

New Results for the PTB-PTS Attack on Tunnelling Gateways

Vincent Roca¹, Ludovic Jacquin³, Saikou Fall¹, and Jean-Louis Roch²

¹ Inria, France, {vincent.roca}@inria.fr

² Inria, Grenoble Université, Grenoble INP, LIG, France, jean-louis.roch@imag.fr,

³ HP Labs, Bristol, England, {ludovic.jacquin@hp.com}

Abstract. This work analyzes the impacts of the "Packet Too Big"- "Packet Too Small" (PTB-PTS) Internet Control Message Protocol (ICMP) based attack against tunneling gateways. It is a follow up of a prior work [2] that detailed how to launch the PTB-PTS attack against IPsec gateways (for secure tunnels) and their consequences, ranging from major performance impacts (additional delays at session establishment and/or packet fragmentation) to Denial of Services (DoS).

In the present work we examine a much wider range of configurations: we now consider the two IP protocol versions (previous work was limited to IPv4, we add IPv6), two operating systems (previous work was limited to Linux Debian, we add a recent Ubuntu distribution as well as Windows 7), and two tunnelling protocols (previous work was limited to IPsec, we add IPIP).

This work highlights the complexity of the situation as different behaviors will be observed depending on the exact configuration. It also highlights Microsoft's strategy when approaching the "minimum maximum packet size" (i.e., minimum MTU) any link technology should support: if Windows 7 mitigates the attack in IPv4 (there is no DoS), however the performance impact is present and the technique is inapplicable to IPv6. Finally, it highlights a fundamental problem: the impossibility to identify illegitimate ICMP error packets coming from the untrusted network.

Keywords: Tunneling gateway, IPsec, IPIP, ICMP "Packet Too Big"

1 Introduction

This work analyzes the impacts of the "Packet Too Big"- "Packet Too Small" (PTB-PTS) Internet Control Message Protocol (ICMP) based attack against tunneling gateways. The idea behind the attack is to take advantage of both:

- the "minimum maximum packet size" any link technology should support (officially called "minimum Maximum Transmission Unit, or "minimum MTU"), namely 576 bytes with IPv4 and 1280 bytes with IPv6;
- the necessary addition of headers at a tunnelling gateway that increases the packet size after encapsulation;

The attack principle is to reduce the MTU along the path at the tunneling gateway down to the "minimum MTU", so that the gateway prevents any client from sending packets of size larger than the "minimum MTU" minus the required size for the tunneling headers. Confusion is then created between the gateway and the clients that are asked to use packets of size lower than the "minimum MTU", hence performance impacts and sometimes deadlocks resulting in a Denial of Service (DoS). This is illustrated in Figure 1 that shows the local client (A) and gateway (G) that is targeted by the attacker (not shown, but located in the Internet).

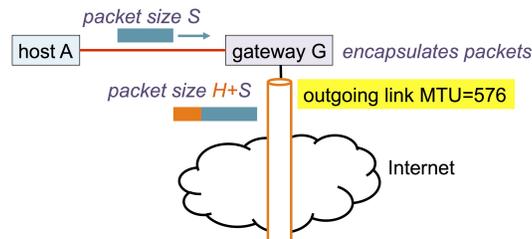


Fig. 1: Attack principles: dealing with an MTU lower than the minimum authorized MTU.

This work is a follow up of a prior work [2] that detailed how to launch the PTB-PTS attack against IPsec gateways (in ESP/tunneling mode) and their consequences in case of a Linux clients. In the present work we examine a much wider range of configurations. In addition to the transport protocol (TCP versus UDP) and path MTU discovery mechanisms (PMTUd [6] versus PLPMTUd [3]) parameters, both already considered in previous work, we now consider:

- two IP protocol versions (previous work was limited to IPv4, we add IPv6);
- two operating systems (previous work was limited to Linux Debian, we add a recent Ubuntu distribution as well as Windows 7);
- and two tunnelling protocols (previous work was limited to IPsec, we add IPIP).

