



OpenStack NFV: Performance with OvS-DPDK for NFV and Connection Tracking

Bhanuprakash Bodireddy(bhanuprakash.bodireddy@intel.com),

Sugesh Chandran(sugesh.chandran@intel.com),

Antonio Fischetti(antonio.fischetti@intel.com)

OpenStack Summit, Boston, May'17

Introduction

- OvS-DPDK, compelling solution to achieve NFV goals
- Zero-Trust environments
- Connection Tracker and performance
- OvS-DPDK for NFV use cases
- Troubleshooting
- Hot Spot Analysis

Connection Tracker



Stateful vs Stateless Firewalls

Software Must Be Easier, Right?



Gerry Smith ♥ Become a fan
 Gerald.Smith@huffingtonpost.com



 Follow

Matt L.
 04 April

White House Hacked In Cyber Attack That Used Spear-Phishing To Crack Unclassified Network

Posted: 10/01/2012 12:35 pm EDT

NSA Admin Offers Rep.

Paul Wagenseil, SecurityNe
 07 June 2011 05:19 PM ET

Experts: Google's 'Aurora' hackers still at it years later

ring 'Major'

Attacks on Google and others three years ago have similarities to new attacks, including use of rare zero-day vulnerabilities.

arfare
 asking if our own infrastructure

Launch an instance with Security Groups

Launch Instance

Details *

Source *

Flavor *

Networks

Network Ports

Security Groups

Key Pair

Configuration

Server Groups

Scheduler Hints

Metadata

Select the security groups to launch the instance in.

▼ Allocated 5

Name	Description	
Internal_access	For internal access	▼
External_access	For external access	▼
DNS	DNS configuration (port 53)	▼
WWW	Allow HTTP/HTTPS (Port 80/443)	▼
My SQL Client	My SQL Client (port 3306)	▼

▼ Available 2

Select one or more

Q

Click here for filters.

x

Name	Description	
default	Default security group	▲
My SQL Server	My SQL server (port 3306)	▲

✕ Cancel

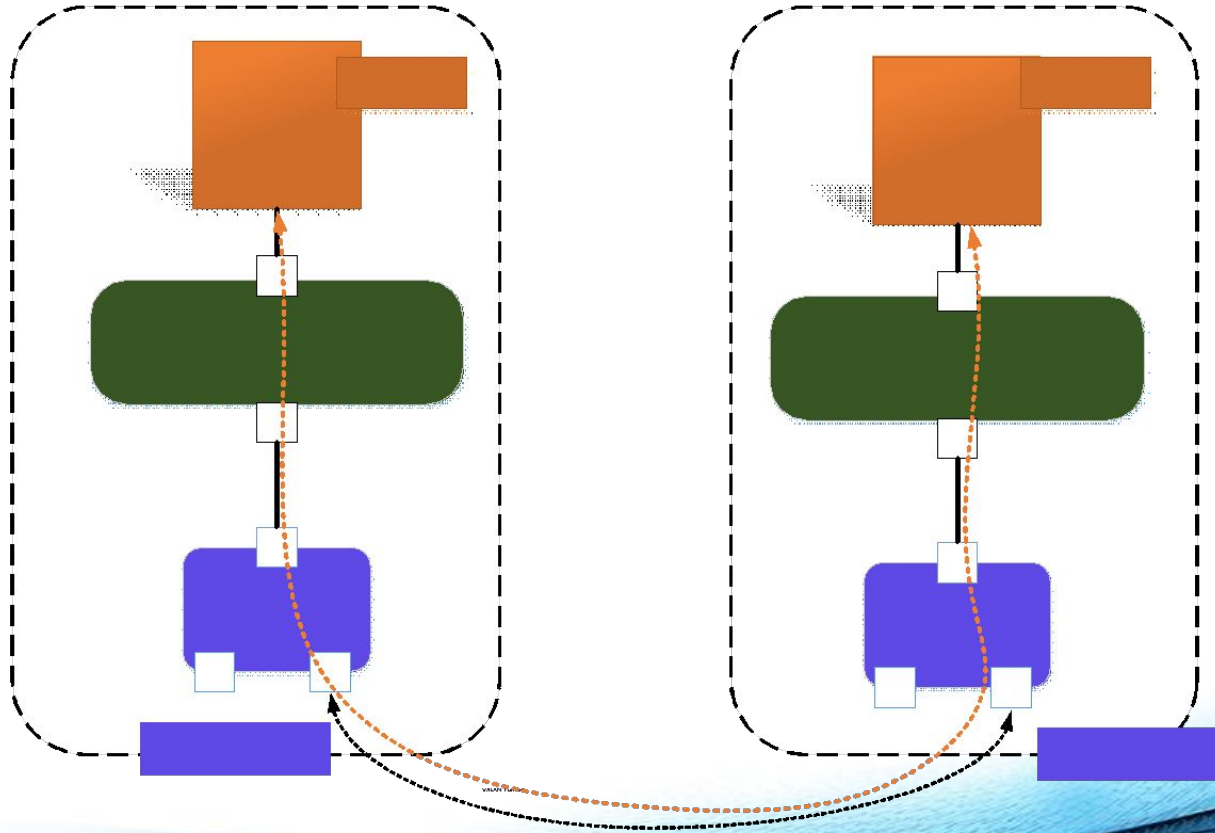
< Back

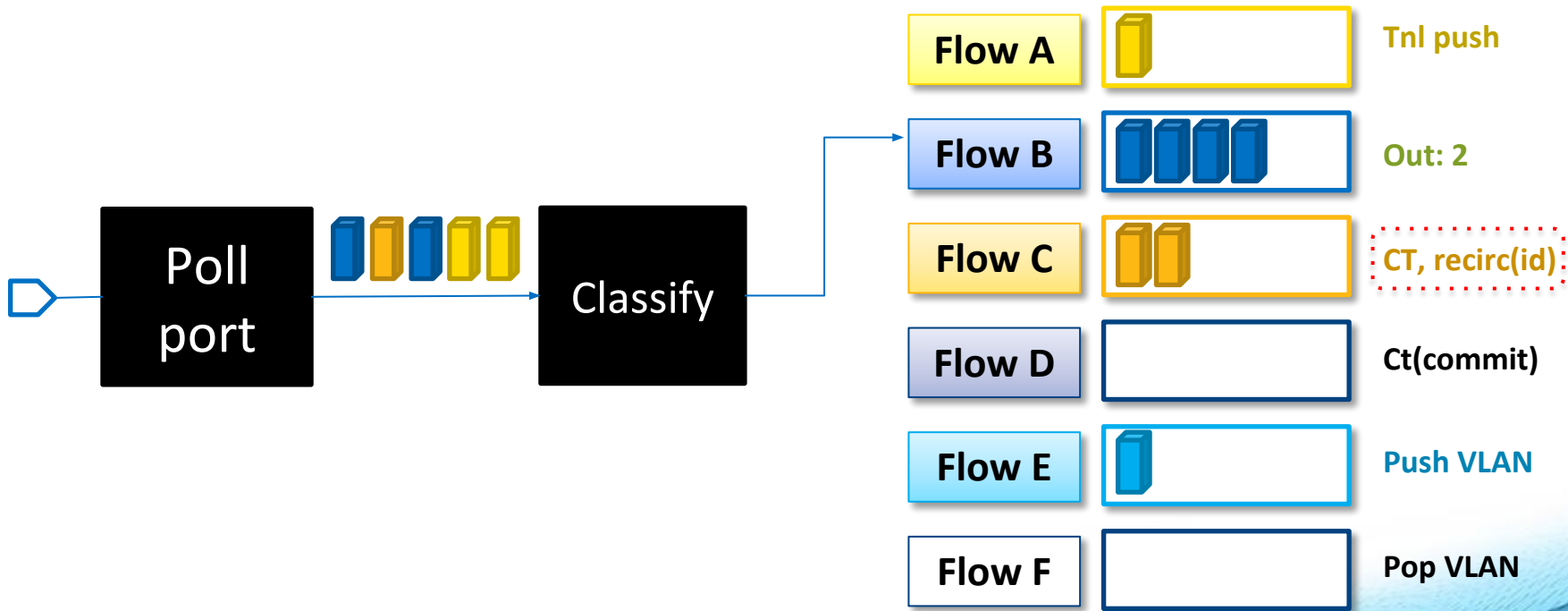
Next >

Launch Instance

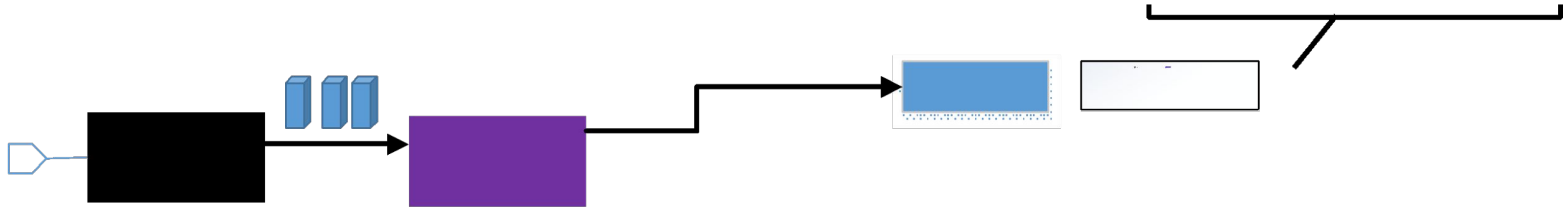


Verifying VM connectivity with SG enabled



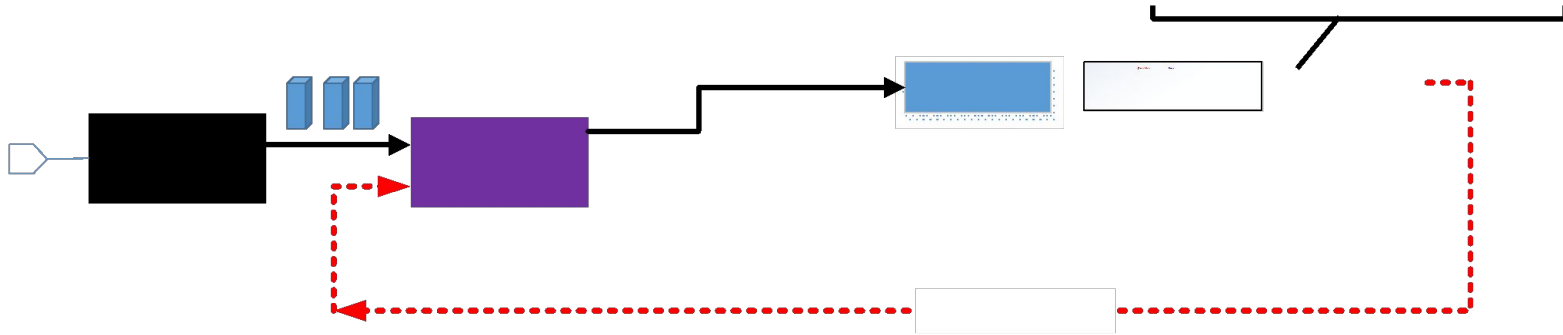


Recirculation explained



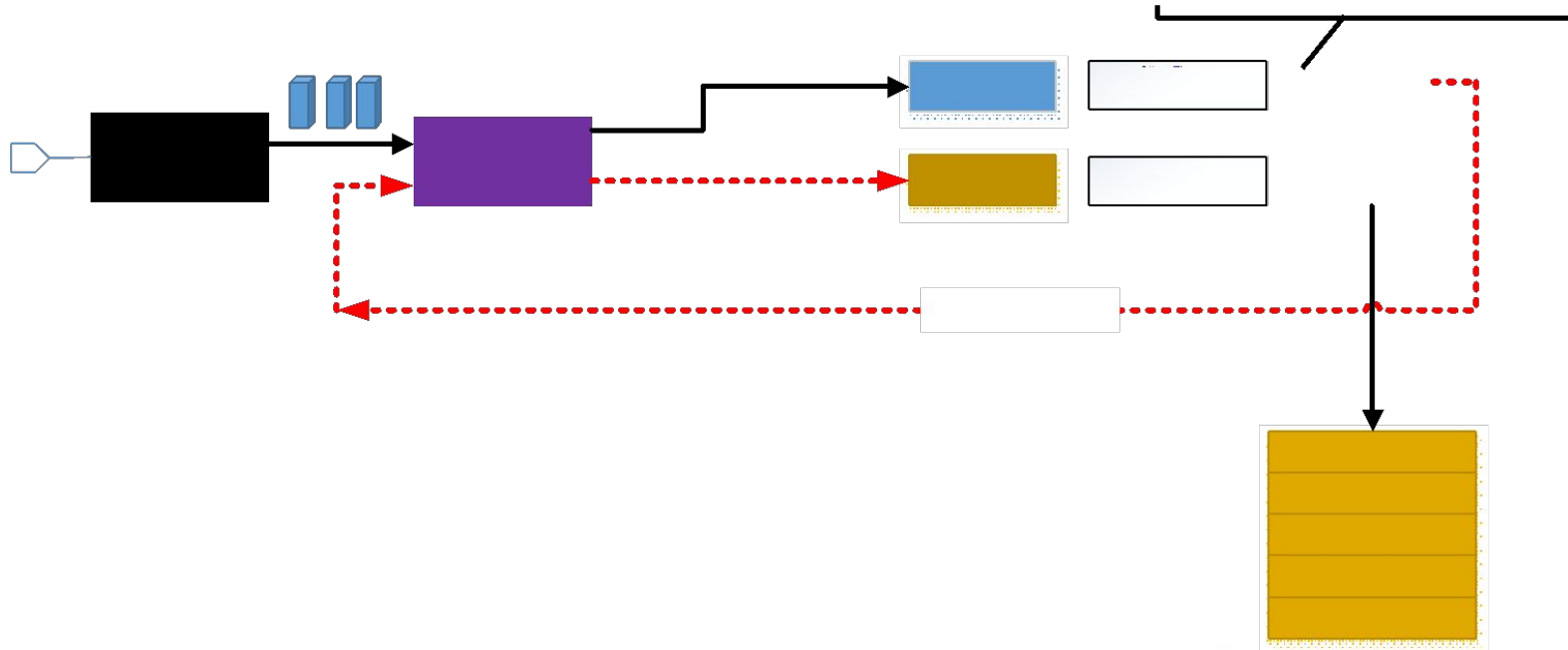
Conntrack sets the connection bits in packet meta data

