strongSwan - Issue #3670

Can routing rules be changed without terminating and re-initiating the tunnel

11.01.2021 13:55 - Rob Singleton

Status:	Feedback		
Priority:	Normal		
Assignee:			
Category:			
Affected version:	5.7.2	**Resolution:**	

Description

I'm developing a very mobile, dynamic network. Let's say nodes A and B open up an IPSec tunnel between them. Users attached to A can communicate with users on B and vice-versa. Then node C comes online and attaches to node B. I have a service that detects node C's presence and adds a swanctl.conf entry for it and the tunnel comes up. B tells C about the subnets for A and B. Now users on B can communicate with users on C. The problem is that A doesn't know about C and traffic will not flow between nodes A and C unless I take down the A-B tunnel, reconfigure it to add the subnet info for C, and bring the tunnel back up. This interferes with the A-B traffic and is very disruptive. I can't use a static configuration because the network is mobile. Node A might drop it's connection to B and physically move closer to C so that it connects to C next. So now the network looks like B-C-A instead of A-B-C.

My question is whether there is a way to reconfigure the tunnel routing rules without taking down, reconfiguring, and re-initializing the tunnel. I tried using "ip route add" commands but that didn't work.

History

#1 - 11.01.2021 15:08 - Tobias Brunner

- Status changed from New to Feedback

> My question is whether there is a way to reconfigure the tunnel routing rules without taking down, reconfiguring, and re-initializing the tunnel. I tried using "ip route add" commands but that didn't work.

Routes have nothing to do with IPsec policies. If the latter change, you could negotiate a new CHILD_SA with new traffic selectors and then terminate the old one (you don't have to terminate that first). But you might want to look into route-based VPNs.

#2 - 12.01.2021 21:09 - Rob Singleton

- File b_ifconfig added

- File b_ip_addr added

- File b_ip_route added

- File b_ipsec0.pcap added

- File b_swanctl.conf added

- File b_swanctl_l added

- File b_wwan0.pcap added

- File d_ifconfig added

- File d_ip_addr added

- File d_ip_route added

Thank-you for the quick response. I have tried a few things but can't get it to work.

| *If the latter change, you could negotiate a new CHILD_SA with new traffic selectors and then terminate the old one (you don't have to terminate that first).*

I don't see how to do this. There are three nodes: B, D, and E. Here were my steps:

1. Node D: swanctl -q
2. Node B: swanctl -q;swanctl -i b-d -c b-d
   The tunnel is up and I can ping D from B and B from D. I want to add node E off of node D. I bring up the tunnel from D to E. Now I want to add E's subnet to B so that I can ping from B through D to E and vice versa.
3. Edit D's swanctl.conf and add E's subnet to child b-d's local_ts

4. Edit B's swanctl.conf and add E's subnet to child b-d's remote_ts
5. Node A: swanctl ~~q;swanctl -i --child b-d <--~~ this failed:
   [IKE] establishing CHILD_SA b-d{11}
   [ENC] generating CREATE_CHILD_SA request 19 [SA No TSi TSr]
   [NET] sending packet: from 192.168.18.206[4500] to 192.168.18.1[4500] (272 bytes)
   [NET] received packet: from 192.168.18.1[4500] to 192.168.18.206[4500] (80 bytes)
   [ENC] parsed CREATE_CHILD_SA response 19 [N(TS_UNACCEPT)]
   [IKE] received TS_UNACCEPTABLE notify, no CHILD_SA built
   [IKE] failed to establish CHILD_SA, keeping IKE_SA
   initiate failed: establishing CHILD_SA 'b-d' failed

If I take down the b-d tunnel and re-add it, it works:

1. Node B: swanctl -t --ike b-d;swanctl -i b-d -c b-d

| *But you might want to look into* route-based VPNs.
I'm running StrongSwan 5.7.2 which can't do xfrm, so I tried vti. What is confusing me is that the documentation says "Only packets that are marked accordingly will match the policies and get tunneled. For other packets the policies are ignored. Whenever a packet is routed to a VTI device it automatically gets the configured mark applied so it will match the policy and get tunneled." I brought up the tunnel with mark_in and mark_out specified, but if I do a simple ping without having done any of the vti steps, **the pings still go through**. I thought they would be dropped without the mark. I added the vti with:
Node B:

1. ip tunnel add ipsec0 local 192.168.18.206 remote 192.168.18.1 mode vti key 18 (I also tried local 192.168.19.1)
2. sysctl -w net.ipv4.conf.ipsec0.disable_policy=1
3. ip link set ipsec0 up
4. ip route add 192.168.18.1 dev ipsec0 src 192.168.19.1
   Node D:
5. ip tunnel add ipsec0 local 192.168.18.1 remote 192.168.18.206 mode vti key 18 (I also tried local 192.168.18.1)
6. sysctl -w net.ipv4.conf.ipsec0.disable_policy=1
7. ip link set ipsec0 up
8. ip route add 192.168.19.1 dev ipsec0 src 192.168.18.1

Node B:
ping ~~-Ipsec0 192.168.18.1 <--~~ no response

Node D:
ping ~~-Ipsec0 192.168.19.1 <--~~ no response

Neither of the pings work. tcpdump shows the pings at the ipsec0 interface, but not at the wwan0 tunnel.

