



# **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

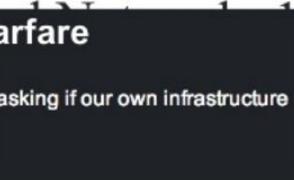


# Stateful vs Stateless Firewalls

## Software Must Be Easier, Right?

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



Matt L. 04 April 2012  **White House Hacked In Cyber Attack That Used Spear-Phishing To Crack Unclassified Network**   
111 111 100  
arfare  
asking if our own infrastructure

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

NSA Audit

**Experts: Google's 'Aurora' Offers Rep.**  **hackers still at it years later**

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

ring 'Major'

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

# Launch an instance with Security Groups

Launch Instance

Select the security groups to launch the instance in.

Allocated 5

Name	Description	Actions
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

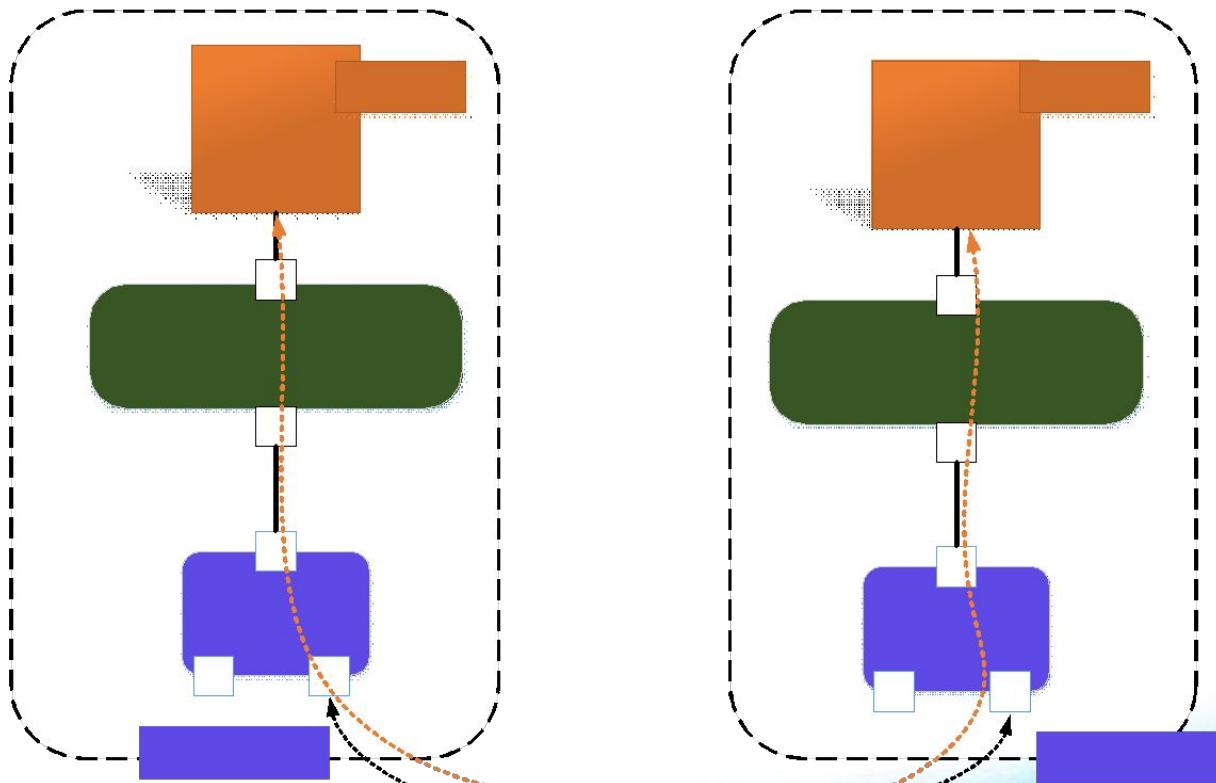
Click here for filters.

