Subject: ARP queris being ignored
Posted by chris on Wed, 23 Apr 2008 06:07:28 GMT

View Forum Message <> Reply to Message

Hi All,

I have been using OpenVZ for a few months now, and I've been having a problem with VE's running for longer then a few days - the hardware node ignores ARP queries for VE's that are running on it. The VE has IP 192.168.0.105.

When things are working, I see a line from 'arp -a' for the 192.168.0.105 host:
```
  [root@sonata ~]# arp -a
  ? (192.168.0.160) at 00:07:E9:5F:BA:60 [ether] on eth0
  ? (192.168.0.253) at 00:1B:2B:2C:C3:4D [ether] on eth0
  ? (192.168.0.105) at * PERM PUP on eth0
```

When things stop working, the 192.168.0.105 disappears from the output of 'arp -a'. An entry for 192.168.0.105 stops appearing in /proc/net/arp as well.

A tcpdump show the arp query arriving on the hardware node.
```
  [root@sonata ~]# tcpdump -nni any arp
  tcpdump: WARNING: Promiscuous mode not supported on the "any" device
  tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
  listening on any, link-type LINUX_SLL (Linux cooked), capture size 96 bytes
  15:30:19.964806 arp who-has 192.168.0.105 tell 192.168.0.253
  15:30:21.973385 arp who-has 192.168.0.105 tell 192.168.0.253
  15:30:23.973453 arp who-has 192.168.0.105 tell 192.168.0.253
  15:30:25.976838 arp who-has 192.168.0.105 tell 192.168.0.253
```

Further info:
 - VE is running with ip 192.168.0.105
```
   [root@sonata ~]# vzlist |grep 192.168.0.105
        105        125 running 192.168.0.105   zimbra-ose.cgb1911.mine.nu
```

 - Centos 5 with a OpenVZ kernel from the yum repo
```
   [root@sonata ~]# cat /etc/redhat-release
   CentOS release 5 (Final)
   [root@sonata ~]# uname -a
   Linux sonata.cgb1911.mine.nu 2.6.18-53.1.13.el5.028stab053.10 #1 SMP Tue Apr 1 14:58:47
MSD 2008 i686 i686 i386 GNU/Linux
```

Anyone have any idea what is going on? Is there something further I can do to debug this?

Thanks in advance for any responses,

Chris Bennett (cgb)

Subject: Re: ARP queris being ignored
View Forum Message <> Reply to Message

On Wed, 2008-04-23 at 15:37 +0930, chris@cgb1911.mine.nu wrote:
> Hi All,
>
> I have been using OpenVZ for a few months now, and I've been having a
> problem with VE's running for longer then a few days - the hardware
> node ignores ARP queries for VE's that are running on it. The VE has
> IP 192.168.0.105.
>
> When things are working, I see a line from 'arp -a' for the 192.168.0.105
> host:
>   [root@sonata ~]# arp -a
>   ? (192.168.0.160) at 00:07:E9:5F:BA:60 [ether] on eth0
>   ? (192.168.0.253) at 00:1B:2B:2C:C3:4D [ether] on eth0
>   ? (192.168.0.105) at * PERM PUP on eth0
>
> When things stop working, the 192.168.0.105 disappears from the output
> of 'arp -a'. An entry for 192.168.0.105 stops appearing in
> /proc/net/arp as well.
>
> A tcpdump show the arp query arriving on the hardware node.
>   [root@sonata ~]# tcpdump -nni any arp
>   tcpdump: WARNING: Promiscuous mode not supported on the "any" device
>   tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
>   listening on any, link-type LINUX_SLL (Linux cooked), capture size 96 bytes
>   15:30:19.964806 arp who-has 192.168.0.105 tell 192.168.0.253
>   15:30:21.973385 arp who-has 192.168.0.105 tell 192.168.0.253
>   15:30:23.973453 arp who-has 192.168.0.105 tell 192.168.0.253
>   15:30:25.976838 arp who-has 192.168.0.105 tell 192.168.0.253

you'll have arp replay iff on the _target_ node/VE
ip route get 192.168.0.105 from 192.168.0.253 iif <incoming_device>
will return either a local route or route to a device other than
incoming.

