DPDK 高性能处理框架 VPP

VPP 环境安装

\$ git clone -b stable/1801 https://github.com/FDio/vpp.git

```
king@ubuntu:~/share$ git clone -b stable/1801 https://github.com/FDio/vpp.git
Cloning into 'vpp'...
remote: Enumerating objects: 121933, done.
remote: Counting objects: 100% (6928/6928), done.
remote: Compressing objects: 100% (2715/2715), done.
remote: Total 121933 (delta 4800), reused 6227 (delta 4171), pack-reused 115005
Receiving objects: 100% (121933/121933), 102.20 MiB | 773.00 KiB/s, done.
Resolving deltas: 100% (92173/92173), done.
Checking connectivity... done.
king@ubuntu:~/share$
```

\$./extras/vagrant/build.sh && make

```
king@ubuntu:~/share/dpdk-linuxapp/yp$ ls build-root/
autowank config.site path_setup vpp-api-java_18.01.2-1~g9b554f3_amd64.deb
bootstrap.sh copying platforms.mk vpp-api-lua_18.01.2-1~g9b554f3_amd64.deb
build-config.mk deb vrpp-api-python_18.01.2-1~g9b554f3_amd64.deb
build-config.mk.README install-vpp_debug-native build-tool-native install-vpp-native build-vpp_debug-native build-vpp_abebug-native build-vpp_native packages vpp_18.01.2-1~g9b554f3_amd64.deb
build-vpp-native packages vpp_18.01.2-1~g9b554f3_amd64.deb
vpp_18.01.2-1~g9b554f3_amd64.deb
vpp_pls.01.2-1~g9b554f3_amd64.deb
vpp_pls.01.2-1~g9b554f3_amd64.deb
vpp_pls.01.2-1~g9b554f3_amd64.deb
vpp_pls.01.2-1~g9b554f3_amd64.deb
vpp_pls.01.2-1~g9b554f3_amd64.deb
vpp_pls.01.2-1~g9b554f3_amd64.deb
```

在编译成功以后,会生成上图红色的 deb 表

```
$ dpkg -i vpp-lib_18.01.2-1~g9b554f3_amd64.deb
```

\$ dpkg -i vpp_18.01.2-1~g9b554f3_amd64.deb

\$ dpkg -i vpp-dev_18.01.2-1~g9b554f3_amd64.deb

\$ dpkg -i vpp-plugins_18.01.2-1~g9b554f3_amd64.deb

执行完成以后,在 /etc/vpp/startup.conf

```
king@ubuntu:~/share/dpdk-linuxapp/vpp$ ls /etc/vpp/startup.conf
/etc/vpp/startup.conf
king@ubuntu:~/share/upuk-linuxapp/vpp$
```

在 /usr/lib/vpp_plugins/

在 conf 文件里面添加了三个 dev 网卡

```
ot@ubuntu:/home/king/share# ifconfig
h0 Link encap:Ethernet HWaddr 00:0c:29:85:2e:88
inet addr:192.168.0.120 Bcast:192.168.0.255 Mask:255.255.25
inet6 addr: fe80::20c:29ff:fe85:2e88/64 Scope:Link
othe
              UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:295 errors:0 dropped:0 overruns:0 frame:0
              TX packets:10 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000
              RX bytes:41851 (41.8 KB) TX bytes:1332 (1.3 KB)
eth1
              Link encap:Ethernet HWaddr 00:0c:29:85:2e:92
              inet addr:192.168.0.109 Bcast:192.168.0.255 Mask:255.255.255.0
             inet6 addr: fe80::20c:29ff:fe85:2e92/64 Scope:Link
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:222 errors:0 dropped:0 overruns:0 frame:0
TX packets:101 errors:0 dropped:0 overruns:0 carrier:0
              collisions:0 txqueuelen:1000
RX bytes:31827 (31.8 KB) TX bytes:12708 (12.7 KB)
              Link encap:Ethernet HWaddr 00:0c:29:85:2e:9c
inet addr:192.168.0.118 Bcast:192.168.0.255 Mask:255.255.255.0
inet6 addr: fe80::20c:29ff:fe85:2e9c/64 Scope:Link
eth2
              UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
              RX packets:295 errors:0 dropped:0 overruns:0 frame:0
              TX packets:10 errors:0 dropped:0 overruns:0 carrier:0
              collisions:0 txqueuelen:1000
              RX bytes:41851 (41.8 KB) TX bytes:1332 (1.3 KB)
eth3
              Link encap:Ethernet HWaddr 00:0c:29:85:2e:7e
inet addr:192.168.232.133 Bcast:192.168.232.255 Mask:255.255.255.0
inet6 addr: fe80::20c:29ff:fe85:2e7e/64 Scope:Link
              UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
              RX packets:1026 errors:0 dropped:0 overruns:0 frame:0
TX packets:733 errors:0 dropped:0 overruns:0 carrier:0
              collisions:0 txqueuelen:1000
              RX bytes:87706 (87.7 KB) TX bytes:112831 (112.8 KB)
lo
              Link encap:Local Loopback
              inet addr:127.0.0.1 Mask:255.0.0.0
inet6 addr: ::1/128 Scope:Host
UP LOOPBACK RUNNING MTU:65536 Metric:1
              RX packets:164 errors:0 dropped:0 overruns:0 frame:0
              TX packets:164 errors:0 dropped:0 overruns:0 carrier:0
              collisions:0 txqueuelen:1
              RX bytes:12080 (12.0 KB) TX bytes:12080 (12.0 KB)
```

