RPR MAC Definition and implementation

Sanjay K. Agrawal

Luminous Networks

sanjay@luminousnetworks.com

Outline

- Class of service support in layered networks
- RPR MAC framework proposal
- RPR MAC hardware implementation

Class of Service support

- Layer 3: IP DSCP specifies 3 bits for CoS, 3 bits for Drop Presidence, 2bits for ECN
- Layer2: 802.1P/Q specifies 3 bits for CoS
- Layer 2.5: MPLS specifies 3 bits of CoS

Diffserv code points are standard and poised to be consistent across layers

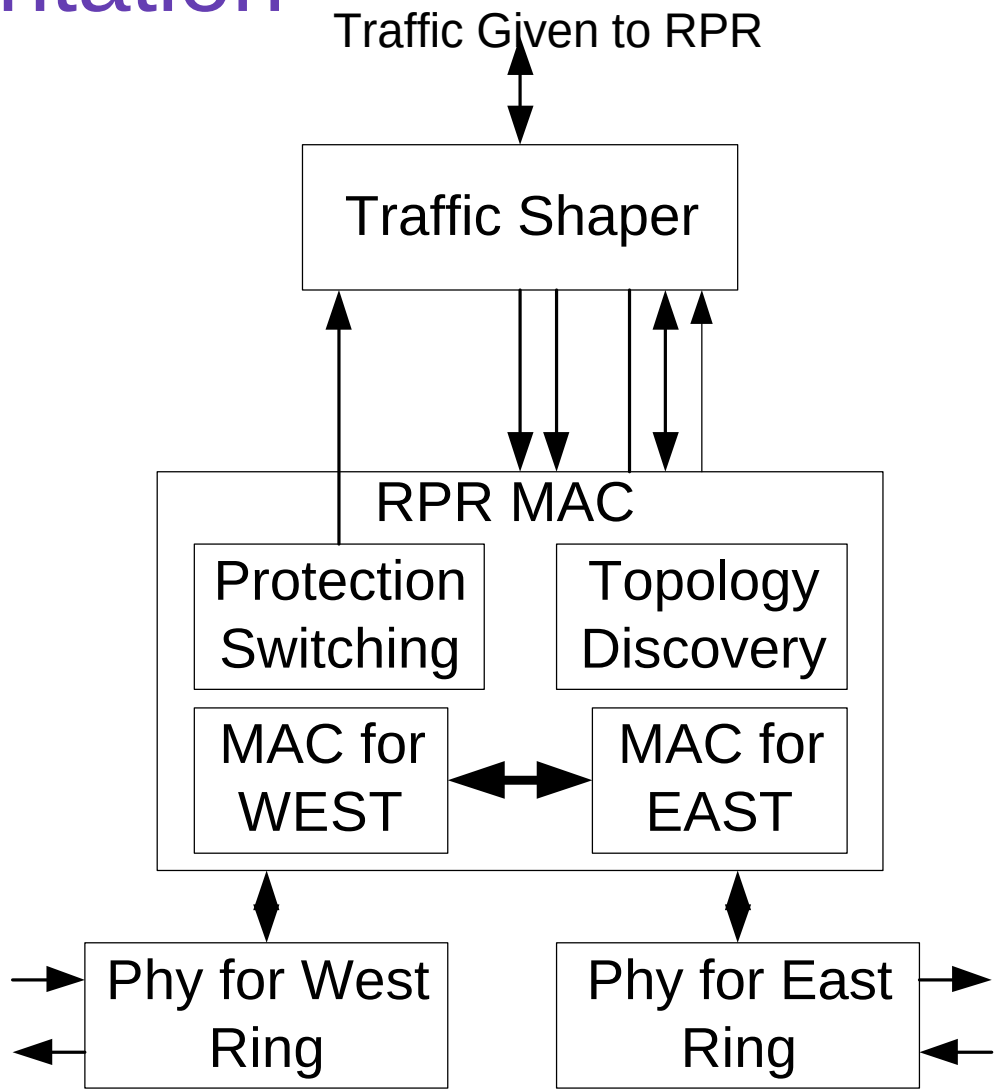
Class of Service

- Express Forwarding (110): Time sensitive committed class
  - Legacy leased line, Voice over IP
- Assured Forwarding 3 (011): Time Sensitive committed class
  - Video
- Assured Forwarding 2(010): Time insensitive committed class
  - Committed data Services, Protected
- Assured Forwarding 1(001) Time insensitive over-committed class
  - Over committed data services
- Best Effort (000): data services