Regards,
 Den

Subject: Re: ARP queris being ignored

Posted by chris on Wed, 23 Apr 2008 13:18:48 GMT

Hi Den,

Thanks for your reply.

> you'll have arp replay iff on the _target_ node/VE
> ip route get 192.168.0.105 from 192.168.0.253 iif <incoming_device>
> will return either a local route or route to a device other than
> incoming.

I don't really get you on this, are you able to elaborate?

Also, if it helps, here's the output of command you suggested above:

Working:
  [root@sonata ~]# ip route get 192.168.0.105
  192.168.0.105 dev venet0 src 192.168.0.224
     cache mtu 1500 advmss 1460 hoplimit 64

Not working:
  [root@sonata ~]# ip route get 192.168.0.105
  192.168.0.105 dev venet0 src 192.168.0.224
     cache mtu 1500 advmss 1460 hoplimit 64

That is, they are exactly the same.

Regards,

Chris Bennett (cgb)

Subject: Re: ARP queris being ignored
Posted by den on Wed, 23 Apr 2008 19:39:58 GMT

On Wed, 2008-04-23 at 22:48 +0930, chris@cgb1911.mine.nu wrote:
> Hi Den,
>
> Thanks for your reply.
>
> > you'll have arp replay iff on the _target_ node/VE
> > ip route get 192.168.0.105 from 192.168.0.253 iif <incoming_device>
> > will return either a local route or route to a device other than
> > incoming.
>
> I don't really get you on this, are you able to elaborate?

>
> Also, if it helps, here's the output of command you suggested above:
>
> Working:
>   [root@sonata ~]# ip route get 192.168.0.105
>   192.168.0.105 dev venet0 src 192.168.0.224
>   cache mtu 1500 advmss 1460 hoplimit 64
>
> Not working:
>   [root@sonata ~]# ip route get 192.168.0.105
>   192.168.0.105 dev venet0 src 192.168.0.224
>   cache mtu 1500 advmss 1460 hoplimit 64

You are using a wrong command. If you are using venet0 you should
execute the command _exactly_ in the form I ask, with "from" and "iif",
in VE0.

There should be a difference.

Regards,
 Den

Hi Again,

> You are using a wrong command. If you are using venet0 you should
> execute the command _exactly_ in the form I ask, with "from" and "iif",
> in VE0.

Ok, I understand you on this now..   I wasn't cloear on what is the
incoming_device, (eth0 or venet0) so I provided both:

eth0 working:
  [root@sonata ~]# ip route get 192.168.0.105 from 192.168.0.253/24 iif eth0
  192.168.0.105 from 192.168.0.253 dev venet0 src 192.168.0.224
      cache <src-direct> mtu 1500 advmss 1460 hoplimit 64 iif eth0

eth0 not working:
  [root@sonata ~]# ip route get 192.168.0.105 from 192.168.0.253/24 iif eth0
  192.168.0.105 from 192.168.0.253 dev venet0 src 192.168.0.224
      cache <src-direct> mtu 1500 advmss 1460 hoplimit 64 iif eth0

venet0 working:

```
[root@sonata ~]# ip route get 192.168.0.105 from 192.168.0.253/24 iif venet0
 192.168.0.105 from 192.168.0.253 dev venet0  src 192.168.0.224
    cache  mtu 1500 advmss 1460 hoplimit 64 iif venet0
```

venet0 not working:
```
 [root@sonata ~]# ip route get 192.168.0.105 from 192.168.0.253/24 iif venet0
 192.168.0.105 from 192.168.0.253 dev venet0  src 192.168.0.224
    cache  mtu 1500 advmss 1460 hoplimit 64 iif venet0
```

> There should be a difference.

In both cases, they are the same. I don't think I mentioned it in the
first email - restarting the VE fixes the problem.

Is there anything else you can suggest I try?

