Change before you have to be claimed

Experience in DPDK-enabled SDN vSwitch and DPDK-enabled VNF with Vhost

Tomoya Hibi, Yoshihiro Nakajima, Hirokazu Takahashi
NTT Network Innovation Labs
What we did

- First experiment with DPDK vSwitch and DPDK VNF with vHost PMD
  - DPDK-enabled vSwitch (DPDK vHost PMD)
  - DPDK-enabled VNF (DPDK virtio-PMD)
  - DPDK 16.04 + patch

- Examine how performance impacts we face only resource assignment
  - CPU and memory assignment
  - VNF and vSwitch assignment

- NFV middleware for scale-out VNFs
  - Thanks to Interop Tokyo 2016 ShowNet!

https://www.facebook.com/interop.shownet
NFV middleware for scale-out VNFs

- Flexible load balance for VNFs with smart hash calculation and flow direction
  - Hash calc: NetFPGA-SUME
    - Hash calculation using IP address pairs
    - Hash value are injected to MAC src for flow direction for VNF
  - Classification and flow direction: Lagopus
    - Flow direction with MAC src lookup

<table>
<thead>
<tr>
<th>hash</th>
<th>dl_src</th>
</tr>
</thead>
<tbody>
<tr>
<td>type1</td>
<td>52:54:00:00:00:01</td>
</tr>
<tr>
<td>type2</td>
<td>52:54:00:00:00:02</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Type 256</td>
<td>52:54:00:00:00:FF</td>
</tr>
</tbody>
</table>
Challenges in vSwitch

- **vNIC between DPDK-enabled vSwitch called Lagopus and DPDK-enabled VNF**
  - vrouter called Virnos provided by IP infusion

- **Many vNICs and flow director (load-balancing)**
  - 8 VNFs and total 18 vNICs
Best resource assignment for vSwitch and VNFs for performance?

- Packet processing workload aware assignment is required for Lagopus and VNF

- Best configuration for resource assignment?
  - Dual Xeon (E5-E2667 v3, Haswell-EP)
  - 8 x 8GB DDR4-2133 memory
  - 1x Dual port Niantic NIC
Resource assign impacts in packet processing performance

4.4Gbps

10Gbps
CPU resource assignment for I/O (1/2)

- DPDK-based system needs CPUs for I/O because polling-based network I/O in DPDK
- Physical I/O is relative intensive compared to vNICs
CPU resource assignment for I/O (2/2)

- Traffic-path-aware CPU assign
- 4 CPU core were assigned to I/O thread of Lagopus
Other optimization in flow-rule reduction

- **512 match rules are required by default**
  - 256 MAC src match
  - Both direction (up link/down link)

- **Only 16 rules cover the above requirements using mask-aware match rule technique**
  - Hash value are injected lower 1byte of MAC address
  - Traffic are distributed by statistical multiplexing effect of the nature of traffic
  - Only 3bit-lookup cover the above requirements

<table>
<thead>
<tr>
<th>in_port</th>
<th>dl_src</th>
<th>action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52:54:00:00:00:01</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>52:54:00:00:00:02</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>52:54:00:00:00:FF</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>in_port</th>
<th>dl_src</th>
<th>action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em><strong>:</strong></em>:<em><strong>:</strong></em>:*0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td><em><strong>:</strong></em>:<em><strong>:</strong></em>:*1</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td><em><strong>:</strong></em>:<em><strong>:</strong></em>:*7</td>
<td>7</td>
</tr>
</tbody>
</table>
Performance evaluation

**Long packet journey**
- Packet-in -> Physical NIC -> Lagopus -> vNIC -> VNF -> vNIC -> Lagopus -> Physical NIC -> Packet-out
- Two major packet copy (vNIC-related copy)

![Graph showing performance evaluation](image-url)
Conclusion

- **Needs more hardware details for performance**
  - CPU, Memory, PCI-exp topology
  - Memory allocation, CPU core assign
  - Ie Enhanced Platform Awareness (EPA)

- **Performance profiling is essential**
  - Needs VNF/vSwitch modeling and benchmark test suite
  - Difficult to know performance degradation point, performance bottleneck
    - Still primitive tools are provided (perf, htop…)

Reference

■ Web
  ● https://lagopus.github.io

■ Github
  ● Lagopus vswitch
    • https://github.com/lagopus/lagopus
  ● Lagopus Book
    • https://github.com/lagopus/lagopus-book

■ Visit IDF16 booth #825 (August 16-18)
  ● Cloud WAN solution using Lagopus vSwitch