6MAN W. Kumari
Internet-Draft Google
Intended status: Best Current Practice J. Jaeggli
Expires: December 17, 2015 Zynga
R. Bonica
Juniper Networks
J. Linkova
Google
June 15, 2015

# Operational Issues Associated With Long IPv6 Header Chains
# draft-wkumari-long-headers-03

## Abstract

This memo specifies requirements for IPv6 forwarders as they process packets with long header chains. It also provides guidance for application developers whose applications might rely on long headers chains.

As background, this memo explains how many ASIC-based IPv6 forwarders process packets and why processing of packets with long header chains might be problematic.

## Requirements Language

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 RFC 2119 [RFC2119].

## Status of This Memo

This Internet-Draft will expire on December 17, 2015.

Copyright Notice

Table of Contents

1. Introduction . . . . . . . . . . . . . . . . . . . . . . . . 2
1.1. Termnology . . . . . . . . . . . . . . . . . . . . . . . 4
2. Forwarder Information Requirements . . . . . . . . . . . . . 4
3. Requirements For IPv6 Forwarders . . . . . . . . . . . . . . 5
4. Recommendations For Application Developers . . . . . . . . . 7
5. IANA Considerations . . . . . . . . . . . . . . . . . . . . . 7
6. Security Considerations . . . . . . . . . . . . . . . . . . . 7
7. Acknowledgements . . . . . . . . . . . . . . . . . . . . . . 8
8. References . . . . . . . . . . . . . . . . . . . . . . . . . 8
8.1. Normative References . . . . . . . . . . . . . . . . . . 8
8.2. Informative References . . . . . . . . . . . . . . . . . 8
Appendix A. Changes / Author Notes. . . . . . . . . . . . . . . . 8
Authors' Addresses . . . . . . . . . . . . . . . . . . . . . . . 9

## 1. Introduction

IPv6 [RFC2460] forwarders can acquire information from the following sources:

- The IPv6 header

- One or more IPv6 extension headers

- An upper-layer header

Section 2 of this document explains how IPv6 forwarders use information from the IPv6 header and IPv6 extension headers to provide traditional forwarding services. It also explains how IPv6 forwarders use information from the upper-layer header to provide enhanced forwarding services.

When a software-based forwarder processes an IPv6 datagram, it parses the header chain, regardless of its length, acquires the required information and makes a forwarding decision. Typically, software-based forwarders process a relatively small number of packets per second. Therefore, they can perform the above mentioned procedure within the constraints of their processing budget.

By contrast, ASIC-based forwarders process many more packets per second. In order to fulfill this requirement, ASIC-based forwarders copy a fixed number of bytes from the beginning of the packet to on-chip memory. Forwarders do this because they can access on-chip memory much more quickly than they can access off-chip memory. Once the beginning of the packet has been transferred to on-chip memory, subsequent processing can proceed very quickly.

The act of copying bytes from the beginning of a packet to on-chip memory consumes:

- Processor cycles

- On-chip memory

- Wall-time

Therefore, the number of bytes copied to on-chip memory must be chosen wisely. If a forwarder copies more bytes than it needs, it wastes resources and adversely impacts performance. If it copies too few bytes, it may not have sufficient information to make a correct forwarding decision.

The IPv6 header chain is a variable-length data structure, whose size can exceed 64 kilobytes. However, packets with header chains exceeding 256 bytes are rarely observed on the Internet. Therefore, most ASIC-based forwarders copy a relatively small number of bytes from the beginning of a packet into on-chip memory. While this small number varies from platform to platform, it is generally much closer to 256 bytes than it is to 64 kilobytes.

IPv6 forwarders MUST behave in a predictable manner when they process a packet whose header chain length exceeds the number of bytes copied to on-chip memory. Section 3 of this memo defines required behaviors.

Application developers should be aware of how ASIC-based forwarders process packets with long extension header chains. Therefore, Section 4 of this document provides guidance to application developers.

### 1.1. Termnology

For the purposes of this document, the terms "header chain" and "upper-layer" header are used as defined in [RFC7112].

