OvS-DPDK Full VXLAN Offload with SR-IOV

Hemal V. Shah, Distinguished Engineer and Architect, Broadcom Inc.
Sriharsha Basavapatna, Principal Software Engineer, Broadcom Inc.
Agenda

- OvS Offload Layer
- OvS-DPDK Tunnel Offload Using Helper APIs
- VXLAN Decap Offload with SR-IOV
- Summary
OvS Offload Layer

- Implements control path for flow offloads
- Device agnostic
- Hidden from ofproto layer

- Split in generic and provider sub-layers
- Enables multiple NIC flow offload APIs
- Flow APIs registered by a specific provider
- Two subsets of flow APIs:
  - Flow Put, Delete, and Stat APIs for a specific flow
  - Flow dump Create, Destroy, etc. APIs for dumps
Tunnel Encap/Decap Offload Using Helper APIs

- **Tunnel encap offload involves a single flow with encap action**
  - Flow (Match: in_port, smac, dmac; Action: Tunnel push and Output to phy port)

- **Tunnel decap involves two flows and recirculation in OvS**
  - Flow-F1 (Match: t_dmac, t_dip, t_proto, t_port; Action: Tunnel pop and Output to tunnel port)
  - Flow-F2 (Match: t_dip, t_sip, t_id, inner eth, Action: Output to VF-Rep)

- **Decap flow offload sequences can be different (F2→F1, F1→F2, F2 only)**
  - PMDs can not assume a specific sequence
  - PMDs need to internally handle all possible sequences

- **RTE_FLOW tunnel helper APIs simplify decap offload**
  - rte_flow_tunnel_decap_set(): validates tun type before F1 offload, PMD may return priv actions
  - rte_flow_get_restore_info(): used on F1 hit/F2 miss, PMD returns packet specific details
  - rte_flow_tunnel_match(): issued right before F2 offload, PMD may return private items
  - rte_flow_tunnel_action_decap_release(): called by OvS after success or failure of F1 offload

- **Decap offload design avoids double counting (HW and OvS) of F1 on F2 miss**
- **Tunnel physical port metadata enables F2 offload on port where flow originated**
VXLAN Decap Offload with SR-IOV

OvS-DPDK

OF-Proto

User-Datapath

Offload-DPDK

rte_flow

rte_eth0

rteEth1 (vf-rep)

PMD

PF

VF

NIC

VM

Solid arrows: Pkt sequence
Dotted arrows: Flow sequence
Summary

- Two bridge model posed significant challenges for tunnel decap offload
- OvS-DPDK encap/decap offload with tunnel helper APIs addressed challenges
- Full VXLAN Encap/Decap offload with SR-IOV is in place with OvS 2.16
- VXLAN Encap/Decap offload design can be leveraged to other tunnel types