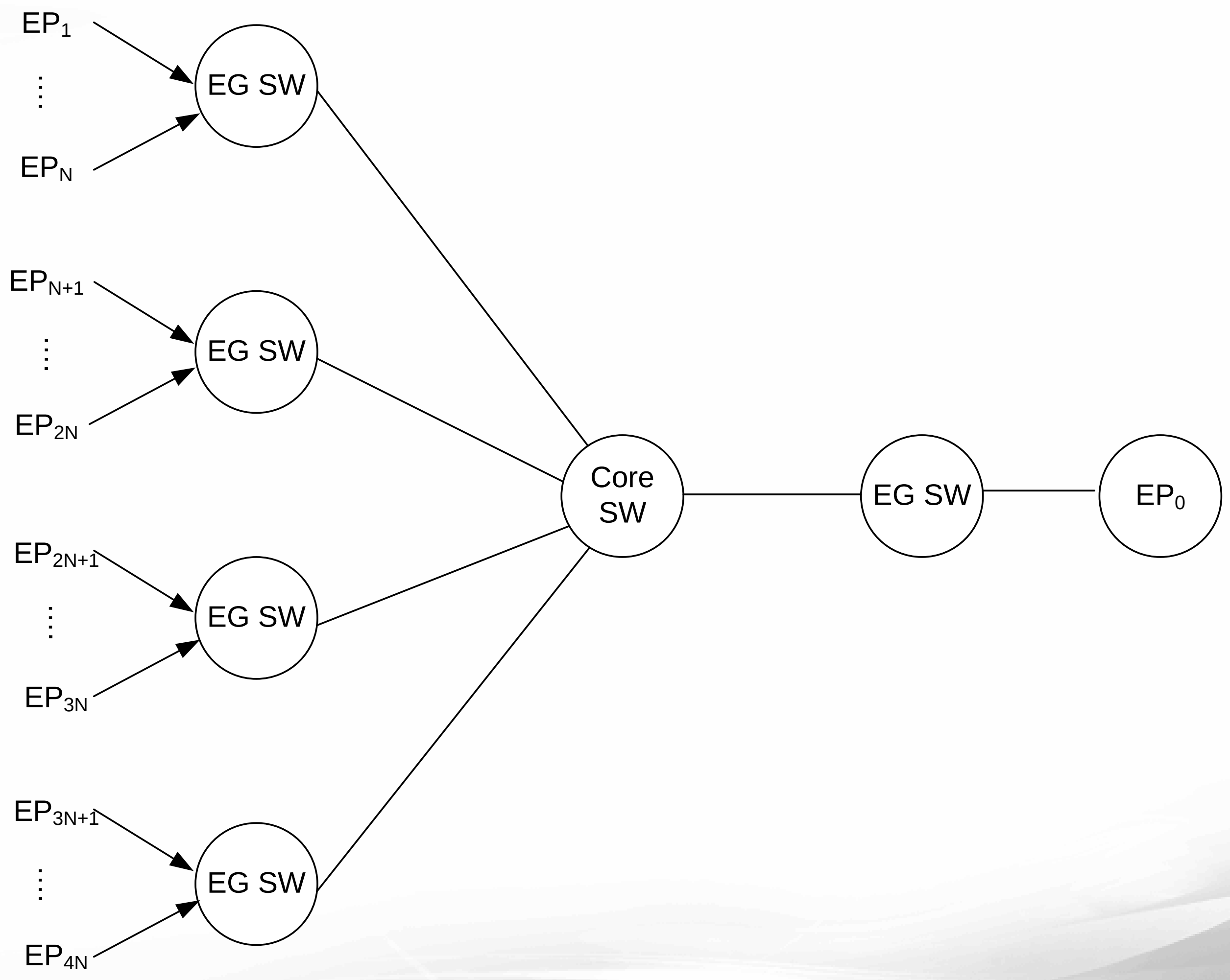
BCN Simulation Results
100 Flows Scenario

