                 TCP Fails To Respect IPV6_USE_MIN_MTU
                 draft-andrews-tcp-and-ipv6-use-minmtu-04

Abstract

   The IPV6_USE_MIN_MTU socket option directs the IP layer to limit the
   IPv6 packet size to the minimum required supported MTU from the base
   IPv6 specification, i.e. 1280 bytes. Many implementations of TCP
   running over IPv6 neglect to check the IPV6_USE_MIN_MTU value when
   performing MSS negotiation and when constructing a TCP segment
   despite MSS being defined to be the MTU less the IP and TCP header
   sizes (60 bytes for IPv6). This leads to oversized IPv6 packets
   being sent resulting in unintended Path Maximum Transport Unit
   Discovery (PMTUD) being performed and to fragmented IPv6 packets
   being sent.

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 April 20, 2016.

Copyright Notice

Table of Contents

1. Introduction

The IPV6_USE_MIN_MTU [RFC3542], Section 11.1, socket option directs the IP layer to limit the IPv6 packet size to the minimum required supported MTU from the base IPv6 specification [RFC2460], i.e. 1280 bytes. Many implementations of TCP running over IPv6 neglect to check the IPV6_USE_MIN_MTU value when performing MSS negotiation and when constructing a TCP segment despite MSS being defined to be the MTU less the IP and TCP header sizes [RFC0879] (60 bytes for IPv6). This leads to oversized IPv6 packets being sent resulting in unintended Path Maximum Transport Unit Discovery (PMTUD) [RFC1191] being performed and to fragmented IPv6 packets being sent.

TCP, when running over IPv6, SHOULD check the value of IPV6_USE_MIN_MTU when performing MSS negotiation. TCP implementations already use learnt PMTU and interface MTU when performing MSS negotiation. This is yet another constraint on the MTU which SHOULD be considered.

TCP, when running over IPv6, SHOULD check the value of IPV6_USE_MIN_MTU when calculating the segment size to send. TCP implementations already use learnt PMTU and interface MTU when calculating the segment size to send.

1.1. Reserved Words

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 [RFC2119].

2. MSS Negotiation

TCP, when running over IPv6, SHOULD check the value of
IPV6_USE_MIN_MTU when performing MSS negotiation. If the value of
IPV6_USE_MIN_MTU is one (1) then the application has requested that
PMTUD not be performed on the socket and that IPv6 packets be sent at
a size no greater then the network minimum MTU of 1280 bytes. This
means that the TCP MSS negotiation size SHOULD be no bigger than 1220
(1280 - 40 - 20) to account for the IPv6 header and the TCP header
and MAY be smaller.

If this negotiation is properly performed then PMTUD of reply traffic
should not normally occur.

3. Segment Size Calculation

TCP, when running over IPv6, SHOULD check the value of
IPV6_USE_MIN_MTU when calculation the next segment to send. If the
value of IPV6_USE_MIN_MTU is one (1) them the maximum segment size
SHOULD be 1220.

If the TCP layer neglects to check the value of IPV6_USE_MIN_MTU and
it is one (1), the packet, when passed to the IPv6 layer, will be
fragmented if the resulting packet is bigger that 1280 octets. This
can result in communications failures due to intermediate nodes not
passing fragmented packets.

4. Current Usage

A example of current usage of IPV6_USE_MIN_MTU=1 and TCP is in DNS
nameservers. This is done as the TCP message streams are normally no
more than a couple of IPv6 packets so there is little benefit in
using maximum sized packet, and no real negative effects from using
smaller packets. There are lots of servers / clients that these
servers talk to and maintaining PMTU knowledge is not effective for
long enough resulting in PMTUD being repeated performed. There are
external time constraints where recovery from lost ICMPv6 PTB will
result in a elapsed transaction time that falls outside of the time
constraint window.

5. Acknowledgements

I would like to thank Havard Eidnes, Sander Steffann, John Leslie,
and Brian E Carpenter for their feedback.

6. Security Considerations

   The document makes no changes that could impact on the security of a
   IPv6 stack.

7. IANA Considerations

   There are no actions for IANA.

8. Normative References

   [RFC0879] Postel, J., "The TCP Maximum Segment Size and Related
              Topics", RFC 879, DOI 10.17487/RFC0879, November 1983,
              <http://www.rfc-editor.org/info/rfc879>.

   [RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191,
              DOI 10.17487/RFC1191, November 1990,
              <http://www.rfc-editor.org/info/rfc1191>.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/
              RFC2119, March 1997,
              <http://www.rfc-editor.org/info/rfc2119>.

   [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6
              (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460,
              December 1998, <http://www.rfc-editor.org/info/rfc2460>.

   [RFC3542] Stevens, W., Thomas, M., Nordmark, E., and T. Jinmei,
              "Advanced Sockets Application Program Interface (API) for
              IPv6", RFC 3542, DOI 10.17487/RFC3542, May 2003,
              <http://www.rfc-editor.org/info/rfc3542>.

Author's Address

   M. Andrews
   Internet Systems Consortium
   950 Charter Street
   Redwood City, CA 94063
   US

   Email: marka@isc.org