DNS and UDP Fragmentation
draft-andrews-dnsext-udp-fragmentation-01.txt

Abstract

This document provides advice to DNS developers about sending DNS UDP messages and Path MTU Discovery.

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 July 25, 2012.

Copyright Notice

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

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

1. Background ................................................................. 3  
   1.1. Reserved Words ..................................................... 3  
2. IPv4 and Fragmentation .................................................. 3  
3. IPv6 and Fragmentation .................................................. 3  
4. IANA Considerations ..................................................... 4  
5. Security Considerations ................................................ 4  
6. Normative References .................................................... 4  

Author’s Address ............................................................ 5
1. Background

IPv6 has changed the dynamics of UDP, Path MTU Discovery (PMTUD) and IP fragmentation. With IPv4 DNS/UDP packets were fragmented by the network and no PMTUD was performed. With IPv6 fragmentation occurs in the sending node and PMTUD is always performed unless the IPv6 packet is fragmented by the sending node using the minimum IPv6 MTU.

DNS/UDP does not work well when PMTUD is performed. If the Packet Too Big (PTB) / Need Fragmentation ICMP messages are not received there is no feedback path in DNS to reduce the size of the fragments like there is with TCP.

Additionally there is no automatic retransmission of UDP packets like there is with TCP in response to a PTB message. The sender needs to send the request after timing out. Not only is this process slow, the resulting traffic patterns can be confused with other common sources of error, resulting from badly configured firewalls, leading to inappropriate remedial action being taken.

This document recommends that all DNS/UDP messages are sent such that they do not trigger PMTUD.

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 [RFC 2119].

2. IPv4 and Fragmentation

There are a number of IP stacks that enable PMTUD for all IP packets by default against the advice of [RFC 1191]. On those IP stacks it is necessary for the application to disable PMTUD on a per socket/packet basis or for the operator to disable it globally if there is no per socket/packet control.

3. IPv6 and Fragmentation

It was realised that IPv6 changed the way PMTUD happened and that there were applications, like DNS, that would not work well with PMTUD. For those applications a socket option called IPV6_USE_MIN_MTU was developed [RFC 3542] which tells the IPv6 stack to fragment packets at the minimum IPv6 MTU rather than use PMTUD to find the actual PMTU.
It is RECOMMENDED that IPV6_USE_MIN_MTU be set to 1 (one) when sending DNS/UDP messages over IPv6. This option can be set at the socket level or it can be set on a per UDP datagram basis.

If the IPv6 stack does not support IPV6_USE_MIN_MTU, then steps should be taken to prevent PMTUD occuring. These include, but are not limited to, setting the MTU of the interface the packets are being sent over to the minimum IPv6 MTU (1280 bytes), or restring DNS/UDP packets to no more than 1280 bytes including IPv6 headers.

It should be noted that even with IPV6_USE_MIN_MTU set to one that a PTB message may still be received [RFC 2460] which requires a IPv6 to add a Fragmentation header to subsequent packets. There is currently no way to avoid this, without using raw sockets, as there is no way for an application to request that a Fragmentation header be added to a packet. [I-D. draft-andrews-6man-force-fragmentation] however has some proposed methods.

4. IANA Considerations

No IANA Considerations.

5. Security Considerations

Failure to prevent PMTUD can lead to denial of service for DNS clients.

Firewalls are often configured to block fragmented IP packets as early IP stacks had fragmentation re-assembly bugs. These bugs were exploited to perform a number of denial of service and other attacks circa 1999.

Such blocks should be relaxed to permit fragmented UDP packets.

6. Normative References

[I-D. draft-andrews-6man-force-fragmentation]
Andrews, M., "Forcing Fragmentation of IPv6 Packets",
draft-andrews-6man-force-fragmentation-01 (work in progress), January 2012.

[RFC 1191]
[RFC 2119]
Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[ RFC 2460 ]

[ RFC 3542 ]

Author’s Address

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

Email: marka@isc.org