November 2, 2006

Topology



EP1
EPN
EG SW
EPN+1
EP2N
EG SW
EP2N+1
EP3N
EG SW
EP3N+1
EP4N
EG SW
Core SW
EG SW
EP0

Topology & Workload

- Short Range, High-Speed Datacenter-like Network
  - Link Capacity = 10 Gbps
  - Egress Port Buffer Size = 150 KB
  - Switch Latency = 1 us
  - Link Length = 100 m (.5 us propagation delay)
  - Endpoint response time = 1 us
- Traffic Load
  - Traffic Type: 100% UDP (or Raw Ethernet) Traffic
  - Frame Size Distribution: Fixed length (1500 bytes) frames
  - Arrival Distribution: Bernoulli temporal distribution
  - Offered Load/Endpoint = 2%
  - N=25, Destination Distribution: EP_1 – EP_{100} send to EP_0
- Simulation Time
  - Each source starts at 5ms, and simulation stops at 200ms

BCN Parameters

- Qeq
  - 16 (1500-byte frames)
  - 375 * 64 byte pages
- Frame Sampling
  - Frames are sampled on average 150 KB received to the egress queue
- W = 2
- Gi = 12.42
  - Computed as (Linerate/10) * [1/((1+2*W)*Q_eq)]
  - Gi = 5.3×10^{-1} * (1500/64) = 12.42
- Maximum rate decrease
  - 0.5, computed as 1/2*[1/((1+2*W)*Q_eq)]
  - 0.95, computed as 0.95*[1/((1+2*W)*Q_eq)]
- Ru = 1 Mbps

BCN(0,0), BCN(MAX), Drift

BCN(0,0) (from Cisco)

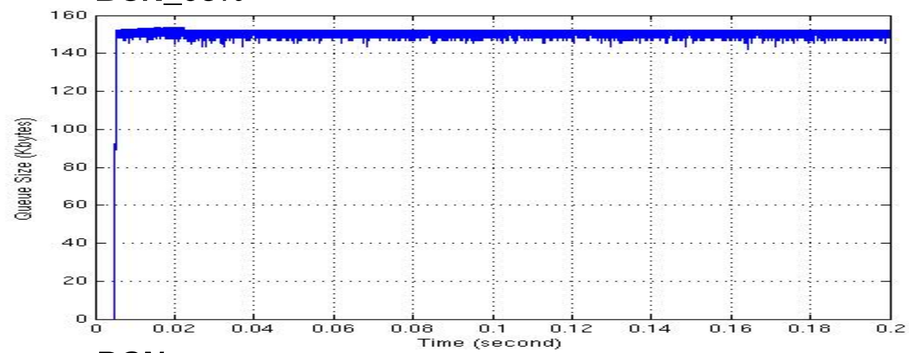
- – Current rate R is set to 0
- – Random timer [0, Tmax]: when timer expires, current rate R set to Rmin
- – Each time Tmax doubled and Rmin halved (exponential backoff)
- – Settings:
  - Qsc = 112.5 KB (75% buffer)
  - Tmax = 100us
  - Rmin = 1 Gbps (10% max rate)

- BCN(MAX):
  - – Instead of BCN(0,0) when Q>Qsc, send BCN(MAX) to decrease the rate by maximum amount
- Drift:
  - – At fixed time intervals Ti, the current rate is incremented by a unit
  - – Never stop drift except timeout in BCN(0,0)
  - – Drift = 1 Mbps every 100us

