ODL SFC with OVS-DPDK, HW accelerated dataplane and VPP

Prasad Gorja, Senior Principal Engineer, NXP
Harish Kumar Ambati, Lead Engineer, NXP
Srikanth Lingala, Lead Engineer, NXP
Agenda

• SFC Introduction
• ODL SFC with OVS-DPDK
• ODL SFC with HW accelerated dataplane and VPP as vNF
OpenDayLight SFC

**SFC**
- SF - Service Function
- SFF - Service Function Forwarder
- SFP - Service Function Path
- RSP – Rendered Service Path

**NSH**
- NSH – Network Service Header
- NSP - Service Function Forwarder
- NSI - Network Service Index
ODL SFC with OVS-DPDK

Legend:
- Green: OpenFlow 1.3/OVSDB
- Red: Service Function Chain Path

ODL SFC with OVS-DPDK and NXP ARM Compute Node

Control Node
- Tacker (OpenStack vNF Manager)
- ODL SFC
- OpenStack
- OVS
- ToR

NXP ARM Compute Node
- SFC
- FW VM
- DPI VM
- Classifier1
- Classifier2
- SFF
- OVS-DPDK (br-int)
- DHCP
- 10.10.10.3
- 10.10.10.4
- 10.10.10.5
- 172.16.0.132
- 112.16.0.4
- 112.16.0.5

Internet

Client

OpenStack

FW VM

DPI VM

Classifier1

SFF

Classifier2

OVSDPDK (br-int)
Tacker and NSH Flows

Vnf-create

1. Tacker and NSH Flows

2. SFC and SFC-Classifier

- cookie=0x1110010000440255, duration=632.734s, **table=11**, n_packets=7, n_bytes=518, tcp
  actions=move:NXM_NX_TUN_ID[0..31]–>NXM_NX_NSH_C2[], push_nsh, load:0x1–>NXM_NX_NSH_MDTYPE[], load:0xc0a80018–>NXM_NX_NSH_C1[], load:0x2c–>NXM_NX_NSP[0..23], load:0xff->NXM_NX_NSI[], load:0x7b7b7b03–>NXM_NX_TUN_IPV4_DST[], load:0x2c–>NXM_NX_TUN_ID[0..31], resubmit(,0)

- cookie=0x1110010000440255, duration=632.734s, **table=11**, n_packets=7, n_bytes=518, tcp, reg0=0x1, tp_src=2000, tp_dst=80
  actions=move:NXM_NX_TUN_ID[0..31]–>NXM_NX_NSH_C2[], push_nsh, load:0x1–>NXM_NX_NSH_MDTYPE[], load:0xc0a80018–>NXM_NX_NSH_C1[], load:0x2c–>NXM_NX_NSP[0..23], load:0xff->NXM_NX_NSI[], load:0x7b7b7b03–>NXM_NX_TUN_IPV4_DST[], load:0x2c–>NXM_NX_TUN_ID[0..31], resubmit(,0)

3. OpenStack Dashboard
## Need for Programmable HW Acceleration

<table>
<thead>
<tr>
<th>General processor are sub-optimal for packet processing</th>
<th>HW Acceleration (Advanced IO Processor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low packet processing locality underutilizes cache and pipeline and suffers from <strong>high DDR latency</strong></td>
<td>Specialized memory hierarchy and explicit DMA operations instead of cache allow <strong>deterministic performance</strong>.</td>
</tr>
<tr>
<td>Increasing single thread performance causes super-linear power increase</td>
<td><strong>Parallelism</strong> with more “small” cores. <strong>Higher Performance/Watt.</strong></td>
</tr>
<tr>
<td>Performance and complexity of software hiding latency of asynchronous access to accelerators</td>
<td><strong>Hardware scheduler</strong> based multitasking environment hides access latency.</td>
</tr>
<tr>
<td>More core cycles are consumed for standard packet processing operations</td>
<td><strong>Hardware accelerators</strong> for common tasks: Lookups, parse, frame operations, timers, statistics, frag and reassembly</td>
</tr>
<tr>
<td>Performance and complexity of software ordering and synchronization</td>
<td>Functions are provided by hardware scheduler</td>
</tr>
<tr>
<td>More cores are engaged when OVS runs in GPP</td>
<td><strong>AIOP Offload offers conservation of cores</strong></td>
</tr>
</tbody>
</table>
HW Accelerators

Classifier & Table Lookup Unit (CTLU)
- Parser & Keygen
- Parse packet based on parser profile
- Read parser results during processing
- The TLU performs table lookup off load acceleration in the AIOP
- Supports - Exact match (EM), Longest Prefix Match (LPM) and Algorithmic ACL (ACL) table types.
- Rule timestamping for activity.

Timer Manager (TMAN) engine
- Offloads the management of millions of timers.
- Manages the timers, and automatically initiates AIOP tasks when timers expire.
- Micro/Milli/Seconds and Periodic or One shot timers
- Timer priority and Timer containers

Frame DMA (FDMA)
- Autonomous initial frame presentation & subsequent presentation to AIOP WS.
- Frame modification, Frame replication, Frame concatenation, split acceleration, Frame store, discard acceleration. Enque frames to the Queue Manager (FQ/QD)
- Direct access to system memory (DMA data).

Context DMA (CDMA)
- Read context buffer contents from external memory (DDR/PEB) and present in workspace memory.
- Write modified contents of context buffer (in WS-MEM) to external memory (PEB/DDR)
- Synchronization mutex locks and RCU
- Context buffer reference counting
Hardware OVS Acceleration

- Conserve cores and increased VM density
- Integrated at well defined ovs-dpif and netdev provider
- Deterministic Performance
ODL SFC with HW accelerated dataplane

Legend:
- OpenFlow 1.3/OVSDB
- Service Function Chain Path

Control Node
- Tacker (OpenStack vNF Manager)
- ODL SFC
- OpenStack
- OVS
- ToR

NXP ARM Compute Node
- SFC
- FW VM
- VPP
- DHCP
- Client
- AIOP HW Accel
- Classifier2
- SFF
- 10.10.10.3
- 10.10.10.4
- 10.10.10.5
- 112.16.0.4
- 172.16.0.132

Internet

OpenStack

Tacker (OpenStack vNF Manager)
NOVA VIF driver

- Create Data Path and connect to SDN Controller in OpenStack Compute Nodes
- Flows pushed from OPENDAYLIGHT SDN Controller.
- While Nova Compute is started all available ports are published to nova in JSON Format.
- Dynamic assignment of ports from JSON format to VMs.
- Create DPRC Containers for VM and DPNI in each in each VM DPRC.
- Connect DPNI in VM DPRC to AIOP DPRC.
- Bind VM DPRC to VFIO and Building libvirt XML with specific DPNI and DPRC IDs.
THANK YOU

prasad.gorja@nxp.com
harish.ambati@nxp.com
srikanth.lingala@nxp.com