

Using Programmable Chip and Open Source SW Toward Disaggregated Network Packet Broker and 5G UPF

P4 Workshop

May 1st, 2019

Chris Sueng-Y. Park, Ph.D.
KulCloud Corp.



INDEX

001. Introduction to SK Telecom Integrated Network Analyzer (TINA)

002. Architecting Network Packet Broker (NPB)

003. Making Disaggregated NPB

004. Utilizing P4 for 5G User Plane Function

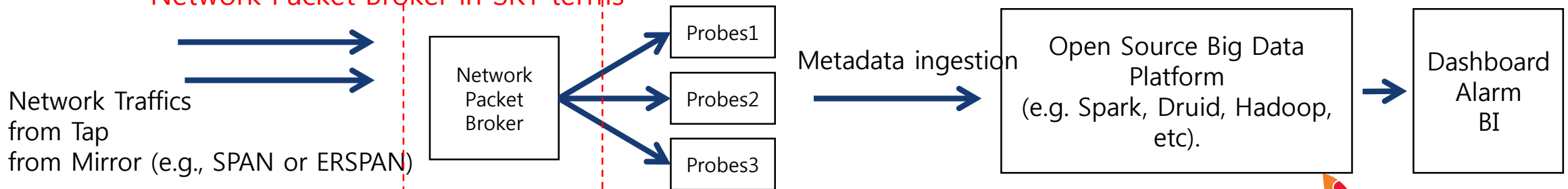
001

Introduction to SK Telecom Integrated Network Analyzer (TINA)

TINA: SK Telecom Integrated Network Analyzer

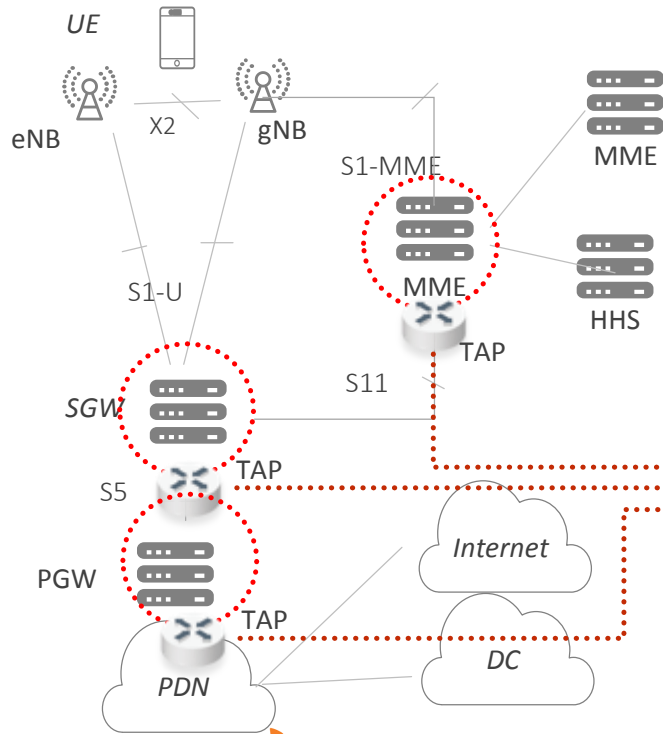
- A high performance NPMD (Network Performance Monitoring and Diagnostics) platform
 - Provides the end-to-end visibility, the network monitoring data and the troubleshooting information for network operators.
 - Collects, processes, and analyzes the network traffic and the link/device status of the physical/virtual networks.
 - Open collaboration with the in-house departments and the engineering partners.
 - Currently analyzes Public/Private cloud services, 4G/5G mobile network services, ISP (SKB IPTV service).