This document also introduces the following terms:

- forwarding service - a service that accepts a packet from one interface and forwards it through another

- traditional forwarding service - a forwarding service in which all parameters to the forwarding algorithm are drawn from the IPv6 header, the hop-by-hop extension header, and the routing extension header

- enhanced forwarding service - a forwarding service in which parameters to the forwarding algorithm can be drawn from any portion of the IPv6 header chain

## 2. Forwarder Information Requirements

When an IPv6 forwarder provides traditional forwarding services, it extracts all information required by the forwarding algorithm from the IPv6 header, the hop-by-hop extension header (if present), and the routing extension header (if present). In the nominal case, the IPv6 header contains all information required by the forwarding algorithm. However, the hop-by-hop and routing extension headers can also impact forwarding behavior.

Section 4.2 of [RFC2460] explains how the hop-by-hop extension header impacts forwarding behavior. When the forwarder processes a hop-by-hop extension header, it examines each option contained by the header. If forwarder encounters an unrecognized hop-by-hop option, and the high-order bits of the option type are "00", the forwarder skips over the option and continues to process subsequent options. However, if an forwarder encounters an unrecognized option, and the high-order bits of the option type are "01", "10" or "11", the forwarder discards the packet.

Section 4.4 of [RFC2460] explains how the routing extension header impacts forwarding behavior. When the forwarder processes a packet whose destination address is local to itself, it scans the header chain, searching for a routing extension header. If the packet contains a routing extension header and the forwarder recognizes the routing header type, it processes the header. If the forwarder does not recognize the routing header type, the required behavior depends upon the Segments Left field. If the Segments Left field is equal to zero, the forwarder ignores the routing extension header. Otherwise,

the forwarder discards the packet. [RFC6275] and [RFC6554] describe currently defined routing extension header types.

Some IPv6 forwarders provide enhanced forwarding services, such as firewall filtering, rate limiting and load balancing. In order to provide these services, the forwarder requires access to an upper layer header. The following are examples of enhanced services that require the forwarder to examine the upper layer header:

- Discard all packets directed to TCP port 25
- Rate limit packets destined for a particular address whose payload is TCP and have the TCP SYN bit set
- Load balance packets across parallel links so that all packet belonging to particular TCP session traverse the same link.

## 3. Requirements For IPv6 Forwarders

The following requirements apply to all IPv6 forwarders:

- REQ-1: By default an IPv6 forwarder SHOULD NOT discard a valid packet because of its header chain length. However, the forwarder MAY support a configuration option that causes it to discard packets whose header chain length exceeds a specified value.
- REQ-2: When processing packet that contains a hop-by-hop extension header, an IPv6 forwarder MUST process the entire hop-by-hop extension header, regardless of its length. The forwarder MUST process each option as specified in Section 4.2 of [RFC2460]. If an IPv6 forwarder is not able to process the entire hop-by-hop extension header, it MUST discard the packet and SHOULD originate an ICMPv6 Parameter Problem message to the packet's source. The forwarder MAY have a configurable policy for sending ICMPv6 messages such as rate limiting or completely disabling them If an IPv6 forwarder is not able to process the entire hop-by-hop extension header, it MUST discard the packet and SHOULD originate an ICMPv6 Parameter Problem message to the packet's source. The forwarder MAY have a configurable policy for sending ICMPv6 messages such as rate limiting or completely disabling them.
- REQ-3: When processing a packet whose destination address is local to itself, an IPv6 forwarder MUST scan the entire header chain, regardless of its length, in order to determine whether the packet contains a routing extension header. If the packet contains a routing extension header, the forwarder MUST process routing extension header as specified in Section 4.4 of [RFC2460]. If an IPv6 forwarder is not able to process the entire routing extension

header, it MUST discard the packet and SHOULD originate an ICMPv6 Parameter Problem message to the packet's source. The forwarder MAY have a configurable policy for sending ICMPv6 messages such as rate limiting or completely disabling them.

The length of the IPv6 header plus the length of the hop-by-hop extension header can exceed the number of bytes that an ASIC-based forwarder copies into on-chip memory. Therefore, in order to support REQ-2, ASIC-based forwarders typically support a special processing mechanism for packets containing hop-by-hop extensions.

