User Plane Hardware Acceleration in Access Networks: Experiences in Offloading Network Functions in Real 5G Deployments

KuVS Fachgespräch “Network Softwarization”, 2022

Ralf Kundel, ralf.Kundel@kom.tu-darmstadt.de
Tobias Meuser
Timo Koppe
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

11. April 2022

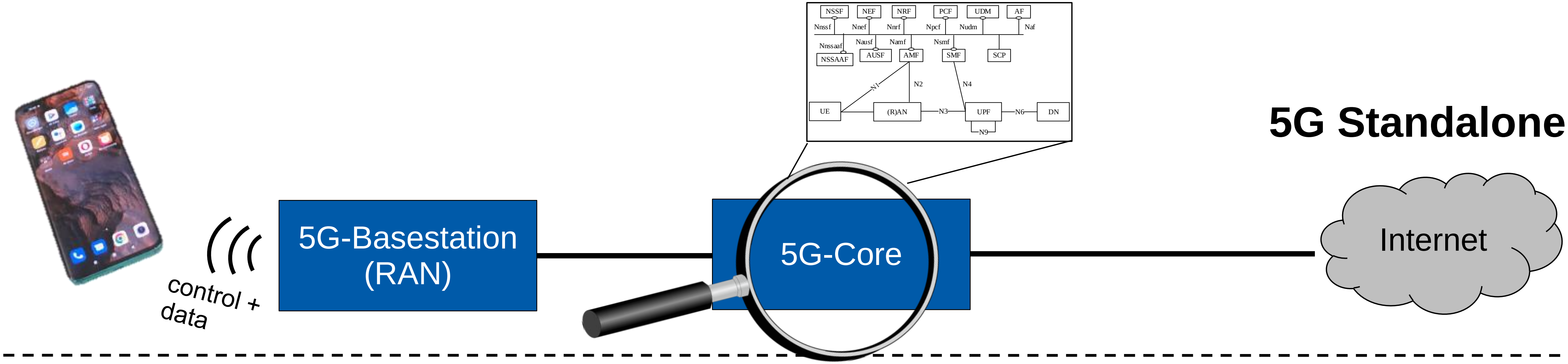
Goal/Challenge

Having an End-to-End
working 5G SA setup for research purposes!

5G Acceleration Goals:
Low Latency (~ 1ms)
High Throughput (> 1 Gbit/s)

5G Standalone vs. 5G Non-Standalone

5G Standalone



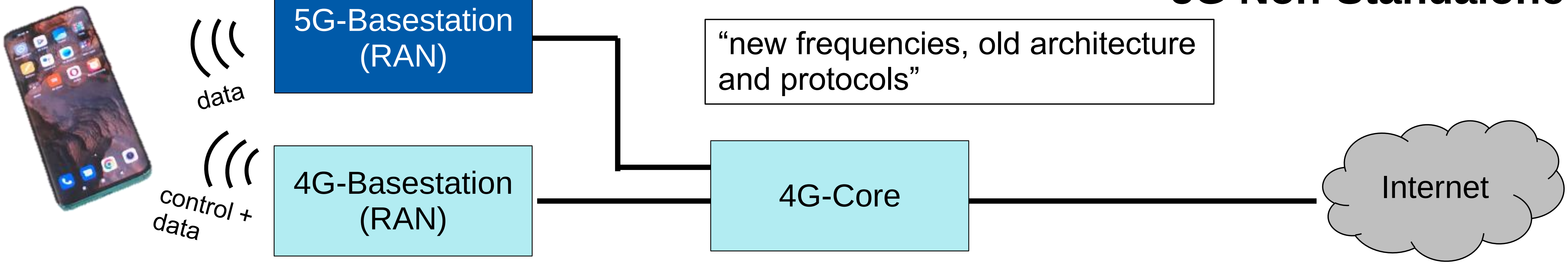
control + data

5G-Basestation (RAN)

5G-Core

Internet

5G Non-Standalone



data

control + data

5G-Basestation (RAN)

4G-Basestation (RAN)

4G-Core

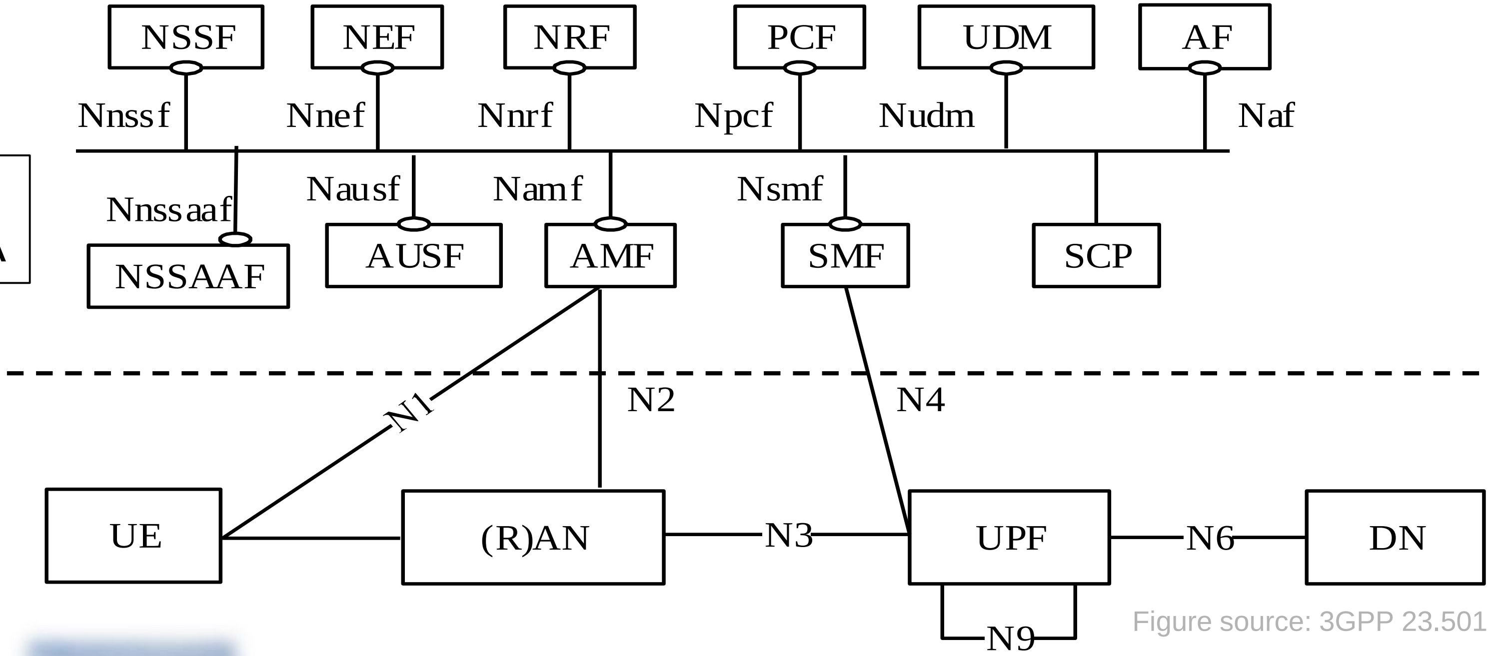
Internet

“new frequencies, old architecture and protocols”

