

# HPCC++: Enhanced High Precision Congestion Control

[draft-pan-tsvwg-hpccplus-00](#)

Rui Miao, Hongqiang Harry Liu, Rong Pan, Jeongkeun Lee, Changhoon Kim, Barak Gafni, Yuval Shpigelman

IETF-108 tsvwg

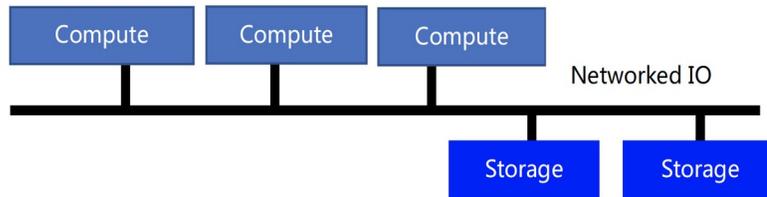
July 2020

# Cloud desires hyper-speed networking

Today, clouds have

- bigger data to compute & store
- faster compute & storage devices
- more types of compute and storage resources

## High-performance storage



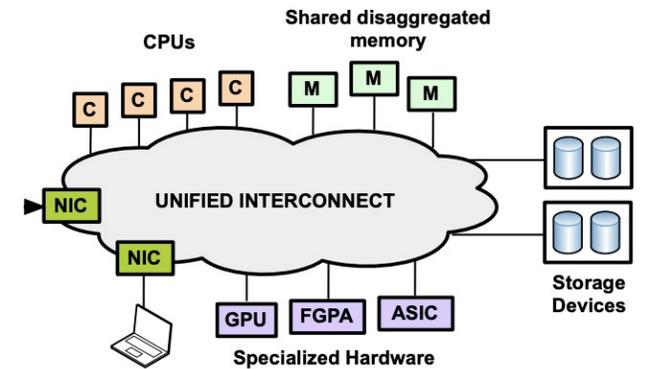
- Storage-compute separation is norm
- HDD → SSD → NVMe
- Higher-throughput, lower latency
- 1M IOPS / 50~100us

## High-performance computation



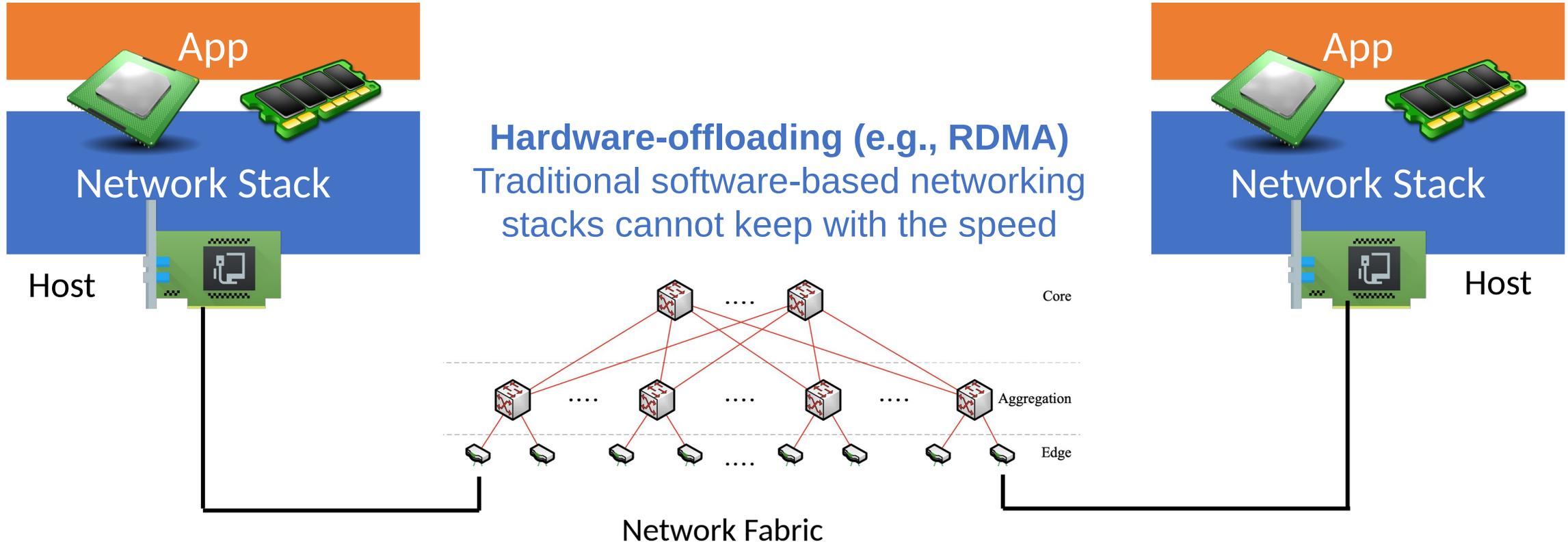
- Distributed deep learning, HPC
- CPU → GPU, FPGA, ASIC
- Faster compute, lower latency
- E.g. latency < 10us

