TOC

Behavior Engineering for Hindrance	R. Denis-Courmont
Avoidance	Nokia
Internet-Draft	March 09, 2009
Intended status: Experimental	
Expires: September 10, 2009	

IPv6 destination header option for IPv4 translator mapping notification
draft-denis-behave-v4v6exthdr-01

Status of this Memo

Copyright Notice

Abstract

This memo defines a new IPv6 Destination header option to convey the transport mapping information from an IPv4-IPv4 protocol translator to the IPv6 end of a protocol-translated packet flow.

Table of Contents

1. Introduction
2. Definitions
3. IPv4-IPv6 Translation
    3.1. Inserting the flow mapping option
        3.1.1. Usage with connection-oriented protocols
        3.1.2. Usage with other protocols
    3.2. Receiving the flow mapping option
4. Option format
5. UNSAF Considerations
    5.1. Exit strategy
    5.2. Interactions with legacy NATs
6. Security Considerations
7. IANA Considerations
8. API Considerations
9. References
    9.1. Normative References
    9.2. Informative References
§ Author's Address

1. Introduction

TOC

To overcome the shortage of IPv4 addresses within the Internet, Network Address and Port Translators (NATs) have been widely deployed, such that multiple IPv4 nodes can share a single IPv4 address. However, that method is known to break certain application protocols, which need to know their own assigned external IP address and/or port number (i.e. the transport address). New solutions are now under consideration which would extend NAT mechanisms such that IPv6 nodes could access the IPv4 Internet.

This memo proposes an in-band method for such a IPv6-IPv4 NAT to notify affected IPv6 applications of the IPv4 transport address associated with any of their active communication flows. A new option for the IPv6 Destination extension header, the Translated Flow Mapping option is hereby defined to carry this information.

2. Definitions

TOC

TBD.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] (Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” March 1997.).

3. IPv4-IPv6 Translation

TOC

An IPv4-IPv6 NAT performs two separate functions:

- It receives IPv4 packets on its IPv4 interface, translates them to IPv6. To that end, for each IPv4 packet, it crafts a new IPv6 header to replace the IPv4 header, may modify the inner transport protocol header. Then, it sends the resulting translated IPv6 packets through its IPv6 interface.
- Reciprocally, it translates IPv6 packets into IPv4 packets.

The details of IPv4-IPv6 translation are beyond the scope of this document, please refer to [whatever IETF ends up specifying for this] instead.

3.1. Inserting the flow mapping option

TOC

When a translator receives an IPv4 packet, following certain conditions, it inserts an IPv6 Destination extension header containing a Translated Flow Mapping option (as defined in the next section). As a general rule, this option MUST NOT be inserted, if the resulting packet would exceed the known MTU to the IPv6 destination, or 1280 bytes if there is no known MTU.

3.1.1. Usage with connection-oriented protocols

TOC

For connection-oriented transport protocols, this option SHOULD be inserted is part of the protocol handshake, and SHOULD NOT be inserted otherwise.

3.1.1.1. Datagram Congestion Control Protocol (DCCP)

TOC

This option SHOULD be inserted within DCCP Sync, DCCP Sync/Ack and DCCP Listen packets. See [RFC4340] (Kohler, E., Handley, M., and S. Floyd, “Datagram Congestion Control Protocol (DCCP),” March 2006.) and

[I-D.ietf-dccp-simul-open] (Fairhurst, G., “DCCP Simultaneous-Open Technique to Facilitate NAT/Middlebox Traversal,” May 2009.).

3.1.1.2. Stream Control Transmission Protocol (SCTP)

TOC

TBD.

3.1.1.3. Transmission Control Protocol (TCP)

TOC

This option SHOULD be inserted within TCP SYN and TCP SYN/ACK packets. See [RFC0793] (Postel, J., “Transmission Control Protocol,” September 1981.).

3.1.2. Usage with other protocols

TOC

So long as a translated packet is small enough (with regards to the MTU rule above), and uses a non-connection-oriented (including UDP and UDP-Lite) or unknown transport protocol, the translator MAY insert the option. If it is known that the packet is one of the first 10 (FIXME: is this OK?) packets translated in the same direction for the corresponding mapping, then the translator SHOULD insert the option.

