Subject: A really strange network issue
Posted by Choco52 on Thu, 16 Feb 2012 09:38:50 GMT
View Forum Message <> Reply to Message

Hello,

I already search in the forum but did not find my answer. Sorry if my question has already been asked.

I'm using OpenVZ on a Debian:

On the HN :
root@milo:/home/choco52# uname -a
Linux milo 2.6.32-5-openvz-amd64 #1 SMP Mon Oct 3 05:12:50 UTC 2011 x86_64 GNU/Linux

I have one VE (CTID = 101) :
root@localhost:/# uname -a
Linux localhost 2.6.32-5-openvz-amd64 #1 SMP Mon Oct 3 05:12:50 UTC 2011 x86_64 GNU/Linux

On 101:

root@localhost:/# ip rule list
0: from all lookup local
32766: from all lookup main
32767: from all lookup default

By the way , i don't really understand this command [if you can explain, could be great =)).

On 101 :

root@localhost:/# ifconfig
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:10 errors:0 dropped:0 overruns:0 frame:0
TX packets:10 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:0
RX bytes:696 (696.0 B) TX bytes:696 (696.0 B)

venet0 Link encap:UNSPEC HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
inet addr:127.0.0.1 P-t-P:127.0.0.1 Bcast:0.0.0.0 Mask:255.255.255.255
UP BROADCAST POINTOPOINT RUNNING NOARP MTU:1500 Metric:1
RX packets:28 errors:0 dropped:0 overruns:0 frame:0
TX packets:14 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:0
RX bytes:2736 (2.6 KiB) TX bytes:1070 (1.0 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:192.168.210.101 P-t-P:192.168.210.101 Bcast:0.0.0.0 Mask:255.255.255.255
UP BROADCAST POINTOPOINT RUNNING NOARP MTU:1500 Metric:1

On 101 :

root@localhost:/# sysctl -a | grep venet0
error: permission denied on key 'net.ipv4.route.flush'
error: permission denied on key 'net.ipv4.route.flush'
net.ipv4.neigh.venet0.mcast_solicit = 3
net.ipv4.neigh.venet0.ucast_solicit = 3
net.ipv4.neigh.venet0.app_solicit = 0
net.ipv4.neigh.venet0.retrans_time = 100
net.ipv4.neigh.venet0.base_reachable_time = 30
net.ipv4.neigh.venet0.delay_first_probe_time = 5
net.ipv4.neigh.venet0.gc_stale_time = 60
net.ipv4.neigh.venet0.unres_qlen = 3
net.ipv4.neigh.venet0.proxy_qlen = 64
net.ipv4.neigh.venet0.anycast_delay = 100
net.ipv4.neigh.venet0.proxy_delay = 80
net.ipv4.neigh.venet0.locktime = 100
net.ipv4.neigh.venet0.retrans_time_ms = 1000
net.ipv4.neigh.venet0.base_reachable_time_ms = 30000
net.ipv4.conf.venet0.forwarding = 1
net.ipv4.conf.venet0.mc_forwarding = 0
net.ipv4.conf.venet0.accept_redirects = 1
net.ipv4.conf.venet0.secure_redirects = 1
net.ipv4.conf.venet0.shared_media = 1
net.ipv4.conf.venet0.rp_filter = 0
net.ipv4.conf.venet0.send_redirects = 1
net.ipv4.conf.venet0.accept_source_route = 1
net.ipv4.conf.venet0.src_valid_mark = 0
net.ipv4.conf.venet0.proxy_arp = 0
net.ipv4.conf.venet0.medium_id = 0
net.ipv4.conf.venet0.bootp_relay = 0
net.ipv4.conf.venet0.log_martians = 0
net.ipv4.conf.venet0.tag = 0
net.ipv4.conf.venet0.arp_filter = 0
net.ipv4.conf.venet0.arp_announce = 0
net.ipv4.conf.venet0.arp_ignore = 0
net.ipv4.conf.venet0.arp_accept = 0
net.ipv4.conf.venet0.arp_notify = 0
net.ipv4.conf.venet0.disable_xfrm = 0

