Network Working Group Internet-Draft Intended status: Informational Expires: January 6, 2013

IPv6 Source Fragmentation for Link Adaptation Avoidance draft-generic-6man-tunfrag-02.txt

Abstract

IPv6 intentionally deprecates fragmentation by routers in the network. Instead, links with restricting MTUs must either drop each too-large packet and return an ICMP Packet Too Big message or perform link-specific fragmentation (also known as "link adaptation") at a layer below IPv6. This latter category of links is often performance-challenged to accommodate steady-state link-specific fragmentation to the point that it would be highly desirable to push the fragmentation burden back to the IPv6 source. A common case that exhibits these link characteristics is seen for IPv6-within-IP tunnels. This document therefore proposes an update to the base IPv6 specification to support link adaptation avoidance.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of \underline{BCP} 78 and \underline{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 <u>http://datatracker.ietf.org/drafts/current/</u>.

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 January 6, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to <u>BCP 78</u> and the IETF Trust's Legal Provisions Relating to IETF Documents (<u>http://trustee.ietf.org/license-info</u>) in effect on the date of

Expires January 6, 2013

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

<u>1</u> .	Introduction \ldots \ldots \ldots \ldots \ldots \ldots 3
<u>2</u> .	Problem Statement
<u>3</u> .	IPv6 Protocol Specification Updates
<u>4</u> .	Additional Fragmentation Considerations
<u>5</u> .	IANA Considerations
<u>6</u> .	Security Considerations
<u>7</u> .	Acknowledgments
<u>8</u> .	References
<u>8</u>	<u>.1</u> . Normative References
<u>8</u>	<u>.2</u> . Informative References
Autl	hor's Address

Expires January 6, 2013 [Page 2]

<u>1</u>. Introduction

IPv6 intentionally deprecates fragmentation by routers in the network. Instead, links with restricting MTUs must either drop each too-large packet and return an ICMP Packet Too Big message or perform link-specific fragmentation (also known as "link adaptation") at a layer below IPv6. This latter category of links is often performance-challenged to accommodate steady-state link-specific fragmentation to the point that it would be highly desirable to push the fragmentation burden back to the IPv6 source. A common case that exhibits these link characteristics is seen for IPv6-within-IP tunnels [I-D.generic-v6ops-tunmtu]. This document therefore proposes an update to the base IPv6 specification to support link adaptation avoidance.

<u>2</u>. Problem Statement

The current "Internet cell size" is effectively 1500 bytes, i.e., the minimum MTU configured by the vast majority of links in the Internet. However, due to issues with Path MTU Discovery (PMTUD) this size can only be accommodated when links with smaller link-layer segment sizes are permitted to perform link adaptation. A common example of such links is seen for IPv6-within-IP tunnels. For those links, the tunnel ingress can perform fragmentation on the outer packet following encapsulation and can instead (or in addition) perform "tunnel fragmentation" via an encapsulation mid-layer inserted between the inner and outer header. In both cases reassembly would be performed by the tunnel egress.

Unfortunately, link-layer fragmentation can present a significant burden to the link endpoints, i.e., especially when the link supports high data rates and/or is located nearer the "middle" of the network instead of nearer the "edge". The third alternative therefore is to ask the original IPv6 source to perform fragmentation on the packet before sending it out, in which case reassembly would be performed by the final destination. This document therefore updates the IPv6 protocol specification [RFC2460] to allow links that perform link adaptation to send advisory messages to the original source as described in the next section.

<u>3</u>. IPv6 Protocol Specification Updates

Section 5 of [RFC2460] states:

"IPv6 requires that every link in the internet have an MTU of 1280 octets or greater. On any link that cannot convey a 1280-octet

packet in one piece, link-specific fragmentation and reassembly must be provided at a layer below IPv6."

This document does not propose to change this requirement, but notes that link-specific fragmentation can be burdensome for some links (e.g., IPv6-within-IP tunnels), to the point that it would be highly desirable for the fragmentation and reassembly to be pushed back to the communication endpoints. In order to accommodate this, when the router at the link ingress performs link adaptation on a packet it should also send an advisory ICMPv6 Packet Too Big (PTB) message back to the original source (subject to rate limiting). This document therefore proposes to add the following specification as a new final paragraph to the end of <u>Section 5</u>:

"In response to an IPv6 packet that is sent to an IPv6 destination located beyond a link that must perform link-specific fragmentation, the originating IPv6 node may receive an ICMP Packet Too Big message reporting a Next-Hop MTU less than 1280. In that case, the IPv6 node should perform IPv6 fragmentation on any subsequent packets that are larger than this MTU value but no larger than the minimum of the source's link MTU and 1500 bytes. Note that these Packet Too Big messages are advisory in nature and do not necessarily indicate packet loss. Note also that the node is permitted to continue to send packets larger than 1500 bytes without fragmentation, but should implement [<u>RFC4821</u>] to ensure that the packets are reaching the final destination."

An example tunnel protocol that invokes this new clause appears in: [<u>I-D.templin-intarea-seal</u>].

<u>4</u>. Additional Fragmentation Considerations

Regardless of whether there is a link that performs link adaptation in the path, when an original source receives a PTB message reporting a Next-Hop MTU value greater than 1280 but less than 1500, the source need not reduce the size of the packets it sends but may instead invoke fragmentation with a maximum fragment size set to the Next-Hop MTU. These fragments will be reassembled by the final destination, which is required by [<u>RFC2460</u>] to configure a minimum reassembly buffer size of 1500 bytes.

5. IANA Considerations

There are no IANA considerations for this document.

Internet-Draft

<u>6</u>. Security Considerations

The security considerations for [RFC2460] apply also to this document.

7. Acknowledgments

This method was inspired through discussion on the IETF v6ops and NANOG mailing lists in the May/June 2012 timeframe.

8. References

8.1. Normative References

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

8.2. Informative References

```
[I-D.generic-v6ops-tunmtu]
    Templin, F., "Operational Issues with Tunnel Maximum
    Transmission Unit (MTU)", draft-generic-v6ops-tunmtu-09
    (work in progress), July 2012.
```

```
[I-D.templin-intarea-seal]
Templin, F., "The Subnetwork Encapsulation and Adaptation
Layer (SEAL)", <u>draft-templin-intarea-seal-42</u> (work in
progress), December 2011.
```

[RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", <u>RFC 4821</u>, March 2007.

Author's Address

Fred L. Templin (editor) Boeing Research & Technology P.O. Box 3707 Seattle, WA 98124 USA

Email: fltemplin@acm.org

Expires January 6, 2013 [Page 5]