Subject: routing from external through HN to loopback on CT
Posted by Christopher McCrory on Fri, 20 Jul 2012 16:25:07 GMT
View Forum Message <> Reply to Message

Hello...

The setup:
x.y.1.1/24 Cisco router
x.y.1.2/24 OpenVZ Hardware Node on CentOS 5 with all updates
x.y.1.3/24 OpenVZ container CT103 using veth
x.y.42.42/32 IP address on container's loopback interface
iptables off on all hosts, ip forwarding enabled everywhere
the router/external host has a route for x.y.42.42/32 via x.y.1.3

I cannot get packets from external hosts to route to the loopback on CT103. packets leave the router, enter the HN, but never exit any interface. This should be working but isn't and the same setup not on a container works. What am I missing?

in the openvz/103.conf file this works
IP_ADDRESS="x.y.1.3 x.y.42.42"

but will not work as I need x.y.42.42 to be on several containers (anycast DNS) and advertised via BGP. I also need to create more CTs that will have a hundreds of IP addresses on loopbacks and putting them in the openvz config will not scale.

Again, what am I missing?

thanks

--
Christopher McCrory
To the optimist, the glass is half full.
To the pessimist, the glass is half empty.
To the engineer, the glass is twice as big as it needs to be.

Subject: Re: routing from external through HN to loopback on CT
Posted by coolcold on Fri, 20 Jul 2012 17:02:17 GMT
View Forum Message <> Reply to Message

On Fri, Jul 20, 2012 at 8:25 PM, Christopher McCrory <chrismcc@gmail.com>wrote:

> Hello...
>

> The setup:
> x.y.1.1/24   Cisco router
> x.y.1.2/24   OpenVZ Hardware Node on CentOS 5 with all updates
> x.y.1.3/24   OpenVZ container CT103 using veth
> x.y.42.42/32 IP address on container's loopback interface
> iptables off on all hosts, ip forwarding enabled everywhere
> the router/external host has a route for x.y.42.42/32 via x.y.1.3
>
> I cannot get packets from external hosts to route to the loopback on
> CT103. packets leave the router, enter the HN, but never exit any
> interface. This should be working but isn't and the same setup not on a
> container works. What am I missing?
>
>
>
> in the openvz/103.conf file this works
> IP_ADDRESS="x.y.1.3 x.y.42.42"
>
> but will not work as I need x.y.42.42 to be on several containers
> (anycast DNS) and advertised via BGP. I also need to create more CTs
> that will have a hundreds of IP addresses on loopbacks and putting them
>

> in the openvz config will not scale.
>
> Again, what am I missing?
>
I think problem is in routing table on HN - as you have that x.y.42.42 on
loopback of VE, routing on HN has no idea what to do with them.

>
> thanks
>
> --
> Christopher McCrory
> To the optimist, the glass is half full.
> To the pessimist, the glass is half empty.
> To the engineer, the glass is twice as big as it needs to be.
>
--
Best regards,
[COOLCOLD-RIPN]

Subject: Re: routing from external through HN to loopback on CT

On Fri, 2012-07-20 at 21:02 +0400, CoolCold wrote:
>
>
> On Fri, Jul 20, 2012 at 8:25 PM, Christopher McCrory
> <chrismcc@gmail.com> wrote:
> Hello...
>
> The setup:
> x.y.1.1/24 Cisco router
> x.y.1.2/24 OpenVZ Hardware Node on CentOS 5 with all
> updates
> x.y.1.3/24 OpenVZ container CT103 using veth
> x.y.42.42/32 IP address on container's loopback interface
> iptables off on all hosts, ip forwarding enabled everywhere
> the router/external host has a route for x.y.42.42/32 via
> x.y.1.3
>
> I cannot get packets from external hosts to route to the
> loopback on
> CT103. packets leave the router, enter the HN, but never exit
> any
> interface. This should be working but isn't and the same
> setup not on a
> container works. What am I missing?
>
>
>
> in the openvz/103.conf file this works
> IP_ADDRESS="x.y.1.3 x.y.42.42"
>
> but will not work as I need x.y.42.42 to be on several
> containers
> (anycast DNS) and advertised via BGP. I also need to create
> more CTs
> that will have a hundreds of IP addresses on loopbacks and
> putting them
>
> in the openvz config will not scale.
>
> Again, what am I missing?
> I think problem is in routing table on HN - as you have that x.y.42.42
> on loopback of VE, routing on HN has no idea what to do with them.
>

Adding a route on the HN does not help (I tried that). Nor should it be

required. The packet should be bridged through the HN to CT103. Then CT103 knows that x.y.42.42 is on itself and can process the packets.
>From what I see using tcpdump the packet never leaves the bridge on the HN. ? ? ?

