CN-SIM: A Common Bridge Model

Davide Bergamasco (davide@cisco.com)

October 12th, 2006

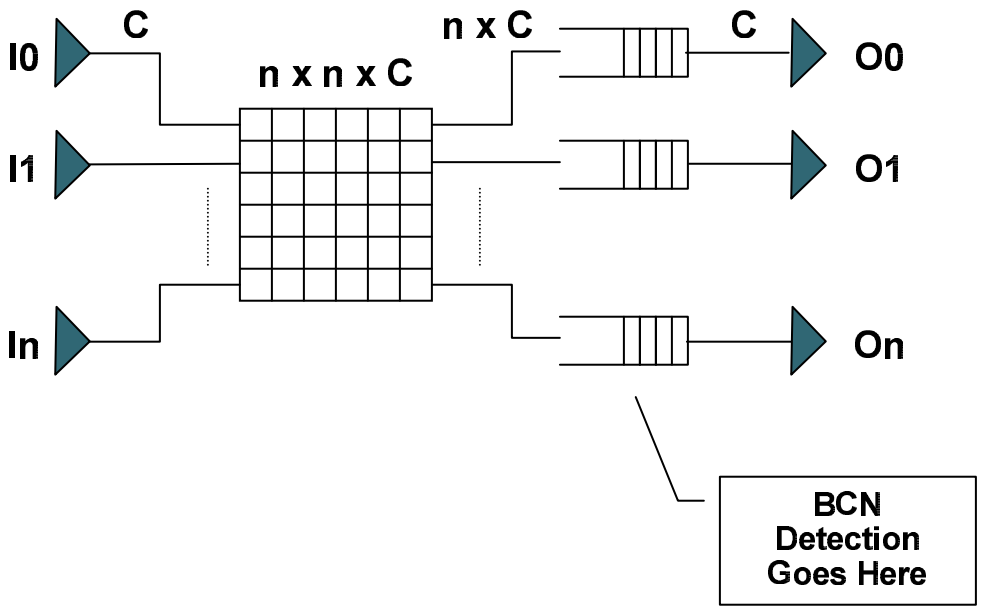
V. 2.0

Why a common bridge model?

- Our goal is to run independent simulations to validate BCN
- So far we have defined a set of common
  - Topologies
  - Traffic Patterns
  - Metrics
- But to ensure comparability of results, we also need a common bridge model