The contributions of the present work are three-fold:

- this work first highlights the **complexity of the situation**, as totally different behaviors are observed depending on the exact configuration;
- it highlights **Microsoft limited solution** when fragmentation is needed. If it mitigates the attack in IPv4 (no DoS), however the performance impact is present and the technique remains inapplicable to IPv6;
- finally, it highlights a **fundamental problem**: the impossibility to identify illegitimate ICMP error packets coming from the untrusted network. If approaches exist to mitigate it, there is no real solution to the core problem.

It differs from the previous attacks on ICMP (a known vector of attacks) by the fact that ICMP is not the target, it is only an intermediate to launch the attack, and a single valid (i.e., correctly formatted) ICMP error packet is sufficient.

The paper is organized as follows: we quickly provide some background and explain the attacker model; the following section is devoted to the PTB-PTS attack, detailing the results with the new configurations; finally we wrap-up and discuss the results. We invite the interested reader to refer to [2] for a detailed background on ICMP, on ICMP "Packet Too Big" error messages, on the two Path MTU discovery mechanisms, and the way the PTB-PTS attack can be launched in case of IPsec. In the following, we will essentially focus on the new results.

2 The attacker model

In this work, we consider that all the attacks are conducted by adversaries located on the external unsecure network, typically the Internet.

In case of IPsec, an IPsec gateway can filter out any ICMP error packet not coming from the path, by checking the copy of the packet that needs to be present in the payload of the ICMP error packet. This is performed by decrypting this packet and checking that it refers to an active IPsec Security Association in the local database. This way, IPsec avoids attacks initiated by attackers that cannot intercept legitimate packets, typically attackers that are not in the path followed by IPsec packets.

Such a security mechanism does not exist with the IPIP tunneling mechanism that we also consider in this work. In that case, any attacker capable of forging and sending an ICMP PTB error packet can launch the attack, which greatly simplifies the attack compared to the IPsec case.

To summarize, in case of IPsec, we assume an attacker can both eavesdrop the traffic in the IPsec tunnel and inject a forged packet (a single packet is sufficient). However the attacker has no way to decrypt packets nor encrypt its own packets (the underlying IPsec cryptographic building blocks and key exchange protocols are considered secure). For instance, the attacker can be located on a compromised router along the path followed by an IPsec tunnel, in the external unsecure black network. But more simply, the attacker can also be attached to the same unsecure WiFi network (e.g., that sends WiFi frames in the clear, without any WPA/WPA2 security) as the target user that connects to his home network through an IPsec VPN.

In case of the IPIP tunneling protocol, the attack is much simpler to launch, and the attacker only has to know the IP address of the IPIP tunnel. We also assume the attacker can inject a forged packet (here also, a single packet is sufficient).

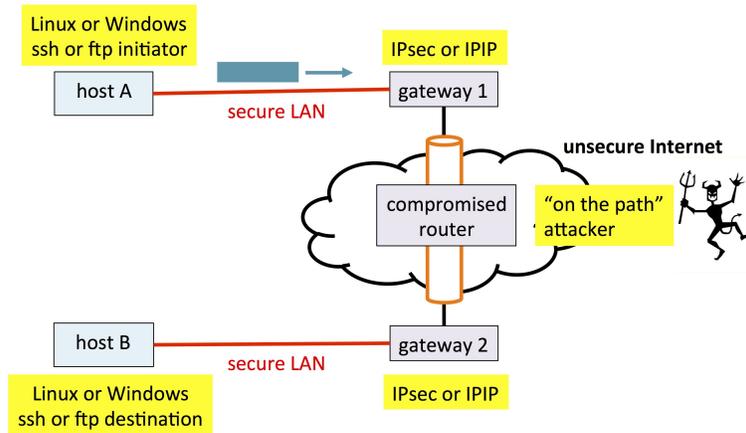


Fig. 2: Test configuration showing the compromised router (attacker), the two tunneling gateways, and the two end-hosts.

