RoCEv2 Enhancements for Large Scale Deployments

Hemal V. Shah and Moshe Voloshin

Data Center Solutions Group (DCSG), Broadcom Inc.

Agenda

- Ethernet for HPC
- MPI and Communication Topologies
- PFC Challenges
- RoCEv2 with Congestion Control
- Congestion Control Enhancements for at Scale Deployment

BROADCOM

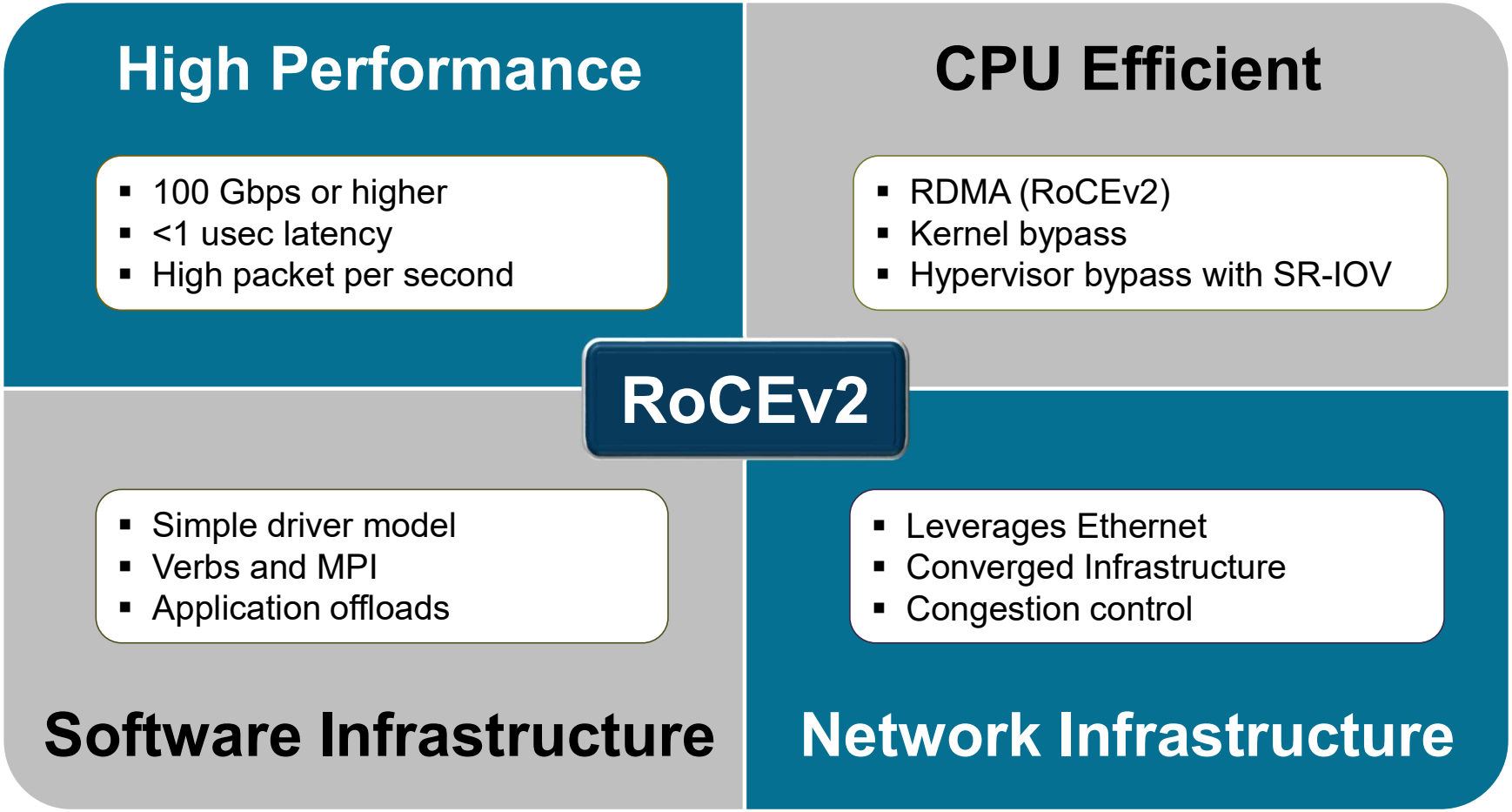
Why Ethernet for HPC? – High Level Analysis

