Workgroup: Network Working Group
Internet-Draft:
draft-peng-6man-tracing-option-01
Published: 30 June 2022
Intended Status: Standards Track
Expires: 1 January 2023
Authors: Y. Yin
China Telecom
S. Peng
Huawei Technologies
R. Zhao
Huawei Technologies

Tracing process in IPv6 VPN Tunneling Networks

Abstract

This document specifies the tracing process in IPv6 VPN tunneling networks for diagnostic purposes. An IPv6 Tracing Option is specified to collect and carry the required key information in an effective manner to correctly construct ICMPv4 and ICMPv6 Time Exceeded messages at the corresponding nodes, i.e. PE and P nodes, respectively.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 1 January 2023.

document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction

ICMPv6 (Internet Control Message Protocol) RFC 4443 [RFC4443] is used by IPv6 nodes to report errors encountered in processing packets and to perform other internet-layer functions, such as diagnostics (ICMPv6 "ping"). RFC 4443 [RFC4443] describes the format of a set of control messages used in ICMPv6, including the Time Exceeded Message. Every ICMPv6 message is preceded by an IPv6 header and zero or more IPv6 extension headers. The ICMPv6 header is identified by a Next Header value of 58 in the immediately preceding header.

If a router receives a packet with a Hop Limit of zero, or if a router decrements a packet's Hop Limit to zero, it MUST discard the packet and originate an ICMPv6 Time Exceeded message with Code 0 to the source of the packet.

In the case of VPN, as shown in Figure 1, where CE1 and CE2 are IPv4, an IPv6 tunnel exists between PE1 and PE2, and all the nodes belong to a single network operator. For diagnostic purposes, CE1 sends out an IPv4 packet with its TTL set to a value. The IPv4 packet is encapsulated within the IPv6 tunnel at PE1. The TTL of the IPv4 packet will be copied, based on which a new value will be set as the Hop Limit in the outer IPv6 tunnel header. The new Hop Limit value depends on the mode configured on PE1, i.e., Uniform mode or Pipe mode RFC 3443 [RFC3443]. If it is the Uniform mode, the Hop Limit will be the TTL value in the received packet minus one. When an intermediate router P decrements the Hop Limit in the outer

tunnel header to zero, an ICMPv6 Time Exceeded Message needs to be sent back to the source, which should be the CE1 via PE1.

The Pipe mode can be used to detect the routing loop. If it is the Pipe mode configured on PE1, the Hop Limit will be set to be the maximum value (e.g., 64). In this case, when an intermediate router P decrements the Hop Limit in the outer tunnel header to zero, it means that the routing loop has happened, and this packet needs to be dropped.

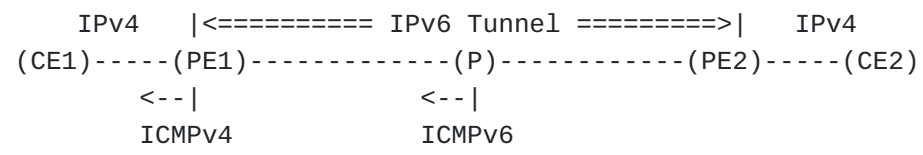


Figure 1. The tracing in IPv6 VPN tunneling networks

In order to construct a correct ICMPv4 Time Exceeded Message at PE1 and send it to CE1, a couple of key information is required:

1) The IPv4 address of the access interface at the P node, which will be taken as the source address of the ICMPv4 Time Exceeded Message.

2) The VPN information, which is used to identify the VPN, either using the VPN ID or the Access Interface ID at the PE1.

However, currently this information is missing and there is still no appropriate way to collect and carry it to the right nodes.

This document specifies the tracing process in IPv6 VPN tunneling networks. An IPv6 Tracing Option is specified to collect and carry the required key information in an effective manner to correctly construct ICMPv4 and ICMPv6 Time Exceeded messages at the corresponding nodes, i.e. CE and P nodes, respectively.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 RFC 2119 [RFC2119] RFC 8174 [RFC8174] when, and only when, they appear in all capitals, as shown here.

