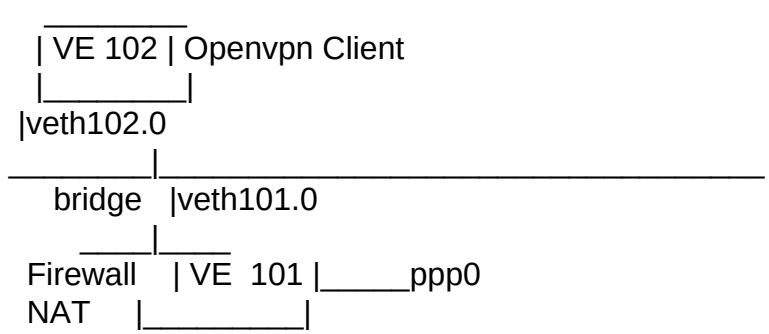
Subject: OpenVPN, IPTables, NAT Issues with UDP Translation after IP Change
Posted by Michael Flaig on Thu, 15 Sep 2011 20:25:23 GMT
View Forum Message <> Reply to Message

Hi Folks,

I have a Debian host running multiple VEs and seeing trouble with udp translations (masq or snat) when the interface ip of ppp0 changes.

A little about the Setup (Stripped down to the problem)

VE0 Squeeze
  VE101 Firewall / NAT / PPPOE
  VE102 Firewall (all accept) / OpenVPN Client



Then there is a OpenVPN gateway on the NET somewhere behind ppp0 that acts as Server and gets connected by VE102's Openvpn CLient.

The Firewall Policy on VE101 is build by ferm - nothing fancy yet.
* Allows forwarding of packets from openvpn client to the outside world
* Does a MASQUERADE to outer-interface ppp0
* Also tried SNAT, same effect.
* The iptables rules are generated right with MASQ (of course) and also on my SNAT test

PPP0 is brought up the debian way with /etc/network/interfaces and ifup/ifdown

The Policy of VE101 looks like (input and output chains stripped):
--- snip ---
@def $IP_WORLD = `ip addr show ppp0 | grep inet | awk '{print $2}'`;

table filter {

[...]

```
    chain FORWARD {
        policy DROP;

        # connection tracking
        mod state state INVALID DROP;
        mod state state (ESTABLISHED RELATED) ACCEPT;

        # connections from the internal net to the internet or to other
        # internal nets are allowed
        interface $DEV_PRIVATE ACCEPT;

        # Reject, the rest is dropped by the above policy
 REJECT;
    }
}

table nat {
    chain POSTROUTING {
        # masquerade private IP addresses
        saddr $NET_PRIVATE outerface $DEV_WORLD MASQUERADE;
 # saddr $NET_PRIVATE outerface $DEV_WORLD SNAT to $IP_WORLD;
    }
}
--- snap ---
```

Now the whole setup works at first. Tunnel comes up. All good.
Until the first IP change. The Tunnel is down but I see connections
incoming on the OpenVPN server.

I do a tcpdump on ppp0 and see outgoing packets that get translated to
the old IP address (X.X.X.X) instead of the new IP address.

09:50:09.628341 IP X.X.X.X.20001 > Y.Y.Y.Y.20001: UDP, length 14

I flushed the rules, reinstalled the rules - not helping.

Now /proc/net/conntrack on VE101 tells me that the old connection is
there in state ASSURED

udp   17 166 src=VE102 dst=Openvpnserver sport=20001 dport=20001
packets=435 bytes=23129 src=Openvpnserver dst=X.X.X.X sport=20001
dport=20001 packets=92 bytes=9701 [ASSURED] mark=0 secmark=0 use=2

Hm ... okay. weird. Shouldn't this be flushed on an IP change?
Well as UDP is stateless and iptables makes it stateful, however source
and destination port do not change, this old conntrack rule is still
used but should not be there anymore IMHO.

So what about deleting it?

ON VE101
conntrack -D -s VE102
conntrack v0.9.14 (conntrack-tools): Operation failed: Connection refused

strace for conntrack:
socket(PF_NETLINK, SOCK_RAW, 12)   = 3
getsockname(3, {sa_family=AF_NETLINK, pid=0, groups=00000000}, [12]) = 0
bind(3, {sa_family=AF_NETLINK, pid=0, groups=00000000}, 12) = 0
getsockname(3, {sa_family=AF_NETLINK, pid=947, groups=00000000}, [12]) =
0
bind(3, {sa_family=AF_NETLINK, pid=947, groups=00000000}, 12) = 0
socket(PF_NETLINK, SOCK_RAW, 12)   = 4
getsockname(4, {sa_family=AF_NETLINK, pid=0, groups=00000000}, [12]) = 0
bind(4, {sa_family=AF_NETLINK, pid=0, groups=00000000}, 12) = 0
getsockname(4, {sa_family=AF_NETLINK, pid=-4269, groups=00000000}, [12])
= 0
bind(4, {sa_family=AF_NETLINK, pid=-4269, groups=00000000}, 12) = 0
sendto(3, "\24\0\0\0\1\1\1\3\270\260qN\0\0\0\0\2\0\0\0", 20, 0,
{sa_family=AF_NETLINK, pid=0, groups=00000000}, 12) = -1 ECONNREFUSED
(Connection refused)
close(4)   = 0
close(3)   = 0
write(2, "conntrack v0.9.14 (conntrack-too"..., 37conntrack v0.9.14
(conntrack-tools):) = 37
write(2, "Operation failed: Connection ref"..., 36Operation failed:
Connection refused) = 36
write(2, "\n", 1
)   = 1
exit_group(1)   = ?

So strace ends on a socket connect with a connection refused error.
Deleting the conntrack rule is not allowed inside VE101. Maybe missing
something in the VE 101 config? Could this error also be the reason for
not flushing those conntrack rules?

--- snip ---
[... non-relevant parts cut ...]
IPTABLES="ip_tables iptable_filter iptable_mangle iptable_nat ipt_limit
ipt_multiport ipt_tos ipt_TOS ipt_REJECT ipt_TCPMSS ipt_tcpmss ipt_ttl
ipt_length ip_conntrack ip_conntrack_ftp ip_conntrack_irc ipt_conntrack
ipt_state ipt_helper ip_nat_ftp ip_nat_irc ipt_REDIRECT xt_mac
ipt_recent ipt_owner"
FEATURES="ppp:on "
DEVICES="c:108:0:rw "
CAPABILITY="NET_ADMIN:on NET_RAW:on SYS_ADMIN:on "
--- snap ---

On the CAP - this VE101 is running quagga.

Oh of course this problem can be fixed stop and start of VE101 - but that should not be the solution here on a 24 hr ip change :-/

Any suggestions? I'm pretty much out of google hints and ideas myself. Oh there are a few forum posts about AF_NETLINK in russian though.

Should conntrack-tools work at all inside a VE and if yes anyone know how to make them work?

Any input greatly appreciated.

Thanks,

Mike