The 5G Standalone Architecture

“The Internet meets SDN”

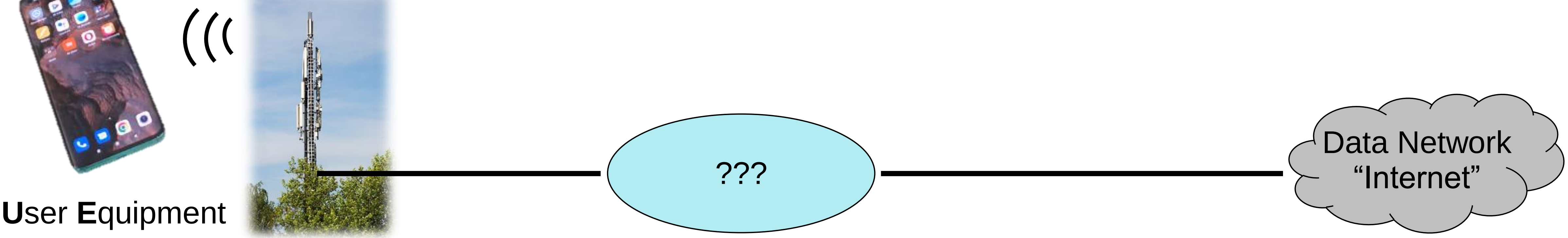
“service based architecture” - SBA



Control Plane

Data/User Plane

Figure source: 3GPP 23.501: Figure 4.2.3



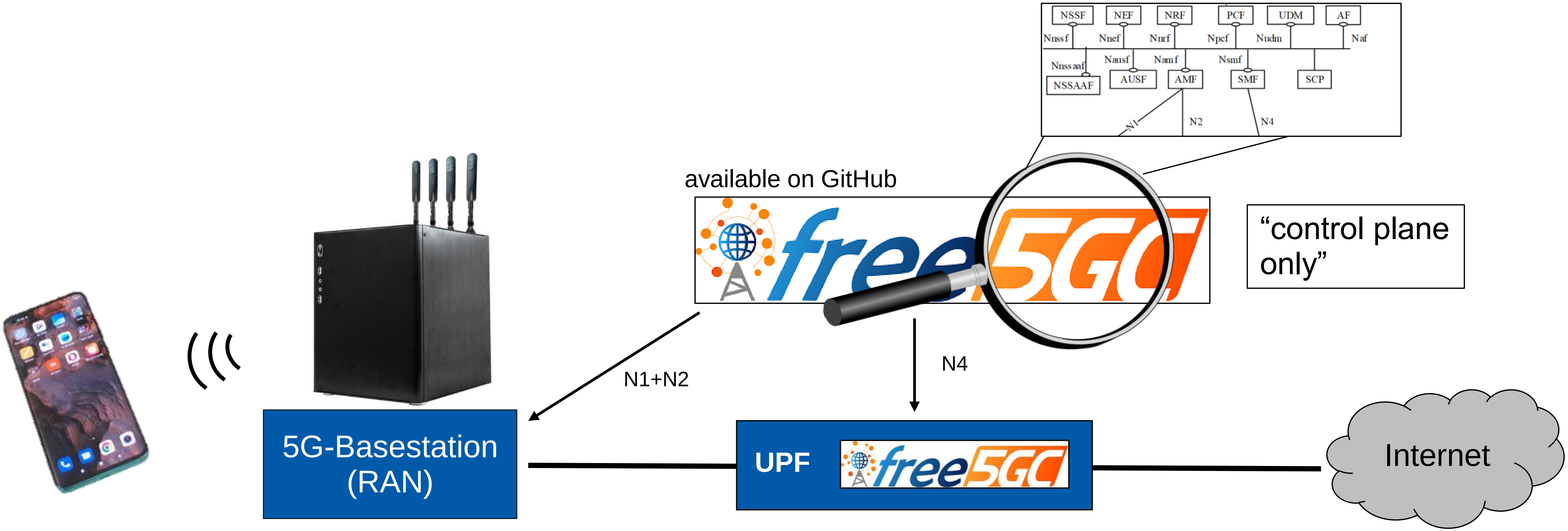
User **E**quipment

1. A 5G end-to-end setup

A basis for research

5G-Setup for data plane research

- 5G Standalone
- Theoretically ~600 Mbit/s (Uplink + Downlink)
- Basis for future research



available on GitHub

“control plane only”

N1+N2

N4

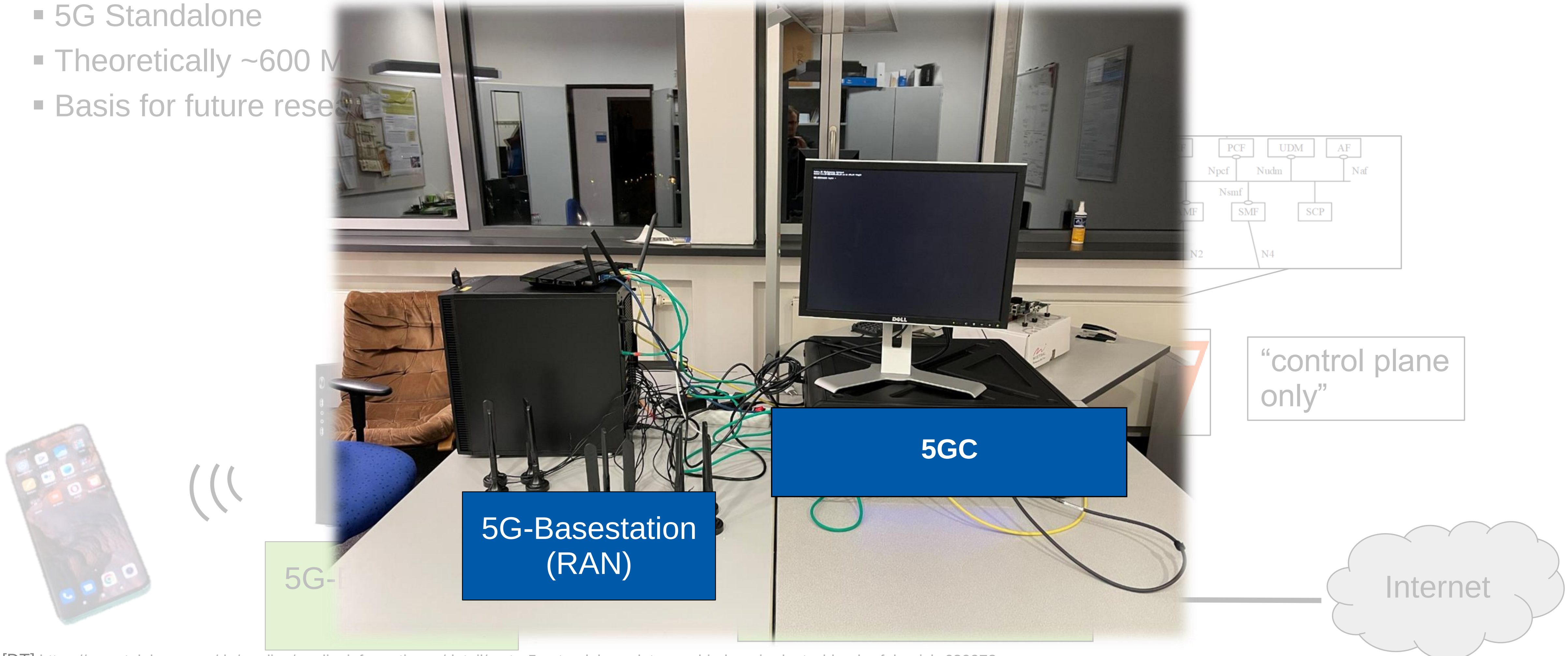
5G-Basestation (RAN)