Startup.conf 里面的 pci 地址

```
root@ubuntu:/home/king/share# lspci | grep Ethernet
02:01.0 Ethernet controller: Intel Corporation 82545EM Gigabit Ethernet Controller (Copper) (rev 01)
03:00.0 Ethernet controller: VMware VMXNET3 Ethernet Controller (rev 01)
0b:00.0 Ethernet controller: VMware VMXNET3 Ethernet Controller (rev 01)
13:00.0 Ethernet controller: VMware VMXNET3 Ethernet Controller (rev 01)
root@ubuntu:/home/king/share#
```

编写 Startup.conf

```
king@ubuntu:~/share/dpdk-linuxapp/vpp$ cat /etc/vpp/startup.conf

cpu {
    main-core 0
}

unix {
    interactive cli-listen 127.0.0.1:5002
    log /tmp/vpp.log
    full-coredump
}

dpdk {
    uio-driver igb_uio
    dev 0000:03:00.0
    dev 0000:13:00.0
    dev 0000:13:00.0
    dev 0000:13:00.0
    dev 0000:13:00.0
    socket-mem 512, 512
}

api-trace {
    on
}
```

如果需要使用 vppctl, 需要 cli-listen /run/vpp/cli.sock

```
unix {
    interactive cli-listen /run/vpp/cli.sock
    log /tmp/vpp.log
    full-coredump
}
```

再 dpdk 的运行环境准备就绪,

```
root@ubuntu:/home/king/share# modprobe uio
root@ubuntu:/home/king/share# modprobe igb_uio
root@ubuntu:/home/king/share#
root@ubuntu:/home/king/share#
root@ubuntu:/home/king/share#
root@ubuntu:/home/king/share# echo 4 > /sys/devices/system/node/node0/hugepages/hugepages-1048576kB/nr_hugepages
root@ubuntu:/home/king/share# echo 1024 > /sys/devices/system/node/node0/hugepages/hugepages-2048kB/nr_hugepages
root@ubuntu:/home/king/share#
```

将 eth0/eth1/eth2 关闭

- \$ ifconfig eth0 down
- \$ ifconfig eth1 down
- \$ ifconfig eth2 down

启动 vpp,

\$ vpp -c /etc/vpp/startup.conf

```
roof@ubuntur/home/king/share/dpdk-linuxapp/vpp# vpp -c /etc/vpp/startup.conf
vlib_plugin_early_init:356: plugin_path /usr/lib/vpp_plugins
load_one_plugin:341: Loaded plugin: acl_plugin.so (Access Control Lists)
load_one_plugin:341: Loaded plugin: dpdk_plugin.so (Data Plane Development Kit (DPOK))
load_one_plugin:341: Loaded plugin: ipom_plugin.so (GPOF)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (Jebound OAM)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (Jebound OAM)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (Loade Interpretation)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (Loade Interpretation)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (Loade Interpretation)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (Hoberone' Address Iranslation)
load_one_plugin:341: Loaded plugin: ipom_plugin.so (PPOF)
load_one_plugin:341: Loaded plugin: i
```

