strongSwan - Issue #3628

Constant `retransmit` while establishing CHILD_SA

12.11.2020 11:09 - Albert Wu

Status:	Feedback		
Priority:	Normal		
Assignee:			
Category:			
Affected version:	5.2.2	**Resolution:**	

Description

Disclaimer: I know that 5.2.2 is a very old version; however, the Unifi Security Gateway comes bundled with 5.2.2. I'd like to know if the issue I experience is due to misconfiguration, or if it is a known issue with 5.2.2 (and how to work around it).

Issue:

I am setting up IPSec between Unifi Security Gateway (VyOS) and AWS with 2 tunnels. Sometimes, both tunnels come up fine. Sometimes, only one tunnel comes up, and the other gets stuck in a `retransmit x of request with message ID 1` loop. Sometimes, both of the two tunnels gets stuck in that loop.

Additionally, when the two tunnels are up, and I run `ipsec restart`, I almost always see both tunnels get stuck in retransmit loop again. However, sometimes, when I run `ipsec stop`, wait 2-3 minutes, and run `ipsec start`, both tunnels can be re-established.

I am behind a NAT device, but I have also tried removing that hop, and I still have the same issue.

```
conn aws-tunnel-vti-0
    left="%any"
    leftsourceip=aaa.aaa.aaa.aaa/30
    leftupdown=/config/ipsec-updown.sh
    right=zzz.zzz.zzz.zzz
    rightid="%any"
    leftsubnet=0.0.0.0/0
    rightsubnet=0.0.0.0/0
    ike=aes128-sha1-modp1024!
    keyexchange=ikev2
    reauth=no
    ikelifetime=28800s
    # ikelifetime=120s
    dpddelay=15s
    dpdtimeout=30s
    # dpdaction=restart
    dpdaction=clear
    esp=aes128-sha1-modp1024!
    keylife=3600s
    rekeymargin=540s
    type=tunnel
    compress=no
    authby=rsasig
    leftrsasigkey=%cert
    rightrsasigkey=%cert
    rightca=%same
    leftcert=/etc/ipsec.d/certs/cert.txt
    mark=9437185
    #auto=route
    auto=start
    keyingtries=%forever
```

```
Nov 12 01:57:00 05[IKE] <aws-tunnel-vti-1|1> retransmit 1 of request with message ID 1
Nov 12 01:57:00 05[NET] <aws-tunnel-vti-1|1> sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy
.yyy.yyy[4500] (1644 bytes)
Nov 12 01:57:00 06[NET] sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy.yyy.yyy[4500]
Nov 12 01:57:01 16[IKE] <aws-tunnel-vti-0|2> retransmit 1 of request with message ID 1
Nov 12 01:57:01 16[NET] <aws-tunnel-vti-0|2> sending packet: from xxx.xxx.xxx.xxx[4500] to zzz.zzz
```

```
.zzz.zzz[4500] (1644 bytes)
Nov 12 01:57:01 06[NET] sending packet: from xxx.xxx.xxx.xxx[4500] to zzz.zzz.zzz.zzz[4500]
Nov 12 01:57:07 03[IKE] <aws-tunnel-vti-1|1> retransmit 2 of request with message ID 1
Nov 12 01:57:07 03[NET] <aws-tunnel-vti-1|1> sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy
.yyy.yyy[4500] (1644 bytes)
Nov 12 01:57:07 06[NET] sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy.yyy.yyy[4500]
Nov 12 01:57:08 04[IKE] <aws-tunnel-vti-0|2> retransmit 2 of request with message ID 1
Nov 12 01:57:08 04[NET] <aws-tunnel-vti-0|2> sending packet: from xxx.xxx.xxx.xxx[4500] to zzz.zzz
.zzz.zzz[4500] (1644 bytes)
Nov 12 01:57:08 06[NET] sending packet: from xxx.xxx.xxx.xxx[4500] to zzz.zzz.zzz.zzz[4500]
Nov 12 01:57:20 11[IKE] <aws-tunnel-vti-1|1> retransmit 3 of request with message ID 1
Nov 12 01:57:20 11[NET] <aws-tunnel-vti-1|1> sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy
.yyy.yyy[4500] (1644 bytes)
Nov 12 01:57:20 06[NET] sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy.yyy.yyy[4500]
Nov 12 01:57:21 15[IKE] <aws-tunnel-vti-0|2> retransmit 3 of request with message ID 1
Nov 12 01:57:21 15[NET] <aws-tunnel-vti-0|2> sending packet: from xxx.xxx.xxx.xxx[4500] to zzz.zzz
.zzz.zzz[4500] (1644 bytes)
```

History

#1 - 12.11.2020 11:44 - Tobias Brunner

- Status changed from New to Feedback

> I am setting up IPSec between Unifi Security Gateway (VyOS) and AWS with 2 tunnels.