Attribute	InfiniBand	Ethernet	Notes
Scalability (# of Nodes)	1000s	1000s	Ethernet as a fabric scales
Throughput	200G	200G	Ethernet network speeds scale
E2E Latency (B2B) – ½ RT	600 ns	1 – 1.2us	Unloaded RoCE latency < 2x IB latency
Switch Latency	~100ns	200-500ns	ETH switch latency in future ~200-300ns
Transport/Network	IBA Transport	RoCEv2/UDP/IP	Ethernet ecosystem not as limited as IB
Verbs API	OFA OFED	OFA OFED	Common host interface
Congestion Management	FECN/BECN	ECN	Standards based congestion control
Availability	Limited sources	Many sources	Large number of Ethernet suppliers

- **Ethernet has advantages in its ubiquitous deployment → Ethernet is Everywhere**
- **Ethernet ecosystem is mature and standards based → Benefits HPC markets**
- **100/200G Ethernet latencies (w/ RoCE) meets HPC requirements**

BROADCOM

RoCEv2 for HPC



BROADCOM

MPI and Communication Topologies

- MPI is widely used in HPC/ML clusters as the communication layer
- A process group in MPI represents a collection of processes
- The number of processes can be 100s per node
- The number of nodes can scale to 1000s in a cluster
- The communication pattern of processes is represented by a logical topology
  - Ring, Binary cube, Tree, etc.
- Selection of logical topologies depends on applications and communication libs
- MPI collectives (Gather, Reduce..) can create congestion in the network

Challenges with PFC without Congestion Control

- Priority Flow Control (PFC) is used for lossless service
- PFC is a point-to-point protocol between two Ethernet endpoints
- PFC can result in congestion spreading
- PFC can create PFC storm due to slow receivers
- PFC may result in transport live-lock

Congestion Control (CC) with RoCEv2

- ECN based CC schemes do not require any additional infrastructure support
- Congestion control without PFC can be sufficient for most of the workloads
- CC with PFC addresses PFC storms & live locks and preserves lossless service
- Even w/ large number of competing flows switch egress queue peak levels are low
- Reaction by sender is quick – few 10s of micro-seconds due to low queue level
  - Even with low marking threshold, network utilization is high
- Low marking threshold delivers low end-to-end latency with minimum interference
- Low marking threshold leaves majority of switch buffer for incast absorption
- Both probabilistic and deterministic marking is possible

BROADCOM

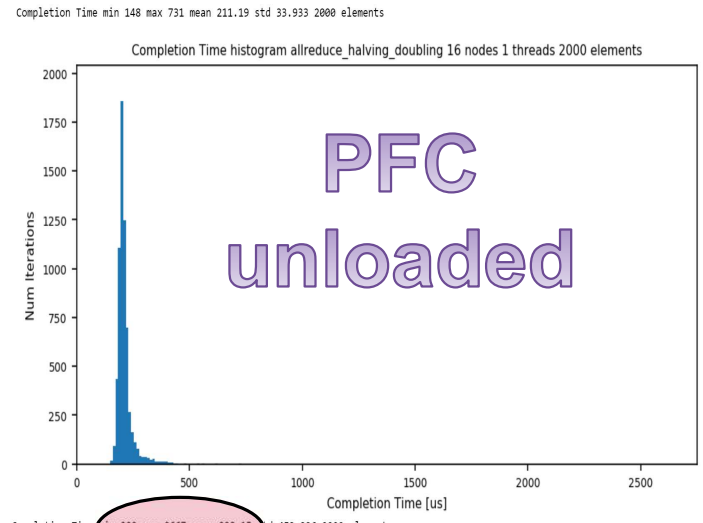
RoCEv2 Application Performance Under Congestion

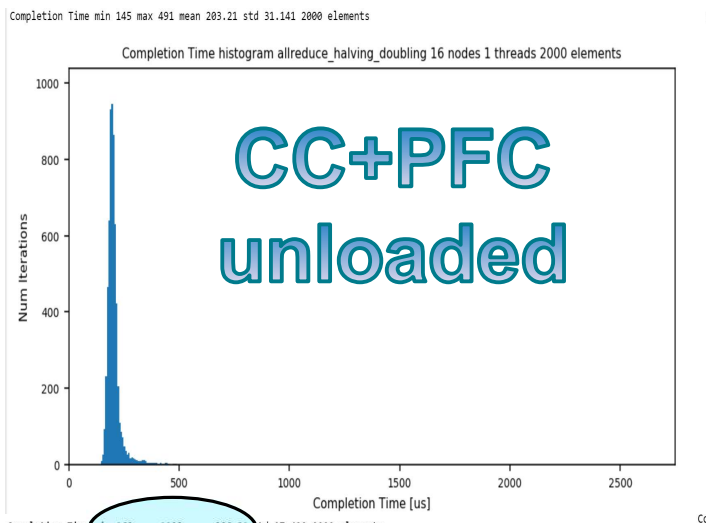
Test Scenario	Overview
Gloo	Open-source collective communication library using for ML applications, includes barrier, broadcast and allreduce Test setup • Two gloo allreduce_halving_doubling tests are run concurrently • One gloo test uses small data size (2000 elements or ~8KB) • Second gloo test uses large data size of 5M elements of ~20MB • Second gloo instance causes network congestion Results are shown for the first gloo instance • Unloaded completion time (running first instance only) • Loaded completion time - Impact of congestion from second gloo instance Modes: PFC only, DC-ROCE, DC-ROCE+PFC
GPCNeT (modified)	Global Performance and Congestion Network Test MPI test designed to measure relative performance under load and congestion Designed for large multi-layer switch network • 20% of nodes w/ test tasks: allreduce, p2p latency, random ring neighbor exchange • 80% of nodes assigned with congestor tasks: All2All, incast, RMA put and get • Nodes would share switch buffering resources and will cross path between switches Results were collected using single switch • The ratio of loaded to baseline for p50 and p99 is reported • Code was modified to make the impact of congestors on test tasks significant

