# Wire-Speed Packet Capture and Replay at 200 Gbps

Lukáš Kekely
CESNET a.l.e.
Prague, Czech Republic
kekely@cesnet.cz

Viktor Puš
Netcope Technologies
Brno, Czech Republic
pus@netcope.com

## ABSTRACT

CESNET (Czech NREN) and Netcope Technologies are ready to demonstrate their NFB-200G2QL smart NIC with Virtex UltraScale+ FPGA specifically designed to push the achievable traffic processing throughput to 200 Gbps in a single card. Unique high-speed DMA engines in the FPGA together with highly optimized Linux drivers enable to achieve 200 Gbps data transfer through two PCIe Gen3 ×16 interfaces with minimal CPU overhead. Captured network traffic can be independently distributed among individual cores of two physical CPUs (NUMA nodes) without utilization of QPI. As a result, wire-speed packet capture to the host memory from two fully saturated 100 Gbps Ethernet interfaces (QSFP28+) is achieved and various network monitoring applications can utilize the power of the latest FPGAs and CPUs for data processing. This is especially useful when traffic of both directions of a single 100GbE link needs to be processed.

The proposed demonstration shows how the packets can be received from two 100 Gbps Ethernet links at full speed and captured to the host memory at 200 Gbps without any loss. The opposite direction of communication is also shown, i.e. how the packets can be transmitted from the host memory towards the two 100GbE network interfaces. Achieved speeds are demonstrated by counters and gauges showing generated, received/transmitted and captured packets. We also show detailed statistics of CPU load during the packet capture/replay for different packet lengths.

**ACM Reference Format:**
Lukáš Kekely and Viktor Puš. 2018. Wire-Speed Packet Capture and Replay at 200 Gbps. In *SIGCOMM '18: ACM Special Interest Group on Data Communication, August 20–25, 2018, Budapest, Hungary*. ACM, New York, NY, USA, 2 pages.


*SIGCOMM '18, August 20–25, 2018, Budapest, Hungary*


## 1 TECHNOLOGY OVERVIEW

100 Gigabit Ethernet (100GbE) was first defined by the IEEE 802.3ba standard [2] from 2010 and currently is the fastest deployed standard of Ethernet for computer networks. It enables for transmitting frames at a rate of 100 Gbps, which translates up to nearly 150 millions frames per second (new frame every 6.7 ns). We demonstrated COMBO-100G card back in 2014 [3] as the first PCI Express adapter card to support 100G Ethernet technology worldwide.

The new low-profile NFB-200G2QL card is shown in Figure 1. It is world's first PCI Express adapter designed to enable wire-speed processing of traffic at 200 Gbps. This hardware-accelerated card with FPGA uses unique DMA modules that can sustain data transfers over PCI Express between the card and memory of the host computer at such high speeds. This feature makes the card ideal for deployment in the fastest backbone networks and in high-throughput data centers. Other main features of the card include:

- two 100GbE QSFP28+ transceiver interfaces (cages),
- powerful Virtex UltraScale+ VU7P FPGA chip,
- three static QDR-IIIe memories (max. 288 Mb each),
- two PCI Express Gen3 interfaces with 16 lanes each,
- PCI Express half-length and low-profile form factor,
- external PPS input for precise timestamps.

**Figure 1: Front view of the NFB-200G2QL FPGA card.**

As a base of our FPGA firmware, we have developed a platform for rapid development of hardware-accelerated applications. The platform includes a set of firmware IP cores, especially blocks for network interfaces and a unique high-performance DMA bus-master connection to the software layer via PCIe bus. The software layer consists of Linux device drivers, tools for card management, and libraries for high-speed data transfers between the card and the host memory (DPDK or proprietary SZE2). The framework also specifies a generic interface to optional traffic processing pipeline in FPGA that can be described using P4 language [4] to perform different operations on passing network data [1]. Our HaNIC solution is an example of such traffic processing in the FPGA firmware. It extends the functionality of a basic NIC by the support of packet parsing, filtering, and configurable hash-based distribution among CPU cores.

