Open vSwitch Offload

Simon Horman
DXDD Europe
Utrecht, 8th June 2017
Overview

- What is Open vSwitch?
- Offload Models
- TC-Flower Based Offload
What is Open vSwitch
What is Open vSwitch (OvS)

- Fully featured virtual switch
- Provides match-action scheme
- Configuration via extended OpenFlow and OVSDB
- Accelerated forwarding of cached flows
- Mega Flows to enhance caching behaviour
- Tunnel termination and origination
- Platform dependent extensions: QoS, Rate-Limiting
Kernel Datapath

User-Space

ovs-vswitchd

Kernel

flows, stats, misses

OVS Datapath

packet

packet
Offload Models
OVS Datapath Hooks

- User-Space: ovs-vswitchd
  - flows, stats, misses

- Kernel: OVS Datapath
  - flows, stats, misses

- SmartNIC: SmartNIC Datapath
  - packet
  - packet
OvS-TC

User-Space

Kernel

SmartNIC

flows, stats

flows, stats

flows, stats

packet

packet

flows, stats, misses

flows, stats, misses

ovs-vswitchd

TC Datapath

OVS Datapath

SmartNIC Datapath

User-Space

Kernel

SmartNIC
TC Flower Based Offload
TC Flower Based Offload

- What is TC Flower?
- Overview of OvS-TC
- Example Modification
- How to Participate
What is TC Flower

- Packet classifier for Linux kernel traffic classification (TC) subsystem
- Allows match on key with a wide number of packet and metadata fields
- TC actions may be used to provide match-action behaviour similar to OvS
Tables and Flows

- OvS Datapath
  - Single table
  - Match on in_port
  - Flows have a wide key and are disjoint
  - And therefore can be partitioned into slices
  - Megaflows are priority independent

- TC Flower
  - Recently gained multi-table (chain) support
  - Attached to in_port
  - Flows have a wide key
  - Only one mask per priority
OvS with Offloaded TC Flower

- ovs-vswitch
- TC Flower
- Driver
- SmartNIC
- User-Space
- Kernel
- Hardware
TC Flow Placement

- May flag TC filters as software-only, hardware-only
- Default is software and if available hardware
Near Term Feature Goals for NFP

- **Stage 1**
  - L2, L3 & L4 matches
  - Drop, output & VLAN actions

- **Stage 2**
  - VXLAN

- **Stage 3 - Requires further user-space enhancements**
  - Set Action
Status

- **OvS-TC**
  - Patches available
  - Hope to see these included in OvS v2.8

- **Kernel**
  - TC Flower and related actions are present
  - More features are being added, more to be added…

- **NFP**
  - Prototyping
  - Plan to post for inclusion in Linux kernel soon
Example TC Flower Modification

- Two Kernel Components Work Together
  - Network stack core: flow dissector
  - TC: flower classifier
- Userspace Utility for Exercising Code
  - Iproute2: tc
How to Participate

- Feedback on feature set
  - OvS has many features
  - OvS-TC is starting with few
  - Plans for adding new features to OvS-TC is not fixed
- Discussion and code review on mailing lists
  - Kernel: netdev@vger.kernel.org
  - Open vSwitch: dev@openvswitch.org
Thank You