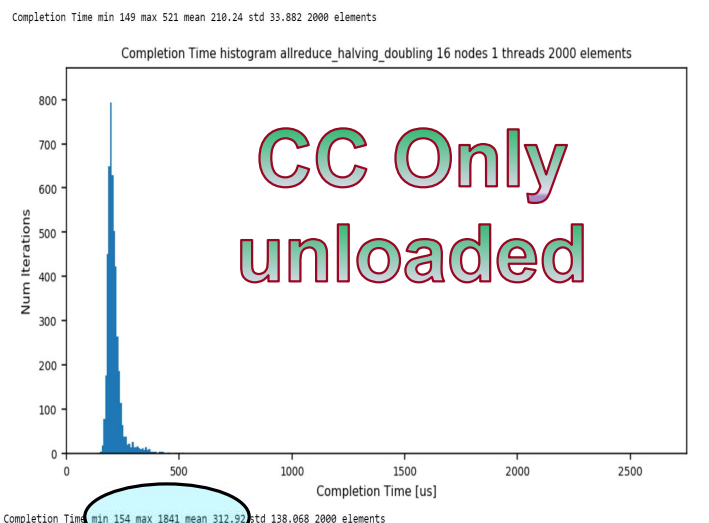
Testing is done in the presence of network congestion

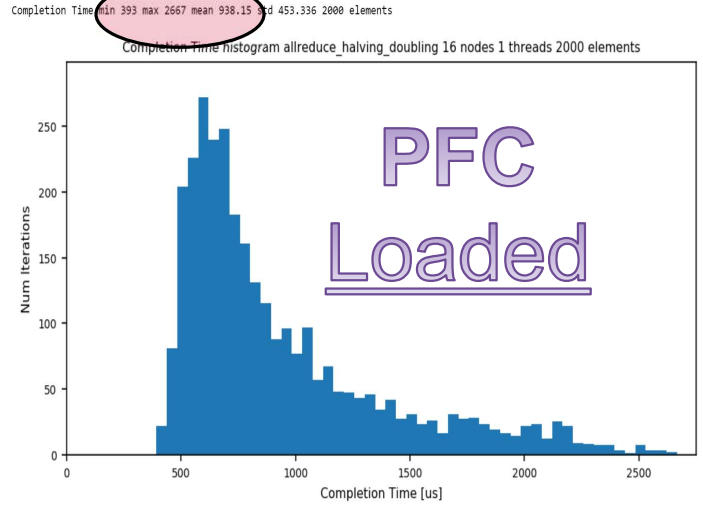
BROADCOM

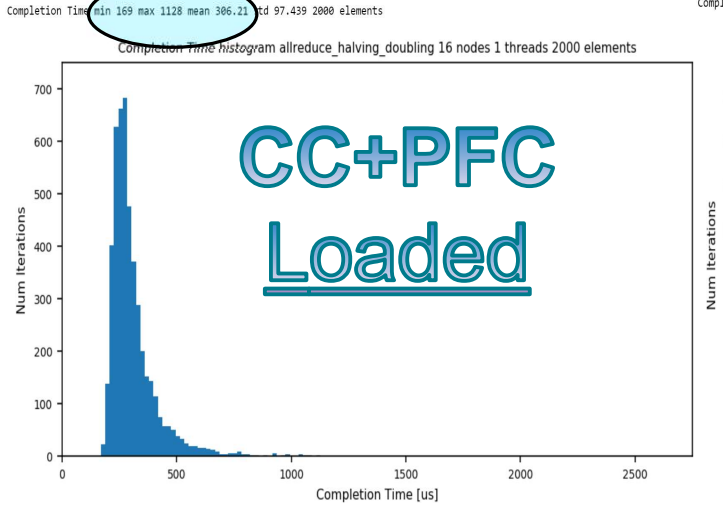
Inter Thread/App Interference due to High End-to-End Latency for Gloo