3 Launching the PTB-PTS attack

3.1 Test configuration

The results collected in this document have been achieved with the configuration depicted in 2. Five virtual machines are created (with VirtualBox). The compromised router and two gateways use Ubuntu 14.04.2 LTS. The IPsec gateways use the Openswan IPsec implementation with its default configuration. The two clients use either Ubuntu 14.04.2 LTS or Windows 7. The MTU on the various (virtual) links is configured to the usual 1500 byte Ethernet value.

TCP connection is tested either through ssh (that implies the transmission of keying material whose size is larger than the MTU) or FTP (that implies the transmission of a large file). With both tools, the three-way TCP connection handshake succeeds but problems may arise later when dealing with larger TCP segments.

Windows 7 host configuration is controlled by the EnablePMTUDiscovery [5] and EnablePMTUBHDetect [4] ("PMTU Black-Hole Detection") registers. The default configuration corresponds to values (1, 0) respectively (i.e., EnablePMTUDiscovery set to 1 and EnablePMTUBHDetect set to 0). We only considered this default configuration in our tests as we do not expect end-users to change it.

In the following sections, we detail the exchanges, the first two steps (1 and 2), common to any IPsec test, being factorized. Even if we do not detail all the configurations (too long), they are at least summarized in tables 2 and 2.

3.2 Step 1 (common): Forging an ICMP PTB packet from the untrusted network

The attacker first has to forge an appropriate ICMP PTB packet (a single packet is sufficient). For instance, with the IPsec configuration, this is done by eavesdropping a valid packet from the IPsec tunnel on the untrusted network. Then the attacker forges an ICMP PTB packet, specifying a very small MTU value equal or smaller than 576 with IPv4 (resp. 1280 with IPv6). The attacker can use 0 for instance as this is checked by the gateway that will revert to 576 (resp. 1280). This packet spoofs the IP address of a router of the untrusted network (in case the source IP address is checked), and in order to bypass the IPsec protection mechanism against blind attacks, it includes as a payload a part of the outer IP packet that has just been eavesdropped. This is the only packet an attacker needs to send, and none of the following steps involve the attacker any more. The process is similar (and even simpler) in case of IPIP.

3.3 Step 2 (common): Reset of the PMTU on the gateway

This ICMP packet is processed by the IPsec gateway. As the packet appears to belong to an active tunnel, the gateway stores the following PMTU value in its SAD: $PMTU_{SAD} = \max(MTU_{ICMPPTB}, 576) = 576$ bytes. It is important to note that the gateway does not store a proposed value smaller than the minimum guaranteed MTU, 576 (resp. 1280) bytes.

At this point, the traffic is not blocked in any way between the targeted gateway and the remote end of the tunnel. Nevertheless the throughput is reduced on the IPsec tunnel as any packet exceeding the `PMTU_SAD` size must be fragmented (usually by the end-host).

3.4 Following steps with Linux, TCP/IPv4 and PMTUd

The following steps depend on the end-host client Operating System (OS), IP version, and MTU discovery protocol. In case of Linux (Ubuntu 14.04.2 LTS is the OS of all the machines, end-hosts and IPsec gateway), TCP (i.e., during an ssh connection attempt to the remote end-host), IPv4, PMTUd, we observe the following.

The TCP 3-way handshake performs normally as all TCP segments are of tiny size, which enables the attacker to intercept a packet on the IPsec tunnel and to perform the above attack (steps 1 and 2). Then a large packet is sent with the IPv4 Don't Fragment (DF) bit turned on.

This packet gets rejected by the IPsec gateway as it exceeds the `PMTU_SAD` value stored in the SAD, and an ICMP PTB error packet is sent back with the following MTU indication: $MTU = PMTU_{SAD} - \text{sizeof}(IP + IPsec + ESPheaders)$. Due to the encapsulation header (whose size depends on the chosen ciphering algorithm), the gateway restricts the MTU value to 502 bytes.