UPF

Internet

5G-Setup @KOM

- 5G Standalone
- Theoretically ~600 M
- Basis for future rese



[DT] https://www.telekom.com/de/medien/medieninformationen/detail/erste-5g-standalone-daten-verbindung-in-deutschland-erfolgreich-620972

It’s running

In detail: Overall Setup

“IMS issue”

5G-core

- Free5gc

NAT/routing/switching

- Standard Linux tools
- “0815”-L2/L3 network

RAN

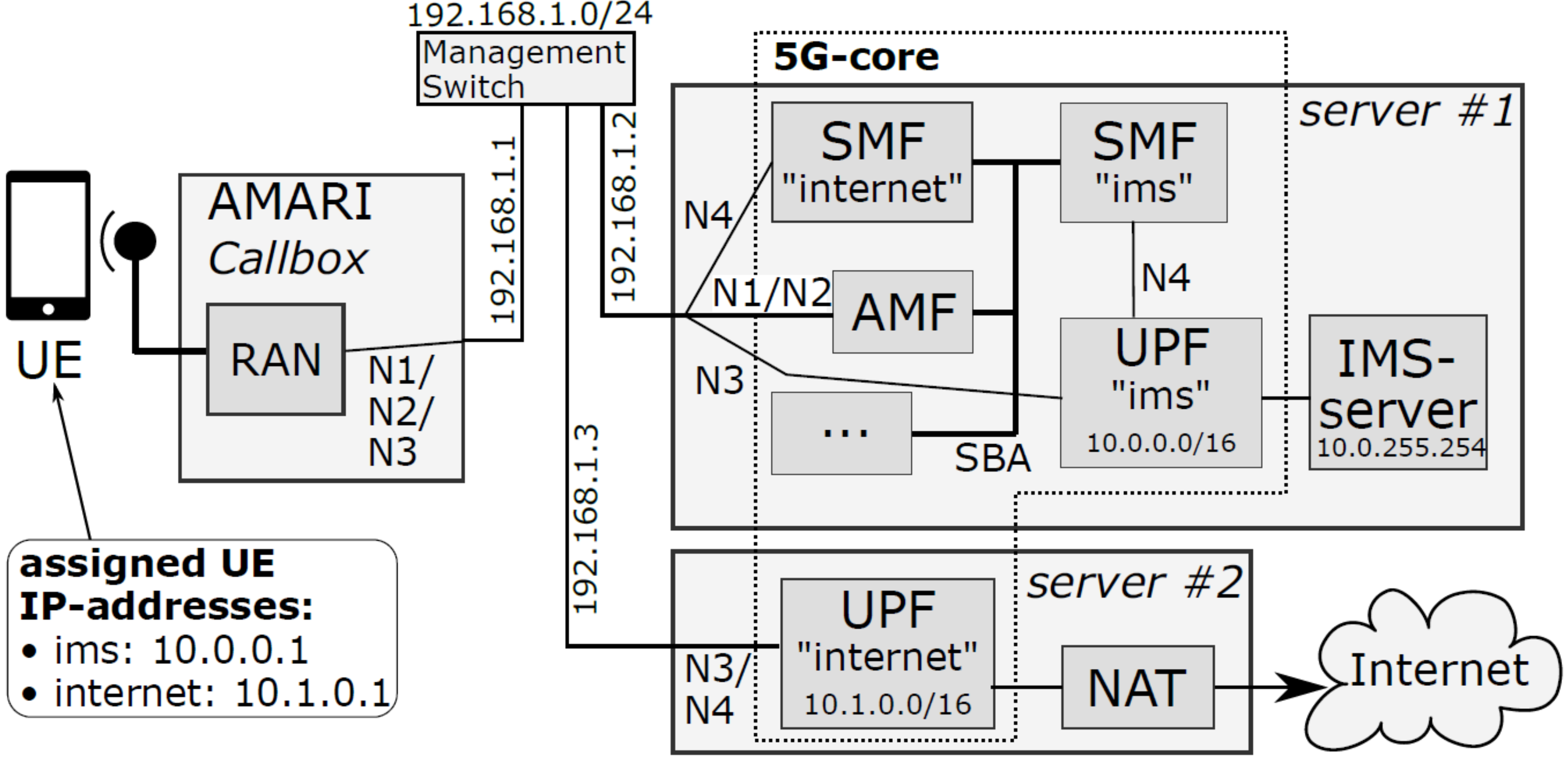
- Amarisoft Callbox

UEs:

- Huawei P40
- Oppo Find X2 Pro
- Oneplus 8T
- Waveshare SIM8200EA-M2-5G

free5GC

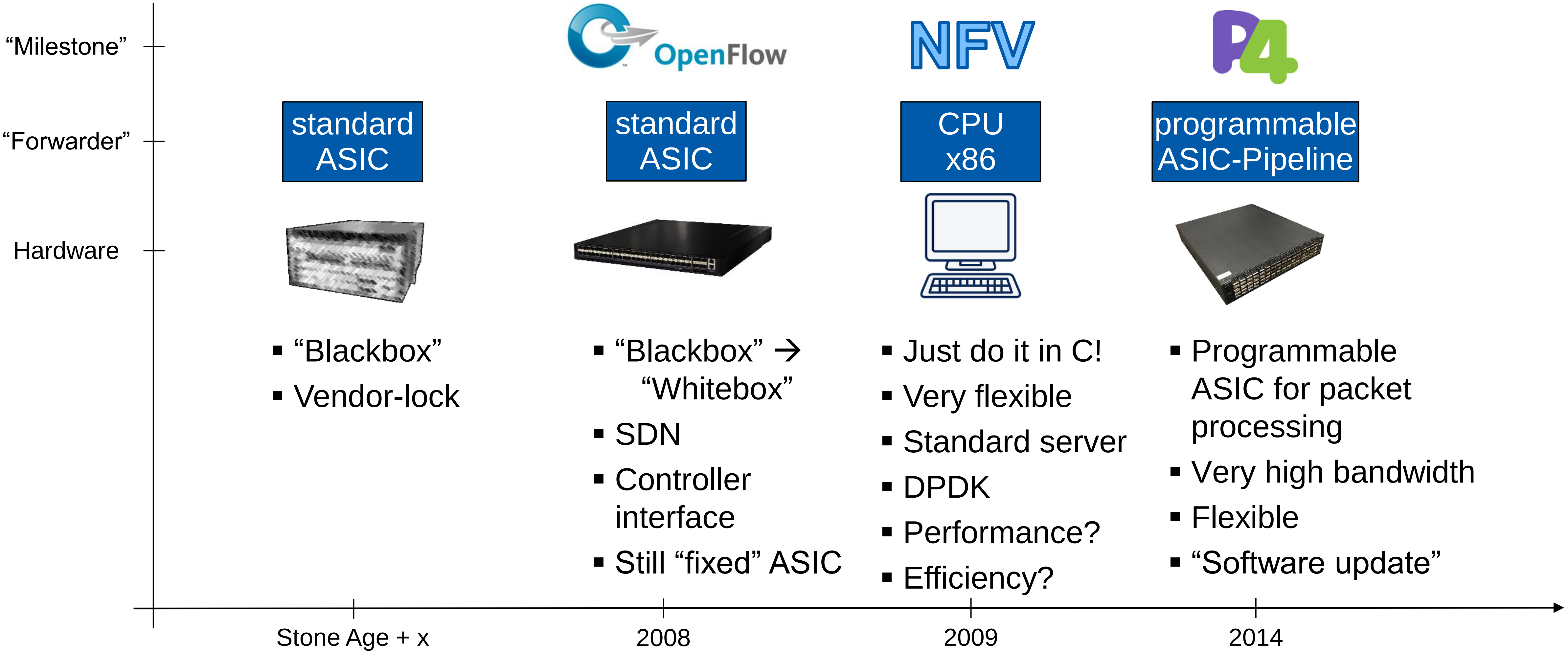
v3.0.6 is working!



