Subject: IPsec packets from VEs sent to wrong interface
Posted by marcusb on Wed, 18 Jun 2008 08:42:44 GMT
View Forum Message <> Reply to Message

Hi,

I'm running OpenSWAN on the host node to provide tunnels to some VEs. The VEs are connected on veth devices that are bridged together to br0. The IPsec tunnel is correctly established, but response traffic from the VE is being sent out on br0, not the external interface eth0. Is there a workaround for this?

Details of the setup:

Server is OpenVZ 2.6.24 (compiled from git), Debian x86_64, OpenSWAN 2.4.12.

Host node interfaces:
eth0: public address 1.2.3.4 server.example.org
br0: bridge, internal address 172.16.1.1/24, only slave interface veth106.0
veth106.0: host end of veth.

VE interfaces:
eth0: veth interface, address 172.16.1.106

Now "ping 172.16.1.1" from the IPsec client (client.example.org with private address 172.16.2.2) works correctly, but "ping 172.16.1.106" shows this:
[host:~]# tcpdump -i br0
10:31:43.238582 IP 172.16.2.2 > 172.16.1.106: ICMP echo request, id 9274, seq 40, length 64
10:31:43.238617 IP server.example.org.4500 > client.example.org.4500: UDP-encap: ESP(spi=0xeee72df0,seq=0x35c), length 132
10:31:44.230477 IP 172.16.2.2 > 172.16.1.106: ICMP echo request, id 9274, seq 41, length 64
10:31:44.230509 IP server.example.org.4500 > client.example.org.4500: UDP-encap: ESP(spi=0xeee72df0,seq=0x35d), length 132

Here the packets destined for client.example.org are only seen on br0, not on the external interface. I have forwarding enabled on both br0 and eth0.

Cheers,

Marcus

Subject: Re: IPsec packets from VEs sent to wrong interface
Posted by Jerry Del Rio on Wed, 18 Jun 2008 21:04:51 GMT
View Forum Message <> Reply to Message

Marcus,

I would have a look at OpenVPN. OpenVPN works

as a client/server application when enabled. OpenVPN will
handle not just the tunnels but the routing for each
client as well as devices behind the clients that wish
to communicate with other clients or devices via the
VPN/tunnel. I haven't used the setup you are trying to do
but I think its feasible. If your VE's are not on the
same subnets then you should use VETH as you are doing,
otherwise a VENET device is fine if your ISP can get
you the block that you need.

Your initial hang up will be the tun device that needs
to be mapped to the VE's like this:

Allow your container to use the tun/tap device:

vzctl set 101 --devices c:10:200:rw --save
vzctl set 101 --capability net_admin:on --save

And create the character device file inside the container:

vzctl exec 101 mkdir -p /dev/net
vzctl exec 101 mknod /dev/net/tun c 10 200
vzctl exec 101 chmod 600 /dev/net/tun

***Anyone else correct me if I am wrong but I recall
somewhere that we can use the same host node tun device
for other VE's. This should enable other VE's that
are not the OpenVPN(server) to VPN as OpenVPN(clients).

Jerry Del Rio
Systems Engineer
949-331-4038 - Mobile
www.qunoc.com

Quoting Marcus Better <marcus@better.se>:

> Hi,
>
> I'm running OpenSWAN on the host node to provide tunnels to some

> VEs. The VEs are connected on veth devices that are bridged together
>  to br0. The IPsec tunnel is correctly established, but response
> traffic from the VE is being sent out on br0, not the external
> interface eth0. Is there a workaround for this?
>
> Details of the setup:
>
> Server is OpenVZ 2.6.24 (compiled from git), Debian x86_64, OpenSWAN 2.4.12.
>
> Host node interfaces:
> eth0: public address 1.2.3.4 server.example.org
> br0: bridge, internal address 172.16.1.1/24, only slave interface veth106.0
> veth106.0: host end of veth.
>
> VE interfaces:
> eth0: veth interface, address 172.16.1.106
>
> Now "ping 172.16.1.1" from the IPsec client (client.example.org with
>  private address 172.16.2.2) works correctly, but "ping
> 172.16.1.106"  shows this:
> [host:~]# tcpdump -i br0
> 10:31:43.238582 IP 172.16.2.2 > 172.16.1.106: ICMP echo request, id
> 9274, seq 40, length 64
> 10:31:43.238617 IP server.example.org.4500 >
> client.example.org.4500: UDP-encap: ESP(spi=0xeee72df0,seq=0x35c),
> length 132
> 10:31:44.230477 IP 172.16.2.2 > 172.16.1.106: ICMP echo request, id
> 9274, seq 41, length 64
> 10:31:44.230509 IP server.example.org.4500 >
> client.example.org.4500: UDP-encap: ESP(spi=0xeee72df0,seq=0x35d),
> length 132
>
> Here the packets destined for client.example.org are only seen on
> br0, not on the external interface. I have forwarding enabled on
> both br0 and eth0.
>
> Cheers,
>
> Marcus
>
>
> --
> This message has been scanned for viruses and
> dangerous content by MailScanner, and is
> believed to be clean.
>

--
This message has been scanned for viruses and
dangerous content by MailScanner, and is
believed to be clean.

Subject: Re: IPsec packets from VEs sent to wrong interface
Posted by marcusb on Thu, 19 Jun 2008 07:43:10 GMT
View Forum Message <> Reply to Message

Hi,

Jerry Del Rio wrote:
> I would have a look at OpenVPN.

Thanks, but for now I'm planning to use IPsec. I'll keep OpenVPN in mind in case this proves impossible.