停止 vpp

\$ service vpp stop

root@ubuntu:/home/king/share# service vpp stop
root@ubuntu:/home/king/share#

Vpp 命令操作

vpp# show int

```
load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/flowprobe_test_plugin.so
dpdk_config:1240: EAL init args: -c 1 -n 4 --huge-dir /run/vpp/hugepages --file-prefix vp
p -w 0000:03:00.0 -w 0000:0b:00.0 -w 0000:13:00.0 --master-lcore 0 --socket-mem 64
EAL: 1024 hugepages of size 2097152 reserved, but no mounted hugetlbfs found for that siz
EAL: VFIO support initialized
EAL: Invalid NUMA socket, default to 0
EAL: Invalid NUMA socket, default to 0
EAL: Invalid NUMA socket, default to 0
DPDK physical memory layout:
Segment 0: IOVA:0x40000000, len:1073741824, virt:0x7f69c0000000, socket_id:0, hugepage_sz
unix_physmem_region_iommu_register: ioctl (VFIO_IOMMU_MAP_DMA): Invalid argument
0: dpdk_ipsec_process:1011: not enough DPDK crypto resources, default to OpenSSL
                                                       Idx
                                                                       State
                                                                                                Counter
                                                                                                                            Count
 GigabitEthernet13/0/0
GigabitEthernet3/0/0
GigabitEthernetb/0/0
                                                                       down
                                                                       down
 local0
 vpp# set interface state GigabitEthernet13/0/0 up
 vpp# show int
                                                                       State
                                                                                               Counter
                                                                                                                            Count
 GigabitEthernet13/0/0
                                                                                        rx packets
                                                                        up
                                                                                        rx bytes
                                                                                       drops
                                                                                        ip6
GigabitEthernet3/0/0
GigabitEthernetb/0/0
 local0
                                                                       down
```

vpp# set interface state GigabitEthernet13/0/0 up

```
pp# set interface state GigabitEthernet13/0/0 up
vpp# show int
                                           State
                                                          Counter
                                                                            Count
GigabitEthernet13/0/0
                                            up
                                                      rx packets
                                                      rx bytes
                                                                                    624
                                                     drops
                                                      ip6
GigabitEthernet3/0/0
GigabitEthernetb/0/0
                                           down
local@
                                           down
```

vpp# set interface ip address GigabitEthernet13/0/0 192.168.0.119/24

```
interface ip address GigabitEthernet13/0/0 192.168.0.119/32
/pp# show int
                                 Idx
                                            State
                                                           Counter
                                                                             Count
GigabitEthernet13/0/0
                                                      rx packets
                                             up
                                                      rx bytes
                                                      drops
                                                                                    749
                                                      ip4
                                                                                    222
GigabitEthernet3/0/0
GigabitEthernetb/0/0
```

配置 vlan

vpp# create bridge-domain 10
vpp# set interface l2 bridge GigabitEthernet3/0/0 10
vpp# create sub-interface GigabitEthernet3/0/0 10 dot1q 10
vpp# show bridge-domain 10 detail

```
pp# create sub-interface GigabitEthernet3/0/0 10 dot1q 10
GigabitEthernet3/0/0.10
vpp# show int
                                           State
                                                           Counter
                                                                            Count
                                 Idx
GigabitEthernet13/0/0
                                            down
GigabitEthernet13/0/0.10
                                            down
GigabitEthernet3/0/0
                                            up
                                                      rx packets
                                                                                   2057
                                                      rx bytes
                                                                                 963054
                                                      tx packets
                                                                                    10
                                                      tx bytes
                                                      drops
                                                                                   2052
                                                      ip4
GigabitEthernet3/0/0.10
GigabitEthernetb/0/0
                                                                                    738
                                                      rx packets
                                                      rx bytes
                                                                                 209532
                                                      drops
                                                                                    488
                                                      ip6
                                                                                    228
```

配置 veth

- \$ ip link add name vpp1out type veth peer name vpp1host
- \$ ip addr show vpp1host
- \$ ip link set dev vpp1out up
- \$ ip link set dev vpp1host up

```
oot@ubuntu:/home/king# ip link add name vpp1out type veth peer name vpp1host
root@ubuntu:/home/king#
root@ubuntu:/home/king# ip addr show vpp1host
6: vpp1host@vpp1out: <BROADCAST,MULTICAST,M-DOWN> mtu 1500 qdisc noop state DOWN group default qlen 1000
link/ether 36:fb:a2:c5:b7:60 brd ff:ff:ff:ff:ff
oot@ubuntu:/home/king#
root@ubuntu:/home/king# ip link set dev vpp1out up
root@ubuntu:/home/king# ip link set dev vpp1host up
root@ubuntu:/home/king#
root@ubuntu:/home/king# ifconfig
eth3 Link encap:Ethernet HWaddr 00:0c:29:85:2e:7e
             inet addr:192.168.232.133 Bcast:192.168.232.255 Mask:255.255.255.0
            inet6 addr: fe80::20c:29ff:fe85:2e7e/64 Scope:Link
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:256387 errors:0 dropped:0 overruns:0 frame:0
             TX packets:333858 errors:0 dropped:0 overruns:0 carrier:0
            collisions:0 txqueuelen:1000
            RX bytes:47129193 (47.1 MB) TX bytes:547739230 (547.7 MB)
            Link encap:Local Loopback
inet addr:127.0.0.1 Mask:255.0.0.0
inet6 addr: ::1/128 Scope:Host
lo
            UP LOOPBACK RUNNING MTU:65536 Metric:1
            RX packets:202 errors:0 dropped:0 overruns:0 frame:0
            TX packets:202 errors:0 dropped:0 overruns:0 carrier:0
             collisions:0 txqueuelen:1
            RX bytes:14318 (14.3 KB) TX bytes:14318 (14.3 KB)
vpp1host Link encap:Ethernet HWaddr 36:fb:a2:c5:b7:60
            inet6 addr: fe80::34fb:a2ff:fec5:b760/64 Scope:Link
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
             RX packets:7 errors:0 dropped:0 overruns:0 frame:0
            TX packets:7 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000
             RX bytes:578 (578.0 B) TX bytes:578 (578.0 B)
vpp1out Link encap:Ethernet HWaddr 36:5a:4c:2e:51:d3
   inet6 addr: fe80::345a:4cff:fe2e:51d3/64 Scope:Link
             UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
            RX packets:7 errors:0 dropped:0 overruns:0 frame:0
TX packets:7 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:1000
             RX bytes:578 (578.0 B) TX bytes:578 (578.0 B)
```