Upon receiving this ICMP PTB packet, the large TCP segment is fragmented. Nevertheless, instead of creating 502 byte long packets as requested by the gateway, TCP chooses to reduce the MSS to 500 bytes only as it considers the

value advertised by the ICMP PTB is below what should be accepted by any link. More precisely, with Linux there is a minimum PMTU configuration parameter (e.g., `cat /proc/sys/net/ipv4/route/min_pmtu` returns 552 on Debian "Squeeze") that is preferred to the value advertised by the ICMP PTB message: $PMTU = \max(MTU_{ICMPPTB}, PMTU_{config})$. So, once the TCP/IP headers are added, the 500 byte long TCP segment results in a 552 byte long packet.

Since it remains too large, the packet is dropped by the gateway and this latter replies with the same ICMP PTB packets, with the same result.

After 2 minutes of failures and a total of 10 re-transmission attempts, the ssh server closes the connection (FIN/ACK exchange quickly followed by a RST). The DoS successfully prevented any ssh setup.

3.5 Following steps with Linux, TCP/IPv4 and PLPMTUd

In case of TCP/IPv4 and PLPMTUd, we observe the following.

The TCP 3-way handshake performs normally. Then a large packet is sent with the IPv4 Don't Fragment (DF) bit turned one.

This packet gets rejected by the IPsec gateway and an ICMP PTB is returned to the end-host that restricts the MTU to 502 bytes.

Although PLPMTUd is not dependant on ICMP, this error message is immediately taken into account and several TCP segments of maximum size 500 bytes are sent. Once the TCP/IP headers are added, the 500 byte long TCP segment results in a 552 byte long packet.

Since it remains too large, the packet is dropped by the gateway and this latter replies with the same ICMP PTB packets, with the same result.

This pattern happens 5 times, generating a total of 6 ICMP PTB packets including the first one.

Then, 6.3 seconds after the TCP connection establishment, the PLPMTUd component decides to drastically reduce the segment size: instead of 500 byte TCP segments, it now sends a sequence of alternatively 256 byte TCP segment followed by a 244 byte TCP segment. Those segment are typically probes, chosen by PLPMTUd in order to test this value.

Since the resulting packets are Small enough (at most $256 + 52 = 308$ bytes), they reach the other side that acknowledges them. The ssh connection finishes after a few additional segments and a prompt appears in the terminal.

To conclude a delay of 6.3s was required for the ssh connection to be setup. Additionally, any packet leaving the host after this initial delay contains at most 256 bytes of TCP payload, which significantly reduces the TCP throughput and consumes more resources in the forwarding nodes.

3.6 Following steps with Linux, TCP/IPv6 and PMTUd

In case of TCP/IPv6 and PMTUd, we observe the following.

The situation is pretty the same as with IPv4. The main difference is the ICMP PTB packet that advertises a MTU of 1198 bytes (i.e., 1280 minus the

various IPsec encapsulation headers). Upon receiving this ICMP PTB packet, the large TCP segment is fragmented into TCP segments of maximum size 1200 bytes. Once the TCP/IPv6 headers are added, it results into packets of maximum size 1256 bytes, which is too large for the IPsec gateway.

After 2 minutes of failures and a total of 10 re-transmission attempts, the ssh server closes the connection (FIN/ACK exchange quickly followed by a RST). The DoS successfully prevented any ssh setup.

3.7 Following steps with Linux, TCP/IPv6 and PLPMTUd

In case of TCP/IPv6 and PLPMTUd, we observe the following.

The situation is pretty the same as with IPv4. However upon receiving the ICMP PTB packet that advertises a MTU of 1198 bytes, the end-host first tries to use TCP MSS=1200 which is too large for the IPsec gateway. A total of 4 re-transmissions happen, generating a total of 5 ICMP PTB packets including the first one.

