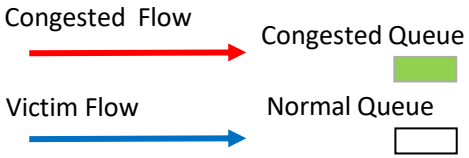
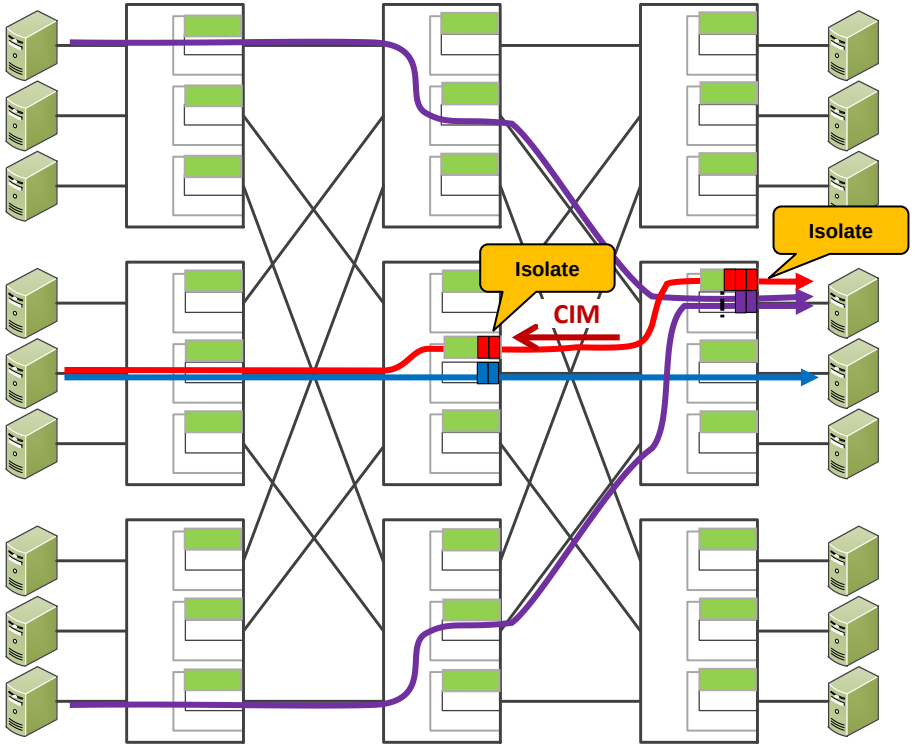
Next Steps for Source PFC and/or Source Flow Control

Paul Congdon
JK Lee
Lily Lv
December 23, 2021

Future 802.1 Congestion Management Tools



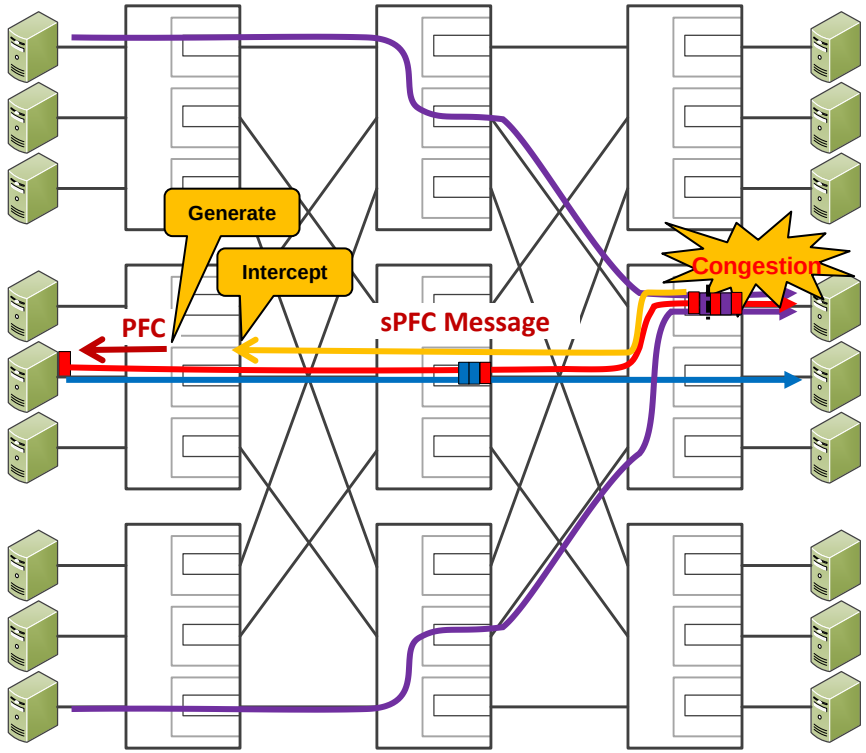
P802.1Qcz - Congestion Isolation



Implementation details

- Congesting flows are isolated locally first
- As queues continue to congest, CIM is generated and sent to upstream bridge/router
- CIM can be L2 or L3 message to support L3 networks (common deployment model).

