Subject: *BUG REPORTED* Network issues inside of VE
Posted by jbreck on Mon, 31 Dec 2007 01:38:31 GMT
View Forum Message <> Reply to Message

I'm using OpenVZ on a CentOS box with kernel:

2.6.18-8.1.8.el5.028stab039.1
(Also using lxlabs hypervm as a management interface)

I'm having a problem where VE's created with the default CentOS(And our custom templates) can ping the Hardware node, but cannot ping out. In addition, they don't show up in the the hardware nodes arp. This is very puzzling.

Here is an example of a working (Ubuntu, 66.103.230.228) VE:

root@ubuntufix:/# ping fark.com
PING fark.com (207.246.126.10) 56(84) bytes of data.

root@ubuntufix:/# 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

Non working node(66.103.230.240):

[root@btimberlake /]# ping fark.com
(hangs)

[root@btimberlake /]# ping 69.39.86.190 <--- Hardware node IP
PING 69.39.86.190 (69.39.86.190) 56(84) bytes of data.
64 bytes from 69.39.86.190: icmp_seq=1 ttl=64 time=0.059 ms
64 bytes from 69.39.86.190: icmp_seq=2 ttl=64 time=0.048 ms

[root@btimberlake /]# route -n
Kernel IP routing table

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
192.0.2.0	0.0.0.0	255.255.255.0	U	0	0	0	venet0
169.254.0.0	0.0.0.0	255.255.0.0	U	0	0	0	venet0
0.0.0.0	192.0.2.1	0.0.0.0	UG	0	0	0	venet0

Interesting, I see arps for all of the working VE's on the box, but not the non working one... I also notice the routes are different, but we are perplexed as to why this is... the box and images where all working a couple of weeks ago. We did update vzctl from 3.0.18 to 3.0.22 to fix a problem with debian and ubuntu images, which no work, but it appears CentOS is now broken (although I'm not sure that's the cause.)

Can anyone assist me in troubleshooting this? I can provide any information you need, I just don't know what to look for.

Subject: Re: Network issues inside of VE
Posted by ugob on Mon, 31 Dec 2007 01:48:01 GMT
View Forum Message <> Reply to Message

Looks like IP forwarding is not on. Did you set sysctl.conf correctly?

Subject: Re: Network issues inside of VE
Posted by jbreck on Mon, 31 Dec 2007 02:16:40 GMT
View Forum Message <> Reply to Message

I believe so... from host:

sysctl.conf:

net.ipv4.ip_forward = 1

and..

```
[root@a2vps2 ~]# sysctl -A | grep forward
net.ipv6.conf.default.forwarding = 0
net.ipv6.conf.all.forwarding = 0
net.ipv6.conf.eth0.forwarding = 0
net.ipv6.conf.lo.forwarding = 0
net.ipv4.conf.venet0.mc_forwarding = 0
net.ipv4.conf.venet0.forwarding = 1
net.ipv4.conf.eth0.mc_forwarding = 0
net.ipv4.conf.eth0.forwarding = 1
net.ipv4.conf.lo.mc_forwarding = 0
net.ipv4.conf.lo.forwarding = 1
net.ipv4.conf.default.mc_forwarding = 0
net.ipv4.conf.default.forwarding = 1
net.ipv4.conf.all.mc_forwarding = 0
net.ipv4.conf.all.forwarding = 1
net.ipv4.ip_forward = 1
```

It does work from some of the VE's on this box... just not the CentOS images all of a sudden.

Subject: Re: Network issues inside of VE
Posted by maratrus on Tue, 01 Jan 2008 19:14:00 GMT
View Forum Message <> Reply to Message

Hi,

Happy New Year!

1. Quote:Interesting, I see arps for all of the working VE's on the box, but not the non working one...
Does it mean that "arp -n" output from HN contains information about all VEs besides non working ones?
Can you please show the "arp -n" output from HN?

2. Also "ip a l" and "ip ro l" from HN (when non working VEs are on).
3. Do you set a nameserver to your VE? Can you ping fark.com using it ip?
4. You can use "tcpdump" to determine the packages behaviour.
When you ping the 207.246.126.10 from the VE please provide us with such information:
- Are ping requests exit the VE?
- Are ping requests come to the HN?
- Are ping requests exit the HN?
- Is ping reply comes to HN?