Recirculation explained



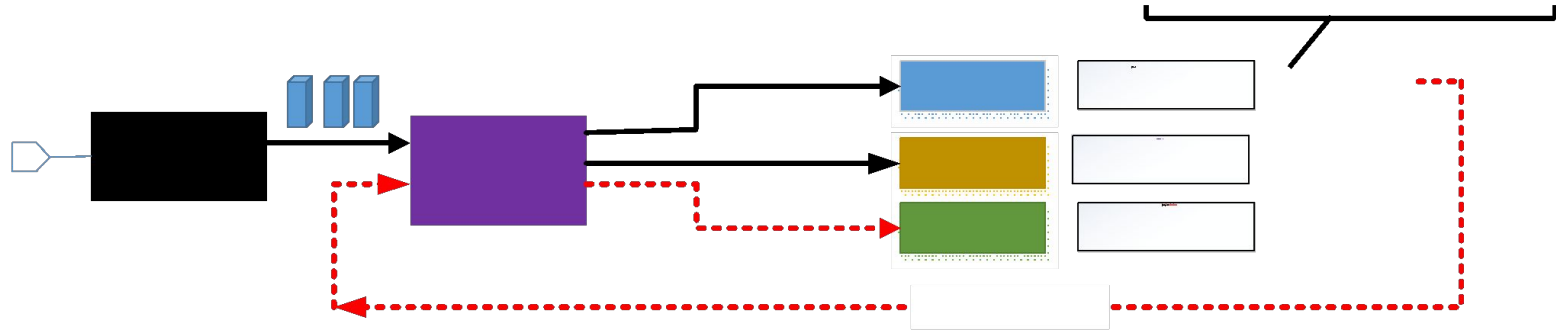
'recirc' flag is set and so the packets are sent back to OvS flow table

Recirculation explained



- Recirculated packet matches the 'flow c' due to connection bits
- ct(commit) - update the connection info in to CT Table
- output:2 - forward the packet on to output port 2

Recirculation explained



Once the connection is established, recirculated packets will match 'Flow D' and get forwarded on output port 2

Compute Node 1 (egress)

Flow Entries

Rule 1: Match: `ct_state(-trk)`, `recirc_id(0)`, `in_port(6)`, `eth(src=fa:16:3e:2a:05:ce)`, `eth_type(0x0800)`, `ipv4(src=10.0.0.3,proto=1,frag=no)`, ...
actions: `ct(zone=4)`, `recirc(0x188)`

Rule 2: Match: `ct_state(-new+est-rel- rpl)`, `recirc_id(0x188)`, `in_port(6)`, `eth(src=fa:16:3e:2a:05:ce,dst=fa:16:3e:9f:55:c1)`, `eth_type(0x0800)`, `ipv4(tos=0/0x3,frag=no)`, ...
actions: `tnl_push(tnl_port(5),header(size=50,type=4,`
`eth(dst=90:e2:ba:6b:1f:61,src=90:e2:ba:2b:49:41,dl_type=0x0800)`,
`ipv4(src=10.237.100.2,dst=10.237.100.3,proto=17,...)`,
`udp(src=0,dst=4789,csum=0x0),vxlan(flags=0x8000000,vni=0x4)),out_port(2))`

Rule 3: Match: `recirc_id(0)`, `in_port(2)`, `eth(src=90:e2:ba:2b:49:41,dst=90:e2:ba:6b:1f:61)`, `eth_type(0x0800)`, `ipv4(frag=no)`,
actions: 3

vHost port

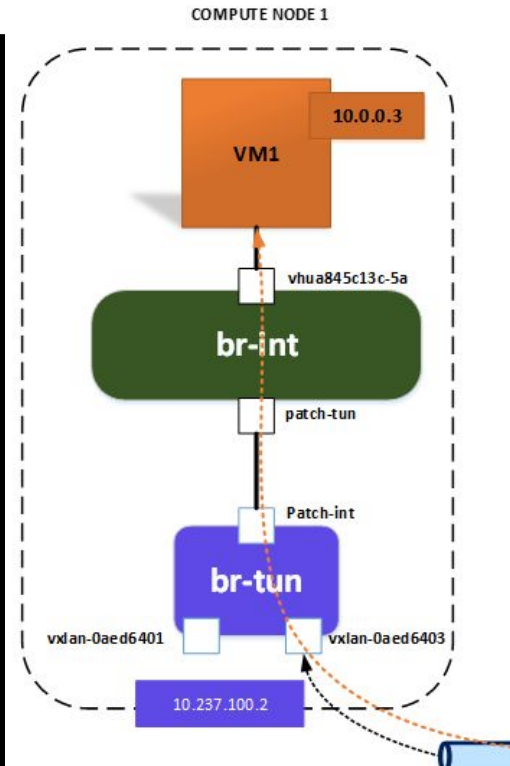
vHost port

Vxlan

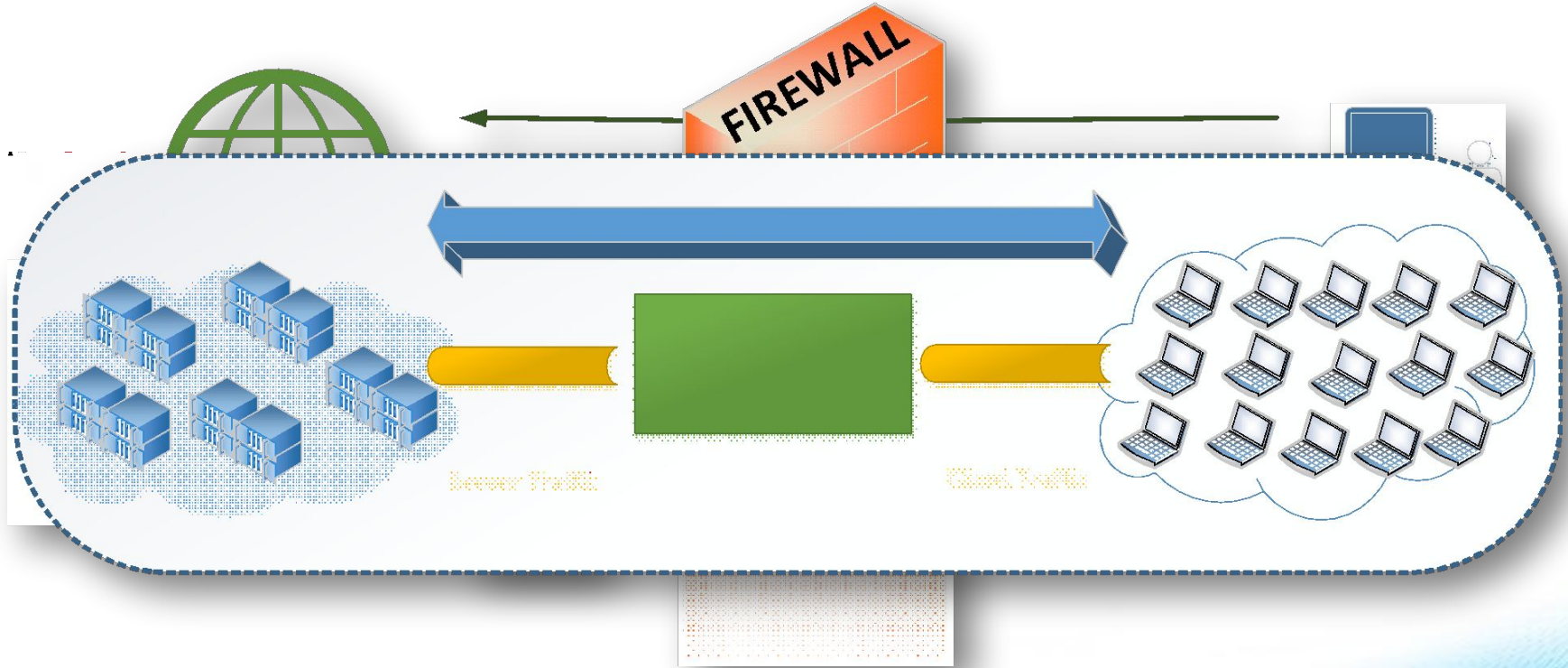
br-ens785f1

br-ens785f1

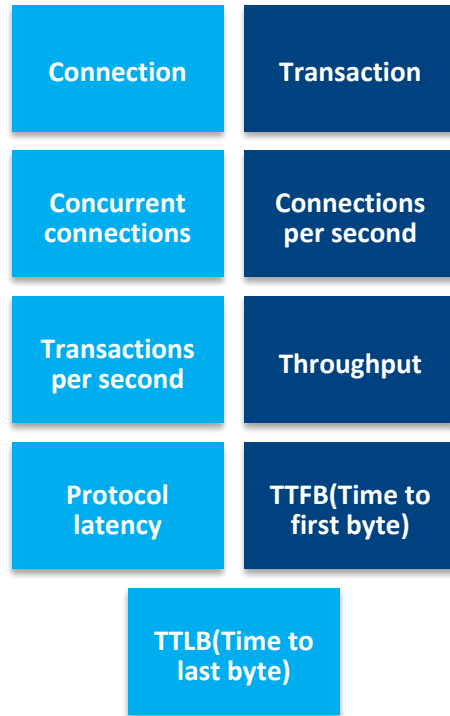
ens785f1



Testing Stateful Firewall, Setup Details



Key Performance Metrics



Desirable Connection

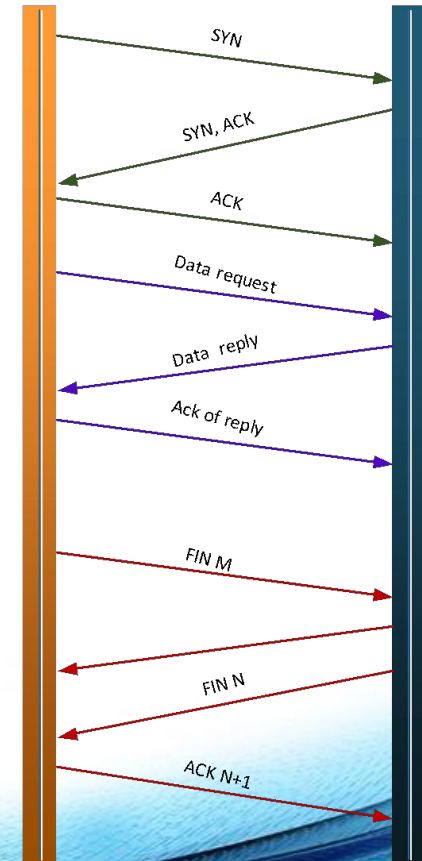
1. TCP connection establishment (SYN, SYN-ACK, ACK)
2. layer 7 transaction(Request, Response)
3. TCP teardown(FIN, ACK).

L4 performance with limited L7 transactions

1. TCP connection establishment (SYN, SYN-ACK, ACK)
2. **Partial/incomplete** layer 7 transaction(only Request)
3. TCP teardown(FIN, ACK).

