strongSwan - Feature #1291

Avoid packet loss during IKEv2 CHILD_SA rekeying

01.02.2016 15:12 - Avinoam Meir

Status:	Closed	**Start date:**	
Priority:	Normal	**Due date:**	
Assignee:	Tobias Brunner	**Estimated time:**	0.00 hour
Category:	libcharon		
Target version:	5.5.3		
Resolution:	Fixed		

Description

I have question/proposal about CHILD SAs rekey:
If I understand correctly, today in rekey task, after creating the new CHILD SA, immediately delete task is created and executed. (see here).

This can cause packets loss If the peer gateway sends ESP packets in parallel to the rekey, so there are some old ESP packet on the network.

Maybe StrongSwan can defer the call to kerne_interface->del_sa() for the inbound CHILD SA (only), so the kernel continue to process esp packets for the old SAs for a while, and prevent the packet loss.

Related issues:

Related to Issue #3310: Packet loss observed at rekey, hence request for patc...	**Closed**	
Has duplicate Issue #2242: Packets get dropped due to no no_sa_found during r...	**Closed**	**04.02.2017**

Associated revisions

Revision f8eb636e - 23.05.2017 18:49 - Tobias Brunner

Merge branch 'avoid-rekey-loss'

This changes the behavior during IKEv2 CHILD_SA rekeyings to avoid traffic loss. When responding to a CREATE_CHILD_SA request to rekey a CHILD_SA the responder already has everything available to install and use the new CHILD_SA. However, this could lead to lost traffic as the initiator won't be able to process inbound packets until it processed the CREATE_CHILD_SA response and updated the inbound SA. To avoid this the responder now only installs the new inbound SA and delays installing the outbound SA until it receives the DELETE for the replaced CHILD_SA. The messages transporting these DELETEs could reach the peer before packets sent with the deleted outbound SAs reach the respective peer. To reduce the chance of traffic loss due to this the inbound SA of the replaced CHILD_SA is not removed for a configurable amount of seconds after the DELETE has been processed.

Fixes #1291.

History

#1 - 02.02.2016 16:36 - Tobias Brunner

- Status changed from New to Feedback

Thomas already addressed some aspects of this on the dev mailing list but this is a very general problem with IPsec rekeying and also described in RFC 7296 section 2.8 (last three paragraphs).

Deleting the old SA a bit later might be an option (would require using state CHILD_REKEYED also for IKEv2) but it would not really solve this for all situations. In particular, the responder of a rekeying still installs and most likely uses (the kernel usually uses the latest SA) the new outbound SA, which could cause the initiator to drop some packets as it can only install the SA after it received the CREATE_CHILD_SA response. We could delay the installation of the SAs too but the responder MUST be prepared to receive inbound packets on the new SA as soon as it sent the CREATE_CHILD_SA response, otherwise a similar situation would ensue. Now you could argue that we could split the installation of the SAs (i.e. install the inbound SA rightaway, but delay the installation of the outbound SA for a while). That would, however, require quite some refactoring and complicating of the code.

Since IP packets may get dropped anyway and IPsec basically has the same semantic it's up to the layers above to compensate for that.

#2 - 03.02.2016 12:01 - Avinoam Meir

Thank you for the detailed response.

#3 - 12.10.2016 11:15 - Emeric Poupon

Tobias Brunner wrote:

> Now you could argue that we could split the installation of the SAs (i.e. install the inbound SA rightaway, but delay the installation of the outbound SA for a while). That would, however, require quite some refactoring and complicating of the code.

I guess we could install the new outbound SA just before deleting the old one in the DELETE message processing? (proposal (2) in the last paragraph of https://tools.ietf.org/html/rfc7296#section-2.8)

You said it requires quite some refactoring and complicates the code: could you please tell what you would do to make this work?

#4 - 07.12.2016 19:05 - Bernhard Kaindl

I and "pradeep kumar nalla" (see this mail) seem to have the same issue on CHILD_SA rekey:
https://groups.google.com/forum/#!topic/strongswan-users/_SpX4fSdyIg

While some packet loss must be expected over the internet and overloaded links, it should never happen on switched (v)LANs when the traffic budget is far from saturated. Direct NIC2NIC full-duplex links not even using a switch is another use case which I also need to solve.

In the latter HW configuration (IPSec over Ethernet NICs connected in full-duplex mode), I reproduced 100% packet loss for 80-100 msec (timed using ping -i0.01 - that is 8-10 packages in a row in that relatively slow speed) with the IKEv2-based net2net KVM test example configs, in both ipsec.conf (with the in this case unrelated strongswan.conf{charion.make_before_break=yes} and the swanctl.conf configs.

Interruptions of traffic in that time span are a problem for e.g. for control applications requiring real-time network connectivity for syncronisation purposes.

Tobias, could you describe what needs to be done at a minimum to keep the old CHILD_SAs until all old packages are received by both ends so that 0% packet loss from CHILD_SA rekeying would be guaranteed?

#5 - 06.02.2017 10:11 - Tobias Brunner

- Has duplicate Issue #2242: Packets get dropped due to no no_sa_found during rekeying added

#6 - 18.04.2017 15:14 - Tobias Brunner

I've implemented some (prototypical) changes that try to avoid packet loss during rekeyings. These include two major behavioral changes. First, the responder of a rekeying will not immediately install the complete CHILD_SA when responding to the CREATE_CHILD_SA exchange. Instead, it installs only the inbound SA and then waits for the delete for the replaced SA, at which point it assumes the initiator installed its inbound SA and it is safe to install the outbound SA. Second, the deleted CHILD_SA is not completely uninstalled immediately (on initiator and responder). Instead, only the outbound SA is uninstalled and the inbound SA is kept around for a few seconds (configurable, the default is 5) to process any delayed messages.

If you are interested, please try the code in the *1291-avoid-rekey-loss* branch and let me know if you find any issues or think something should be changed.

#7 - 27.04.2017 10:17 - Emeric Poupon

Tobias Brunner wrote:

> If you are interested, please try the code in the *1291-avoid-rekey-loss* branch and let me know if you find any issues or think something should be changed.

Hello,
As I said in the dev ML:
- it seems there is no more packet loss during the CHILD SA rekeying :)
- I noticed some drop during the IKE SA reauthentication, despite the make_before_break option set to yes.

Thanks again for your support

#8 - 24.05.2017 15:37 - Tobias Brunner

- Tracker changed from Issue to Feature

- Subject changed from Packets loss during rekey to Avoid packet loss during IKEv2 CHILD_SA rekeying

- Status changed from Feedback to Closed

- Assignee set to Tobias Brunner

- Target version set to 5.5.3

- Resolution set to Fixed

#9 - 14.01.2020 16:08 - Tobias Brunner

- Related to Issue #3310: Packet loss observed at rekey, hence request for patch of Feature #1291 added