Enhancing Network Visibility Based On Open-Converged Network Appliance

Dongheon Lee, Junho Suh
SW-Defined Infra Lab, SK telecom
Telco Services
- Ultra High Data Rate
- Low Latency
- Massive Connectivity

**ATSCALE**
- Software-Defined RAN
- Network & Service Slicing
- Next-Generation OSS (TANGO)

**COSMOS**
- Software-Defined Infrastructure
- Open Hardware and Software
- Universal Platform for Diverse Applications

ATSCALE: Scalable, Cognitive, Automated, Lean, E2E
COSMOS: Composable, Open, Scalable, Mission-critical Optimized System
All-IT Network Architecture and Challenges

Performance and reliability
- Open source software performance and reliability
- Provisioning and live migration speed
- SDN controller and switch performance

Bandwidth-on-demand service
- Data center interconnection - stretching SDN coverage to the transport layer
- Data center backup and virtual machine migration

Operational Intelligence
- E2E monitoring and analytics for global orchestration and management
- Disaster prediction and recovery
- Security

Static → Dynamic
Manual → Automated
You Can't Manage What You Can't See

Network visibility is extremely important to network providers for managing All-IT Infra.

- Troubleshooting application performance issues
- Monitoring application performance and reliability
- Ensuring network scalability
- Protecting and securing the network
- Managing complex network infrastructure

TiNA for COSMOS Network Monitoring

TiNA: SKT integrated Network Analyzer

- Unified network monitoring and operating solution for software-defined data center
TiNA: SKT integrated Network Analyzer

Unified network monitoring and operating solution which includes essential systems and tools for network health monitoring, traffic monitoring, packet analysis, session analysis, and troubleshooting.

Service-Centric Network Monitor
- Tennant traffic monitoring

3D Network Visualization
- Physical/virtual network visualization
- End-to-end flow monitoring

Network Analyzer
- Packet-Level: Packet into flow aggregation
- Flow-Level: Tennant traffic classification
- Connection performance analysis
- Top-N conversations
- E2E Flow Path Calculation

Network Probe
- SNMP
- sFlow, NetFlow

Network Packet Broker
- On-demand packet brokering
- w/ filtering/replication/aggregation

Packet Capture Appliance
- Flow-aware high speed packet dump

Fast Packet Processing
- DPDK based I/O, support L2-L4, VXLAN, ERSPAN, ….
T-CAP is a converged network appliance, which integrates high performance Xeon server with data center switch.

**T-CAP Use Cases**

- **Network Applications on ToR Switch**
  - Load Balancing, Firewall, 3rd Party Applications

- **Data Backup & Network Analytics**
  - Network Traffic Backup and Monitoring

- **Overlay Gateway for Virtual Network**
  - VTEP (VxLAN Tunnel End Point) Node

**Convergence of switch and server**

- Open hardware
- High computing on a switch
- Lower cost due to convergence
- Simpler deployment and operation

**T-CAP**

Converged Network Appliance

**Server (Dual E5 Xeon)**

**Switch (10/40/100G)**

**Storage**

OCP Contribution as a open hardware

Co-operation with high performance N/W Controller
Developed by using open source and open hardware

TiNA & T-CAP Development

SK Telecom CNA-SSX2RC
(Service-Optimized Hybrid Network Appliance)
TiNA consists of 6 systems in which each system can be deployed as a standalone appliance or systems can be bundled together according to the use-case.

- **Network Packet Broker**
  - Brokering network traffic from multiple mirror or taps according to the demand of multiple monitoring devices

- **Network Probe**
  - Supporting network and traffic monitoring protocol such as SNMP and sFlow

- **Network Analyzer**
  - Packet-level: Aggregates packets into flows and measures various TCP statistics
  - Flow-level: Tenant traffic classification, Top-N Conversations, E2E flow path calculation

- **Packet Capture & Analysis Appliance**
  - High speed packet capture appliance supporting up to 40Gbps traffic

- **3D-Based Network Monitor**
  - Supporting intuitive network monitoring and operation by efficiently visualizing complex network topology and traffic flow

- **Service Centric Network Monitor**
  - Classifying tenant traffic and collecting various network traffic statistics of cloud data center
Flow-level network analyzer
- support analyzing millions of information of nodes, links, and traffic flows by using spark streaming engine

Packet-level network analyzer
- TCP connection analyzer support up to 1 million TCP flows
- Packet into Flow Aggregation: support up to 200K new flow arrivals per second

Network Packet Broker
- Virtual and physical NPB, vNPB supports 40 Gbps packet filtering & forwarding performance

Packet Capture Appliance
- 40 Gbps packet capture system with 24 HDDs

Low Cost
- implemented on x86 commodity hardware by leveraging DPDK fast packet processing framework
Network Monitoring with TiNA

- Service-Centric Network Monitor
- 3D Network Visualization
- Packet Capture Appliance
- Packet-Level Network Analyzer
- Network Probe
- Network Packet Broker
Packet-Flow Analysis with T-CAP