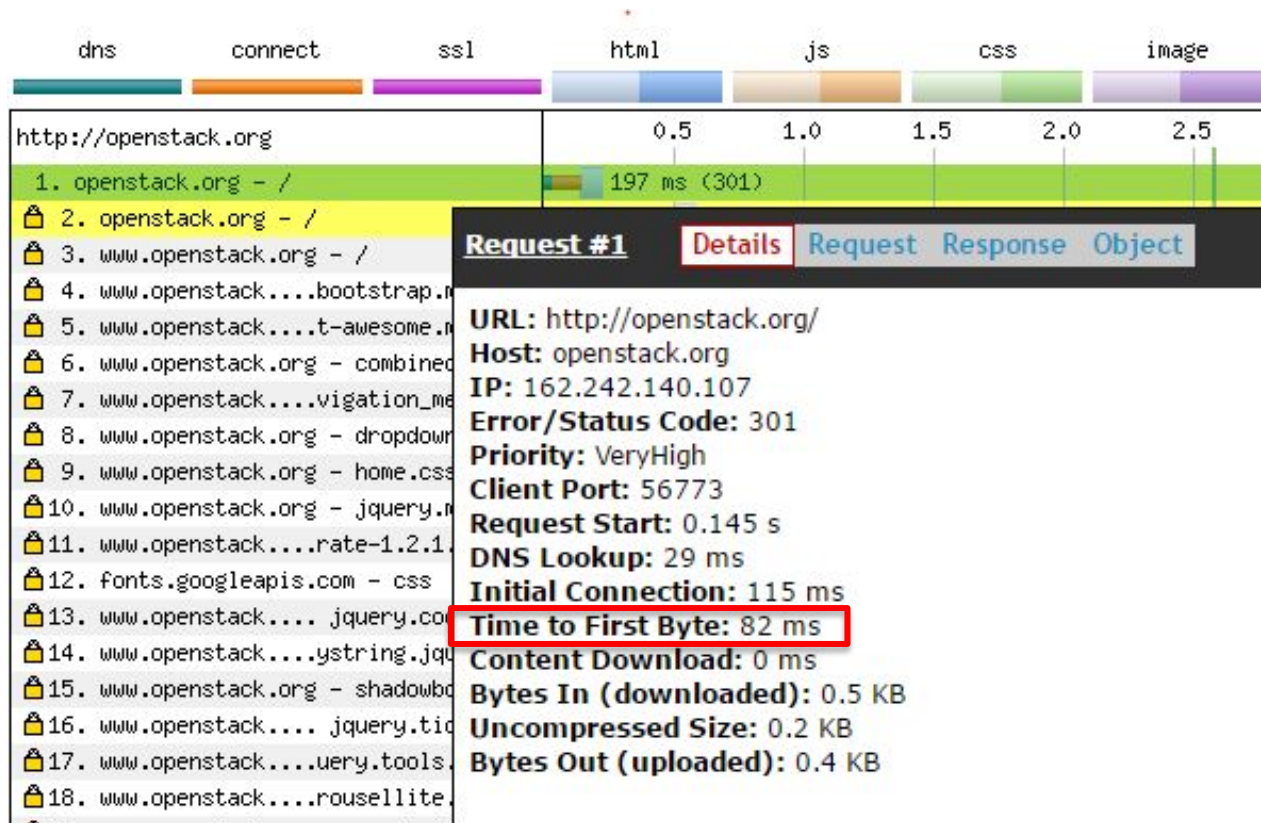
Forceful connection teardowns

1. TCP connection establishment (SYN, SYN-ACK, ACK)
2. Partial/incomplete layer 7 transaction(Request)
3. **Forced TCP teardown**(RST).



Waterfall view of response time

**Results for
openstack.org website
retrieved using
Webpagetest.org**



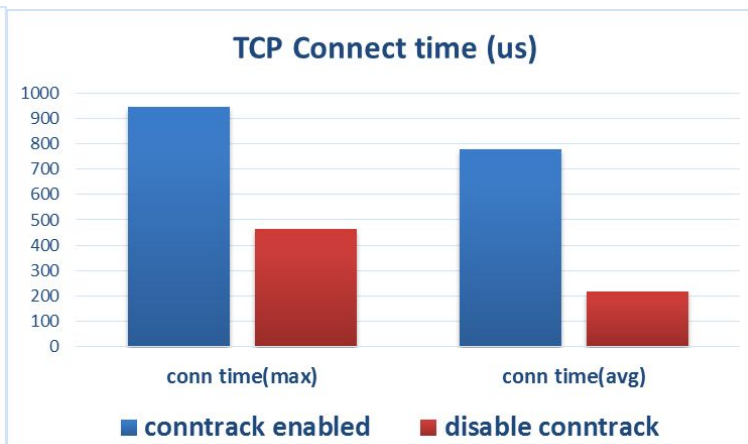
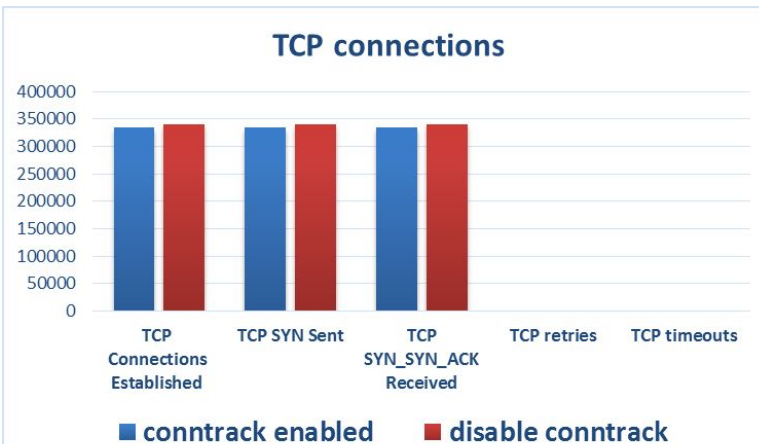
An aerial photograph of a dense city skyline, featuring numerous skyscrapers. Four of these buildings are topped with large, square Intel Xeon processors, which are silver with black and gold accents. The processors are positioned on the roofs of four different buildings, creating a visual metaphor for high performance in a city environment. The word "Performance" is overlaid in large white text across the center of the image.

Performance

HTTP client/server – Throughput test



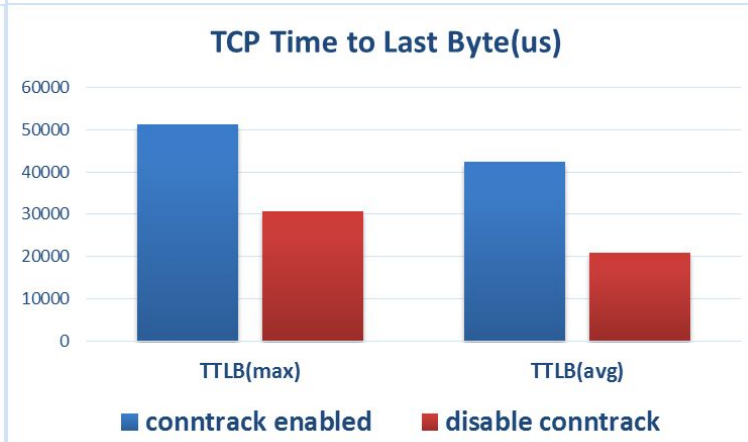
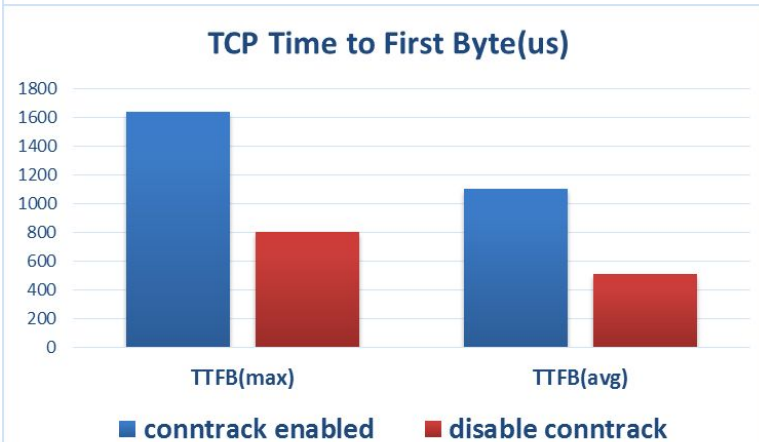
The Higher
the better



The Lower
the better

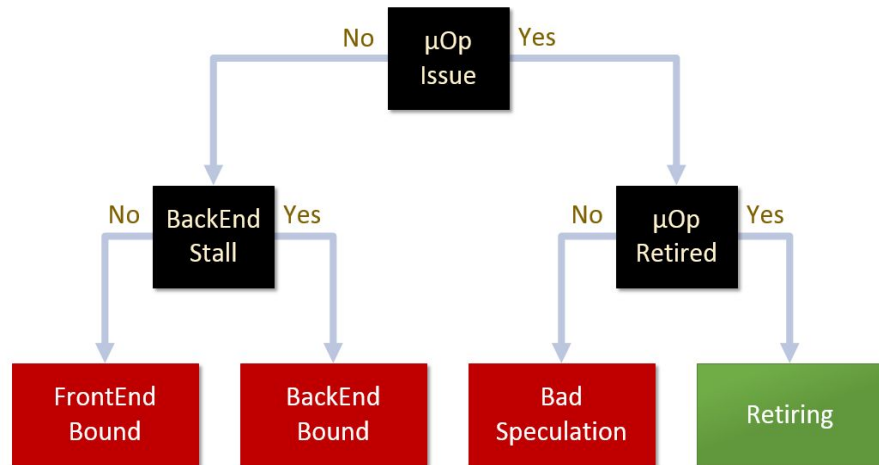


The Lower
the better



The Lower
the better

Analyzing Bottlenecks – Throughput test



Category

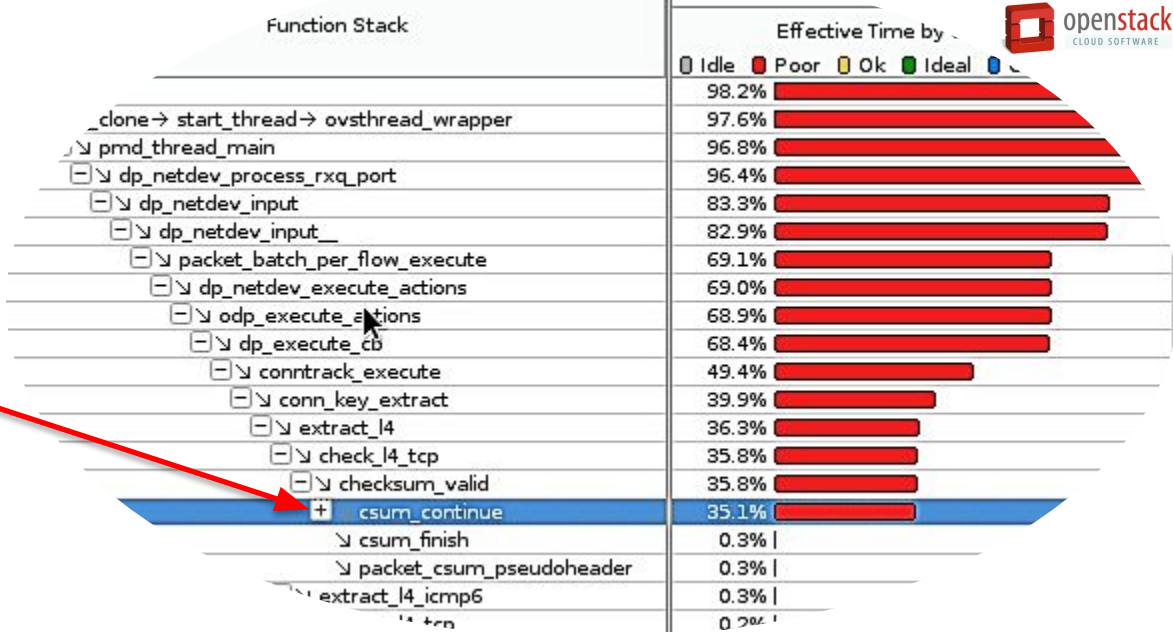
Category	Acceptable Values for HPC applications		OvS-DPDK <u>conntrack</u> (http client/server throughput test)
	↑ The Higher the better	↓ The Lower the better	
Front-End Bound	5 – 10%	↓	15.2%
Bad Speculation	1 – 5%	↓	2.8%
Back-End Bound	20 – 40%	↓	35.9%
Retiring	30 – 70%	↑	46.1%

⑤ CPU Time[?]: 122.500s

Total Thread Count: 14
Paused Time ^(?): 0s

This section lists the most active functions in your a

Function	Module	CPU Time ^②
<u>csum_continue</u>	ovs-vswitchd	42.831s
<u>40e_rcv_pkts_vec</u>	ovs-vswitchd	6.530s
<u>clock_gettime</u>	libc.so.6	3.500s
<u>miniflow_extract</u>	ovs-vswitchd	3.350s
<u>tcp_conn_update</u>	ovs-vswitchd	2.550s
[Others]	N/A*	63.738s

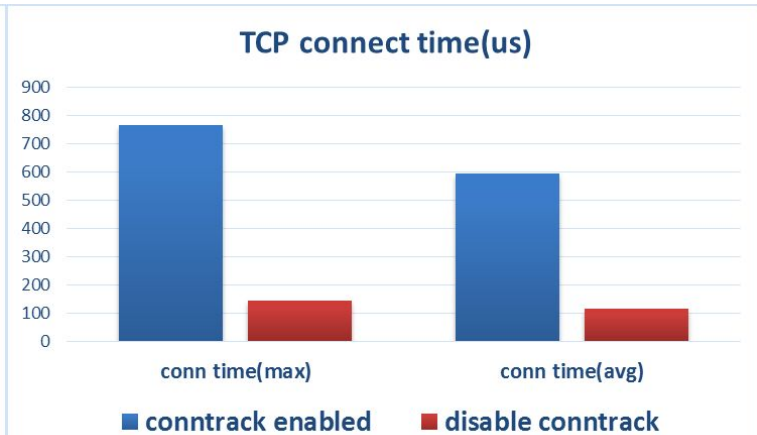
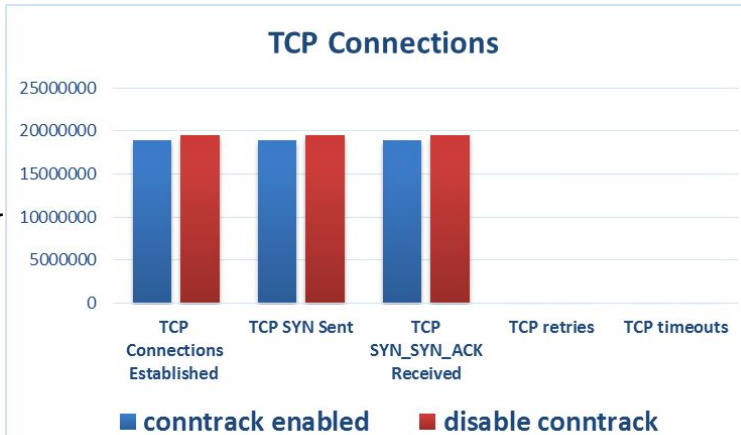