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

Cancel Back Next Launch Instance



# Verifying VM connectivity with SG enabled

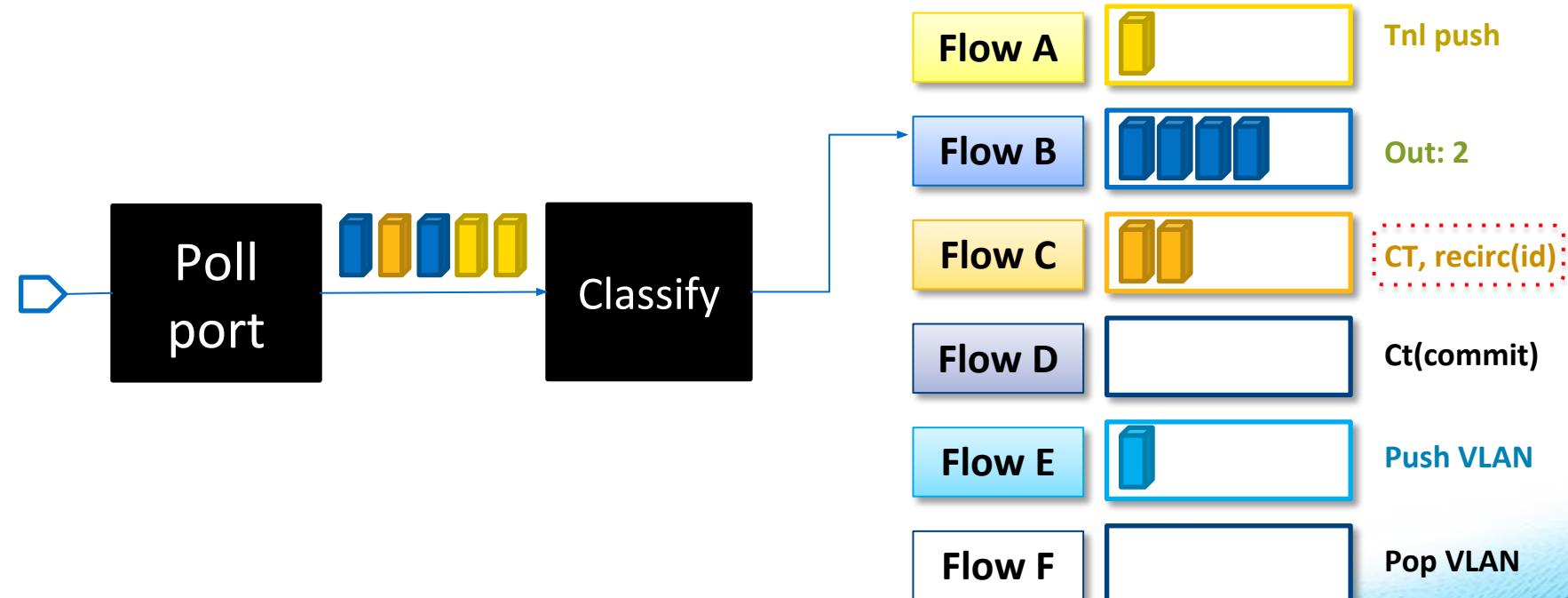


Read max 32 packets

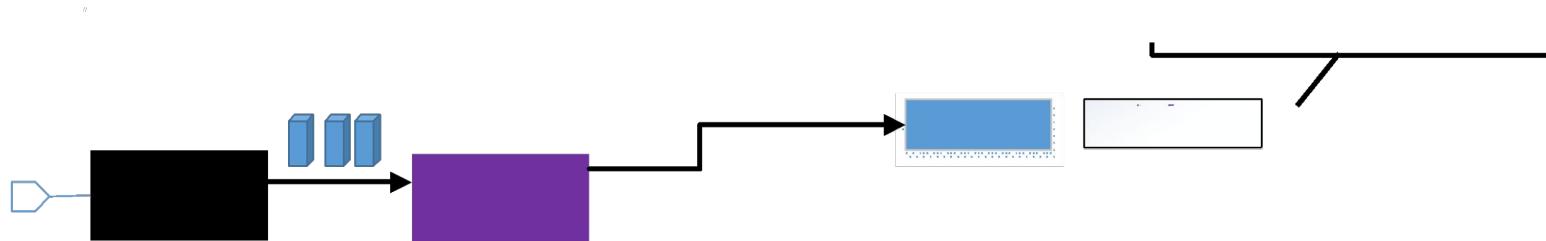
classification

Batch Packets by flows

Execute Actions

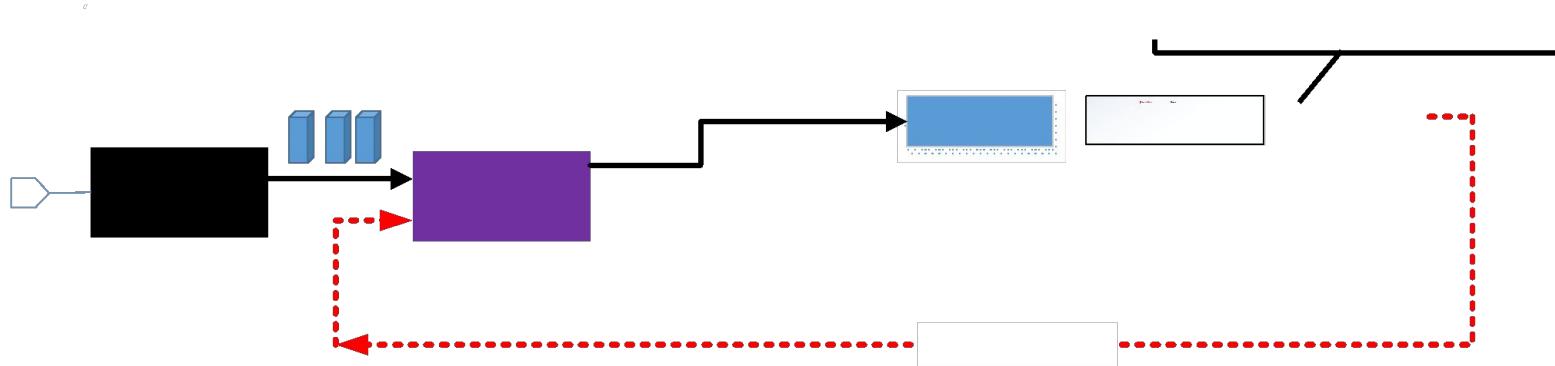


# Recirculation explained



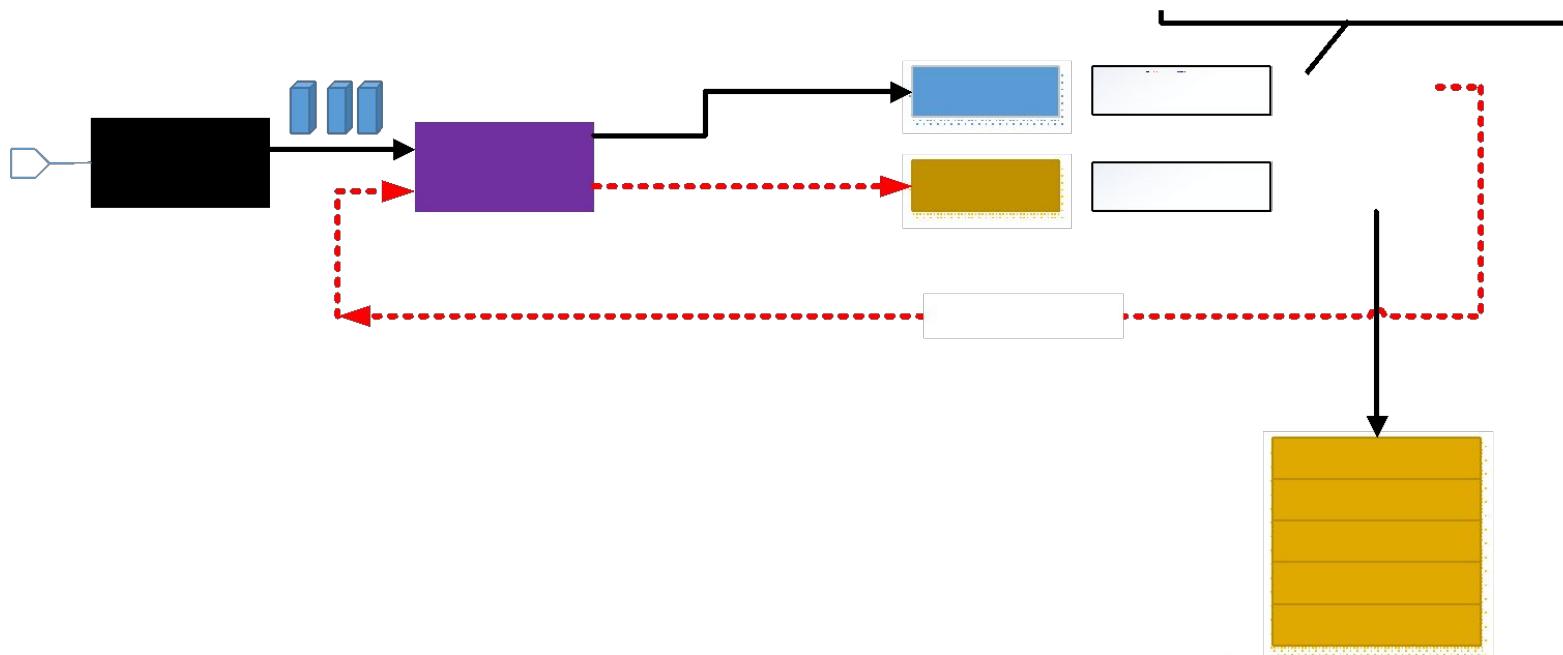
Conntrack sets the connection bits in packet meta data

# Recirculation explained



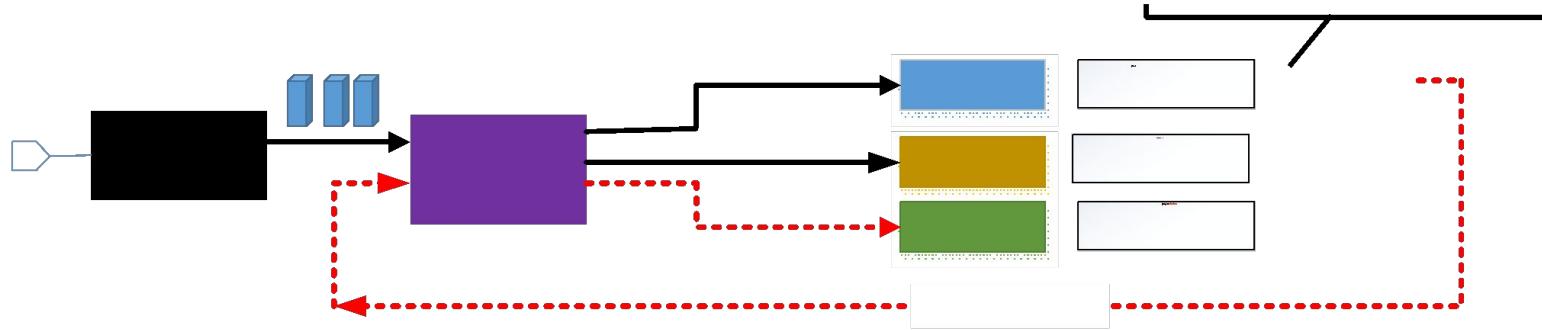
**'recirc' flag is set and so the packets are sent back to OvS flow table**

# 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

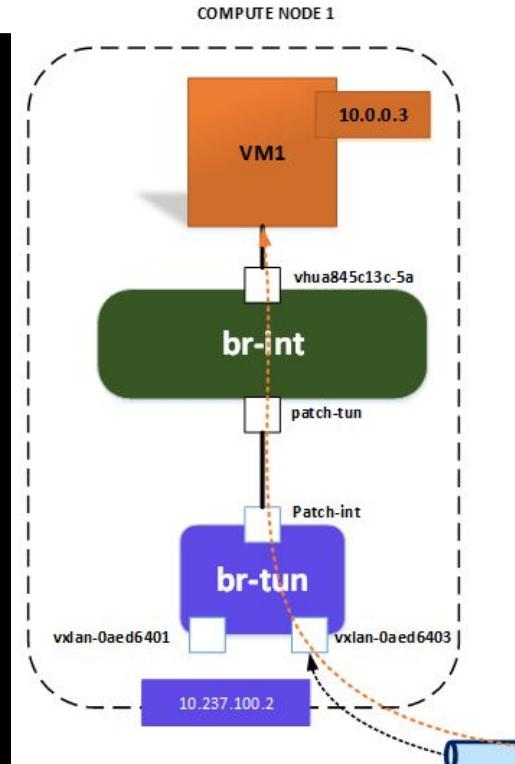
# Recirculation explained



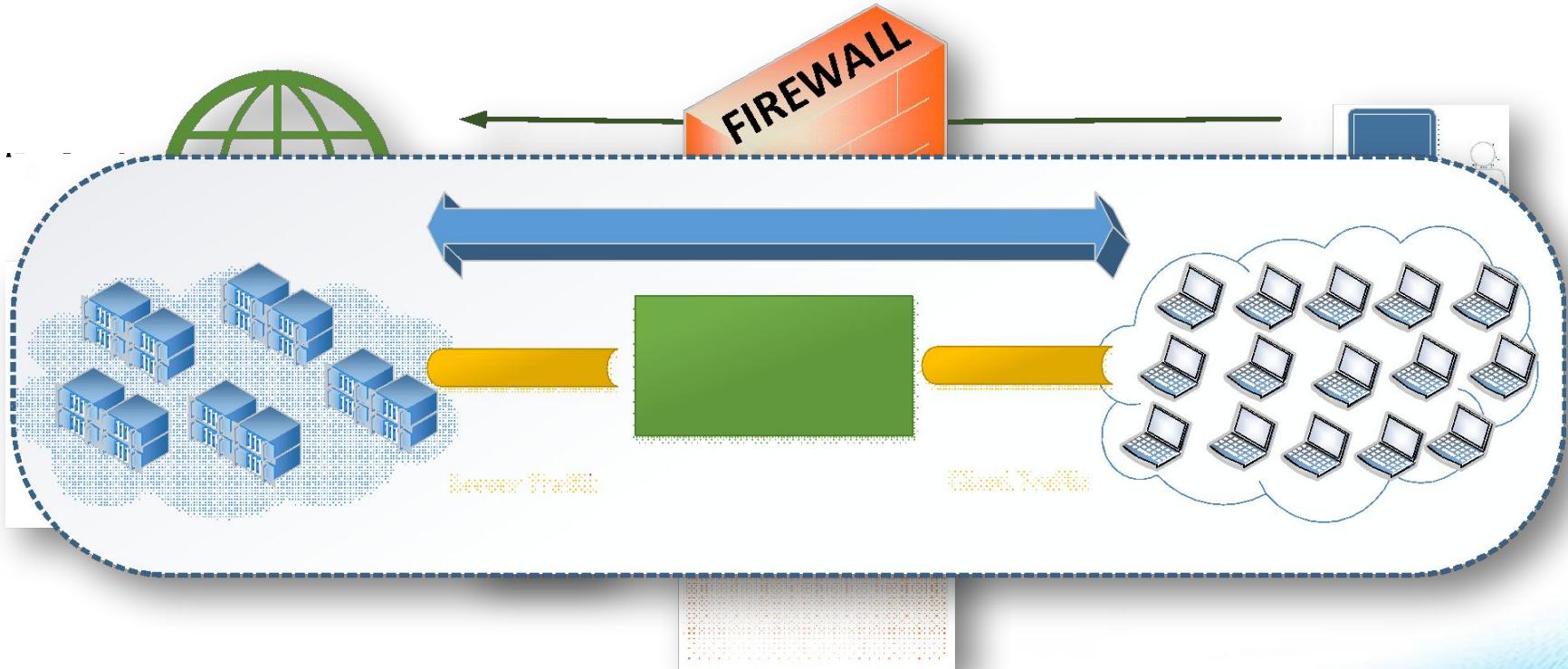
Once the connection is established, recirculated packets will match  
**'Flow D'** and get forwarded on output port 2

# Compute Node 1 (egress)

## Flow Entries



# Testing Stateful Firewall, Setup Details



# Key Performance Metrics

Connection	Transaction
Concurrent connections	Connections per second
Transactions per second	Throughput
Protocol latency	TTFB(Time to first byte)
TTLB(Time to last byte)	

## Desirable Connection

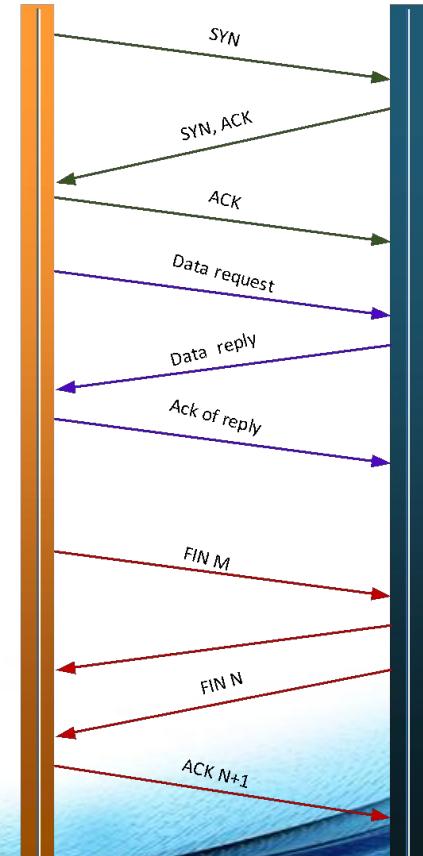
1. TCP connection establishment (SYN, SYN-ACK, ACK)
2. layer 7 transaction(Request, Response)
3. TCP teardown(FIN, ACK).

