Mark Andrews (CSIRO)
Paul Vixie (ISC)

INTERNET-DRAFT
<draft-andrews-dns-more-01.txt>
June 1996

Updates: RFC 1035

Large Responses to DNS Queries (DNS MORE)

Status of This Memo

Abstract

DNS messages are limited to 64 kilobytes in size. At times it is necessary to send a message that is greater that 64 kilobytes. This is currently not possible. AXFR is the one exception. This document describes how to send a sequence of messages, the total length which may be greater than 64 kilobytes, by extending the protocol.

In addition average message sizes are increasing and the 512 byte payload limit for UDP is now too small. This document describes how servers can identify when they can send bigger messages without necessarily resorting to TCP.

1 - Protocol

This extension uses one of the RESERVED flags bits from DNS header [RFC1035 4.1.1] to indicate when a server can send the extended response. This flag bit shall be known as MORE.

The MORE flag's semantics depend upon the underlying transport protocol.

This document only defines the use of the MORE flag with the opcode QUERY.

1.1 - TCP Usage

When using TCP a resolver sets the MORE flag to indicate that it is capable of receiving a multi message response (which we call a ``message sequence'').

To indicate that the message sequence is not complete the server shall set the RCODE to CONTINUED (TBA) in all but the last message of the message sequence.

The order of resource records in a multi message response MUST be the same as if the response could have been sent is a single response. The Questions first followed by, the Answer RRs, Authority RRs and Additional RRs.

Each message in a sequence will contain a header with the same ID value, flags, opcode. Only the count fields and the rcode are permitted to change. The counts shall represent the number of resource records in this message. MORE MUST cleared in the response.

1.1.1 - TCP Example

The following example show how to send an answer with one question, 10 answer records, 14 authority records and 5 additional records. The answer is split up across 3 messages.

```
MESSAGE 1: QCOUNT=1, ANCOUNT=10, AUCOUNT=0,
           ADCOUNT=0, RCODE=CONTINUED
MESSAGE 2: QCOUNT=0, ANCOUNT=0, AUCOUNT=11,
           ADCOUNT=0. RCODE=CONTINUED
MESSAGE 3: QCOUNT=0, ANCOUNT=0, AUCOUNT=3,
           ADCOUNT=5, RCODE=NOERROR
```

1.2 - UDP Usage

When using UDP, a resolver may set the MORE flag in a QUERY request to indicate that its receive buffer is greater than 512 bytes in size, rather than the 512 byte size given in [RFC1035 3.2.4]. The resolver is expected to set this flag only if it knows that the host's reassembly buffer is large enough to accommodate datagrams of the size indicated.

The new size is indicated by the RCODE is the query. The receive buffer is 512 * (2 ^ RCODE) bytes in size.

RCODE	SIZE
0	512
1	1024
2	2048
3	4096
4	8192
5	16384
6	32768
7	65536
8	131072
9	262144
10	524288
11	1048576
12	2097152
13	4194304
14	8388608
15	16777216

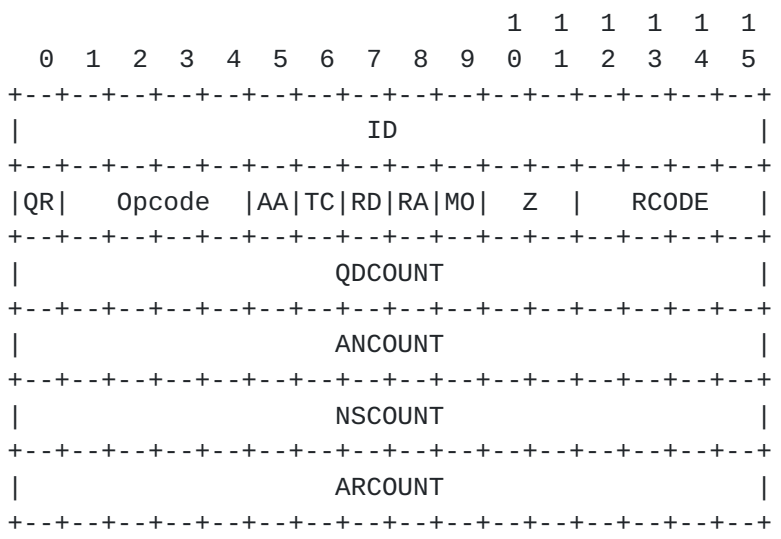
A server receiving a QUERY request with the MORE flag set is allowed to transmit a response of up to the size indicated. If the response will not fit in size indicated, then the rules given in [RFC1035 4.1.1, 4.2.1, 6.2] apply.

The server MUST clear the MORE flag in the response.

The server SHOULD disable path MTU discovery on the UDP response packet resulting in host fragmentation.

2 - Header Format

The header format is that described in [RFC1035 4.1.1] with the MORE flag added:



Where MO is the MORE flag.

3 - Security Considerations

Though DNS is related to several security problems, no attempt is made to fix them in this document.

This document is believed to introduce no additional security problems to the current DNS protocol.

References

[RFC1035]P. Mockapetris, ``Domain Names - Implementation and Specification,'' RFC 1035, USC/Information Sciences Institute, November 1987.

Authors' Addresses

Mark Andrews
CSIRO - Division of Mathematics and Statistics
Locked Bag 17
North Ryde NSW 2113
AUSTRALIA
+61 2 325 3148
<Mark.Andrews@dms.csiro.au>

Paul Vixie
Internet Software Consortium
Star Route Box 159A
Woodside, CA 94062
USA
+1 415 747 0204
<paul@vix.com>