3.2. Receiving the flow mapping option

TOC

Processing of the flow mapping option is optional. In fact, an IPv6 implementation that does not support the flow mapping option MUST ignore it, according to [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.) (this is not a new requirement for IPv6 implementation).
The content of the flow mapping option is merely informational. Hence, there are no particular requirements as regards its processing. An IPv6 stack that implements the flow mapping option MAY store and or forward the flow mapping informations, as it sees fit. For instance, it might forward the informations to the application (see below for an example API) if it requests them.

TOC

4. Option format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Option Type  | Option Length |          Mapped Port          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Mapped IPv4 Address                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Remote IPv4 Address                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Translated Flow Mapping option

The Translated Flow Mapping option format is defined as follow:

> **Option Type:** XXX (TBD: IANA)
>
> **Option Length:** 10 (10 bytes worth of data)
>
> **Mapped Port:** If the type of the first header that is not an IPv6 extension header is DCCP, SCTP, TCP, UDP or UDP-Lite, the transport protocol mapped port number. This is the destination port number found in the original IPv4 packet that was translated into the IPv6 packet containing this option. Otherwise, this must be set to zero by sender, and ignored by receivers.
>
> **Mapped IPv4 Address:** Destination IPv4 address, as found in the origin IPv4 packet before translation.
>
> **Remote IPv4 Address:** Source IPv4 address, as found in the origin IPv4 packet before translation.

The Translated Flow Mapping option requires a 4n alignment (as defined per [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.) section 4.2). In particular, if it is the only non-padding option in an IPv6 extension header, it will be preceded by two bytes of padding. That is normally achieved through a single PadN option with a zero-length payload.

TOC

5. UNSAF Considerations

The Translated Flow Mapping option can be inserted by translators and received by IPv6 nodes.

5.1. Exit strategy

TOC

It is expected that any applicable translation mechanism will define its own UNSAF Considerations, at least as regards the translators. Those should be referred to when it comes to inserting the Flow Mapping option. In particular, such a specification shall narrow down the scope of the translation scheme, define an exit strategy and longer term solutions (e.g. complete translation-free native IPv6 networking). See [RFC3424] (Daigle, L. and IAB, "IAB Considerations for UNilateral Self-Address Fixing (UNSAF) Across Network Address Translation," November 2002.) for further references.

However, a dedicated exit strategy is required for the IPv6 nodes that would be capable of parsing the Translated Flow Mapping option.

When applicable translator deployments are being phased out, parsing the option becomes increasingly irrelevant, as the option will be absent from any received packets. At that point, IPv6 implementations can stop recognizing and parsing the option. They can instead return an error to any IPv6 application that would still try to use of the Flow Mapping option. IPv6 applications MUST be prepared to deal with IPv6 implementations that do not support this specification.

5.2. Interactions with legacy NATs

TOC

Legacy NATs do not support this option. This situation can normally be detected by the absence of the Translated Flow Mapping option.

Problems may occur if a translator that implements this specification is located behind a legacy NAT. In this case, the Translated Flow Mapping option may contain incorrect informations. This can most often be detected by verifying that the embedded IPv4 address is a globally unique one rather than a private one (as defined by [RFC1918] (Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets," February 1996.) and [RFC3927] (Cheshire, S., Aboba, B., and E. Guttman, "Dynamic Configuration of IPv4 Link-Local Addresses," May 2005.)).

However, any application using this extension SHOULD be prepared to fail gracefully if incorrect informations are received. Indeed, a legacy NAT could internally use public address space. Or the (non-legacy) translator could be deployed in a closed network using private IPv4 addresses, even in the absence of legacy NATs.

6. Security Considerations

TOC

By maliciously inserting or altering a Translated Flow Mapping option to an IPv6 packet, an attacker could cause manipulate IP and transport addressing informations to be received.
This may specifically allow an IPv6 attacker to refer the victim recipient node to an arbitrary IPv4 third party. As usual, IP nodes should not make assumptions to lightly as regard the IP address information they get. This problem is very similar to that of an IPv6 node handling a source-spoofed IPv6 packet, and the same precautions applies. In particular, proper transport or application-layer congestion control mechanisms need to be used, to prevent a distributed denial-of-service attack. Also, in security-sensitive cases, adequate security protocols are needed, such as TLS or IPsec.
The Translated Flow Mapping option can also cause a victim recipient to assume an incorrect arbitrary IPv4 self-referral address. TBD: Do we need to fix this? How?