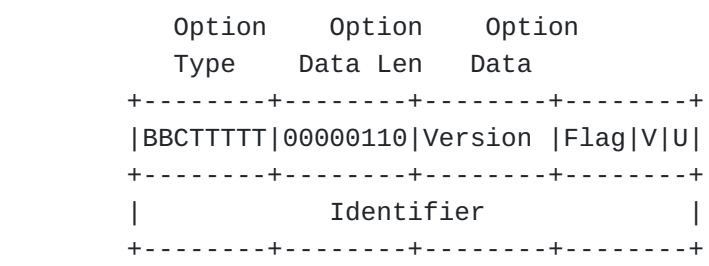
3. Terminologies

TTL: Time To Live

ICMP: Internet Control Message Protocol

4. IPv6 Tracing Option

The tracing option has the following format.



Option Type (see Section 4.2 of [RFC8200]):

BB 00 Skip over this option and continue processing.

C 0 Option data can not change en route to the packet's final destination.

TTTTT TBD Option Type to be assigned from IANA.

Length 6 8-bit unsigned integer indicates the length of the option Data field of this option, in octets. The value of Opt Data Len of the IPv6 Tracing option SHOULD be set to 6.

Version n 8 bits. It indicates the version of this mechanism.

Flag n 8 bits, where:

U n 1 bit. U-Flag. If set by the ingress PE it indicates that the Uniform mode is configured on the ingress PE. Otherwise, the ingress PE is on the pipe mode.

V n 1 bit. V-Flag. If set by the ingress PE it indicates that the carried following Identifier is a VPNID. Otherwise, it is the Access Interface ID.

Identifier n 4 octets. It is used to identify the VPN, either using the VPN ID or the Access Interface ID, as indicated by the V flag.

5. Tracing Process in different modes of the ingress PE

The diagnostic IPv4 packet sent by CE is encapsulated within the IPv6 tunnel at the ingress PE. The TTL of the IPv4 packet is copied, based on which a new value is set as the Hop Limit in the outer IPv6 tunnel header.

The ingress PE can be configured in two modes, that is, Uniform mode and Pipe mode. The new Hop Limit value depends on the mode configured on PE1. If it is the Uniform mode, the Hop Limit will be the TTL value in the received packet minus one. If it is the Pipe mode, the Hop Limit will be set to be the maximum value (e.g., 255). More details are described below.