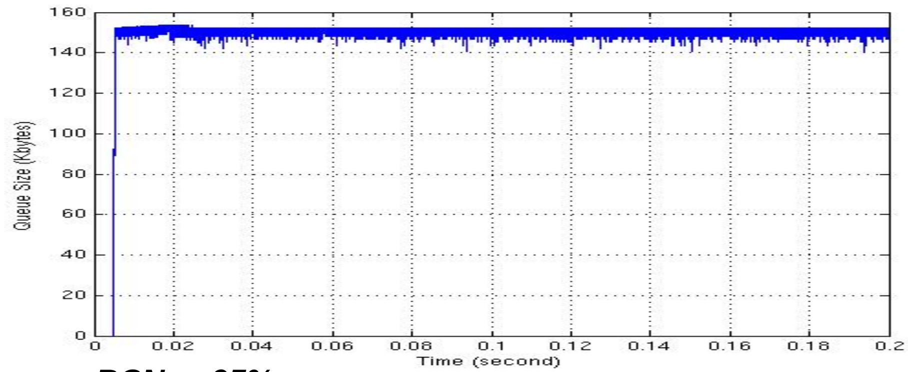
Non-BCN(0,0) Variants : CS Queue (150K Sampling)

Severe Congestion Behavior	Max Rate Decrease Percentage	CS Packet Loss
None	95%	162947
BCN(MAX)	50%	137099
BCN(MAX)	95%	12363

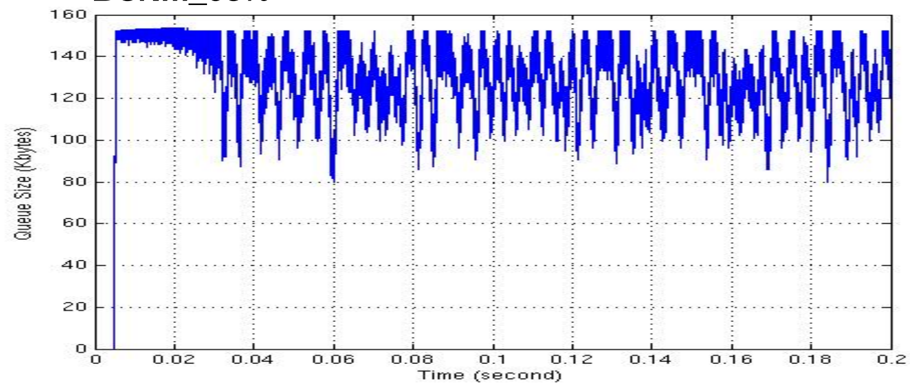
BCN_95%



BCNm



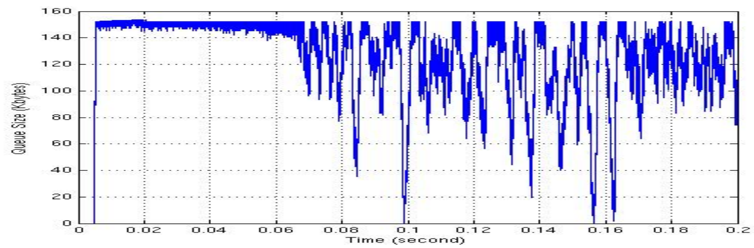
BCNm_95%



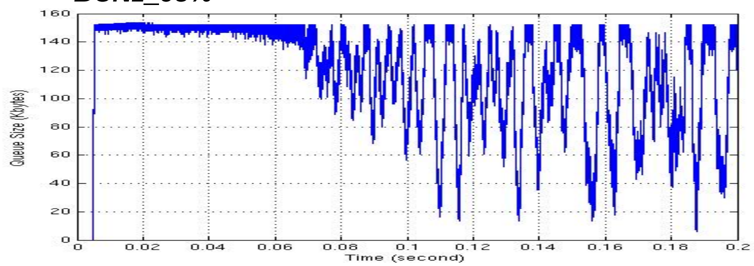
BCN(0,0) Variants : CS Queue (150K Sampling)

Max Rate Decrease Percentage	Reset Behavior	CS Packet Loss
50%	Always	39093
95%	Always	37863
50%	At least 1 ms after timeout	39185
50%	Never	37194