Today's talk covers
 Network Visibility Node, a.k.a.
 Network Packet Broker in SKT terms

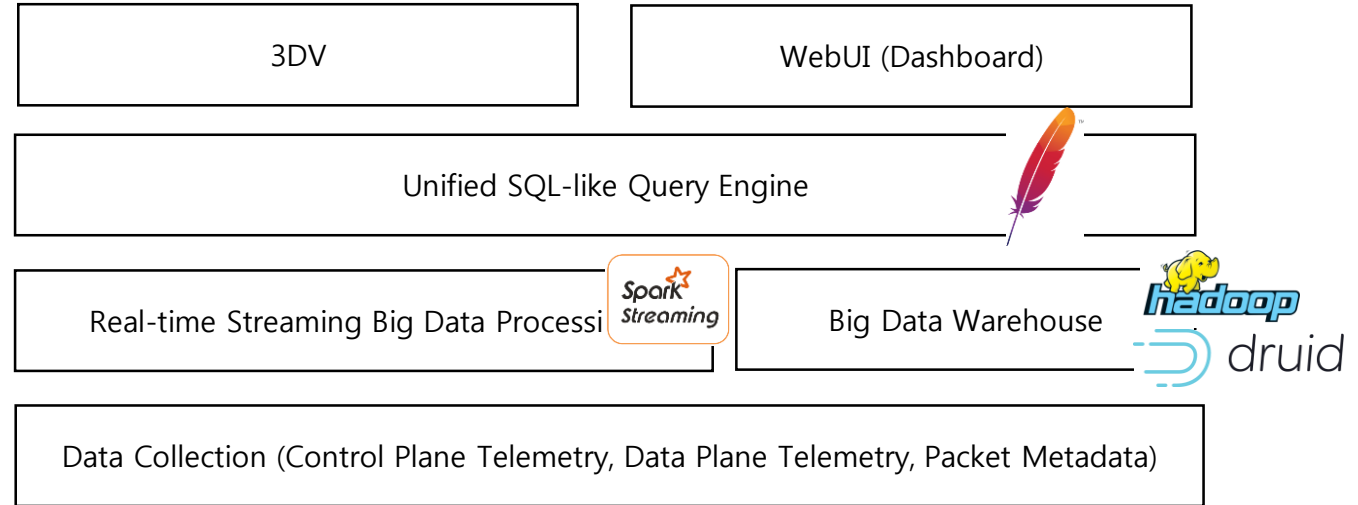


TINA: Architecture

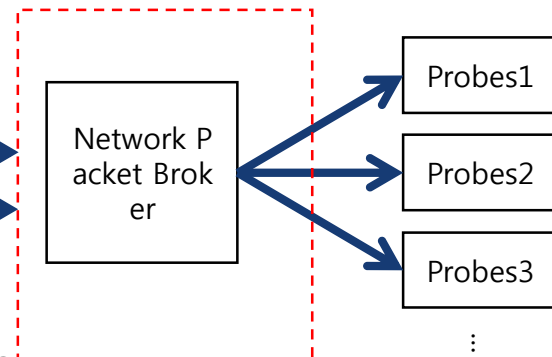
Monitors SGW, PGW, and Core Network NFVs.



Big Data Analysis Engine



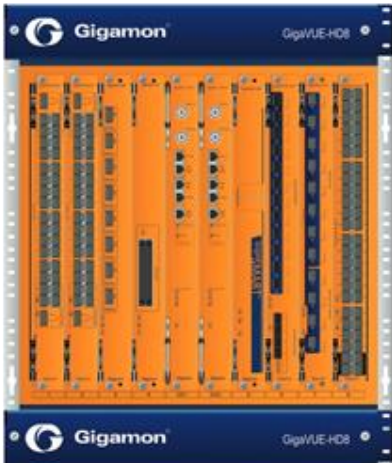
Network Traffic from Tap from Mirror (e.g., SPAN or ERSPAN)



e.g., IDS, Firewall, Netflow generator

Problem Statement: Network Packet Broker (NPB)

- Aggregating, filtering, load balancing traffics from tap/mirror to the monitoring tools in the probes.
 - Usually multi-million dollar hardware (e.g., Gigamon, Netscout, Ixia, ...)