## L4 performance with limited L7 transactions

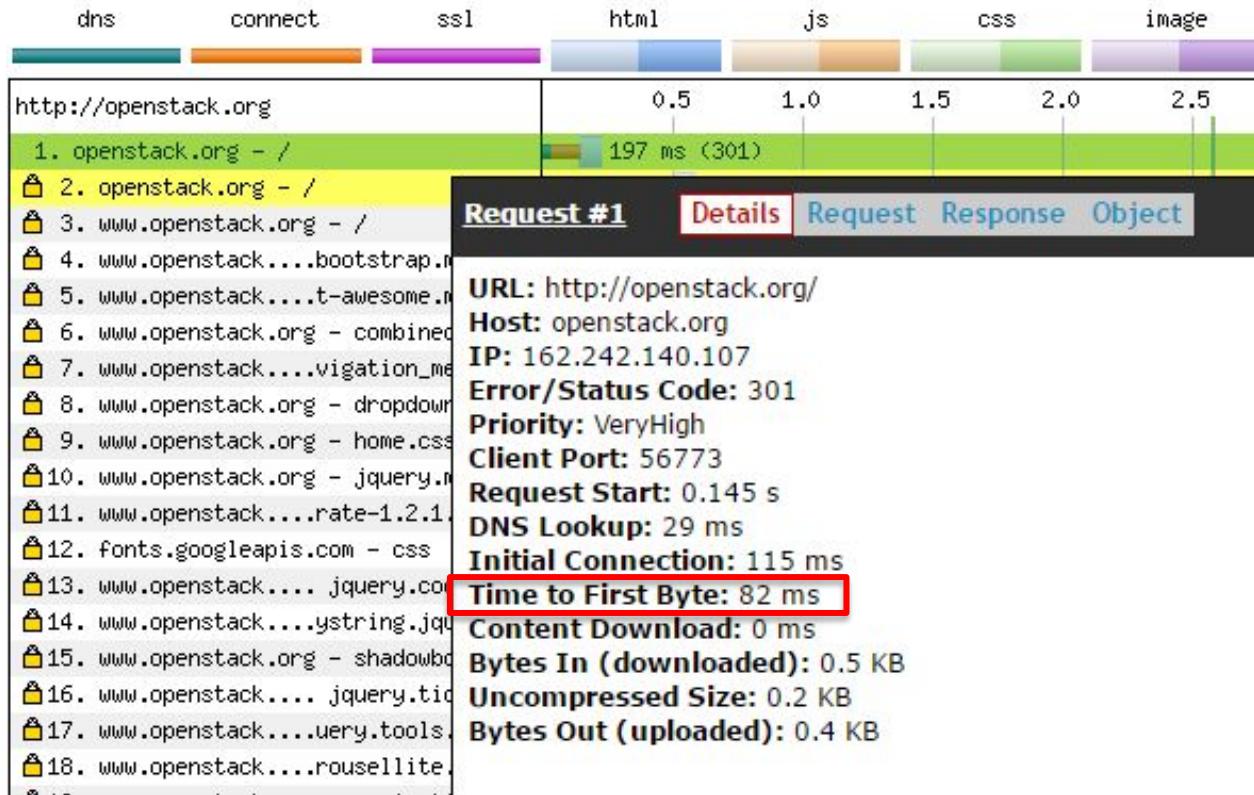
1. TCP connection establishment (SYN, SYN-ACK, ACK)
2. **Partial/incomplete** layer 7 transaction(only Request)
3. TCP teardown(FIN, ACK).

## Forceful connection teardowns

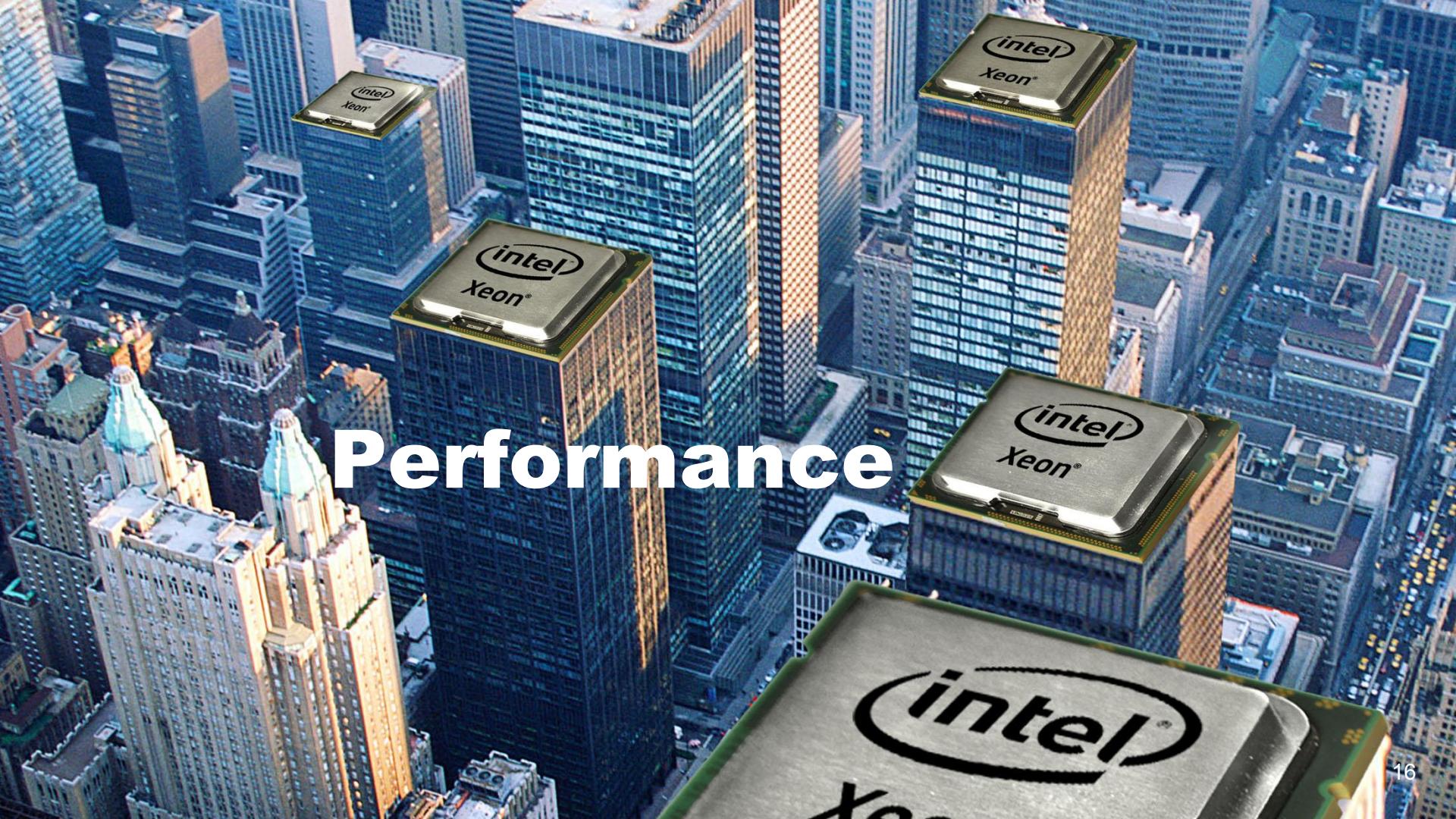
1. TCP connection establishment (SYN, SYN-ACK, ACK)
2. Partial/incomplete layer 7 transaction(Request)
3. **Forced TCP teardown(RST).**



# Waterfall view of response time

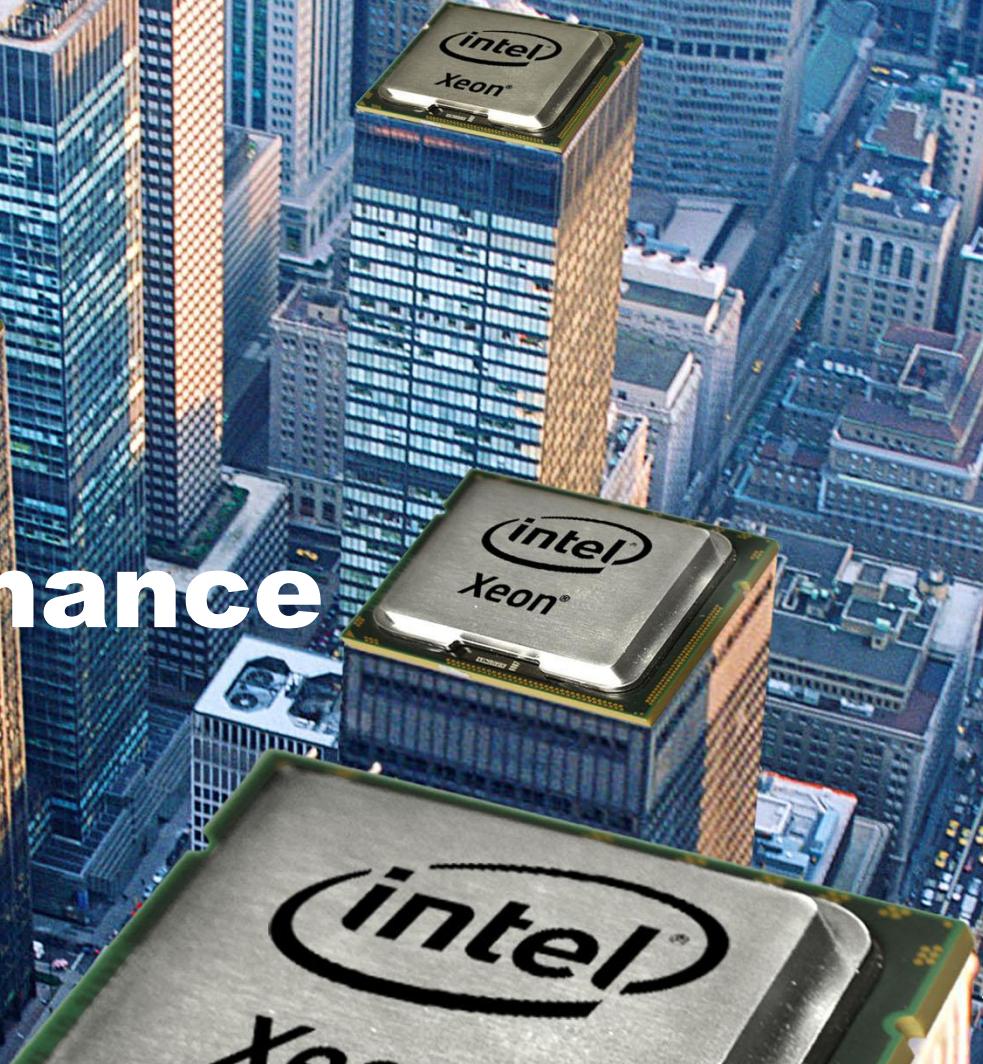


**Results for  
openstack.org website  
retrieved using  
Webpagetest.org**



The background is an aerial photograph of a city skyline, likely New York City, featuring many tall skyscrapers.

