P4-BNG

Central Office Network Functions on Programmable Packet Pipelines

@ *Conference on Network and Service Management (CNSM) 2019*

TECHNISCHE
UNIVERSITÄT
DARMSTADT

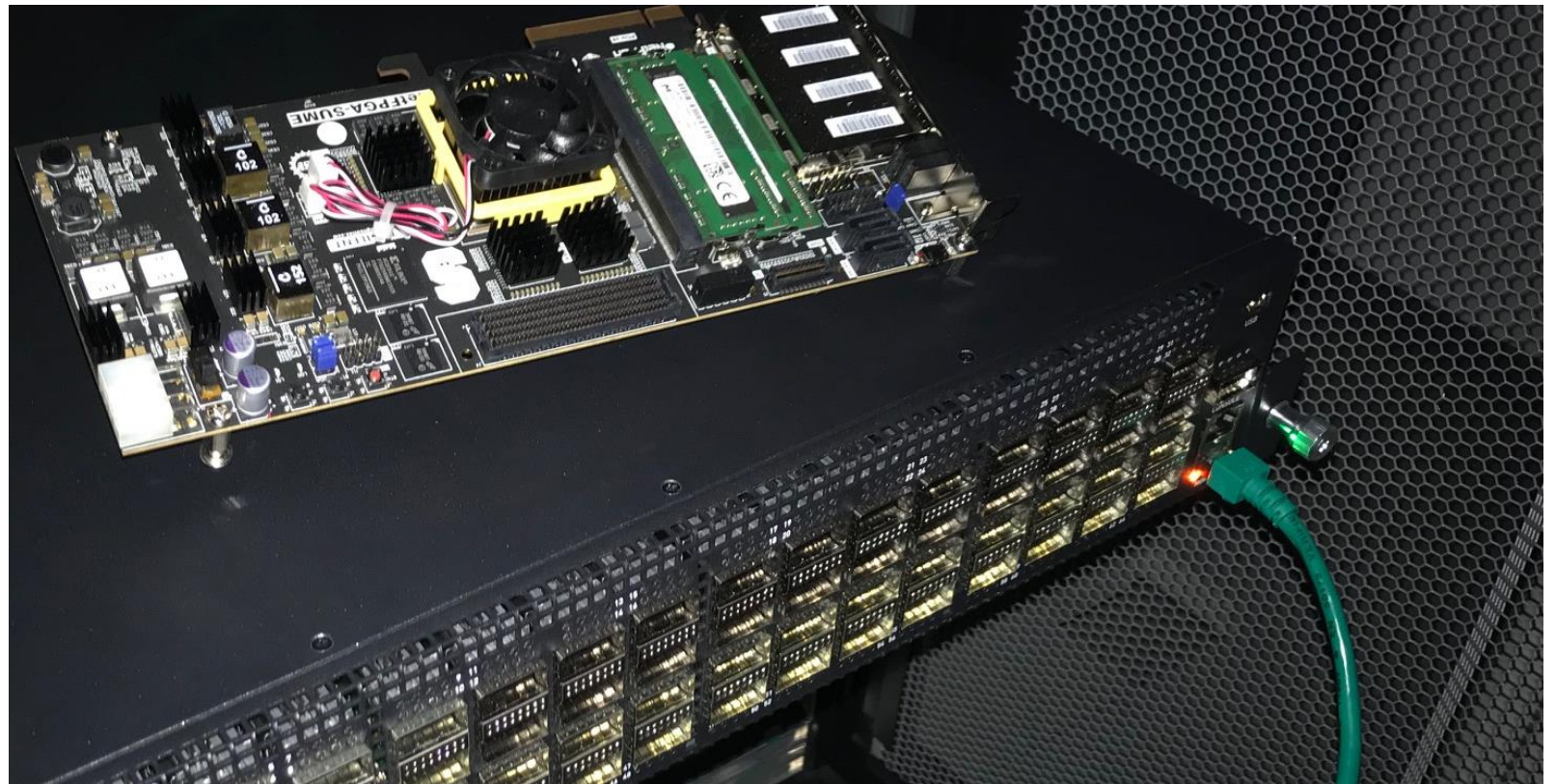**KuVS Fachgespräch "Network Softwarization"**
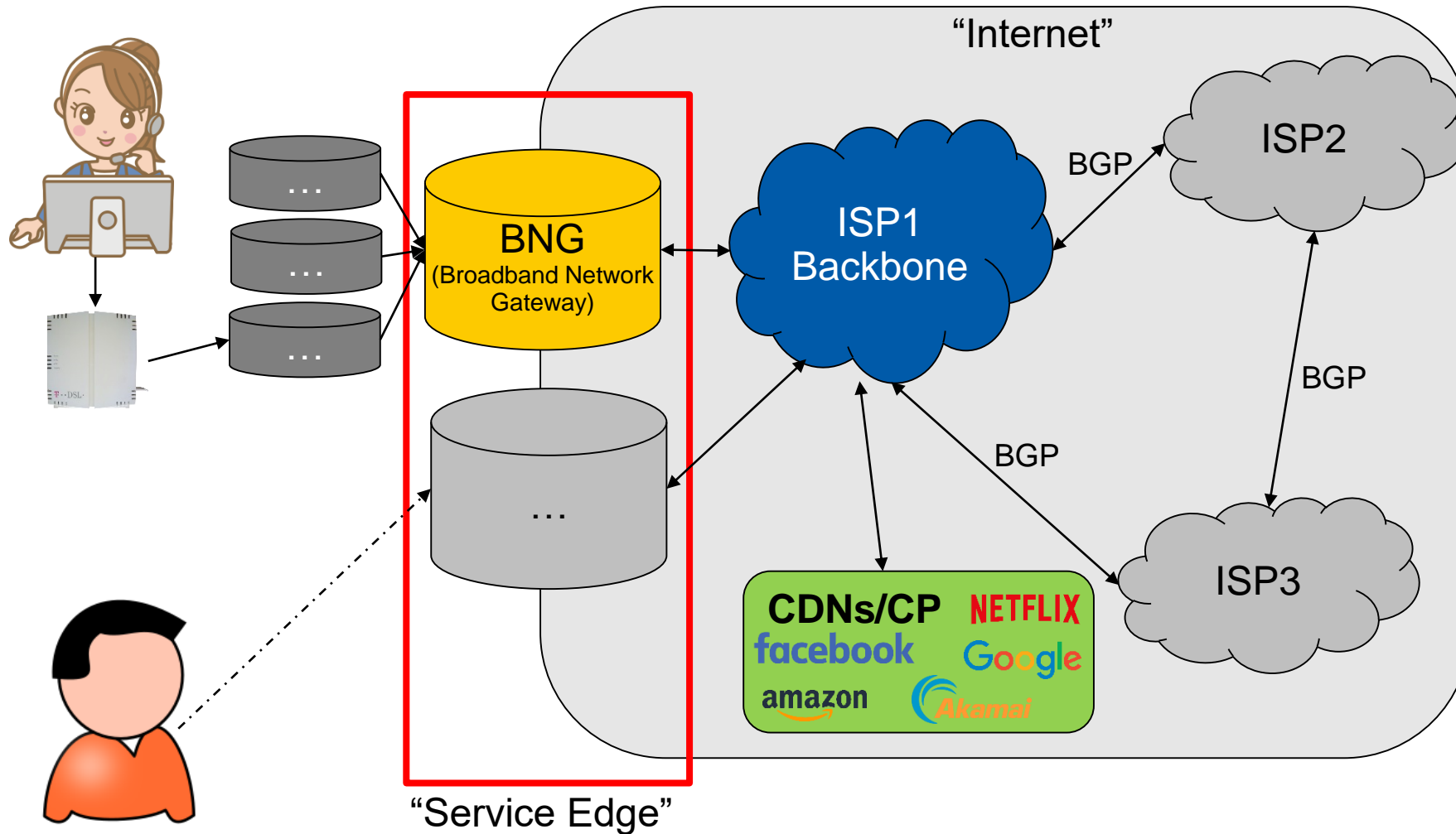April 2nd, 2020