2. Accelerate the user plane

How to make 5G fast

History of Data Planes



Stone Age + x – standard ASIC

- "Blackbox"
- Vendor-lock

2008 – OpenFlow – standard ASIC

- "Blackbox" → "Whitebox"
- SDN
- Controller interface
- Still "fixed" ASIC

2009 – NFV – CPU x86

- Just do it in C!
- Very flexible
- Standard server
- DPDK
- Performance?
- Efficiency?

2014 – P4 – programmable ASIC-Pipeline

- Programmable ASIC for packet processing
- Very high bandwidth
- Flexible
- "Software update"

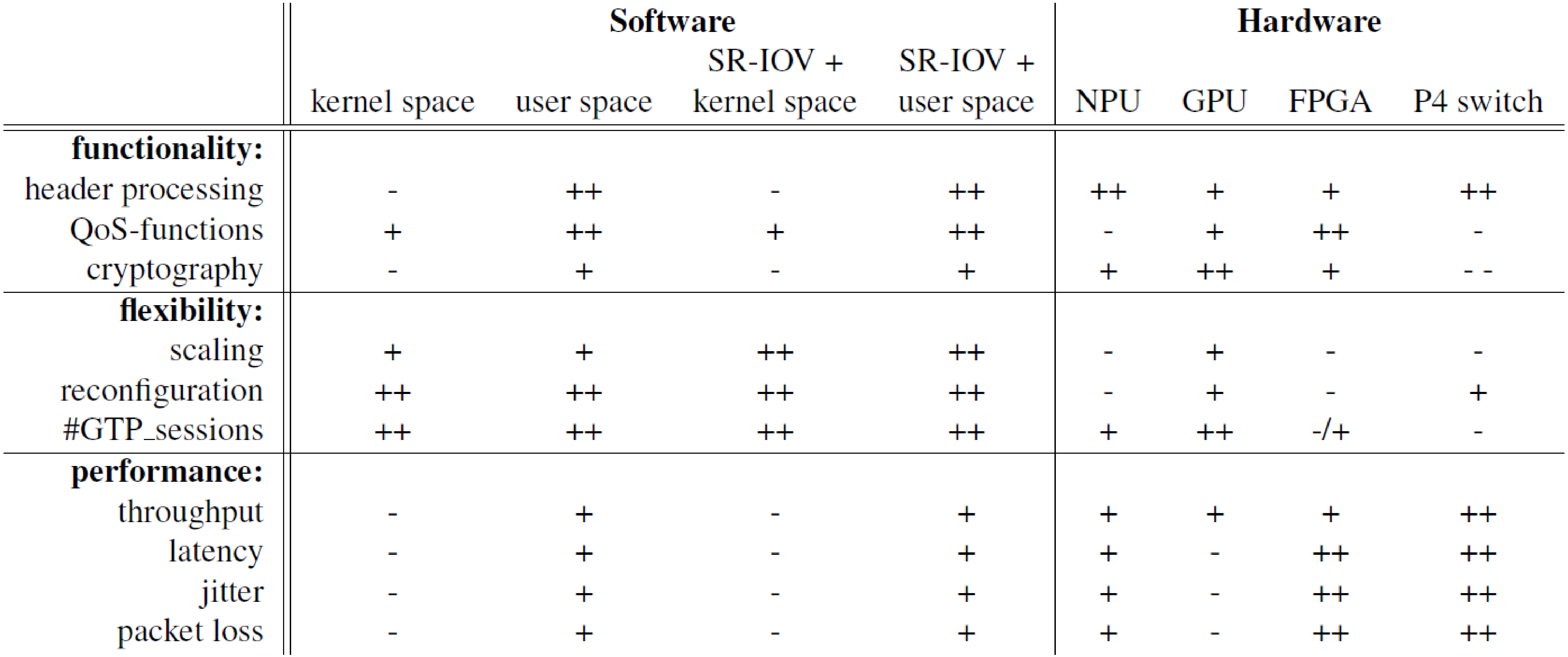
ASIC = „Application Specific Integrated Circuit" = purpose-build chip

P4

Programming **P**rotocol Independent **P**acket **P**rocessors

You all know it ;-)

Acceleration Techniques



	Software				Hardware			
	kernel space	user space	SR-IOV + kernel space	SR-IOV + user space	NPU	GPU	FPGA	P4 switch
functionality:								
header processing	-	++	-	++	++	+	+	++
QoS-functions	+	++	+	++	-	+	++	-
cryptography	-	+	-	+	+	++	+	- -
flexibility:								
scaling	+	+	++	++	-	+	-	-
reconfiguration	++	++	++	++	-	+	-	+
#GTP_sessions	++	++	++	++	+	++	-/+	-
performance:								
throughput	-	+	-	+	+	+	+	++
latency	-	+	-	+	+	-	++	++
jitter	-	+	-	+	+	-	++	++
packet loss	-	+	-	+	+	-	++	++

3. A P4 UPF

User Plane Function: PFCP Interface

Attach UPF <-> SMF

- SMF Registers at UPF
- In general: SMF is the “master”