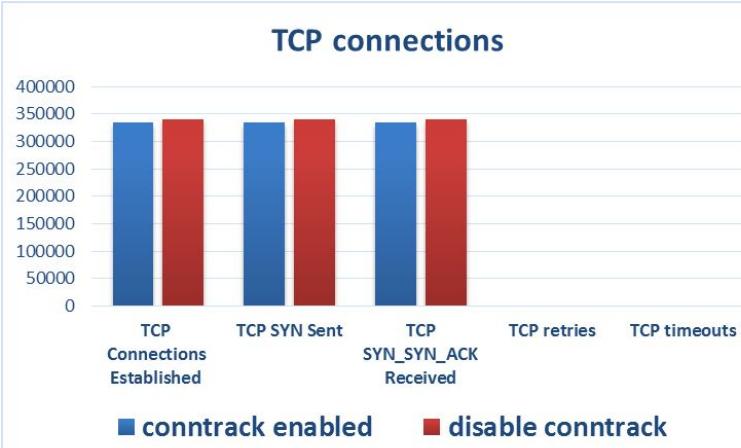
# Performance



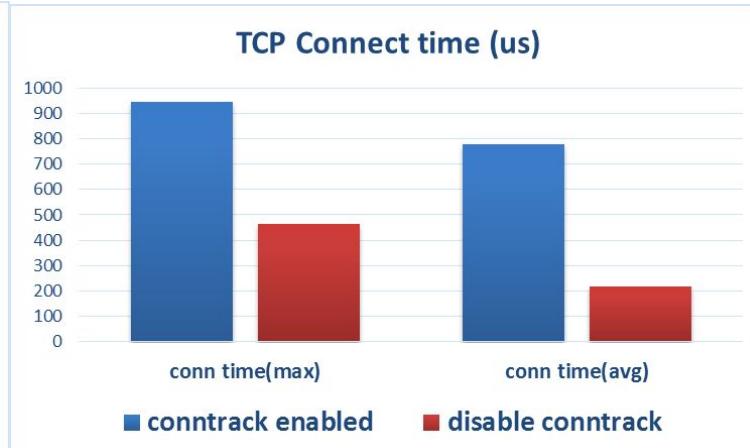
# HTTP client/server – Throughput test



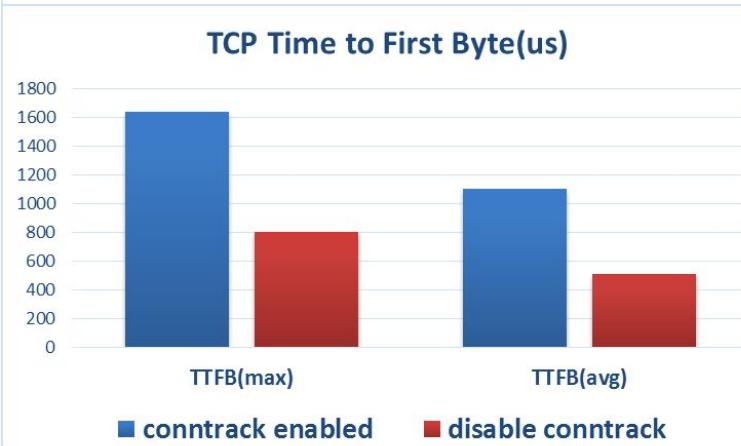
The Higher  
the better



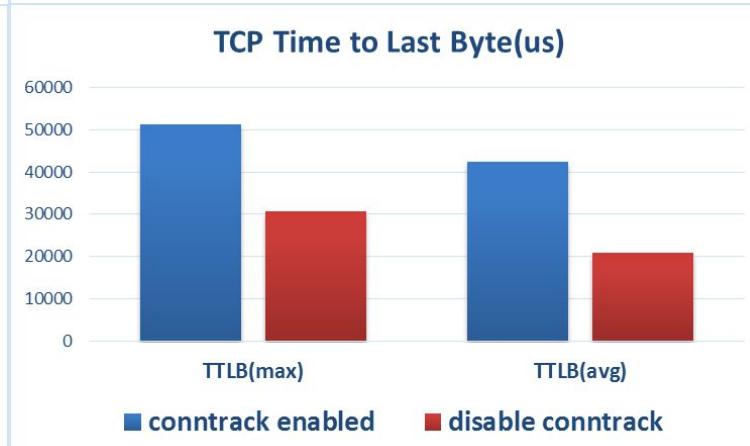
The Lower  
the better



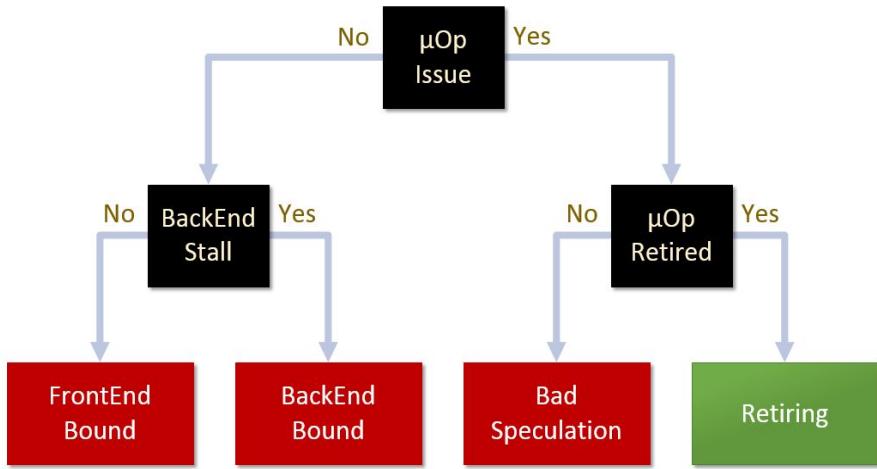
The Lower  
the better



The Lower  
the better



# Analyzing Bottlenecks – Throughput test



Category

	Acceptable Values for HPC applications ↑ The Higher the better ↓ The Lower the better	OvS-DPDK conntrack (http client/server throughput test)
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% ↑

Elapsed Time: 120.011s

③ CPU Time ⑦: 122.500s

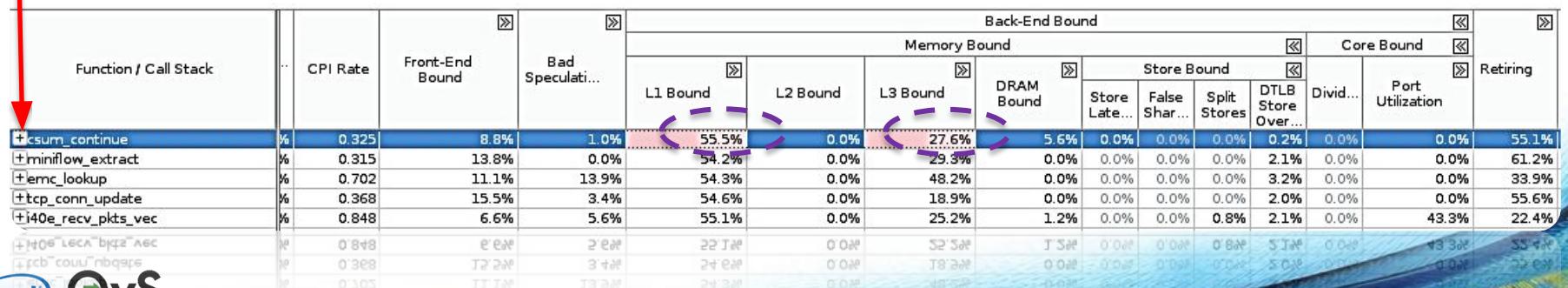
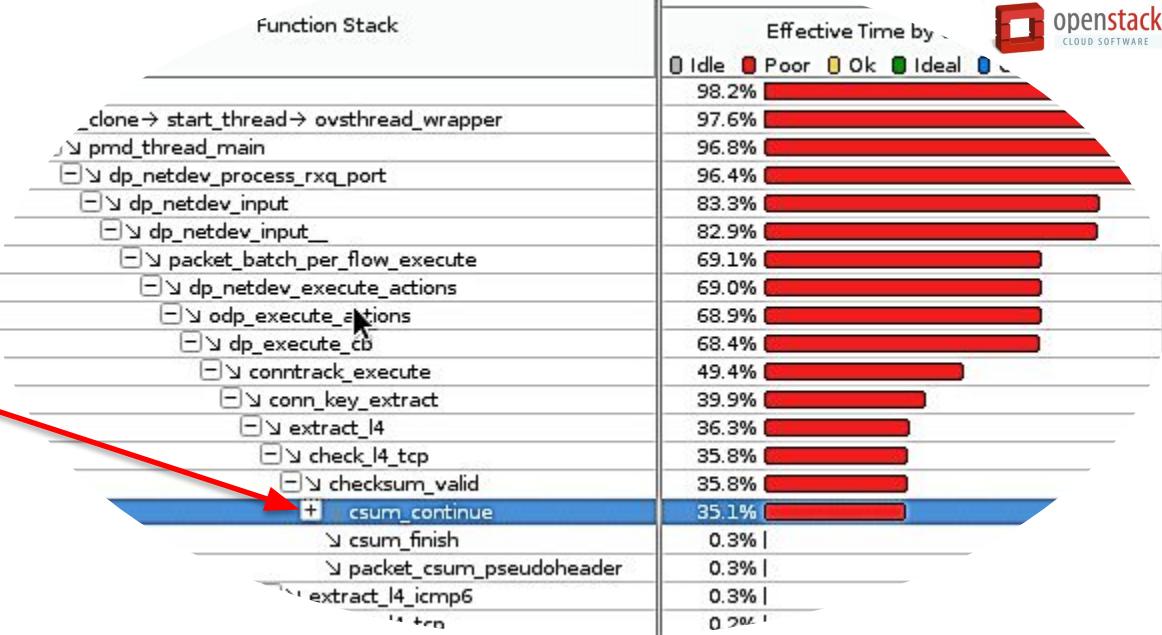
Total Thread Count: 14

Paused Time<sup>②</sup>:

## Top Hotspots

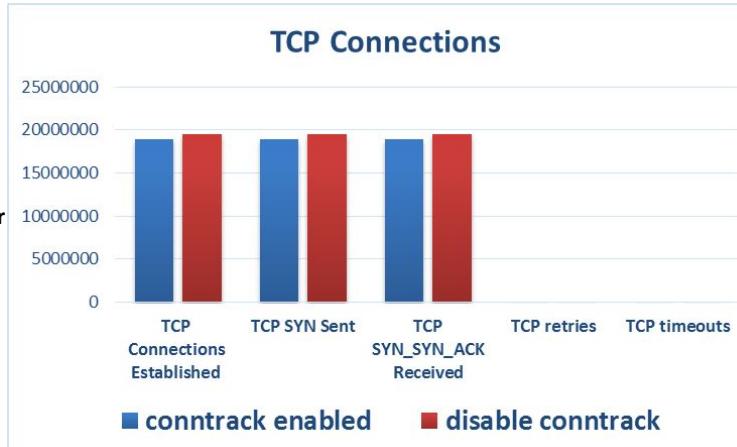
This section lists the most active functions in your application.

Function	Module	CPU Time
<u>csum_continue</u>	ovs-vswitchd	42.831s
<u>i40e_recv_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

