BCN Simulation Results
100 Flows Scenario
November 2, 2006
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: \( EP_1 \rightarrow EP_{100} \) 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 \(\frac{\text{Linerate}}{10} \times \frac{1}{(1+2W)Q_{eq}}\)
  - \(Gi = 5.3 \times 10^{-1} \times \frac{1500}{64} = 12.42\)

• **Maximum rate decrease**
  - 0.5, computed as \(\frac{1}{2} \times \frac{1}{(1+2W)Q_{eq}}\)
  - 0.95, computed as \(0.95 \times \frac{1}{(1+2W)Q_{eq}}\)

• **Ru = 1 Mbps**
BCN(0,0), BCN(MAX), Drift

BCN(0,0) (from Cisco)
- Current rate $R$ is set to 0
- Random timer $[0, T_{max}]$: when timer expires, current rate $R$ set to $R_{min}$
- Each time $T_{max}$ doubled and $R_{min}$ halved (exponential backoff)
- Settings:
  - $Q_{sc} = 112.5 \text{ KB (75\% buffer)}$
  - $T_{max} = 100\text{us}$
  - $R_{min} = 1 \text{ Gbps (10\% max rate)}$

• BCN(MAX):
  - Instead of BCN(0,0) when $Q > Q_{sc}$, send BCN(MAX) to decrease the rate by maximum amount

• Drift:
  - At fixed time intervals $T_i$, 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)

<table>
<thead>
<tr>
<th>Severe Congestion Behavior</th>
<th>Max Rate Decrease Percentage</th>
<th>CS Packet Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>95%</td>
<td>162947</td>
</tr>
<tr>
<td>BCN(MAX)</td>
<td>50%</td>
<td>137099</td>
</tr>
<tr>
<td>BCN(MAX)</td>
<td>95%</td>
<td>12363</td>
</tr>
</tbody>
</table>
### BCN(0,0) Variants: CS Queue (150K Sampling)

<table>
<thead>
<tr>
<th>Max Rate Decrease Percentage</th>
<th>Reset Behavior</th>
<th>CS Packet Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>Always</td>
<td>39093</td>
</tr>
<tr>
<td>95%</td>
<td>Always</td>
<td>37863</td>
</tr>
<tr>
<td>50%</td>
<td>At least 1 ms after timeout</td>
<td>39185</td>
</tr>
<tr>
<td>50%</td>
<td>Never</td>
<td>37194</td>
</tr>
</tbody>
</table>

![Graphs](BCNz, BCNz_95%, BCNzd, BCNzp)
## Non-BCN(0,0) Variants: CS Queue (25K Sampling)

<table>
<thead>
<tr>
<th>Severe Congestion Behavior</th>
<th>Max Rate Decrease Percentage</th>
<th>CS Packet Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>95%</td>
<td>19657</td>
</tr>
<tr>
<td>BCN(MAX)</td>
<td>50%</td>
<td>5071</td>
</tr>
<tr>
<td>BCN(MAX)</td>
<td>95%</td>
<td>851</td>
</tr>
</tbody>
</table>

**BCN 95%**

- CS Packet Loss: 19657

**BCNm**

- CS Packet Loss: 5071

**BCNm_95%**

- CS Packet Loss: 851
## BCN(0,0) Variants: CS Queue (25K Sampling)

<table>
<thead>
<tr>
<th>Max Rate Decrease Percentage</th>
<th>Reset Behavior</th>
<th>BCNz</th>
<th>BCNz_95%</th>
<th>BCNzd</th>
<th>BCNzp</th>
<th>CS Packet Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>Always</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
<td>6052</td>
</tr>
<tr>
<td>95%</td>
<td>Always</td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
<td>2571</td>
</tr>
<tr>
<td>50%</td>
<td>At least 1 ms after timeout</td>
<td><img src="image9.png" alt="Graph" /></td>
<td><img src="image10.png" alt="Graph" /></td>
<td><img src="image11.png" alt="Graph" /></td>
<td><img src="image12.png" alt="Graph" /></td>
<td>4647</td>
</tr>
<tr>
<td>50%</td>
<td>Never</td>
<td><img src="image13.png" alt="Graph" /></td>
<td><img src="image14.png" alt="Graph" /></td>
<td><img src="image15.png" alt="Graph" /></td>
<td><img src="image16.png" alt="Graph" /></td>
<td>2725</td>
</tr>
</tbody>
</table>
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