Function / Call Stack	CPI Rate	Front-End Bound	Bad Speculati...	Back-End Bound										Retiring	
				Memory Bound				Store Bound				Core Bound			
				L1 Bound	L2 Bound	L3 Bound	DRAM Bound	Store Late...	False Shar...	Split Stores	DTLB Store Over...	Divid...	Port Utilization		
+tsum_continue	%	0.325	8.8%	1.0%	55.5%	0.0%	27.6%	5.6%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	55.1%
+miniflow_extract	%	0.315	13.8%	0.0%	54.2%	0.0%	29.3%	0.0%	0.0%	0.0%	0.0%	2.1%	0.0%	0.0%	61.2%
+emc_lookup	%	0.702	11.1%	13.9%	54.3%	0.0%	48.2%	0.0%	0.0%	0.0%	0.0%	3.2%	0.0%	0.0%	33.9%
+tcp_conn_update	%	0.368	15.5%	3.4%	54.6%	0.0%	18.9%	0.0%	0.0%	0.0%	0.0%	2.0%	0.0%	0.0%	55.6%
+i40e_recv_pkts_vec	%	0.848	6.6%	5.6%	55.1%	0.0%	25.2%	1.2%	0.0%	0.0%	0.8%	2.1%	0.0%	43.3%	22.4%
+H06"leca"bktz"Aec	%	0.848	6.6%	5.6%	55.1%	0.0%	25.2%	1.2%	0.0%	0.0%	0.8%	2.1%	0.0%	43.3%	22.4%
+tcp_conn_update	%	0.368	15.5%	3.4%	54.6%	0.0%	18.9%	0.0%	0.0%	0.0%	0.0%	2.0%	0.0%	0.0%	55.6%
+H06"leca"bktz"Aec	%	0.848	6.6%	5.6%	55.1%	0.0%	25.2%	1.2%	0.0%	0.0%	0.8%	2.1%	0.0%	43.3%	22.4%

HTTP client/server – 65k connections/sec



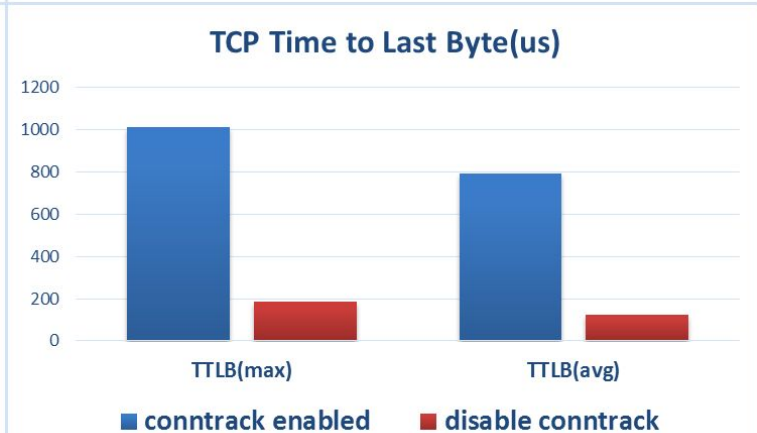
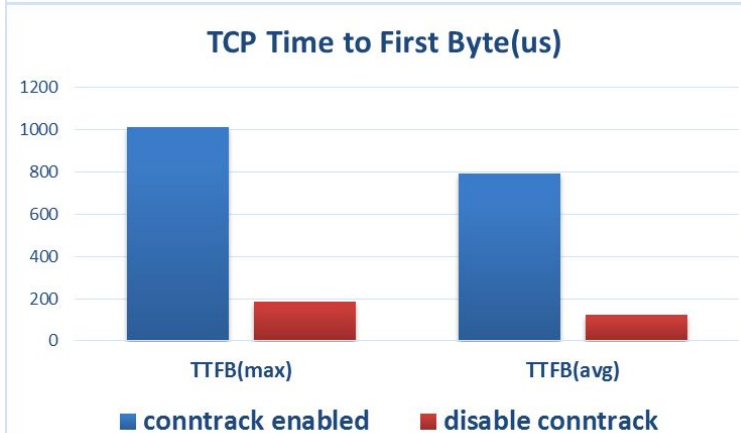
The Higher
the better



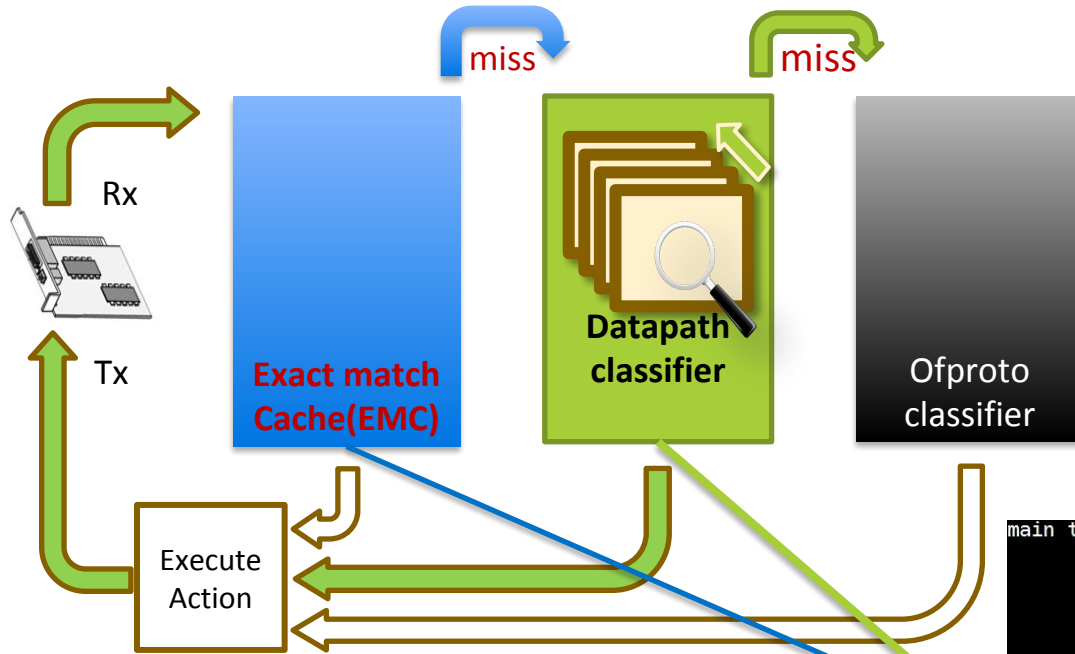
The Lower
the better



The Lower
the better



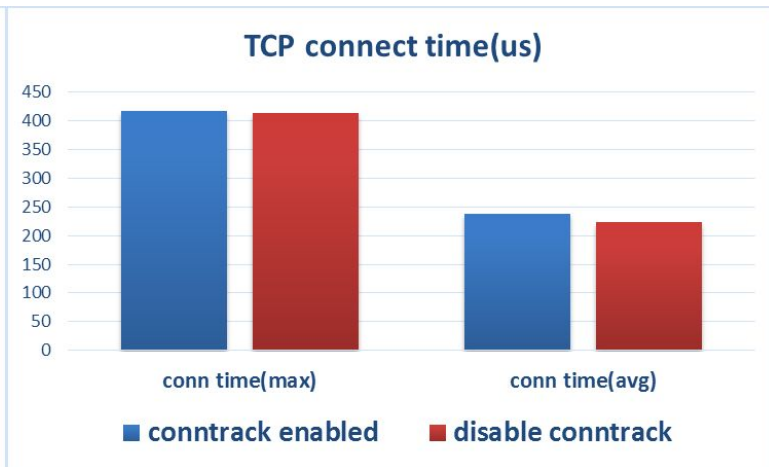
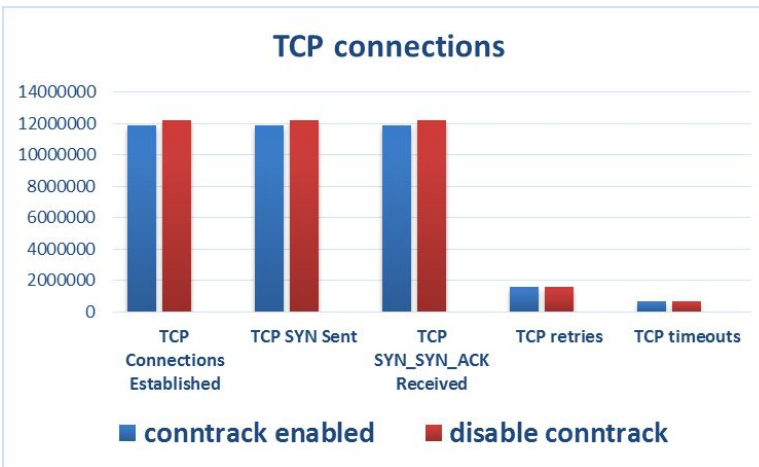
The Lower
the better



```
main thread:
  emc entries:0 (0.00% full)
  emc hits:0
  megaflow hits:0
  avg. subtable lookups per hit:0.00
  miss:0
  lost:0
  polling cycles:572403 (100.00%)
  processing cycles:0 (0.00%)
  sub thread numa_id 0 core_id 3:
    emc entries:8192 (100.00% full)
    emc hits:173
    megaflow hits:4960084
    avg. subtable lookups per hit:1.57
    miss:0
    lost:0
    polling cycles:3884380448 (39.72%)
    processing cycles:5895049012 (60.28%)
    avg cycles per packet: 766.74 (9779429460/12754628)
    avg processing cycles per packet: 462.19 (5895049012/12754628)
```

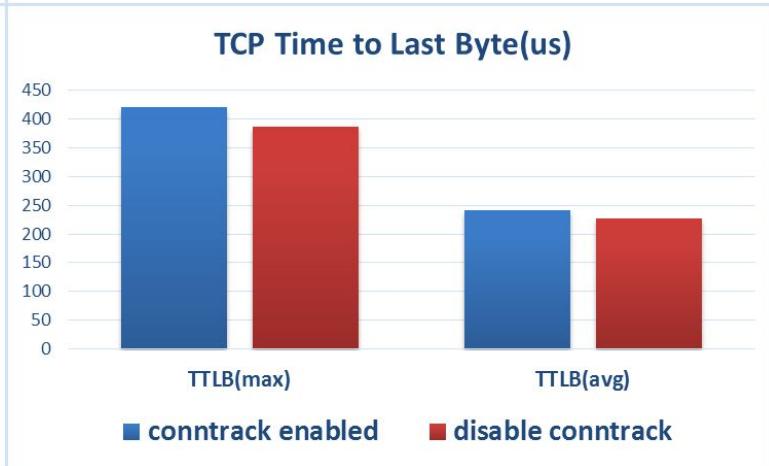
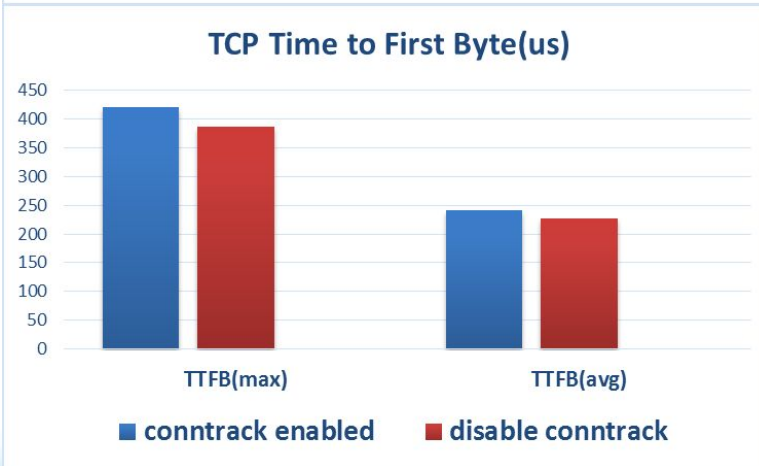
HTTP client/server -800k concurrent connections

