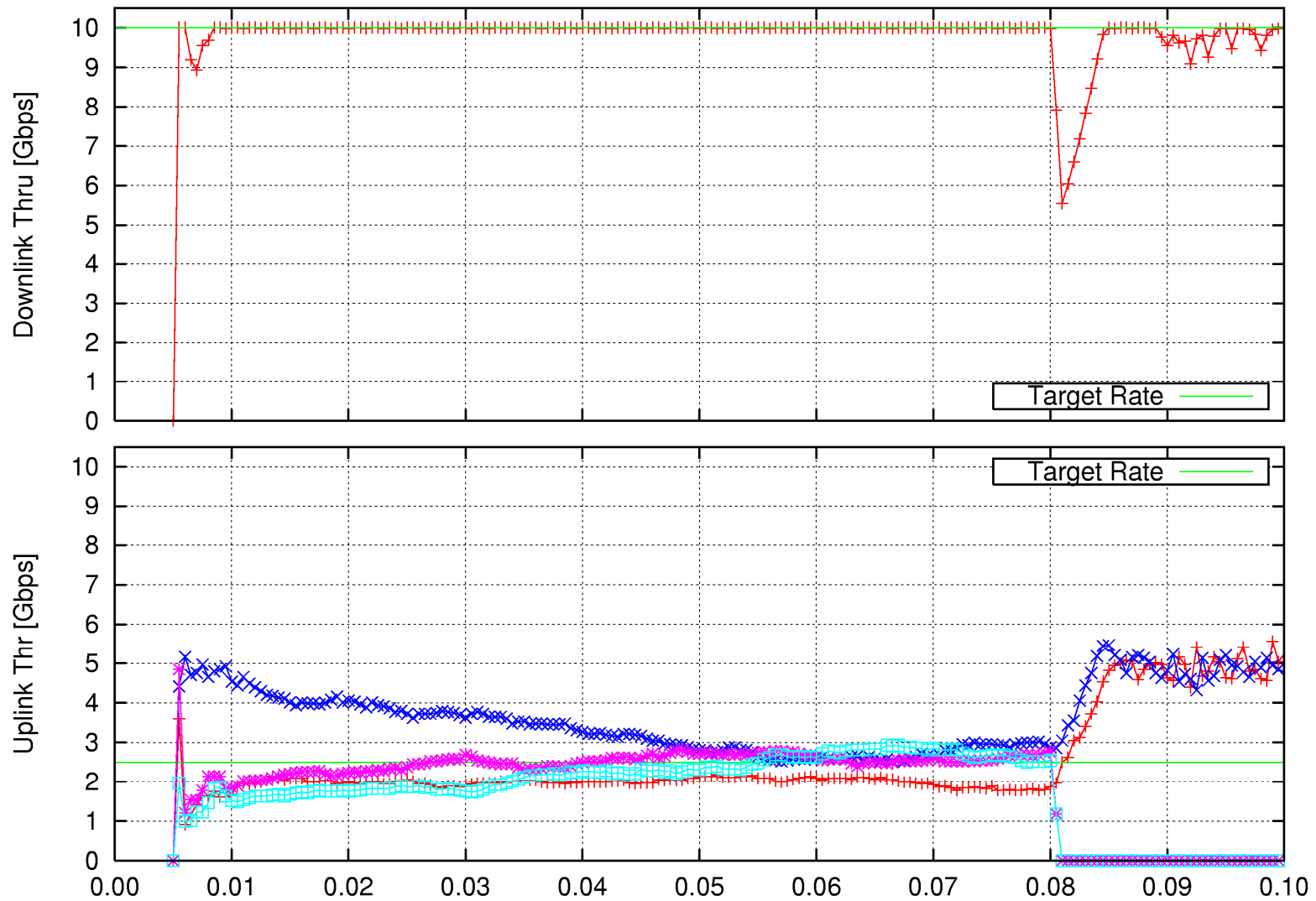
**$Q_{sc} = 112.5 \text{ KB}$  (75% buffer)**

