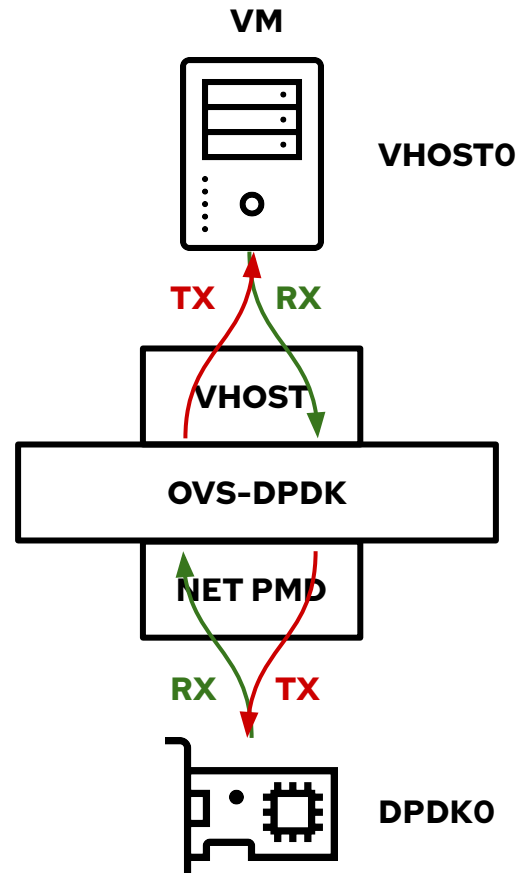


OVS-DPDK life of a packet

OVS/OVN Conference 2019

Kevin Traynor
Senior Software Engineer

Eelco Chaudron
Principal Software Engineer



On the receive path packets can be dropped for multiple reasons

- ▶ `ovs_rx_qos_drops`, packets dropped by ingress policer
- ▶ `rx_dropped` (`ovs_rx_qos_drops` + `rte_stats.rx_nombuf`)
- ▶ `rx_missed_errors` for PMDs, `rte_stats.imissed`

Total of RX packets dropped by the HW, because there are no available buffers

```
$ ovs-vsctl get Interface dpdk0 statistics
{mac_local_errors=0, mac_remote_errors=0, ovs_rx_qos_drops=0, ovs_tx_failure_drops=0,
ovs_tx_mtu_exceeded_drops=0, ovs_tx_qos_drops=0, rx_1024_to_1522_packets=0,
rx_128_to_255_packets=0, rx_1523_to_max_packets=0, rx_1_to_64_packets=1180076160,
rx_256_to_511_packets=0, rx_512_to_1023_packets=0, rx_65_to_127_packets=0,
rx_broadcast_packets=0, rx_bytes=34760355000, rx_crc_errors=0, rx_dropped=600736910,
rx_errors=0, rx_fragmented_errors=0, rx_illegal_byte_errors=0, rx_jabber_errors=0,
rx_length_errors=0, rx_mac_short_dropped=0, rx_mbuf_allocation_errors=0,
rx_missed_errors=1118323666, rx_oversize_errors=0, rx_packets=61752494,
rx_undersized_errors=0, tx_1024_to_1522_packets=0, tx_128_to_255_packets=0,
tx_1523_to_max_packets=0, tx_1_to_64_packets=61752480, tx_256_to_511_packets=0,
tx_512_to_1023_packets=0, tx_65_to_127_packets=0, tx_broadcast_packets=0,
tx_bytes=3705148800, tx_dropped=0, tx_errors=0, tx_link_down_dropped=0,
tx_multicast_packets=0, tx_packets=61752480}
```

On the transmit path packets can be dropped for multiple reasons

- ▶ `ovs_tx_failure_drops`, packet that could not be sent by the PMD driver
- ▶ `ovs_tx_mtu_exceeded_drops`, packets too big to be sent to the PMD driver
- ▶ `ovs_tx_qos_drops`, packets dropped by egress policer
- ▶ `tx_dropped`, total of all type of transmit drops

Other transmit counters

- ▶ `ovs_tx_retries`, when virtqueue is full it will retry x times (default 8) to enqueue the packet(s)

```
$ ovs-vsctl get Interface vhost0 statistics
{ovs_rx_qos_drops=0, ovs_tx_failure_drops=282123, ovs_tx_mtu_exceeded_drops=0,
ovs_tx_qos_drops=0, ovs_tx_retries=125483, rx_1024_to_1522_packets=0,
rx_128_to_255_packets=0, rx_1523_to_max_packets=0, rx_1_to_64_packets=61752480,
rx_256_to_511_packets=0, rx_512_to_1023_packets=0, rx_65_to_127_packets=0,
rx_bytes=3705148800, rx_dropped=0, rx_errors=0, rx_packets=61752480,
tx_bytes=3705148800, tx_dropped=282123, tx_packets=61752480}
```

- ▶ When VM is not in polling mode, kernel needs to be notified packet(s) have arrived
 - This is done by the Linux `eventfd_write()` system call
 - This call could stall the PMD thread, and it would be nice to know
- ▶ Multiple PMD threads might need to write to the same VHOST queue
 - When this happens a spinlock is taken

```
$ ovs-appctl coverage/show | grep -E "vhost|Event"
Event coverage, avg rate over last: 5 seconds, last minute, last hour, hash=2fe71230:
vhost_notification      522.2/sec      43.917/sec      0.7319/sec      total: 8691
vhost_tx_contention     38203.0/sec   4019.717/sec    66.9953/sec     total: 241183
```

- ▶ Currently the [Improved Packet Drop Statistics in OVS](#) patch from Anju/Rohith/Keshav is under review, which will add a lot of more coverage counters for internal OVS drop reasons:

*datapath_drop_invalid_port, datapath_drop_invalid_tnl_port, datapath_drop_lock_error,
datapath_drop_meter, datapath_drop_nsh_decap_error, datapath_drop_recirc_error,
datapath_drop_rx_invalid_packet, datapath_drop_sample_error, datapath_drop_tunnel_pop_error,
datapath_drop_tunnel_push_error, datapath_drop_upcall_error, datapath_drop_userspace_action_error,
drop_action_bridge_not_found, drop_action_congestion, drop_action_forwarding_disabled,
drop_action_invalid_tunnel_metadata, drop_action_no_recirculation_context, drop_action_of_pipeline,
drop_action_recirculation_conflict, drop_action_recursion_too_deep, drop_action_stack_too_deep,
drop_action_too_many_mpls_labels, drop_action_too_many_resubmit,
drop_action_unsupported_packet_type*

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

 [linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)

 [youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)

 [facebook.com/redhatinc](https://www.facebook.com/redhatinc)

 twitter.com/RedHat