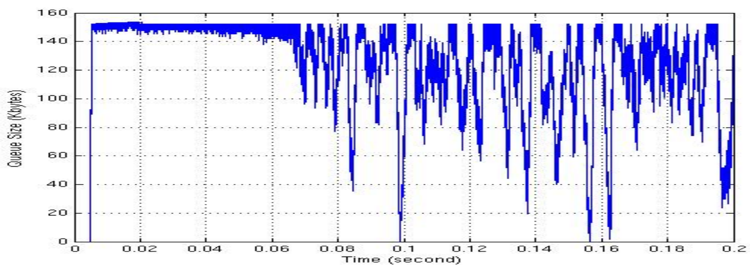
BCNz



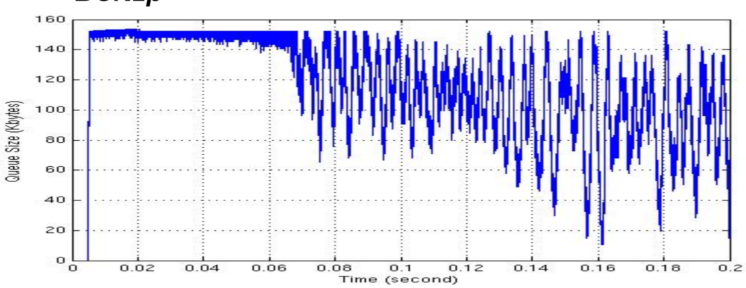
BCNz_95%



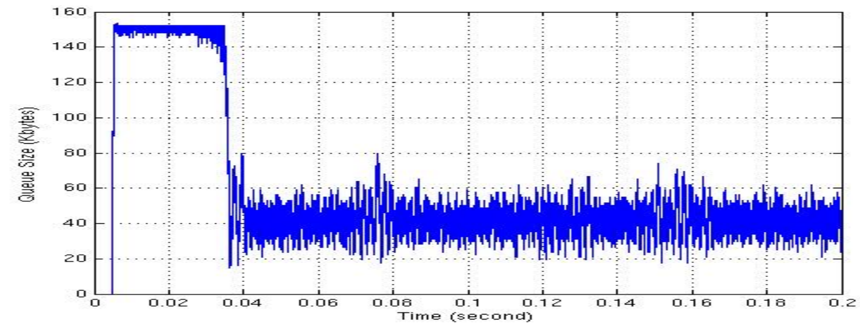
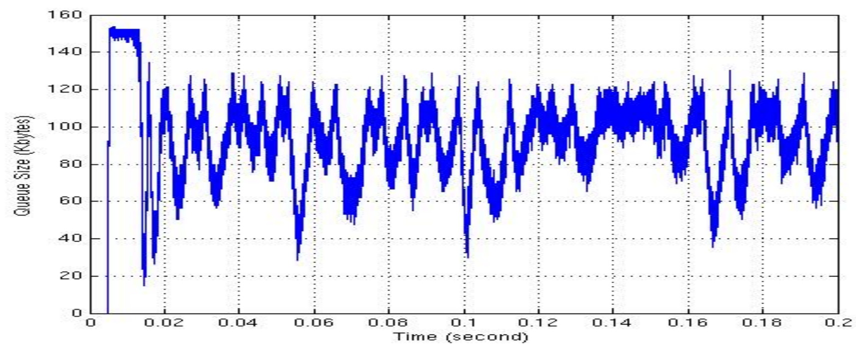
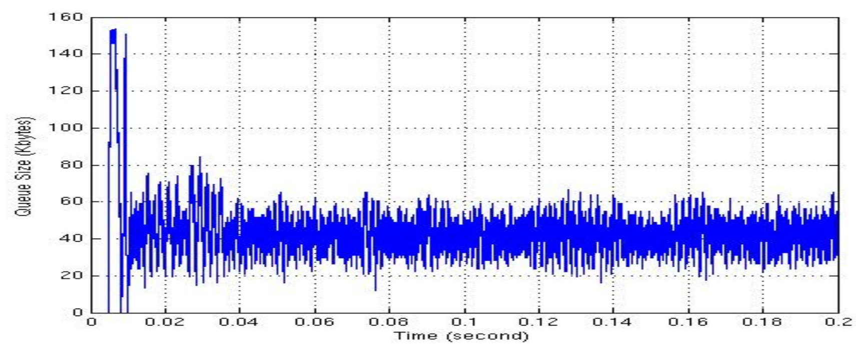
BCNzd