5.1. Tracing Process in Uniform mode

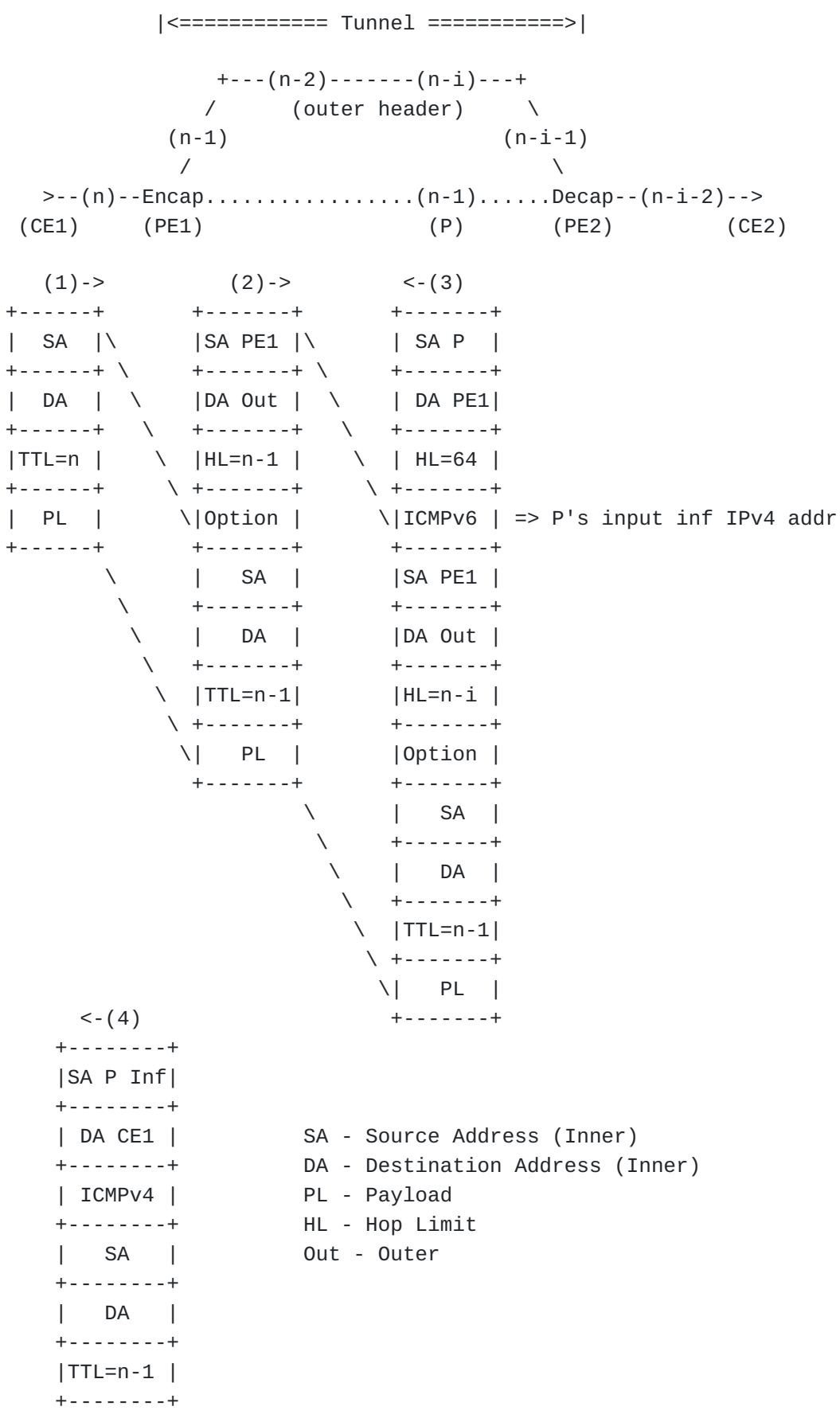
When the ingress PE is configured in Uniform Mode, the inner and outer TTLs of the packets are synchronized at tunnel ingress (PE1) and egress (PE2).

Figure 2 shows the tracing process in the Uniform Mode. When an IP packet (shown as (1) in the figure and with TTL = n) reaches the ingress PE (PE1), it is encapsulated by the ingress PE into a newly created IPv6 header and an extension header (Hop-by-Hop Options Header or Destination Options Header RFC 8200 [RFC8200]) carrying the IPv6 Tracing Option defined in this document. The Hop Limit is set to be n - 1, shown as (2) in the figure.

When the Hop Limit becomes zero, the P node will check whether the IPv6 Tracing Option is carried. If carried, the information in the IPv6 Tracing Option will trigger the following actions.

If the U-flag is set, it means that the ingress PE is in the Uniform Mode, so an ICMPv6 packet (shown as (3) in the figure) will be sent back to the PE1. The SA of the packet is the IPv6 address of the P node, while the DA is the IPv6 address of the PE1. The ICMPv6 Error Message carries the IPv4 address of the input port interface of the packet entering the P node, which will be taken as the source address of the ICMPv4 message to be sent by the ingress PE towards CE1.