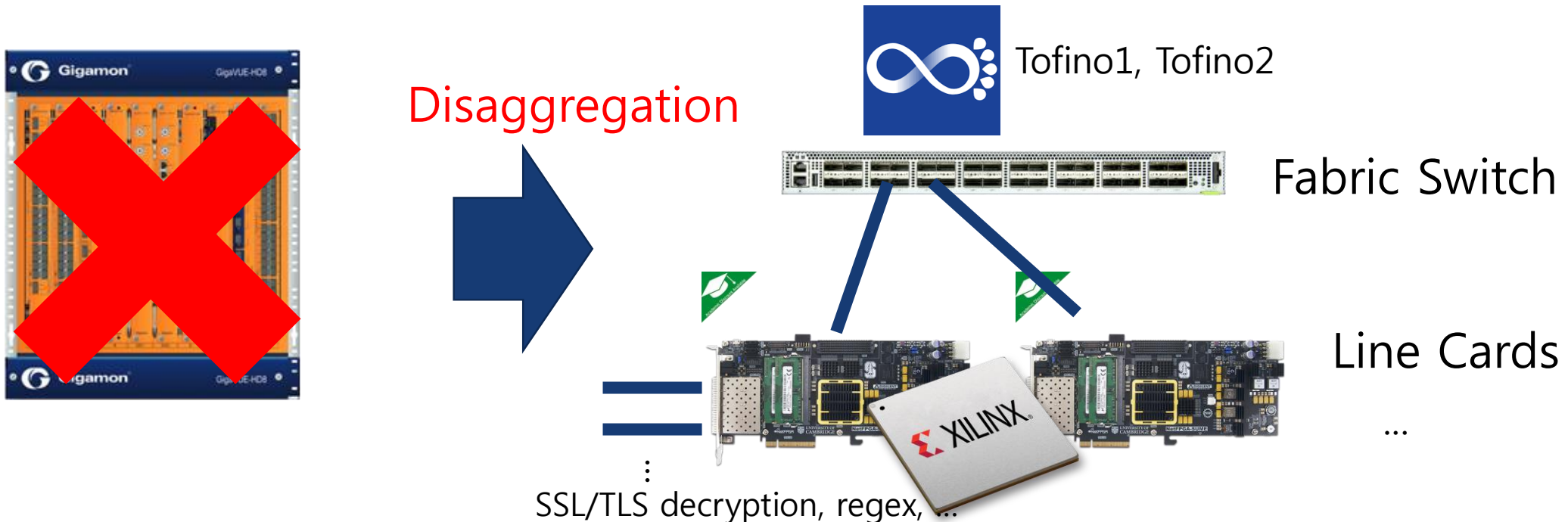


Limitations of legacy NPB

- Telco system is a very complex integration of various devices.
 - Telco asks for the customizations of NPB. But they are either very costly or impossible to implement.
- A long lead time to deliver the high performance device.
 - Telco needs 25G/100G per port interface now. But they are not available from vendors.
- Restrictions on certain features.
 - Restrictions on data storage options , APIs etc.

Problem Statement: Network Packet Broker (NPB)

- SKT decided to design a highly customizable and scalable disaggregated NPB with the Tofino programmable data plane.
 - Applied the well-studied Leaf-Spine architecture to smartNIC-Tofino.



002

Architecting Disaggregated Network Packet Broker

Workflow of NPB

Classification & Forwarding

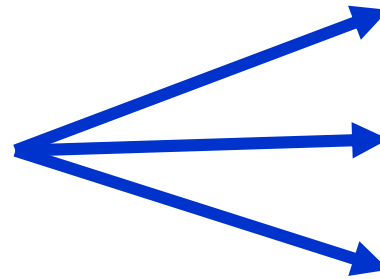
L4+ Processing
Context Processing



Tofino1, Tofino2



Fabric Switch



Workflow of NPB

Classification &
Forwarding



Tofino1, Tofino2



Fabric Switch

Classify by:

- Port / Port Group.
- User-defined ingress/egress filters.

Forward to egress ports by:

- Simple forwarding/Load balanced forwarding (normal, symmetric, resilient)/Replicated forwarding.
- IEEE 1588 Time stamping (nsec granularity).
 - 64-bit time stamp (32b second, 32b nanosecond).
 - ERSPAN type III or INT (In-band Network Telemetry).
- GTP session-aware traffic steering*.

Manage & Report by:

- C/C++ API.
- Metadata generation* (eg. Netflow etc).

*: To do list for SKT

L4+ Processing Context Processing*



Workflow of NPB

L4+ Processing:

- regex matching and filtering on SmartNIC.
 - Application aware?

