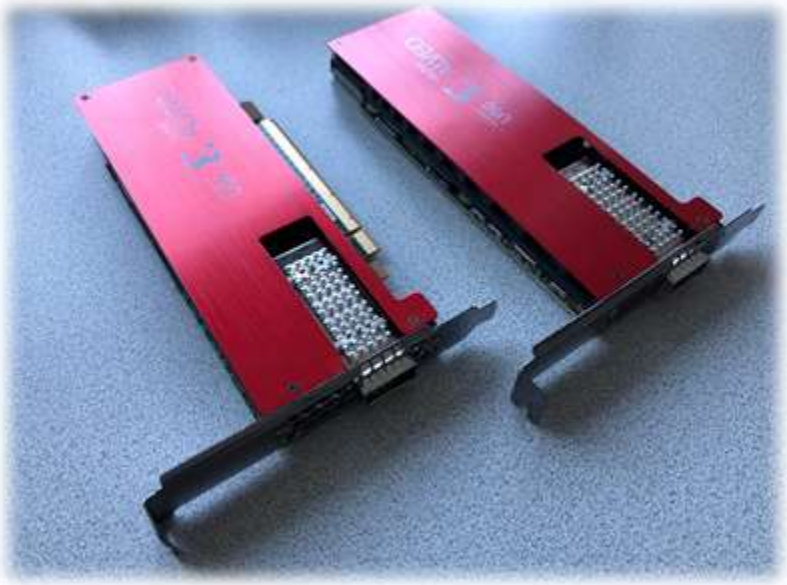
Host Bypassing: Direct Data Piping from the Network to the Hardware Accelerator

MCSoC, 2021



Ralf Kundel, ralf.Kundel@kom.tu-darmstadt.de
Kadir Eryigit
Jonas Markussen
Carsten Griwodz
Osama Abboud
Rhaban Hark
Ralf Steinmetz

KOM - Multimedia Communications Lab
Prof. Dr.-Ing. Ralf Steinmetz (Director)
Dept. of Electrical Engineering and Information Technology
Dept. of Computer Science (adjunct Professor)
TUD – Technische Universität Darmstadt
Merckstr. 25, D-64283 Darmstadt, Germany
Tel.+49 6151 166150, Fax. +49 6151 166152
www.KOM.tu-darmstadt.de

