# strongSwan - Bug #436

## ip xfrm policy creation

03.11.2013 18:32 - Ralf Rüther

| | | | |
|---|---|---|---|
| **Status:** | Closed | **Start date:** | 03.11.2013 |
| **Priority:** | Normal | **Due date:** | |
| **Assignee:** | Tobias Brunner | **Estimated time:** | 0.00 hour |
| **Category:** | | | |
| **Target version:** | 5.1.2 | | |
| **Affected version:** | 5.1.1 | **Resolution:** | Fixed |

**Description**

I've problems with my configuration since version 5.1.0. I currently use version 5.1.1. My setup did work as aspected until version 5.0.4.

My config consists of three hosts which all should communicate witch each other.

HostA <----IPSEC Tunnel----> HostB <----IPSEC Tunnel----> HostC

Since version 5.1.0 HostA can't communicate with HostC and vice versa. HostA and HostC can each ping HostB and HostB can ping the two other hosts.

I did a packet inspection with wireshark, including packet decryption, on host A. When I try to ping HostC from HostA, I didn't see any icmp packets for HostC leaving HostA.
In my opinion the problem is related to wrong ip xfrm policies. When I manually remove the transport polices templates, the icmp messages are leaving HostA and are tunneld to HostB.

Is it a bug or is my config incomplete for newer versions of strongswan?

### Associated revisions

**Revision 6b955657 - 23.01.2014 10:31 - Tobias Brunner**

Merge branch 'ipcomp'

Fixes compatibility issues between firewall rules (leftfirewall=yes)
and IPComp (compress=yes), plus issues with IPComp when used with
multiple subnets in left|rightsubnet.

Fixes #436.

### History

**#1 - 03.11.2013 18:36 - Ralf Rüther**

I forgot to say: The hosts A and C are natted.

**#2 - 05.11.2013 12:11 - Tobias Brunner**

*- Status changed from New to Feedback*

*- Assignee set to Tobias Brunner*

The policies look that way because of *compress=yes*. The tunneling is already handled by the IPComp SA so the ESP SA is installed in transport mode.

Are strongSwan versions the only thing you changed? Or did you also use a different kernel with 5.0.4? It looks like there haven't been any changes that would affect the installed policies since 5.0.4, so I don't think they'd look different if you used 5.0.4 (to see more details about policies and SAs make sure you use ip -s xfrm state|policy). If you actually changed kernel versions then it might be a kernel issue related to IPComp.

Does the setup actually work as expected if you set *compress=no*?

**#3 - 05.11.2013 20:29 - Ralf Rüther**

*- File hostA_504.txt added*

*- File hostA_511.txt added*

First off all, disabling the compression resolves the problem under 5.1.1.

The only thing I changed, was the version of strongSwan. To analyze the policy creation, I switched back to version 5.0.4 (with compression enabled) and checked the results. I saw that transport policies are only created unter 5.1.1.
The created policies for host A are attached for version 5.0.4 and 5.1.1, the setup is the same as above.

#### #4 - 06.11.2013 11:05 - Tobias Brunner

Ah, now I see. In releases before 5.1.0 IPComp was not allowed on natted connections. As you can see in the output of ip -s xfrm state in hostA_504.txt there are no IPComp SAs installed, therefore the policy does not have an IPComp template. You should also see a log message saying that IPComp is disabled due to the NAT. So you were basically running with *compress=no* in earlier releases.

Since 5.1.0 this restriction has been removed (44d9970f4). But it might be that it was there for a reason after all. Our IPComp test scenario ikev2/compress does not use NAT so we did not notice anything amiss.

I will do some tests later today to try to determine what the exact problem is and if there is a fix.

#### #5 - 08.11.2013 12:39 - Tobias Brunner

*- Tracker changed from Issue to Bug*

*- Status changed from Feedback to Closed*

*- Target version set to 5.1.2*

*- Resolution set to Fixed*

Alright, I have some more information on the IPComp issue. There are actually two issues that might affect this situation.

First, there is a peculiarity of the Linux kernel when sending small packets over an IPComp SA (e.g. default-sized pings). For packets smaller than a certain threshold the kernel does not do any compression. If tunnel mode is used the kernel uses an implicitly created IPIP tunnel. You can see those SAs listed in ip xfrm state with proto 4 (you'll also note that the kernel uses the IPIP tunnel only for inbound traffic, for outbound traffic the IPComp SA is used, it simply does no compression for small packets). The thresholds are statically defined in net/xfrm/xfrm_algo.c for each algorithm. For the default, *deflate*, it is 90 bytes. For IPv6 the kernel also creates such SAs (proto 41) but does not seem to use them at all.

What the above means is that if you use a default DROP policy on your hosts, the firewall rules installed by the *updown* script with *leftfirewall=yes* are not in all cases enough to allow such traffic. That's because after decryption of such packets the firewall sees them with the tunnel IPs. So the rules installed by the script, that contain the *tunneled* IPs, won't match these packets and so they'll get dropped. This is the case for host-net and net-net tunnels but even for host-host tunnels if a NAT is involved. While the host behind the NAT would not need any additional rules, on the public host the firewall rule covers only the address behind the NAT and not the address of the NAT device which the header will contain after decryption, so the packets get dropped.

The second issue was caused by how the *kernel-netlink* plugin installed tunnel mode SAs with IPComp. As I mentioned earlier, only the IPComp SA is installed in tunnel mode, the ESP SA is switched to transport mode. The problem with that is that the plugin usually configures a traffic selector with transport mode SAs. But if the original SA was in tunnel mode it could have more than one selector associated with it (take your case as an example, a host-host and a host-net policy that both use the same IPsec SA). Since only one selector can be installed on an SA, traffic that would match traffic selectors that are not installed would get dropped.

I pushed fixes for both issues (and some test scenarios) to the *ipcomp* branch of our repository.

### Files

| | | | |
|---|---|---|---|
| hostA_ipsec.conf | 694 Bytes | 03.11.2013 | Ralf Rüther |
| hostB_ipsec.conf | 1 KB | 03.11.2013 | Ralf Rüther |
| hostC_ipsec.conf | 649 Bytes | 03.11.2013 | Ralf Rüther |
| hostA_ip_xfrm_policy.txt | 2.02 KB | 03.11.2013 | Ralf Rüther |
| hostB_ip_xfrm_policy.txt | 3.68 KB | 03.11.2013 | Ralf Rüther |
| hostC_ip_xfrm_policy.txt | 1.82 KB | 03.11.2013 | Ralf Rüther |
| hostA_504.txt | 11.6 KB | 05.11.2013 | Ralf Rüther |
| hostA_511.txt | 17.3 KB | 05.11.2013 | Ralf Rüther |