Regards,

Chris Bennett (cgb)

Subject: Re: ARP queris being ignored
Posted by Christian Hofstaedtle on Fri, 25 Apr 2008 07:54:45 GMT
View Forum Message <> Reply to Message

Hello,

* chris@cgb1911.mine.nu <chris@cgb1911.mine.nu> [080423 23:36]:
> [ip route stuff]
> > There should be a difference.
>
> In both cases, they are the same. I don't think I mentioned it in the
> first email - restarting the VE fixes the problem.
>
> Is there anything else you can suggest I try?

Try: arp -vn

It should show a line with your VE IP address, like this:

10.10.150.244   *   <from_interface>   MP   eth0

If you don't have such a line, the host node will not respond to ARP
queries (except if you turn on proxy_arp, but I'd advise against
this).

The VE start script sets this up (so a restart will fix it), but
I've also seen cases where these entries get lost after some time.

Haven't been able to produce a test case, though.

Any chance that you are using a firewall on the host node which
fiddles with the ARP stuff? (shorewall?)

Christian

Subject: Re: ARP queris being ignored
Posted by chris on Sat, 26 Apr 2008 15:08:10 GMT
View Forum Message <> Reply to Message

Hi Christian,

I didn't see your post on openvz-users (I only got the personally
addressed copy). I'll address this to openvz-users. Thanks for your
reply and ideas also.

> Try: arp -vn
>
> It should show a line with your VE IP address, like this:
>
> 10.10.150.244   *   <from_interface>   MP   eth0
>
> If you don't have such a line, the host node will not respond to ARP
> queries (except if you turn on proxy_arp, but I'd advise against
> this).

Unfortunately things still seem wrong. To illustrate, here's what I
get for vzlist:

[root@sonata ~]# vzlist
     VEID   NPROC STATUS IP_ADDR   HOSTNAME
      10   68 running 192.168.0.10   trixbox.cgb1911.mine.nu
     105   124 running 192.168.0.105   zimbra-ose.cgb1911.mine.nu
     106   24 running 192.168.0.106   otrs.cgb1911.mine.nu
     107   26 running 192.168.0.107   ipplan.cgb1911.mine.nu
     108   6 running 192.168.0.108   freeside.cgb1911.mine.nu
     109   8 running 192.168.0.109   enomalism.cgb1911.mine.nu
     110   65 running 192.168.0.110   virtualmin.cgb1911.mine.nu
     111   77 running 192.168.0.111   ispconfig.cgb1911.mine.nu
     112   33 running 192.168.0.112   splunk.cgb1911.mine.nu
     115   77 running 192.168.0.115   zimbra4.cgb1911.mine.nu
     116   28 running 192.168.0.116   cacti-test.cgb1911.mine.nu
  100001   1 running -   -

Now, you'd expect exactly that same number of entries in the arp

table, one for each IP. Unfortunately, only entries for 105 and 10
exist (they are ones I've had to restart to 'make work').

```
[root@sonata ~]# arp -vn
Address              HWtype  HWaddress        Flags Mask          Iface
192.168.0.221          ether   00:11:93:98:67:45   C              eth0
192.168.0.160          ether   00:07:E9:5F:BA:60   C              eth0
192.168.0.253          ether   00:1B:2B:2C:C3:4D   C              eth0
192.168.0.43           ether   00:19:D1:69:FD:2E   C              eth0
192.168.0.105          *       *                MP              eth0
192.168.0.10           *       *                MP              eth0
Entries: 6     Skipped: 0     Found: 6
```

> The VE start script sets this up (so a restart will fix it), but
> I've also seen cases where these entries get lost after some time.
> Haven't been able to produce a test case, though.

Yeah, I hadn't determined what triggers the problem as yet. Some
further digging has revealed that a script:

 /usr/share/vzctl/scripts/vpsnetclean

runs every 5 minutes in cron. It has all the characteristics of a
program that would be breaking arp entires for my running VE's. I
just need to work out if and then why it determines a VE is 'stopped'
and calls clear_ve_net to tidy up IP/ARP entries. Will reply if I
have anything more.

> Any chance that you are using a firewall on the host node which
> fiddles with the ARP stuff? (shorewall?)

Nope, not using anything like that. Just standard Centos 5 base
system with OVZ kernel and management tools.

Regards,

Chris Bennett (cgb)