Ralf Kundel   **ralf.kundel@kom.tu-darmstadt.de**
Leonhard Nobach
Jeremias Blendin
Hans-Joerg Kolbe
Georg Schyguda
Vladimir Gurevich
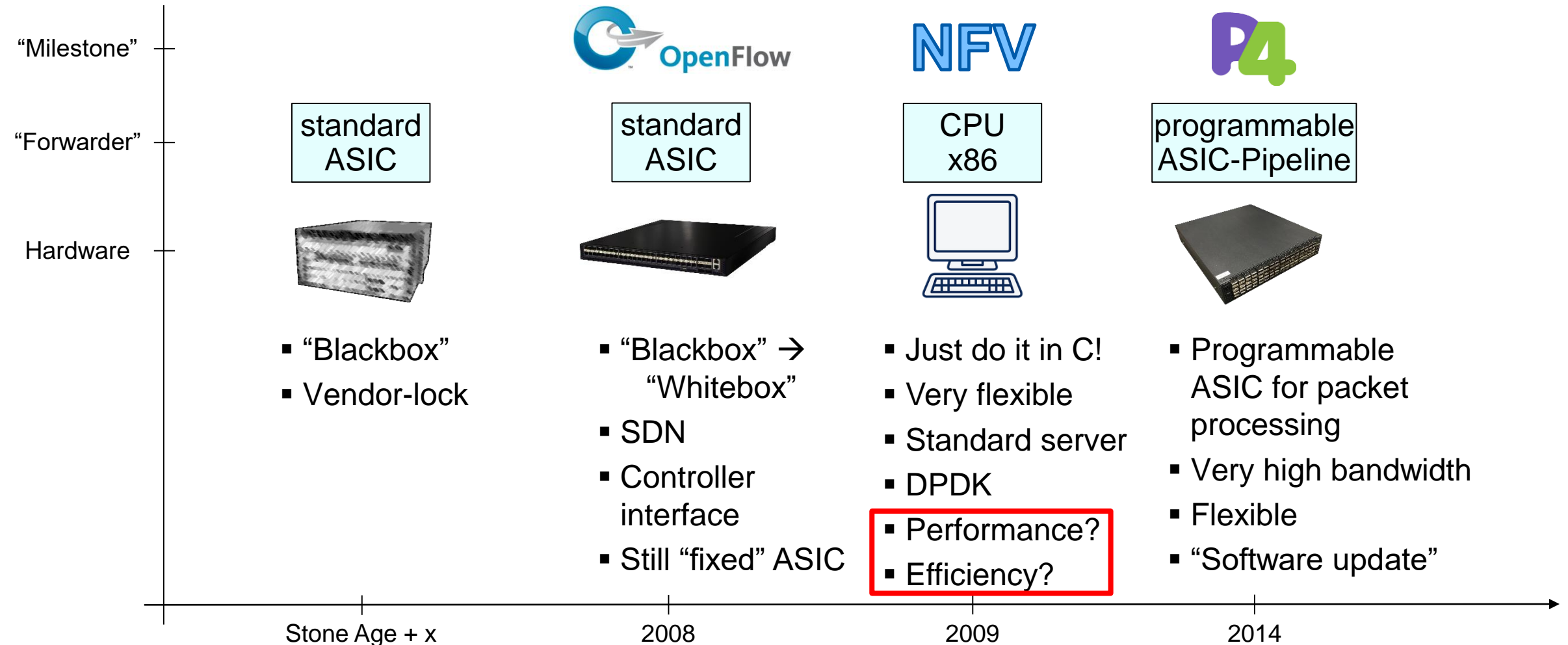Boris Koldehofe
Ralf Steinmetz

ISP (access) networks

"Internet"