↑
The Higher
the better



↓
The Lower
the better

↓
The Lower
the better

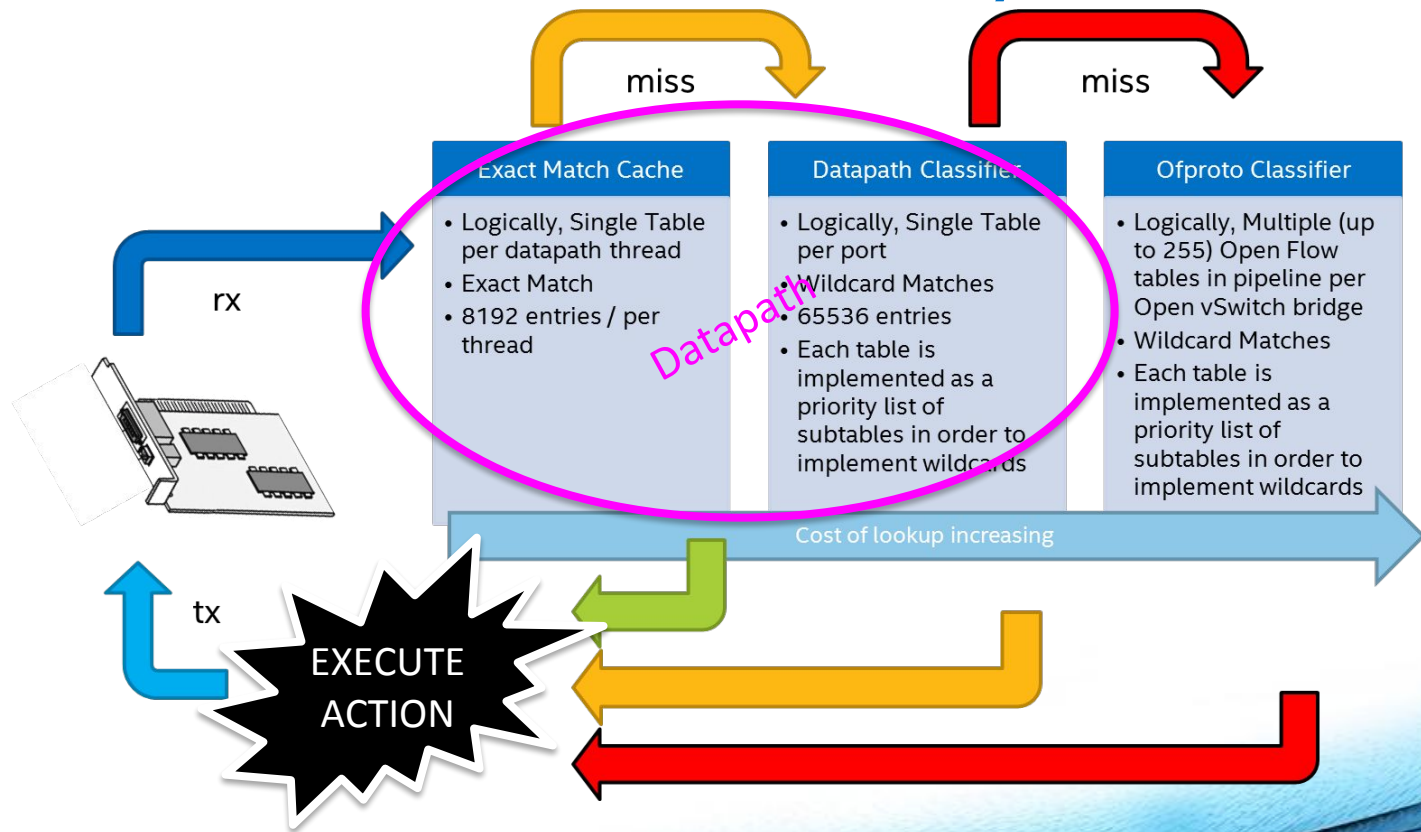


↓
The Lower
the better

OVS-DPDK in real world telco deployments

- Not to showcase OVS-DPDK performance figures
- OVS-DPDK for Telco use cases.
- Debugging and performance fine-tuning of OVS-DPDK
- OVS-DPDK hotspot analysis

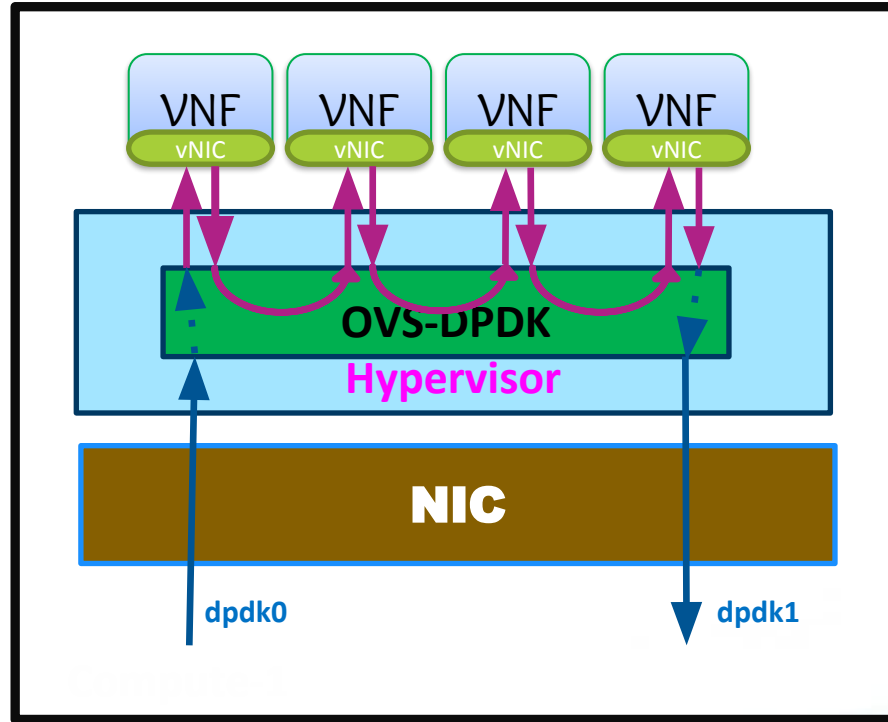
OVS-DPDK Table Hierarchy



OvS-DPDK Troubleshooting

A person is sitting on the floor, completely surrounded by a massive, chaotic tangle of network cables. The cables are of various colors, including black, white, yellow, and blue, and are densely packed, creating a complex web around the person. The person is wearing light-colored trousers and dark shoes. The scene is set in what appears to be a server room or a data center, with the floor being a light-colored carpet. The overall image conveys a sense of being overwhelmed or stuck in a complex technical environment.

East-West traffic test



Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-ports-desc br0 |more
OFPST_PORT_DESC reply (xid=0x2):
```

```
1(dpdk0): addr:90:e2:ba:2b:49:40
  config:      0
  state:       0
  current:     10GB-FD
  speed: 10000 Mbps now, 0 Mbps max
2(dpdk1): addr:90:e2:ba:2b:49:41
  config:      0
  state:       0
  current:     10GB-FD
  speed: 10000 Mbps now, 0 Mbps max
3(dpdkvhostuser0): addr:00:00:00:00:00:00
  config:      0
  state:       0
  speed: 0 Mbps now, 0 Mbps max
4(dpdkvhostuser1): addr:00:00:00:00:00:00
  config:      0
  state:       0
  speed: 0 Mbps now, 0 Mbps max
5(dpdkvhostuser2): addr:00:00:00:00:00:00
  config:      0
  state:       0
```

**ovs-ofctl dump-ports-desc
**

--More--

Detailed information about the openflow ports in a switch.

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-ports-desc br0 |more
OFPST_PORT_DESC reply (xid=0x2):
```

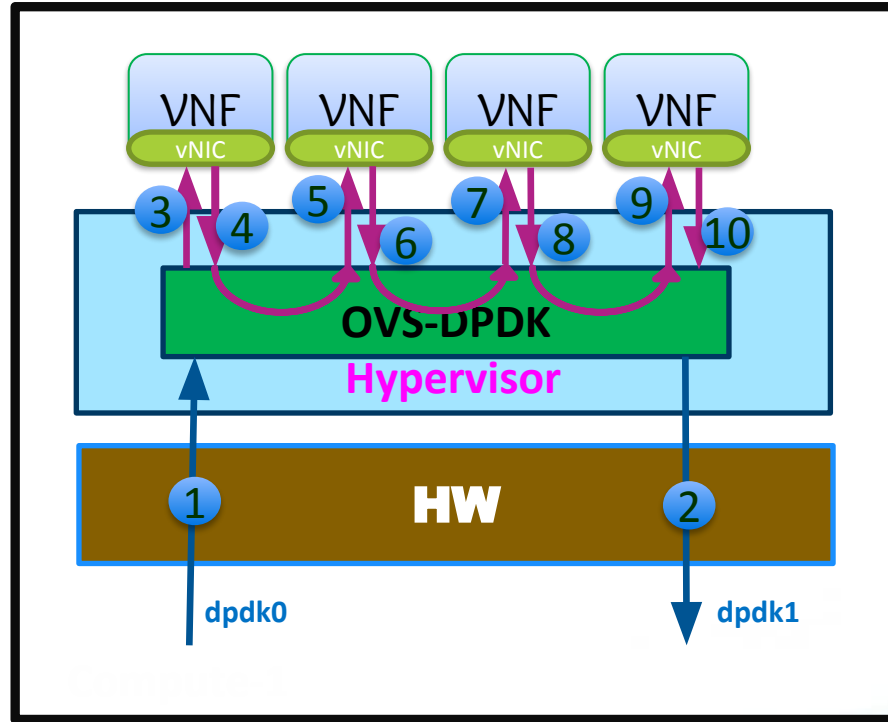
```
1(dpdk0): addr:90:e2:ba:2b:49:40
  config:      0
  state:       0
  current:     10GB-FD
  speed: 10000 Mbps now, 0 Mbps max
2(dpdk1): addr:90:e2:ba:2b:49:41
  config:      0
  state:       0
  current:     10GB-FD
  speed: 10000 Mbps now, 0 Mbps max
3(dpdkvhostuser0): addr:00:00:00:00:00:00
  config:      0
  state:       0
  speed: 0 Mbps now, 0 Mbps max
4(dpdkvhostuser1): addr:00:00:00:00:00:00
  config:      0
  state:       0
  speed: 0 Mbps now, 0 Mbps max
5(dpdkvhostuser2): addr:00:00:00:00:00:00
  config:      0
  state:       0
```

**ovs-ofctl dump-ports-desc
**

--More--

Detailed information about the openflow ports in a switch.

Debugging



● Openflow port

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-flows br0
NXST_FLOW reply (xid=0x4):
 cookie=0x0, duration=177.460s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=1, dl_src=00:00:00:00:00:ff, dl_dst=00:00:00:00:00:01 actions=output:3
 cookie=0x0, duration=177.412s, table=0, n_packets=38, n_bytes=7008, idle_age=12, in_port=3 actions=output:1
 cookie=0x0, duration=177.360s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=4 actions=output:5
 cookie=0x0, duration=177.320s, table=0, n_packets=41, n_bytes=7659, idle_age=3, in_port=5 actions=output:4
 cookie=0x0, duration=177.272s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=6 actions=output:7
 cookie=0x0, duration=177.228s, table=0, n_packets=40, n_bytes=7152, idle_age=0, in_port=7 actions=output:6
 cookie=0x0, duration=177.176s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=8 actions=output:9
 cookie=0x0, duration=177.136s, table=0, n_packets=37, n_bytes=6394, idle_age=14, in_port=9 actions=output:8
 cookie=0x0, duration=177.068s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=10 actions=output:2
 cookie=0x0, duration=177.024s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=2 actions=output:10
 sugeshch@silpixa00389816:~/repo/ovs_master$
```

**ovs-ofctl dump-flows
**

List all openflow flow entries in switch tables.

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-flows br0
NXST_FLOW reply (xid=0x4):
 cookie=0x0, duration=177.460s, table=0, n_packets=0, n_bytes=0 idle_age=177, in_port=1, dl_src=00:00:00:00:00:ff, dl_dst=00:00:00:00:00:01 actions=output:3
 cookie=0x0, duration=177.412s, table=0, n_packets=38, n_bytes=7008, idle_age=12, in_port=3 actions=output:1
 cookie=0x0, duration=177.360s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=4 actions=output:5
 cookie=0x0, duration=177.320s, table=0, n_packets=41, n_bytes=7659, idle_age=3, in_port=5 actions=output:4
 cookie=0x0, duration=177.272s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=6 actions=output:7
 cookie=0x0, duration=177.228s, table=0, n_packets=40, n_bytes=7152, idle_age=0, in_port=7 actions=output:6
 cookie=0x0, duration=177.176s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=8 actions=output:9
 cookie=0x0, duration=177.136s, table=0, n_packets=37, n_bytes=6394, idle_age=14, in_port=9 actions=output:8
 cookie=0x0, duration=177.068s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=10 actions=output:2
 cookie=0x0, duration=177.024s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=2 actions=output:10
sugeshch@silpixa00389816:~/repo/ovs_master$
```

**ovs-ofctl dump-flows
**

List all openflow flow entries in switch tables.