**$T_{max} = 100 \text{ us}$**

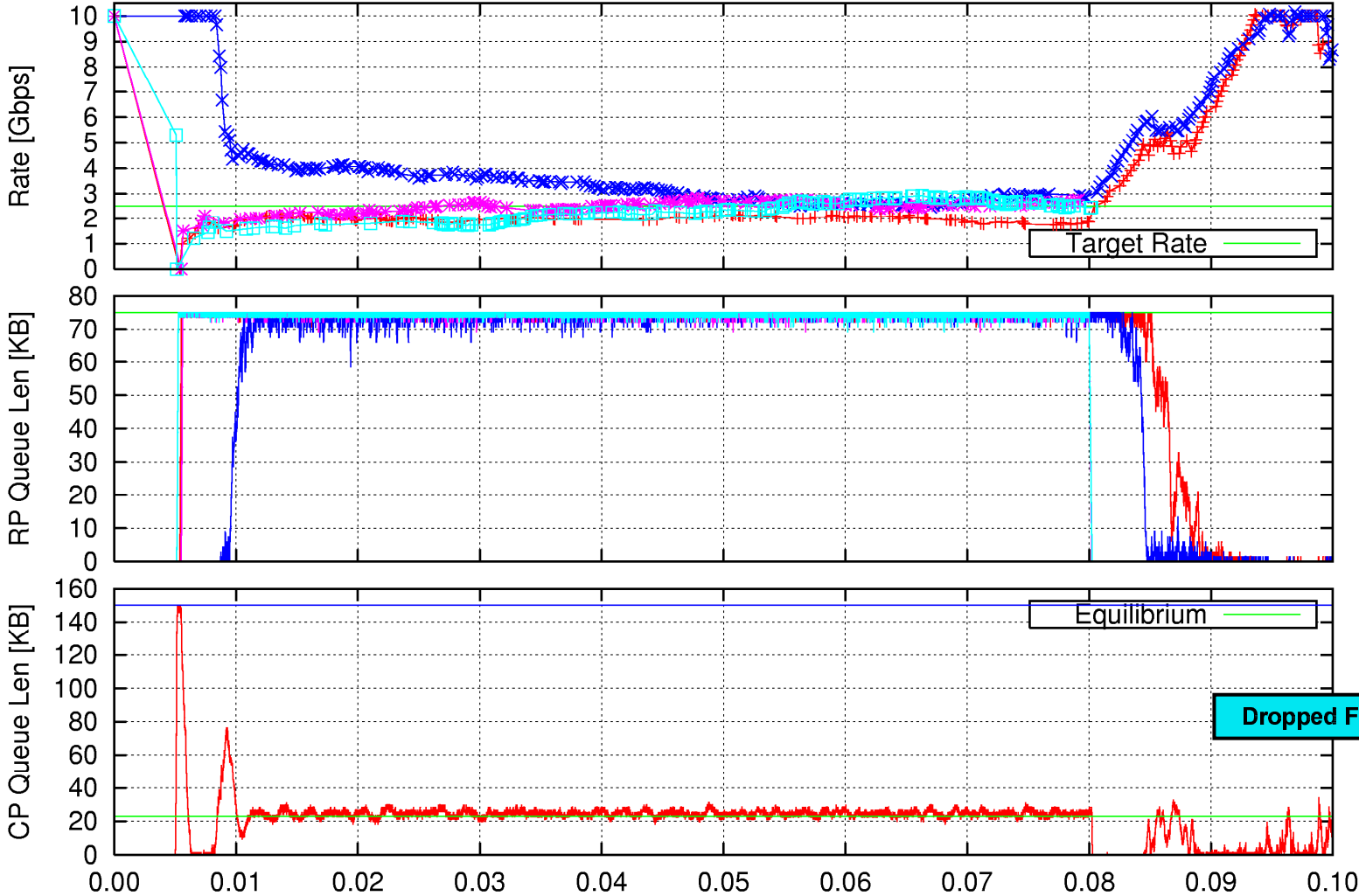
**$R_{min} = 1 \text{ Gbps}$  (10% link capacity)**

