

INTERNET-DRAFT
Updates [RFC 1035](#)
Expires December 1998

Donald E. Eastlake, 3rd
CyberCash
June 1998

Bigger Domain Name System UDP Replies

Donald E. Eastlake 3rd

Status of This Document

This draft, file name [draft-ietf-dnsind-udp-size-02.txt](#), is intended to become a Proposed Standard RFC. Distribution of this document is unlimited. Comments should be sent to the DNS mailing list <namedroppers@internic.net> or to the author.

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months. Internet-Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet-Drafts as reference material or to cite them other than as a ``working draft'' or ``work in progress.''

To view the entire list of current Internet-Drafts, please check the "lid-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), ftp.nordu.net (Northern Europe), ftp.nis.garr.it (Southern Europe), munnari.oz.au (Pacific Rim), ftp.ietf.org (US East Coast), or ftp.isi.edu (US West Coast).

Abstract

The Domain Name System defaults to using UDP for queries and replies with a DNS payload limit of 512 bytes. Larger replies cause an initial truncation indication leading to a subsequent handling via TCP with substantially higher overhead. An extension to DNS UDP requests is specified which frequently permits larger UDP responses thus reducing the need for use of TCP.

INTERNET-DRAFT

Bigger DNS UDP Replies

Acknowledgements

Paul Vixie originated the basic idea specified herein.

Some errors notice by Chris Thompson in version -00 have been fixed.

Additional suggestions were made by James Gilroy.

Table of Contents

| | |
|--|-------------------|
| Status of This Document..... | 1 |
| Abstract..... | 1 |
| Acknowledgements..... | 2 |
| Table of Contents..... | 2 |
| 1 . Introduction..... | 3 |
| 2 . Permitting Larger DNS UDP Packets..... | 3 |
| 3 . Compatibility Discussion..... | 5 |
| 4 . Security Considerations..... | 5 |
| References..... | 6 |
| Author's Addresses..... | 6 |
| Expiration and File Name..... | 6 |

INTERNET-DRAFT

Bigger DNS UDP Replies

1. Introduction

The global Internet Domain Name System (DNS) is documented in [RFC 1034](#), 1035, and numerous additional Requests for Comment. It provides a distributed hierarchical database with redundant servers. Recently security features have been added to the DNS [[RFC 2065](#)].

DNS can transfer data via both UDP and TCP. Some requests that are very likely to have big responses, most commonly zone transfers, just use TCP. However, the vast majority of requests are initially sent via UDP which causes the response to be via UDP.

DNS over UDP is constrained to one packet for the request, which is normally no problem as requests are usually small, and one packet for response, which can be a problem. The DNS data portion of DNS UDP packets is currently limited to 512 bytes. The standard states that if the data required to be in the response to a UDP request does not fit in 512 bytes, a truncation flag bit is set in the response and the resolver must try again using TCP with TCP's substantially higher set up and tear down overhead.

As signatures and/or keys are included in more responses due to DNS security [[RFC 2065](#)] and average domain names get longer and larger addresses for IPv6 [[RFC 1886](#)] come into use and there are increasing numbers of instances of larger RRsets, the old UDP response size limit will increasingly be exceeded. Yet the bulk of the network has MTUs on the order of the Ethernet MTU or larger (in some cases simulated by link adaptation layers that disguise a smaller physical MTU) and all modern IP stacks can handle buffering of that size or larger.

[2. Permitting Larger DNS UDP Packets](#)

Efforts are under way to define an additional information resource record that can be used to communicate exact buffer sizes and many other options and extensions. However, many older DNS servers ignore any request with a non-empty additional information section resulting in a requirement for probing and maintaining per server state for optimal performance. Pending deployment of some such more exact and comprehensive solution, the following change is made in the DNS over UDP protocol.

No change is made in the size limit on UDP queries. It remains at 512 bytes.

