Subject: Strange Behavior with multiple venet0 interfaces
Posted by glic3rinu on Thu, 09 Jul 2009 09:14:53 GMT
View Forum Message <> Reply to Message

Hi all,
I have Debian 5.0 based server with various network interfaces connected on different networks (eth0:77.178.25.0/24, eth1:10.0.0.0/24,eth2:10.0.10.0/24, eth3:10.10.0.0/24, and more...). Today I have been doing tests with an OpenVZ virtual machine with multiple venet0 interfaces (vente0:0, venet0:1 and venet0:2), each interface belonging to a different network (venet0:0 77.178.25.145/24, venet0:2 10.10.0.155/24, venet0:1 10.0.0.155/24).

I observe a strange behavior when I ping from OpenVZ VE to another machines on the network. The source IP address of the output packets for the VE is in all the cases 77.178.25.145, in spite of being 10.10.0.155 or 10.0.0.155

In the worst case it causes packet losing because some machines of the private network doesn't know how/to route packets to 77.178.25.145.

Whats the reason of this behavior? is this a bug? Is there any way to fix it? Perhaps a clue to solve the problem is that the interfaces are called venet0:0,venet0:1, venet0:2 instead of venet0, venet1 and venet2?

I know this problem disappears when using the Veth interfaces instead of Venet0, but I like to use the Venet0 system because it offer more performance and security.

I leave an example to ping 10.0.0.10 from the virtual machine.

1) Tcpdump show as on the hardware node packets leave by the correct interface (eth1: network 10.0.0.0/24) but with the source address changed, 77.178.25.145 instead of 10.0.0.155

tcpdump -i eth1
14:20:13.939721 IP 77.178.25.145 > 10.0.0.10: ICMP echo request, id 63234, seq 2, length 64
14:20:14.939208 IP 77.178.25.145 > 10.0.0.10: ICMP echo request, id 63234, seq 3, length 64

Then the answers returns for another interface (eth0: network 77.178.25.0/24)

tcpdump -i eth0

14:22:46.075225 IP 10.0.0.10 > 77.178.25.145: ICMP echo reply, id 63490, seq 6564, length 64
14:22:46.075473 IP 10.0.0.10 > 77.178.25.145: ICMP echo reply, id 63490, seq 6565, length 64

finally the packets are relayed to the virtual machine.

tcpdump -i venet0

12:24:45.211486 IP 10.0.0.10 > 77.178.25.145: ICMP echo reply, id 10500, seq 21, length 64
12:24:46.211202 IP 77.178.25.145 > 10.0.0.10: ICMP echo request, id 10500, seq 22, length 64

Configuration of the Virtual Machine - OpenVZ

```
test:/# route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
192.0.2.1       0.0.0.0         255.255.255.255 UH    0      0        0 venet0
0.0.0.0         192.0.2.1       0.0.0.0         UG    0      0        0 venet0


test:/# ifconfig
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:16436  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)

venet0    Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          inet addr:127.0.0.1  P-t-P:127.0.0.1  Bcast:0.0.0.0 Mask:255.255.255.255
          UP BROADCAST POINTOPOINT RUNNING NOARP  MTU:1500  Metric:1
          RX packets:177951 errors:0 dropped:0 overruns:0 frame:0
          TX packets:179295 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:14947840 (14.2 MiB)  TX bytes:15060736 (14.3 MiB)

venet0:0  Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          inet addr:77.178.25.145  P-t-P:77.246.179.123  Bcast:0.0.0.0 Mask:255.255.255.255
          UP BROADCAST POINTOPOINT RUNNING NOARP  MTU:1500  Metric:1

venet0:1  Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          inet addr:10.0.0.155  P-t-P:10.0.0.155  Bcast:0.0.0.0 Mask:255.255.255.255
          UP BROADCAST POINTOPOINT RUNNING NOARP  MTU:1500  Metric:1

venet0:2  Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          inet addr:10.10.0.155  P-t-P:10.10.0.155  Bcast:0.0.0.0 Mask:255.255.255.255
          UP BROADCAST POINTOPOINT RUNNING NOARP  MTU:1500  Metric:1
```

Hardware Node configuration.

```
backup:~# route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
77.178.25.145   0.0.0.0         255.255.255.255 UH    0      0        0 venet0
10.0.0.155      0.0.0.0         255.255.255.255 UH    0      0        0 venet0
10.10.0.155     0.0.0.0         255.255.255.255 UH    0      0        0 venet0
77.178.25.0     0.0.0.0         255.255.255.0   U     0      0        0 vmbr0
10.0.0.0        0.0.0.0         255.255.255.0   U     0      0        0 eth1
10.10.0.0       0.0.0.0         255.255.255.0   U     0      0        0 eth3
10.0.10.0       0.0.0.0         255.255.255.0   U     0      0        0 eth2
10.0.10.0       0.0.0.0         255.255.255.0   U     0      0        0 eth4
0.0.0.0         77.178.25.1     0.0.0.0         UG    0      0        0 vmbr0
```

/etc/vz/conf/101.conf

```
# CPU fair sheduler parameter
CPUUNITS="1000"
CPUS="1"
VE_ROOT="/var/lib/vz/root/$VEID"
VE_PRIVATE="/var/lib/vz/private/$VEID"
OSTEMPLATE="debian-5.0-standard_5.0-1_i386"
ORIGIN_SAMPLE="pve.auto"
IP_ADDRESS="77.178.25.145 10.0.0.155 10.10.0.155"
HOSTNAME="test.xxxx.org"
DESCRIPTION="Debian 5.0 (standard)"
NAMESERVER="208.67.222.222"
SEARCHDOMAIN="xxxxx.org"
```

Thanks very much!
Marc