strongSwan - Bug #2183

Initiator ends up with no route if IPv6 DAD for the virtual IP failed

08.12.2016 17:46 - Petre Rodan

Status:	Closed	**Start date:**	08.12.2016
Priority:	Normal	**Due date:**	
Assignee:	Tobias Brunner	**Estimated time:**	0.00 hour
Category:	kernel-interface		
Target version:	5.5.2		
Affected version:	5.5.1	**Resolution:**	Fixed

Description

two problems will be described:

- the bug appears to resemble issue #1189, but the kernel used here is much newer and in 30% of cases it works properly
- 'ipsec up' command ends in success regardless if the default gateway has been set or not in table 220. this can be confusing to users since they think the vpn connection is up and routing their packets

tested with strongswan-5.5.1, kernels 4.6.4 and 4.8.10 (on both endpoints)
configure arguments:

```
./configure --prefix=/usr --build=x86_64-pc-linux-gnu --host=x86_64-pc-linux-gnu --mandir=/usr/sha
re/man --infodir=/usr/share/info --datadir=/usr/share --sysconfdir=/etc --localstatedir=/var/lib -
-disable-dependency-tracking --disable-silent-rules --libdir=/usr/lib64 --disable-static --enable-
ikev1 --enable-ikev2 --enable-swanctl --enable-socket-dynamic --with-capabilities=libcap --disable
-curl --enable-constraints --disable-ldap --disable-leak-detective --disable-dhcp --enable-eap-sim
 --enable-eap-sim-file --enable-eap-simaka-sql --enable-eap-simaka-pseudonym --enable-eap-simaka-r
eauth --enable-eap-identity --enable-eap-md5 --enable-eap-aka --enable-eap-aka-3gpp2 --enable-md4
--enable-eap-mschapv2 --enable-eap-radius --enable-eap-tls --enable-xauth-eap --enable-farp --enab
le-gmp --enable-gcrypt --disable-mysql --disable-nm --enable-openssl --disable-xauth-pam --disable
-pkcs11 --disable-sqlite --with-systemdsystemunitdir=/usr/lib/systemd/system --disable-eap-gtc --e
nable-vici --enable-rdrand
```

initiator ipsec.conf:

```
config setup
    uniqueids = no

conn %default

conn generic
    leftfirewall=yes
    leftauth=eap
    eap_identity=CENSORED
    rightauth=pubkey
    right=2001:4d80:0:42:1501::1
    rightid=@bucharest-s15-i01.cg-dialup.net
    dpdaction=clear
    dpddelay=5s
    dpdtimeout=20s

conn v4
    also=generic
    left=%any
    leftsourceip=%config
    rightsubnet=0.0.0.0/0
    auto=add

conn both
    also=generic
    left=%any
```

```
    leftsourceip=%config,%config6
    rightsubnet=::/0,0.0.0.0/0
    auto=add

conn v6
    also=generic
    left=%any6
    leftsourceip=%config6
    rightsubnet=::/0
    auto=add

conn pass
    leftsubnet=192.168.1.121/32
    rightsubnet=192.168.1.0/24
    type=passthrough
    authby=never
    auto=route
```

ipsec up v6 can end up in two ways:

- in 70% of cases it ends up with:

```
authentication of 'bucharest-s15-i01.cg-dialup.net' with EAP successful
IKE_SA v6[1] established between 2001:4d80:0:40:b889:86ff:fefb:fe51[2001:4d80:0:40:b889:86ff:f
efb:fe51]...2001:4d80:0:42:1501::1[bucharest-s15-i01.cg-dialup.net]
scheduling reauthentication in 9884s
maximum IKE_SA lifetime 10424s
installing DNS server CENSORED to /etc/resolv.conf
installing new virtual IP 2001:4d80:0:42:1501:240:0:2
received netlink error: Invalid argument (22)
unable to install source route for 2001:4d80:0:42:1501:240:0:2
CHILD_SA v6{1} established with SPIs cb1a9bab_i c439f9ad_o and TS 2001:4d80:0:42:1501:240:0:2/
128 === ::/0
connection 'v6' established successfully
 ~ # echo $?
0

 ~ # ip -6 a s eth0
4: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 state UP qlen 1000
    inet6 2001:4d80:0:42:1501:240:0:2/128 scope global tentative deprecated dadfailed
       valid_lft forever preferred_lft 0sec
    inet6 2001:4d80:0:40:b889:86ff:fefb:fe51/64 scope global mngtmpaddr dynamic
       valid_lft 2337495sec preferred_lft 350295sec
    inet6 2001:4d80:0:40:a001::1/64 scope global
       valid_lft forever preferred_lft forever
    inet6 fe80::b889:86ff:fefb:fe51/64 scope link
       valid_lft forever preferred_lft forever
 ~ # ip -6 r show table 220
[empty]
```

