OVS-DPDK Practices in Meituan cloud

Huai, Huang
MOS: experiences large-scale business practices

A public cloud provider based on the world's largest O2O platform
CONTENTS

- Introducing OVS-DPDK
- Performance
- Our works
- Further works
OVS & OVS-DPDK
C ONTENTS

- OVS-DPDK
- Performance
- Our works
- Further works
NIC-BR-NIC

- E5-2650 v2 @ 2.6G
  - 1 core
- 128GB mem
- 10-Gigabit X540-AT2
- kernel 3.10.0
- Open vSwitch 2.4.90
- qemu 2.6.0
- 64bytes

![Graph showing performance comparison between OVS-kernel and OVS-DPDK]

```plaintext
in_port=2, action=output:1
```
NIC-BR-VM

```
ip,in_port=2,nw_src=192.168.102.0/27 actions=output:3
ip,in_port=2,nw_src=192.168.102.32/27 actions=output:4
ip,in_port=2,nw_src=192.168.102.64/27 actions=output:5
ip,in_port=2,nw_src=192.168.102.96/27 actions=output:6
...
```
CONTENTS

- OVS-DPDK
- Performance
- Our works
- Further
Question 1: stability

- stability problem
  - OVS-DPDK, DPDP, qemu
  - fix ovs, dpdk, qemu 16 bugs
- debug status
  - rte resource: rte_heap
  - vhost: virtqueue, fd_set, reconnect_list, virtio_dev...
  - NIC: link status, bond status...
Question 2: Function Defect

- raise
- vhost recovery
- live migration
- windows vm run
- finished by Intel
Problem 3: Compatibility

- DPDK-BOND port
- dump
- QoS
Problem 3: Compatibility

- dump
- ovs-filter framework
  - filter hook
  - load/unload module
- kernel module transplant
- L4 session
Problem 4: LOCAL capability

- DPDK current method
  - TAP
  - KNI
  - 2-3Gbps big packets
Problem 4: LOCAL capability

- Local port
  - forward bottlenecks
- Case.
  - Distributed file system (EBS)
  - I/O -> network
  - Offline calculation, Data collection
SR-IOV Approach

- SRIOV
  - PF: DPDK NIC
  - VF: VM
  - eswitch

- Switch
  - controll
  - user isolation
  - live migration
LOCAL capability

- SR-IOV method
  - linux PF + DPDK VF
  - DPDK PF + linux VF
- CHANGES
  - DPDK VF code
  - ixgbevf code
  - PF, VF queue set
  - bond4
CONTENTS

- OVS-DPDK
- Performance
- Our works
- Further
Problem 5: OVS -> OVS-DPDK

- old OVS update OVS-DPDK
  - Method 1: native update
  - Method 2: offline migration
  - question: VM need shutdown
- live migration
  - vhost-kernel vm -> vhost-user vm
Problem 6: Hot Update

- **Question**
  - ovs-vswitchd restart 2-3min
  - last time section
    - hugepage init
    - nic port restart
      - link auto negotiation
      - bond4 handshake
    - vhost-user port reinit

  - **step 1:** 2+min -> 10- s

  - **step 2:** 10- s -> 2- s

- **Solution.**
Thanks!!

huanghuai@meituan.com