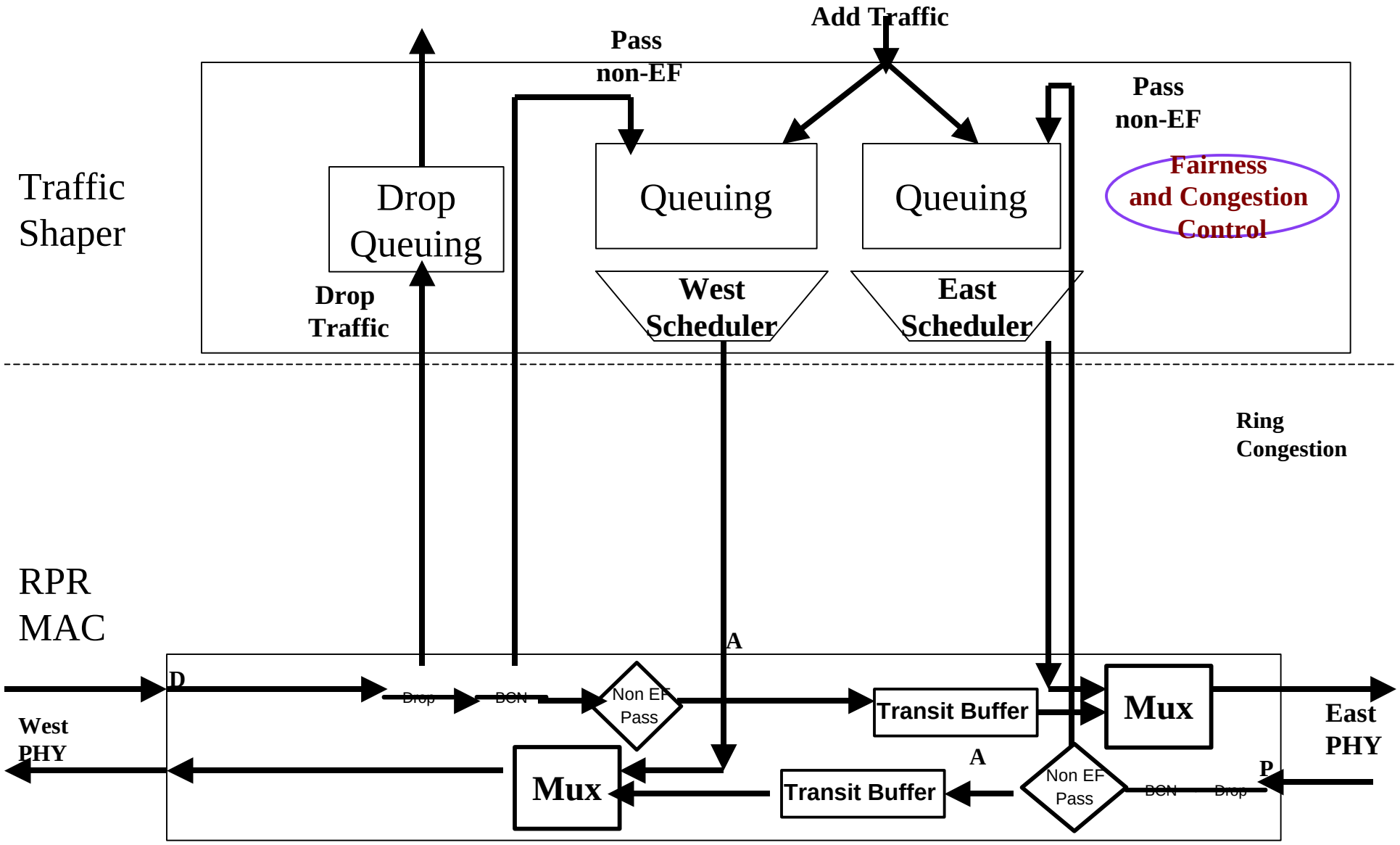
Common RPR MAC functional Requirements

- Class of Service Support
- Backward Congestion Notification using Internode signaling
- Using class of service simultaneous support for
  - non-preemptable cut through traffic
    - to minimize latency for high priority class
  - Store and forward traffic
    - to allow low priority pass traffic to be stored while high priority add is admitted