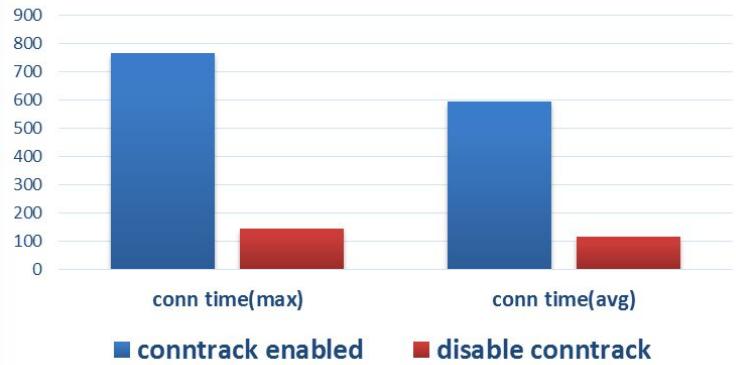


# HTTP client/server – 65k connections/sec

  
The Higher  
the better

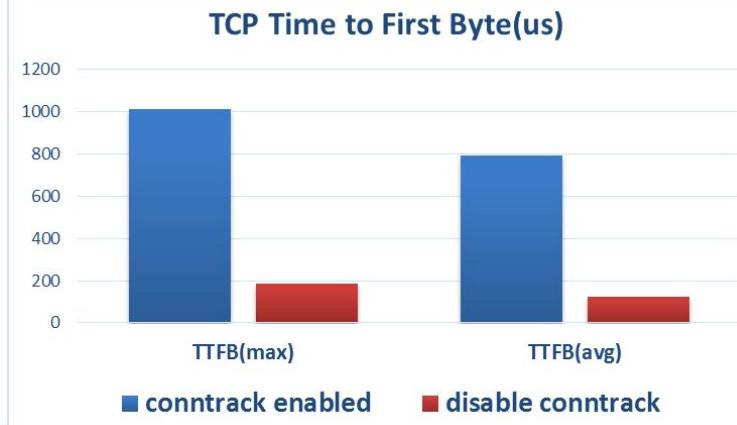


**TCP connect time(us)**

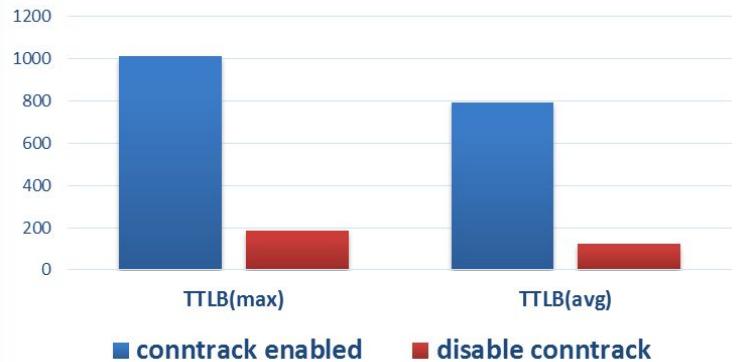


  
The Lower  
the better

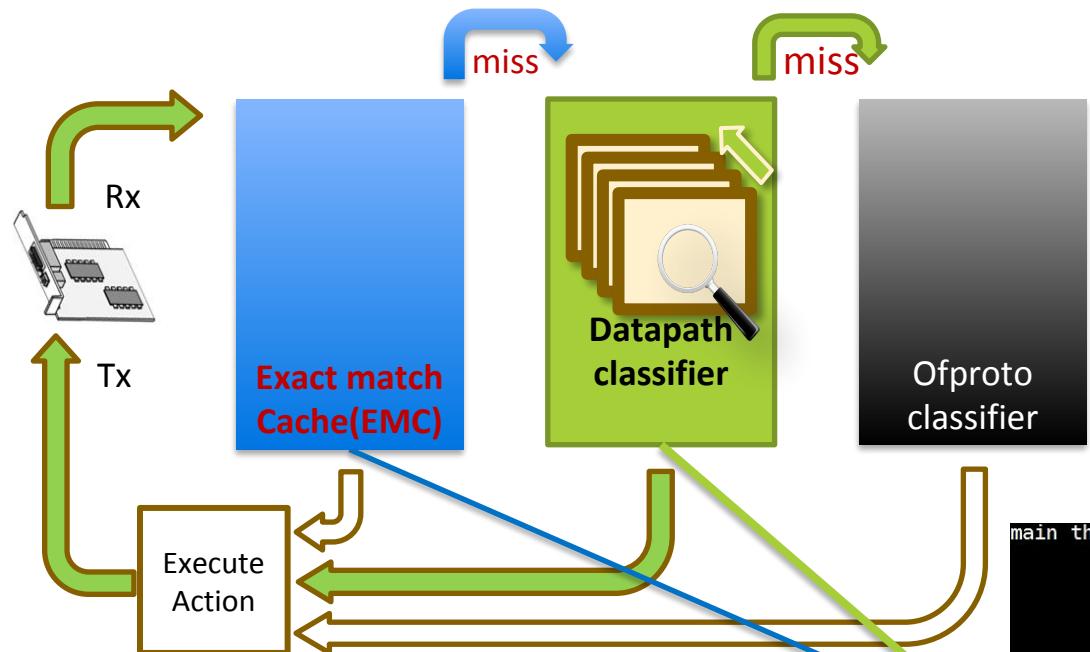
  
The Lower  
the better



**TCP Time to Last Byte(us)**



  
The Lower  
the better



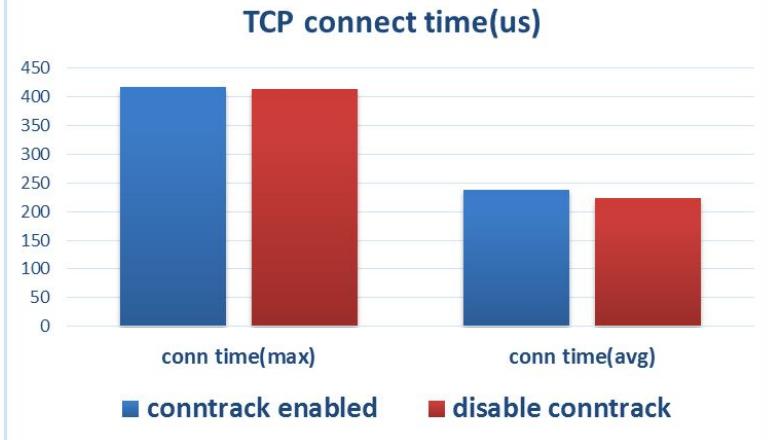
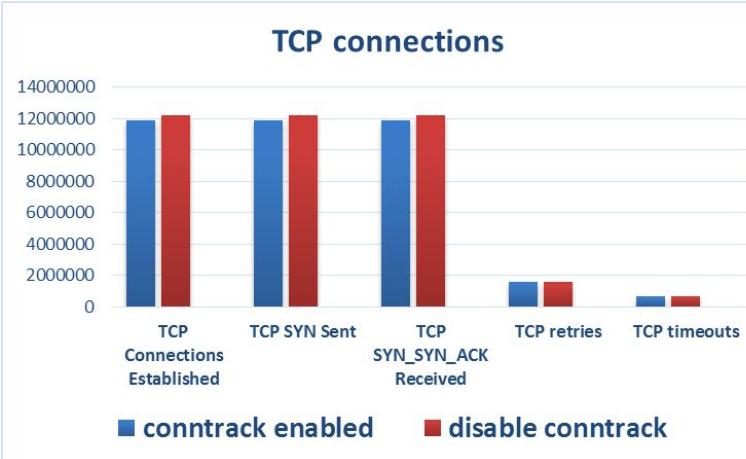
```

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%)
  end 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)

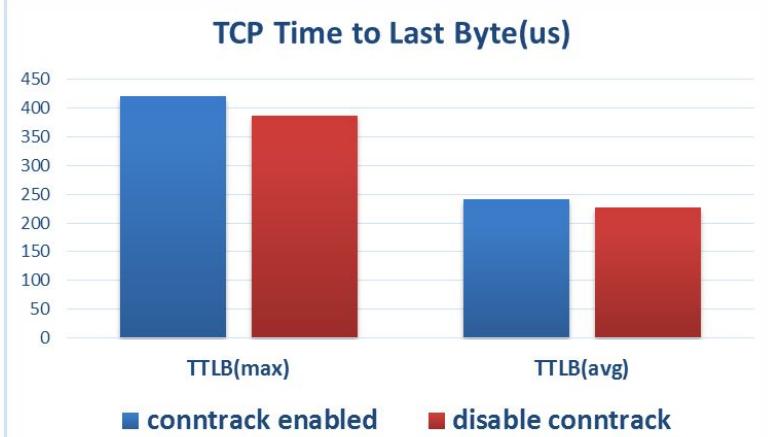
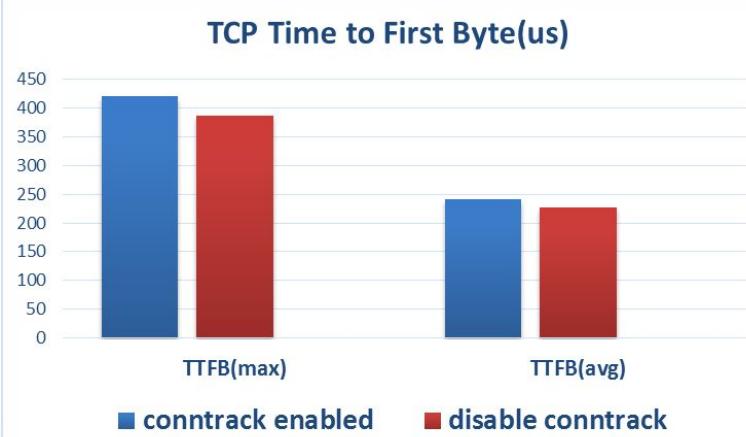
```

