redhat.

OVS-DPDK: Embracing NUMA

OVS-DPDK params for NUMA

Kevin Traynor
ktraynor@redhat.com
16th November 2017

Introduction

- We want multi-NUMA
- Performance needs careful configuration, but...
- ...backup - let's get some packets flowing first
  - Device NUMA association
  - Provision cpu
  - Provision memory

redhat.

Physical ports NUMA association

Find NICs

```
# lspci | grep Network
01:00.0 Ethernet controller: Intel Corporation 82599ES
10-Gigabit SFI/SFP+ Network Connection (rev 01)
01:00.1 Ethernet controller: Intel Corporation 82599ES
10-Gigabit SFI/SFP+ Network Connection (rev 01)
```

Find NUMA node

```
# lspci -vmms 01:00.0 | grep NUMANode
NUMANode:    0
```


redhat.

Virtual ports NUMA association

Identify cpu NUMA layout

```
# numactl -H
available: 2 nodes (0-1)
node 0 cpus: 0 2 4 6 8 10 12 14
node 1 cpus: 1 3 5 7 9 11 13 15
```

Virtual NICs

```
<cputune>
    <vcpupin vcpu='0' cpuset='2'/>
    <vcpupin vcpu='1' cpuset='4'/>
    <vcpupin vcpu='2' cpuset='6'/>
    <emulatorpin cpuset='8'/>
  </cputune>
```


redhat.

Provision cpu

Identify cpu layout

```
# numactl -H
available: 2 nodes (0-1)
node 0 cpus: 0 2 4 6 8 10 12 14
node 1 cpus: 1 3 5 7 9 11 13 15
```

Set cpu for both NUMA nodes

```
# ovs-vsctl set Open_vSwitch . other_config:pmd-cpu-mask=0xF0
```


redhat.

Provision memory

Identify memory layout

```
    # cat /proc/meminfo | grep Hugepagesize
    Hugepagesize:    1048576 kB

    # cat /sys/devices/system/node/node*/meminfo | grep Huge
    Node 0 HugePages_Total:  16
    Node 0 HugePages_Free:   16
    ...
    Node 1 HugePages_Total:  16
    Node 1 HugePages_Free:   16
```

Allocate memory for both NUMA nodes

```
    # ovs-vsctl --no-wait set Open_vSwitch .
other_config:dpdk-socket-mem=1024,1024
```


redhat.

Debug

other_config

```
# ovs-vsctl get Open_vSwitch . other_config
    {dpdk-init="true", dpdk-lcore-mask="0x4",
dpdk-socket-mem="1024,1024", pmd-cpu-mask="F0"}
```

Debug messages

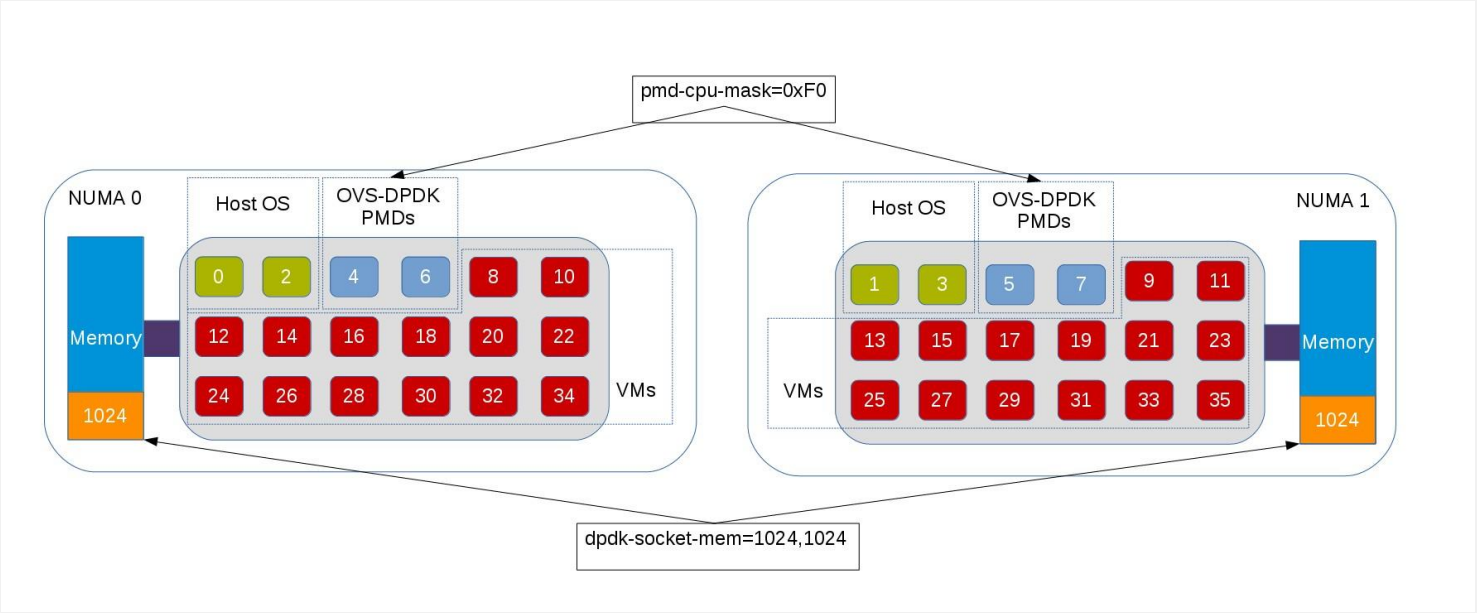
```
|dpdk|INFO|EAL ARGS: ovs-vswitchd -c 0x4 --socket-mem 1024,1024

|dpif_netdev|INFO|PMD thread on numa_id: 0, core id:  4 created.
|dpif_netdev|INFO|PMD thread on numa_id: 1, core id:  5 created.
|dpif_netdev|INFO|PMD thread on numa_id: 0, core id:  6 created.
|dpif_netdev|INFO|PMD thread on numa_id: 1, core id:  7 created.

|dpif_netdev|WARN|There's no available pmd thread on numa node 0
```


redhat.

Summary



Blog: https://developers.redhat.com/blog/2017/06/28/ovs-dpdk-parameters-dealing-with-multi-numa/

redhat.

redhat.

THANK YOU

plus.google.com/+RedHat

facebook.com/redhatinc

linkedin.com/company/red-hat

twitter.com/RedHatNews

youtube.com/user/RedHatVideos