Then, 3.3 seconds after the beginning, the end-host tries with TCP MSS=504. Once the TCP/IPv6 headers are added, it results into packets of maximum size 560 bytes, which is acceptable for the IPsec gateway. The ssh connection finishes after a few additional segments and a prompt appears in the terminal.

To conclude a delay of 3.3s was required for the ssh connection to be setup (compared to 6.3 in case of IPv4). Additionally, any packet leaving the host after this initial delay contains at most 504 bytes of TCP payload, far below the 1280 minimum MTU guaranteed by IPv6. This behavior significantly reduces the TCP throughput and consumes more resources in the forwarding nodes.

3.8 Following steps with Windows 7, TCP/IPv4 and default configuration

In case of TCP/IPv4 and PMTU-1-0 (default configuration), we observe the following.

The TCP 3-way handshake performs normally. Then a large packet is sent with the IPv4 Don't Fragment (DF) bit turned one. This packet gets rejected by the IPsec gateway which returns an ICMP PTB with the MTU value to 502 bytes.

Upon receiving this ICMP PTB packet, the Windows 7 end-host sends a smaller TCP segment, of size 556 bytes (instead of 502 bytes). Once the TCP/IP headers are added, this TCP segment results in a 596 byte long packet.

This packet is still too large for the IPsec gateway, however the IPv4 DF bit is now turned off, which authorizes the IPsec gateway to perform IP fragmentation. It therefore gets fragmented by IP within the IPsec gateway, then reassembled on the other side of the tunnel, and acknowledged.

Upon receiving this TCP acknowledgment, the PLPMTUd mechanism starts (Windows 7 merges PMTUd and PLPMTUd like mechanisms together). The following TCP segment is now of size 1112 (i.e., 1152 with TCP/IPv4 headers),

here also with the DF bit turned off. This large packet therefore gets fragmented by IP within the IPsec gateway, then reassembled on the other side of the tunnel, and acknowledged by the remote TCP.

The process continues, with TCP segment sizes that progressively increase (we observed a maximum value of 63,940 bytes!), always with the DF bit turned off. All of them are IP fragmented into small IP datagrams of size 548 bytes or 120 bytes. The traffic on the IPsec tunnel is therefore composed of a many tiny packets (never more than 548 bytes long), which creates a huge performance penalty in case of high rate data flows.

3.9 Following steps with Windows 7, TCP/IPv6 and default configuration

In case of TCP/IPv6 and PMTU-1-0 (default configuration), we observe the following.

The TCP 3-way handshake performs normally. Then a large packet is sent. This packet gets rejected by the IPsec gateway which returns an ICMP PTB with the MTU value set to 1198 bytes (as in Section 3.6).

Upon receiving this ICMP PTB packet, TCP segments have maximum size 1212 bytes (1276 bytes with the TCP/IPv6 headers), which is too large for the IPsec gateway. However the end-host, upon receiving the same ICMP PTB packet, keeps on using MSS=1212, resulting in the same problem.

After 10 transmission attempts and 21 seconds, the ftp client closes the connection (with a RST). The DoS successfully prevented any ssh setup.

3.10 Summary of the results

Tables 1 and 2 summarize the consequences of a PTB-PTS attack on a end-host depending on the exact configuration and tunnelling protocol in use. It highlights the large diversity of consequences for the same attack (i.e., the attacker follows the same approach and issues the same forged ICMP PTB packet (v4 or v6) in all IPsec (resp. IPIP) attacks.