Source PFC

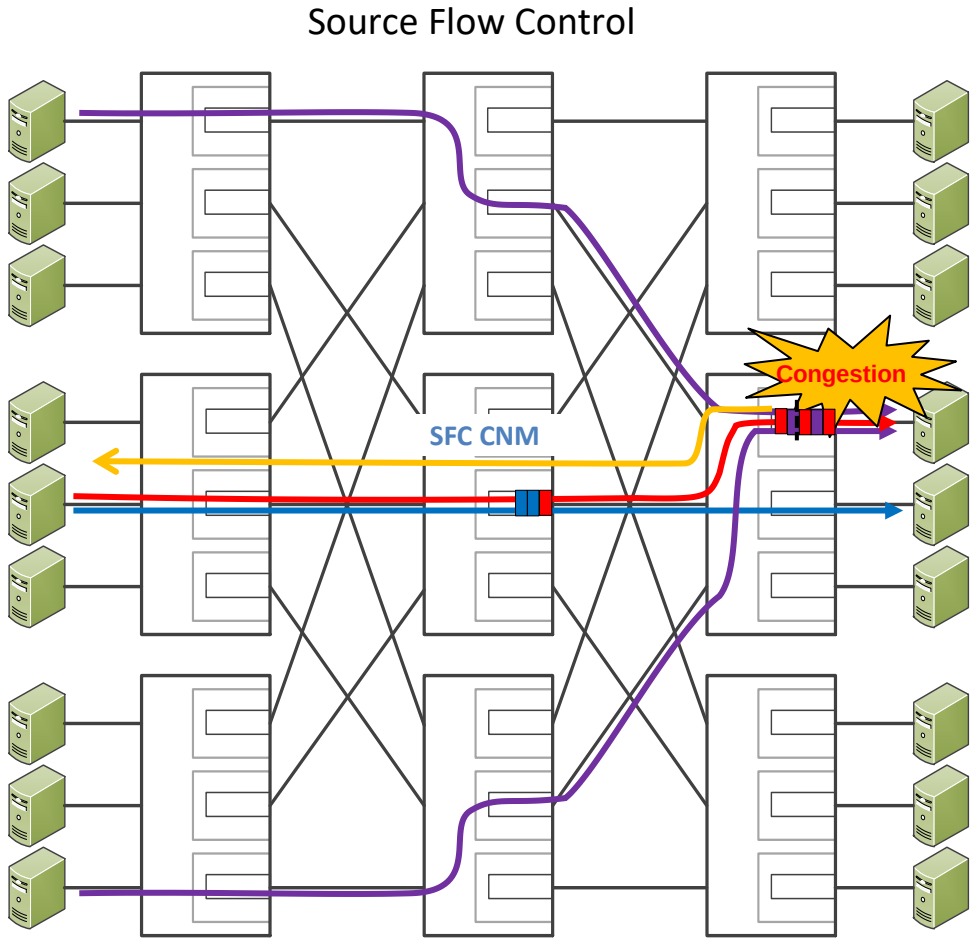


Details

- Can be combined with Congestion Isolation
- If congestion persists, Edge-to-edge signaling using L3 message
- Existing PFC generated at last hop
- NOTE: signaling message could pass to end-station directly if supported.

Source PFC vs Source Flow Control

- sPFC = remote generation of PFC at the source ToR
- SFC = pause at the flow level
- sPFC signaling message direct to end-point
- SFC/sPFC in the perspective
  - SFC/sPFC backward signaling ~= a L3 version of 802.1Qau (L3-QCN)
  - SFC/sPFC control = pause/flow control unlike QCN congestion control
  - NOTE: RoCEv2 DCQCN is a L3 adoption of QCN, using the ECN end-to-end congestion control loop



Progress To Date

- Public presentations of the concept and data at P4 Workshops (Apr'20, May'21) and Open Fabrics Alliance (Mar'21)
  - https://opennetworking.org/wp-content/uploads/2020/04/JK-Lee-Slide-Deck.pdf (slide 12)
  - https://www.openfabrics.org/wp-content/uploads/2021-workshop-presentations/503_Lee_flatten.pdf
  - https://opennetworking.org/wp-content/uploads/2021/05/2021-P4-WS-JK-Lee-Slides.pdf (slide 14)
- Previous Nendica presentations
  - https://mentor.ieee.org/802.1/dcn/21/1-21-0055-00-ICne-source-flow-control.pdf - 9/16/2021
  - https://mentor.ieee.org/802.1/dcn/21/1-21-0061-00-ICne-source-remote-pfc-test.pdf – 10/14/2021
  - https://mentor.ieee.org/802.1/dcn/21/1-21-0067-00-ICne-source-remote-pfc-status-update.pdf - 11/04/2021
  - https://mentor.ieee.org/802.1/dcn/21/1-21-0077-00-ICne-consideration-of-spfc-sfc-issues-when-leveraging-qcz.pdf - 12/16/2021
- IETF Awareness
  - Topic raised at IEEE 802 / IETF Coordination call – 10/25/2021
  - https://datatracker.ietf.org/meeting/112/materials/slides-112-iccrg-source-priority-flow-control-in-data-centers-00 - 11/08/2021
- Related References
  - Lui, S; Ghalayini, A; Alizadeh, M; Prabhakar, B; Rosenblum, M; and Sivaraman, A, "Breaking the Transience-Equilibrium Nexus: A New Approach to Datacenter Packet Transport", 18th USENIX Symposium on Networked Systems Design and Implementation (NSDI 21), https://www.usenix.org/system/files/nsdi21-liu.pdf

Planned Next steps

- Ongoing technical discussions
  - ✓ Address issues related to leveraging Qcz – see Nendica 12/16/21
- Analysis of impact on 802.1Q for an amendment
  - \+ Clause by clause examination of anticipated changes
- Move to authorize PAR & CSD development after March 2022 Plenary

January Interim Plans?

- Continue Nendica presentation or Initial TSN presentation or Both?
- Choices:
  1. Nendica presentation on 802.1Q Amendment changes
  2. Longer TSN introductory presentation including 802.1Q Amendment changes
  3. Separate Nendica presentation on amendment changes and shorter Introductory presentation to TSN

Contributor preference: 2 or 3 – in order to prepare for potential March motion, time to expand the technical audience.