Are these separate IKE_SAs? Or two CHILD_SAs in a single IKE_SA? You already negotiate 0.0.0.0/0 as traffic selectors here (I guess for a route-based setup via VTI), so how does the other tunnel differ? Could it be that the system sends IKE packets over one of the tunnel interfaces?

> Sometimes, both tunnels come up fine. Sometimes, only one tunnel comes up, and the other gets stuck in a `retransmit x of request with message ID 1` loop. Sometimes, both of the two tunnels gets stuck in that loop.

How does the routing table look like if that happens? Are there still VTIs?

> Additionally, when the two tunnels are up, and I run `ipsec restart`, I almost always see both tunnels get stuck in retransmit loop again. However, sometimes, when I run `ipsec stop`, wait 2-3 minutes, and run `ipsec start`, both tunnels can be re-established.

Sounds weird, but might be due to some router/firewall resetting whatever it is doing after a while or packets taking a different route.

```
Nov 12 01:57:00 05[IKE] <aws-tunnel-vti-1|1> retransmit 1 of request with message ID 1
Nov 12 01:57:00 05[NET] <aws-tunnel-vti-1|1> sending packet: from xxx.xxx.xxx.xxx[4500] to yyy.yyy.yyy.yyy
[4500] (1644 bytes)
```

From the size I'd guess this is an IKE_AUTH message. Is that the case? And is it always the same? Can you capture traffic on the peer to see if it receives anything? Due to the size it might be an IP fragmentation issue. Maybe try setting *fragmentation=yes* if the peer supports it. If not, there are other ways to reduce the size (e.g. not sending the certificate if the peer allows a local configuration of the same, or using ECDSA instead of RSA for smaller keys/certificates if both peers support it).

#2 - 13.11.2020 19:41 - Albert Wu

To address your questions:
- Two separate IKE_SAs. One vti for each tunnel.
- No, there are not vtis when the tunnel gets stuck in `retransmit` loop. I configure the vti in a custom updown script.
- Interesting. I could try looking into the IP fragmentation issue. How do I find out what packet size is supported?

Now on to new information:

The issue is indeed something to do with the NAT device in between.
- When I had tested with the NAT device in between, I had the tunnels sometimes working and sometimes not.
- When I had tested with the NAT device in between, but with IP passthrough: that is, external WAN IP shared with my gateway and a promise of no firewall settings applied, I still had the issue.
- However, when I moved my gateway to another building, I was able to connect IPSec and `ipsec restart` constantly without issue.

Do you know what kind of settings/behavior I should be looking for in the NAT device that could cause this? Is it most likely the IP fragmentation issue?

Thanks.

#3 - 16.11.2020 10:14 - Tobias Brunner

> - Interesting. I could try looking into the IP fragmentation issue. How do I find out what packet size is supported?

The MTU for Ethernet (without jumbo frames) is 1500, so any IP message larger than that gets fragmented (for IPv6, the sender has to do the fragmenting, with IPv4 any router can). With fragmentation enabled, strongSwan will fragment IKEv2 messages to 1280 bytes each (configurable via *charon.fragment_size*), which is the minimum MTU that must be supported for IPv6.

> Do you know what kind of settings/behavior I should be looking for in the NAT device that could cause this? Is it most likely the IP fragmentation issue?

If it's always the IKE_AUTH message, a fragmentation issue seems likely. But it's strange that it would only happen sometimes (that could be caused by switching routes beyond the NAT device, i.e. one route dropping fragments, the other not). What often causes problems on NAT devices are IPsec passthrough features/rules that treat IKE/ESP packets specially and sometimes incorrectly.