Tracing process in IPv6 VPN Tunneling Networks

draft-peng-6man-tracing-option-03

Shuping Peng, Ranxiao Zhao

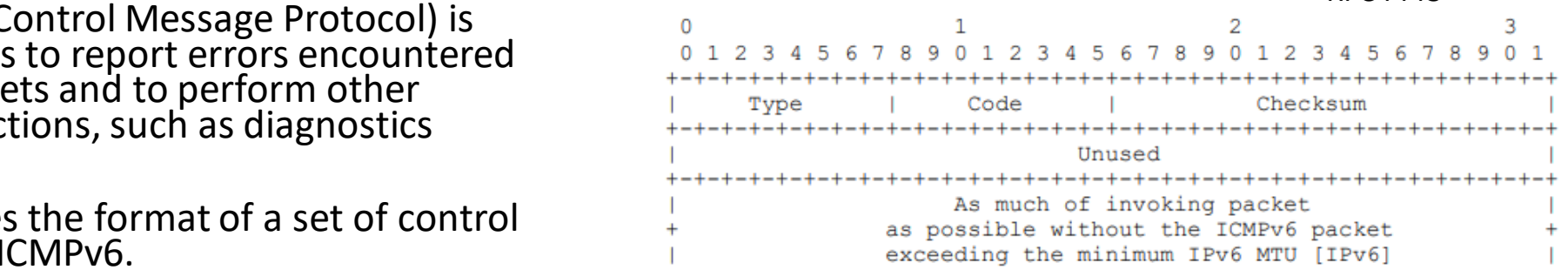
Huawei Technologies

Background

- ICMPv6 (Internet Control Message Protocol) 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 describes the format of a set of control messages used in ICMPv6.
- 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.

RFC4443

3.3. Time Exceeded Message



IPv6 Fields:

Destination Address
Copied from the Source Address field of the invoking packet.

ICMPv6 Fields:

Type 3

Code
0 - Hop limit exceeded in transit
1 - Fragment reassembly time exceeded

Unused
This field is unused for all code values. It must be initialized to zero by the originator and ignored by the receiver.

Tracing in IPv6 VPN Tunneling Networks

- In the case of VPN, an example as shown in Figure 1, where CE1 and CE2 are IPv4 (could also v6), 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 [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.
- If it is the **Pipe mode** configured on PE1, the Hop Limit will be set to be the maximum value (e.g., 255). 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.
  - The router P only sees HL = 0, so it needs a mechanism to determine whether HL = 0 is caused by a loop or a normal traceroute.

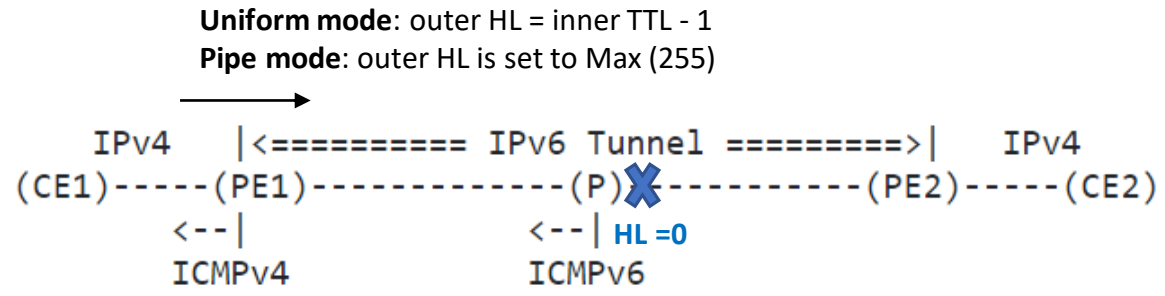


Figure 1. The tracing in IPv6 VPN tunneling networks

Necessary information and mechanism are needed

- In order to construct a correct ICMPv4/v6 Time Exceeded Message at PE1 and send it to CE1, the following key information is required:

  1) **The IPv4/v6 address of the access interface at the P node**, which will be taken as the source address of the ICMPv4/v6 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 an appropriate way is desired to collect and carry it to the right nodes.
- A **mechanism for** the router P to determine whether HL=0 is caused by a loop or a normal traceroute is also needed.