4 Discussion

This work highlights two issues:

- **Issue 1: dealing with a small Path MTU in presence of a tunnel:**
When the Path MTU advertised to a tunneling gateway through an ICMP PTB packet approaches the "minimum MTU" (i.e., 576 or 1280 bytes), problems can arise as IPsec and IPIP tunneling need to add encapsulation headers. There are two sides to the problem:
 - on the one side the IPsec gateway should not accept a request to reduce its Path MTU if, after encapsulation, the MTU advertised to end-hosts is below the "minimum MTU". This is typically a situation where an alarm should be sent to the IPsec gateway administrator, which is not the case today;

<i>IPsec tunnel</i>	<i>Results of a PTB-PTS attack</i>
TCP/IPv4, PMTUd	DoS: no ssh connection setup (TCP close after 2 mn)
TCP/IPv4, PLPMTUd	Major performance impacts: initial 6.3 sec delay, then tiny packets (TCP MSS=256)
UDP/IPv4, PMTUd	Major performance impacts: tiny packets
TCP/IPv6, PMTUd	DoS: no ssh connection setup (TCP close after 2 mn)
TCP/IPv6, PLPMTUd	Important performance impacts: initial 3.3 sec delay, then small packets (TCP MSS=504)
<i>IPIP tunnel</i>	<i>Results of a PTB-PTS attack</i>
TCP/IPv4, PMTUd	Important performance impacts: initial 7 min. delay, then small packets
TCP/IPv4, PLPMTUd	Important performance impacts: initial 6.7 sec delay, then small packets

Table 1: Results for Linux Ubuntu 14.04.02 LTS clients.

<i>IPsec tunnel</i>	<i>Results of a PTB-PTS attack</i>
TCP/IPv4, default (1-0)	Functional, but performance impacts (548 and 120 byte IP fragments)
TCP/IPv6, default (1-0)	DoS: no ftp transfer possible (TCP close after 21 sec)
<i>IPIP tunnel</i>	<i>Results of a PTB-PTS attack</i>
TCP/IPv4, default (1-0)	DoS: no ssh connection setup (TCP close after 35 sec)

Table 2: Results for Windows 7 clients.

- on the other side, the end-host is ignorant of the need to keep room for the encapsulation headers at the gateway and it rejects any request to go below the "minimum MTU". For instance, a Linux host simply ignores any Path MTU advertised by the gateway if it is smaller than the minimum MTU configured locally. If the end-host uses PMTUd, the compliance to the minimum MTU is strict and a DoS results. If the end-host uses PLPMTUd, it will take time for TCP to find an appropriate segment size (which remains significantly smaller than the one advertised by the gateway). With Windows 7, the end-host decides not to set the IPv4 DF bit any more when it encountered an error, which is of course only valid with IPv4 (there is no IPv6 datagram fragmentation possible). In any case, even if a DoS is avoided, there is a major performance impact due to packet fragmentation (at end-host or at the gateway);

- **Issue 2: determining the legitimacy of untrusted ICMP PTB packets:**
The security measures W.R.T. the processing of ICMP PTB packets, namely the outer header verification and payload verification, are essential to avoid blind attacks, but not sufficient if the attacker is on the path followed by the IPsec tunnel (and this can be pretty easy in case a client is connected through an open WiFi network, see [2]). And the situation is even worse with IPIP, as there is way to filter out attacks coming from attackers that are not on the path followed by the IPIP tunnel. No totally reliable solution exists to distinguish legitimate from illegitimate ICMP PTB error messages, which is a fundamental issue.

Let us now discuss several potential counter-measures.

4.1 Trivial unsatisfying counter-measures

A trivial counter measure to mitigate the attack consists in configuring the IPsec gateway so that IPv4 packets are fragmented regardless of the original DF bit setting. This is feasible (and recommended) with Cisco IOS 12.2(11)T and above ([1], "DF Bit Override Functionality with IPsec Tunnels" section). However, as mentioned in [1], "a significant performance impact occurs at high data rate", and ignoring the DF bit cannot be considered as a valid approach. And in any case, this is inapplicable to IPv6 flows.

