Network Working Group
Internet-Draft
Intended status: Informational
Expires: November 19, 2012

F. Templin, Ed.
Boeing Research & Technology

May 18, 2012

6rd Tunnel MTU

draft-foo-v6ops-6rdmtu-00.txt

Abstract

The 6rd tunnel MTU is currently recommended to be set to 1480. This is to avoid IPv4 fragmentation within the tunnel, but requires the 6rd tunnel ingress interface to drop any IPv6 packet larger than 1480 bytes and return an ICMPv6 Packet Too Big (PTB) message. Concerns for operational issues with both IPv4 and IPv6 Path MTU Discovery point to the possibility of MTU-related black holes when a packet is dropped due to an MTU restriction, so dropping packets is considered highly undesirable. Fortunately, the "Internet cell size" is 1500 bytes, i.e., the minimum MTU configured by the vast majority of links in the Internet. This document therefore presents a method to boost the 6rd MTU to 1500 bytes.

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 http://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 November 19, 2012.

Copyright Notice

Table of Contents

1. Introduction

The 6rd tunnel MTU is currently recommended to be set to 1480 [RFC5969]. This is to avoid IPv4 fragmentation within the tunnel [RFC0791], but requires the 6rd tunnel ingress interface to drop any IPv6 packet larger than 1480 bytes and return an ICMPv6 Packet Too Big (PTB) message [RFC2460]. Concerns for operational issues with both IPv4 and IPv6 Path MTU Discovery [RFC1191][RFC1981] point to the possibility of MTU-related black holes when a packet is dropped due to an MTU restriction, so dropping packets is considered highly undesirable. Fortunately, the "Internet cell size" is 1500 bytes, i.e., the minimum MTU configured by the vast majority of links in the Internet, such that 1500 byte or smaller packets are likely to be delivered without loss to MTU limitations in the vast majority of cases. This document therefore presents a method to boost the 6rd tunnel MTU to 1500 bytes.

Pushing the 6rd tunnel MTU to 1500 bytes is met with the challenge that the addition of the IPv4 encapsulation header would cause a 1500 byte IPv6 packet to appear as a 1520 byte IPv4 packet on the wire. This can result in the packet being either fragmented or dropped by an IPv4 router that configures a 1500 byte link, depending on the setting of the "Don't Fragment" (DF) bit in the IPv4 header. Using the approach outlined in this document, the 6rd tunnel avoids this issue by performing IPv6 fragmentation on the inner IPv6 packet before IPv4 encapsulation. The approach is outlined in the following sections.

2. Setting the 6rd MTU to 1500 Bytes

Setting the 6rd MTU to 1500 bytes is accomplished via the following algorithm:

1. set the 6rd tunnel interface MTU to 1500
2. for IPv6 packets to be admitted into the 6rd tunnel, do the following:

   a) drop the packet and send an ICMPv6 PTB if it is 1501 or more
   b) encapsulate the packet in an IPv4 header and send it to the tunnel far end if it is 1280 or less
   c) if the packet is between 1281 - 1500:
      - break it into 2 equal-sized pieces and insert a fragment header on both pieces
      - choose a random 32-bit value and write the value in the Identification field in both pieces
      - encapsulate both pieces in an IPv4 header and send them to the tunnel far end
3. the IPv6 destination host gets to reassemble if necessary

3. Discussion

In the algorithm given in Section 2, the 6rd tunnel interface MTU for 6rd CPE routers and/or Border Routers (BRs) can be set to 1500 bytes independently of other 6rd interfaces within the operator network, i.e., the approach is incrementally deployable.

The algorithm in Section 2 further ignores the IPv6 requirement that routers in the network must not fragment IPv6 packets, i.e. fragmentation must be performed only by hosts. However, we observe that the 6rd CPE is close enough to the IPv6 host such that its fragmentation behavior is very close to that of a host. We further observe that the 6rd BR is tied through stateless address mapping to a single CPE that provides access to the 6rd site such that there would not be multiple paths into the site via different CPEs with widely varying delay characteristics.

The algorithm in Section 2 requires that 6rd ingress tunnel endpoint perform IPv6 fragmentation (when necessary), but the 6rd egress tunnel endpoint does not perform reassembly; hence, the 6rd tunnel endpoints remain stateless. Instead, the final destination IPv6 host may be obliged to reassemble IPv6 packets up to 1500 bytes in length, but hosts are required to reassemble at least that much by the IPv6 standard [RFC2460].

Since the 6rd tunnel endpoints set the Identification field in fragmented IPv6 packets to a random 32-but value, there is no possibility for "serialization" in which an IPv6 host might find the

number of distinct outstanding Identification values reduced due to the 6rd tunnel endpoint applying a single serial Identification value for all IPv6 hosts. And, the use of a random 32-bit value would still allow for far more than enough outstanding IPv6 packet reassemblies without risk of colliding Identification values. This is especially true since 6rd sites do not connect to their ISPs via high data rate interfaces, i.e., their interface data rates are normally O(10Mb/sec) and not O(10Gbps).

4. IANA Considerations

There are no IANA considerations for this document.

5. Security Considerations

The security considerations for 6rd apply also to this document.

6. Acknowledgments

This method was inspired through discussion on the IETF v6ops list in the May 2012 timeframe.

7. References

7.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

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

[RFC5969] Townsley, W. and O. Troan, "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) -- Protocol Specification", RFC 5969, August 2010.

7.2. Informative References

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

Author's Address

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

Email: fltemplin@acm.org