UE Registration

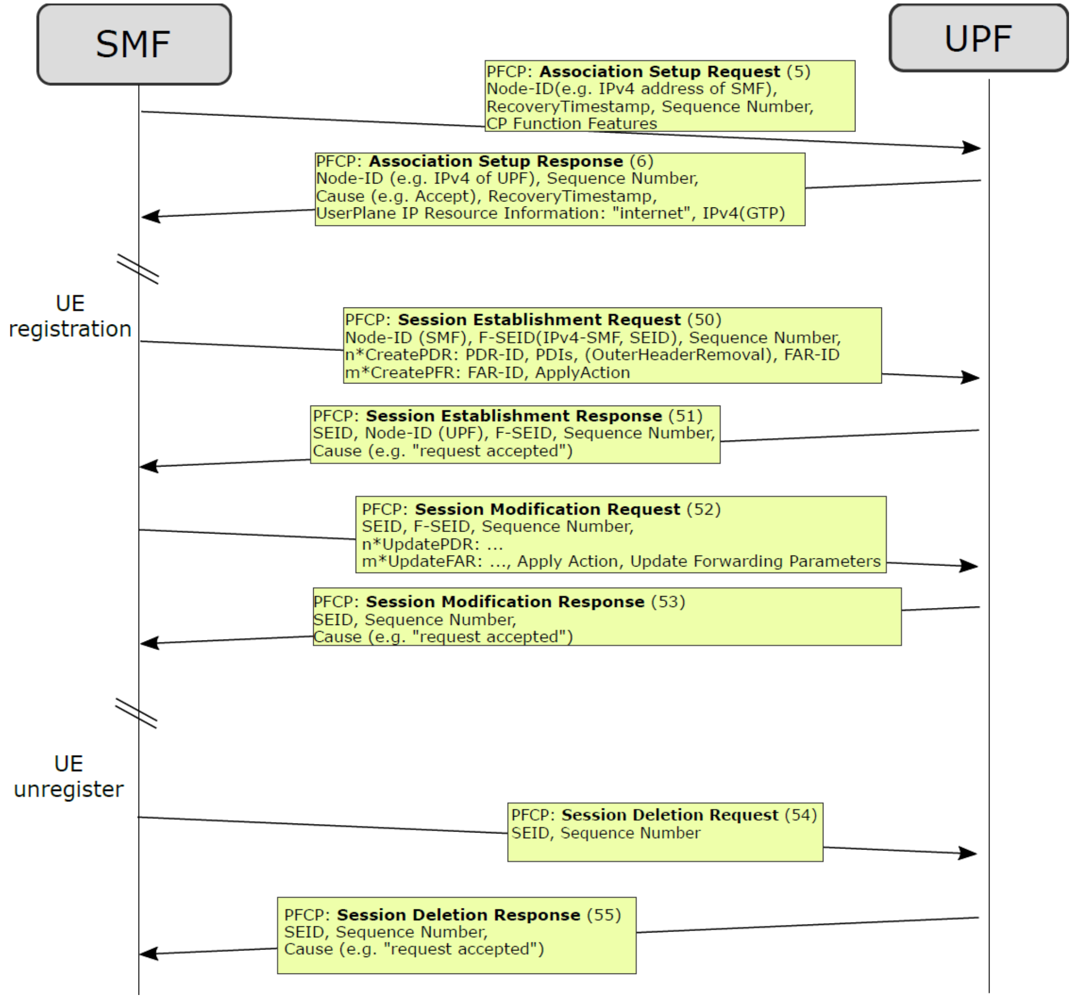
- Establish PDU Session

UE Modification

- Update PDU Session

UE Release

- Delete PDU Session



PFCP: **Association Setup Request** (5)
Node-ID(e.g. IPv4 address of SMF),
RecoveryTimestamp, Sequence Number,
CP Function Features

PFCP: **Association Setup Response** (6)
Node-ID (e.g. IPv4 of UPF), Sequence Number,
Cause (e.g. Accept), RecoveryTimestamp,
UserPlane IP Resource Information: "internet", IPv4(GTP)

PFCP: **Session Establishment Request** (50)
Node-ID (SMF), F-SEID(IPv4-SMF, SEID), Sequence Number,
n*CreatePDR: PDR-ID, PDIs, (OuterHeaderRemoval), FAR-ID
m*CreatePFR: FAR-ID, ApplyAction

PFCP: **Session Establishment Response** (51)
SEID, Node-ID (UPF), F-SEID, Sequence Number,
Cause (e.g. "request accepted")

PFCP: **Session Modification Request** (52)
SEID, F-SEID, Sequence Number,
n*UpdatePDR: ...
m*UpdateFAR: ..., Apply Action, Update Forwarding Parameters

PFCP: **Session Modification Response** (53)
SEID, Sequence Number,
Cause (e.g. "request accepted")

PFCP: **Session Deletion Request** (54)
SEID, Sequence Number

PFCP: **Session Deletion Response** (55)
SEID, Sequence Number,
Cause (e.g. "request accepted")

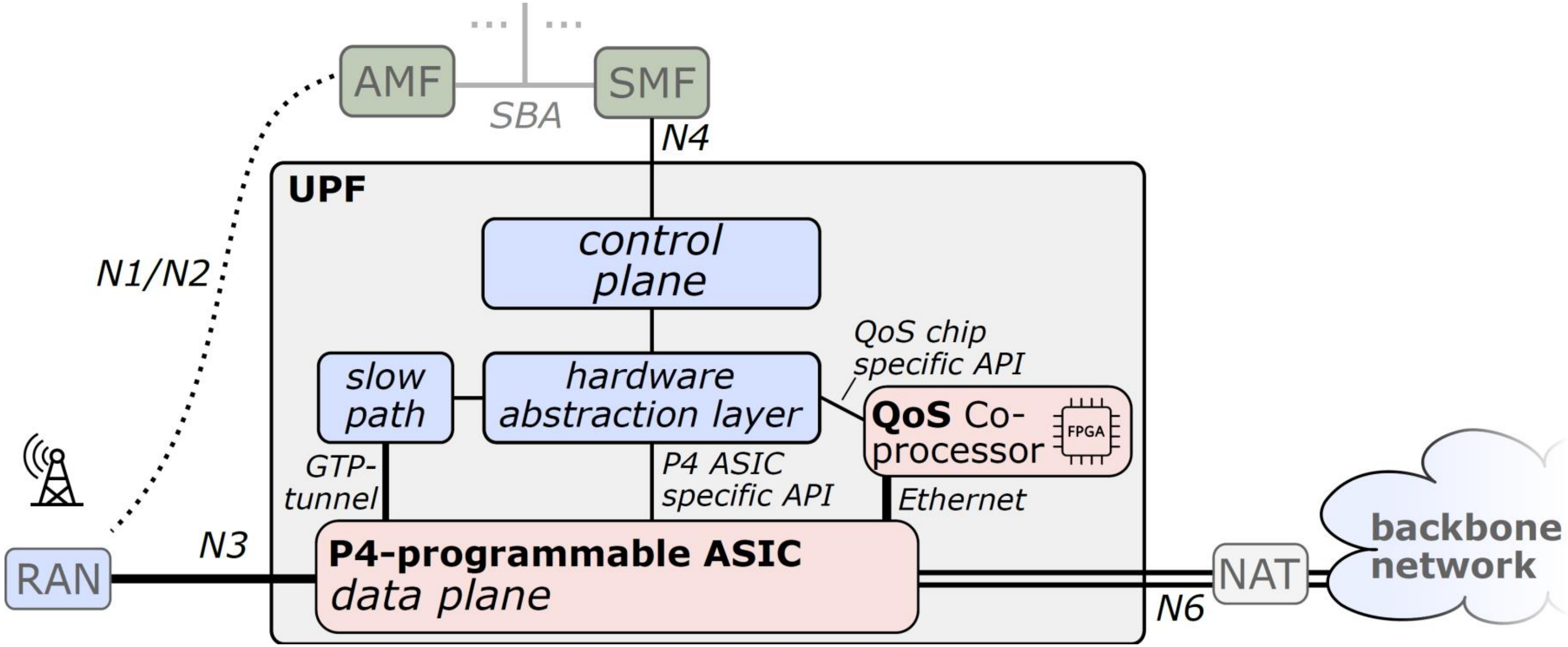
P4 UPF

Make the UPF fast

- Terminate GTP sessions in hardware
- Predictable good performance
- Low latency