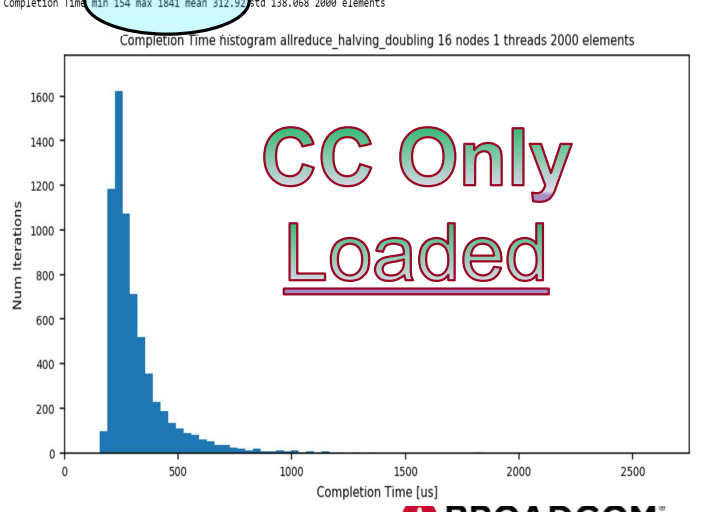






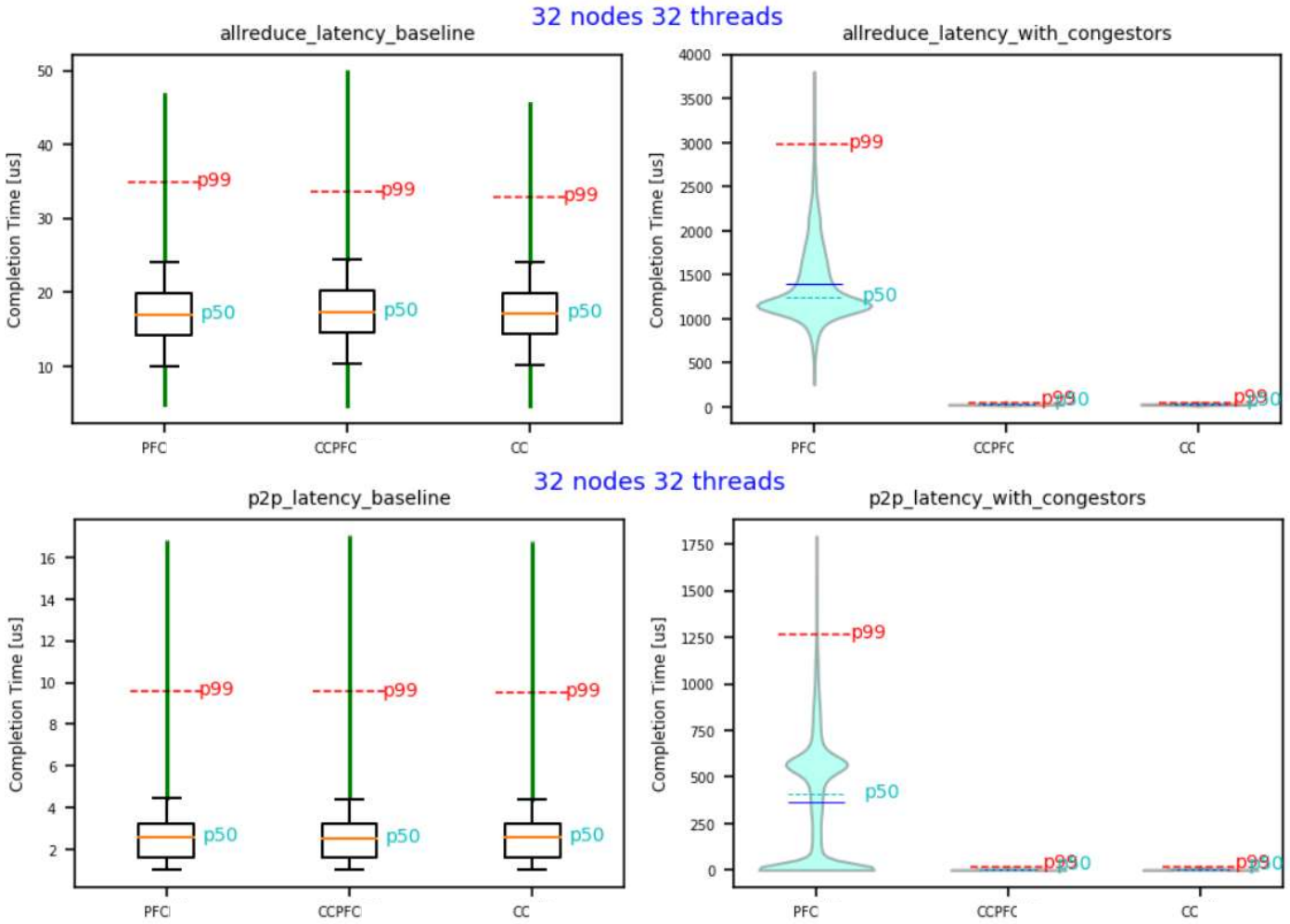






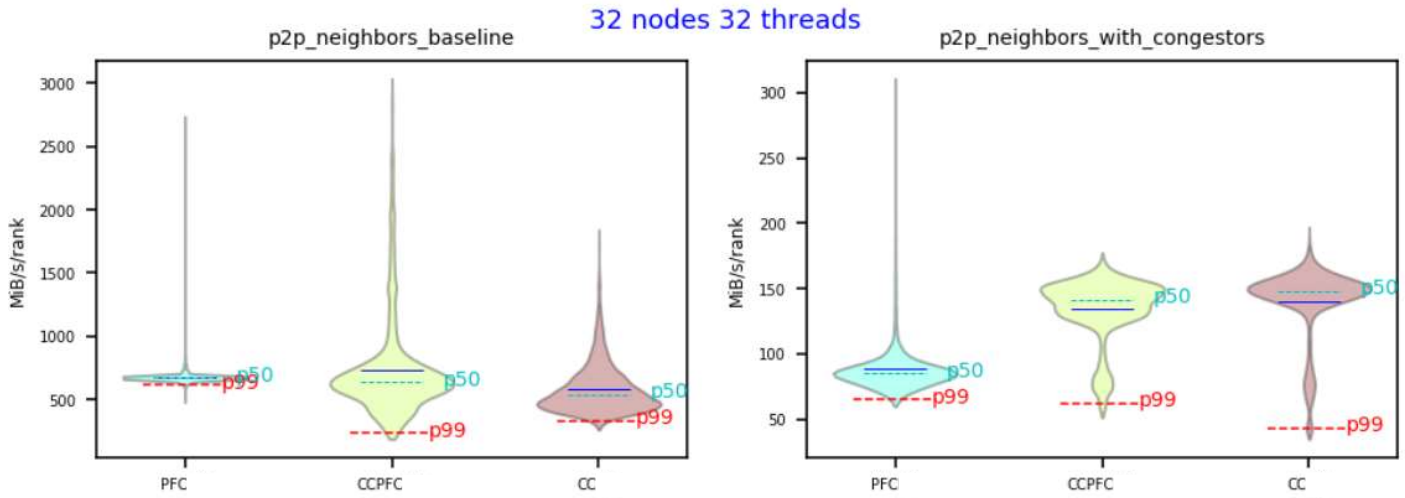
BROADCOM

GPCNeT results – Results on 32 Nodes



BROADCOM

GPCNeT results – Results on 32 Nodes



CCONLY CC+PFC PFC

Network Tests running with Congestion Tests - Key Results			ith Congestion Tests - Key Results		ith Congestion Tests - Key Results	
Name	Congestion		Congestion Impact		Congestion Impact Factor	
	Avg		Avg		Avg	99%
RR Two-sided Lat (8 B)	1.8X		1.8X		120.9X	132.5X
RR Two-sided BW+Sync (131072 B)	3.9X		4.6X		7.9X	9.4X
Multiple Allreduce (8 B)	1.7X		1.5X		79.8X	87.1X

BROADCOM

Congestion Control (CC) Enhancements for RoCEv2

- RoCEv2 is designed to scale – no inherent limitation at the protocol level
- RoCEv2 with PFC provides lossless service with a limited scaling of nodes
  - PFC without congestion control should be avoided
- CC with or without PFC is essential for node and process scaling
  - ECN marking in switches enable Congestion Notifications to minimize congestion
  - CC algorithms are evolving to provide better congestion avoidance & faster congestion reaction
- CC enhancements in NICs further improve performance & scalability of RoCEv2
  - ECN marking/CNP generation
  - Hardware-based congestion control

BROADCOM

Thank You


BROADCOM®
connecting everything®