- **Packet-Level Network Analyzer**
- **Network Probe**
- **vNPB**
  - Packet into Flow Aggregation
  - TCP Connection Analysis
  - SNMP/sFlow Collector
  - Virtual network packet broker

- **T-CAP**
- **Commodity Server**
  - x86

- **Server** (Dual E5 Xeon)
- **Switch** (10/40/100G)
- **Storage**

- **Connection Analyzer**
  - Packet into Flow Aggregation
  - TCP Connection Analysis
  - SNMP/sFlow Collector
  - OpenFlow based packet broker

- **NPB Controller**
- **NPB Fabric**
TCP Connection Performance Analysis

- TCP Connection Setup Time, Connection per Sec, SYN/FIN Count, Retransmission Rate, Zero Window, App Response Time
- Determine whether the network problem is the application or the network
NPB, Network Probe, Network Analyzer Usage
High Speed Packet Capture and Retrieval System

- Essential tool for troubleshooting and network forensics
- Pipelined parallel packet processing based on DPDK
- Support both packet and flow-based capture

Specifications

- 20 cores (Intel Xeon CPU E5-2650 v3 @ 2.30 GHz, 2ea)
- 128 GB RAM (DDR4 16 GB DIMM 2,133 MHz, 8ea)
- 192 TB HDD (SATA 6 Gb/s 3.5" 7.2K RPM 8 TB ENT HDD, 24ea)
- Support 40Gbps line rate packet capture to disk
- With 192 TB storage
  - 40 Gbps * 11 hour
  - 10 Gbps * 43 hour
  - 1 Gbps * 427 hour
SONA: Simplified Overlay Network Architecture presented in ONS ‘16

- Consolidated physical/virtual network based on real-time flow stats
- L3-based leaf-spine networking to maximize link utilization

Use-Case: Traffic Engineering with TiNA and SONA
Use-Case: Traffic Engineering with TiNA and SONA
Cloud Data Center Multi-Tenancy

- Cloud needs to be managed and monitored based on application
- Per-tenant traffic monitoring for SLA management and pricing
- Classifying tenant traffic based on IP flow, VLAN, VXLAN

Service-Centric Network Monitor

TiNA Multi-Tenant Traffic Monitoring System

- Mirror or sFlow
  - Aggregating packets into flows
- JSON format conversion
  - Classifying flows into tenants
- Register tenants and hosts
  - Monitor tenant traffic
- Service traffic statistics monitoring per-tenant
- Per service/host/flow traffic statistics, Top-N conversations
Efficiently visualize multi-layer network topology providing intuitive and interactive user interface

3D UI/UX
- Display physical/logical/virtual network topology efficiently
- Drill down from a global datacenter to a traffic flow in one second
- Support Android, Windows Client

Visualizing real-time network analysis results
- Collect and analyze millions of information of nodes, links, and traffic flows by using spark streaming engine
- Diagnose network events with game animation within a few seconds
- End-to-end traffic flow path visualization
TiNA 3D-Based Network Monitor

Physical Network Topology

Logical Network Topology

E2E Flow Path Visualization

Virtual Network Topology
Network Operation

Traffic Engineering with TiNA, SONA, 3DV
- Subscriber and tunnel end-point based packet filtering
- VoLTE Traffic Analysis: throughputs, packet loss, jitter, latency
- H/W NPB Time-stamping by T-CAP S/W silicon
- GTP Correlation
T-CAP is a converged network appliance, which integrates high performance Xeon server with data center switch.

**T-CAP**

**Convergence of switch and server**
- Open hardware
- High computing on a switch
- Lower cost due to convergence
- Simpler deployment and operation

**T-CAP Use Cases**

- **Network Applications on ToR Switch**
  - Load Balancing, Firewall, 3rd Party Applications

- **Data Backup & Network Analytics**
  - Network Traffic Backup and Monitoring

- **Overlay Gateway for Virtual Network**
  - VTEP(VxLAN Tunnel End Point) Node

**OCP Contribution**
as a open hardware

**Co-operation with**
high performance N/W Controller
T-CAP Hardware Specification

36-port Ethernet with 4 PCIe Gen3 Ethernet Controller
- Up to 200Gbps B/W toward hosts

Flexible Frame Processing Pipeline
- L2/L3/L4/OpenFlow Forwarding
- ACLs & NAT, Stateless Load Balancing

Tunneling Engine
- 64k Exact Match Table
- NVGRE/Geneve/VXLAN-GRE/NSH

Unparalleled HW Latency
- 300 ns network latency
- 1000 ns host-network latency

http://files.opencompute.org/oc/public.php?service=files&t=6a542c4983c8180f5c6f6e7da198e60a
CPU: Intel Rangeley/Avoton
Memory: up to 4 GB DDR3/4
Storage
- 8 GB storage embedded
- SATA-DOM