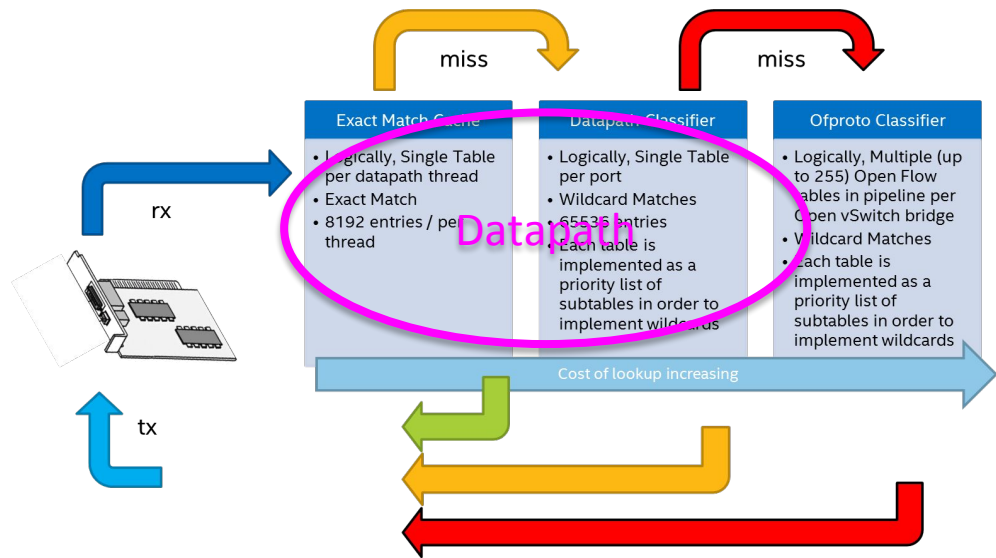
Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-flows br0
NXST_FLOW reply (xid=0x4):
 cookie=0x0, duration=177.460s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=1, dl_src=00:00:00:00:00:ff, dl_dst=00:00:00:00:00:01 actions=output:3
 cookie=0x0, duration=177.412s, table=0, n_packets=38, n_bytes=7008, idle_age=12, in_port=3 actions=output:1
 cookie=0x0, duration=177.360s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=4 actions=output:5
 cookie=0x0, duration=177.320s, table=0, n_packets=41, n_bytes=7659, idle_age=3, in_port=5 actions=output:4
 cookie=0x0, duration=177.272s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=6 actions=output:7
 cookie=0x0, duration=177.228s, table=0, n_packets=40, n_bytes=7152, idle_age=0, in_port=7 actions=output:6
 cookie=0x0, duration=177.176s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=8 actions=output:9
 cookie=0x0, duration=177.136s, table=0, n_packets=37, n_bytes=6394, idle_age=14, in_port=9 actions=output:8
 cookie=0x0, duration=177.068s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=10 actions=output:2
 cookie=0x0, duration=177.024s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=2 actions=output:10
 sugeshch@silpixa00389816:~/repo/ovs_master$
```

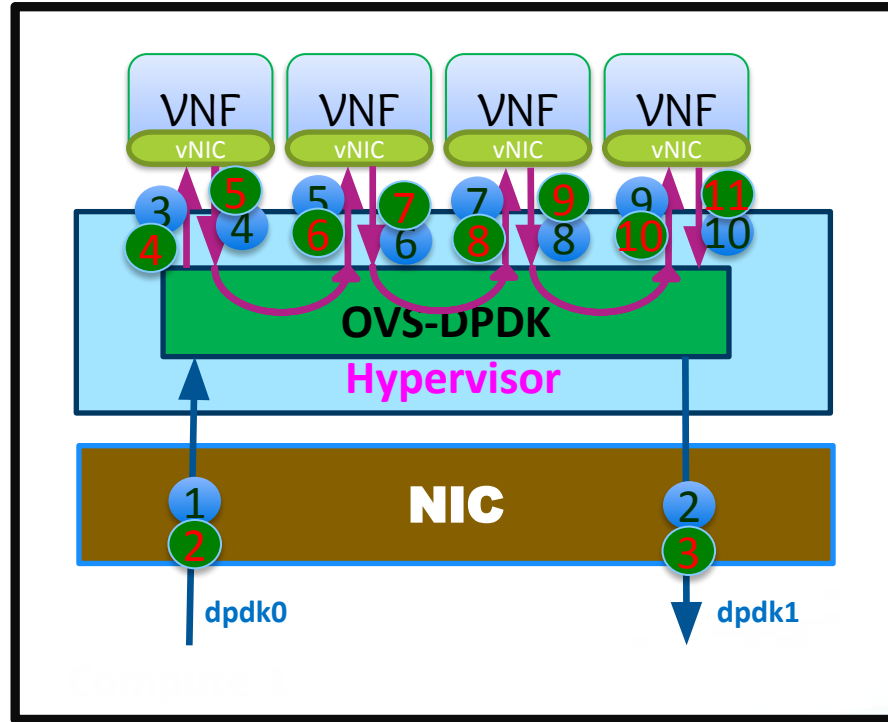
**ovs-ofctl dump-flows
**

List all openflow flow entries in switch tables.

Datapath Debugging



Debugging



● Openflow port

● Datapath port

Debugging

```
netdev@ovs-netdev: hit:17543389635 missed:188
br0:
    br0 65534/1: (tap)
        dpdk0 1/2: (dpdk: configured_rx_queues=1, configured_rxq_descriptors=2048, configured_tx_queues=5, configured_txq_descriptors=2048, mtu=1500, requested_rx_queues=1, requested_rxq_descriptors=2048, requested_tx_queues=5, requested_txq_descriptors=2048, rx_csum_offload=true)
        dpdk1 2/3: (dpdk: configured_rx_queues=1, configured_rxq_descriptors=2048, configured_tx_queues=5, configured_txq_descriptors=2048, mtu=1500, requested_rx_queues=1, requested_rxq_descriptors=2048, requested_tx_queues=5, requested_txq_descriptors=2048, rx_csum_offload=true)
        dpdkvhostuser0 3/4: (dpdkvhostuser: configured_rx_queues=1, configured_tx_queues=1, mtu=1500, requested_rx_queues=1, requested_tx_queues=1)
        dpdkvhostuser1 4/5: (dpdkvhostuser: configured_rx_queues=1, configured_tx_queues=1, mtu=1500, requested_rx_queues=1, requested_tx_queues=1)
        dpdkvhostuser2 5/6: (dpdkvhostuser: configured_rx_queues=1, configured_tx_queues=1, mtu=1500, requested_rx_queues=1, requested_tx_queues=1)
        dpdkvhostuser3 6/7: (dpdkvhostuser: configured_rx_queues=1, configured_tx_queues=1, mtu=1500, requested_rx_queues=1, requested_tx_queues=1)
        dpdkvhostuser4 7/8: (dpdkvhostuser: configured_rx_queues=1, configured_tx_queues=1, m
--More--
```

Summary of datapath with configured port information(Openflow port no, Datapath port no and Type)

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpctl/show -s |more
netdev@ovs-netdev:
  lookups: hit:13798785283 missed:188 lost:0
  flows: 1
  port 0: ovs-netdev (tap)
    RX packets:0 errors:0 dropped:0 overruns:0 frame:0
    TX packets:0 errors:0 dropped:0 aborted:0 carrier:0
    collisions:0
    RX bytes:0 TX bytes:0
  port 1: br0 (tap)
    RX packets:0 errors:0 dropped:0 overruns:0 frame:0
    TX packets:0 errors:0 dropped:0 aborted:0 carrier:0
    collisions:0
    RX bytes:0 TX bytes:0
  port 2: dpdk0 (dpdk: configured_rx_queues=1, configured_rxq_descriptors=2048, configured_tx_queues=5, configured_txq_descriptors=2048, mtu=1500, requested_rx_queues=1, requested_rxq_descriptors=2048, requested_tx_queues=5, requested_txq_descriptors=2048, rx_csum_offload=true)
    RX packets:13798789152 errors:0 dropped:1803077122 overruns:? frame:?
    TX packets:107 errors:0 dropped:0 aborted:? carrier:?
    collisions:?
--More--
```

Summary of datapath with list of data-path-ports. -s shows the stats

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpctl/show -s |more
netdev@ovs-netdev:
lookups: hit:13798785283 missed:188 lost:0
flows: 1
port 0: ovs-netdev (tap)
  RX packets:0 errors:0 dropped:0 overruns:0 frame:0
  TX packets:0 errors:0 dropped:0 aborted:0 carrier:0
  collisions:0
  RX bytes:0 TX bytes:0
port 1: br0 (tap)
  RX packets:0 errors:0 dropped:0 overruns:0 frame:0
  TX packets:0 errors:0 dropped:0 aborted:0 carrier:0
  collisions:0
  RX bytes:0 TX bytes:0
port 2: dpdk0 (dpdk: configured_rx_queues=1, configured_rxq_descriptors=2048, configured_tx_queues=5, configured_txq_descriptors=2048, mtu=1500, requested_rx_queues=1, requested_rxq_descriptors=2048, requested_tx_queues=5, requested_txq_descriptors=2048, rx_csum_offload=true)
  RX packets:13798789152 errors:0 dropped:1803077122 overruns:? frame:?
  TX packets:107 errors:0 dropped:0 aborted:? carrier:?
  collisions:?
--More--
```

Summary of datapath with list of data-path-ports. -s shows the stats

Debugging

ovs-appctl dpctl/dump-flows

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpctl/dump-flows
flow-dump from pmd on cpu core: 2
recirc_id(0),in_port(2),eth(src=00:00:00:00:00:0f,dst=00:00:00:00:00:01),eth_type(0x0800),ipv4(frag=no), packets:35128942175, bytes:2107736530500, used:0.000s, actions:drop
```

Display all active flow entries in a datapath

Debugging

ovs-appctl dpctl/dump-flows

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpctl/dump-flows
flow-dump from pmd on cpu core: 2
recirc_id(0), in port(2), eth(src=00:00:00:00:00:0f,dst=00:00:00:00:00:01), eth_type(0x0800), ipv4(frag=no), packets:35128942175, bytes:2107736530500, used:0.000s, actions:drop
```

Display all active flow entries in a datapath

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-flows br0
NXST_FLOW reply (xid=0x4):
 cookie=0x0, duration=177.460s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=1, dl_src=00:00:00:00:00:ff, dl_dst=00:00:00:00:00:01 actions=output:3
 cookie=0x0, duration=177.412s, table=0, n_packets=38, n_bytes=7008, idle_age=12, in_port=3 actions=output:1
 cookie=0x0, duration=177.360s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=4 actions=output:5
 cookie=0x0, duration=177.320s, table=0, n_packets=41, n_bytes=7659, idle_age=3, in_port=5 actions=output:4
 cookie=0x0, duration=177.172s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=6 actions=output:7
 cookie=0x0, duration=177.228s, table=0, n_packets=40, n_bytes=7152, idle_age=0, in_port=7 actions=output:6
 cookie=0x0, duration=177.176s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=8 actions=output:9
 cookie=0x0, duration=177.136s, table=0, n_packets=37, n_bytes=6394, idle_age=14, in_port=9 actions=output:8
 cookie=0x0, duration=177.068s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=10 actions=output:2
 cookie=0x0, duration=177.004s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=2 actions=output:10
 sugeshch@silpixa00389816:~/repo/ovs_master$
```

Pkt dl_src = 00:00:00:00:00:0f

ovs-ofctl dump-flows

List all openflow flow entries in switch tables.

Debugging

- Flow configured wrongly.
 - set "dl_src = **00:00:00:00:00:0F** "

```
ovs-ofctl add-flow br0 in_port=1,dl_src=00:00:00:00:00:0F,  
dl_dst=00:00:00:00:00:01,action=output:3
```