I am attaching the relevant files.

#3 - 13.01.2021 10:01 - Tobias Brunner

- Affected version changed from 5.9.1 to 5.7.2

1. Edit D's swanctl.conf and add E's subnet to child b-d's local_ts
2. Edit B's swanctl.conf and add E's subnet to child b-d's remote_ts

I assume you successfully reloaded the connections on both ends? Did you change anything else? If not, this should replace the child configs in the existing peer configs (i.e. also on the active IKE_SA). Are the subnets correctly listed in swanctl --list-conns on both ends?

> [ENC] parsed CREATE_CHILD_SA response 19 [N(TS_UNACCEPT)]
> [IKE] received TS_UNACCEPTABLE notify, no CHILD_SA built

As always, if you receive an error notify, read the log on the other end (increase the log level for cfg to 2 there to see more details about the TS negotiation).

> I brought up the tunnel with mark_in and mark_out specified, but if I do a simple ping without having done any of the vti steps, **the pings still go through**.

That doesn't sound right. Check that the marks are applied on the policies/SAs via ip -s xfrm state|policy (also check the use time and packet counters to see if they are actually used and that the packets are not sent unencrypted somehow).

1. ip tunnel add ipsec0 local 192.168.18.206 remote 192.168.18.1 mode vti key 18 (I also tried local 192.168.19.1)

You have to use the actual source IP of the IPsec SA (presumably 192.168.18.206).

> Neither of the pings work. tcpdump shows the pings at the ipsec0 interface, but not at the wwan0 tunnel.

What policies were negotiated between these endpoints? Just the two 192.168 subnets or 0.0.0.0/0? Did you still configure an interface in swanctl.conf? (If so, don't.)

#4 - 04.08.2021 20:58 - Rob Singleton

I got this working by running a GRE tunnel over IPSec, and changing the routing rules outside of the IPSec configuration. That's working well now and five nodes in a mesh network can communicate with each other.

My new problem is that I want to take the connection down and contact the same host via unencrypted means. I take down the IPSec on both endpoints with:
swanctl -t -c ab
and
swanctl -t --ike ab
Then I remove the configuration in the /etc/swanctl/swanctl.conf files on both endpoints and do
swanctl -q
But then I can't ping my host. tcpdump shows no packets going out the interface. "ip route" shows that there's a route for that IP through the correct interface. There are no routes in table 220. The firewall is default ACCEPT all. It seems like Linux still believes data for this interface needs to go through IPSec but I'm not sure where to look for this. Any ideas?

#5 - 12.08.2021 14:29 - Tobias Brunner

> My new problem is that I want to take the connection down and contact the same host via unencrypted means.

Does it work before initiating the IPsec connection?

> I take down the IPSec on both endpoints with:
> swanctl -t -c ab
> and
> swanctl -t --ike ab

The latter is enough, terminating an IKE_SA will terminate all associated CHILD_SAs.

> Then I remove the configuration in the /etc/swanctl/swanctl.conf files on both endpoints and do
> swanctl -q

I guess that's not really necessary (unless you use trap policies, in which case you should probably do this before terminating the running SAs as there would otherwise be a chance of them getting initiated again).

> But then I can't ping my host. tcpdump shows no packets going out the interface. "ip route" shows that there's a route for that IP through the correct interface. There are no routes in table 220. The firewall is default ACCEPT all. It seems like Linux still believes data for this interface needs to go through IPSec but I'm not sure where to look for this. Any ideas?

What hosts and interfaces are you referring to? It's quite unclear how this correlates with GRE/IPsec exactly. So please provide details about the configs, routes, interfaces, IP addresses etc. (see HelpRequests).

Files

b_ifconfig	2.65 KB	12.01.2021	Rob Singleton
b_ip_addr	2.11 KB	12.01.2021	Rob Singleton
b_ipsec0.pcap	1.17 KB	12.01.2021	Rob Singleton
b_ip_route	2.17 KB	12.01.2021	Rob Singleton
b_swanctl.conf	528 Bytes	12.01.2021	Rob Singleton
b_swanctl_l	636 Bytes	12.01.2021	Rob Singleton
b_wwan0.pcap	656 Bytes	12.01.2021	Rob Singleton
d_ifconfig	2.66 KB	12.01.2021	Rob Singleton
d_ip_addr	2.12 KB	12.01.2021	Rob Singleton
d_ip_route	2.17 KB	12.01.2021	Rob Singleton