The presently unused RCODE field in UDP queries is redefined to specify the resolver imposed limit on the DNS data in the UDP response. This four bit field is presently specified as zero. Such

Donald E. Eastlake, 3rd

[Page 3]

INTERNET-DRAFT

Bigger DNS UDP Replies

non-zero RCODE values in requests will be ignored by older DNS servers that will continue to use the old UDP size limit for responses. Thus server probing and state maintenance are not required. Values from 0 through 10 are defined as follows:

| RCODE | DNS reply data limit in bytes |
|-------|---|
| 0 | 512 (old default) |
| 1 | 768 |
| 2 | 1,280 (new default) |
| 3 | 1,920 |
| 4 | 3,200 (appropriate for UDP entirely via FDDI) |
| 5 | 4,800 |
| 6 | 8,000 |
| 7 | 12,000 |
| 8 | 20,000 |
| 9 | 30,000 |
| 10 | 50,000 |
| 11-15 | -reserved |

A resolver should take into account its local buffer space and any knowledge it has about the local network MTU (maximum transmission unit) or the PMTU (path MTU) to the server it is querying. Making a

reasonable allowance for IP headers that may be added by the server, the resolver should then pick an RCODE value from the above table. A value that might be expected to cause a reply packet to fragment into two pieces is still preferable to using TCP. In the absence of any information, the value 2 should be used.

The resolver should not do PMTU discovery just to provide a more accurate RCODE. The additional packets that might be required for PMTU discovery would defeat the purpose of avoiding the additional packets required by TCP.

A server, on receiving a query with a non-zero RCODE, MUST limit its DNS response message to the size specified but may need to limit it to a lower amount due to buffer space available. It SHOULD also limit it based on local network MTU or the PMTU to the resolver, if known, less a reasonable allowance for IP headers.

1280 bytes of DNS data is chosen as the new default to provide a generous allowance for IP headers and still be within the highly prevalent approximately Ethernet size or larger MTU and buffering generally available today.

An IPv6 server should enable fragmentation on UDP replies. While fragmentation will not be frequent if the above guidelines are followed, it may occur on occasion. In principle, IPv6 headers and options could be huge, resulting in a very large UDP packet even though the DNS payload is limited, but this should not occur in

practice.

[3.](#) Compatibility Discussion

No cases are known where the above change will cause a problem for non-recursive queries. Old servers will ignore the RCODE field of the UDP query and should return 512 or fewer bytes, possibly with a truncation indication. Servers with this feature included should use the RCODE value to determine a ceiling on the size of response they will send. Non-zero values of RCODE will permit them to send larger UDP responses if local conditions are appropriate.

There is a potential problem with recursive queries. If (1) an

updated recursive query specifies bigger UDP responses with a non-zero RCODE to an old server and (2) that server in turn issues a corresponding query into which it blindly copies the RCODE field and (3) this corresponding query goes to an updated server that honors the non-zero RCODE field and (4) the updated server response DNS data is actually larger than 512 bytes as permitted by the RCODE in the query, then the intermediate old DNS server may be confused by the larger than 512 byte DNS response it receives. However, there are already DNS implementations out there on the Internet that send back larger than 512 byte responses in violation of the old standard and DNS implementations are being deployed which protect themselves against and are not confused by larger than expected responses.

Should the above problem manifest itself, it can be cured by making the queries be TCP based or non-recursive or by upgrading the intermediate DNS server to which the recursive queries are being sent to implement this bigger UDP packet feature. There are cases, such as resolvers behind a firewall that can only get outside DNS information via a recursive server and changing to non-recursive queries is not possible. Upgrading the DNS server is the strongly recommended solution.

[4. Security Considerations](#)

General DNS security issues are considered in [RFC 2065](#).

In the absence of request security [[RFC 2065](#)], the request RCODE could be modified in transit. If set lower, this might result in unnecessary TCP. If set higher, this might result in unnecessary fragmentation.

Larger packets may make it easier to cause some forms of denial of service due to fragment loss.

References

[RFC 1034](#) - P. Mockapetris, "Domain names - concepts and facilities", 11/01/1987.

[RFC 1035](#) - P. Mockapetris, "Domain names - implementation and specification", 11/01/1987.

[RFC 2065](#) - D. Eastlake, C. Kaufman, "Domain Name System Security Extensions", 01/03/1997.

Author's Addresses

Donald E. Eastlake 3rd
CyberCash, Inc.
318 Acton Street
Carlisle, MA 01741 USA

Telephone: +1 978 287 4877
 +1 703 620-4200 (main office, Reston, Virginia)
FAX: +1 978 371 7148
EMail: dee@cybercash.com

Expiration and File Name

This draft expires December 1998.

Its file name is [draft-ietf-dnsind-udp-size-02.txt](#).