TECHNISCHE UNIVERSITÄT DARMSTADT

5G: A exemplary use-case

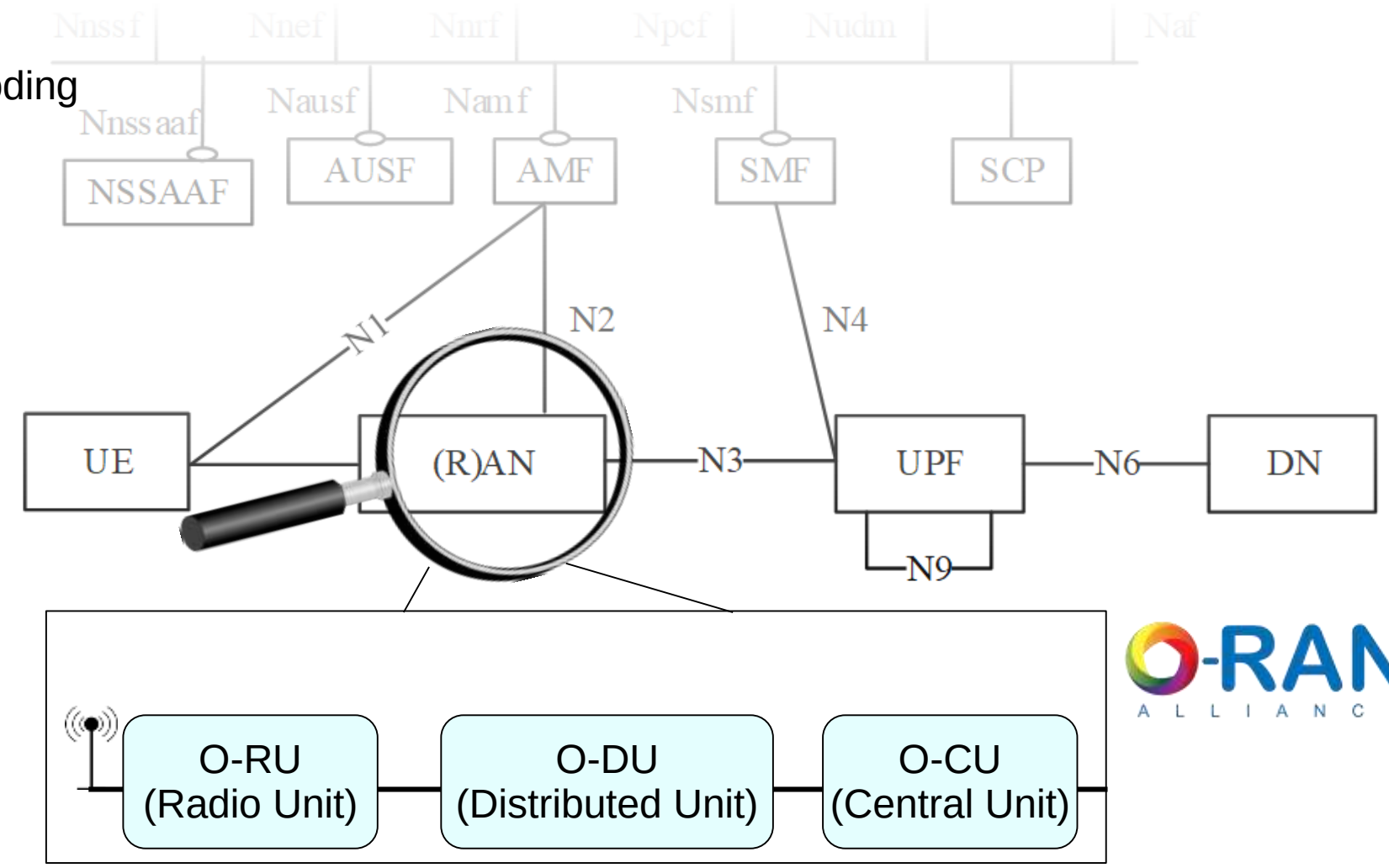
Acceleration along the Data Path

Disaggregated RAN:

- Channel Encoding/Decoding
- Rate Matching
- Layer Mapping
- PTP Timing
- …

Proprietary 4G base band unit

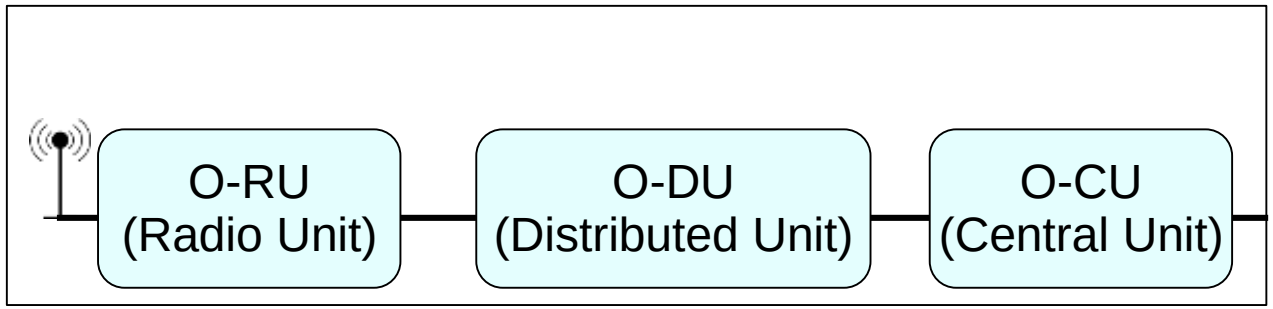
Figure source: https://www.xilinx.com/publications/presentations/xilinx-5g-telco-accelerator-cards.pdf



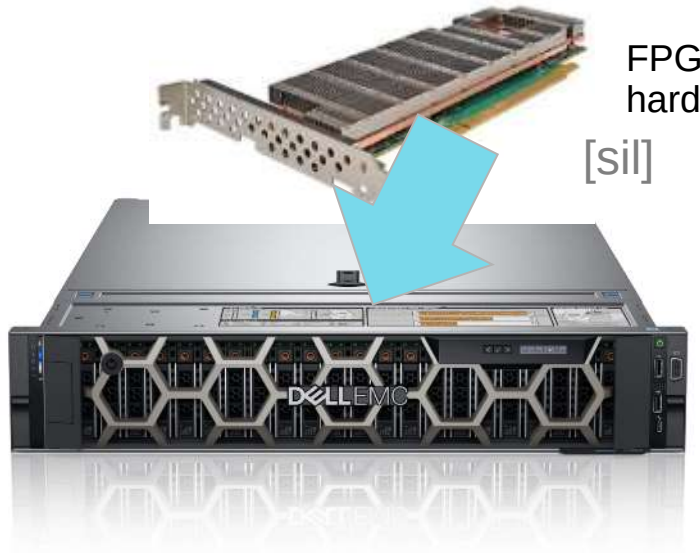
Distributed Unit

Technical Realization

- DU & CU can be combined (but must not be)
- Very compute intensive
  - Each bit must be touched
  - Latency/Jitter sensitive