Peripherals
- 2x USB2.0
- 1x External Console (RJ45 Type)
- 1x GbE for Management

http://files.opencompute.org/oc/public.php?service=files&t=6a542c4983c8180f5c6f6e7da198e60a
T-CAP Hardware Specification

- Dual Intel Xeon E5-2600 v3 CPU (Haswell)
- Up to 512GB memory
- 4X 2.5" SATA SSD
- 1+1 Redundant PSU (Power Supply Unit)
- Front Loading Switch Port Module
  - 10G 12 Port
  - 10G 4 Port + 40G 2 Port

http://files.opencompute.org/oc/public.php?service=files&t=6a542c4983c8180f5c6f6e7da198e60a
T-CAP Software Stack

Legend

- SKT projects
- FOSS
- SKT maintained
- Intel proprietary

Linux Kernel

- Device
  - fm10k

Net Device

- SwitchD
  - IES SDK
- fm10k

Networking Stack

- OVS
  - Linux Bridge
- Quagga
- 3rd Party
- Platform Manager
- SONA Fabric SKT’s NW Fabric Auto. Tool
- T-O&M SKT’s Op. & Mgmt. Tool

Unified Network Config. Data Model & API (e.g., OpenConfig)

- i2c
- SFP/QSFP+
- LED
- PWD
- Sensors

Legend

- FOSS
- SKT maintained
- Intel proprietary

SKT's Op. & Mgmt. Tool

SKT's NW Fabric Auto. Tool

Platform Manager

Unified Network Config. Data Model & API (e.g., OpenConfig)
Real deployment requirements for a single cluster

- Performance for packet filtering and actions
  - ✓ ~ # of policy rules
  - ✓ ~ # of vantage points
- Low Latency ~ 400ns

<table>
<thead>
<tr>
<th>Feature Required</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet Filtering</td>
<td>L2 Filter(src/dst mac, ip, vlan id, eth type …), L3 Filter(src/dst ip, ip protocol ToS …), L4 Filter (src/dst Port, VNI)</td>
</tr>
<tr>
<td>Action</td>
<td>Forwarding, Aggregation, Replication, Load Balancing, Drop</td>
</tr>
<tr>
<td>Source Port Labeling</td>
<td>Tagging source port number in packet headers</td>
</tr>
<tr>
<td>Header Stripping</td>
<td>Strip unnecessary packet headers to analyze traffic such as VLAN Tagging/VXLAN,MPLS</td>
</tr>
<tr>
<td>Packet Slicing</td>
<td>Trim packets</td>
</tr>
<tr>
<td>HW Timestamp</td>
<td>Time synchronization between analysis server farms and NPB H/Ws through PTP</td>
</tr>
<tr>
<td>Tunneling</td>
<td>Transfer to remote branch</td>
</tr>
</tbody>
</table>

T-CAP Software Stack for Hardware Network Packet Broker
Build and Package Management From Kernel to Applications

Yocto Project (DIY/Roll-Your-Own Linux Distro.)

Openembedded Architecture Workflow
- Upstream Source
- Metadata/Inputs
- Build system
- Process steps (tasks)
- Output Image Data

Package Feeds
- .rpm Generation
- .deb Generation
- Jpk Generation
- Image Generation
- SDK Generation
- Images
- Application Development SDK

Source Mirror(s)
- Upstream Project Releases
- Local Projects
- SCMs (optional)

User Configuration
- Metadata (bb + patches)
- Machine (BSP) Configuration
- Policy Configuration

ONIE image
CPP board (Intel x86)
Mininet extension to support feature validation test for physical switches

- **Scalability** test for control-plane protocols support on multiple physical hosts or cloud
- **Reusable test scenario** regardless of vendors
- **Interop** test between multiple NOS vendors by plugins implemented
Work Flow for NetDevOps

1. branch & sync
2. clone
3. push
4. merge
5. build
6. test pass or fail
6. code quality check
7. notification if failed
8. release if pass
9. update noti. & package update
SKT New Infra Architecture for 5G and Beyond

COSMOS – All-IT Undelay Architecture

- Local NFV Orchestrator
- Transport Infra Orchestrator

Virtualized Network Functions
- RNF
- CNF
- TNF
- ESF
- CSF
- OSF

Resource Abstraction Layer
- Edge DC: Open & Programmable H/W
- Central DC: Open & Programmable H/W

Service Orchestration and Exposure
- Low Latency Service
- Immersive Media
- Telco Service

End-to-end Network Orchestration
- Unified DC

Transport Infrastructure
- Transport Open H/W
- POTN

Mobile Connectivity Functions
- 4G
- 5G

Network Service Functions
- SDRAN
- vCore
- uCTN

Fronthaul
- L1/L2
- Remote Unit

WAN

(E2E Resource Mgmt., Cognitive & Intelligent Automation)
End of Document

dongheon@sk.com
junho.suh@sk.com