Context Processing:

- GTP session aware correlation/load balancing.
- Encryption/decryption of MACsec, IPsec tunnels for the system overlay.
- Packet deduplication (possible? open for opinions)*.

Synchronization:

- 1588 time synchronization between SmartNICs, and time stamp 64b UTC TOD appended on the end of payload.
- Fabric switch supports Transparent Clock.

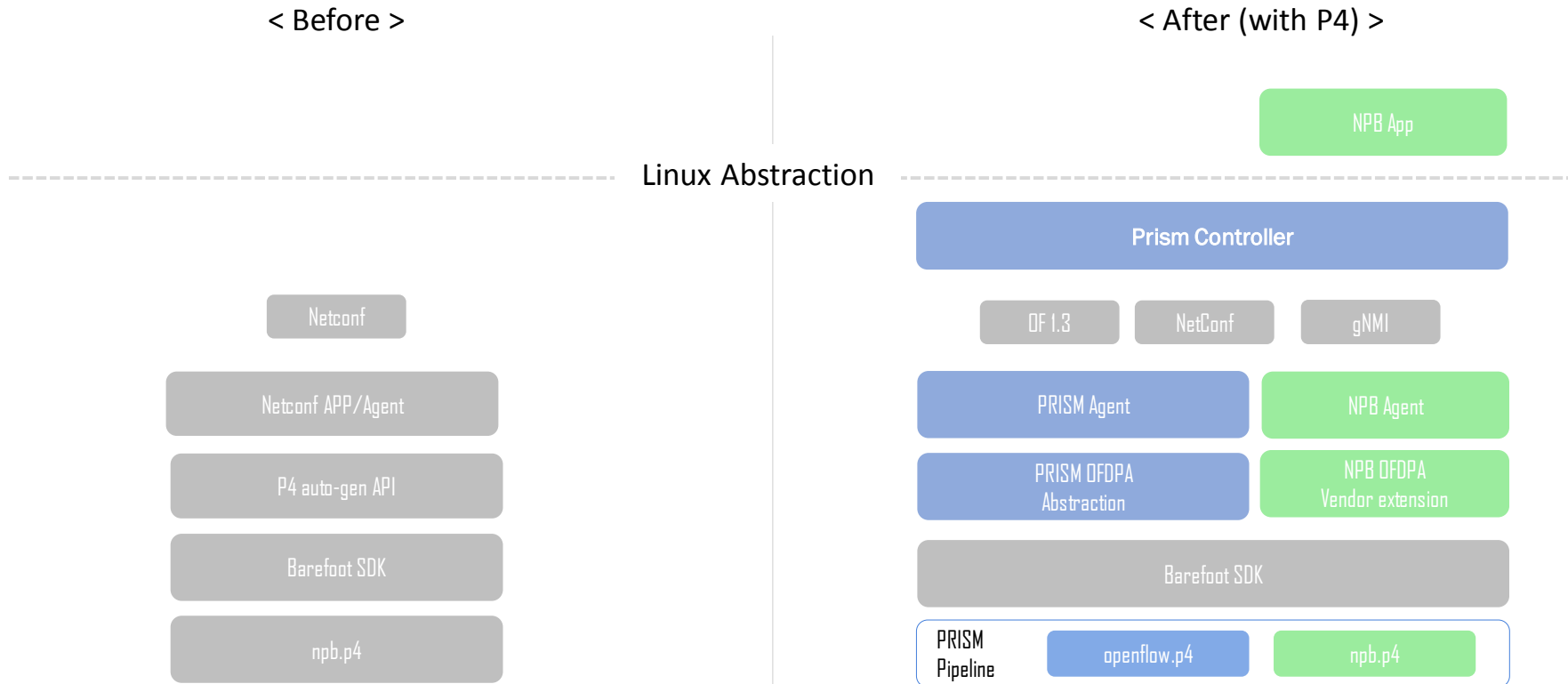
*: To do list for SKT

003

Making Disaggregated Network Packet Broker

Why P4?