Software based realizations:
"Decoder speed of up to 0.183 Gbit/s" [xil]



FPGAs as hardware accelerators [sil]

[xil] https://www.xilinx.com/publications/presentations/xilinx-5g-telco-accelerator-cards.pdf
[sil] https://www.silicom-usa.com/pr/server-adapters/4g-5g-acceleration-adapters/acc100_fec_accelerator_server_adapter/

TECHNISCHE UNIVERSITÄT DARMSTADT

Host Bypassing

PCIe Host Bypassing

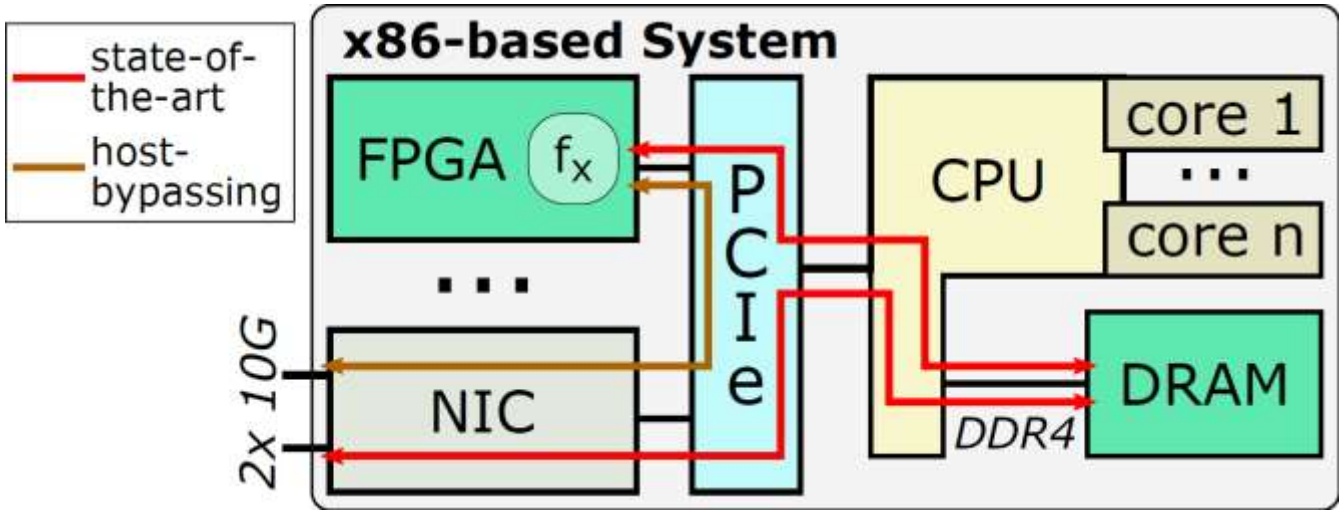
TECHNISCHE UNIVERSITÄT DARMSTADT

Problem:

- Currently four (4) memory copies needed
- Software necessarily involved

Approach:

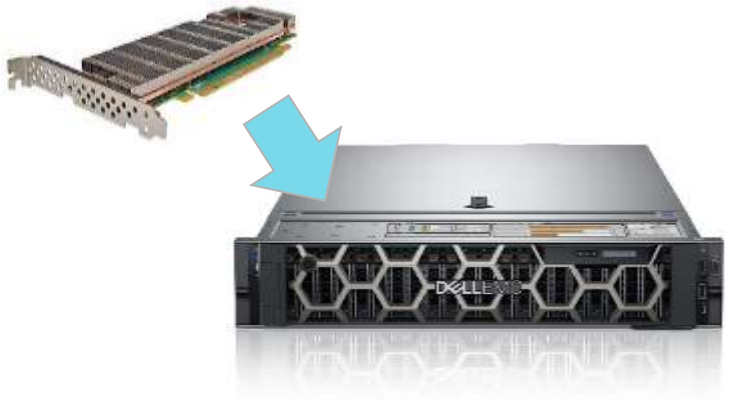
- Receive data directly on FPGA/GPU
- PCIe P2P DMA needed



NIC = Network Interface Card

Presented Design:

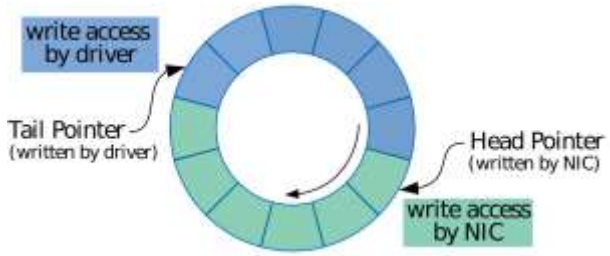
- Intransparent to Network Interface Card (NIC)
  - Commodity NICs, e.g., Intel 82599
- FPGA support (Intel and Xilinx)