Needed functionality:
- Packet forwarding
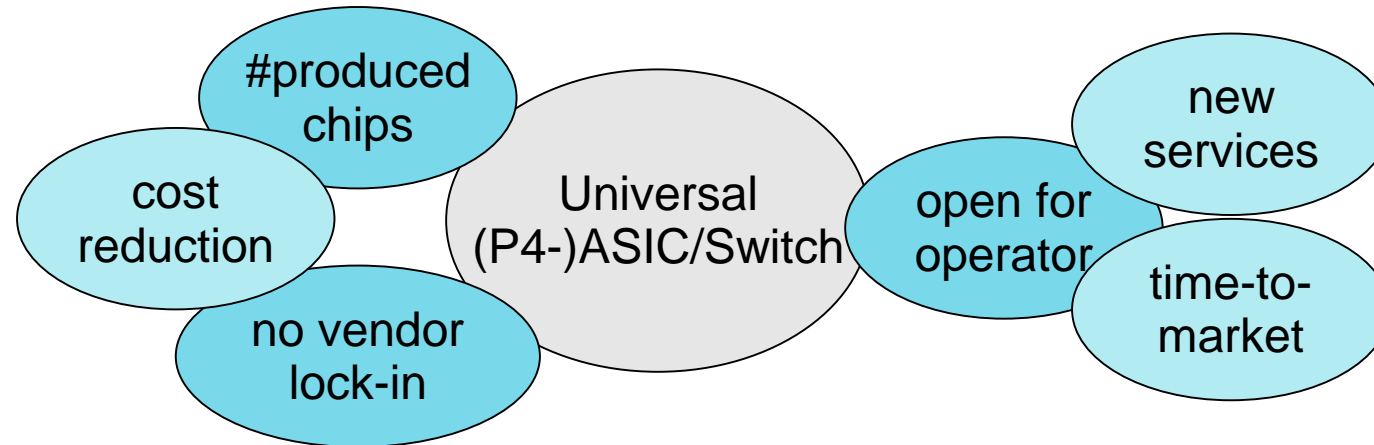- Authentication
- Tunneling
- Accounting
- Multicast
- QoS
- …

ISP1 Backbone

BGP

ISP2

BGP

ISP3

BGP

BNG (Broadband Network Gateway)

…

CDNs/CP NETFLIX facebook Google amazon Akamai

"Service Edge"

What do Telco's need at the service edge?

History of Data Planes

"Milestone"	**OpenFlow**	**NFV**	**P4**
"Forwarder"	standard ASIC / standard ASIC	CPU x86	programmable ASIC-Pipeline
Hardware			

standard ASIC
- "Blackbox"
- Vendor-lock

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

CPU x86
- Just do it in C!
- Very flexible
- Standard server
- DPDK
- Performance?
- Efficiency?

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

| Stone Age + x | 2008 | 2009 | 2014 |

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

Why innovation friendly Data Planes?

How to benefit from programmable hardware:

cost reduction · #produced chips · Universal (P4-)ASIC/Switch · no vendor lock-in · open for operator · new services · time-to-market

Challenge	"Blackbox" / fixed ASIC	Programmable DP
Time to market	month → years	"software update"
Special requirements	few products	"software update"
New features	buy new hardware	"software update"
Cost	feature dependent	constant low

Keep up with technology

Can BNG functionality be realized with programmable data plane hardware?

WHAT DOES RELATED WORK SAY?

Open CORD

L.Peterson et al.: "Central Office Re-architected as a Datacenter", 2016

TECHNISCHE
UNIVERSITÄT
DARMSTADT

The CORD Concept

Performance?

NFV

- Mini data centers with SDN & NFV at the network edge
- Use commodity data center technology to increase efficiency
- Initial approach: OpenFlow switches and virtual machines on x86
- Operate mobile and residential network access in a single CORD data center
  - R-CORD: Broadband Network Gateway (BNG)
  - M-CORD: Serving Gateway (SGW) + Packet Gateway (PGW) = SPGW
  - E-CORD: Enterprise Customers

"economies of scale"

agility of data centers

AT&T Open GPON
48-port, 1RU, I/O Box

OPEN

Leaf-Spine Fabric Built from White-box Switches

GPON (Access)

PON MACs

PON OLT MACs

ROADM (Core)

Commodity Servers, Storage, Switches, and I/O
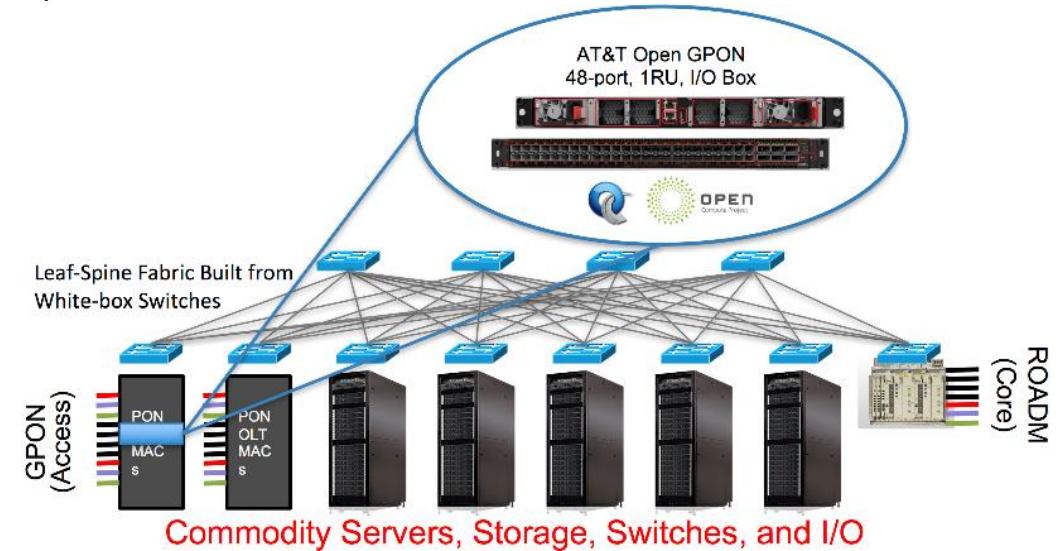
Bare-Metal Switches And Their Customization …