- **[Programmability]** New data plane functions (not possible with Openflow).
- **[Extensibility]** Easy to make & extend.
- **[Synergy with PRISM, the in-house solution]** Offload Linux network apps to Tofino with P4.

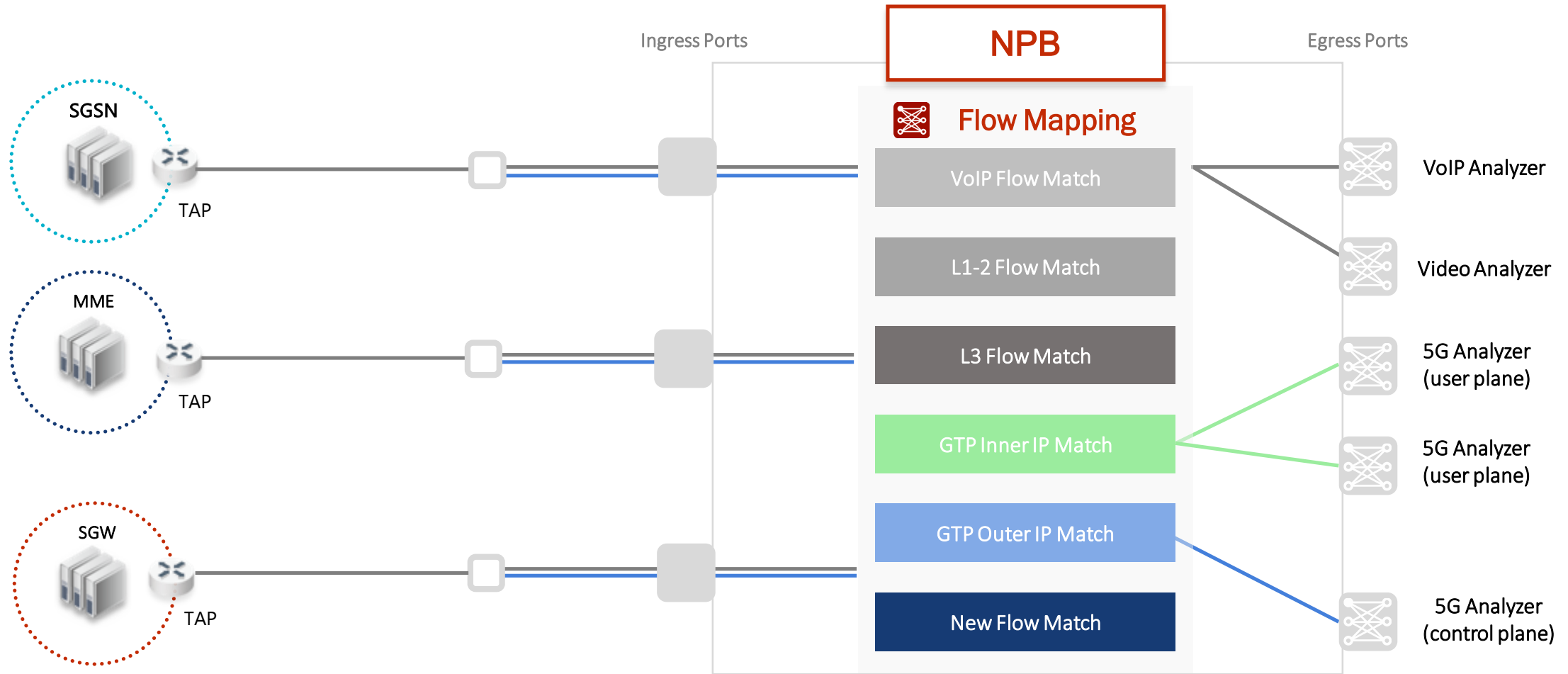


How to make a disaggregated NPB ?

- **PRISM Controller**
 - SDN controller providing out-of-box SDN NPB applications on top of Linux.
- **NPB Agent**
 - Production grade agent providing north-bound Openflow, Netconf, gNMI and CLI interface for NPB app.
- **NPB OFDPA Abstraction**
 - It is a thin lightweight abstraction layer written on top of switch API/BF Runtime to provide OF-DPA object abstraction.
- **npb.p4**
 - It is based on reference switch.p4 from BF SDE.
 - Contains additional changes to support full NPB pipelines.