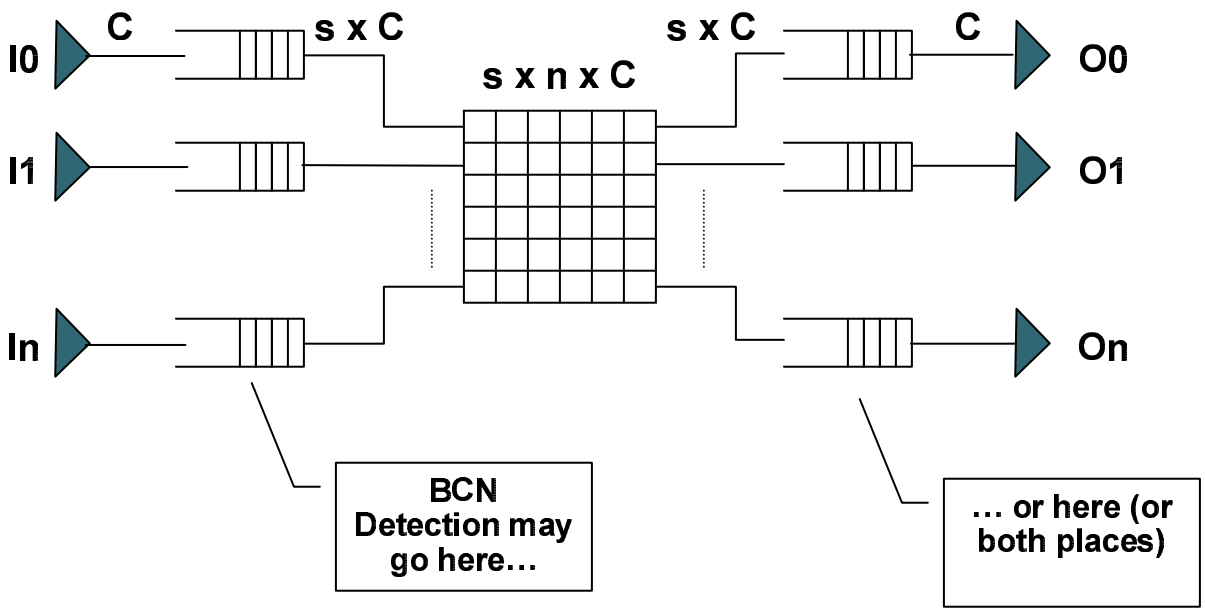
Possible Bridge Models

- **Output Queued Bridge (as defined by 802.1)**



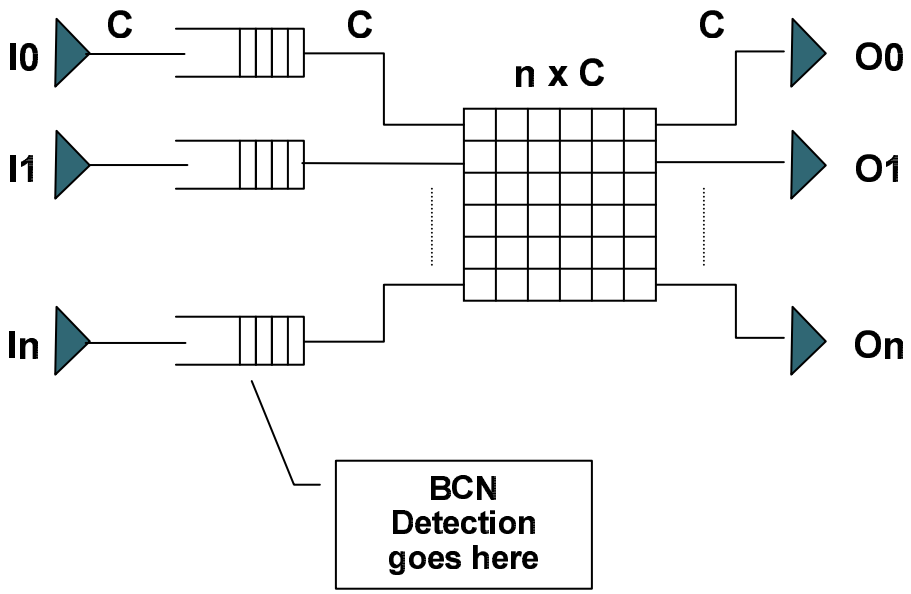
Possible Bridge Models

- **Combined Input-Output Queued Bridge (CIOQ)**



Possible Bridge Models

- **Input Queued Bridge (probably not interesting…)**



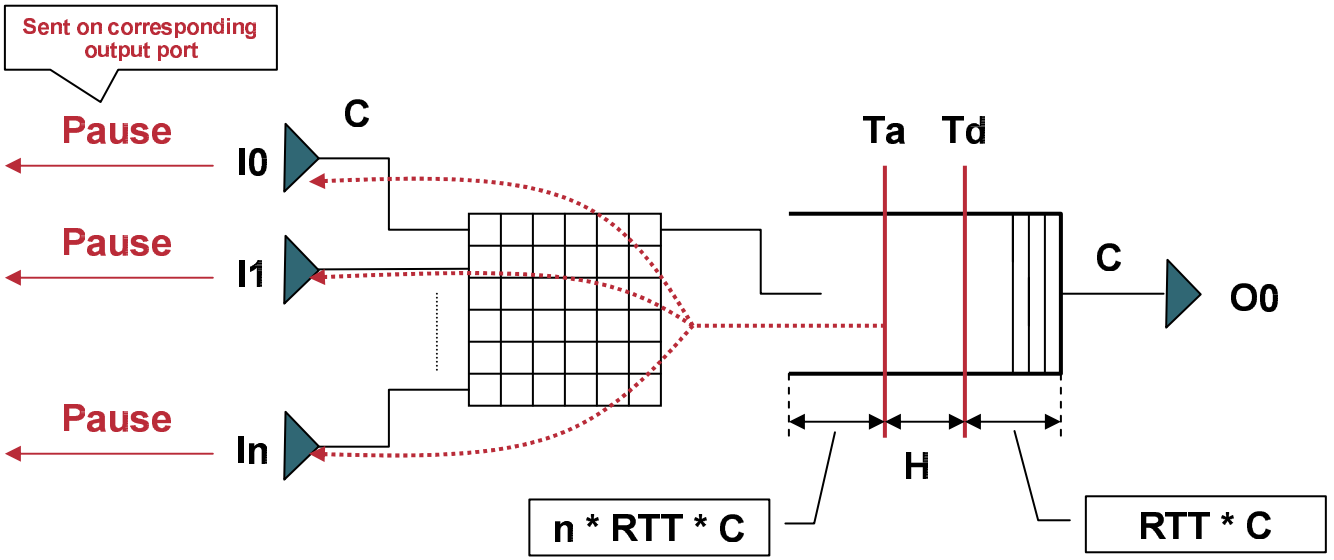
Possible Bridge Models

- **Other models?**
  - **Parallel Packet Switch**
  - **Modular**
  - **etc…**
- **Conclusion:**
  - **IEEE 802.1 defines an Output Queued bridge**
  - **CN-SIM must focus its effort on such architecture**
  - **Vendors will be responsible of mapping CN-SIM results on their architectures**