- in 30% of cases it ends up with:

```
authentication of 'bucharest-s15-i01.cg-dialup.net' with EAP successful
IKE_SA v6[6] established between 2001:4d80:0:40:b889:86ff:fefb:fe51[2001:4d80:0:40:b889:86ff:f
efb:fe51]...2001:4d80:0:42:1501::1[bucharest-s15-i01.cg-dialup.net]
scheduling reauthentication in 9849s
maximum IKE_SA lifetime 10389s
installing DNS server CENSORED to /etc/resolv.conf
installing new virtual IP 2001:4d80:0:42:1501:240:0:2
CHILD_SA v6{6} established with SPIs ce41a7bc_i cf33d714_o and TS 2001:4d80:0:42:1501:240:0:2/
128 === ::/0
received AUTH_LIFETIME of 10130s, scheduling reauthentication in 9590s
peer supports MOBIKE
connection 'v6' established successfully
```

```
 ~ # ip -6 a s eth0
4: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 state UP qlen 1000
    inet6 2001:4d80:0:42:1501:240:0:2/128 scope global deprecated
       valid_lft forever preferred_lft 0sec
    inet6 2001:4d80:0:40:b889:86ff:fefb:fe51/64 scope global mngtmpaddr dynamic
       valid_lft 2337358sec preferred_lft 350158sec
    inet6 2001:4d80:0:40:a001::1/64 scope global
       valid_lft forever preferred_lft forever
    inet6 fe80::b889:86ff:fefb:fe51/64 scope link
       valid_lft forever preferred_lft forever
 ~ # ip -6 r show table 220
default via 2001:4d80:0:42:1501::1 dev eth0  proto static  src 2001:4d80:0:42:1501:240:0:2  me
tric 1024  pref medium
```

Associated revisions

Revision 6d45ca92 - 25.01.2017 17:36 - Tobias Brunner

kernel-netlink: Set NODAD flag for virtual IPv6 addresses

The Optimistic Duplicate Address Detection (DAD) seems to fail in some cases (`dadfailed` in `ip addr`) rendering the virtual IP address unusable.

Fixes #2183.

Revision b062d3cc - 06.02.2017 11:10 - Tobias Brunner

kernel-netlink: Set NODAD flag for virtual IPv6 addresses

The Optimistic Duplicate Address Detection (DAD) seems to fail in some cases (`dadfailed` in `ip addr`) rendering the virtual IP address unusable.

Fixes #2183.

History

#1 - 09.12.2016 11:50 - Tobias Brunner

- Description updated

- Category deleted (charon)

- Status changed from New to Feedback

Could be a problem with NDP, which you might have to passthrough.

If the failure to install the route is caused by this:

```
inet6 2001:4d80:0:42:1501:240:0:2/128 ... dadfailed
```

I guess that could also be avoided by setting the IFA_F_NODAD flag on the virtual IPs.

> 'ipsec up' command ends in success regardless if the default gateway has been set or not in table 220.

The stroke command (as called by the ipsec script) does only return exit codes != 0 if there was an internal failure. There is no feedback from the *stroke* plugin whether a command was successful. If you want that try swanctl. But even so, the failure to install the route is currently not handled as fatal error so the CHILD_SA is still established successfully.

#2 - 09.12.2016 14:34 - Petre Rodan

Tobias Brunner wrote:

> Could be a problem with NDP, which you might have to passthrough.

I tried adding this exact passthrough:

```
conn icmpv6
```

```
    right=::1 # so this connection does not get used for other purposes
    leftsubnet=::/0[ipv6-icmp/%any]
    rightsubnet=::/0[ipv6-icmp/%any]
    type=passthrough
    auto=route
```

but I get the exact same error.
here is some debug output, maybe it helps:

when it fails:

```
[KNL] getting iface name for index 4
[KNL] using 2001:4d80:0:49:1501::1 as nexthop and eth0 as dev to reach 2001:4d80:0:49:1501::1/128
[KNL] installing route: ::/0 via 2001:4d80:0:49:1501::1 src 2001:4d80:0:49:1501:240:0:1 dev eth0
[KNL] getting iface index for eth0
[KNL]EWROUTE 210: => 96 bytes @ 0x734058ce5e50
[KNL]    0: 60 00 00 00 18 00 05 06 D2 00 00 00 CE 26 00 00  `............&..
[KNL]   16: 0A 00 00 00 DC 04 00 01 00 00 00 00 14 00 01 00  ................
[KNL]   32: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  ................
[KNL]   48: 14 00 07 00 20 01 4D 80 00 00 00 49 15 01 02 40  .... .M....I...@
[KNL]   64: 00 00 00 01 14 00 05 00 20 01 4D 80 00 00 00 49  ........ .M....I
[KNL]   80: 15 01 00 00 00 00 00 01 08 00 04 00 04 00 00 00  ................
[KNL] received (2) 210: => 116 bytes @ 0x428113f0750
[KNL]    0: 74 00 00 00 02 00 00 00 D2 00 00 00 CE 26 00 00  t............&..
[KNL]   16: EA FF FF FF 60 00 00 00 18 00 05 06 D2 00 00 00  ....`...........
[KNL]   32: CE 26 00 00 0A 00 00 00 DC 04 00 01 00 00 00 00  .&..............
[KNL]   48: 14 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00  ................
[KNL]   64: 00 00 00 00 14 00 07 00 20 01 4D 80 00 00 00 49  ........ .M....I
[KNL]   80: 15 01 02 40 00 00 00 01 14 00 05 00 20 01 4D 80  ...@........ .M.
[KNL]   96: 00 00 00 49 15 01 00 00 00 00 00 01 08 00 04 00  ...I............
[KNL]  112: 04 00 00 00                                      ....
[KNL] received netlink error: Invalid argument (22)
[KNL] unable to install source route for 2001:4d80:0:49:1501:240:0:1
[KNL] policy ::/0 === 2001:4d80:0:49:1501:240:0:1/128 in already exists, increasing refcount
[KNL] updating policy ::/0 === 2001:4d80:0:49:1501:240:0:1/128 in [priority 167232, refcount 2]
```