Applied to Live Mobile Network 3G/4G/5G*



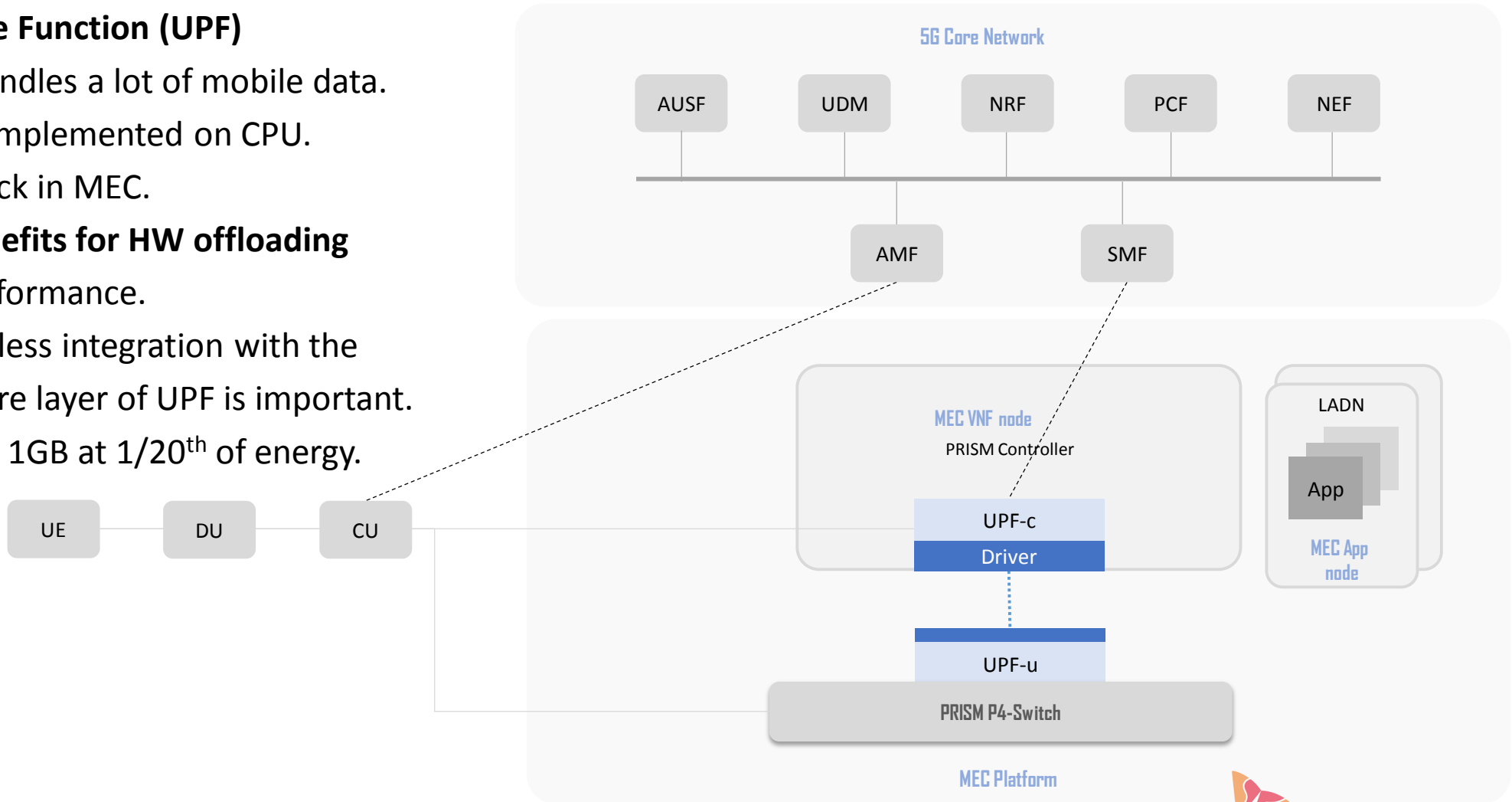
* Application to live 5G network is a work-in-progress. Done by YR 2019.

