Subject: Network settings - differnet address range for containers
Posted by naro on Fri, 11 Mar 2011 19:17:22 GMT
View Forum Message <> Reply to Message

Hello,

I'm trying to set up public IP for my container. HN has address XXX.71.138.11 (eth0). Container should get address from range:
YYY.132.159.194 - 222 (gateway is set up by ISP to YYY.132.159.193 and is sending packets to my HW node).

I tried to set YYY.132.159.200 to HN as eth0:0 and it works fine, it means, ISP routing is fine.

There are no iptables rules.

I've assigned YYY.132.159.194 to container and now I need to set up routing on hardware node. I checked Source based routing wiki, added the rule, but

/sbin/ip route add default dev eth0 via YYY.132.159.193 table 10
failed with RTNETLINK: No such file or directory. This is caused by missing YYY.132.159.193 route in the main table probably, but I don't know how to set it up.

HN# ip route list
YYY.132.159.194 dev venet0 scope link
XXX.71.138.0/27 dev eth0 proto kernel scope link src XXX.71.138.11
default via XXX.71.138.1 dev eth0

HN has assigned XXX.71.138.11 on eth0 only. No other public address.

How should I setup routing on HN to correctly pass packets from YYY.132.159.194-222 to YYY.132.159.193 via eth0 ?

Thanks

Edit:
forgot to mention sysctl settings:
net.ipv4.ip_forward = 1
net.ipv4.conf.default.proxy_arp = 0
net.ipv4.conf.all.rp_filter = 1
net.ipv4.conf.default.send_redirects = 1
net.ipv4.conf.all.send_redirects = 0

I have tried to assign 192.168.200.101 to container and configured iptables postrouting on HN and it works fine, but I'm still not able to get public address work.

Subject: Re: Network settings - differnet address range for containers

Posted by nuno on Tue, 15 Mar 2011 07:28:53 GMT
View Forum Message <> Reply to Message

Maybe it's easier to setup a bridge. That way all the configuration is done inside the container.

Don't forget to install brctl (bridge-utils) in HN.

Regards,

Subject: Re: Network settings - differnet address range for containers
Posted by naro on Tue, 15 Mar 2011 07:58:39 GMT
View Forum Message <> Reply to Message

I solved my problem by modifying ARP table on HN:

arp -i eth0 -Ds YYY.132.159.194 eth0 pub

YYY.132.159.194 is IP address of my container.

I wrote a vps.mount/vps.umount hook scripts which adds record to the arp table on VPS start and removes it on VPS shutdown. Works fine for all containers (the script reads particular IP from IP_ADDRESS environment variable).

Subject: Re: Network settings - differnet address range for containers
Posted by pilpelet on Thu, 04 Aug 2011 21:18:11 GMT
View Forum Message <> Reply to Message

Hi,
Would be great if you can post that scripts.
We also need to solve the arp issue when working with 2 subnets.

Tnx,

Subject: Re: Network settings - differnet address range for containers
Posted by curx on Tue, 09 Aug 2011 22:21:25 GMT
View Forum Message <> Reply to Message

Hi,

naro wrote on Tue, 15 March 2011 08:58I solved my problem by modifying ARP table on HN:

arp -i eth0 -Ds YYY.132.159.194 eth0 pub

YYY.132.159.194 is IP address of my container.

I wrote a vps.mount/vps.umount hook scripts which adds record to the arp table on VPS start and removes it on VPS shutdown. Works fine for all containers (the script reads particular IP from IP_ADDRESS environment variable).

can you please show the OpenVZ config option NEIGHBOUR_DEVS

% grep NEIGHBOUR_DEVS /etc/vz/vz.conf

Bye,
Thorsten

Subject: Re: Network settings - differnet address range for containers
Posted by naro on Wed, 10 Aug 2011 06:26:42 GMT
View Forum Message <> Reply to Message

Please replace 11.22.33 with your own common part of IP address range. This is just raw string comparsion so no subnet checking etc.

There is one problem I did not solve yet. The ARP cache. If you move VPS to another hardware node, you must wait about 10 minutes until ARP cache on the gateway is cleared. The gateway is not under my control so I can't clear the cache manually

File: /etc/vz/conf/vps.mount

```
#!/bin/bash

# This script source VPS configuration files in the same
# order as vzctl does
# if one of these files does not exist then something is
# really broken

[ -f /etc/vz/vz.conf ] || exit 1
[ -f $VE_CONFFILE ] || exit 1

# source both files. Note the order, it is important
. /etc/vz/vz.conf
. $VE_CONFFILE

if [ -z "$IP_ADDRESS" ]
then
   # no IP for this container
   echo "No IP address for this container"
   exit 0
```

```
fi

for IP in $IP_ADDRESS
do
   if `echo $IP | grep "11.22.33" 1>/dev/null 2>&1`
   then
      # this container uses our IP address range. Check ARP
      if `arp -n | grep $IP 1>/dev/null 2>&1`
      then
         # command above return 1 if item is not found
         echo "$IP is already in ARP table. Weird."
      else
         # IP is not in local ARP table. Add. it
         echo "Adding $IP to ARP table"
         arp -i eth0 -Ds $IP eth0 pub
      fi
   fi
done

exit 0
```

File: /etc/vz/conf/vps.umount

```
#!/bin/bash

# This script source VPS configuration files in the same
# order as vzctl does
# if one of these files does not exist then something is
# really broken

[ -f /etc/vz/vz.conf ] || exit 1
[ -f $VE_CONFFILE ] || exit 1

# source both files. Note the order, it is important
. /etc/vz/vz.conf
. $VE_CONFFILE

if [ -z "$IP_ADDRESS" ]
then
   # no IP for this container
   echo "No IP address for this container"
   exit 0
fi

for IP in $IP_ADDRESS
do
```

```
    if `echo $IP | grep "11.22.33" 1>/dev/null 2>&1`
    then
      # this container uses our IP address. Check ARP
      if `arp -n | grep $IP 1>/dev/null 2>&1`
      then
         # IP is in local ARP table. Remove it
         echo "Removing $IP from ARP table"
         arp -i eth0 -d $IP
      else
         echo "$IP is not in ARP table. Weird."
      fi
    fi
done

exit 0
```

Subject: Re: Network settings - differnet address range for containers
Posted by naro on Wed, 10 Aug 2011 06:28:11 GMT
View Forum Message <> Reply to Message

curx wrote on Tue, 09 August 2011 18:21Hi,

can you please show the OpenVZ config option NEIGHBOUR_DEVS

NEIGHBOUR_DEVS=detect

Subject: Re: Network settings - differnet address range for containers
Posted by pilpelet on Wed, 10 Aug 2011 11:00:31 GMT
View Forum Message <> Reply to Message

Hi naro,

Thanks for the script, i was wondering how did you solve the routing issue? im having the same issue and the max i could achive is pinging from the container out but not from the world in.
I did that by adding
iptables -t nat -A POSTROUTING -d 62.90.150.0/26 -j MASQUERADE -o eth0
on NH
where 62.90.150.0/26 is the subnet that doesnt configured on the NH eth0.
I still not fully understand the routing procedure special the default gateway/route on the NH specialy when its not possible to add default gayway on container working with venet.

maybee the ARP table on NH will solve the issue(it does make sense) i havnt tested it yet.