6MAN
Internet-Draft
Intended status: Informational
Expires: 16 April 2026

Y. Liu
ZTE
Y. Liu
China Mobile

13 October 2025

Problem Statement with Aggregate Header Limit for IPv6
draft-liu-6man-aggregate-header-limit-problem-04

Abstract

This document first proposes the concept of "Aggregate header limit for IPv6"(IPv6-AHL) to indicate the total header size that a router is able to process at full forwarding rate for IPv6 packets. Then this document describes the problems for path calculation and function enablement without the awareness of IPv6-AHL, and the considerations for IPv6-AHL collection are also included.

Table of Contents

1. Introduction

The introduction of IPv6 extension headers has been increasing the total packet header chain size greatly, which may cause inefficient packet processing due to the header chain size exceeding the processing limit of the router.

Some hardware devices implement a parsing buffer of a fixed size to process packets. The parsing buffer is expected to contain all the headers that a device needs to examine. If the aggregate length of headers in a packet exceeds the size of the parsing buffer, a device will either discard the packet or defer processing to a software slow path. [RFC9098] also mentions that due to packet-forwarding engine constraints, if an IPv6 header chain is sufficiently long such that it exceeds the packet lookup capacity of the router, the router might be unable to determine how the packet should be handled and thus could resort to dropping the packet. And some packet-forwarding engines manage IPv6 header chains using recirculation, but recirculation can impact the forwarding capacity of hardware, as each packet will pass through the processing engine multiple times.

Before discussing the problem, this document proposes the concept of Aggregate Header Limit for IPv6 as below:

"Aggregate Header Limit for IPv6 (IPv6-AHL) is the total header size(i.e., from the IPv6 header chain up to any headers that are part of network encapsulation, e.g., the innermost transport layer, ethernet frame, or another IP header) that a router is able to process at full forwarding rate(e.g., at fast path). For some devices designed with parsing buffer, this limit is related with its buffer size and buffer design."

The different between IPv6-AHL and the existing concepts:

* Difference with Aggregate Header Limit(AHL) introduced in [RFC8883]: There’s not a clear definition for AHL in [RFC8883]. If AHL is understood as the total header size that a router is able to process at full forwarding rate, the values of AHL and IPv6-AHL may be the same or may be difference based on router design.

* Difference with extension header chain size limit, the extension header chain size is no bigger than aggregate header size since the packet may contain upper layer headers.As mentioned in [RFC9098], the intermediate nodes/systems may need to process Layer 3/Layer 4 information to make a forwarding decision, in this case, even the extension header chain size limit is not exceeded, an intermediate node may drop the packet due to IPv6-AHL exceeding.

This document describes the problems for path calculation and function enablement without the awareness of IPv6-AHL, and the considerations for IPv6-AHL collection are also included.

2. Conventions used in this document

2.1. Terminology

MSD: Maximum SID Depth as in [RFC8491].

IPv6-AHL: Aggregate header limit for IPv6. It’s the total header size(i.e., from the IPv6 header chain as well as any headers that are part of network encapsulation up to the innermost transport layer) that a router is able to process at full forwarding rate(e.g., at fast path). For some devices designed with parsing buffer, this limit is related with its buffer size and buffer design.

The terminology defined in [RFC9673] are used in this document as below:

Forwarding Plane: IPv6 routers exchange user or applications data through the forwarding plane. Routers process fields contained in IPv6 packet headers. However, they do not process information contained in packet payloads.

Control Plane: Routers exchange management and routing information. They also exchange routing information with one another. Management and routing information are processed by its recipient. Management and control information can be forwarded by a router that process fields contained in packet headers.

Fast Path: A path through a router that is optimized for forwarding packets. The Fast Path might be supported by Application Specific Integrated Circuits (ASICS), Network Processor (NP), or other special purpose hardware. This is the usual processing path within a router taken by the forwarding plane.

Slow Path: A path through a router that is capable of general purpose processing and is not optimized for any particular function. This processing path is used for packets that require special processing or differ from assumptions made in Fast Path heuristics or to process router control protocols used by the control plane.

Full Forwarding Rate: This is the rate that a router can forward packets without adversely impacting the aggregate forwarding rate. For example, a router could process packets at a rate that allows it to maintain the full speed on its outgoing interfaces, which is sometimes called "wire speed".

2.2. 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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Problem Statement

The introduction of IPv6 extension headers has been increasing the total packet header chain size greatly, which may cause inefficient packet processing due to the header chain size exceeding the processing limit of the router. And the possibility of the combination of IPv6 extension headers and different TLVs in the extension headers would make total header size even bigger.

Normally, there’re different models/versions of network devices(i.e., switches, routers) from multiple vendors in an operator’s network, different devices may have different aggregate header limits and different behaviors after aggregate header limit exceeding. As more and more IPv6 functions are superimposed in the operator’s network, packet dropping or rate limiting due to IPv6-AHL exceeding is a potential risk, which makes it difficult to manage the network.

Path calculation, whether by the controller or the headend, without the awareness of IPv6-AHL of the nodes in the network and the prediction on which features would be enabled along the path, may result in a path with nodes with lower AHLs than required. If the controller is aware of aforesaid information, the controller would be able to reserve space for the IPv6 extension headers and EH TLVs to be insered in the packet header to ensure that the packet header size wouldn’t exceed the IPv6-AHLs of the intermediate segment endpoints along the list.