**$\text{Drift} = 10 \text{ Gbps} / \text{s}^2$  (1 Mbps every 100 us)**

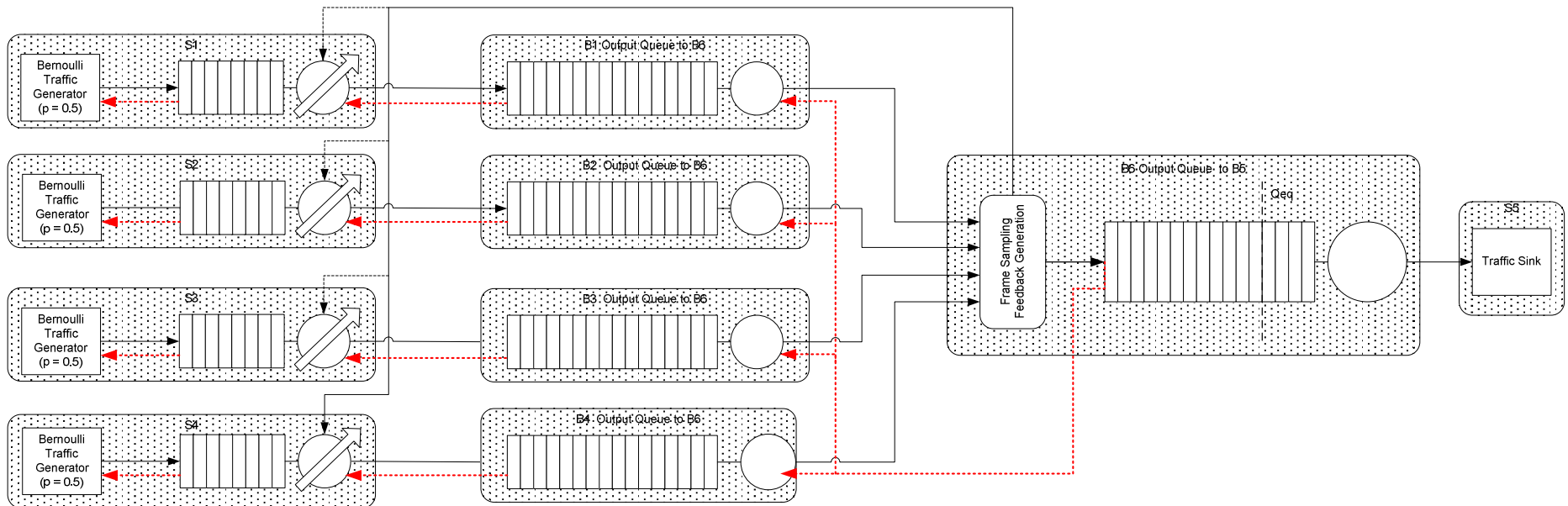
# Throughput



# Buffer