net.ipv4.conf.venet0.disable_policy = 0
net.ipv4.conf.venet0.force_igmp_version = 0
net.ipv4.conf.venet0.promote_secondaries = 0
error: permission denied on key 'net.ipv6.route.flush'
net.ipv6.conf.venet0.forwarding = 0
net.ipv6.conf.venet0.hop_limit = 64
net.ipv6.conf.venet0.mtu = 1500
net.ipv6.conf.venet0.accept_ra = 1
net.ipv6.conf.venet0.accept_redirects = 1
net.ipv6.conf.venet0.autoconf = 1
net.ipv6.conf.venet0.dad_transmits = 1
net.ipv6.conf.venet0.router_solicitations = 3
net.ipv6.conf.venet0.router_solicitation_interval = 4
net.ipv6.conf.venet0.router_solicitation_delay = 1
net.ipv6.conf.venet0.force_mld_version = 0
net.ipv6.conf.venet0.use_tempaddr = 0
net.ipv6.conf.venet0.temp_valid_lft = 604800
net.ipv6.conf.venet0.temp_prefered_lft = 86400
net.ipv6.conf.venet0.regen_max_retry = 5
net.ipv6.conf.venet0.max_desync_factor = 600
net.ipv6.conf.venet0.max_addresses = 16
net.ipv6.conf.venet0.accept_ra_defrtr = 1
net.ipv6.conf.venet0.accept_ra_pinfo = 1
net.ipv6.conf.venet0.accept_ra_rtr_pref = 1
net.ipv6.conf.venet0.router_probe_interval = 60
net.ipv6.conf.venet0.accept_ra_rt_info_max_plen = 0
net.ipv6.conf.venet0.proxy_ndp = 0
net.ipv6.conf.venet0.accept_source_route = 0
net.ipv6.conf.venet0.optimistic_dad = 0
net.ipv6.conf.venet0.mc_forwarding = 0
net.ipv6.conf.venet0.disable_ipv6 = 0
net.ipv6.conf.venet0.accept_dad = -1
net.ipv6.neigh.venet0.mcast_solicit = 3
net.ipv6.neigh.venet0.ucast_solicit = 3
net.ipv6.neigh.venet0.app_solicit = 0
net.ipv6.neigh.venet0.retrans_time = 250
net.ipv6.neigh.venet0.base_reachable_time = 30
net.ipv6.neigh.venet0.delay_first_probe_time = 5
net.ipv6.neigh.venet0.gc_stale_time = 60
net.ipv6.neigh.venet0.unres_qlen = 3
net.ipv6.neigh.venet0.proxy_qlen = 64
net.ipv6.neigh.venet0.anycast_delay = 100
net.ipv6.neigh.venet0.proxy_delay = 80
net.ipv6.neigh.venet0.locktime = 0
net.ipv6.neigh.venet0.retrans_time_ms = 1000
net.ipv6.neigh.venet0.base_reachable_time_ms = 30000

My HN is directly connected to internet with the IP A.B.C.D

As you can see, everything seems alright (however i'm beginner, so if there are some strange things, tell me).

THERE IS NO IPTABLES RULES

On 101

```
root@localhost:/# iptables -L
Chain INPUT (policy ACCEPT)
target     prot opt source               destination

Chain FORWARD (policy ACCEPT)
target     prot opt source               destination

Chain OUTPUT (policy ACCEPT)
target     prot opt source               destination
```

On the HN :

```
root@milo:/home/choco52# iptables -L
Chain INPUT (policy ACCEPT)
target     prot opt source               destination

Chain FORWARD (policy ACCEPT)
target     prot opt source               destination

Chain OUTPUT (policy ACCEPT)
target     prot opt source               destination
```

Same for nat and mangle (clean).

//

I can now expose my weird problem.

The facts :

-My HN can ping 101.

-I also did create another VE (102) with a different IP (like 10.0.0.1/8) and 102 can ping 101.

