Subject: time to live exceeded
Posted by romeor on Wed, 28 Apr 2010 10:51:46 GMT
View Forum Message <> Reply to Message

Hello,

First of all - openvz is great stuff, thank you.

here goes my question (i'm sure that there is some miss in my knowlege, so please could you help me?)

i've got host machine with openvz installed and running 3 containers.

host has 2 eths in different subnets one in public network and one in local one here is conf:

```
[root@watcher ~]# ifconfig
eth0      Link encap:Ethernet  HWaddr 00:C0:9F:3C:D6:22
          inet addr:193.40.142.225  Bcast:193.40.142.255  Mask:255.255.255.0
          inet6 addr: fe80::2c0:9fff:fe3c:d622/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:6962 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1979 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:100
          RX bytes:607673 (593.4 KiB)  TX bytes:280738 (274.1 KiB)
          Base address:0xecc0 Memory:fe100000-fe120000

eth1      Link encap:Ethernet  HWaddr 00:04:75:DA:4C:1B
          inet addr:172.16.0.25  Bcast:172.16.0.255  Mask:255.255.255.0
          inet6 addr: fe80::204:75ff:feda:4c1b/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:82347 errors:0 dropped:0 overruns:1 frame:0
          TX packets:71013 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:9093171 (8.6 MiB)  TX bytes:10801807 (10.3 MiB)
          Interrupt:185 Base address:0x6000

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:8 errors:0 dropped:0 overruns:0 frame:0
          TX packets:8 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:560 (560.0 b)  TX bytes:560 (560.0 b)

venet0    Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          UP BROADCAST POINTOPOINT RUNNING NOARP  MTU:1500  Metric:1
          RX packets:37309 errors:0 dropped:0 overruns:0 frame:0
          TX packets:17915 errors:0 dropped:0 overruns:0 carrier:0
```

collisions:0 txqueuelen:0
RX bytes:6980845 (6.6 MiB) TX bytes:2348837 (2.2 MiB)

i've read some faqs and figured out that i've got to set up some additional routing to get my containers to work, so i did:

ip rule add from 172.16.0.26 table 6
ip rule add from 172.16.0.21 table 6
ip route add default dev eth1 via 172.16.0.254 table 6

and here is the result:

[root@watcher ~]# ip rule
0:	from all lookup 255
32764: from 172.16.0.21 lookup 6
32765: from 172.16.0.26 lookup 6
32766: from all lookup main
32767: from all lookup default

[root@watcher ~]# ip route
172.16.0.21 dev venet0 scope link
193.40.142.130 dev venet0 scope link
193.40.142.226 dev venet0 scope link
172.16.0.26 dev venet0 scope link
172.16.0.0/24 dev eth1 proto kernel scope link src 172.16.0.25
193.40.142.0/24 dev eth0 proto kernel scope link src 193.40.142.225
169.254.0.0/16 dev eth1 scope link
default via 193.40.142.129 dev eth0

route:
[root@watcher ~]# route
Kernel IP routing table
Destination   Gateway   Genmask   Flags Metric Ref   Use Iface
172.16.0.21   *   255.255.255.255 UH   0   0   0 venet0
dns2.tlulib.ee *   255.255.255.255 UH   0   0   0 venet0
monitor.tlulib. *   255.255.255.255 UH   0   0   0 venet0
172.16.0.26   *   255.255.255.255 UH   0   0   0 venet0
172.16.0.0   *   255.255.255.0   U   0   0   0 eth1
193.40.142.0   *   255.255.255.0   U   0   0   0 eth0
169.254.0.0   *   255.255.0.0   U   0   0   0 eth1
default   193.40.142.129 0.0.0.0   UG   0   0   0 eth0

so i'm runing fine network for my host and containers (2 containers with 172.16.0.0/24 network and 1 with 193.40.142.0/24)

BUT!

here goes the problem i need to solve>
both containers do not speak to each other, when i ping from one to another, i get
From 193.40.142.225 icmp_seq=629 Time to live exceeded
for both.
what should i do with that? any help is welcomed.