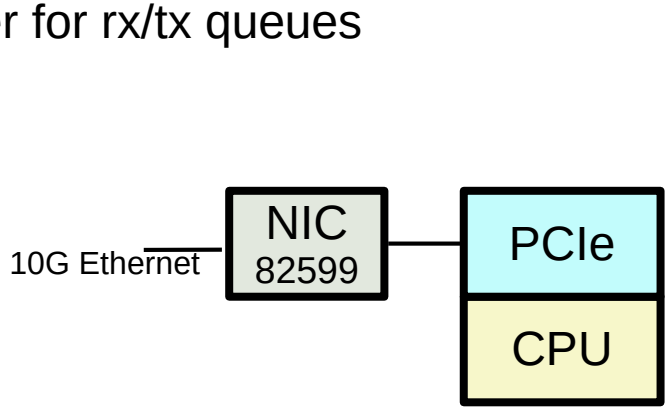


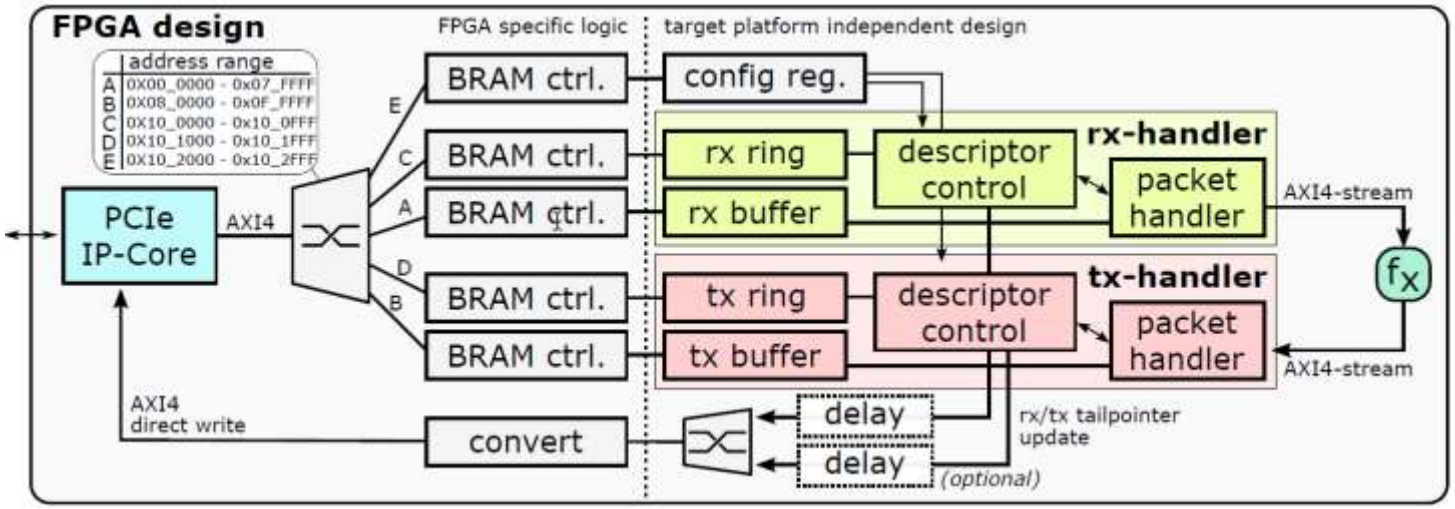
Realization

Emulate SW-driver in FPGA

- No interrupts
- Shared memory for communication
  - NIC can write in main memory of CPU (or FPGA)
  - CPU (or FPGA) can write into NIC registers
- Ring buffer for rx/tx queues





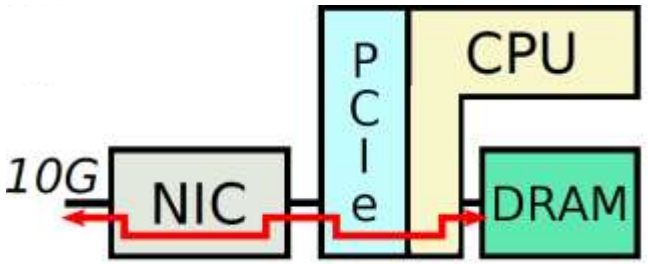


CPU tasks:

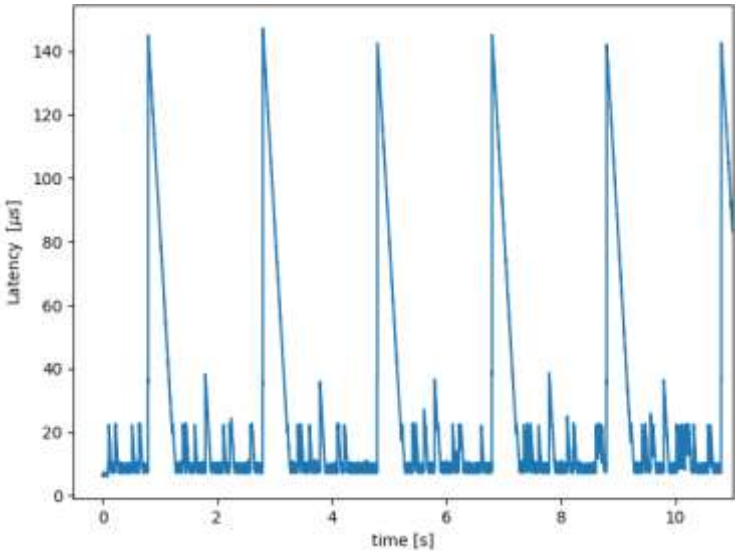
- Configure the NIC to the physical FPGA address
- Configure the FPGA to the physical NIC address