Utilization



# What happens with Pause?



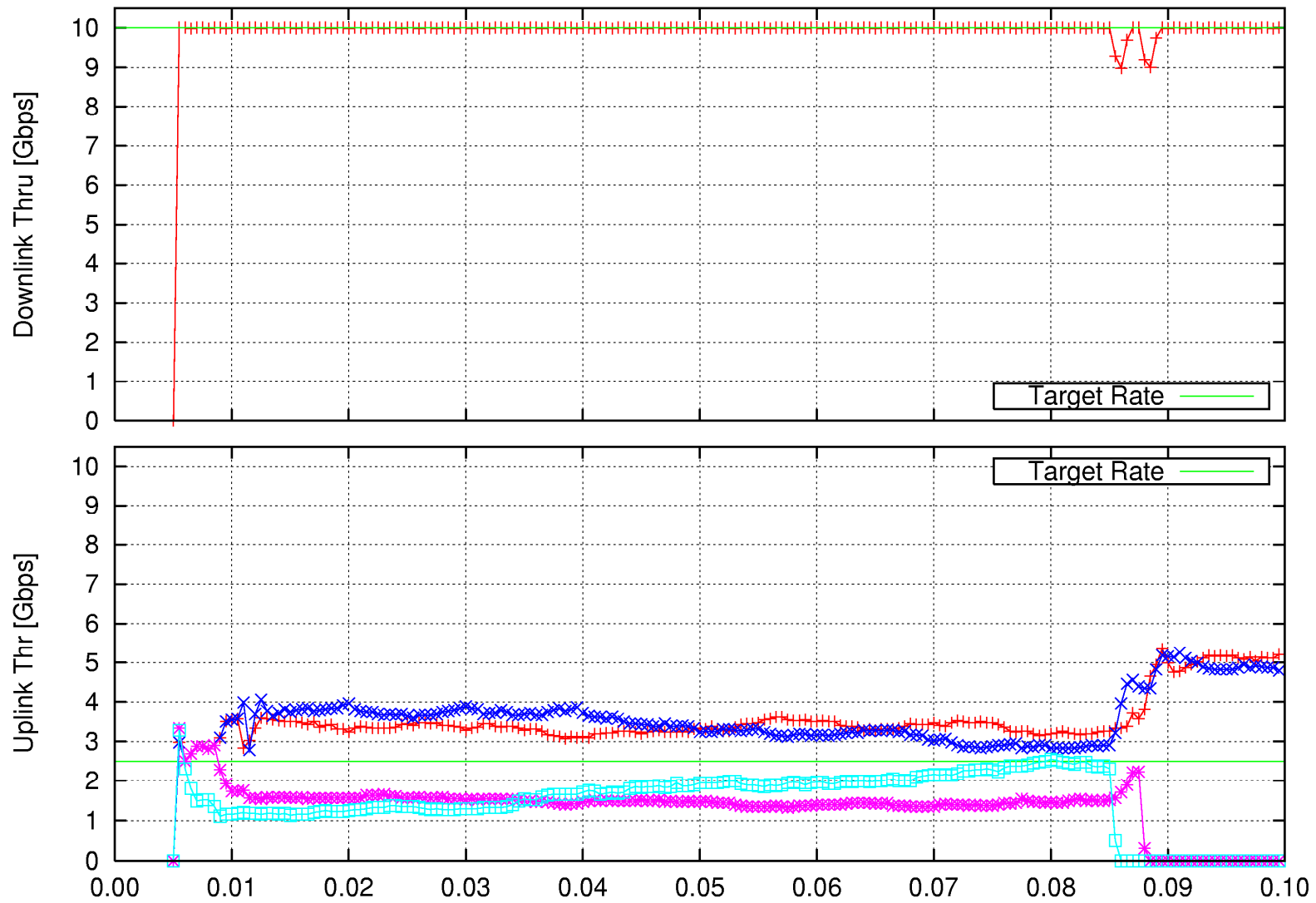
- **Settings**

**Pause Assert Threshold = 140 