004

Utilizing P4 for 5G User Plane Function (UPF)

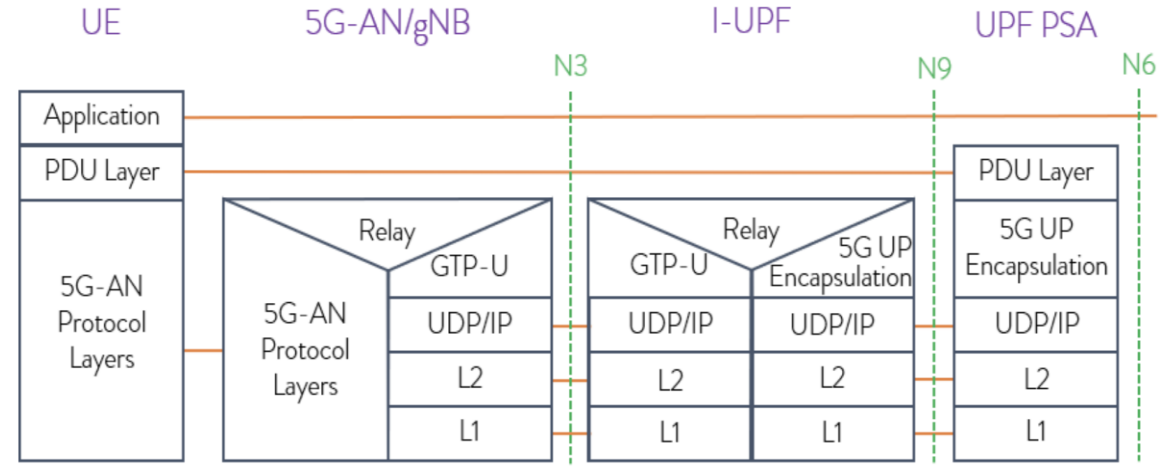
PRISM-P4 5G Use case

- **5G User Plane Function (UPF)**
 - A GW handles a lot of mobile data.
 - Usually implemented on CPU.
 - Bottleneck in MEC.
- **Expected Benefits for HW offloading**
 - High performance.
 - Seamless integration with the software layer of UPF is important.
 - Switches 1GB at 1/20th of energy.



PRISM-P4 5G Use case

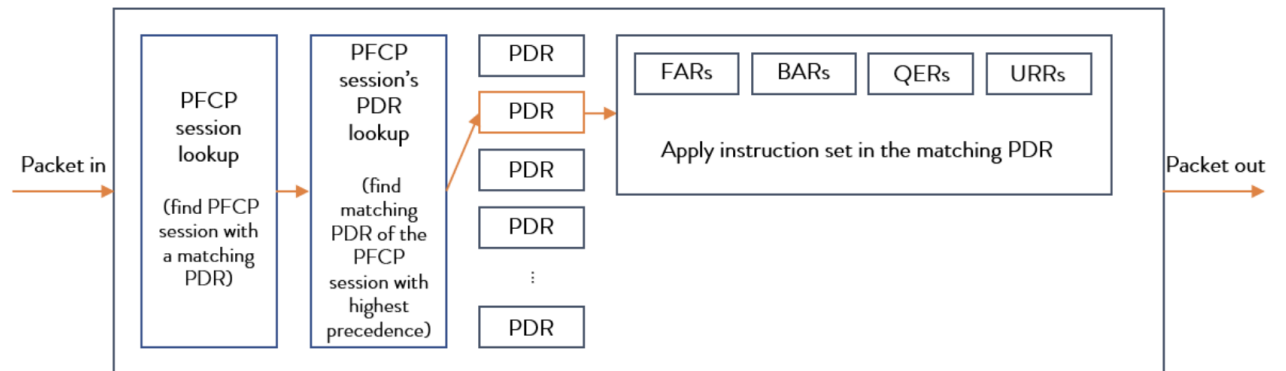
- PFCP's (Packet forwarding control plane) scope is similar to that of Openflow.
- However, it was engineered for the particular use cases of mobile Core Networks (CN).
- Most of the UPF packet processing can be implemented in Barefoot using P4.
- Only BAR (Buffering action rules) can not be directly offloaded to Tofino.



