Subject: OpenVZ scaling - Advanced network concepts
Posted by kingneutron on Tue, 29 May 2007 06:24:01 GMT
View Forum Message <> Reply to Message

See: [http://forum.openvz.org/index.php?t=rview&goto=13504#msg _13504]

--Ok, so far (with a LOT of experimenting and head-banging-against-the-wall) I have been completely able to get by without using iptables/NAT on the Linux host. What I've got looks like this:

NETWORK TOPOLOGY:

o Cable modem (DHCP from ISP)

oo Edge router / NAT box / DHCP server (D-Link box)
oo Acts as DNS Nameserver @ 192.168.2.1
oo Local DHCP-assigned net 192.168.2.0 // 255.255.255.0

ooo Squid cache running on laptop @ 192.168.2.250 and 10.0.0.4

ooo Static local net 10.0.0.0 // 255.0.0.0 for "my stuff"
(server, laptops, etc)

ooo OpenVZ host is connected thru Switches:
o eth0 @ 10.0.0.3 and
o eth1 @ DHCP + 192.168.2.226 static

========

Host ifconfig:

```
eth0      Link encap:Ethernet  HWaddr
          inet addr:10.0.0.3  Bcast:10.255.255.255  Mask:255.0.0.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:28227 errors:0 dropped:0 overruns:0 frame:0
          TX packets:12626 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:28030407 (26.7 MiB)  TX bytes:1146267 (1.0 MiB)
          Base address:0x2000 Memory:92100000-92120000

eth1      Link encap:Ethernet  HWaddr
          inet addr:192.168.2.28  Bcast:192.168.2.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:7690 errors:0 dropped:0 overruns:0 frame:0
          TX packets:422 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:2469697 (2.3 MiB)  TX bytes:45798 (44.7 KiB)
          Interrupt:185 Base address:0x2800
```

```
eth1:0    Link encap:Ethernet  HWaddr
          inet addr:192.168.2.226  Bcast:192.255.255.255  Mask:255.0.0.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          Interrupt:185 Base address:0x2800

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:25 errors:0 dropped:0 overruns:0 frame:0
          TX packets:25 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:1992 (1.9 KiB)  TX bytes:1992 (1.9 KiB)

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:2510 errors:0 dropped:0 overruns:0 frame:0
          TX packets:2384 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:139142 (135.8 KiB)  TX bytes:366027 (357.4 KiB)

veth101.0 Link encap:Ethernet  HWaddr 00:12:34:56:78:01
          inet addr:172.16.0.1  Bcast:172.16.255.255  Mask:255.255.0.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:259 errors:0 dropped:0 overruns:0 frame:0
          TX packets:52 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:24692 (24.1 KiB)  TX bytes:3360 (3.2 KiB)


BEGIN VE guest 101 /etc/network/interfaces

# Auto generated venet0 interfaces
auto venet0 lo eth0
# eth0 IS necessary for class B 172.16

iface lo inet loopback

iface eth0 inet static
        address 172.16.1.1
        netmask 255.255.0.0 up

iface venet0 inet static
        address 127.0.1.1
        netmask 255.255.255.255
        broadcast 0.0.0.0
        up route add -net 192.168.2.0 netmask 255.255.255.255 dev venet0
```

```
        up route add default gw 192.168.2.0
# ^ This is the edge router/NAT box

# Orig:
#       up route add -net 192.0.2.1 netmask 255.255.255.255 dev venet0
#       up route add default gw 192.0.2.1

auto venet0:0
#auto venet0:1
# We're gonna run out of class C addrs in a hurry if we leave this in.
# But activate it if we need name-resolution in-guest w/o going thru the
# proxy.

# Yes, these are actually necessary.  Without, cannot ping localnet.

# NOTE - venet0:0 has to match the HOST...
#   If we change .100.226 to ANYTHING else, it stops working!
# ( goes thru ((venet0)) -> [eth0] -> 10.0.0.4 )

# AHA - Because we said use .226 in vzctl, on the HOST!!!
# ( 'route -n' on host revealed this )

iface venet0:0 inet static
        address 10.0.100.226
        netmask 255.0.0.0
        broadcast 0.0.0.0
#iface venet0:1 inet static
#       address 192.168.2.227
#       netmask 255.255.255.0
#       broadcast 0.0.0.0

# For some reason this not work, even if we have eth1:1 as 172.16 on host
#iface venet0:2 inet static
#       address 172.16.1.1
#       netmask 255.255.0.0
#       broadcast 0.0.0.0


--NOTE: Guest /etc/resolv.conf IS set to 192.168.2.1 in case it's needed.

BEGIN guest 101 ' ifconfig ':

eth0      Link encap:Ethernet  HWaddr 00:12:34:56:78:9D
          inet addr:172.16.1.1  Bcast:172.16.255.255  Mask:255.255.0.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:52 errors:0 dropped:0 overruns:0 frame:0
          TX packets:259 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
```

RX bytes:3360 (3.2 KiB) TX bytes:24692 (24.1 KiB)

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:0 errors:0 dropped:0 overruns:0 frame:0
TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:0
RX bytes:0 (0.0 b) TX bytes:0 (0.0 b)

venet0 Link encap:UNSPEC HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
inet addr:127.0.1.1 P-t-P:127.0.1.1 Bcast:0.0.0.0 Mask:255.255.255.255
UP BROADCAST POINTOPOINT RUNNING NOARP MTU:1500 Metric:1
RX packets:2384 errors:0 dropped:0 overruns:0 frame:0
TX packets:2510 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:0
RX bytes:366027 (357.4 KiB) TX bytes:139142 (135.8 KiB)

venet0:0 Link encap:UNSPEC HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
inet addr:10.0.100.226 P-t-P:10.0.100.226 Bcast:0.0.0.0 Mask:255.0.0.0
UP BROADCAST POINTOPOINT RUNNING NOARP MTU:1500 Metric:1

--Using these configurations:

o Ping host 10.0.0.3 <--> guest 10.0.100.226 works, both ways
o Ping guest 172.16.1.1 <--> host 172.16.0.1 (Class B) works
o Ping from Squid proxy 10.0.0.4 -> guest 10.0.100.226 works!

o Guest can see 10.0.0.4 (squid) and obtain updates if proxy vars are set:

ftp_proxy=http://10.0.0.4:3128
http_proxy=http://10.0.0.4:3128

- Guest cannot do name-resolution on its own (cannot ping google, etc); has to go thru proxy. If needed however, this functionality can be activated on an ad-hoc basis by bringing up the class C interface "venet0:1". (This not $bug, is $feature.)

+ Guest is isolated from the Internet without explicit forwarding being set up. (This is good for security.)

+ The Squid box can ' nmap 10.0.100.226 ' (VE guest) and it reveals SSH running, as expected.

- The Squid box cannot see the 172.16 host-only network. (But this also is good.)

-- The VE guest cannot obtain a DHCP address over eth0. This was thought to _not_ be good, but can be worked around with the Class A and Class-B static IP network scheme.

--Thoughts, comments, advice? TIA