Discussion of Objectives for Congestion Isolation

IEEE 802.1 Interim

Geneva

January 2018

Paul Congdon (Huawei)

Carmi Arad (Marvell)

Objective Categories

- Functionality
- Compatibility
- Performance
- Scale
- Implementation (Cost/Complexity)
- Manageability

Functional Objectives

- With high probability, identifies flows that are causing congestion
- Quickly adjusts transmission scheduling of offending flows
- Avoids head-of-line blocking by signaling to upstream neighbor to also adjust transmission scheduling.
- Reduces frequency of PFC usage to create lossless environments

Compatibility Objectives

- Works in legacy environments
  - Is not enabled unless peer bridges are compatible
  - Does not require network wide upgrade
- Works with existing PFC deployments
  - Does not require additional traffic classes
- Works in conjunction with end-to-end congestion control schemes (e.g. ECN, BBR, RoCEv2 CNM, QCN)
- Works with load balancing techniques

Performance Objectives

- Metrics to measure performance gains
  - Average flow completion time (mice vs elephants)
  - Reduction in pause time if PFC is enabled
  - Reduction in frame loss if PFC is not enabled
  - Reduction in number of victim flows from HoLB
  - Reduction in overall congestion signaling
  - Increased link utilization

Correctness Objectives

- Does not increase probability of out-of-order packets
- Does not increase deadlock vulnerabilities
- Provides mechanisms to avoid starvation
- Resilient to loss of control messages

Scale Objectives

- Works in arbitrary data-center topologies with a mix of link speeds
- Limits messaging overhead
  - Does not require message propagation beyond hop-by-hop
  - Does not increase frequency of messages over existing approaches (e.g. QCN)
- Limits flow table size requirements
  - Flow entries are aged
  - Only offending flows are required to be stored

Implementation Objectives

- Limits impact on traffic selection implementations
  - Can be implemented with existing algorithms, but can be improved with specified enhancements
- Limits buffer size growth requirements
  - Can be implemented using existing traffic classes and buffer allocations
- Limits flow table size requirements
  - Can be implemented by only registering offending flows in flow table

Management Objectives

- Limits configuration requirements
  - Does not require additional tuning
- Provides auto discovery of peer capability
  - LLDP CI Discovery TLV