Also, the combined length of all headers preceding the routing extension header may exceed the number of bytes that an ASIC-based forwarder copies into on-chip memory. Therefore, in order to support REQ-3, ASIC-based forwarders typically support a special processing mechanism for packets whose IPv6 destination address is local to the forwarder. This forwarding mechanism is capable of processing the routing extension header, even if it begins beyond of the portion of the packet that was copied to on-chip memory.

The following requirements apply to IPv6 forwarders that provide enhanced forwarding services:

o REQ-4: If a forwarder's ability to deliver enhanced services is limited in any way by extension header length, that limitation MUST be reflected in user documentation. For example, assume that a forwarder provides a load balancing service, and that it acquires information required by the service from the IPv6 header and the upper-layer header. If the service behaves in one manner when all required information is contained by the first N bytes of the header chain and in another manner when all required information is not contained by the first N bytes of the header chain, user documentation MUST reflect both behaviors as well as the value of N.

o REQ-5: If a forwarder's ability to deliver an enhanced service is limited by extension header length, the policy specification language used to configure the enhanced service MUST be sufficiently robust to address the limitation. For example, assume that the forwarder provides a firewall service. The firewall service is capable of filtering packets directed to a particular TCP port, but only if the TCP header is contained by the first N bytes of the header chain. In this case, it MUST be possible to configure one policy for packets directed to the specified port, another policy for packet not directed to the specified port, and a third policy for packets whose TCP destination port is unknown.

## 4. Recommendations For Application Developers

Applications developers should be aware that many ISPs and enterprises filter or severely rate limit packets containing long header chains. They do this because of limitations imposed by the ASIC-based forwarders deployed at their edges. ISPs and enterprises accept these limitations as part of an engineering trade off, in which high-speed forwarding is achieved at the cost of limiting enhanced services for packets with long extension headers.

For example, assume that an enterprise deploys the following firewall filtering policy at its edge:

- Permit all packets whose destination is TCP port 80
- Discard all packets whose destination is not TCP port 80
- Discard all packets whose header chain is so long that TCP port information is not accessible to the filtering function

In this case, the enterprise discards all packets whose destination cannot be determined by the filtering function.

Aside from the issue of header chain length, operators may filter packets containing extension headers that may either compromise the network's security posture or require inordinate processing resources.

This memo does not specify a maximum header chain length. However, this memo does note that at the time of its publication, the number of bytes that ASIC-based forwarders copy from the beginning of a packet to on-chip memory varies from platform to platform. Typical platforms copy between 128 and 384 bytes. Therefore, application developers should avoid sending packets who header chain length is in that range, unless they have some assurance that their packets will not be discarded.

## 5. IANA Considerations

This document makes no requests of the IANA

## 6. Security Considerations

TBD

## 7. Acknowledgements

The authors wish to thank Paul Hoffman, KK and Fernando Gont. The authors also express their gratitude to an anonymous donor, without whom this document would not have been written.

## 8. References

### 8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[RFC7112] Gont, F., Manral, V., and R. Bonica, "Implications of Oversized IPv6 Header Chains", RFC 7112, January 2014.

### 8.2. Informative References

[RFC6275] Perkins, C., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, July 2011.

[RFC6554] Hui, J., Vasseur, JP., Culler, D., and V. Manral, "An IPv6 Routing Header for Source Routes with the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6554, March 2012.

## Appendix A. Changes / Author Notes.

[RFC Editor: Please remove this section before publication ]

Template to -00

- Initial submission.

-

-00 to -01

- Added maximum header chain recommendation.

- Rewrite the forwarding description.

-02 to -03

- Updating REQ2 and REQ3 with sending ICMPv6 messages part.

Authors' Addresses

Warren Kumari
Google
1600 Amphitheatre Parkway
Mountain View, CA  94043
US

Email: warren@kumari.net

Joel Jaeggli
Zynga
675 East Middlefield
Mountain View, CA
USA

Email: jjaeggli@zynga.com

Ronald P Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, VA
USA

Email: rbonica@juniper.net

J. Linkova
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
USA

Email: furry@google.com