-My VE can ping the HN

-If i do masquerading on the HN, 101 can ping google and aptitude works fin etc..

BUT

-When i try to ping 101 from a host which is on another subnet (let's say 192.168.200.0/24), the ping will fail.

But the detail which is killing me is that my VE receive the ICMP REQUEST but decide to not answer it. It really drives me crazy.

This what the VE receive :

```
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on any, link-type LINUX_SLL (Linux cooked), capture size 65535 bytes
09:35:18.352704 IP 192.168.200.253 > 192.168.210.101: ICMP echo request, id 10, seq 0, length
80
        0x0000:  0000 ffff 0000 0000 0000 0000 0000 0800  ................
        0x0010:  4500 0064 002f 0000 fd01 a0b5 c0a8 c8fd  E..d./..........
        0x0020:  c0a8 d265 0800 6f05 000a 0000 0000 0000  ...e..o.........
        0x0030:  0d37 0204 abcd abcd abcd abcd abcd abcd  .7..............
        0x0040:  abcd abcd abcd abcd abcd abcd abcd abcd  ................
        0x0050:  abcd abcd abcd abcd abcd abcd abcd abcd  ................
        0x0060:  abcd abcd abcd abcd abcd abcd abcd abcd  ................
        0x0070:  abcd abcd                                ....
09:35:20.349820 IP 192.168.200.253 > 192.168.210.101: ICMP echo request, id 10, seq 1, length
80
        0x0000:  0000 ffff 0000 0000 0000 0000 0000 0800  ................
        0x0010:  4500 0064 0030 0000 fd01 a0b4 c0a8 c8fd  E..d.0..........
        0x0020:  c0a8 d265 0800 6734 000a 0001 0000 0000  ...e..g4........
        0x0030:  0d37 09d4 abcd abcd abcd abcd abcd abcd  .7..............
        0x0040:  abcd abcd abcd abcd abcd abcd abcd abcd  ................
        0x0050:  abcd abcd abcd abcd abcd abcd abcd abcd  ................
        0x0060:  abcd abcd abcd abcd abcd abcd abcd abcd  ................
        0x0070:  abcd abcd
```

And it goes like this for a long time, no icmp reply from my VE..

This the same issue if i use another protocol like TCP or UDP, my VE will not answer at all.

I tried to ping from different hosts, the issue is the same.

Why does the VE decide to not answer this ping ???
It's so strange.

I did some researches, i looked at the rp_filter, but it seems not to be the problem (as you can see i set it to 0 not only for venet0 but for all..).

By the way, all my ping from the another network come from a ipsec tunnel configured directly in the HN (and another server). I begin to i believe that there is maybe a bug with ipsec.

If you need more information, just ask.
If you have any idea, any thought please tell me i really need some help here :/

Thx.

Subject: Re: A really strange network issue
Posted by Choco52 on Mon, 20 Feb 2012 15:37:50 GMT
View Forum Message <> Reply to Message

Hi there, no ideas lol ?

By the way i just tried with another distro for my VE and the problem still persist.. (i changed from a minimal Debian to a normal debian-6.0-x86_64.tar.gz debian).

I think it is maybe a IPSec issue(my HN is conencted to another network through an ipsec tunnel), even if i don't see why..

Subject: Re: A really strange network issue
Posted by Choco52 on Sat, 25 Feb 2012 21:00:19 GMT
View Forum Message <> Reply to Message

Hi,

After some researches i did find my problem.

I also saw that some others dudes have the same problem.

In fact a solution is to disable the ipsec policy and encryption for the interface.

I don't really understand why this is a problem but disabling these two thing did solve my problem.

net.ipv4.conf.all.disable_xfrm = 1
net.ipv4.conf.all.disable_policy = 1

FYI:

disable_policy
    BOOLEAN
    Disable IPSEC policy (SPD) for this interface

disable_xfrm
    BOOLEAN
    Disable IPSEC encryption on this interface, whatever the policy

Maybe the ipsec policy is apllied twice when using a VE ? thus causing lot of trouble..