When the packet (3) is received by PE1, the PE1 will construct an ICMPv4 packet (4) and send it to CE1. At the PE1, the information in the carried IPv6 Tracing Option will be read and the VPN using which to continue to forwarding the packet to the corresponding CE will be identified using the V-Flag and the value of the Identifier in the IPv6 Tracing Option.


|<========== Tunnel ==========>|
+---(n-2)-------(n-i)---+
/ (outer header) \
(n-1) (n-i-1)
/ \
>--(n)--Encap.....(n-1).....Decap--(n-i-2)-->
(CE1) (PE1) (P) (PE2) (CE2)
(1)-> (2)-> <-(3)
SA
DA
TTL=n
PL
SA PE1
DA Out
HL=n-1
Option
SA P
DA PE1
HL=64
ICMPv6
=> P's input inf IPv4 addr
SA
DA
TTL=n-1
PL
SA PE1
DA Out
HL=n-i
Option
SA
DA
TTL=n-1
PL
<-(4)
SA P Inf
DA CE1
ICMPv4
SA
DA
TTL=n-1
SA - Source Address (Inner)
DA - Destination Address (Inner)
PL - Payload
HL - Hop Limit
Out - Outer

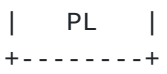


Figure 2. The tracing process in the Uniform Mode

5.2. Tracing Process in Pipe mode

When the ingress PE is configured in Pipe Mode, the inner and outer TTLs of the packets will not be synchronized at tunnel ingress (PE1) and egress (PE2). The tunnel will be taken as one hop by the inner packet, as shown in Figure 3.

The Hop Limit will be set to be the maximum value (e.g., 64) at the ingress PE. Since it is set to the maximum value, in normal case, the Hop Limit will not become zero at any P node. So the only reason when the Hop Limit becomes zero is that a routing loop is detected. In this case, the packet needs to be dropped.

If the U-flag is not set, it means that the ingress PE is in the Pipe Mode, and the packet (i.e. (2) as shown in Figure 2) will be dropped when the Hop Limit becomes zero either at the P node (no ICMPv6 packet (i.e. (3) as shown in Figure 2)) or the PE1 node when the P node does not have the dropping capability.

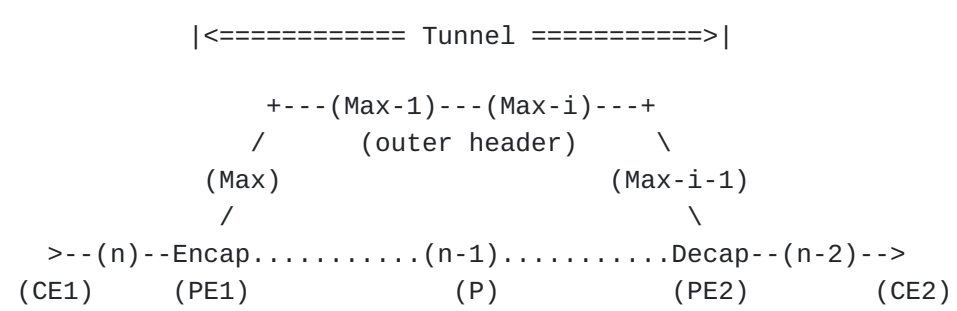


Figure 3. The tracing process in the Pipe Mode

6. IANA Considerations

IANA is requested to allocate one new option type from "Destination Options and Hop-by-Hop Options" registry.

Value	Name	Reference
TBD1	IPv6 Tracing Option	This ID

7. Acknowledgements

The authors would like to thank the careful reviews and valuable comments from Stefano Previdi.

8. Security Considerations

TBD

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3443] Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks", RFC 3443, DOI 10.17487/RFC3443, January 2003, <https://www.rfc-editor.org/info/rfc3443>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

9.2. Informative References

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

Authors' Addresses

Yuanyang Yin
China Telecom
Guangzhou
China

Email: yinyuany@chinatelecom.cn

Shuping Peng
Huawei Technologies
Beijing
China

Email: pengshuping@huawei.com

Ranxiao Zhao
Huawei Technologies
Beijing
China

Email: zhaoranxiao@huawei.com