UPDATE:

SEEMS TO BE SOLVED!!
found some info in forums, i dont really know what these lines do, but they did help.
if some1 can, please explain !!!!

added theese lines:

ip route add 172.16.0.21 dev venet0 table 6
ip route add 172.16.0.26 dev venet0 table 6

Wel,, but now i cant ping these 2 ip addresses from other pcs ...
something weird...

Subject: Re: time to live exceeded
Posted by maratrus on Fri, 30 Apr 2010 13:11:19 GMT
View Forum Message <> Reply to Message

Hello,

look, I'd better try to explain you what's going on. I hope it'll shed some light on your problem.

As you might have noticed each of your VE uses venet0 interface to communicate with external world. venet0 is just nothing but point-to-point connection. So, everything is put inside venet0 on the HN goes to the proper VE (venet0 driver reads destination ip address and puts a network packet to the proper VE).
Let's check an outgoing traffic. If a VE issues a network packet it goes directly to venet0 interface (look at the default gateway inside your VE). But on the HN you have a table 6 which should catch all traffic going from 172.16.0.26 and 172.16.0.21. There is a singe route inside this table ("ip route add default dev eth1 via 172.16.0.254 table 6"). Hence all network packets going from the VE (172.16.0.21 or .26) are to be sent to 172.16.0.254.
At the same moment there are the following routes on your HN
Quote:
[root@watcher ~]# ip route
172.16.0.21 dev venet0 scope link
193.40.142.130 dev venet0 scope link
193.40.142.226 dev venet0 scope link
172.16.0.26 dev venet0 scope link

that means that traffic which is intended to your VEs should go through venet0 interface.

Let's consider your case. You try to communicate from VE 193.40.142.130 to VE 172.16.0.26. The network packet goes to HN and then according to routing table it should pass to venet0 again ("172.16.0.26 dev venet0 scope link"). The answer from the VE goes to HN but at this moment the routing decision should obey the routing rule in the table 6 and network packet goes to 172.16.0.254. I don't know what's going on on that node but eventually it routes the packet somewhere else and the packet never reaches 193.40.142.130 => the VE cannot catch the answer.
I suppose the same situation occurs when you try to ping 172.16.0.26 from 172.16.0.21.

P.S. please use tcpdump utility to trace network traffic. It really helps you find out where the packet is lost.

Subject: Re: time to live exceeded
Posted by romeor on Wed, 26 May 2010 12:15:37 GMT
View Forum Message <> Reply to Message

thanx for explanation, but i still do not understand something

look:
VE net conf:
```
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:2812 errors:0 dropped:0 overruns:0 frame:0
          TX packets:4462 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:544476 (544.4 KB)  TX bytes:275033 (275.0 KB)

venet0:0  Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          inet addr:172.16.0.26  P-t-P:172.16.0.26  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:193.40.142.226  P-t-P:193.40.142.226  Bcast:0.0.0.0  Mask:255.255.255.255
          UP BROADCAST POINTOPOINT RUNNING NOARP  MTU:1500  Metric:1
```

now i try to ping internal network:
```
root@monitor:/# ping 172.16.0.254
PING 172.16.0.254 (172.16.0.254) 56(84) bytes of data.
64 bytes from 172.16.0.254: icmp_seq=1 ttl=63 time=2.22 ms
```
ok

now i try to ping external network:
```
root@monitor:/# ping 212.7.0.1
PING 212.7.0.1 (212.7.0.1) 56(84) bytes of data.
```

--- 212.7.0.1 ping statistics ---
nothing
now i try to ping, using the interface with external addres:

root@monitor:/# ping -I 193.40.142.226 212.7.0.1
PING 212.7.0.1 (212.7.0.1) from 193.40.142.226 : 56(84) bytes of data.
64 bytes from 212.7.0.1: icmp_seq=1 ttl=59 time=1.87 ms

why
