Finding the Right Way for High-Performance NFV

Ryota Kawashima
Nagoya Institute of Technology, Japan
Agenda

1. Backgrounds
2. D-Plane Technologies
3. Performance
4. Research Directions
Virtualizing Everything

Client Computers

1st-gen virtualization (Before NFV)

Edge Functions

2nd-gen virtualization (Early-stage NFV)

Core Functions

3rd-gen virtualization (Future NFV)
Core Network Traffic

Marcus K. Weldon, The Future X Network

Hardware

400G Ethernet
(IEEE 802.3bs)

Software

Next-gen. VNF?
Agenda

1. Backgrounds
2. D-Plane Technologies
3. Performance
4. Research Directions
D-Plane Technologies in NFV-nodes

**OpenDataPlane**

**DPDK**

**Netmap**

**XDP**

**SPP**

**NetVM**

**openNetVM**

**vhost-user**

**SR-IOV**

**BESS**

**Open vSwitch**

**Iagopus**

**Chelsio**

**Mellanox Technologies**

**QLogic**

**Netronome**

**FastClick**

**FastPath**

**Packet I/O Engine**

**VNF (VM)**

**NW Stack**

**Virtual Network I/O**

**Virtual Switch**

**Packet I/O Engine** (Pass-through)

**x86 Server**

**Physical NIC**

**BESS**
**x86 Server Architecture (Intel Xeon Scalable Processor)**

**Intel® Xeon® Scalable Processor**
- **1.7-3.6 GHz**
- 170 Gbps (DDR4-2666)

**Intel® Xeon® Scalable Processor**
- **1.7-3.6 GHz**
- 170 Gbps (DDR4-2666)

**Intel® C620 Series Chipset**
- DMIx4

**1x 100G Mini-Path Fabric ✽**
- 128 Gbps (uni, x16)

**1x 100G Mini-Path Fabric ✽**
- 128 Gbps (uni, x16)

Fast Packet I/O Engine

DPDK
DATA PLANE DEVELOPMENT KIT

- Pure Polling
- Pre-allocated Memory
- User-space DMA
- Packet Batching

100% CPU usage!
Bridging VNFs and the Host

- VNF (Container)
- VNF (VM)
- VNF (VM)

vhost-user

vSwitch

SR-IOV

Packet copy!

Inflexible!
User-space Network Stack

Vector Packet Processing (VPP)

- Full network stack
- Click-like modular design
- Vectorized processing

Packet Vector

(N: 4 - 256)
Agenda

1. Backgrounds
2. D-Plane Technologies
3. Performance
4. Research Directions
Evaluation

Environment

<table>
<thead>
<tr>
<th>VM</th>
<th>Machine 1</th>
<th>Machine 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 cores</td>
<td>Core i7 6700K 4.0 GHz (4 cores w HT)</td>
<td>Xeon E5-2630 v4 2.2 GHz (2x10 cores)</td>
</tr>
<tr>
<td>4 GB</td>
<td>32 GB (DDR4-2133)</td>
<td>128 GB (DDR4-2133)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NIC</th>
<th>Traffic Generator (MoonGen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intel XL710</td>
<td>40 GbE</td>
</tr>
<tr>
<td>PCIE 3.0 (x8)</td>
<td></td>
</tr>
</tbody>
</table>

LFMT (Lock-free Multi-threading)

Run-to-Completion

Pipeline

Lock free!
Scale out!
Phy-to-Phy Throughput (Baremetal)

Machine 1 (4.0 GHz)

Machine 2 (2.2 GHz)

Poor throughput for short packet traffic

(Ref.) Intel XL710 vs. Mellanox ConnectX-5

Intel XL710 (40 GbE)

Mellanox ConnectX-5 (100 GbE)

CPU is bottleneck

Single core can support a mere 30 Mpps
Phy-to-Phy Latency/Jitter (Baremetal)

Machine 1 (4.0 GHz)

Machine 2 (2.2 GHz)

10x latency/jitter to hardware switches

Evaluation Summary

- **High-clock speed** CPUs
- Far from **40G** throughput
- **A few μs** latency per node
Research Direction 1

Scale-out Approach

- Aggregate Throughput
- Per-datapath Throughput
- Latency/Jitter
- Power Consumption

Boost Single DP Performance!
Approach 1: Hardware Offloading

(Smart NIC)

PHY-to-VM Throughput (40G)

https://www.netronome.com/
Approach 2: Packet Aggregation

PA-Flow†

VNF (VM) → Tx Queue (vhost-user) → Packet Vector

Aggregated Packet

Traffic amount increases ...

Aggregation rate increases!

Datapath Throughput
70% up!

Datapath Latency
0.2µs decreases!

Research Direction 2

What about Service Chaining?

Reduce No. Enqueue Times!
Approach: Local Service Chaining
Fast Inter-VNF Communication

Soft Patch Panel (SPP)


Dynamic VNF Fusion (ultimate)

Advantages
- Single DP Performance
- Modularity

Challenges
- Process Fusion Tech.

Dynamic Fusion
- Fusing processes
- Merging VPP graphs

Single Datapath (no sync queue)
Summary

Boost Single DP Performance!

(Hardware Offloading)
Packet Aggregation

Reduce No. Enqueue Times!

Local Service Chaining
Dynamic VNF Fusion