Back-pressure & Pause with OQ Bridges

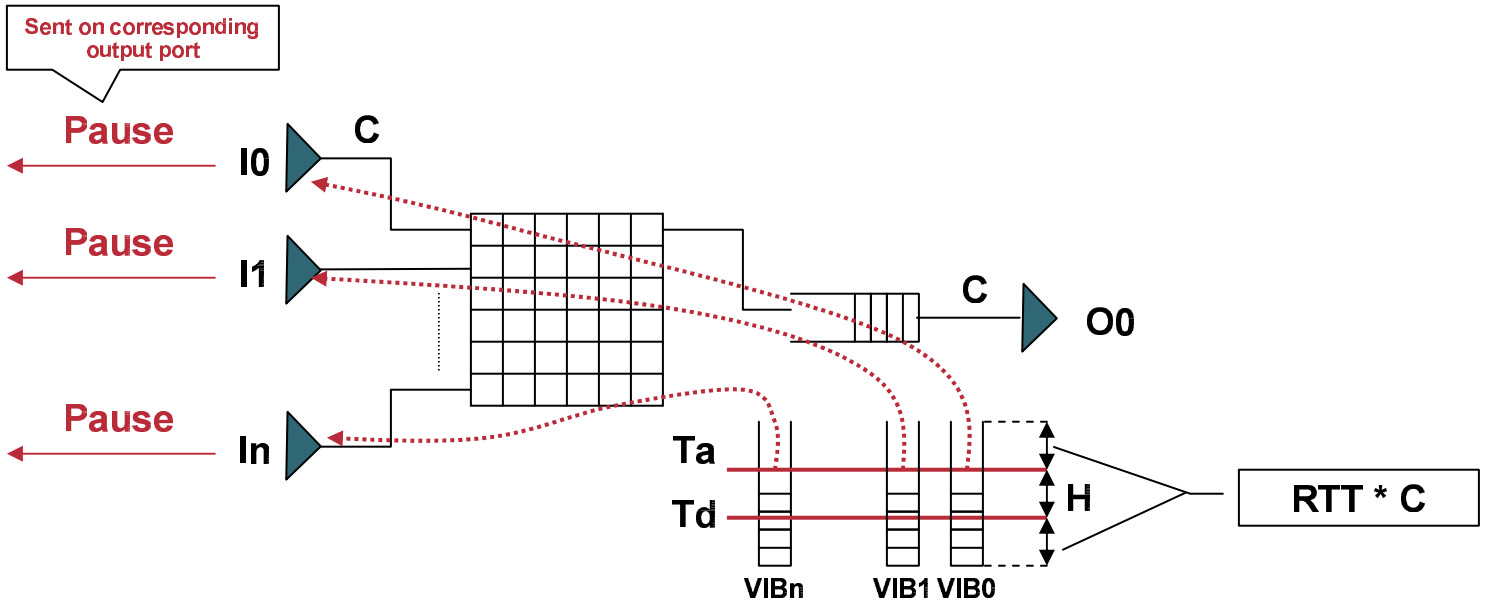
- Considering two alternatives for Back-pressure & Pause
  - 1 – Global Pause on single shared buffer
    - Two two thresholds: assert and de-assert
    - Global back-pressure signal sent to all input ports
    - All Inputs generate Pause simultaneously
  - 2 – Selective Pause on Virtual Input Buffers
    - Output buffer keeps a per-input consumption accounting
    - Each input is allowed to consume up to a certain amount of output buffer (Virtual Input Buffer)
    - On each VIB two thresholds are defined: assert and de-assert
    - Per-input back-pressure signal triggered by VIB occupancy
    - Each input generates Pause independently

Global Pause



- **Thresholds settings**
  - To avoid frame loss, Ta must be set such that backpressure is asserted when there is still enough buffer to store in-flight traffic, i.e., n * max (RTT * C, MTU)
  - To avoid throughput loss, Td must be set such that backpressure is de-asserted when in the buffer there is still enough frames to cover the un-Pause latency, i.e., max(RTT * C, MTU)
  - H = Ta – Td, size of the hysteresis cycle (H >= 2 * Qeq to make BCN work)
- **Minimum Buffer requirements**
  - B_{min} = (n + 1) * max(RTT * C, MTU) + H = (n+1) * max(RTT * C, MTU) + 2 * Qeq

Selective Pause



- **Thresholds settings**
  - To avoid frame loss, Ta must be set such that backpressure is asserted when there is still enough buffer to store in-flight traffic, i.e., max (RTT * C, MTU)
  - To avoid throughput loss, Td must be set such that backpressure is de-asserted when in the buffer there is still enough frames to cover the un-Pause latency, i.e., max(RTT * C, MTU)
  - H = Ta – Td, size of the hysteresis cycle (ΣH >= 2 * Qeq to make BCN work)
- **Minimum Buffer requirements**
  - B_{min} = 2n * max(RTT * C, MTU) + Σ H = 2 *[max(RTT * C, MTU) + Qeq]