when it's fine:

```
[KNL] getting iface name for index 4
[KNL] using 2001:4d80:0:49:1501::1 as nexthop and eth0 as dev to reach 2001:4d80:0:49:1501::1/128
[KNL] installing route: ::/0 via 2001:4d80:0:49:1501::1 src 2001:4d80:0:49:1501:240:0:1 dev eth0
[KNL] getting iface index for eth0
[KNL]EWROUTE 318: => 96 bytes @ 0x7340594e6e50
[KNL]    0: 60 00 00 00 18 00 05 06 3E 01 00 00 CE 26 00 00  `.......>....&..
[KNL]   16: 0A 00 00 00 DC 04 00 01 00 00 00 00 14 00 01 00  ................
[KNL]   32: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  ................
[KNL]   48: 14 00 07 00 20 01 4D 80 00 00 00 49 15 01 02 40  .... .M....I...@
[KNL]   64: 00 00 00 01 14 00 05 00 20 01 4D 80 00 00 00 49  ........ .M....I
[KNL]   80: 15 01 00 00 00 00 00 01 08 00 04 00 04 00 00 00  ................
[KNL] received (2) 318: => 36 bytes @ 0x73403c0015e0
[KNL]    0: 24 00 00 00 02 00 00 00 3E 01 00 00 CE 26 00 00  $.......>....&..
[KNL]   16: 00 00 00 00 60 00 00 00 18 00 05 06 3E 01 00 00  ....`.......>...
[KNL]   32: CE 26 00 00                                      .&..
[KNL] policy ::/0 === 2001:4d80:0:49:1501:240:0:1/128 in already exists, increasing refcount
[KNL] updating policy ::/0 === 2001:4d80:0:49:1501:240:0:1/128 in [priority 167232, refcount 2]
```

> If the failure to install the route is caused by this:
> inet6 2001:4d80:0:42:1501:240:0:2/128 ... dadfailed

I never heard of this NODAD flag, where could I have seen the error you talk about above?

> I guess that could also be avoided by setting the IFA_F_NODAD flag on the virtual IPs.

I see that iproute2 can set this flag when a new address is set, but the virtual IPs are set by strongswan via %config6. which makes me even more confused.

#3 - 09.12.2016 15:13 - Tobias Brunner

> If the failure to install the route is caused by this:

> inet6 2001:4d80:0:42:1501:240:0:2/128 ... dadfailed

> I never heard of this NODAD flag, where could I have seen the error you talk about above?

This is from the output of ip addr you posted. Could you please try to verify that every time it fails dadfailed is seen in that output for the virtual IP.

#4 - 13.12.2016 13:16 - Petre Rodan

Tobias Brunner wrote:

> Could you please try to verify that every time it fails dadfailed is seen in that output for the virtual IP.

I can confirm that every time the source route setup fails I get that "dadfailed" in ip's output. this is happening from within a kvm image (the test servers are virtualized). connecting from the same subnet from a bare metal machine does work without the errors above.

I tried to change the network interface type from the host (e1000, virtio) but with no improvement as far as this ticket is concerned.

how can I help to fix this?

#5 - 13.12.2016 17:33 - Tobias Brunner

> > Could you please try to verify that every time it fails dadfailed is seen in that output for the virtual IP.
>
> I can confirm that every time the source route setup fails I get that "dadfailed" in ip's output. this is happening from within a kvm image (the test servers are virtualized). connecting from the same subnet from a bare metal machine does work without the errors above.

OK, as mentioned, setting the NODAD flag on the address might help. You may try the patch in the *2183-nodad* branch.

#6 - 14.12.2016 05:36 - Petre Rodan

thanks for the patch!
it works 100% of the time after your fix was applied. please merge it into stable when possible.

#7 - 14.12.2016 11:39 - Tobias Brunner

- Tracker changed from Issue to Bug

- Target version set to 5.5.2

OK, I'll line this up for the next release.

#8 - 25.01.2017 17:49 - Tobias Brunner

- Subject changed from ipsec initiator ends up with no ipv6 default gateway to Initiator ends up with no route if IPv6 DAD for the virtual IP failed

- Category set to kernel-interface

- Status changed from Feedback to Closed

- Assignee set to Tobias Brunner

- Resolution set to Fixed