Proposed RPR MAC Implementation



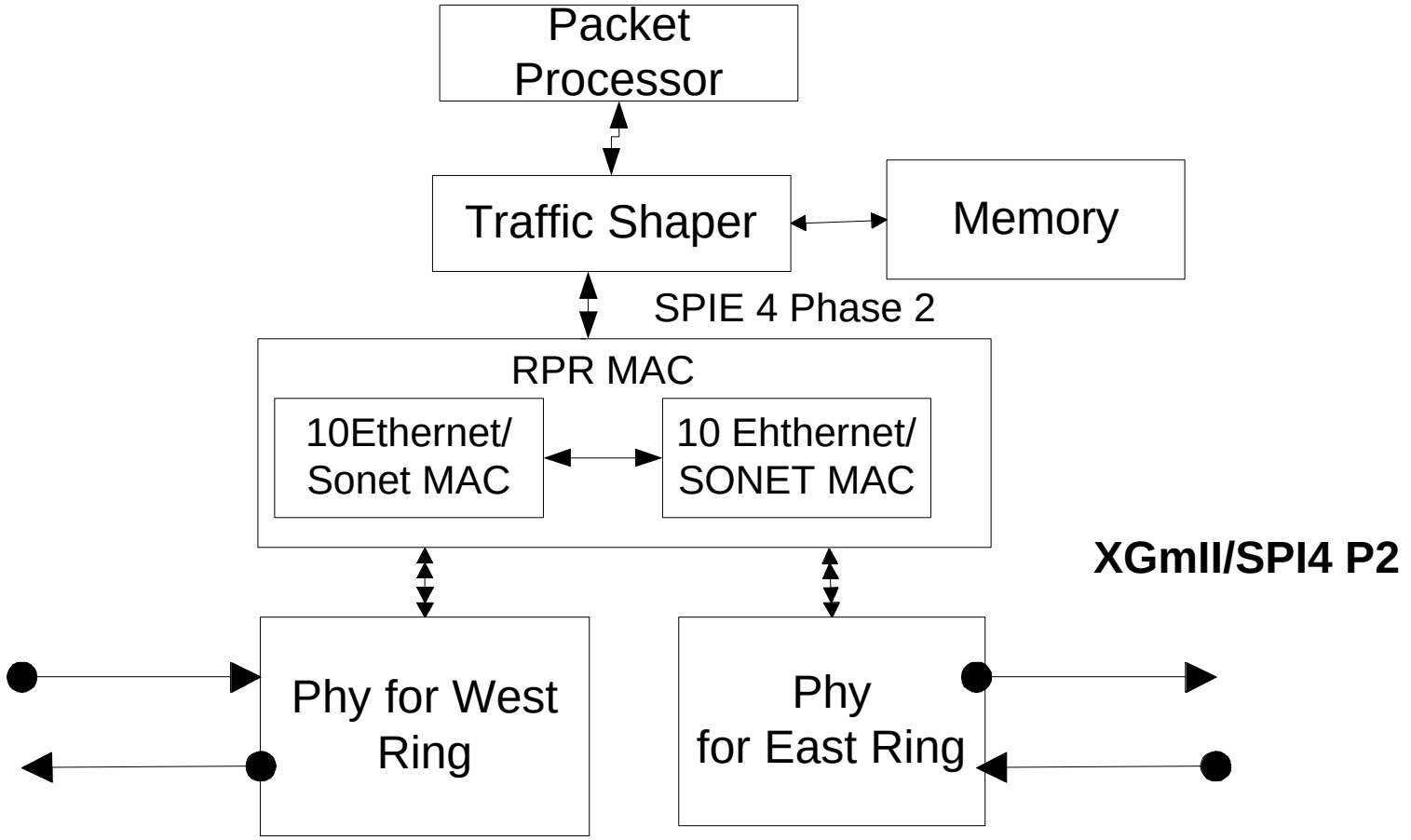
RPR System Architecture



RPR MAC Hardware Implementation Requirements

- Rate adaptation for the drop traffic
- Minimal buffering in the MAC chip
  - Only on board buffers ~ 8Mbits (.8msec@10Gig)
  - External memory interface increases the MAC pin count by 128 pins
- Use of standard interfaces high speed interfaces