7. IANA Considerations

TOC

The Translated Flow Mapping option requires an IPv6 Option number.
IPv6 Option Number [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.):

HEX	act	chg	rest	
---	---	---	-----	
XX	00	0	XXXXX	Translated Flow Mapping

The first two bits indicate that the IPv6 node may skip over this option and continue processing the header if it doesn't recognize the option type, and the third bit indicates that the Option Data may not change en-route.
This document should be listed as the reference document.

8. API Considerations

TOC

This section is non-normative. It defines a potential API to retrieve the flow mapping information as an extension to the Advanced IPv6 socket API [RFC3542] (Stevens, W., Thomas, M., Nordmark, E., and T.

Jinmei, “Advanced Sockets Application Program Interface (API) for IPv6,” May 2003.).
The flow mapping informations shall be passed to applications using a structure defined in <netinet/in.h>, and containing at least the following fields:

```
     struct in6_ipv4flowmapping {
       struct uint16_t i4fm6_mapped_port;
       struct in_addr  i4fm6_mapped_addr;
       struct in_addr  i4fm6_remote_addr;
     };
```

Flow mapping structure

For datagram (type SOCK_DGRAM) and raw (type SOCK_RAW) sockets, a socket option can configure receiving the flow information as ancilliary data on a per-packet basis, using recvmsg. This socket option shall be set to 0 (off) by default. Setting it to 1 (on) shall enabled flow mapping infos reception. Setting it to -1 (default) shall disable it. When enabled, an ancilliary data with level IPPROTO_IPV6, type IPV6_IPV4FLOWMAPPING shall be returned to the application, if a Flow Mapping option was found in the received packet.

```
     int on = 1;

     setsockopt(fd, IPPROTO_IPV6, IPV6_RECVIPV4FLOWMAPPING,
                &yes, sizeof(yes));
```

Per-packet socket option

For a connected socket, a read-only socket option may be used to fetch the flow mapping information if known (i.e. if at least one packet with a Flow Mapping Option was received). If unknown, the returned structure shall contain all zeroes.

```
     struct in6_ipv4flowmapping val;

     getsockopt(fd, IPPROTO_IPV6, IPV6_IPV4FLOWMAPPING,
                &val, sizeof(val));
```

Connected socket option

9. References

TOC

9.1. Normative References

TOC

[I-D.ietf-dccp-simul-open]	Fairhurst, G., “DCCP Simultaneous-Open Technique to Facilitate NAT/Middlebox Traversal,” draft-ietf-dccp-simul-open-08 (work in progress), May 2009 (TXT).
[RFC0793]	Postel, J., “Transmission Control Protocol,” STD 7, RFC 793, September 1981 (TXT).
[RFC1918]	Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, “Address Allocation for Private Internets,” BCP 5, RFC 1918, February 1996 (TXT).
[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC2460]	Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” RFC 2460, December 1998 (TXT, HTML, XML).
[RFC3424]	Daigle, L. and IAB, “IAB Considerations for UNilateral Self-Address Fixing (UNSAF) Across Network Address Translation,” RFC 3424, November 2002 (TXT).
[RFC3927]	Cheshire, S., Aboba, B., and E. Guttman, “Dynamic Configuration of IPv4 Link-Local Addresses,” RFC 3927, May 2005 (TXT).
[RFC4340]	Kohler, E., Handley, M., and S. Floyd, “Datagram Congestion Control Protocol (DCCP),” RFC 4340, March 2006 (TXT).

9.2. Informative References

TOC

[RFC3542]	Stevens, W., Thomas, M., Nordmark, E., and T. Jinmei, “Advanced Sockets Application Program Interface (API) for IPv6,” RFC 3542, May 2003 (TXT).

Author's Address

TOC

Rémi Denis-Courmont
Nokia Corporation
P.O. Box 407
NOKIA GROUP 00045
FI
Phone: +358 50 487 6315
Email: remi.denis-courmont@nokia.com