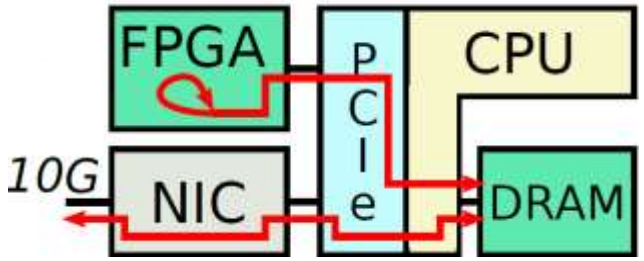
Preliminary Results

300 byte UDP packets, 9.99 Gbit/s input rate



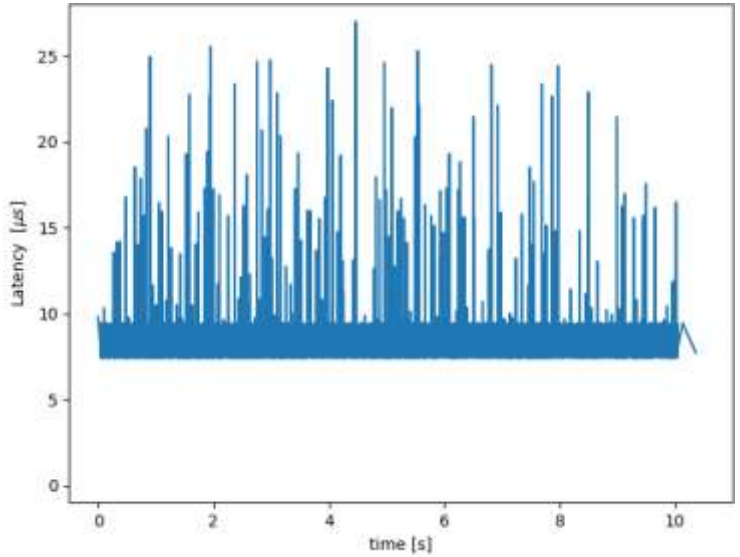
0 packet loss @ 9.99 Gbit/s

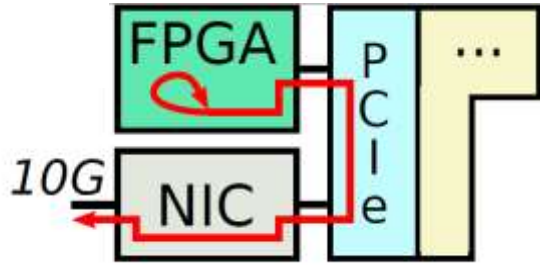




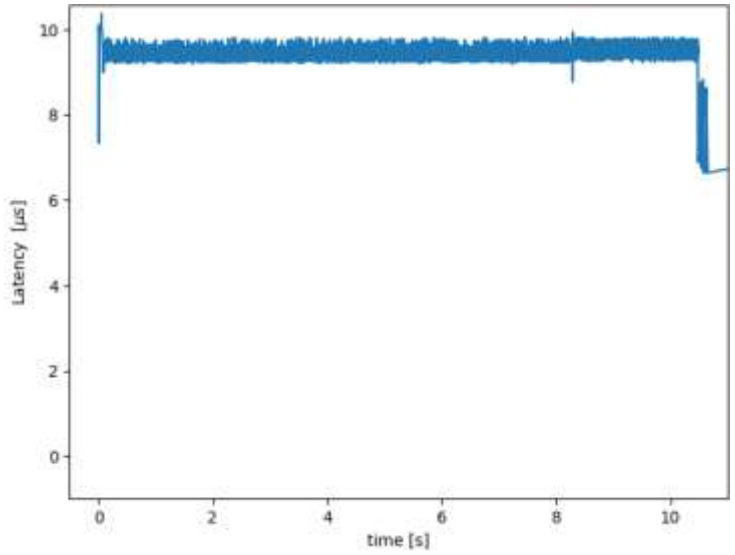
~75% packet loss @ 9.99 Gbit/s

0 packet loss @ 650 Mbit/s



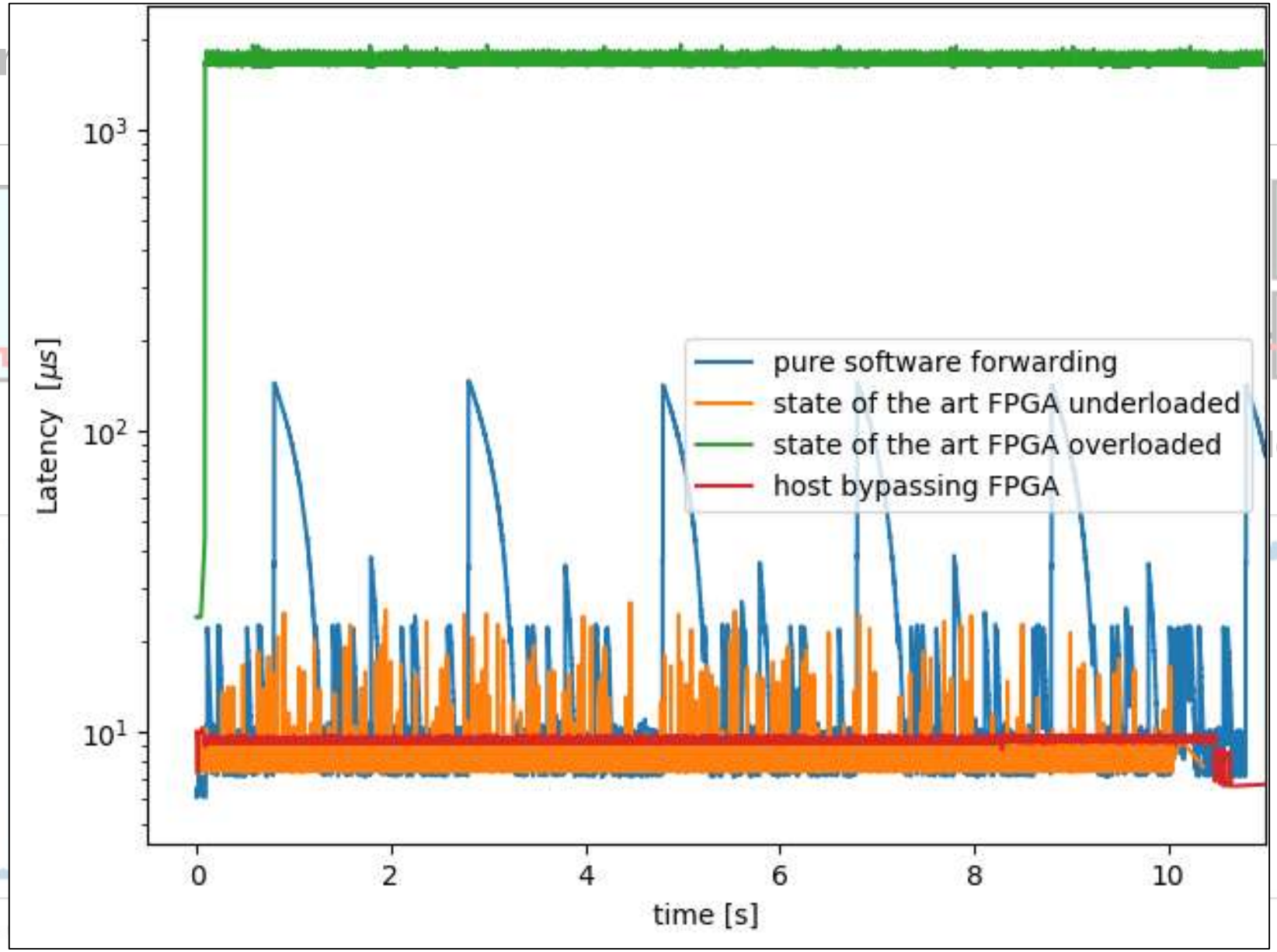