KB**

**Pause De-assert Threshold = 130 KB**

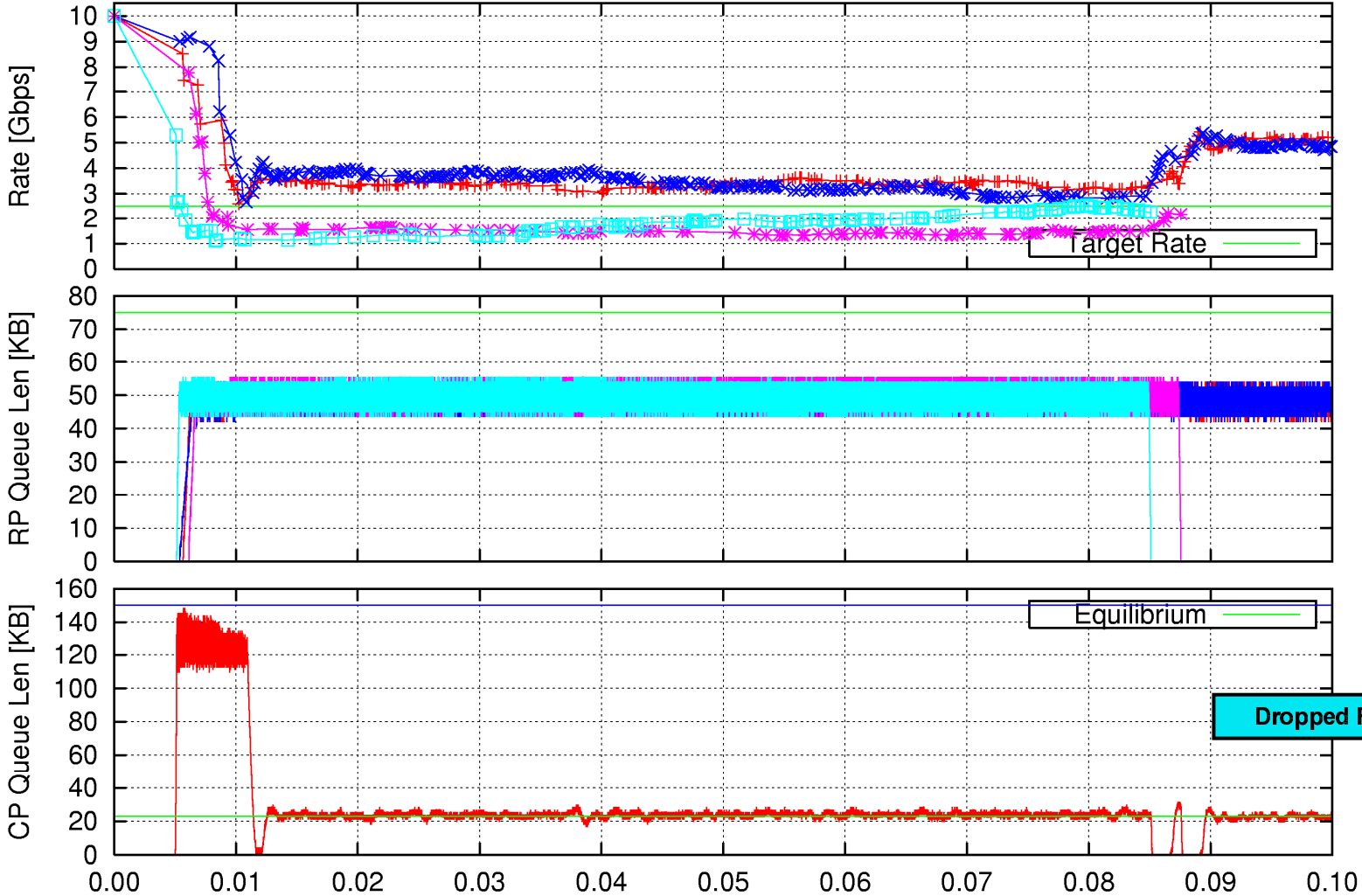
**BCN CP Detection disabled @ bridges B1-B4**