## Resource disaggregation



- More network load
- Need ultra-lower latency: 3-5us, > 40Gbps (Gao Et.al. OSDI'16)

# Hyper-speed network chips != hyper-speed networking



**Hardware-offloading (e.g., RDMA)**  
Traditional software-based networking stacks cannot keep with the speed

## **Congestion control (CC)**

Since, end hosts are aggressive, network is more vulnerable to congestion & packet loss

# Realistic challenges in current CC in RDMA networks

- Operation challenge-1: PFC storm & deadlock
  - **Disabling PFC causes bad performance!!!**

- Operation challenge-2: running multiple applications
  - **QoS queues are scarce resources!!!**

- Operation challenge-3: complex parameter tuning
  - **DCQCN has at least **15** parameters to tune!!!**

## Challenges in current CC

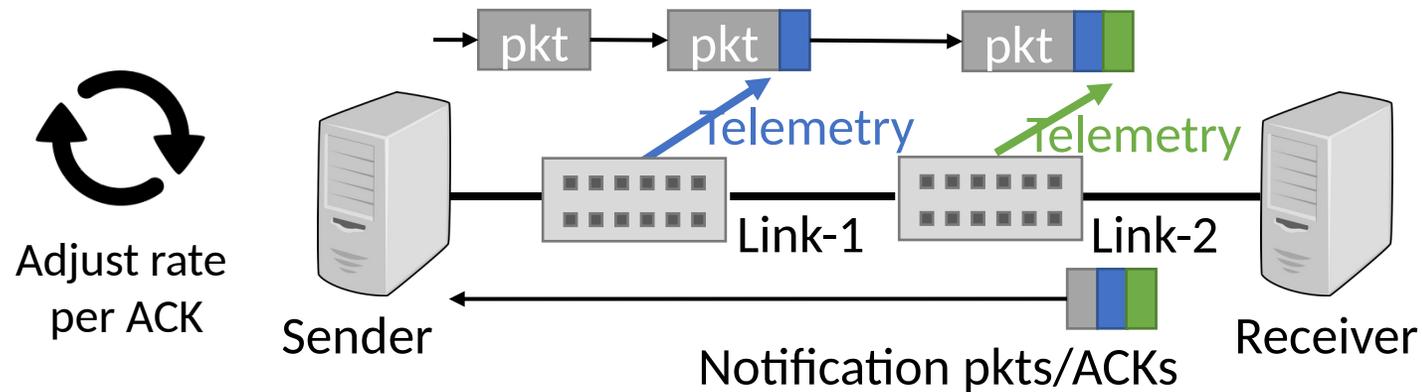
Challenge-1:  
Slow Convergence

Challenge-2:  
Standing queue

Challenge-3:  
Heuristics in CC

# HPCC++: Enhanced High Precision Congestion Control (SIGCOMM'19)

- New commodity ASICs have in-band telemetry ability
- Use in-band telemetry as precise feedback for congestion control