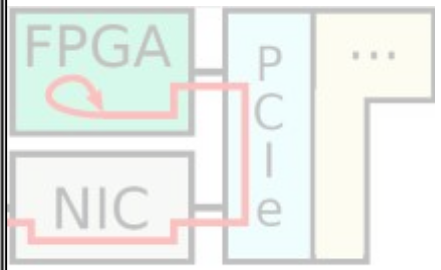
0 packet loss @ 9.99 Gbit/s



Preliminar

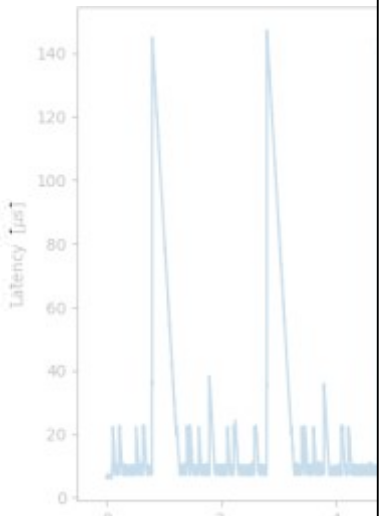
300 byte UDP





0 packet loss

loss



The point of failure

Baseline:

- “dpdk”

State of the Art:

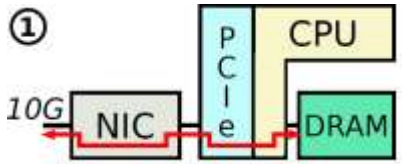
- “dpdk + fpga”

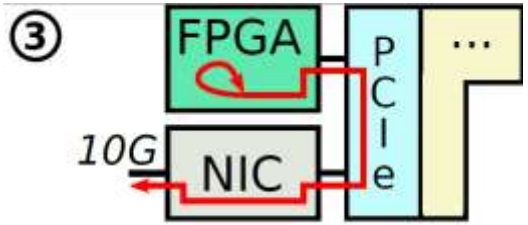
Intel/AMD:

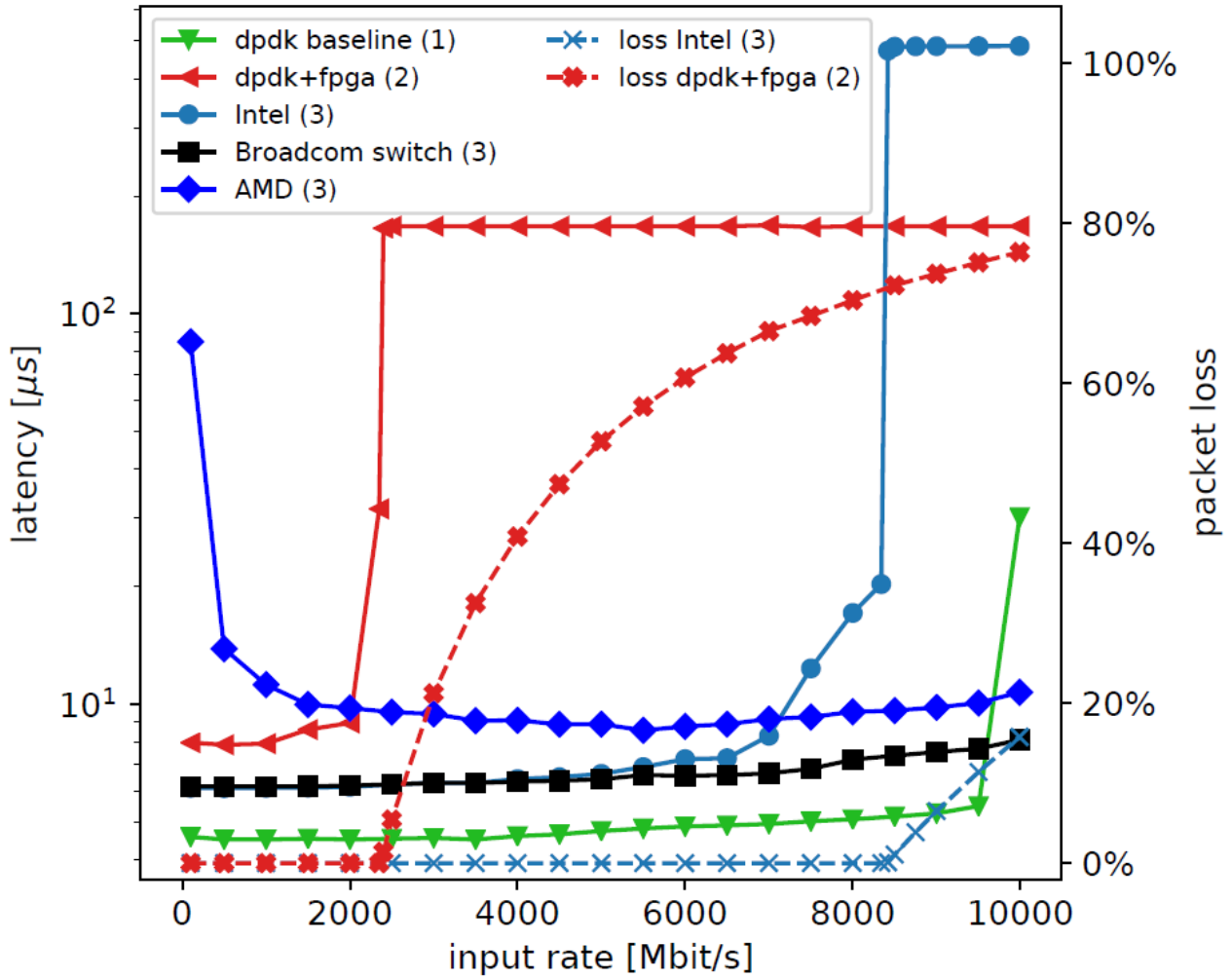
- Xeon 4110
- Epyc 7402

Broadcom switch:

- PEX8747







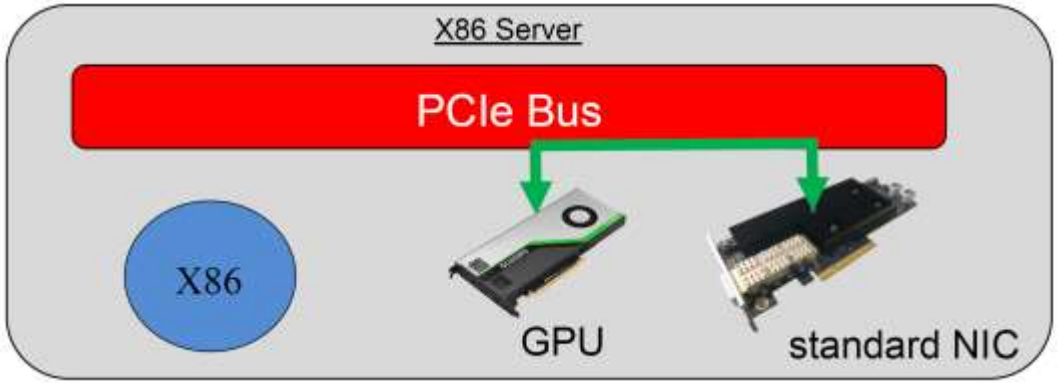
Next Steps

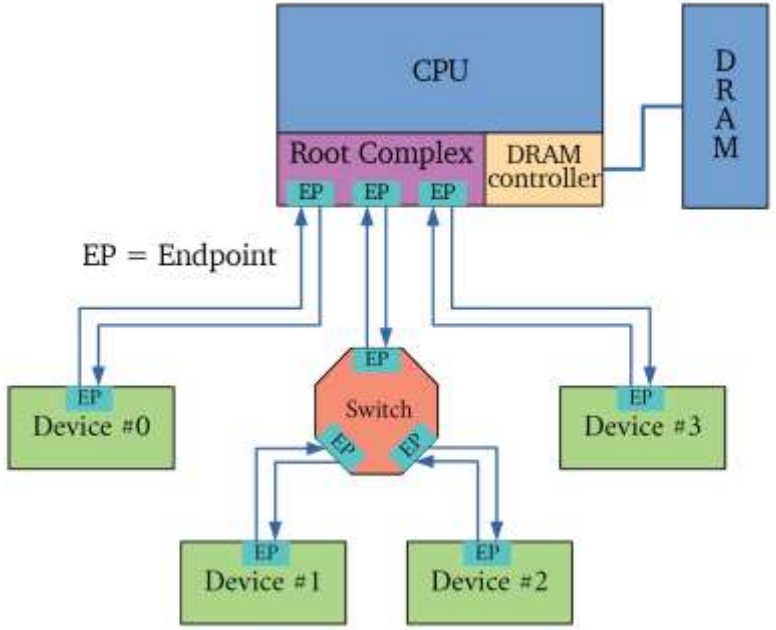
To be investigated:

- GPU bypassing
- Investigate more CPU types
- Offloading a 5G network function
- 40G/100G NICs

CPU vendors:

- There could be more possible …
- Maybe in next generation of CPUs ;-)





Interested in more?

Ralf Kundel et al., **“Host Bypassing: Direct Data Piping from the Network to the Hardware Accelerator”**
to appear in Proceedings of MCSoC-2021


Host Bypassing: Direct Data Piping from the Network to the Hardware Accelerator

There is also related work ;-)

Many thanks for your attention!

TECHNISCHE UNIVERSITÄT DARMSTADT

KOM – Multimedia Communications Lab

Funded by:

DFG CRC 1053 - MAKI
Multi-Mechanisms Adaptation for the Future Internet

software campus