5G User Plane Protocol Stack, employing STP with 5G header extensions

Scaling and other concerns:

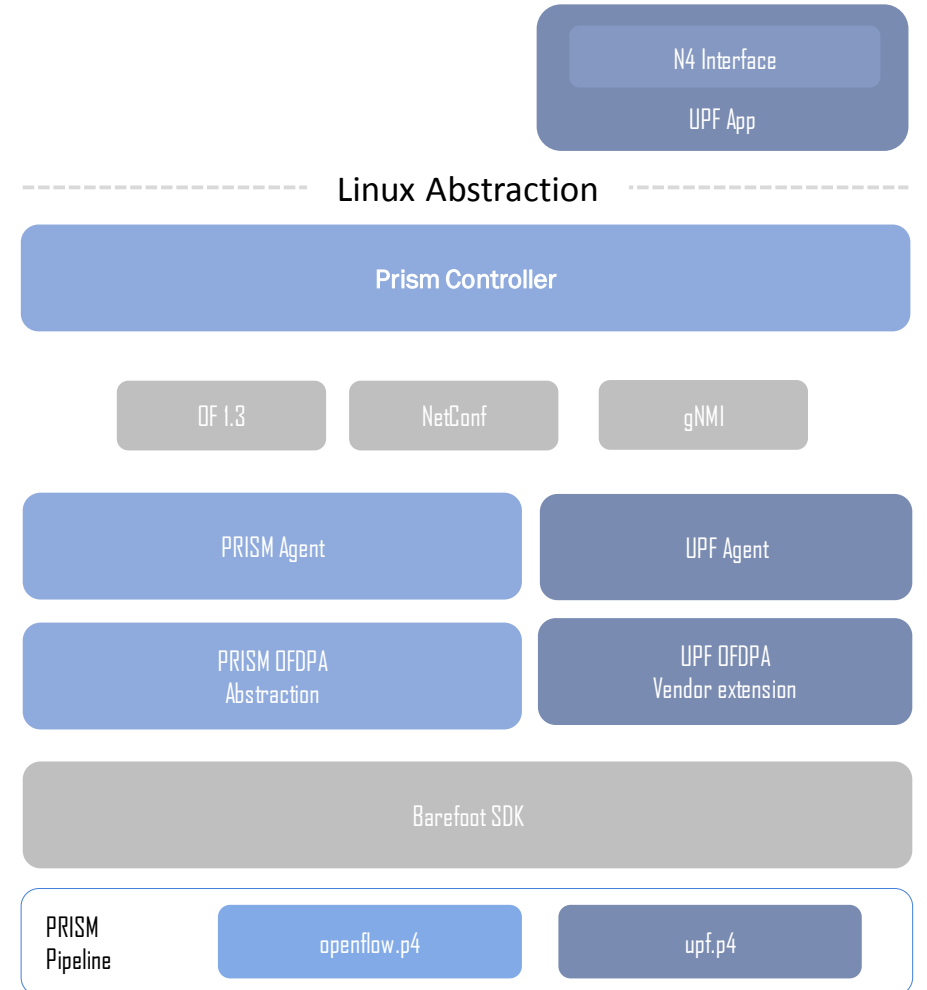
- How many sessions can be actively maintained in Tofino ?
 - Max. TCAM/Exact match table sizes ?
- How many counters and meters ?
- How to efficiently implement Buffering ?
 - Combination of Tofino with Agent Control plane ?



Packet processing flow in the User Plane Function (per 3GPP TS29.244)

How to make a disaggregated UPF ?

- **PRISM Controller**
 - SDN controller providing out-of-box SDN UPF applications(supporting N4 Interface).
- **UPF Agent**
 - Production grade agent providing north-bound Openflow, Netconf, gNMI and CLI interface for UPF rule configuration.
- **UPF OF-DPA Abstraction**
 - It is thin lightweight abstraction layer written on top of switch API/BF Runtime to provide UPF object abstraction.
- **upf.p4**
 - It is based on reference switch.p4 from BF SDE.
 - Contains additional changes to support full UPF pipelines.



References

FPGA Test Environment