The situation is similar for packet encapsulation triggered by function enablement, whether on the headend or the intermediate nodes, packets may be encapsulated with larger header size than the downstream nodes able to process. If IPv6-AHL information of the nodes/path can be obtained in advance, when the node needs to attach extra data along the existing path, and the IPv6-AHL of the downstream nodes along the path are not sufficient to process the headers, the node may choose not to use the related function and log an error.

4. IPv6-AHL Collection Considerations

4.1. PMTUD-style Collection Mechanism

As per [RFC8883], an ICMPv6 Destination Unreachable error with code for "Headers too long" SHOULD be sent when a node discards a packet because the aggregate length of the headers in the packet exceeds the processing limits of the node. Based on this definition, obtaining the minimum AHL along the path can be achieved by sending detection messages of a certain size and receiving the ICMPv6 error messages, which is similar with path MTU discovery for IPv6 in [RFC8201].

This mechanism may work for small networks with static paths in it. But there may be some problems in the following scenarios:

* When the number of paths increases, more and more detection messages need to be sent, and the burden of processing the received ICMPv6 error messages also increases.

* For SR dynamic candidate paths [RFC9256], the segment lists of the paths may change over time, which makes it more difficult to detect the IPv6-AHLs.

4.2. HBH-style Collection Mechanism

[RFC9268] leverages a Hop-by-Hop Option to collect the limit(i.e., minimum path MTU) along the path. A Hop-by-Hop Option for IPv6-AHL collection purpose might be another option following the example of [RFC9268]. This mechanism may not work well for all the cases, e.g., there might be transit routers that just forward the SRv6 packets like normal IPv6 packets without inspecting into the extension header chain. Even if the AHLs of these nodes have smaller AHLs than other nodes, the packets would still be processed normally along the path. In this case, if a mechanism like [RFC9268] is used, the AHL of the path collected would be one of the AHLs of these transit nodes, but actually packets with larger header chain size could be sent and processed normally.

4.3. Signaling-style Collection Mechanism

Signaling could be another option. Considering that there’re already mechanisms like IGP-MSD [RFC8491][RFC8476] to advertise certain size limit on the per-node and per-link basis. The mechanism for advertising IPv6-AHL is similar to IGP-MSD. In the inter-domain scenario, the BGP signaling may help as well. For the controller, it can get the AHLs of the nodes in the network via BGP-LS, YANG or other south-north mechanisms. The details of signaling mechanism is out of the scope of the document and would be discussed in separate documents.

5. IANA Considerations

This document makes no request of IANA.

6. Security Considerations

If the IPv6-AHL-collection message (whether as the ICMPv6 error message, the Hop-by-Hop Option for IPv6-AHL or the routing protocol message) is sent to a third party, it allows network reconnaissance, a third party may speculate on the system design of the node based on the AHL information, such as the size of the buffer, and it may indirectly speculate the hardware configuration and system version of the node as well.

The IPv6-AHL-collection function SHOULD only be used within a trusted domain and proper filtering or authorization method should be considered to prevent the AHL-collection message from being sent to the untrusted third party.

The security considerations described above primarily apply to in-band signaling mechanisms (e.g., ICMPv6, Hop-by-Hop options, or routing protocols). For out-of-band signaling mechanisms, such as NETCONF or RESTCONF used in network management systems, the risk of reconnaissance is significantly reduced due to the use of encrypted and authenticated channels, as well as access control within trusted management networks.

7. Acknowledgement

The authors would like to thank Tom Herbert, Alvaro Retana, Eric Vyncke, Jeff Tantsura, Sasha Vainshtein and Acee Lindem for their helpful comments and suggestions.

8. References

8.1. 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>.

[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>.

[RFC8883] Herbert, T., "ICMPv6 Errors for Discarding Packets Due to Processing Limits", RFC 8883, DOI 10.17487/RFC8883, September 2020, <https://www.rfc-editor.org/info/rfc8883>.

[RFC9673] Hinden, R. and G. Fairhurst, "IPv6 Hop-by-Hop Options Processing Procedures", RFC 9673, DOI 10.17487/RFC9673, October 2024, <https://www.rfc-editor.org/info/rfc9673>.

8.2. Informative References

[RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6", STD 87, RFC 8201, DOI 10.17487/RFC8201, July 2017, <https://www.rfc-editor.org/info/rfc8201>.

[RFC8476] Tantsura, J., Chunduri, U., Aldrin, S., and P. Psenak, "Signaling Maximum SID Depth (MSD) Using OSPF", RFC 8476, DOI 10.17487/RFC8476, December 2018, <https://www.rfc-editor.org/info/rfc8476>.

[RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg, "Signaling Maximum SID Depth (MSD) Using IS-IS", RFC 8491, DOI 10.17487/RFC8491, November 2018, <https://www.rfc-editor.org/info/rfc8491>.

[RFC9098] Gont, F., Hilliard, N., Doering, G., Kumari, W., Huston, G., and W. Liu, "Operational Implications of IPv6 Packets with Extension Headers", RFC 9098, DOI 10.17487/RFC9098, September 2021, <https://www.rfc-editor.org/info/rfc9098>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

[RFC9268] Hinden, R. and G. Fairhurst, "IPv6 Minimum Path MTU Hop-by-Hop Option", RFC 9268, DOI 10.17487/RFC9268, August 2022, <https://www.rfc-editor.org/info/rfc9268>.

Authors' Addresses

Yao Liu
ZTE
China
Email: liu.yao71@zte.com.cn

Yisong Liu
China Mobile
China
Email: liuyisong@chinamobile.com