# HTTP client/server – 800k concurrent connections

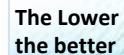
  
The Higher  
the better



  
The Lower  
the better



  
The Lower  
the better

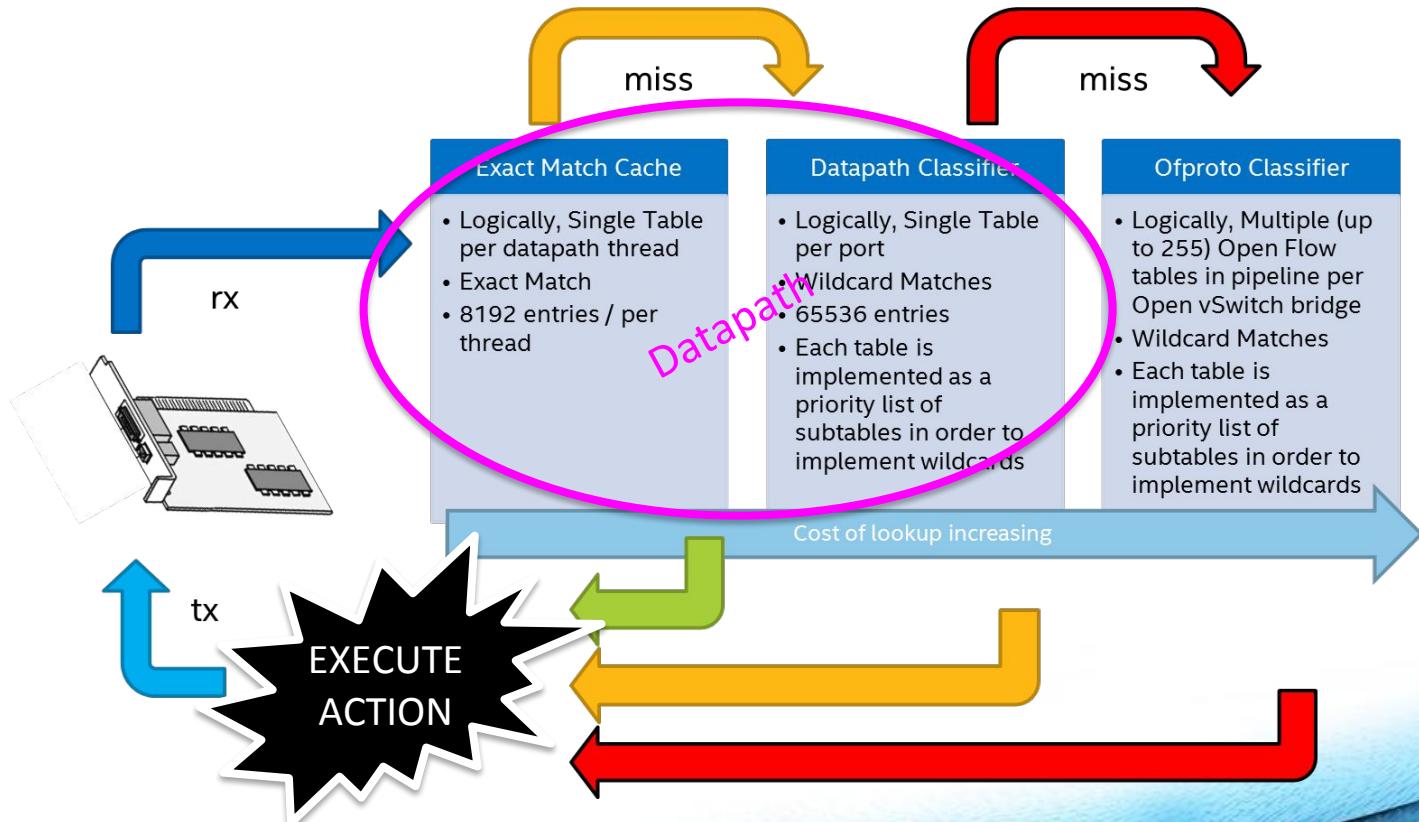


The Lower  
the better

# 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

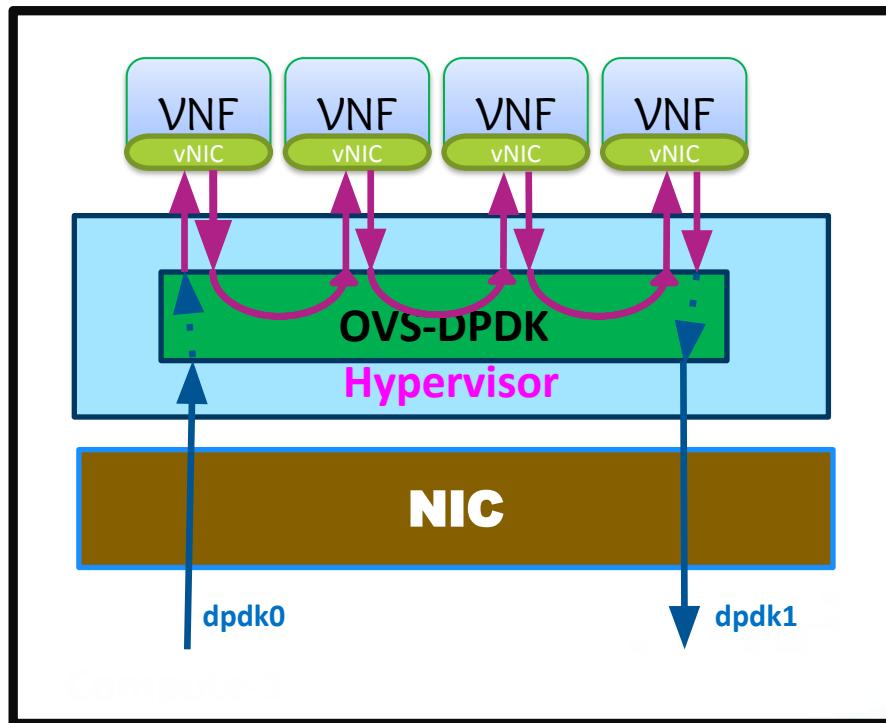
# OVS-DPDK Table Hierarchy



# OvS-DPDK Troubleshooting



# East-West traffic test



# Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-ports-desc br0 |more
0FPST_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
--More--
```

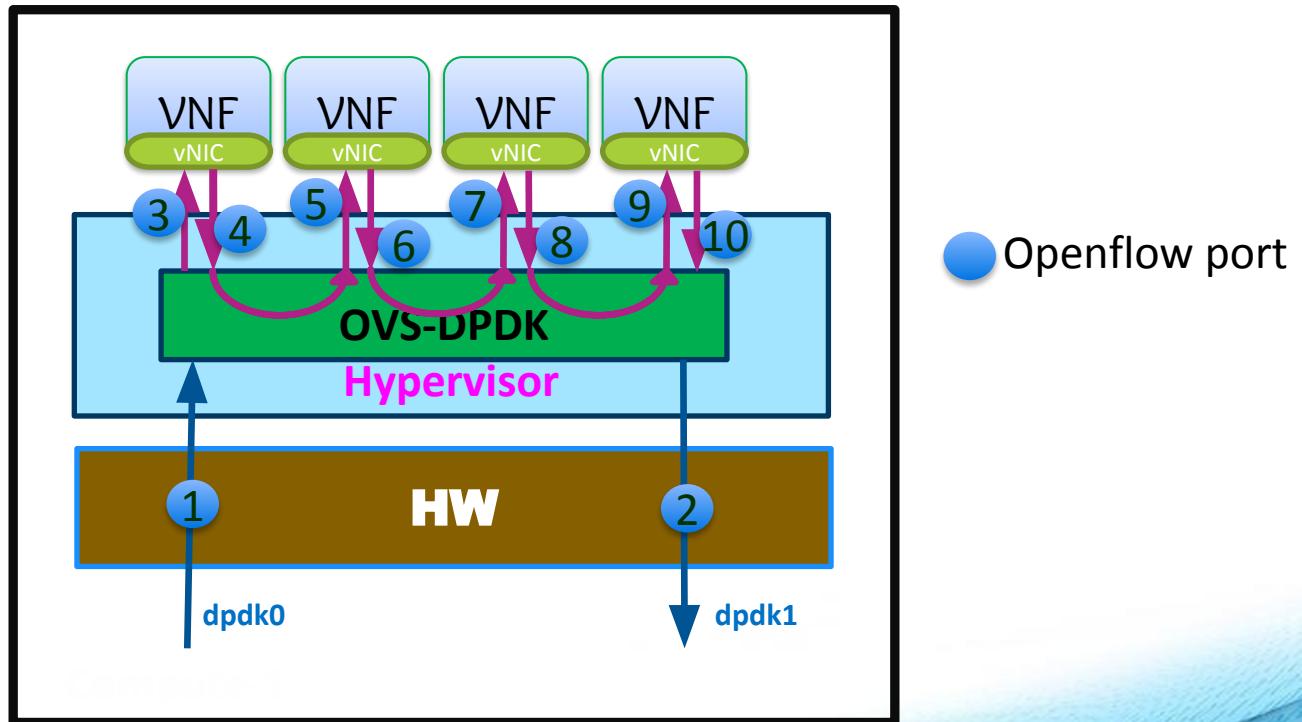
Detailed information about the openflow ports in a switch.

# Debugging

```
sugeshch@silpixa00389816:~/repo/ovs_master$ sudo ./utilities/ovs-ofctl dump-ports-desc br0 |more
0OFPT_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
--More--
```

Detailed information about the openflow ports in a switch.

# Debugging



# 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=o
utput:1
  cookie=0x0, duration=177.360s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=4 actions=outp
ut:5
  cookie=0x0, duration=177.320s, table=0, n_packets=41, n_bytes=7659, idle_age=3, in_port=5 actions=ou
tput:4
  cookie=0x0, duration=177.272s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=6 actions=outp
ut:7
  cookie=0x0, duration=177.228s, table=0, n_packets=40, n_bytes=7152, idle_age=0, in_port=7 actions=ou
tput:6
  cookie=0x0, duration=177.176s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=8 actions=outp
ut:9
  cookie=0x0, duration=177.136s, table=0, n_packets=37, n_bytes=6394, idle_age=14, in_port=9 actions=o
utput:8
  cookie=0x0, duration=177.068s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=10 actions=outp
ut:2
  cookie=0x0, duration=177.024s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=2 actions=outp
ut:10
sugeshch@silpixa00389816:~/repo/ovs_master$
```

ovs-ofctl dump-flows <br>

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=o
utput:1
  cookie=0x0, duration=177.360s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=4 actions=outp
ut:5
  cookie=0x0, duration=177.320s, table=0, n_packets=41, n_bytes=7659, idle_age=3, in_port=5 actions=ou
tput:4
  cookie=0x0, duration=177.272s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=6 actions=outp
ut:7
  cookie=0x0, duration=177.228s, table=0, n_packets=40, n_bytes=7152, idle_age=0, in_port=7 actions=ou
tput:6
  cookie=0x0, duration=177.176s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=8 actions=outp
ut:9
  cookie=0x0, duration=177.136s, table=0, n_packets=37, n_bytes=6394, idle_age=14, in_port=9 actions=o
utput:8
  cookie=0x0, duration=177.068s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=10 actions=outp
ut:2
  cookie=0x0, duration=177.024s, table=0, n_packets=0, n_bytes=0, idle_age=177, in_port=2 actions=outp
ut:10
sugeshch@silpixa00389816:~/repo/ovs_master$
```

ovs-ofctl dump-flows <br>

List all openflow flow entries in switch tables.

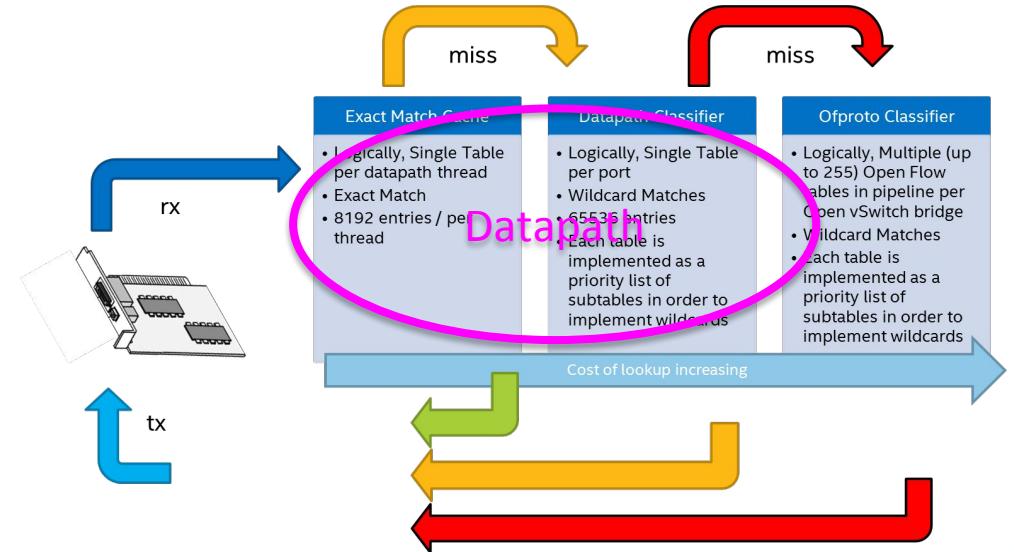
# Debugging

```
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$
```

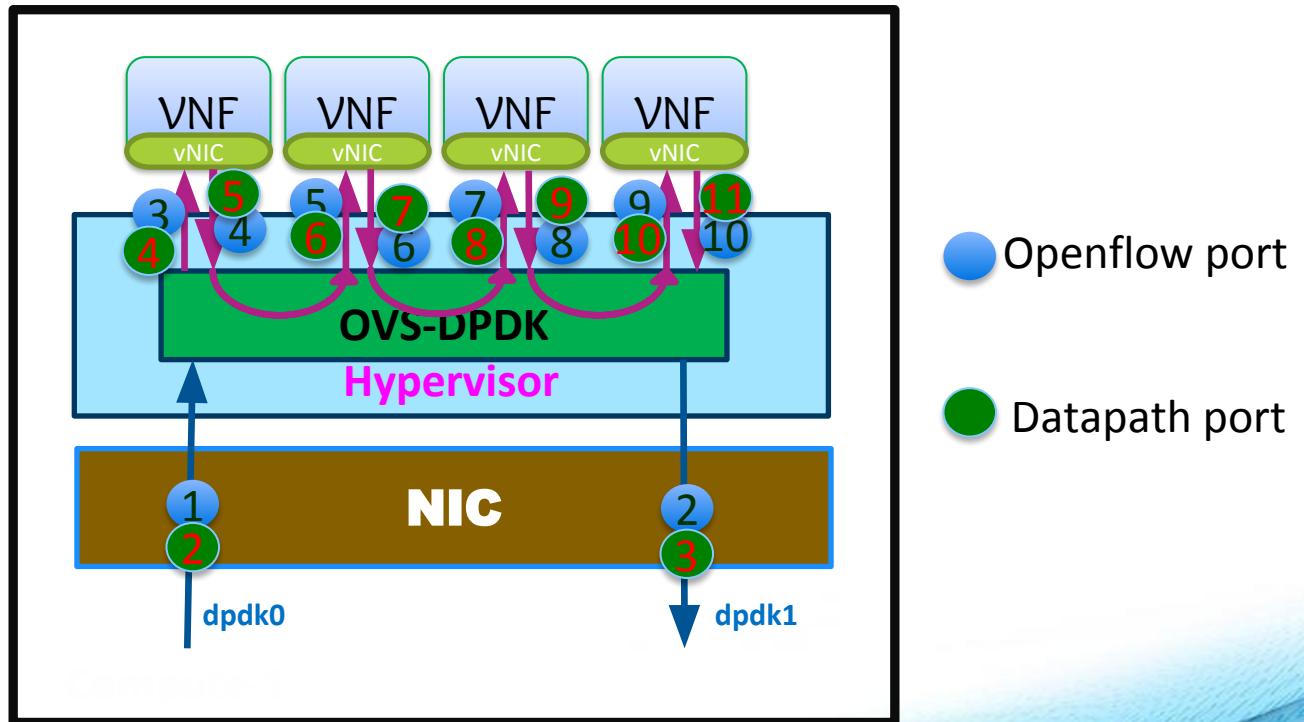
ovs-ofctl dump-flows <br>

List all openflow flow entries in switch tables.

# Datapath Debugging



# Debugging



# Debugging