>
> thanks
>
> --
> Christopher McCrory
> To the optimist, the glass is half full.
> To the pessimist, the glass is half empty.
> To the engineer, the glass is twice as big as it needs to be.
>
> --
> Best regards,
> [COOLCOLD-RIPN]
--
Christopher McCrory
To the optimist, the glass is half full.
To the pessimist, the glass is half empty.
To the engineer, the glass is twice as big as it needs to be.

Subject: Re: routing from external through HN to loopback on CT [solved]
Posted by Christopher McCrory on Mon, 23 Jul 2012 21:57:10 GMT
View Forum Message <> Reply to Message

Hello...

I got this working now. Thanks for all the responses.

This was the platform and problem(s):

I was using two different platforms, one is a proxmox/ubuntu setup and a older pure openvz setup from several months ago on CentOS 5.x. In both cases I was using containers that were already running and I was adding loopback interfaces.

In the docs on the openvz site about using veth instead of venet there are a couple places that (for me at least) have ambiguous terms. CT, CT0 (hardware node) vs CTxx (virtual container) and what/where eth0 means.

So what I ended up with was correctly seeing the vethxx.y device on the

hardware node, but seeing venet on the virtual containers with the ip addresses originally setup. What is not stated in the veth docs is that you must remove any existing ip addresses from the vz${CTID}.conf file. So I had a hybrid veth/venet setup.

suggestions:

from http://wiki.openvz.org/Virtual_Ethernet_device
ifname is the Ethernet device name in the CT
mac is its MAC address in the CT
host_ifname is the Ethernet device name on the host (CT0)
host_mac is its MAC address on the host (CT0), if

maybe the above should be:

ifname is the Ethernet device name in the CT
mac is its MAC address in the CT
host_ifname is the Ethernet device name on the host (HN)
host_mac is its MAC address on the host (HN), if

And maybe a line or two about removing any ip addresses if the CT used venet previously.

or maybe 'vzctl start ${CTID}' could output a warning if vz${CTID}.conf contains both venet (IP_ADDRESS=...) and veth (NETIF=...) stanzas. Or even better setup the IP bits in the
correct /etc/sysconfig/network-scripts/ifcfg-eth0 (for rhel/fedora/centos) instead
of /etc/sysconfig/network-scripts/ifcfg-venet0

thanks again for the help.

On Fri, 2012-07-20 at 21:02 +0400, CoolCold wrote:
>
>
> On Fri, Jul 20, 2012 at 8:25 PM, Christopher McCrory
> <chrismcc@gmail.com> wrote:
> Hello...
>
> The setup:
> x.y.1.1/24 Cisco router

> x.y.1.2/24 OpenVZ Hardware Node on CentOS 5 with all
> updates
> x.y.1.3/24 OpenVZ container CT103 using veth
> x.y.42.42/32 IP address on container's loopback interface
> iptables off on all hosts, ip forwarding enabled everywhere
> the router/external host has a route for x.y.42.42/32 via
> x.y.1.3
>
> I cannot get packets from external hosts to route to the
> loopback on
> CT103. packets leave the router, enter the HN, but never exit
> any
> interface. This should be working but isn't and the same
> setup not on a
> container works. What am I missing?
>
>
>
> in the openvz/103.conf file this works
> IP_ADDRESS="x.y.1.3 x.y.42.42"
>
> but will not work as I need x.y.42.42 to be on several
> containers
> (anycast DNS) and advertised via BGP. I also need to create
> more CTs
> that will have a hundreds of IP addresses on loopbacks and
> putting them
>
> in the openvz config will not scale.
>
> Again, what am I missing?
> I think problem is in routing table on HN - as you have that x.y.42.42
> on loopback of VE, routing on HN has no idea what to do with them.
>
>
> thanks
>
> --
> Christopher McCrory
> To the optimist, the glass is half full.
> To the pessimist, the glass is half empty.
> To the engineer, the glass is twice as big as it needs to be.
>
> --
> Best regards,
> [COOLCOLD-RIPN]
--
Christopher McCrory

To the optimist, the glass is half full.
To the pessimist, the glass is half empty.
To the engineer, the glass is twice as big as it needs to be.
