Open. Together.
Cost-effective Approach for Telco Network Analysis in 5G

Junho Suh, Research Engineer, SK Telecom
Agenda

• Introduction to Project TINA
• Software
  • Programmable Network Packet Broker (NPB)
  • DPDK-based Probes
  • DPDK-based Network Packet Capture (FloX)
• Hardware
  • 1st-gen network appliance hardware design (T-CAP)
  • Disaggregation approach
Project TINA (SKT integrated Network Analytics)

Network Visibility Platform
- Provide end-to-end visibility to Telco network operators
- Reduce TCO of Telco network monitoring tools
- Open hardware/software approach
- Launched @2017
High-Level Design

This Talk

Production Data Center Network

Cloud Service Monitoring
DDoS Attack
Traffic Anomaly Detection
Telco KPI Monitoring

Open Source Big Data Analysis Platform

OSS/BSS

DCN Probe
Mobile Networks Probe
Packet Capture Record

Tap
Mirror (SPAN/ERSPAN)

Open. Together.
Use Case #1 - Network Performance Monitoring for B2C/B2B Services

Running 200+ B2C/B2B services including T-View (CCTV Cloud), NUGU (AI-based Voice Assistant), T-Map (Mobile Navigation app.), etc

What to measure?

• Total traffic volume in bytes, packets, and number of flows per each service
• TCP session context
  • TCP seq. number, SYN, ACK, RST, …
• Anomaly detection using ML technique
  • Abrupt traffic explosion/reduction
  • Volume-based DDoS attack
Use Case #2 - IPTV Service Quality Monitoring

Serving IPTV content across 600+ domains

What to measure?
• DNS resolution result from GSLB
• Where is the location of content server?
Use Case #3 - LTE/5G Network Quality Monitoring and Troubleshooting

<table>
<thead>
<tr>
<th>Call Flow Types</th>
<th>Call Flow Types</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4G / 5G NSA</strong></td>
<td><strong>5G NSA</strong></td>
</tr>
<tr>
<td>• Initial Attach</td>
<td>• Secondary Node Addition</td>
</tr>
<tr>
<td>• Attach after Paging</td>
<td>• Secondary Node Change</td>
</tr>
<tr>
<td>• Mobile-triggered Service Request</td>
<td>• Inter-Master Node Handover with/without Secondary Node Change</td>
</tr>
<tr>
<td>• Network-triggered Service Request</td>
<td>• Master Node to eNB/gNB Change</td>
</tr>
<tr>
<td>• Paging</td>
<td>• eNB/gNB to Master Node Change</td>
</tr>
<tr>
<td>• Tracking Area Update</td>
<td></td>
</tr>
<tr>
<td>• Dedicated Bearer Setup</td>
<td></td>
</tr>
<tr>
<td>• S1 Handover</td>
<td></td>
</tr>
<tr>
<td>• X2 Handover</td>
<td></td>
</tr>
<tr>
<td>• Detach</td>
<td></td>
</tr>
</tbody>
</table>
Agenda

• Introduction to Project TINA
• Software
  • Programmable Network Packet Broker (NPB)
  • DPDK-based Probes
  • DPDK-based Network Packet Capture (FloX)
• Hardware
  • T-CAP
  • Disaggregation approach
Network Packet Broker (NPB)

Requirements
• L2~L4 based filter
• Forward/Load balance/Replicate
• For Telco specific requirements
  • GTP protocol support
  • GTP inner user packet headers
  • GTP load balancing
  • IEEE 1588 time synchronization
  • Hardware time stamping
Lesson Learned from using Fixed Function Ethernet Switching ASIC

ASIC limitation
- Parser depth (up to 128B)
- Unnecessary packet processing behavior (e.g., L2/L3/multicast)
  - Lead to increase engineering complexity
- New protocols support (GTP, ERSPAN)
- Header insertion
- Fixed table size
- Limited # of mirror sessions

SDK limitation
- Packet replication implementation
- Only available for multicast traffic (MAC addr, IP addr)
- Hashing fields selection (either outer or inner, or both)
- Load balancing algorithm
Programable Ethernet Switching ASIC and P4 Language