```
netdev@ovs-netdev: hit:17543389635 missed:188
    br0:
        br0 65534/1: (tap,
                        ovs-appctl dpif/show
                        dpdk0 1/2: (apak: 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, mtu=1500, requested_rx_queues=1, requested_tx_queues=1)
--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--
```

**ovs-appctl dpctl/show -s**

RX packets:13798789152 errors:0 dropped:1803077122 overruns:? frame:?

TX packets:107 errors:0 dropped:0 aborted:? carrier:?  
collisions:?

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--
```

**ovs-appctl dpctl/show -s**

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=n
o), 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=n
o), 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=177.272s, 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$
```

Pkt dl\_src = 00:00:00:00:00:0f

**ovs-ofctl dump-flows <br>**

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=00:00:00:00:0f, dst=00:00:00:00:01),eth_type(0x0800),ipv4(frag=n
o), 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:51650732040,
used:0.001s, actions:3
recirc_id(0),in_port(2),eth(src=00:00:00:00:0f, dst=00:00:00:00:01),eth_type(0x0800),ipv4(frag=n
o), packets:1205917247, bytes:72355034820, used:0.000s, actions:4
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$
```

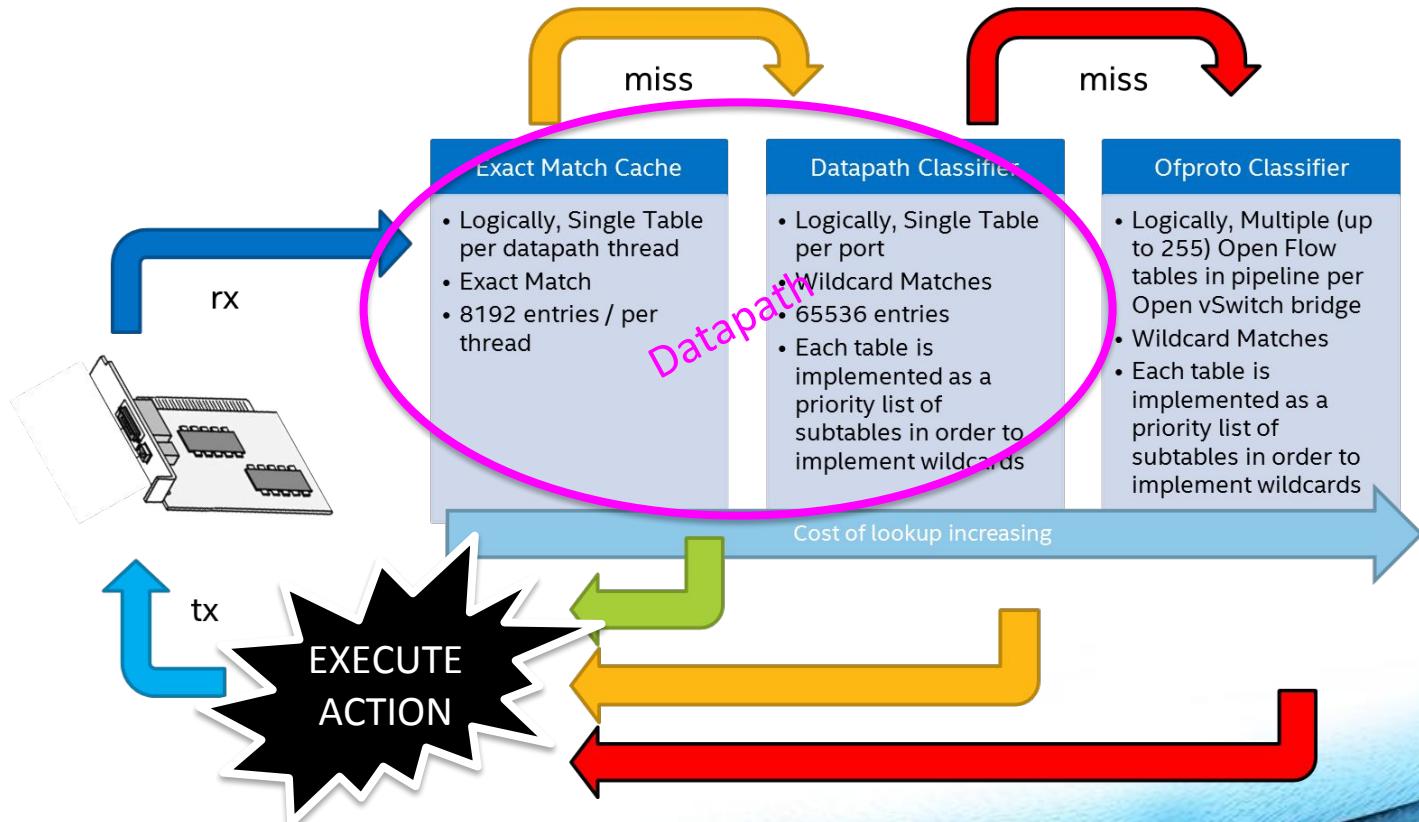


Display all active flow entries in a datapath



# 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:19453002510076 (97.46%)
    processing cycles:50634952804 (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:19453002510076 (97.46%)
    processing cycles:50634952804 (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:19453002510076 (97.46%)
    processing cycles:50634952804 (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--
```

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:19453002510076 (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--
```