L.Nobach et al. @ Local Computer Networks (LCN) 2017

Bare-metal switch

- "not bundled with an operating system"
- Edge-Core AS5712-54x
- Broadcom Trident2
- OF-DPA

Implementing BNG (BRAS) functionality

- "80% possible"
- E.g. PPPoE missing
- Limited to silicon possibilities
- Very good latency/throughput

> standard ASIC
> with fixed functionality

P4
Programming **P**rotocol Independent **P**acket **P**rocessors

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Goal: An open source language allowing the specification of packet processing logic

- Protocol independence
- Reconfigure P4 devices by remote "software update"
- No predefined logic!
- "Hardware independent"

P4-BNG?

P4.org
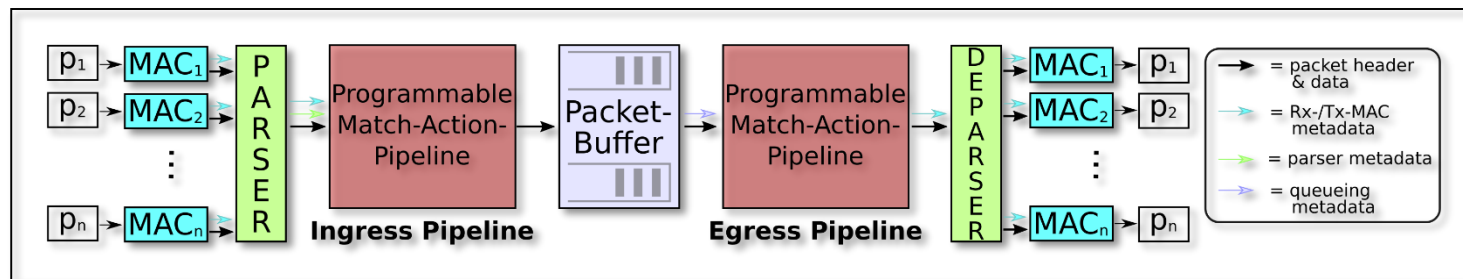
P4-ASIC

P4-SmartNIC

P4-NetFPGA

What we did:

Analysis and Survey on BNG Requirements

Implement BNG on P4-programmable Data Planes

Evaluation Results
- FPGA
- SmartNIC
- P4-ASIC

Disclaimer:
Vary from country to country!

Functional Requirements

BNG

Authentication, Authorization, Accounting

Access control, Security Assurance, DP/CP separation

Who?

- Residential Gateway (RG) <-> BNG

When?

- RG startup
- "blackout"
- Periodically?

How?
- Forward to CP
- Install DP rules
- RPF
- Traffic Separation

Why?

- "Bills payed?"
- Bandwidth Limits?
- Accounting
- "Regulations"
- Customer Separation

BNG-CP

BNG-DP
(Broadband Network Gateway)

ISP1
Core-Network

RG   AN
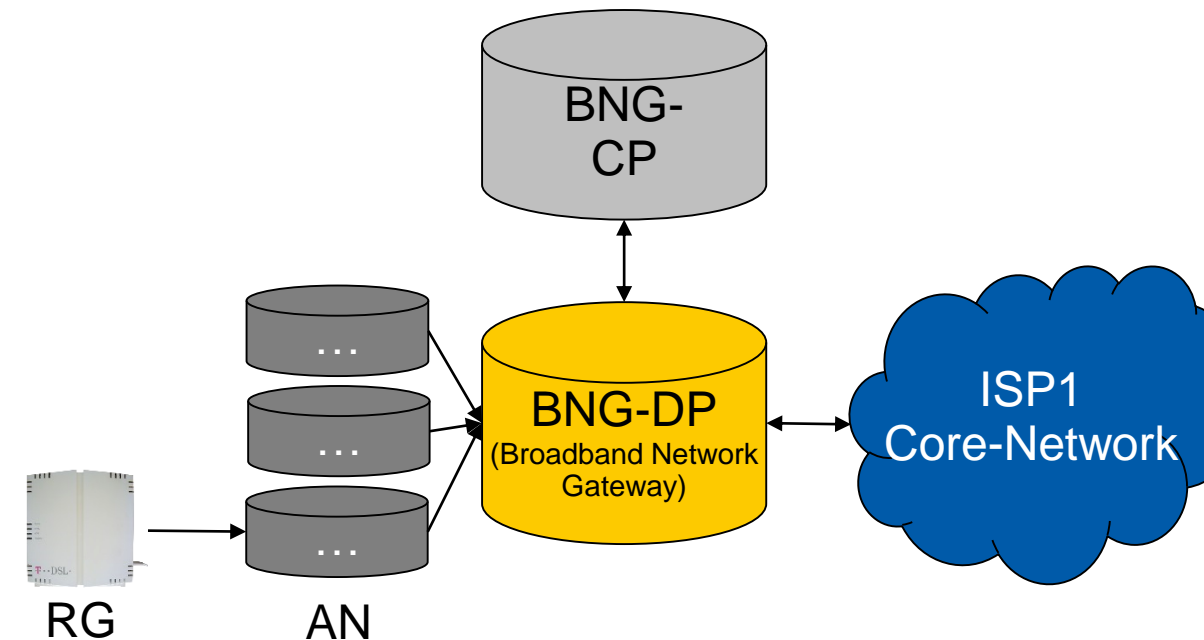
Customer Tunneling

Who?
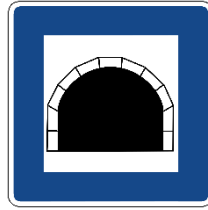
- Residential Gateway (RG) <-> BNG
- "virtual wire" = subscriber line

When?

