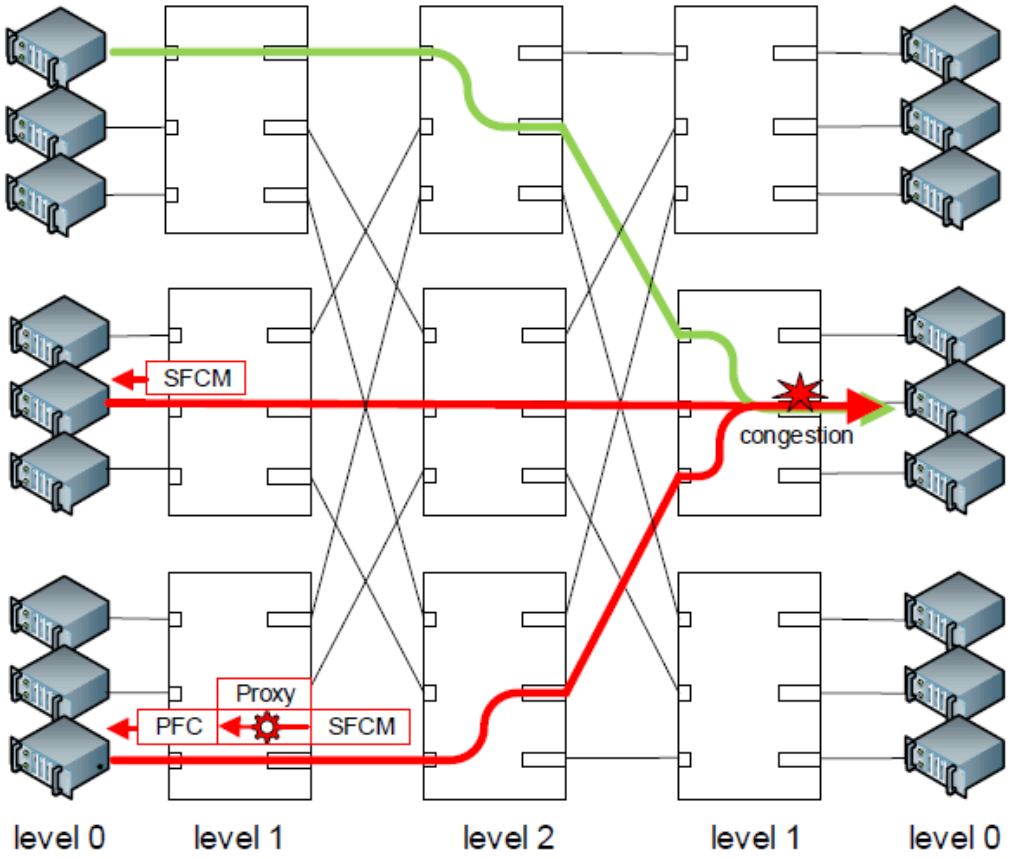
Source Flow Control, Recap and restart?

Lihao Chen (lihao.chen@huawei.com)

HUAWEI

Recap

- Discussion started: 17 Jan 2022
  - > HPE, Huawei, Intel, …
  - > Congestion, in particular incast (AI DC networks)
- PAR approved: 21 Sep 2022
  - > Scope: … for the signaling and remote invocation of flow control at the source of transmission in a data center network… to allow bridges at the edge of the network to intercept and convert signaling messages to existing Priority-based Flow Control (PFC) frames…
  - > Precise PFC, quick reaction, easy adoption.
- Last updated: 19 Jan 2023



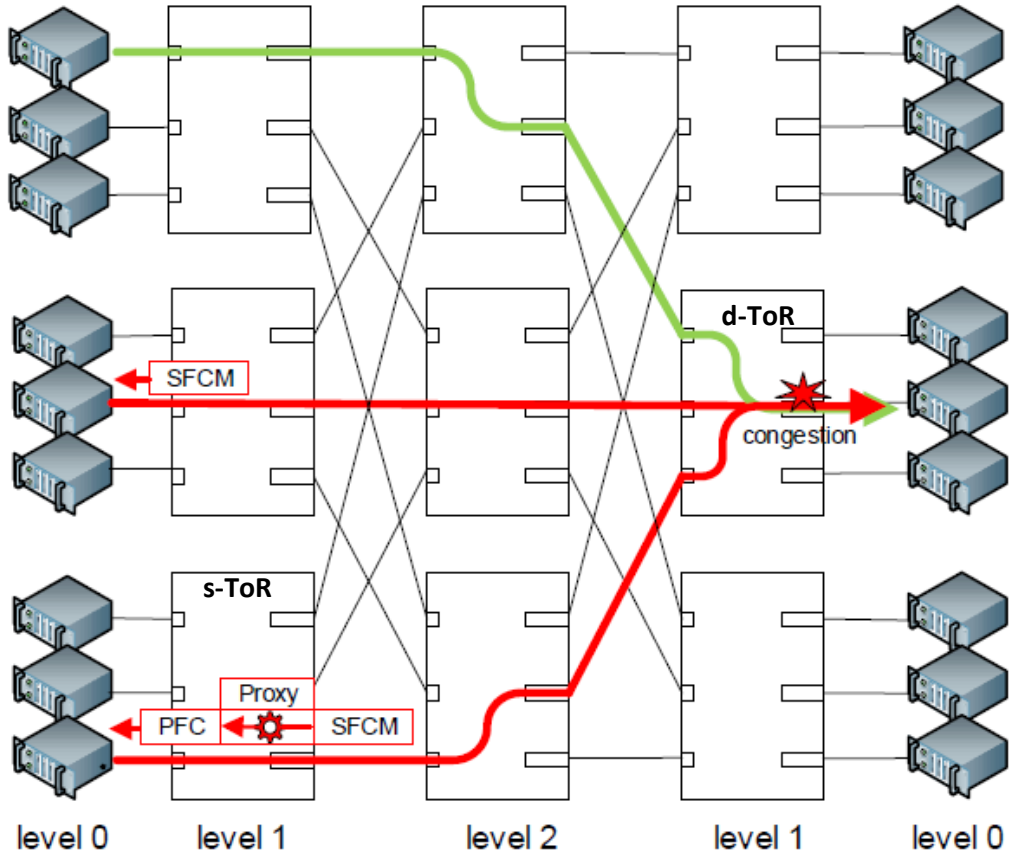
https://www.ieee802.org/1/files/public/docs2022/dw-congdon-individual-text-1122-v01.pdf

HUAWEI

Technical overview

- d-ToR (with SFC enabled):
  > After congestion detection, using the source IP address of the congestion packet as the destination IP address to send the SFCM. Flow control information including the class and pause duration are encapsulated in the SFCM.

- s-ToR (with SFC enabled):
  > Identify the SFCM (e.g., by a specific UDP port Num.) and decapsulate it to get flow control information. Then translate it into a PFC frame and send to the sender server.



https://www.ieee802.org/1/files/public/docs2022/dw-congdon-individual-text-1122-v01.pdf

d-ToR: destination Top of Rack
s-ToR: source Top of Rack

HUAWEI

Restart

- The author wishes to continue advancing this project.
  > The need for the project still stands.
  > Detailed text contributions will be provided.

- Thoughts?