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$
```

# 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: dpdkvhostuser3    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

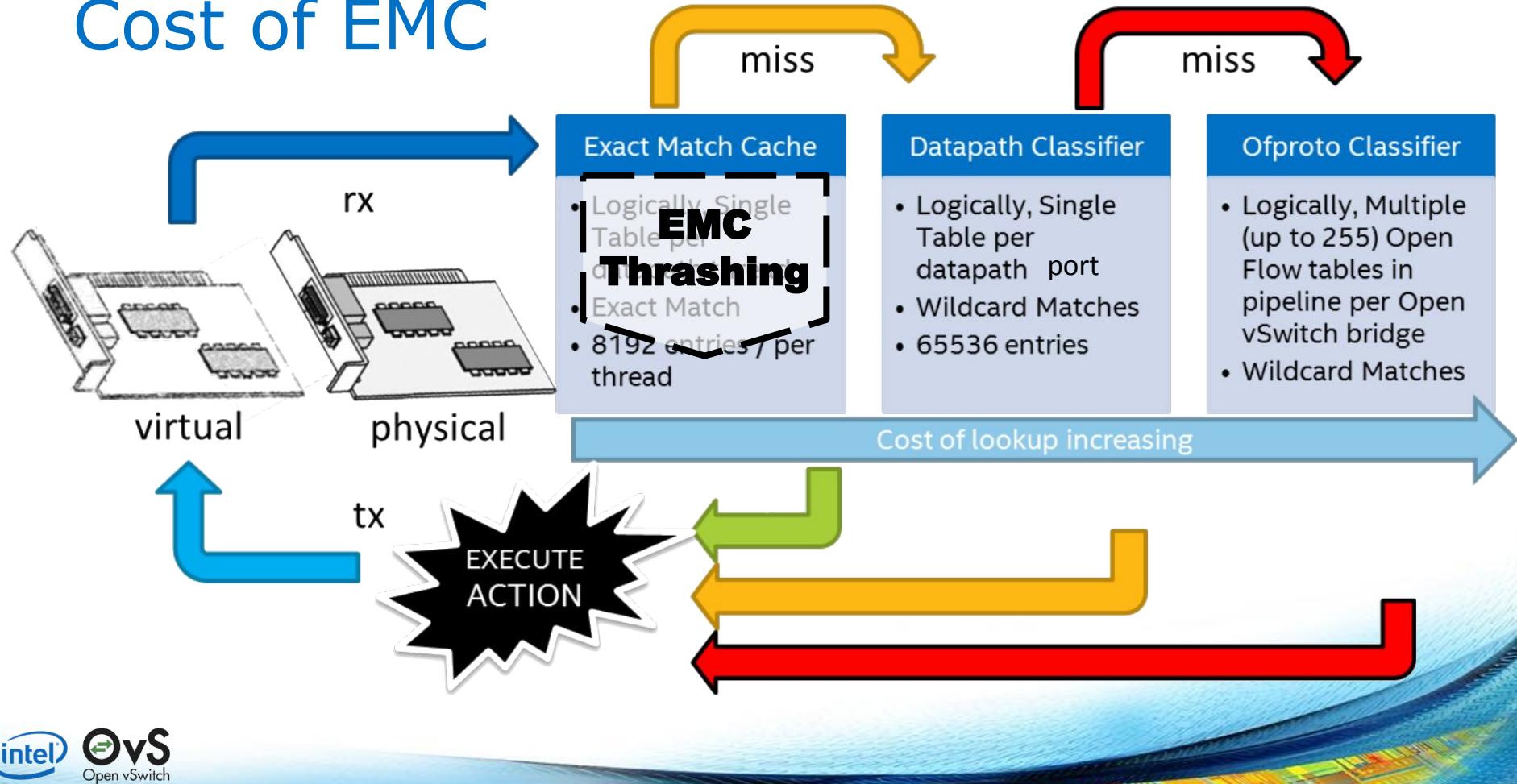
```
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$ sudo ./ovs-appctl dpif-netdev/pmd-rxq-show
pmd thread numa_id 0 core_id 3:
    isolated : true
    port: dpdkvhostuser0      queue-id: 0
    port: dpdkvhostuser3      queue-id: 0
    port: dpdkvhostuser6      queue-id: 0
pmd thread numa_id 0 core_id 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
sugeshch@silpixa00389816:~/repo/ovs_master/utilities$
```

Shows dp ports/rxq assignment to PMD threads

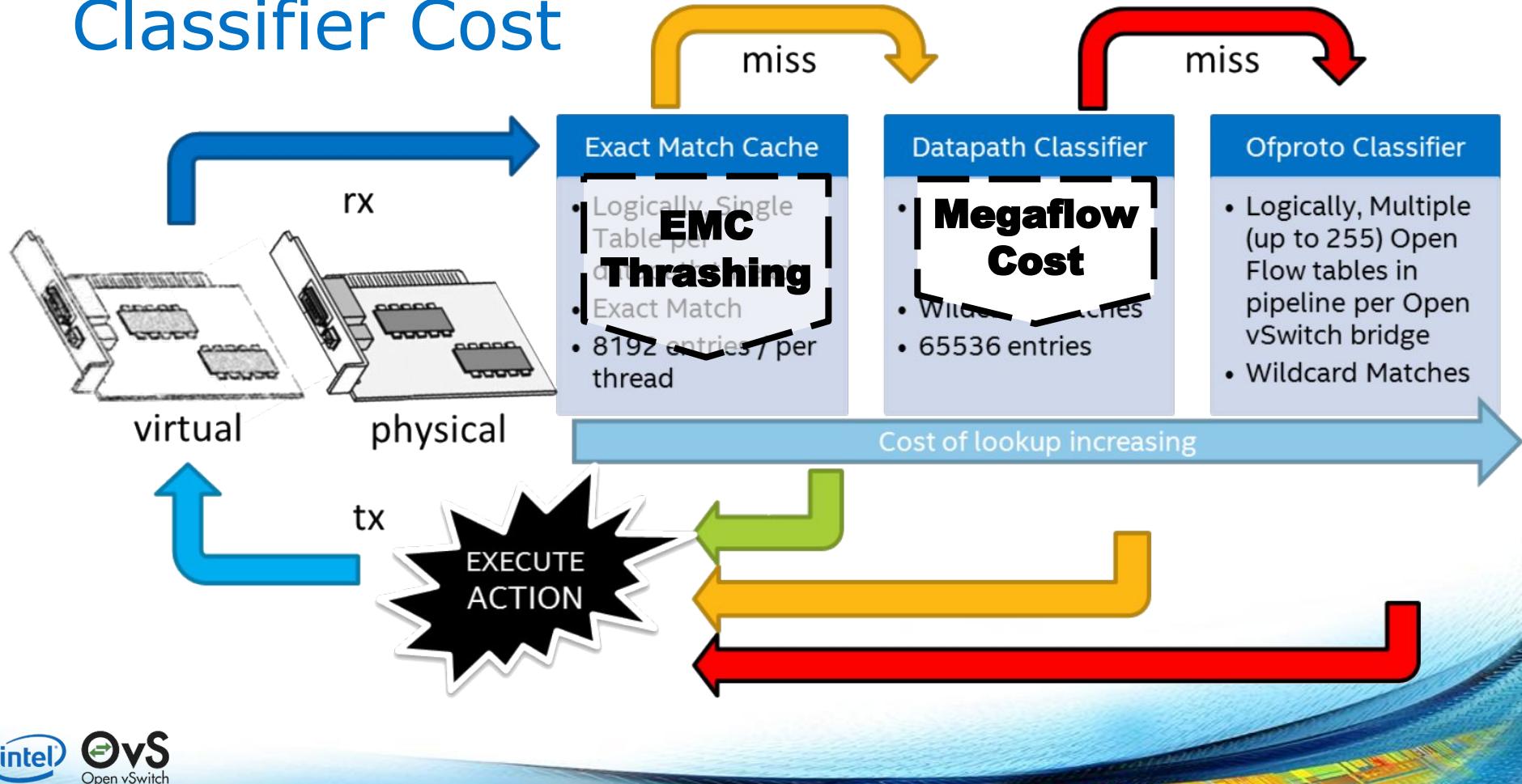
# OVS-DPDK Hotspots



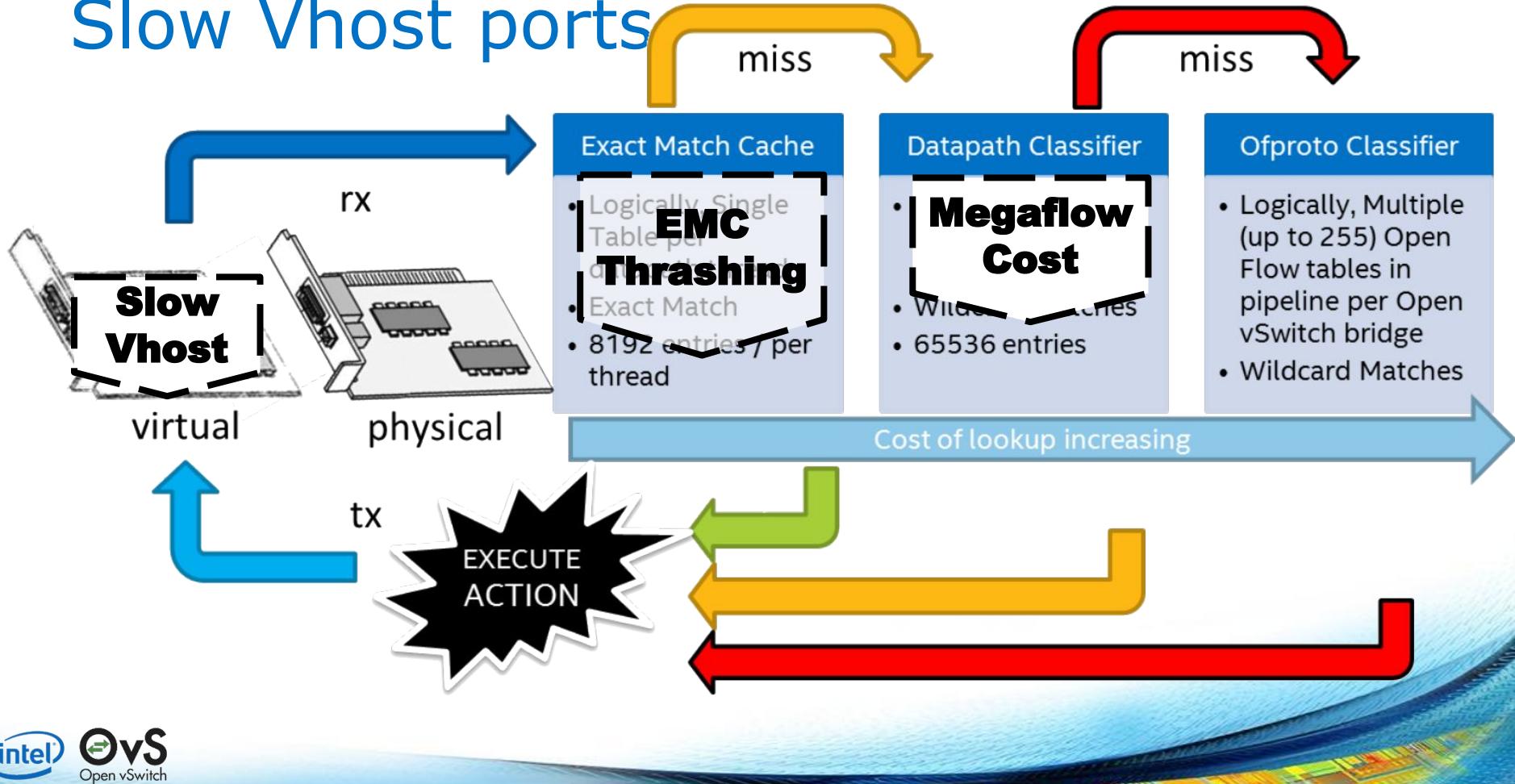
# Cost of EMC



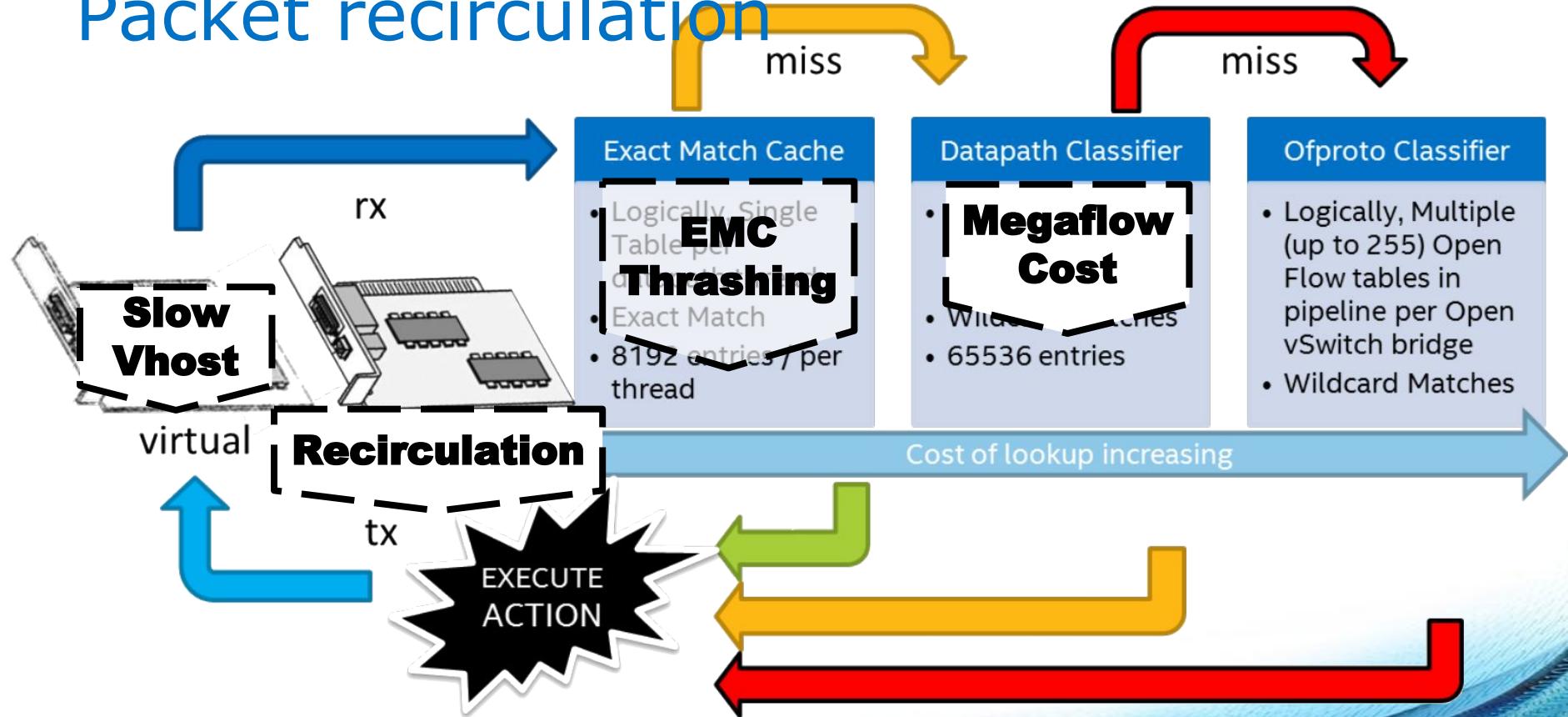
# Classifier Cost



# Slow Vhost ports



# Packet recirculation

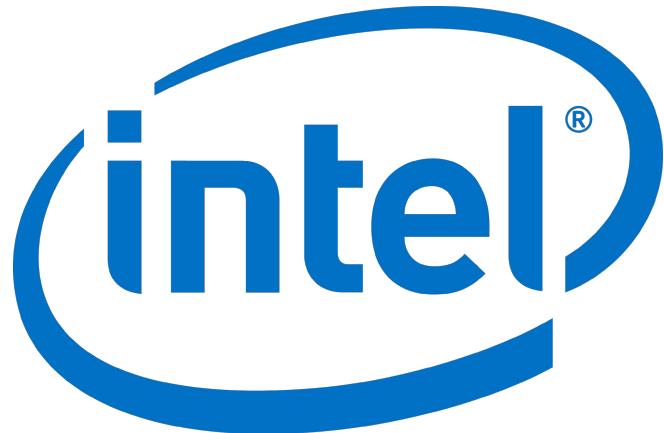


# Summary

- 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.