BCNzp



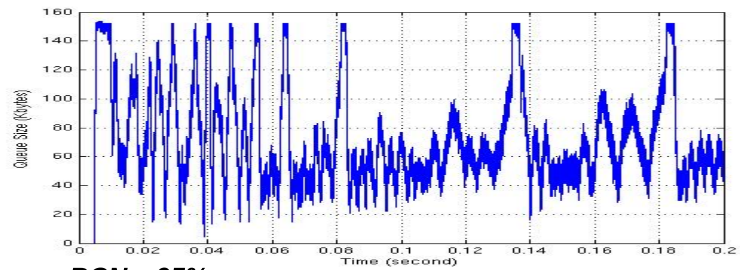
Non-BCN(0,0) Variants : CS Queue (25K Sampling)

Severe Congestion Behavior	*Max Rate Decrease Percentage*		*CS Packet Loss*
None	*95%*	*BCN 95%* 	*19657*
BCN(MAX)	*50%*	*BCNm* 	*5071*
BCN(MAX)	*95%*	*BCNm_95%* 	*851*

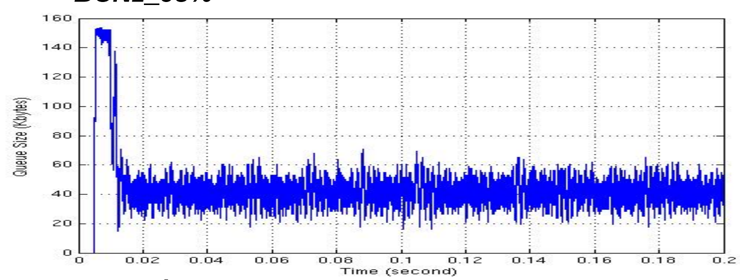
BCN(0,0) Variants : CS Queue (25K Sampling)

Max Rate Decrease Percentage	Reset Behavior	CS Packet Loss
50%	*Always*	*6052*
95%	*Always*	*2571*
50%	*At least 1 ms after timeout*	*4647*
50%	*Never*	*2725*

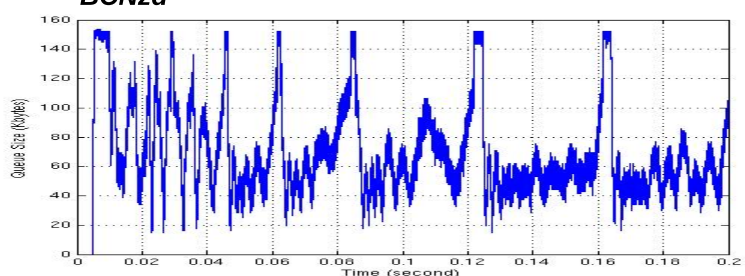
BCNz



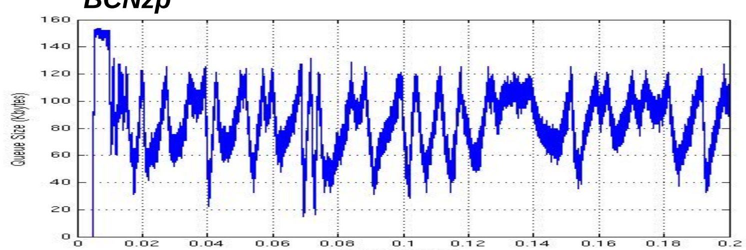
BCNz_95%



BCNzd



BCNzp



Observations

- Current parameters perform poorly for large numbers of flows even if individual rates are small.
  - Loop latency for an RP increases with increasing number of flows
  - Decreasing sampling rate improves performance
  - Other parameter and behavior changes help too
- Trade-offs
  - Some changes may reduce throughput (over control of flow rate) for small number of flows.
  - Some changes produce more BCN traffic.
- Further work:
  - Identify parameters and behaviors that work well for large numbers of flows
  - Verify impact on behavior with small numbers of flows