Debugging

ovs-appctl dpctl/dump-flows

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpctl/dump-flows
flow-dump from pmd on cpu core: 2
recirc_id(0),in_port(2),eth(src=00:00:00:00:00:0f,dst=00:00:00:00:00:01),eth_type(0x0800),ipv4(frag=no), packets:35128942175, bytes:2107736530500, used:0.000s, actions:drop
```



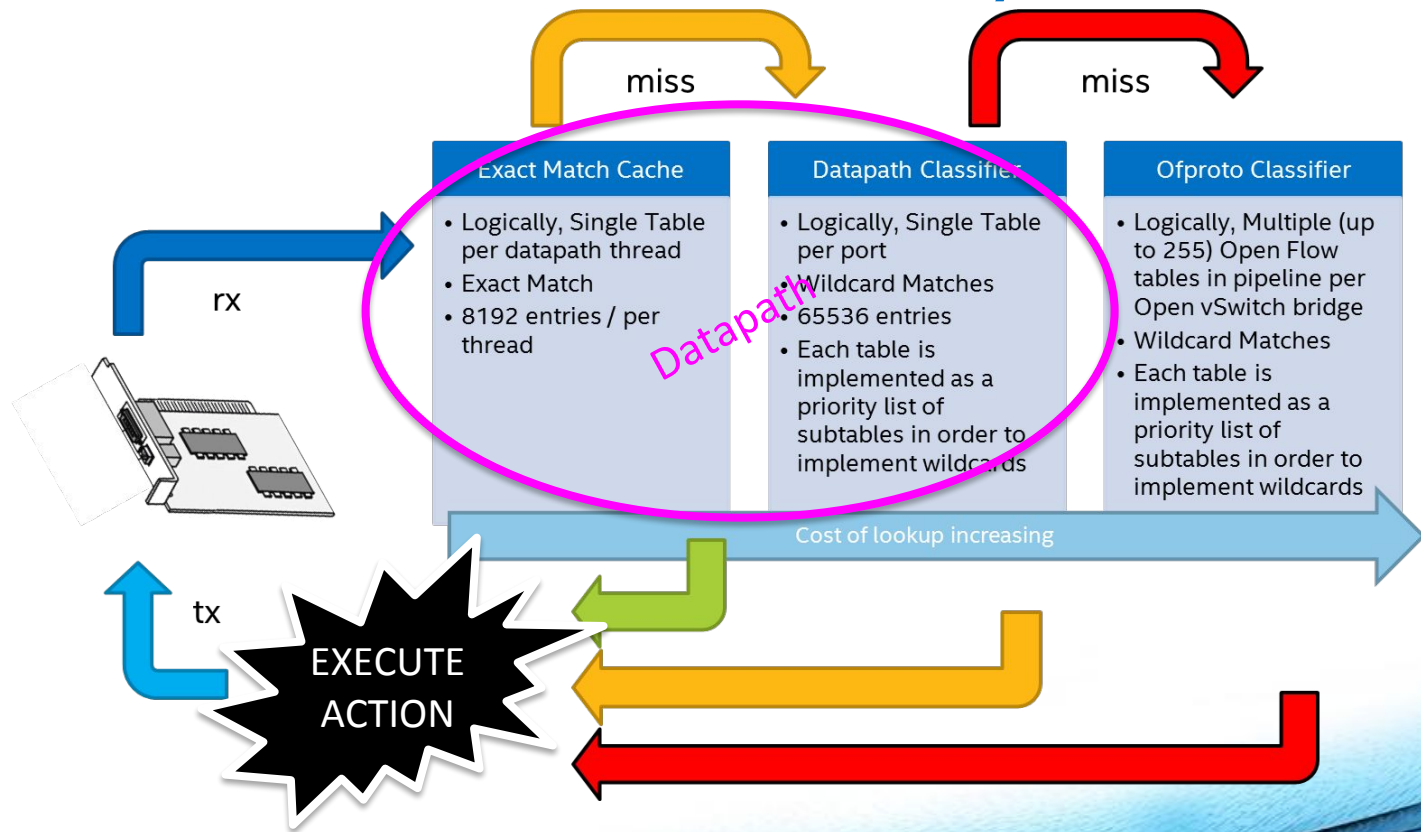
```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpctl/dump-flows
flow-dump from pmd on cpu core: 2
recirc_id(0),in_port(9),eth_type(0x0800),ipv4(frag=no), packets:847435896, bytes:50846153760, used:0.001s, actions:10
flow-dump from pmd on cpu core: 1
recirc_id(0),in_port(5),eth_type(0x0800),ipv4(frag=no), packets:860845534, bytes:51650732040, used:0.002s, actions:6
recirc_id(0),in_port(7),eth_type(0x0800),ipv4(frag=no), packets:831788423, bytes:49907305380, used:0.002s, actions:8
flow-dump from pmd on cpu core: 0
recirc_id(0),in_port(11),eth_type(0x0800),ipv4(frag=no), packets:842759867, bytes:49907305380, used:0.001s, actions:3
recirc_id(0),in_port(2),eth(src=00:00:00:00:00:0f,dst=00:00:00:00:00:01),eth_type(0x0800),ipv4(frag=no), packets:1205917247, bytes:72355034820, used:0.000s, actions:4
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$
```



Display all active flow entries in a datapath

Debugging OVS-DPDK performance issues

OVS-DPDK Table Hierarchy



Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-stats-show |more
```

```
pmd thread numa_id 0 core_id 3:
```

```
emc hits:819011745
megaflow hits:329984399
avg. subtable lookups per hit:1.00
miss:48
lost:0
```

```
polling cycles:19458902510676 (97.46%)
processing cycles:506349652804 (2.54%)
avg cycles per packet: 13499.33 (19965252163480/1478980596)
avg processing cycles per packet: 342.36 (506349652804/1478980596)
```

```
pmd thread numa_id 0 core_id 2:
```

```
emc hits:62786086187
megaflow hits:240496054
avg. subtable lookups per hit:1.00
miss:5
lost:0
```

```
polling cycles:10619005555584 (53.82%)
processing cycles:9112873249736 (46.18%)
avg cycles per packet: 311.88 (19731878805320/63267078300)
avg processing cycles per packet: 144.04 (9112873249736/63267078300)
```

```
main thread:
```

```
emc hits:0
```

```
--More--
```

ovs-appctl dpif-netdev/pmd-stats-show

pmd thread numa id 0 core id 3:

Shows performance stats for each pmd thread of the datapath

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-stats-show |more
pmd thread numa_id 0 core_id 3:
  emc hits:819011745
  megaflow hits:329984399
  avg. subtable lookups per hit:1.00
  miss:48
  lost:0
  polling cycles:19458902510676 (97.46%)
  processing cycles:506349652804 (2.54%)
  avg cycles per packet: 13499.33 (19965252163480/1478980596)
  avg processing cycles per packet: 342.36 (506349652804/1478980596)
pmd thread numa_id 0 core_id 2:
  emc hits:62786086187
  megaflow hits:240496054
  avg. subtable lookups per hit:1.00
  miss:5
  lost:0
  polling cycles:10619005555584 (53.82%)
  processing cycles:9112873249736 (46.18%)
  avg cycles per packet: 311.88 (19731878805320/63267078300)
  avg processing cycles per packet: 144.04 (9112873249736/63267078300)
main thread:
  emc hits:0
--More--
```

ovs-appctl dpif-netdev/pmd-stats-show

emc hits : 819011745
Megaflow hits : 329984399

Shows performance stats for each pmd thread of the datapath

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-stats-show |more
pmd thread numa_id 0 core_id 3:
    emc hits:819011745
    megaflow hits:329984399
    avg. subtable lookups per hit:1.00
    miss:48
    lost:0
    polling cycles:19458902510676 (97.46%)
    processing cycles:506349652804 (2.54%)
    avg cycles per packet: 13499.33 (19965252163480/1478980596)
    avg processing cycles per packet: 342.36 (506349652804/1478980596)
pmd thread numa_id 0 core_id 2:
    emc hits:62786086187
    megaflow hits:240496054
    avg. subtable lookups per hit:1.00
    miss:5
    lost:0
    polling cycles:10619005555584 (53.82%)
    processing cycles:9112873249736 (46.18%)
    avg cycles per packet: 311.88 (19731878805320/63267078300)
    avg processing cycles per packet: 144.04 (9112873249736/63267078300)
main thread:
    emc hits:0
--More--
```

ovs-appctl dpif-netdev/pmd-stats-show

polling cycles : 53.82%
processing cycles : 46.18%

Shows performance stats for each pmd thread of the datapath

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-stats-show |more
pmd thread numa_id 0 core_id 3:
    emc hits:819011745
    megaflow hits:329984399
    avg. subtable lookups per hit:1.00
    miss:48
    lost:0
    polling cycles:19458902510676 (97.46%)
    processing cycles:506349652804 (2.54%)
    avg cycles per packet: 13499.33 (19965252163480/1478980596)
    avg processing cycles per packet: 342.36 (506349652804/1478980596)
pmd thread numa_id 0 core_id 2:
    emc hits:62786086187
    megaflow hits:240496054
    avg. subtable lookups per hit:1.00
    miss:5
    lost:0
    polling cycles:10619005555584 (53.82%)
    processing cycles:9112873249736 (46.18%)
    avg cycles per packet: 311.88 (19731878805320/63267078300)
    avg processing cycles per packet: 144.04 (9112873249736/63267078300)
main thread:
    emc hits:0
--More--
```

ovs-appctl dpif-netdev/pmd-stats-show

avg. subtable lookups per hit :
1.00

Shows performance stats for each pmd thread of the datapath

Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-stats-show |more
pmd thread numa_id 0 core_id 3:
  emc hits:0
  megaflow hits:0
  avg. subtable lookups per hit:0.00
  miss:0
  lost:0
  polling cycles:72481927540 (100.00%)
  processing cycles:0 (0.00%)
pmd thread numa_id 0 core_id 2:
  emc hits:35667360
  megaflow hits:3825812
  avg. subtable lookups per hit:1.00
  miss:1
  lost:0
  polling cycles:58549367676 (80.61%)
  processing cycles:14085077144 (19.39%)
  avg cycles per packet: 1676.73 (72634444820/43318985)
  avg processing cycles per packet: 325.15 (14085077144/43318985)
main thread:
  emc hits:0
  megaflow hits:0
  avg. subtable lookups per hit:0.00
  miss:0
  lost:0
  polling cycles:742880 (100.00%)
```

ovs-appctl dpif-netdev/pmd-stats-show

Debugging

ovs-appctl dpif-netdev/pmd-rxq-show

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-rxq-show
pmd thread numa_id 0 core_id 3:
    isolated : true
    port: dpdkvhostuser0    queue-id: 0
pmd thread numa_id 0 core_id 2:
    isolated : true
    port: dpdkvhostuser5    queue-id: 0
pmd thread numa_id 0 core_id 1:
    isolated : false
    port: dpdkvhostuser1    queue-id: 0
    port: dpdkvhostuser2    queue-id: 0
    port: dpdkvhostuser3    queue-id: 0
    port: dpdkvhostuser4    queue-id: 0
pmd thread numa_id 0 core_id 0:
    isolated : false
    port: dpdk0             queue-id: 0
    port: dpdk1             queue-id: 0
    port: dpdkvhostuser6    queue-id: 0
    port: dpdkvhostuser7    queue-id: 0
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$
```

Shows dp ports/rxq assignment to PMD threads

Debugging

Affinitizing datapath Rx queues for uniform load distribution.

```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-rxq-show
pmd thread numa_id 0 core_id 3:
    isolated : true
    port: dpdkvhostuser0    queue-id: 0
pmd thread numa_id 0 core_id 2:
    isolated : true
    port: dpdkvhostuser5    queue-id: 0
pmd thread numa_id 0 core_id 1:
    isolated : false
    port: dpdkvhostuser1    queue-id: 0
    port: dpdkvhostuser2    queue-id: 0
    port: dpdkvhostuser3    queue-id: 0
    port: dpdkvhostuser4    queue-id: 0
pmd thread numa_id 0 core_id 0:
    isolated : false
    port: dpdk0             queue-id: 0
    port: dpdk1             queue-id: 0
    port: dpdkvhostuser6    queue-id: 0
    port: dpdkvhostuser7    queue-id: 0
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$
```

**ovs-vsctl set interface dpdkvhostuser2
other_config:pmd-rxq-affinity="0:2"**

Debugging

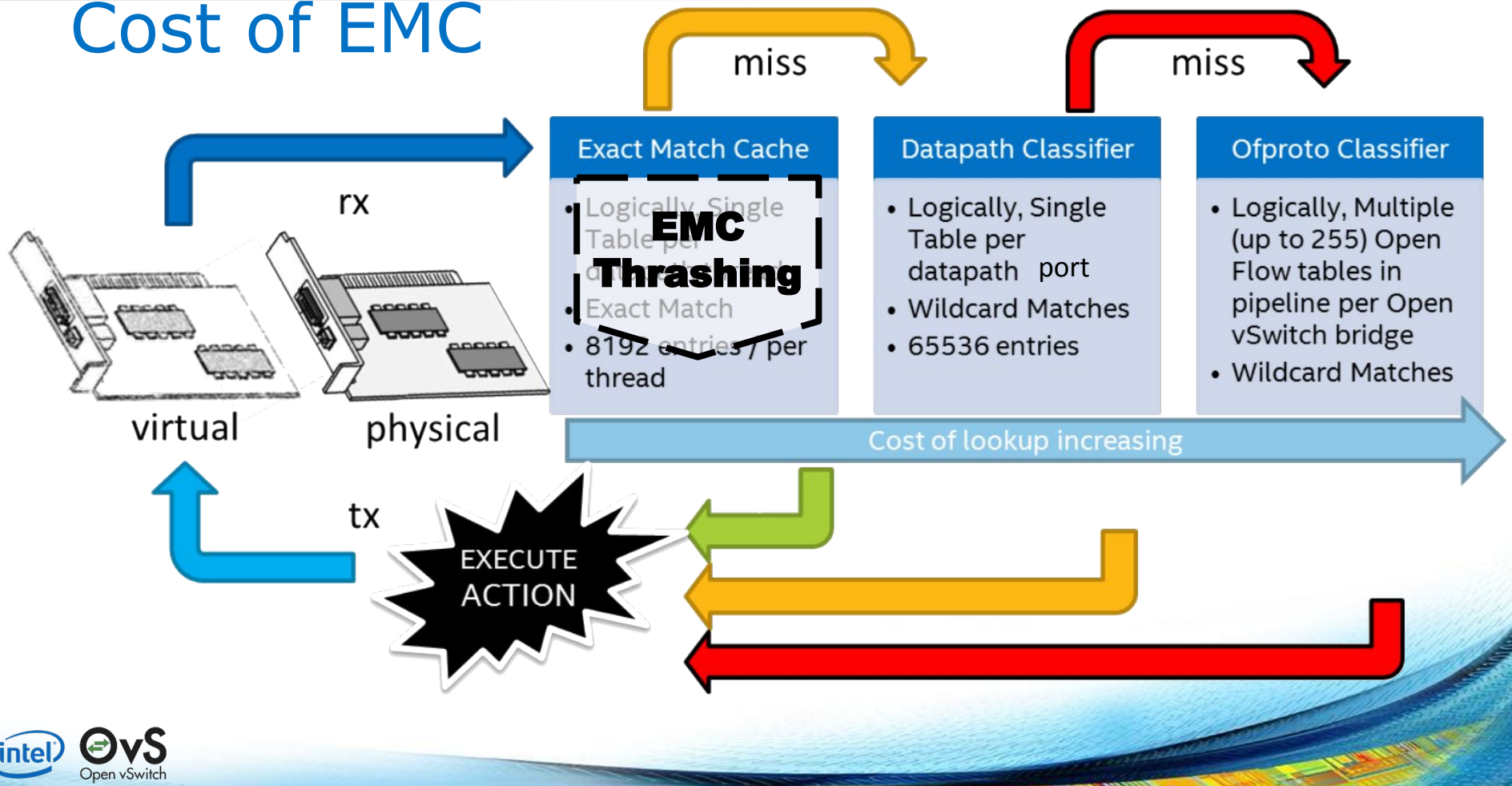
ovs-appctl dpif-netdev/pmd-rxq-show