# Throughput





# Buffer Utilization



# What about Pause and BCN(0,0)?

- **Settings**

**Pause Assert Threshold = 140 KB**

**Pause De-assert Threshold = 130 KB**

**BCN CP Detection disabled @ bridges B1-B4**

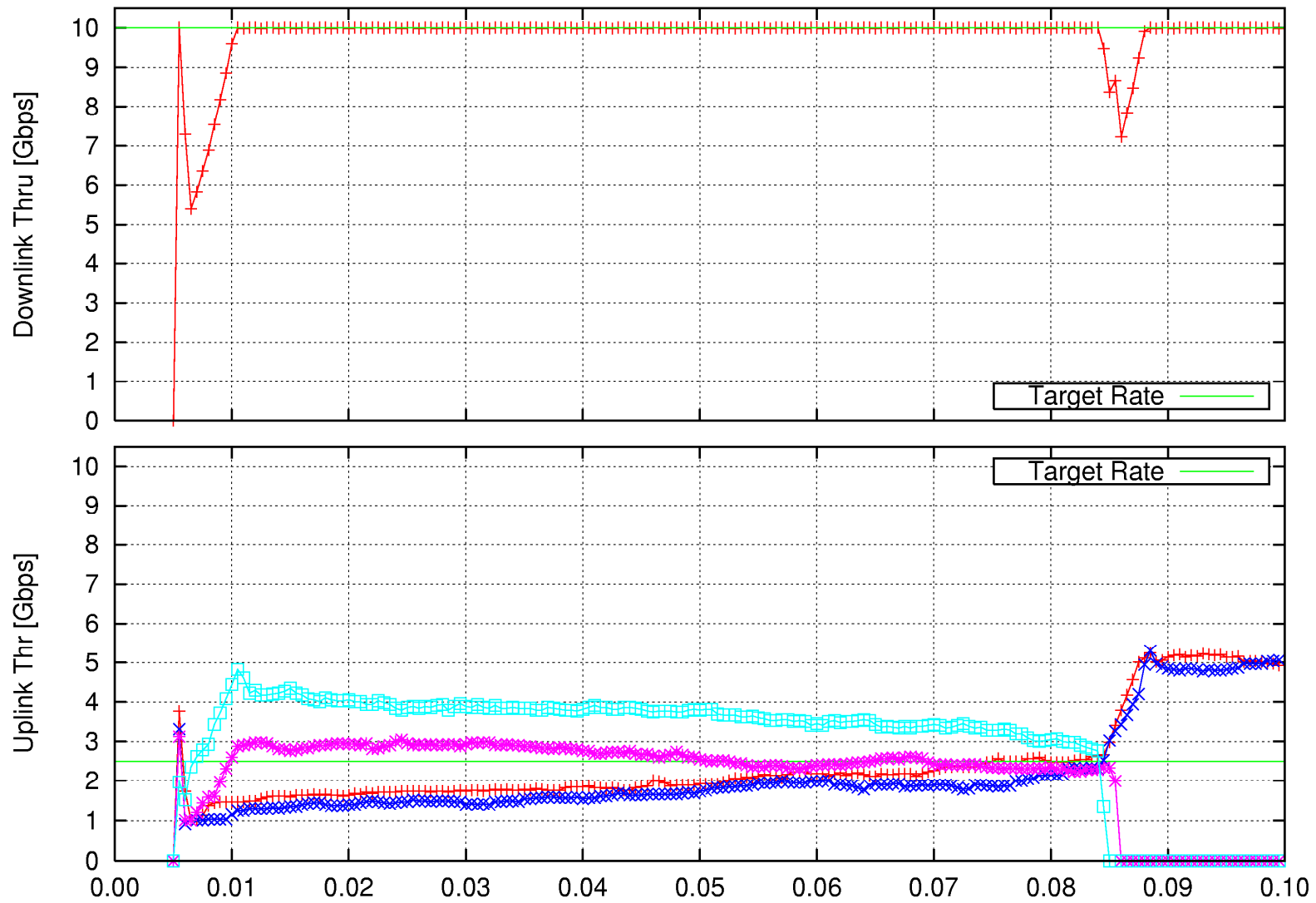
**Qsc = 112.5 KB (75% buffer)**

**Tmax = 100 us**

**Rmin = 1 Gbps (10% max rate)**

**Drift = 10 Gbps / s<sup>2</sup> (1 Mbps every 100 us)**

# Throughput



# Buffer Utilization