Thank You!

Subject: Re: Network issues inside of VE
Posted by Avi Brender on Thu, 03 Jan 2008 01:43:05 GMT
View Forum Message <> Reply to Message

Hi There,

 We're having this issue with CentOS as well (latest ovz kernel and utils).

The malfunctioning VPS ip is NOT listed in arp -a but it IS listed in 'ip ro l'

This is happening only when the main ip of the vps is on a different subnet from the HN main ip. If we change the IP to be one in the subnet of the HN then it works.

Our debugging with tcpdump indicates that if the VPS pings a foreign IP, the following happens:

1) Traffic goes from VPS to HN
2) Traffic goes from HN to destination foreign IP
3) Foreign IP returns traffic to network and the "ARP TELL" for the non-working IP hits the HN but the HN doesn't reply.   Again, the entry is missing from 'arp -a' but is in 'ip ro l'

Subject: Re: Network issues inside of VE
Posted by Avi Brender on Thu, 03 Jan 2008 02:28:48 GMT
View Forum Message <> Reply to Message

Okay, we did some digging and found the bug in the latest version of vzctl scripts

in /usr/lib/vzctl/scripts/vps-functions on roughly line 126 (we did some debugging so we may have added/removed a few lines) there is:

    dev="$(${IP_CMD} route get "$1" |sed -ne '/ via /! s/^.* dev \([^]\+\) \+src \+.*$/\1/p;Q')"

Unfortunately that doesn't parse properly when the argument $1 is outside of the HN main ip range

(HN = Centos 4.6)

Downgrading to the previous version of vzctl solved it.

Subject: Re: Network issues inside of VE
Posted by vaverin on Thu, 03 Jan 2008 07:51:11 GMT
View Forum Message <> Reply to Message

Could you please describe the issue in more details (some example would be great) and file the bug to openVZ bugzilla:
http://bugzilla.openvz.org/

thank you for bugreport,
Vasily Averin

Subject: Re: Network issues inside of VE
Posted by Avi Brender on Sat, 05 Jan 2008 04:45:56 GMT
View Forum Message <> Reply to Message

Done http://bugzilla.openvz.org/show_bug.cgi?id=771

Thanks

Subject: Re: *BUG REPORTED* Network issues inside of VE
Posted by jbreck on Sun, 06 Jan 2008 02:28:54 GMT
View Forum Message <> Reply to Message

Interesting.

Oddly enough, despite all the warnings not to enable proxy_arp... we did enable it, an found it fixed it. Previously, we had an ugly manually entering arp entries into proxy that was our workaround.

Subject: Re: *BUG REPORTED* Network issues inside of VE
Posted by Lorddusty on Sun, 06 Jan 2008 20:30:00 GMT
View Forum Message <> Reply to Message

Hi,

I ran into the same problem today. Well, there is a simple way to fix this issue, which is for me as network-engineer the only and best way:

Simply set up a device route for the network your VEs are in:

For example:
ip route add 192.168.0.0/24 dev eth0

A network you can reach without gateway you always should try to reach directly. Else the router would send you a icmp-redirect anyways.

This way open-vz knows which device to set the arp entry at and you're pretty fine.

Kind regards
Dusty

Subject: Re: *BUG REPORTED* Network issues inside of VE
Posted by kir on Wed, 09 Jan 2008 15:32:33 GMT
View Forum Message <> Reply to Message

If you want old (vzctl <= 3.0.18) behavior) which uses all interfaces, set NEIGHBOUR_DEVS=all in /etc/vz/vz.conf

Here is the description of this option, from vz.conf(8) man page:

Quote:NEIGHBOUR_DEVS[="detect"]

Controls on which interfaces to add/remove ARP records for a VE IP, also which interfaces to use to query/announce ARP. If this is set to detect, the right network interface (the one which is in the

same subnet as a VE IP) will be chosen automatically. Any other value restores old (as of vzctl 3.0.19 or older) behavior, when all the possible interfaces were used.