## 2 DEMO DESCRIPTION

The goal of the proposed demo is to present unique performance and features of our low-profile NFB-200G2QL card. We want to especially stress out the ability of the card to: 1) operate both 100 GbE interfaces at wire-speed, 2) transfer all received data via PCIe into the host memory at full 200 Gbps regardless of the frame length, and 3) transfer data from the host memory via PCIe at full 200 Gbps regardless of the frame length. Illustration of the demo architecture is shown in the Figure 2. The NFB-200G2QL card is connected into PCIe slots of standard server with two multicore CPUs and fully filled memory banks (for maximal throughput). Inside the card's FPGA there is our HaNIC firmware configured to capture all of the incoming traffic and distribute it among available CPU cores. Both Ethernet ports of the card are connected to a tester device that can generate and receive (analyze) 100GbE network traffic. Since conventional hardware testers supporting 100 GbE ports (e.g. Spirent TestCenter) are too large and heavy to transport, we can instead implement the required traffic generation and capture capabilities inside our FPGA firmware and connect the optic cables in a loopback. Described demo architecture can operate in two basic modes: packet capture and packet replay.

In packet capture mode, packets of configurable length are generated at the maximum allowed rate and sent over the fiber into both 100 Gbps Ethernet ports. The packets are then received by on-card PMA, PCS and MAC engines, distributed into multiple DMA channels and transferred via PCIe into the ring buffers inside server's main memory. In the memory, the packets are accessed and counted. Processing is this simple because we want to demonstrate that the card is capable of delivering the 200 Gbps of data into the software and not the performance of some specific advanced packet handling in the CPUs. Finally, live packet capture

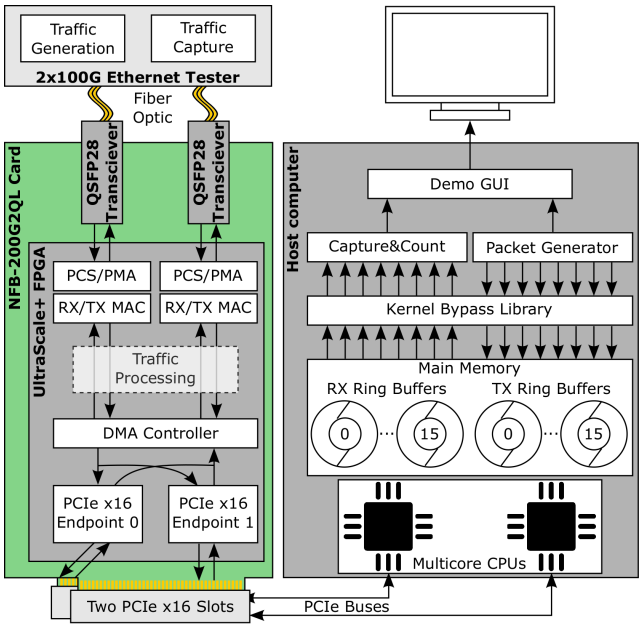


**Figure 2: Demo system architecture illustration.**

performance statistics are shown in the GUI on the screen. This mode corresponds to typical network monitoring scenarios, where traffic of both directions of a tapped 100GbE link needs to be processed.

In packet replay mode, packets of configurable length are prepared by CPUs and copied into multiple DMA ring buffers. From there, they are picked up by the DMA controllers in the card's FPGA and transferred via PCIe into its local memory. They are then transferred using standard Ethernet layers onto two 100 GbE lanes. Finally, live performance statistics are shown in the GUI. This mode corresponds to data center deployment, where large amounts of data are transferred.

## ACKNOWLEDGMENT

This work is supported by the Ministry of the Interior of the Czech Republic project VI20172020064, by the project Reg. No. CZ.02.1.01/0.0/0.0/16_013/0001797, and by the Technology Agency of the Czech Republic project TH02010214.

## REFERENCES

[1] Pavel Benáček, Viktor Puš, Hana Kubátová, and Tomáš Čejka. 2018. P4-To-VHDL: Automatic generation of high-speed input and output network blocks. *Microprocessors and Microsystems* 56 (2018), 22–33.

[2] IEEE Computer Society. 2010. Amendment 4: Media Access Control Parameters, Physical Layers, and Management Parameters for 40 Gb/s and 100 Gb/s Operation. *IEEE Standard 802.3ba-2010* (June 2010), 1–457.

[3] Lukáš Kekely, Jan Kořenek, Viktor Puš, and Milan Dvořák. 2014. 100G Packet Capture Live Demo. *24th International Conference on Field Programmable Logic and Applications (Demo Night)* (September 2014).

[4] Pat Bosshart et al. 2014. P4: Programming Protocol-independent Packet Processors. *SIGCOMM Computer Communication Review* 44, 3 (2014).