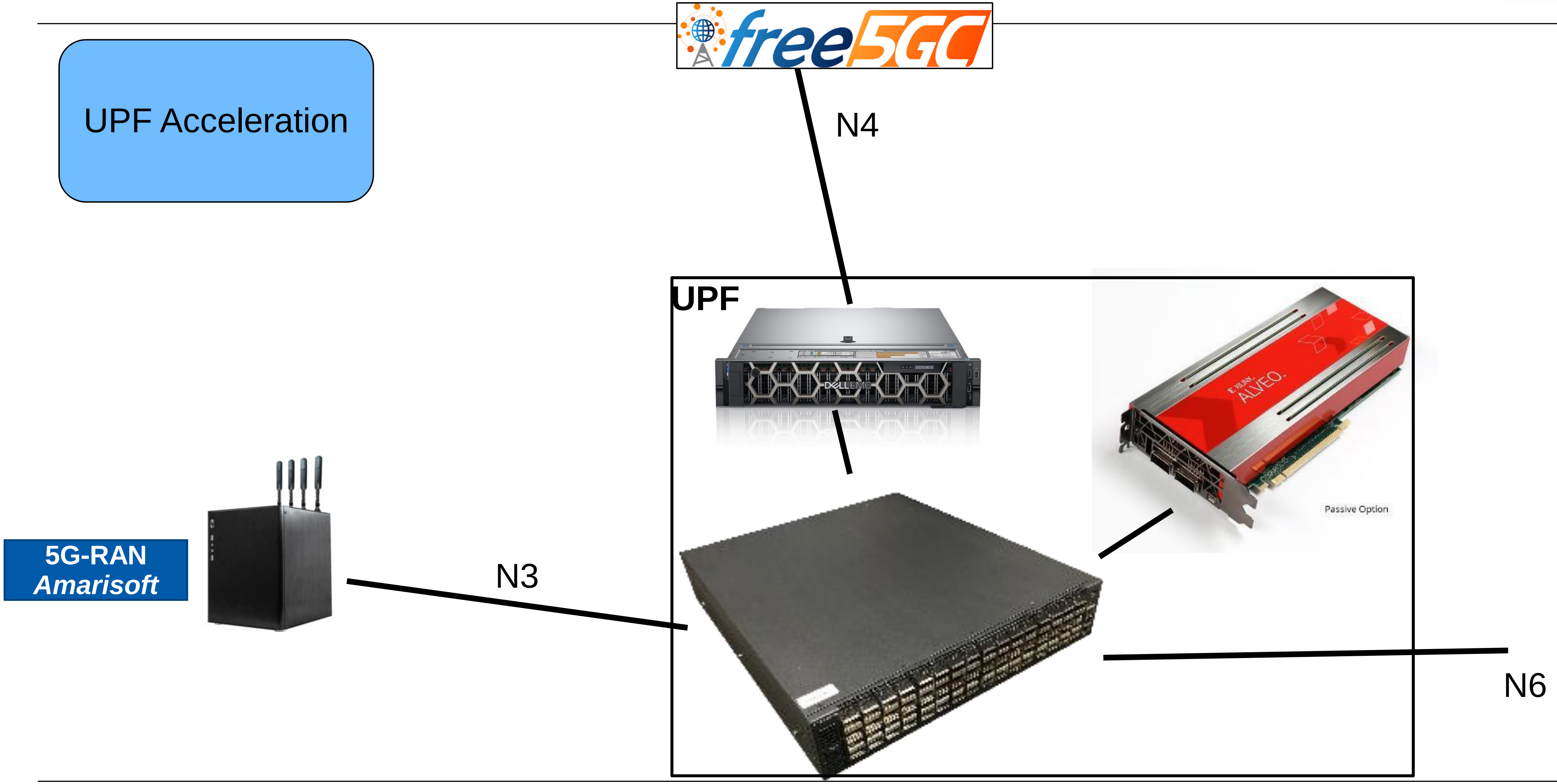
P4-UPF design

- Golang UPF-Controller
- P4-based data plane
- Future Work:
  - FPGA-based QoS chip
  - Hardware abstraction layer