配置 vpp1host ip 地址

\$ ip addr add 192.168.3.128/24 dev vpp1host

\$ ip addr show vpp1host

```
root@ubuntu:/home/king# ip addr add 192.168.3.128/24 dev vpp1host
root@ubuntu:/home/king# ip addr show vpp1host
6: vpp1host@vpp1out: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default qlen 1000
    link/ether 36:fb:a2:c5:b7:60 brd ff:fff:ff:ff:ff
    inet 192.168.3.128/24 scope global vpp1host
        valid_lft forever preferred_lft forever
    inet6 fe80::34fb:a2ff:fec5:b760/64 scope link
        valid_lft forever preferred_lft forever
root@ubuntu:/home/king#
```

VPP 添加插件 plugin

实现一个打印包头信息的插件,源代码如下:

```
root@ubuntu:/home/king/share/dpdk-linuxapp/vpp# ls src/plugins/pktdump
pktdump.c pktdump.h pktdump_node.c
root@ubuntu:/home/king/share/dpdk-linuxapp/vpp#
```

修改 src/configure.ac

修改 src/plugins/Makefile.am

```
#### Update By King
if ENABLE_KINGSAMPLE_PLUGIN
include kingsample.am
endif
if ENABLE_PKTDUMP_PLUGIN
include pktdump.am
endif
```

- # make wipe
 # make build
 # make run
- load one plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/ioam_pot_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/stm_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/kubeproxy_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/loam_trace_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/bb_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/poppoe_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/soam_export_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/soam_export_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/memif_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/flowprobe_test_plugin.so
 load_one_plugin:63: Loaded plugin: /usr/lib/vpp_api_test_plugins/flowprobe_test_plugin.so
 vlib_poi_bind_to_uio: Skipping PCI device 0000:02:01.0 as host interface eth3 is up
 dpdk_config:1240: EAL init args: -c 1 -n 4 --huge-dir /run/vpp/hugepages --file-prefix vpp -b 0000:0
 2:01.0 --master-lcore 0 --socket-mem 64
 EAL: 1024 hugepages of size 2097152 reserved, but no mounted hugetlbfs found for that size
 EAL: VFIO support initialized
 EAL: Invalid NUMA socket, default to 0
 DPDK physical memory layout:
 Segment 0: IOVA:0x40000000, len:1073741824, virt:0x7fbf80000000, socket_id:0, hugepage_sz:1073741824
 , nchannel:0, nrank:0
 unix_physmem_region_iommu_register: ioctl (VFIO_IONMU_MAP_DMA): Invalid argument
 0: dpdk_ibsec_process:1011: not enough DPDK crypto resources, default to 0penSSL
 0: sym_client_scan_this_region_nolock:1181: /global_vm: cleanup ghost pid 4367
 0: sym_client_scan_this_region_nolock:1181: /global_vm: cleanup ghost pid 4367

```
set int ip address GigabitEthernet3/0/0 192.168.0.120/24
DBGvpp#
DBGvpp#
OBGvpp# show int
             Name
                                 Idx
                                            State
                                                           Counter
                                                                             Count
GigabitEthernet13/0/0
                                            down
GigabitEthernet3/0/0
                                                      rx packets
                                                      rx bytes
                                                                                  31608
                                                      drops
                                                      ip4
GigabitEthernetb/0/0
                                            down
local0
                                            down
DBGvpp# show int address
GigabitEthernet13/0/0 (dn):
GigabitEthernet3/0/0 (up):
 192.168.0.120/24
SigabitEthernetb/0/0 (dn):
local0 (dn):
```

执行结果:

Windows 物理机 ping 192.168.0.120

```
GigabitEthernetb/e/e 2 down
DoGGVpps show int address
GigabitEthernet13/e/e (dn):
Giga
```

插件 plugin 的原理