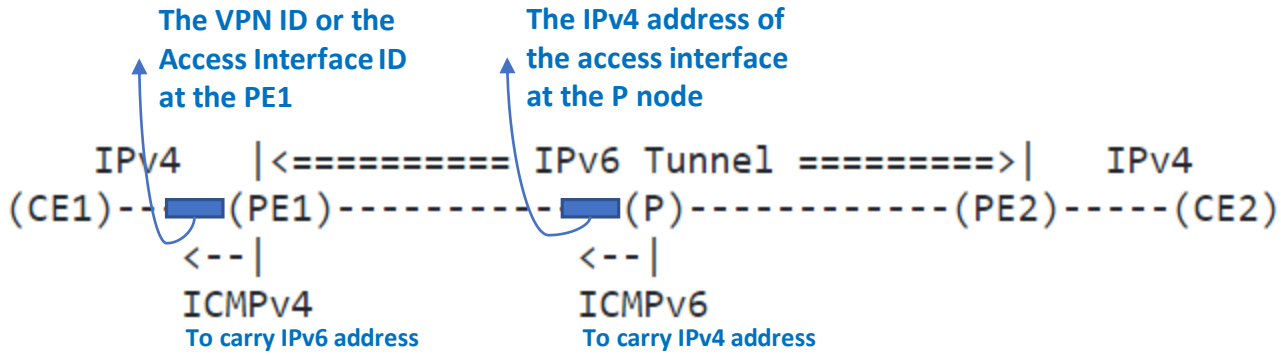


Figure 1. The tracing in IPv6 VPN tunneling networks

Summary of the mailing list discussions

- VPN information could be extracted from the original invoking packets
  - The VPN information cannot be obtained from some tunnel types like SRv6
- Source PE has low capability to extract this information from the ICMPv6 message
  - The control plane is used to process the trace packets
- New ICMPv6 and ICMP (v4) message types need to be defined to carry
  - IPv4 address of the P router in the ICMPv6 generated by the P router
  - IPv6 address of the P router in the ICMP(v4) generated by the source PE
- Corresponding clarifications have been updated in the draft

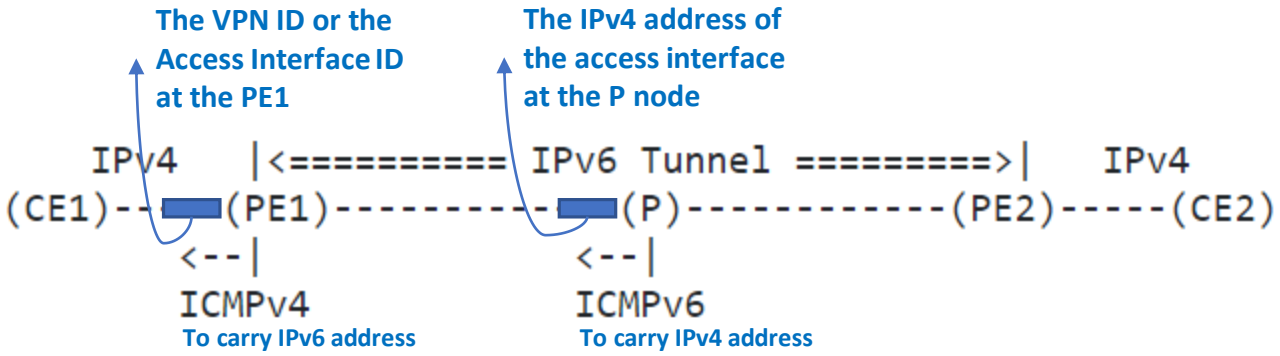


Figure 1. The tracing in IPv6 VPN tunneling networks

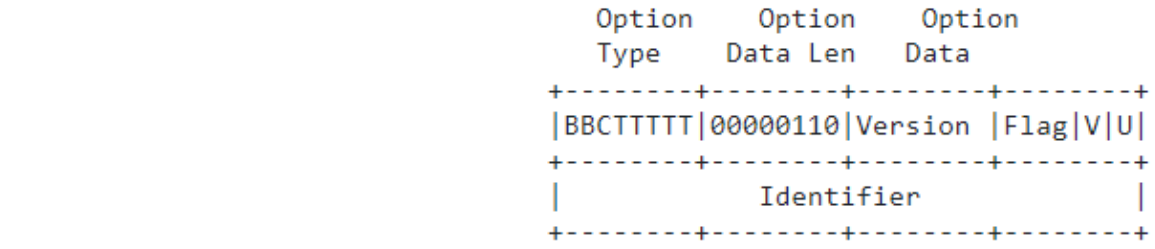
A proposal

- draft-peng-6man-tracing-option-03 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/v6 and ICMPv6 Time Exceeded messages at the corresponding nodes, i.e. CE and P nodes, respectively.

- New ICMPv6 and ICMP (v4) message types will be defined in future versions

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.

Your comments and suggestions are appreciated. Thank you!