P4 is a formal language describing packet processing behavior

• Open
• Target independent (e.g., Barefoot Tofino, NPU, FPGA)

Easy to develop/test/bug fix/proof future features more fast than ever and without hardware upgrade

Easy to optimize memory resources (especially for TCAM) for each use case

Can create SDK based on auto-generated P4 program access APIs (e.g., P4runtime)
npb.p4-16

npb.p4-16
• Port / Port Group
• Ingress/Egress Filter
• Forward/Load balance (normal, symmetric, resilient)/Replicate
• IEEE 1588 based Time stamping (nanosecond granularity) on ERSPAN type III & INT (In-band Network Telemetry) spec
• H/W NetFlow-like generation*

→ Under field test on our production central office (Seoul region)
Software Stack

Management Daemons
CLI, NetConf

System Daemons
Port/Filter

HW Support Daemons
H/W monitoring

YANG-based configuration and operational state data store

API

bf_runtime

bf_drv.ko

Tofino

npb.p4

npb.yang

User

Kernel

Device

Open. Together.
SD-Probe

Requirements

• Capture and parse raw packets
• Extract and generate metadata
  • NetFlow
  • XDR (for Telco)
• In high speed
  • 100Gbps / site
SD-Probe

Probe can achieve 80Gbps throughput and generate 5M flows/sec with ~100% utilization of 24 cores

- Utilize DPDK to pump up raw packet to applications
- Leverage RSS (Receive Side Scaling) to load balance packets to CPU cores assigned (CPU pinning)
  - -> this is heuristic
Lesson Learned from DPDK based Probe

• Can’t achieve more than 100Gbps with deterministic performance
  • Depends on traffic pattern and CPU pinning config which is heuristic
  • Complex computation (e.g., regex, DPI)

-> Programmable hardware chip (i.e., FPGA)
Agenda

• Introduction to Project TINA
• Software
  • Programmable Network Packet Broker (NPB)
  • DPDK-based Probes
  • DPDK-based Network Packet Capture (FloX)
• Hardware
  • First generation hardware design (T-CAP)
  • Disaggregation approach
First-gen Hardware Platform (T-CAP)

SKT Converged Appliance Platform

• Server board
  • Dual Intel Xeon E5-2600 v3 CPU (Haswell)
  • Up to 512GB RAM
  • 4X 2.5` SATA SSD

• Network switch board
  • Intel RRC (Red Rock Canyon) fixed function ethernet switching ASIC
  • x4 ethernet controller w/ PCIe gen3 (up to 200Gbps) connected to server part

https://www.slideshare.net/JunhoSuh/specification-skt-cna-ssx2rc-20160821
Whitebox Switch for npb.p4-16

Barefoot Tofino (2T, 6.5T) w/ whitebox (BSP) vendor collaboration

**KEY FEATURES**

- 48x 25GbE + 8x 100GbE in 1RU chassis
- 2.0 Tbit fully P4 programmable Barefoot Tofino ASIC
- Supporting speeds from 1GbE to 100GbE
- First ever Barefoot Tofino based switch with built in time-synchronisation

- Transparent and boundary clock support (1588v2)
- Strong 8-core x86 CPU, 128GB SSD and 32GB of RAM
- Implement any protocol or feature that your network requires with P4
- Redundant, hot-swappable power and air-flow design for mission critical use-cases

**KEY FEATURES**

- 64x 40/50/100GbE in 1RU chassis
- 6.4 Tbit fully P4 programmable Barefoot Tofino ASIC
- Supporting speeds from 1GbE to 100GbE
- First ever Barefoot Tofino based switch with built in time-synchronisation

- Transparent and boundary clock support (1588v2)
- Strong 8-core x86 CPU, 128GB SSD and 32GB of RAM
- Implement any protocol or feature that your network requires with P4
- Redundant, hot-swappable power and air-flow design for mission critical use-cases
Junho Suh
(junho.suh@sk.com)