Transient pseudo-NAT attacks or
how NATs are even more evil than you believed

Abstract

When a "NAT traversal" capability is added to a class of signaling
protocols which can control some traffic aggregation points, an
attack based on a temporary access to the path followed by messages
exists.

traversal [6], including the IKEv2 [7] proposal, are potentially
affected by this kind of attacks.

This document claims this vulnerability is an intrinsic property
of the NAT traversal capability, and so is another point where the
usage of NATs is very damaging.
1. Introduction

A Network Address Translator (NAT [8]) is a device which rewrites the source address or/and destination address as well as usually the transport protocol ports of a communication. Many kinds of NATs [9] exist but in this document the term NAT will be used for any device which modifies at least one of the IP header addresses (a pseudo-NAT when this is done for an attack, i.e., we will call pseudo-NAT an attacker spoofing a NAT device).

NAT traversal capability consists in a NAT resilient transport protocol, usually UDP, and in address "agility", i.e., addresses in the header of packets are taken as they are, without control, especially the source address (packets with a fake destination address are likely to not reach their intended recipient).

A traffic aggregation point is a place where traffic from many sources and/or many destinations is aggregated and sent to the same destination. The traffic usually arrives from the same source (the traffic aggregation point) through a tunnel. Home agents in Mobile IP and security gateways in IPsec [4] are typical examples of such traffic aggregation points (which are not necessary for the attack to work but increase its impact).

2. The Transient Pseudo-NAT Attack

An attacker acting as a NAT (i.e., a pseudo-NAT) may:
- redirect packets to another node
- make the intended recipient to not receive packets (first form of Denial-of-Service (DoS) attack)
- flood a third party with the hijacked packets (second form of DoS attack, perhaps the most serious)

To perform the attack, the attacker must be on the path of packets during the attack.

When there is a traffic aggregation point, the effects of the attack are amplified when the attack is done "at the outgoing side" of the aggregation point.

When a signaling protocol manages the direction followed by the traffic, the attacker only has to spoof the addresses in the headers of some messages of the protocol in order to hijack the traffic during a long period (i.e., until an error is detected and the correct path re-established). Since the attacker has to stay on the path only for a short moment this attack is named the "transient" pseudo-NAT attack.
3. Attack Examples

3.1 Mobile IP

For Mobile IP the traffic aggregation point of choice is the home agent and the target signaling protocol is the binding update (the binding acknowledgment exchange). If the NAT traversal capability is enabled, the care-of address of the mobile may not be protected, and therefore may be easily spoofed.

If no binding acknowledgment is required, the attack can be reduced to the modification in transit of only one packet. Thus we recommend to always require acknowledgment when NAT traversal is enabled (as a weak form of return-routability check).

3.2 IKE

The context of IKE is a bit different: because of an under-specification in IKE documents, there is no standard provision for address protection and most implementations fix this security flaw in ways which clearly interfere with NAT traversal features.

The attack against IKE is worse because IKE is supposed to ensure a high level of security, unfortunately bypassed by NAT traversal which is the first short-term work item of the IETF ipsec working group charter [5]...

The attack follows the same scheme: addresses in headers of IKE exchange messages are spoofed and the traffic is hijacked.

Any improvement to the IKE protocol makes the attack easier (a very unpleasant property of this attack). For instance if an implementation supports an address change between two "phases", (something desirable and supported via the SPI of the phase one) then spoofing the two or three messages of a quick mode exchange is enough to perform the attack. In IKEv2 only one packet of a CREATE-CHILD-SA exchange is necessary to do so.

Again there is no easy way to keep the NAT traversal capability and to achieve a good level of security at the same time. For instance the protection of the header addresses (which is very easy to provide in the IKE framework) cannot work with the NAT traversal capability.

4. Security Considerations

The Mobile IP NAT traversal document has a long description of this attack [10,5]. We believe the ipsec working group will examine in details which features can help mobility or/and NAT traversal and what are their consequences for security.
The architectural implications of the NAT document [11] do not describe this attack but it can be considered a result of the violation of the end-to-end principle.

5. Acknowledgments

Maryline Maknavicius-Laurent drew my attention on this attack at the IP Cellular Network 2002 conference. Phil Roberts encouraged me to point out this attack in the IETF mobileip WG mailing-list ASAP. I’d like to thank a well known NAT hater who’d like to stay anonymous for his help to write this document. Mohan Parthasarathy helped us to clarify the context of IKE.

6. Normative References


7. Informative References


8. Authors’ Addresses

Francis Dupont
ENST Bretagne
Campus de Rennes
2, rue de la Chataigneraie
CS 17607
35576 Cesson-Sevigne Cedex
FRANCE
Fax: +33 2 99 12 70 30
EMail: Francis.Dupont@enst-bretagne.fr

Jean-Jacques Bernard
AFNIC / NIC France
Immeuble International
2 rue Stephenson
78181 Saint-Quentin-en-Yvelines Cedex
FRANCE
Fax: +33 1 39 30 83 01
EMail: Jean-Jacques.Bernard@nic.fr