Proposed RPR MAC Hardware Implementation



Traffic Shaper Architecture: Buffering vs. BCN

- Backward Congestion Notification (BCN)
  - Avoids buffering in the intermediate nodes in the rings.
  - Propagates congestion to source nodes.
  - Need per RPR node virtual queuing to avoid head of a node blocking
  - Flow control signaling frequency and span distances may be issue
  - Add traffic requires 50-100ms buffering
- Class based queueing
  - Avoids internode signaling.
  - Well tested
  - Requires 50-100ms buffering

Traffic Shaper Functions

- Rate adaptation for the drop traffic.
- BCN based architecture
  - Per RPR Node queuing to minimize head of the line blocking
  - For Class based queuing
    - Class based queuing for the add traffic for each node in the ring: 64*8 = 512 queues
    - Class based queuing for drop traffic
  - Per flow queuing
    - Queues for each flow: millions of flows
    - Queue for each flow for drop
- Class based Queuing Architecture
  - Avoids internode signaling
  - Single Class based queuing structure for add and pass traffic
  - Single class based queuing structure for drop traffic

Conclusion

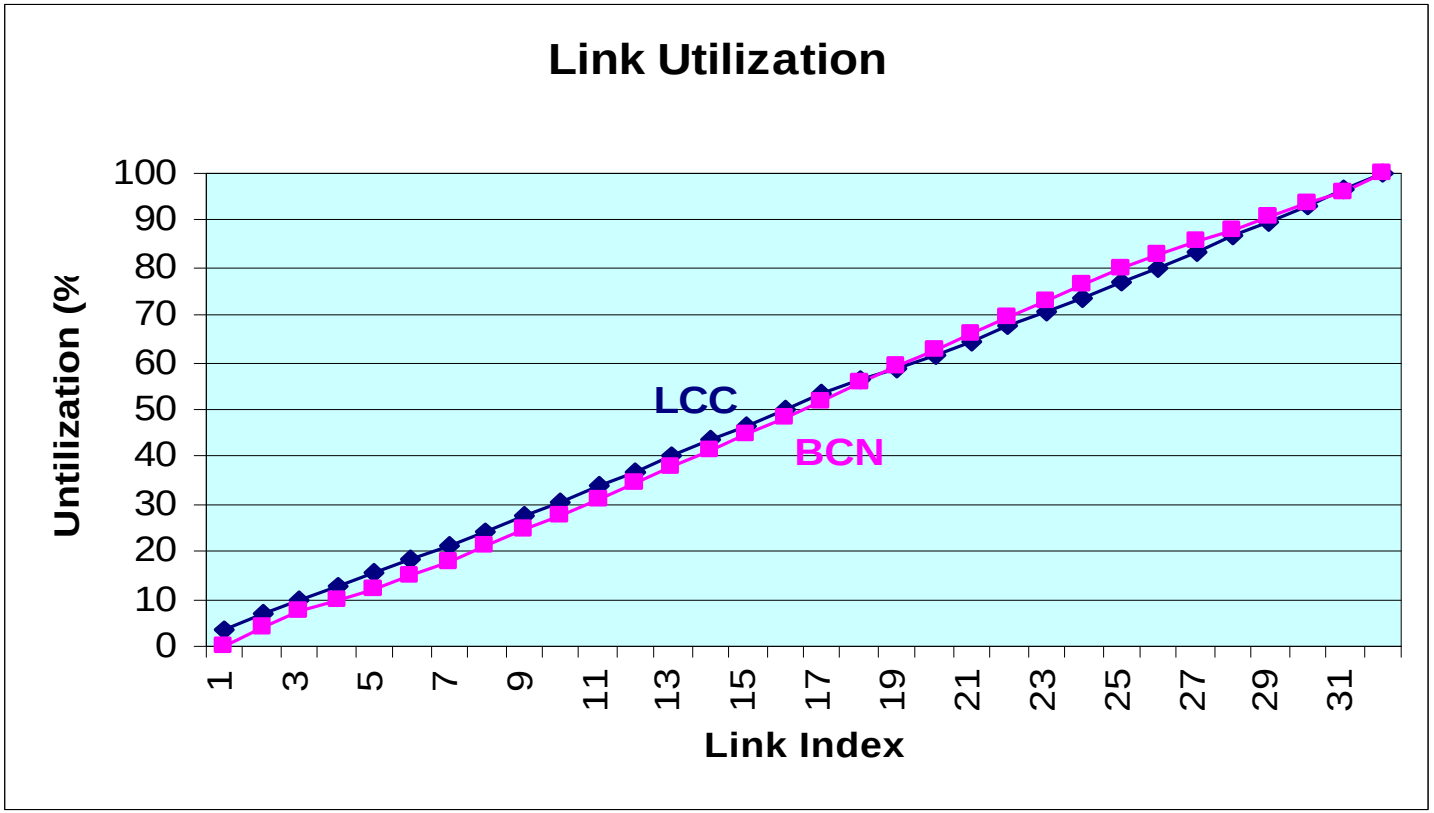
- Support many service scenarios
- Proposal not tied to a particular implementation that addresses only a set of needs.
- Allows vendor differentiation while insuring interoperability
- Hardware proposal supports most of the proposals with minimum cost of implementation