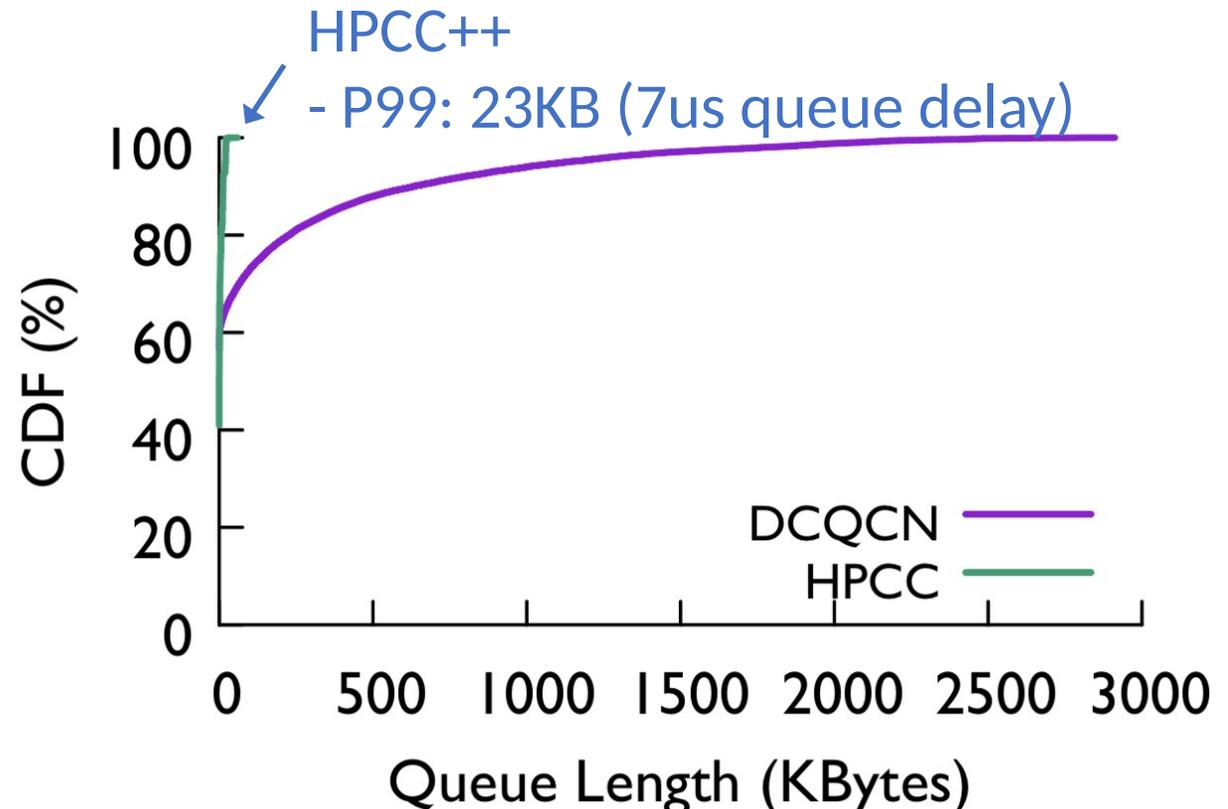
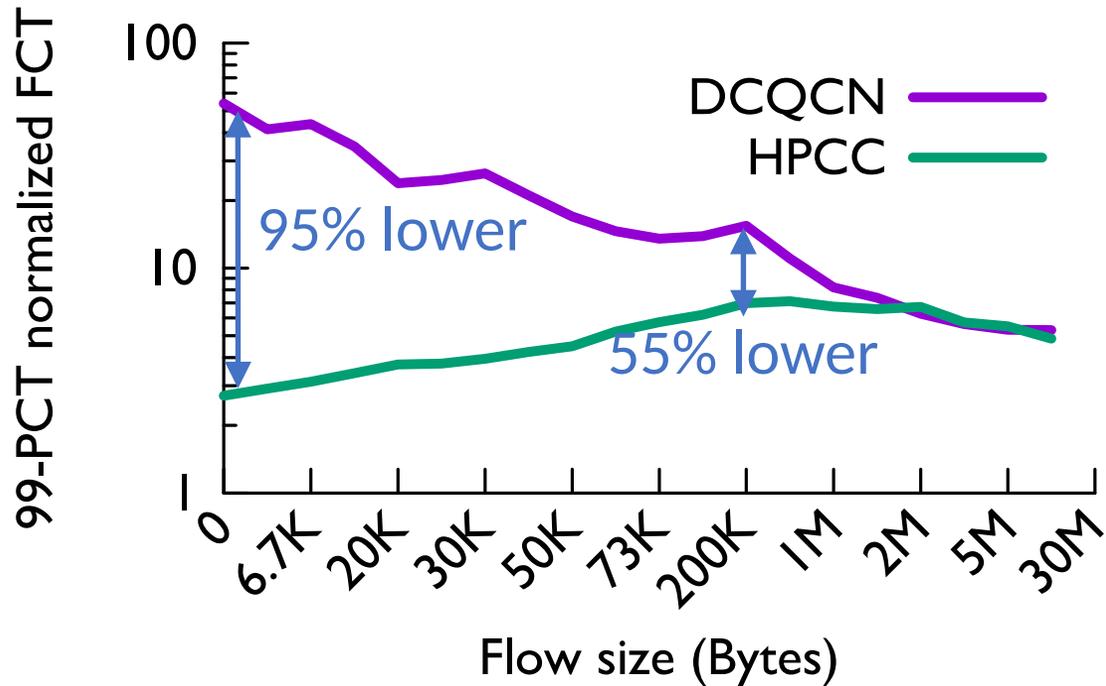
# HPCC solves the 3 problems

## Using INT as the precise feedback

- **Fast convergence**
  - Sender knows the precise rate to adjust to, on every ACK
- **Near-zero queue**
  - Feedback does not rely on queue
- **Few parameters**
  - Precise feedback, so no need for heuristics which requires many parameters

# HPCC++ achieves lower FCT and near-zero queue

- In testbed, vs. DCQCN (hardware-based, widely used in industry)
  - Web search traffic at 50% load
- Vs. other CC (unavailable in HW) in simulation. HPCC performs better



Thank You