Lab Setup – P4+FPGA UPF

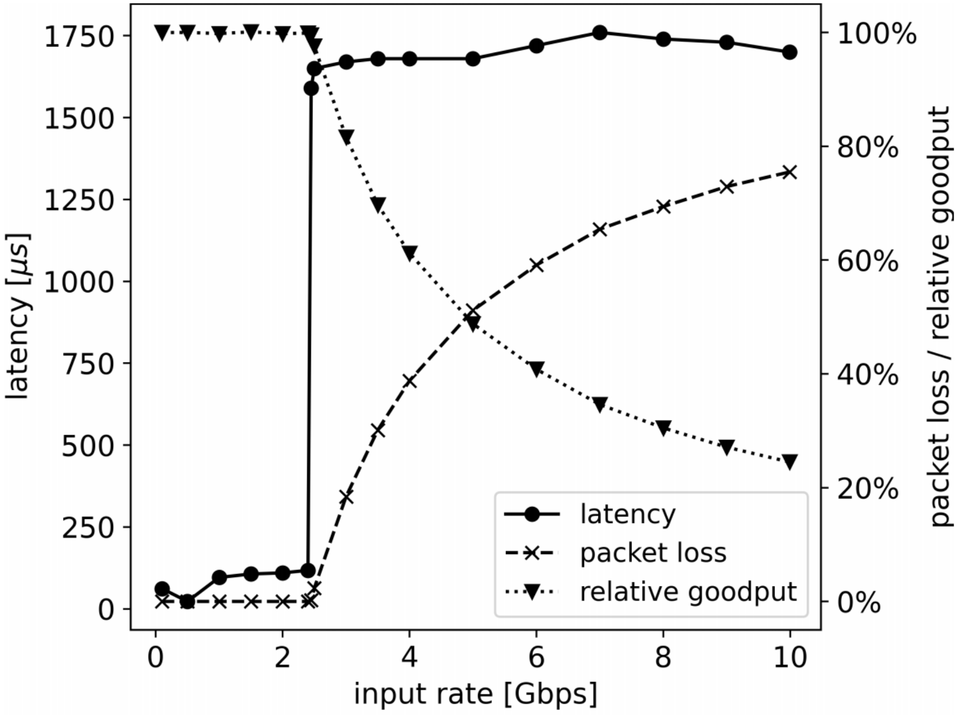
UPF Acceleration

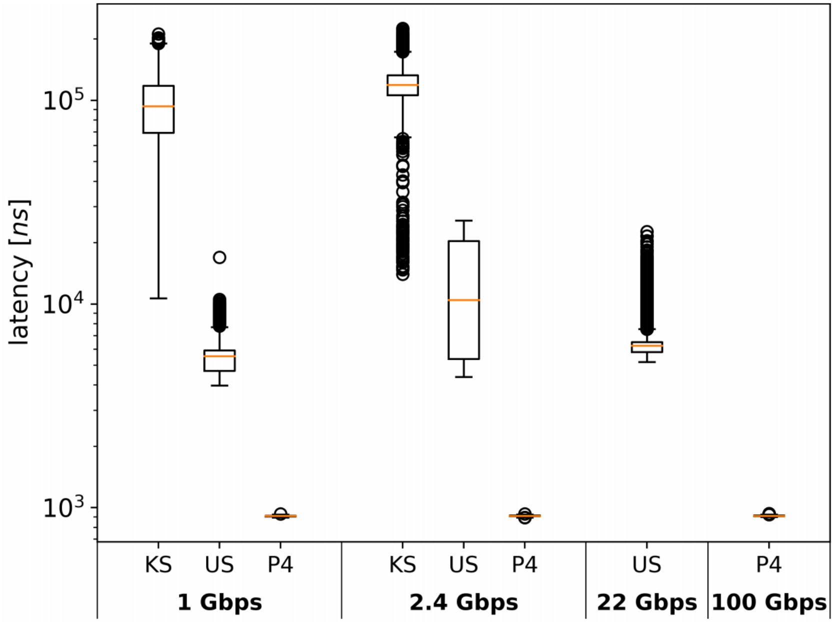


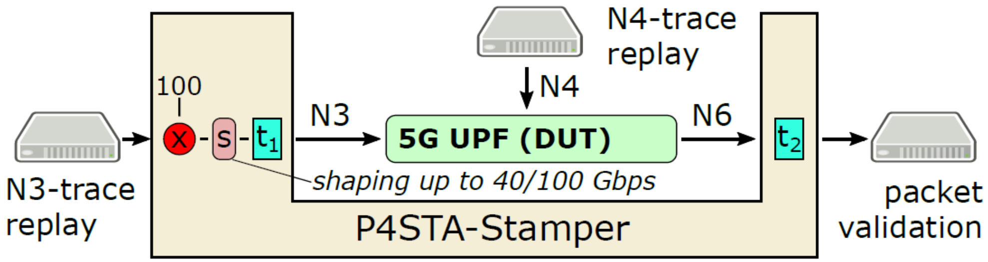
How to accelerate the User Plane Function (UPF)

The point of “failure”

- “non optimized” kernel space UPF of free5gc
- Xeon E5-2670v3





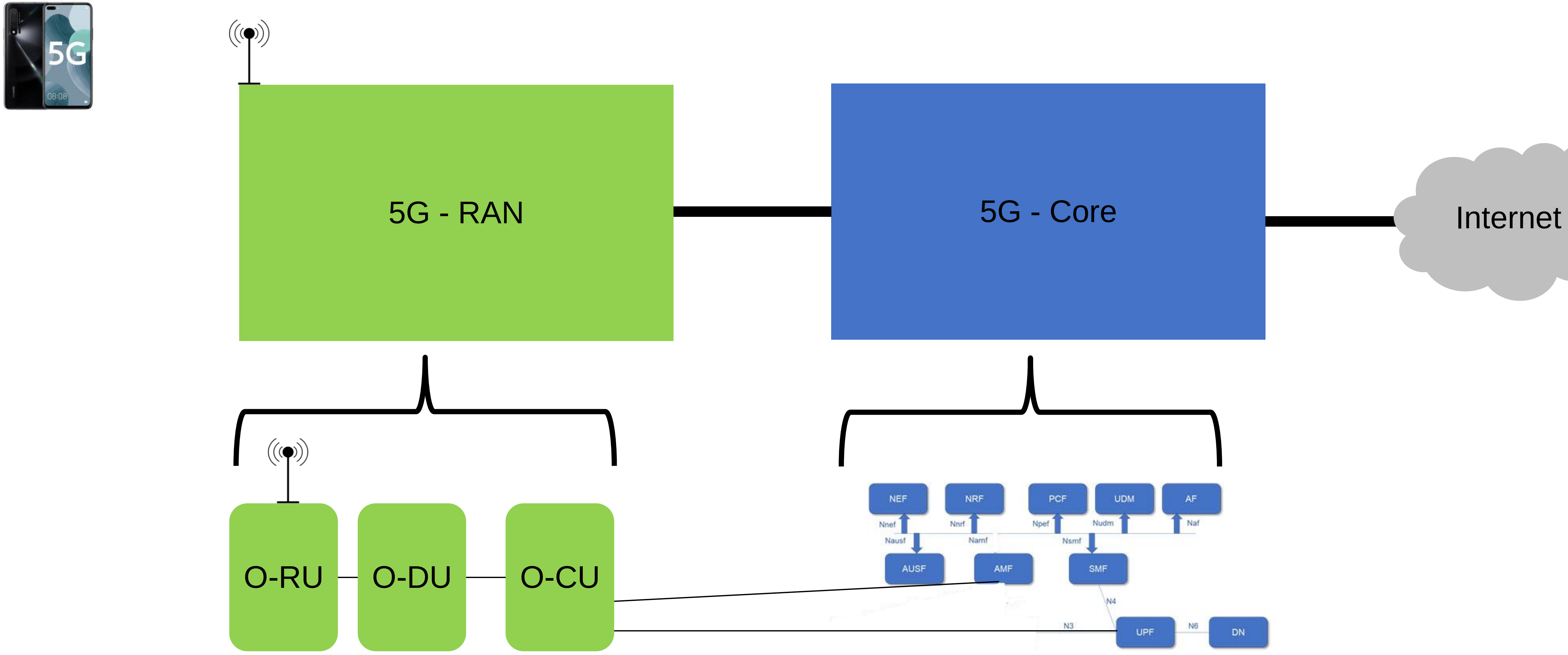


4. O-RAN

… going beyond the paper

What is a O-RAN user plane?

O-RAN 7.2 split



5G - RAN

5G - Core

Internet

O-RU

O-DU

O-CU

O-RAN @TU-Darmstadt

O-RAN

- Any Server(s)
- FPGA(s) for “DU”
- N * “blackbox” Remote Radio Units (RRU)
- Fronthaul Network with PTP timesync

Functionality

- Kubernetes (K8S)
- “just software”
- “Deploy RAN in 10 seconds”
- Replace RAN functionality
- Analyze behavior

Figure source: https://www.fibrolan.com/Falcon-RX,
https://benetel.com/wp-content/uploads/2020/12/Benetel-5G-O-RU-RAN550-Product-Specification-Sheet-Dec-2020.pdf

Next Steps

Make it fast!