```
sugeschch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-rxq-show
pmd thread numa_id 0 core_id 3:
    isolated : true
    port: dpdkvhostuser0    queue-id: 0
    port: dpdkvhostuser3    queue-id: 0
    port: dpdkvhostuser6    queue-id: 0
pmd thread numa_id 0 core_id 2:
    isolated : true
    port: dpdkvhostuser4    queue-id: 0
    port: dpdkvhostuser5    queue-id: 0
    port: dpdkvhostuser7    queue-id: 0
pmd thread numa_id 0 core_id 1:
    isolated : false
    port: dpdk0             queue-id: 0
    port: dpdkvhostuser2    queue-id: 0
pmd thread numa_id 0 core_id 0:
    isolated : false
    port: dpdk1             queue-id: 0
    port: dpdkvhostuser1    queue-id: 0
sugeschch@silpixa00389816:~/repo/ovs_master/utilities$
```

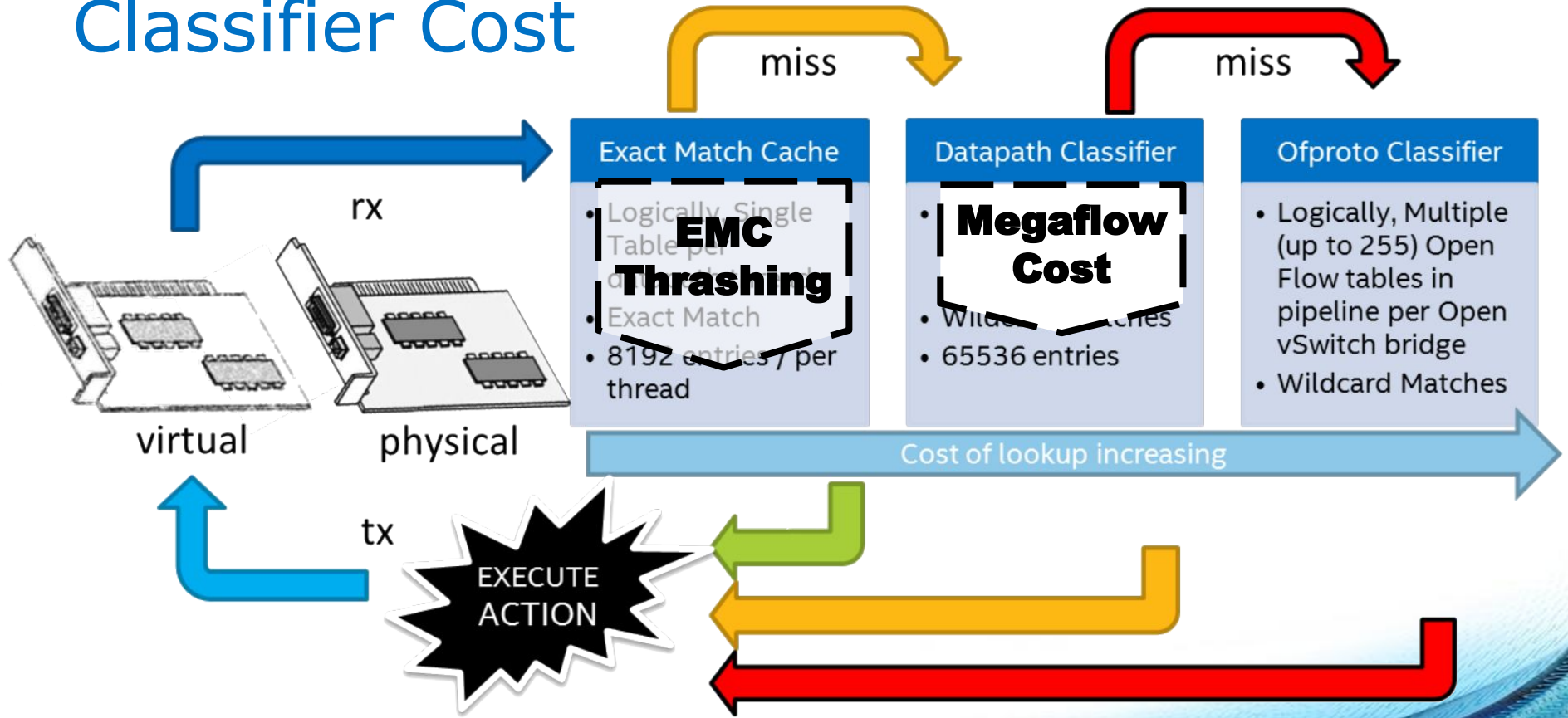

OVS-DPDK Hotspots



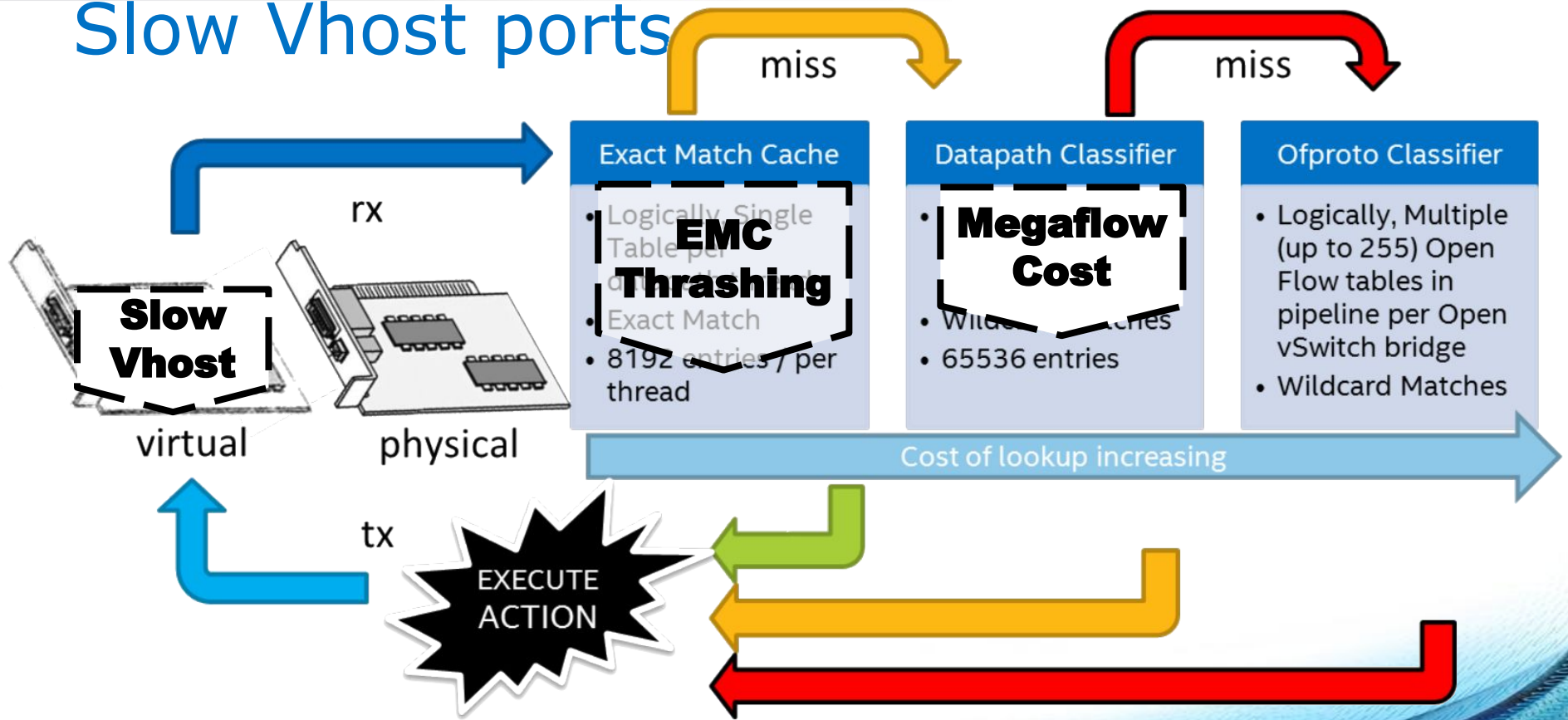
Cost of EMC



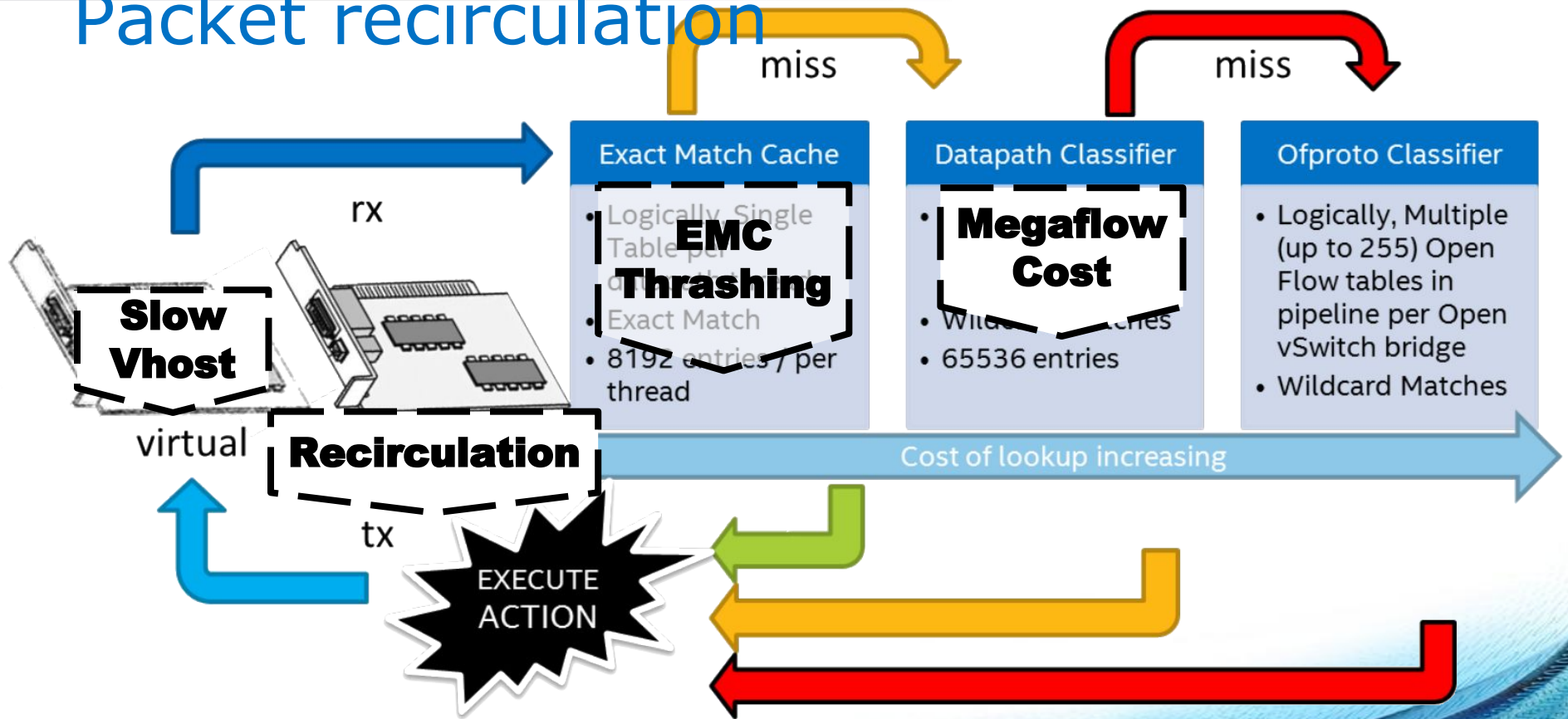
Classifier Cost



Slow Vhost ports



Packet recirculation

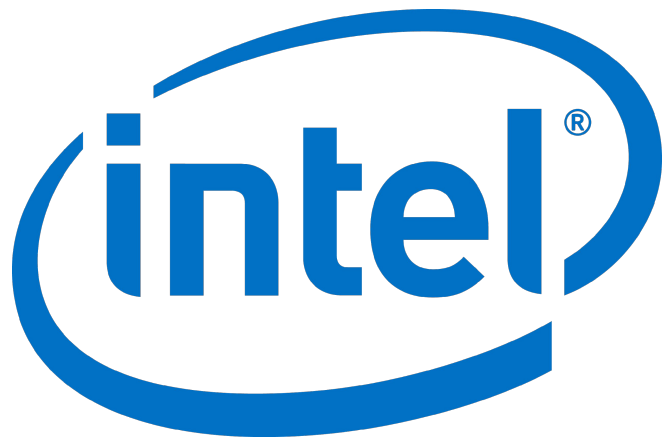


Summary

- OVS-DPDK can offer SR-IOV comparable performance

<http://www.intel.com/content/dam/www/public/us/en/documents/technology-briefs/sr-iov-nfv-tech-brief.pdf>

- OVS-DPDK offers
 - Fully fledged vswitch features.
 - Good Performance
 - Production Quality
- Understanding OVS commands helps to get better performance.
- Latest OVS-DPDK always offers better performance.



Q&A

Legal Disclaimers

- No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.
- Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.
- This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.
- The products and services described may contain defects or errors known as errata which may cause deviations from published specifications. Current characterized errata are available on request.
- Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or by visiting [**www.intel.com/design/literature.htm**](http://www.intel.com/design/literature.htm).
- Intel, the Intel logo, Xeon, VTune are trademarks of Intel Corporation in the U.S. and/or other countries.
- *Other names and brands may be claimed as the property of others
- Copyright © 2017 Intel Corporation. All rights reserved.