- Result of AAA

Why?

- Customer Separation
- Prohibit Misuse
- Fairness
- Accounting

How?
- Encapsulation
  - E.g. PPPoE

RG   AN

BNG (Broadband Network Gateway)

ISP1 Core-Network
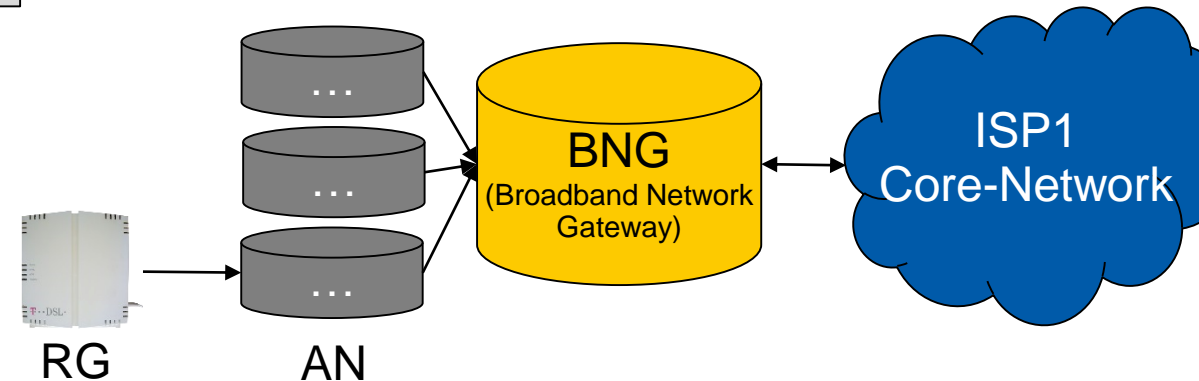
Traffic rate enforcement

What?

- Downstream traffic
- Upstream traffic

Where?

- At the BNG
  - Downstream shaping
  - Upstream Policing
- RG
  - Upstream shaping?

Why?

- Link speed
- Customer Contract

Bandwidth Limit

RG   AN   BNG (Broadband Network Gateway)   ISP1 Core-Network
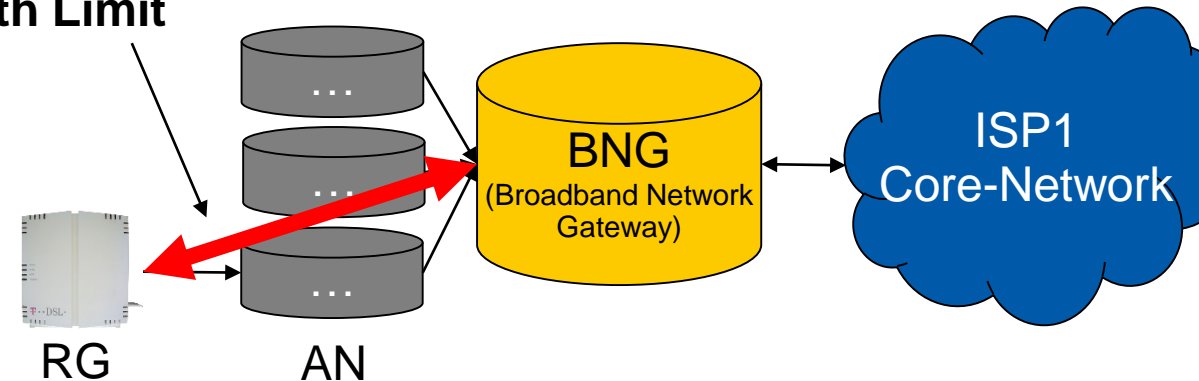
Multicast

What?

- Live TV

Where?

- At BNG

- Included in Bandwidth limits

How?

- Useful for live TV only
- MC-groups
  - Customer RG forwards subscriptions to BNG
  - Forward only requested streams
- Offered by ISP

RG

AN

BNG
(Broadband Network Gateway)

ISP1
Core-Network
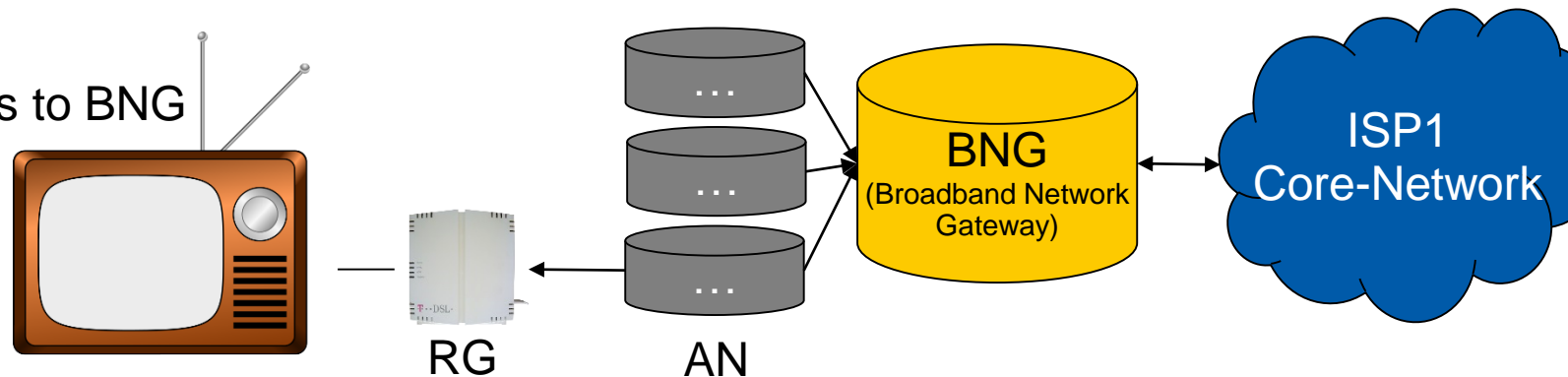
Implementation

P4-BNG

P4-BNG header processing – Implemented Design
after session setup – "normal case"