Another trivial counter measure consists in ignoring all ICMP PTB packets coming from the unsecure network. However, this choice compromises PMTUd that the use of PLPMTUd within end-hosts cannot totally compensate (e.g., PLPMTUd is only applicable to protocols like TCP relying on acknowledgements, no to UDP).

4.2 Two complementary counter-measures

A more interesting choice consists in refusing to reduce the MTU below 576 (resp. 1280) after subtracting the encapsulation headers (e.g., an alarm should be sent to the gateway administrator when this happens). Doing so, the MTU advertised to end-host in ICMP PTB messages will always be at least equal to the "minimum MTU" which avoids the initial delays and DoS discussed in this work. This is not totally satisfying though, given that the attack succeeds at least in reducing the PMTU, sometimes significantly.

Therefore, in addition to refusing to reduce the MTU below the minimum MTU, a second complementary approach could be the following: since the legitimacy of an untrusted ICMP packet cannot be determined, the IPsec gateway should try to confirm the information with a side mechanism, a PLPMTUd-like probing. It could work as follows. The gateway generates a probing packet of a certain size that is sent inside the IPsec or IPIP tunnel. Upon reception, the remote gateway acknowledges this probe, otherwise a timeout occurs at the gateway that sent the probe. Therefore, if the probing mechanism does not confirm the ICMP PTB information received from the unsecure Internet, this ICMP packet

is simply ignored. For this mechanism to be effective, the attacker should not be able to identify and discard selectively the probing packets sent in the tunnel. Even if this is hard to guaranty, we can note that being able to selectively drop some packets goes far beyond the attacker model considered in this work (Section 2), making the attack significantly more complex. Finally, an IPsec gateway level probing mechanism, done periodically, cannot be as reactive as the PMTUD approach (we assume ICMP packets are not filtered out) in case of path MTU change, for instance after a route change. Therefore we believe that both mechanisms could safely work in parallel.

5 Conclusions

This work provides new results for the "Packet Too Big"- "Packet Too Small" (PTB-PTS) attack against tunneling gateways. It goes far beyond the results presented in our previous work [2] by considering several new configurations, and in particular by extending the client OS to Windows 7 in addition to a recent distribution of Linux.⁴ This work highlights the complexity of the situation, many different behaviours being observed, the limits of current solutions that never solve the performance impacts of the attack, and the fundamental problem of identifying illegitimate ICMP error packets. Some counter-measures exist and need to be deployed and tested in order to assess their practical benefits in mitigating the attack. In any case we observe that this problem is currently largely overlooked.

References

1. Cisco: IPsec Data Plane Configuration Guide, Cisco IOS Release 15M&T. Cisco Systems, Inc., http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/sec_conn_dplane/configuration/15-mt/sec-ipsec-data-plane-15-mt-book.pdf (2012)
2. Jacquin, L., Roca, V., Roch, J.L.: Too Big or Too Small? The PTB-PTS ICMP-based Attack against IPsec Gateways. In: IEEE Global Communications Conference (GLOBECOM'14). IEEE (Dec 2014), <https://hal.inria.fr/hal-01052994>
3. McCann, J., Deering, S., Mogul, J.: Path mtu discovery for ip version 6. IETF Request For Comments 1981, <http://datatracker.ietf.org/doc/rfc1981/> (August 1996)
4. Microsoft: EnablePMTUBHDetect. Microsoft TechNet, <https://technet.microsoft.com/en-us/library/cc960465.aspx>
5. Microsoft: EnablePMTUDiscovery. Microsoft TechNet, <https://technet.microsoft.com/en-us/library/cc957539.aspx>
6. Mogul, J., Deering, S.: Path mtu discovery. IETF Request For Comments 1191, <http://datatracker.ietf.org/doc/rfc1191/> (November 1990)

⁴ Note that we are currently testing with Windows 10 end-hosts. As far as we can tell, Windows 10 behaves exactly in the same way as Windows 7. However, the work being on progress, it is not considered in the present paper.