

Data Center Congestion Control – Where's the best fit in IETF/IRTF?

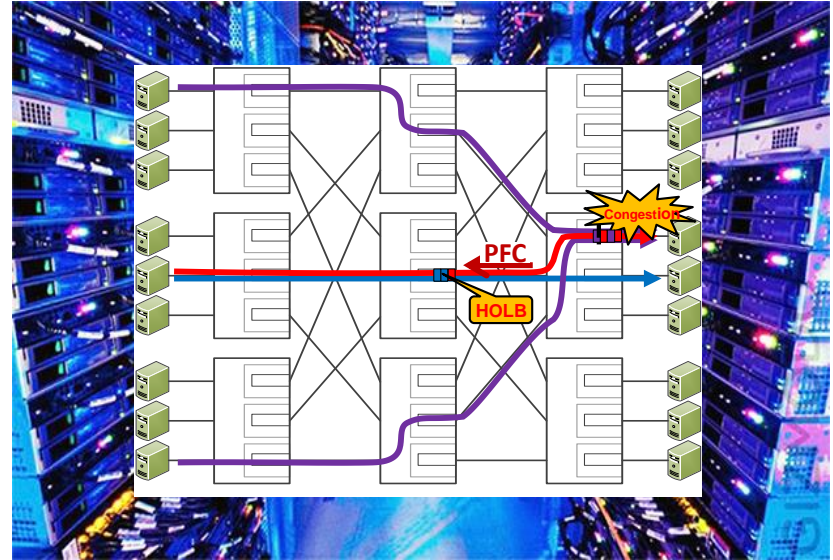
Paul Congdon (Tallac Networks)

Data center congestion is unique

The Internet



The High-Performance Data Centers



Data centers have...

- A much different bandwidth-delay product
- Different switch implementations and buffer configurations than Internet Routers
- More homogeneity with the network design and topology
- A high concentration of high-speed links, compute and storage
- Different traffic profiles with a higher degree of correlation
- Fewer management domains (typically a single management)

Congestion in the DCN environment is different than in the Internet

Where to consider DCN CC Research/New-Work

- ICCRG Charter can be interpreted to include DCN
 - “...The ICCRG may also consider congestion and protocol performance problems in general IP networks, i.e., not only on the global Internet. One example of such IP networks are multi-tenant, heterogeneous datacenters,...”
- Congestion control work is on-going in TSVWG
 - However, nothing particularly DCN focused
- Perhaps a new IRTF group is appropriate
- Let’s discuss this and status of contributions in our side-meeting

Questions about Congestion Control

in the HPC/RDMA/AI DataCenter Network

- What is needed from NICs for better CC?
 - An open framework to negotiate capabilities and algorithms – OpenCC
 - <https://datatracker.ietf.org/doc/draft-zhuang-tsvwg-open-cc-architecture/>
- How can the Network participate?
 - An AI model for parameter tuning
 - <https://datatracker.ietf.org/doc/draft-zhuang-tsvwg-ai-ecn-for-dcn>
 - Fast feedback from the network
 - <https://tools.ietf.org/html/draft-even-iccr-g-dc-fast-congestion-00>
- Other interesting topics
 - Performance metrics for HPC/RDMA/AI networks

Join us for further discussion

- Non-WG IETF Mailing list rdma-cc-interest@ietf.org
 - Subscribe at:
<https://www.ietf.org/mailman/listinfo/rdma-cc-interest>
- Side Meeting: Tuesday 8:30AM – 9:45AM – VIP-A
 - NOTE on side meetings:
 - Open to all
 - Meeting minutes will be posted to rdma-cc-interest@ietf.org
 - Not under NDA of any form

Fast feedback from the Network

- <https://tools.ietf.org/html/draft-even-iccr-g-dc-fast-congestion-00>
- Describes the current state of flow control and congestion for Data Centers and proposes future directions.
- Questions for discussion
 - Is the current IOAM approach sufficient for CC in the DCN?
 - How can the network provide more information about congestion state along the path?
 - How does the network represent signals for sender driven CC?
 - How to notify reaction points as soon as possible?

NIC and Network CC negotiation

- <https://datatracker.ietf.org/doc/draft-zhuang-tsvwg-open-cc-architecture/>
- An open framework to negotiate capabilities and algorithms – OpenCC
- Questions for discussion
 - Can we build a transport agnostic congestion controller?
 - What is the method for negotiating capabilities and supported algorithms?

AI Models for CC Parameter Tuning?

- <https://datatracker.ietf.org/doc/draft-zhuang-tsvwg-ai-ecn-for-dcn>
- An AI framework to configure and tune the CC parameters in the DCN
- Questions for discussion
 - What performance measures should be input to the model?
 - What triggers the need to reconfigure and tune?
 - What are the CC parameters to adjust?