P4 control flow

https://github.com/opencord/p4se

Parser

0. What is the outer ethernet destination address and ethertype? (t_cptap_outer_ethernet)

MPLS, MAC==BNG

1. Is it an upstream or downstream labelport? (t_usds)

Down-stream

D2. Is the destination IP belonging to any subscriber network registered to a session? (t_ds_routev4)

CP

Control Plane

Yes

Upstream

Yes

U3. Does a PPPoE header exist and is it auth/control traffic? (t_pppoe_cpdp)

Yes

U2. Does a registered line for the VLAN and MPLS fields in the packet header exist and set the line ID? (t_line_map)

D3. Is it priority, best effort traffic, or should it be dropped? (t_ds_acl_qos_v4/v6)

Yes, user/IP traffic

U4. Is the PPPoE session ID and RG MAC address corresponding to an active session on this line? (t_antispoof_mac)

U6. Based on the destination IP address, drop or route the traffic to a specific destination? (t_us_routev4/v6)

Yes

Yes

U5. Does the source IP address belong to a network assigned to this session? (t_antispoof_ipv4/v6)

Final actions: Set new MPLS labels, source/dest. MAC decrement TTL

Deparser

Final actions: Set access header stack, source/dest MAC, decrement TTL, set PPPoE payload length

RESULTS

How to Measure a high performance BNG?

Ingredients

- Device Under Test (P4-BNG)
- P4-Switch as "Stamper"
- N load generation servers

Load Generation

- N servers
- Aggregated by „Stamper"
- Checksum verification can be done in SW

Measurement

- Timestamping in HW
  - t_1: before DUT
  - t_2: after DUT
- Loss detection in HW
- Result processing in SW
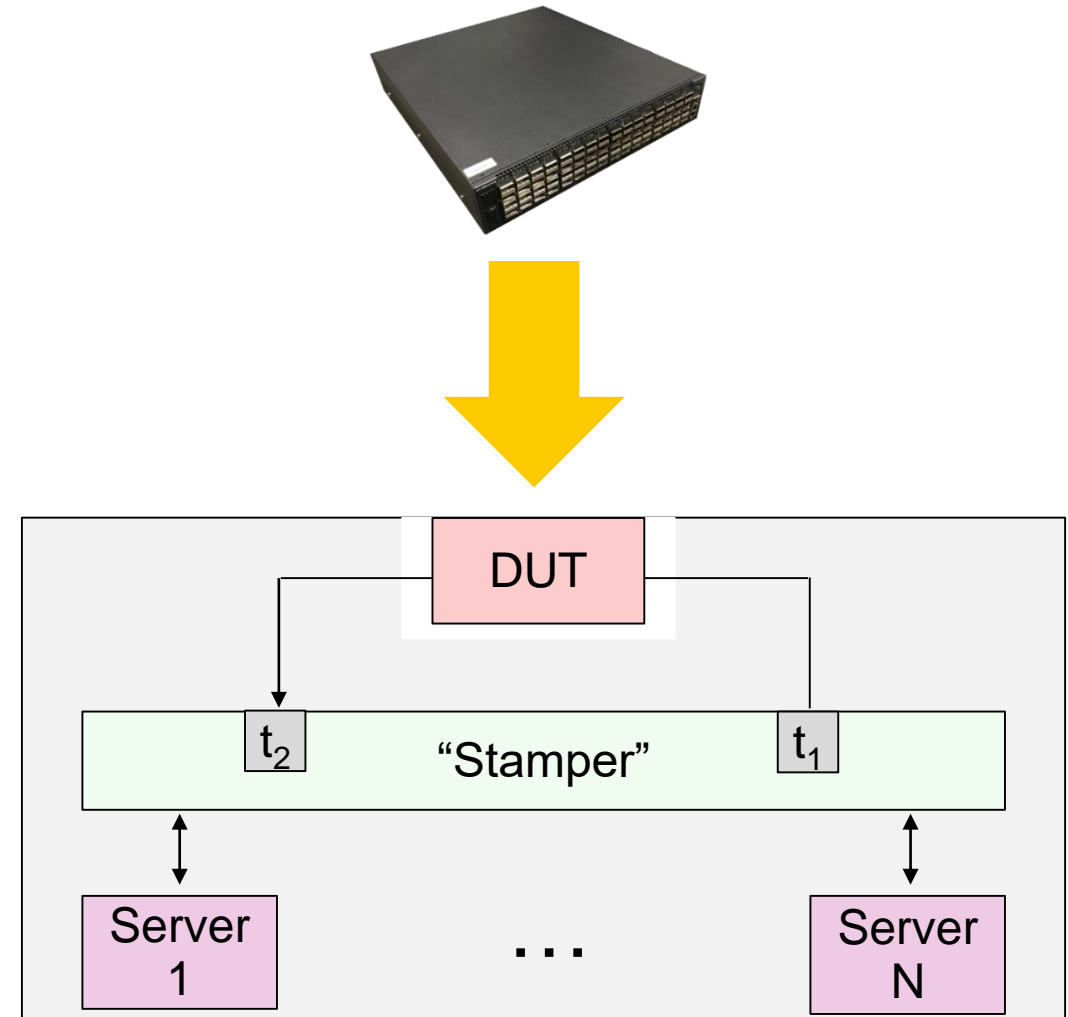
P4STA: A OpenSource Load Generation Framework

Current State

- OpenSource since 09/2019
- Currently „NDA" parts missing
- Supported Stamper devices
  - Barefoot Tofino
  - Netronome SmartNICs

Next Steps:

- Further Stamper
  - FPGAs, …
- …

https://github.com/ralfkundel/P4STA

Packet loss in P4

- What's that?
- Except for > 9.99 Gbit/s & "special packet sizes"
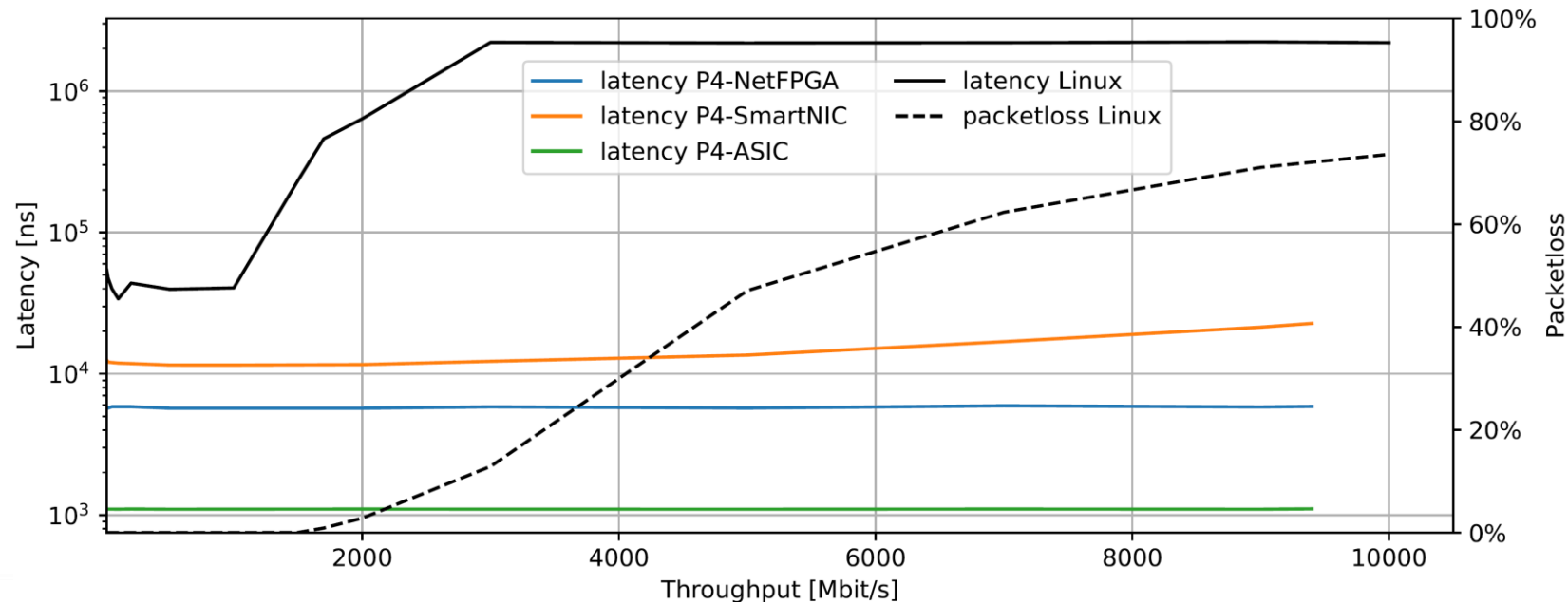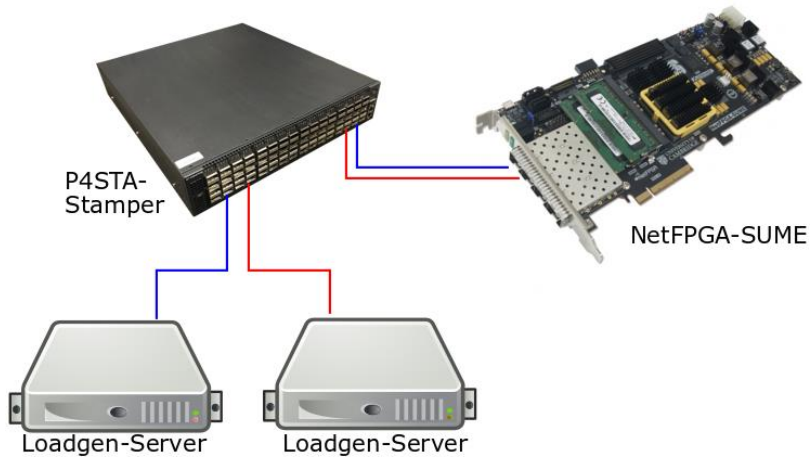
Latency Evaluation
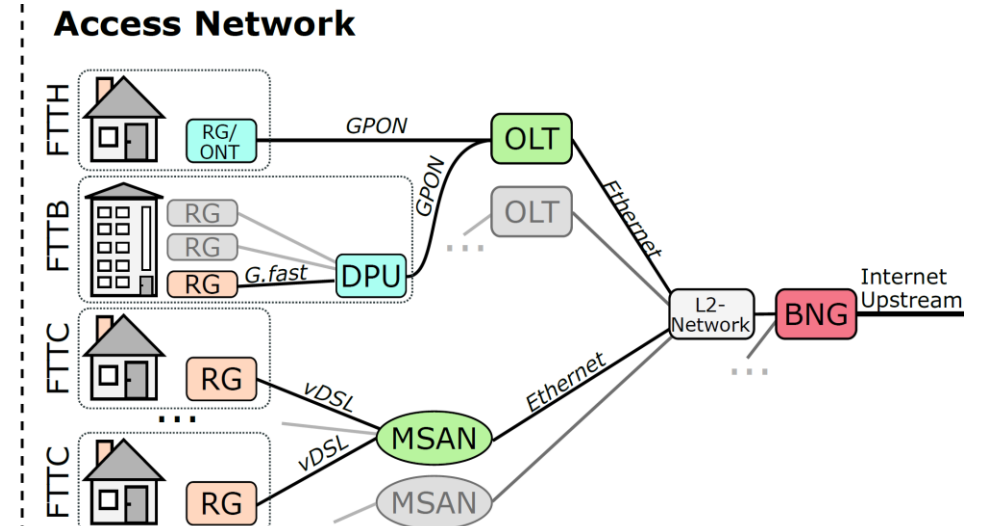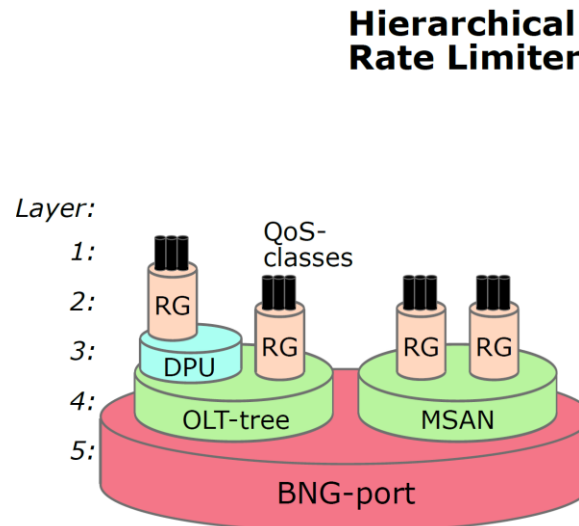## Number Subscribers
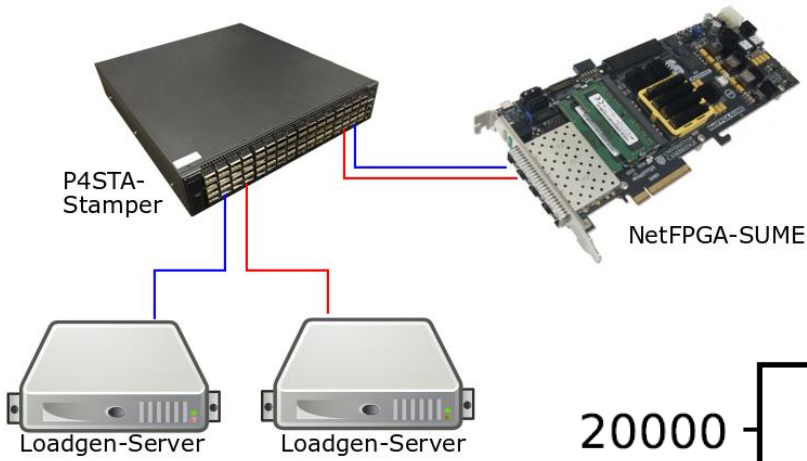
"Hardware either works or not!"
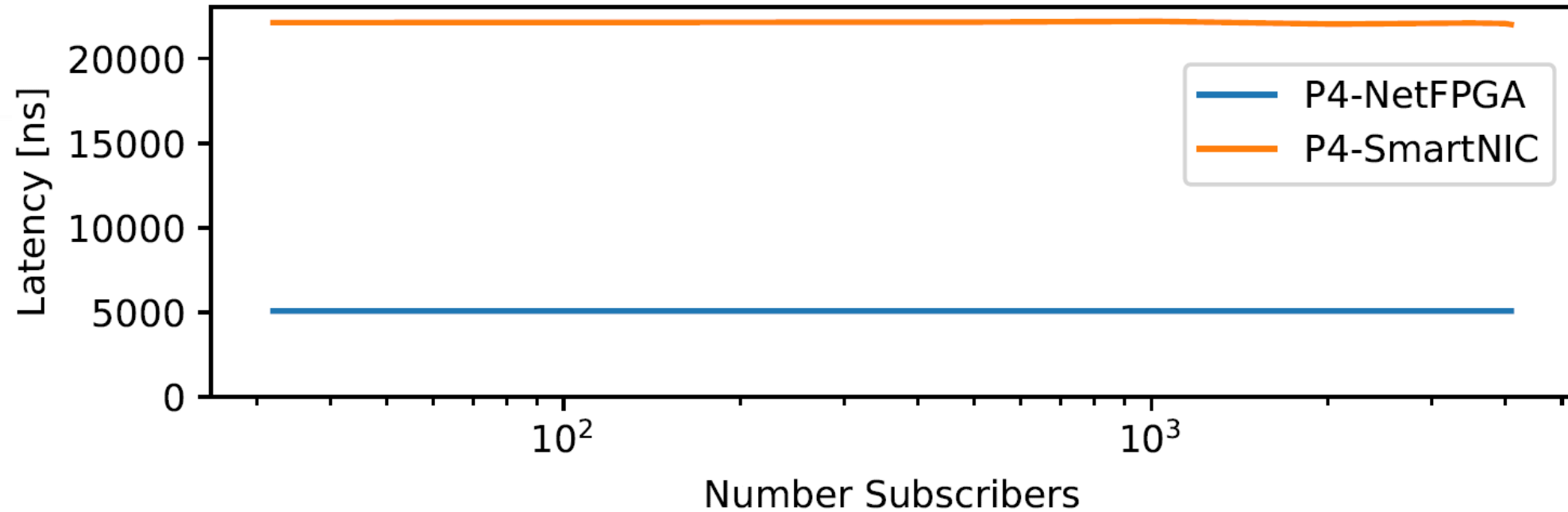
P4-BNG: production ready?

Needed functionality:

- Packet forwarding ✅
- Authentication ✅
- Tunneling ✅
- Accounting ✅
- Multicast ✅
- (H)QoS ❌
  - Queues not P4-programmable

Hierarchical Rate Limiter

Access Network

Summary and Future Work

Summary

- ISPs have very special requirements
- P4-BNG: It's possible!
- OpenSource available on GitHub

Future Work

- (H)QoS
  - Massive Queueing
  - AQM
  - Hierarchical Scheduling
  - ….

Questions & Contact

Deutsche Telekom

TECHNISCHE UNIVERSITÄT DARMSTADT

Ralf Kundel

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

TECHNISCHE
UNIVERSITÄT
DARMSTADT

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

https://github.com/opencord/p4se