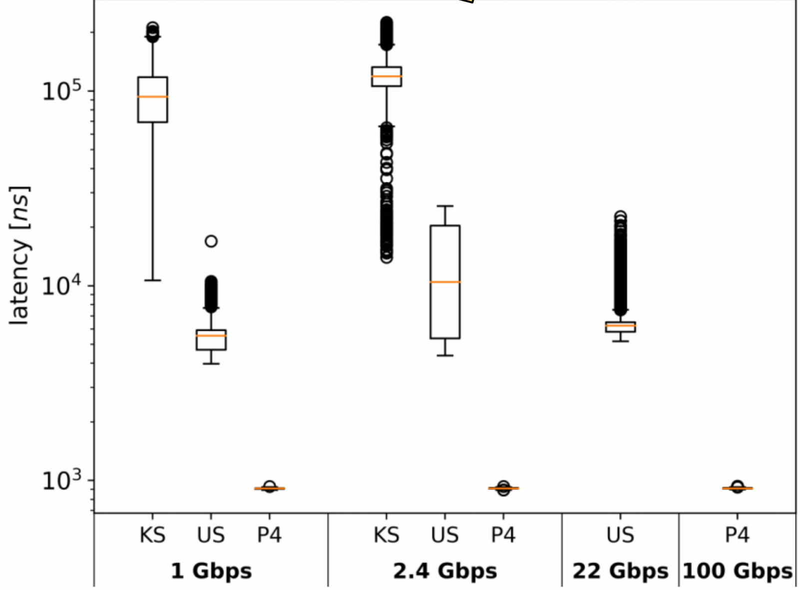
Make it open source!

RAN

- OpenAirInterface5G
  - Integration testing
- “Large Scale” O-RAN deployment
  - Free5gc control plane
  - P4-based UPF
- Accelerate the RAN
  - Current O-RAN has only very basic acceleration

Make the UPF Open Source

- Plan to release P4 UPF in Q2/2022
- Tofino +FPGA



Interested in more?

Ralf Kundel et al., **“User Plane Hardware Acceleration in Access Networks: Experiences in Offloading Network Functions in Real 5G Deployments,”** Proceedings of the 55th Hawaii International Conference on System Sciences, 01/2022.

User Plane Hardware Acceleration in Access Networks: Experiences in Offloading Network Functions in Real 5G Deployments

Ralf Kundel, Tobias Meuser, Timo Koppe, Rhaban Hark, Ralf Steinmetz
Techincal University of Darmstadt
{firstname.lastname}@tu-darmstadt.de

Abstract

Fulfilling the ambitious Quality of Service demands of today's wireless networks, especially low latency, high bandwidths and availability, is a big challenge for researchers, network architects, and operators. Each networking component on the data path between the user equipment and the destination data network, e.g., the Internet, must provide the highest performance to meet these requirements. This work demonstrates how different network elements of the user plane, describing the whole path of user traffic, can be sped up with different hardware acceleration technologies. For that, we demonstrate how to build up a 5G standalone campus network for evaluation, working end-to-end with real user equipment and open-source software components. Further, we analyze the user plane network functions of 5G networks from the radio access network to the core. Based on our real 5G setup, the practical evaluation of the analysis results shows up how the 5G user plane hardware can be accelerated best.

1. Introduction & Background

One main objective of 5G access networks is an increased Quality of Service (QoS), especially high throughput and low latency. All involved network functions and the underlying hardware must perform at the highest possible level to accomplish this.

One enabler for innovation in mobile access networks and computer networks, in general, is Software Defined Networking (SDN). SDN describes a paradigm for disaggregating network functionality [1]. Network switches are divided into a user plane (also known as data plane) and a control plane. The user plane is responsible for forwarding network packets based on simple rules only, while the control plane is responsible for computing and managing these rules. By introducing well-defined interfaces between control and user plane, the replacement of only a single user plane function or control function is eased.

Achieving the postulated QoS requirements in 5G networks (*e.g.*, one millisecond round trip time from devices to the core) requires focusing on the user plane for acceleration. In most of today's mobile networks, including modern 4G and 5G deployments, all network functions are realized in software environments fulfilling the performance goals neither in terms of throughput nor latency. However, this performance can be increased by offloading some functionality on programmable hardware accelerators. Besides the increase in performance, hardware acceleration raises new challenges and not every approach is suitable for every network function to be accelerated.

5G networks can operate either in non-standalone mode with a 4G-based anchor cell for authentication or as a pure 5G-based network, named 5G standalone. We focus on 5G standalone networks in the following. Figure 1 shows the 5G standalone architecture according to the 3GPP specification [2]. The functionality can be divided according to the previously introduced SDN terminology into a user plane and a control plane. The User Equipment (UE) connects with the Data Network (DN) (typically the Internet) through the 5G network.

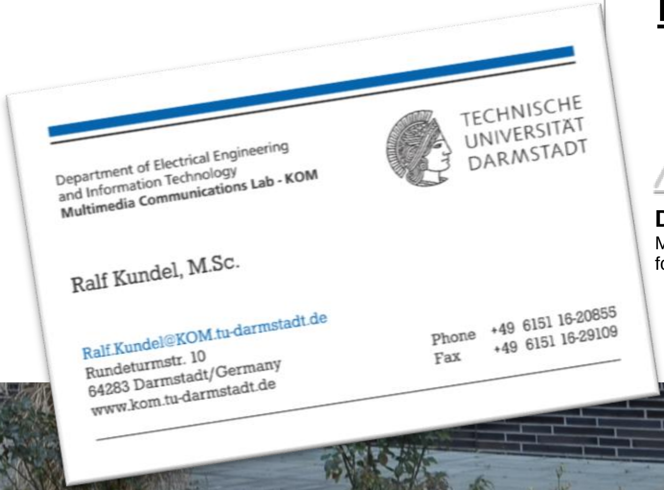
For that, the UE connects via the Radio Access Network (RAN) with the 5G-core at the specified reference points N1, N2 for the control plane and N3 for the user plane. The RAN functionality can be further divided into sub-functions, often named as “split”. In the terminology of the Open RAN Alliance (O-RAN) and the “7.2-split”, this is the Radio Unit (RU), Distributed Unit (DU) and Central Unit (CU). In other split scenarios, the DU and CU are combined to a CU/DU or BaseBand Unit (BBU) in the 4G terminology. However, the overall functionality remains unchanged.

As part of the 5G-core, the User Plane Function (UPF) terminates the Protocol Data Unit (PDU) session of many UEs and forwards the packets to and from the data network. On the N6 side of the UPF, standard IPv4/6 protocols are used; on the N3 side, for each PDU session a GPRS

Many thanks for your attention!

TECHNISCHE
UNIVERSITÄT
DARMSTADT

KOM – Multimedia
Communications Lab



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Department of Electrical Engineering
and Information Technology
Multimedia Communications Lab - KOM

Ralf Kundel, M.Sc.

Ralf.Kundel@KOM.tu-darmstadt.de
Rundeturmstr. 10
64283 Darmstadt/Germany
www.kom.tu-darmstadt.de

Phone +49 6151 16-20855
Fax +49 6151 16-29109

Funded by:

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

software
campus