BCN vs. No BCN Traffic Simulation

Source:

Sink:

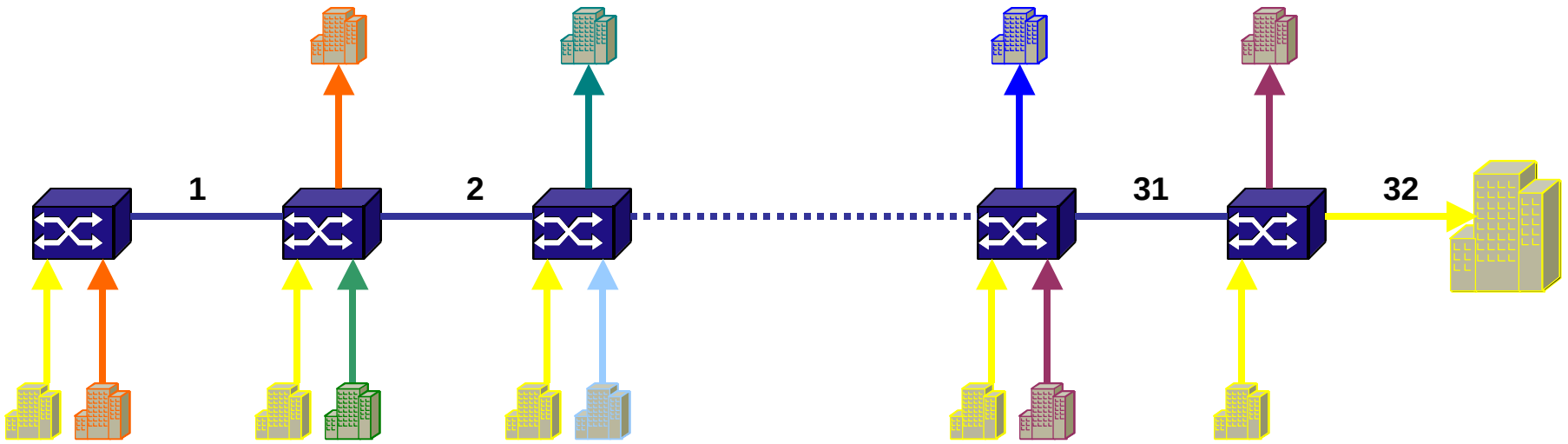
BCN vs. No BCN Traffic Simulation



BCN = Backward Congestion Notification

LCC = Local Congestion Control

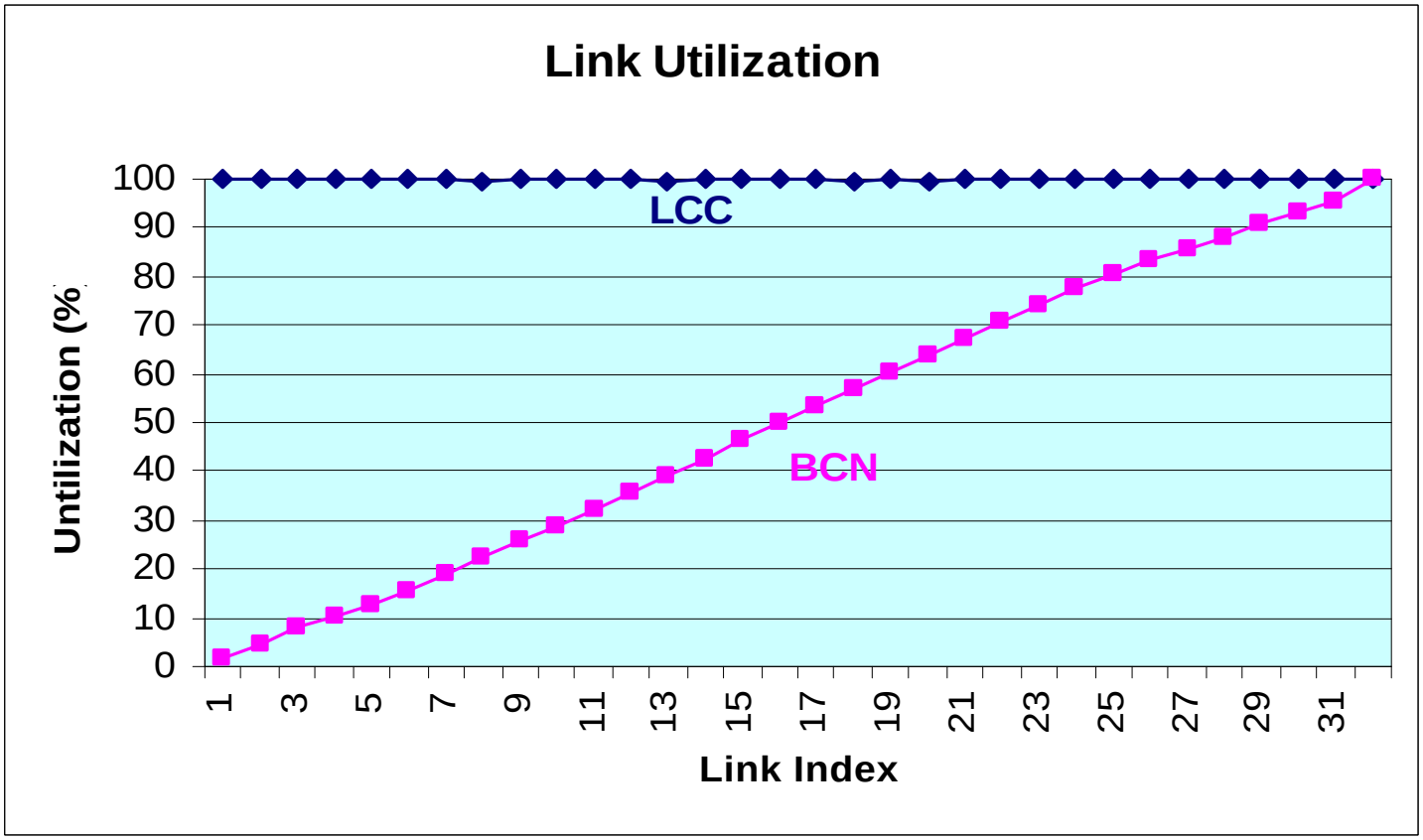
BCN vs. No BCN Traffic Simulation



Source:

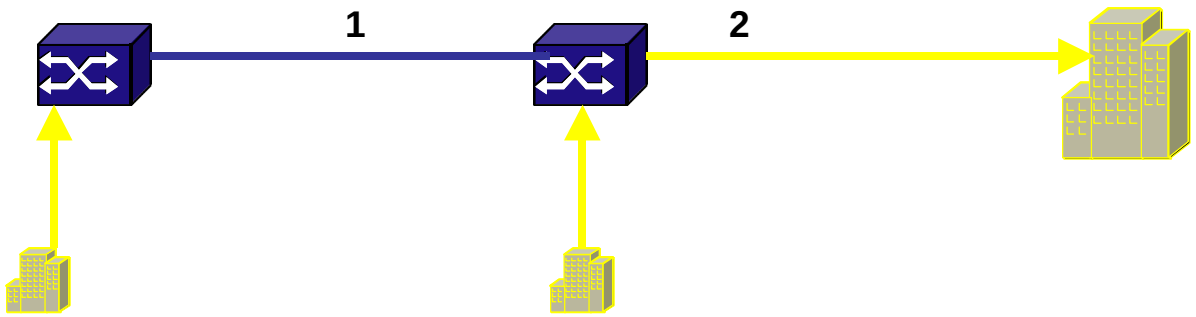
Sink:

BCN vs. No BCN Traffic Simulation



BCN = Backward Congestion Notification
LCC = Local Congestion Control

Cut through vs. Store and Forward



- Link 2 is congested
- Measure Max Q Delay for Host Traffic in Link 2
- Variable: Span Propagation Delay

Cut through vs. Store and Forward High priority Add Max Q Delay