Cheers,

Marcus

Subject: Re: IPsec packets from VEs sent to wrong interface
Posted by marcusb on Mon, 23 Jun 2008 12:15:37 GMT
View Forum Message <> Reply to Message

-----BEGIN PGP SIGNED MESSAGE-----
Hash: SHA1

Marcus Better wrote:
> IPsec tunnel is correctly established, but response traffic from the VE is
> being sent out on br0, not the external interface eth0.

FYI I haven't resolved this yet, and upstream is unwilling to look at OpenVZ-specific problems. Any help with trying to reproduce it in a mainline kernel would be appreciated.

Basically it would suffice to configure a veth pair without a full OpenVZ virtual environment on the other side. Is that possible?

Marcus
-----BEGIN PGP SIGNATURE-----
Version: GnuPG v1.4.6 (GNU/Linux)

iD8DBQFIX5PwXjXn6TzcAQkRAkGoAJ4+8jt0e6R7rAtTaRnYszvHRgM1cACgqEhr
v42RBwPux+QBDKs5qXfpECM=
=RqmU
-----END PGP SIGNATURE-----

---

Subject: Re:  Re: IPsec packets from VEs sent to wrong interface
Posted by den on Mon, 23 Jun 2008 13:07:10 GMT
View Forum Message <> Reply to Message

I think the problem is in you routing setup.
ip route get
with proper source, destination, incoming interface will help you in
this task.

On Mon, 2008-06-23 at 14:15 +0200, Marcus Better wrote:
> -----BEGIN PGP SIGNED MESSAGE-----
> Hash: SHA1
>
> Marcus Better wrote:
> > IPsec tunnel is correctly established, but response traffic from the VE is
> > being sent out on br0, not the external interface eth0.
>
> FYI I haven't resolved this yet, and upstream is unwilling to look at OpenVZ-specific problems. Any help with trying to reproduce it in a mainline kernel would be appreciated.
>
> Basically it would suffice to configure a veth pair without a full OpenVZ virtual environment on the other side. Is that possible?
>
> Marcus
> -----BEGIN PGP SIGNATURE-----
> Version: GnuPG v1.4.6 (GNU/Linux)
>
> iD8DBQFIX5PwXjXn6TzcAQkRAkGoAJ4+8jt0e6R7rAtTaRnYszvHRgM1cACgqEhr
> v42RBwPux+QBDKs5qXfpECM=
> =RqmU
> -----END PGP SIGNATURE-----
>
>

---

Subject: Re: Re: IPsec packets from VEs sent to wrong interface
Posted by marcusb on Tue, 24 Jun 2008 07:39:11 GMT
View Forum Message <> Reply to Message

Hi,

Denis V. Lunev wrote:
> I think the problem is in you routing setup.

I've checked but don't see anything suspicious. The routing setup is very simple.

```
[host:~]# ip route
172.16.2.2 dev eth0  scope link  src 172.16.1.1
172.16.1.101 dev venet0  scope link
x.y.z.0/25 dev eth0  proto kernel  scope link  src x.y.z.w
172.16.1.0/24 dev br0  proto kernel  scope link  src 172.16.1.1
default via x.y.z.1 dev eth0

[host:~]# ip route get 172.16.2.2
172.16.2.2 dev eth0  src 172.16.1.1
    cache  expires 21334342sec mtu 1500 advmss 1460 hoplimit 64
```

(Public IP addresses have been altered.)

ping from 172.16.2.2 to 172.16.1.1 works, but the other direction does  not. When trying, the host node sends out ARP requests for 172.16.2.2 unencrypted on eth0, ignoring IPsec policy.

```
[host:~]# setkey -DP
172.16.2.2[any] 172.16.1.0/24[any] any
        in ipsec
        esp/tunnel/a.b.c.d-x.y.z.w/unique#16397
        created: Jun 24 09:03:28 2008  lastused: Jun 24 09:16:11 2008
        lifetime: 0(s) validtime: 0(s)
        spid=1656 seq=1 pid=2200
        refcnt=1
172.16.1.0/24[any] 172.16.2.2[any] any
        out ipsec
        esp/tunnel/x.y.z.w-a.b.c.d/unique#16397
        created: Jun 24 09:03:45 2008  lastused: Jun 24 09:33:20 2008
        lifetime: 0(s) validtime: 0(s)
        spid=1673 seq=2 pid=2200
        refcnt=3
172.16.2.2[any] 172.16.1.0/24[any] any
        fwd ipsec
        esp/tunnel/a.b.c.d-x.y.z.w/unique#16397
        created: Jun 24 09:03:28 2008  lastused:
        lifetime: 0(s) validtime: 0(s)
        spid=1666 seq=3 pid=2200
        refcnt=1
```

Cheers,

Marcus

Subject: Re:  Re: Re: IPsec packets from VEs sent to wrong interface
Posted by den on Tue, 24 Jun 2008 08:28:10 GMT
View Forum Message <> Reply to Message

On Tue, 2008-06-24 at 09:39 +0200, Marcus Better wrote:
> Hi,
>
> Denis V. Lunev wrote:
> > I think the problem is in you routing setup.
>
> I've checked but don't see anything suspicious. The routing setup is very simple.
>
> [host:~]# ip route
> 172.16.2.2 dev eth0  scope link  src 172.16.1.1
> 172.16.1.101 dev venet0  scope link
> x.y.z.0/25 dev eth0  proto kernel  scope link  src x.y.z.w
> 172.16.1.0/24 dev br0  proto kernel  scope link  src 172.16.1.1
> default via x.y.z.1 dev eth0
>
> [host:~]# ip route get 172.16.2.2
> 172.16.2.2 dev eth0  src 172.16.1.1
>     cache  expires 21334342sec mtu 1500 advmss 1460 hoplimit 64

you are doing a wrong thing. You should ask
 ip route get ADDRESS from ADDRESS iif STRING
with both from and iif specified to get a clue.

Regards,
 Den