DNS Extensions Working Group
Internet-Draft
Intended status: Standards track
Expires: February 2010

G. Barwood

25 August 2009

EDNS Page Option
draft-barwood-dnsext-edns-page-option-04

Status of this Memo

Copyright Notice

Abstract

Describes an EDNS option to allow large DNS responses to be sent using small UDP packets.

Table of Contents

1. Introduction 3

2. Protocol 4

2.1 Initial request 4

2.2 Server response 4

2.3 Follow-up request 5

3. Compatibility 6

4. Security Considerations 6

5. IANA Considerations 7

6. Acknowledgments 7

7. Informative References 7

1. Introduction

DNSSEC implies that DNS responses may be large, possibly larger than the de facto ~1500 byte internet MTU. The IP protocol specifies a means by which large IP packets are split into fragments and then re-assembled.

Fragmented UDP responses are undesirable for several reasons:

(1) Fragments can easily be spoofed. The DNS ID and port number are only present in the first fragment, and the IP ID is usually easy for an attacker to predict.

(2) In practise fragmentation is not reliable, and large UDP packets may fail to be delivered.

(3) If a single fragment is lost, the entire response must be re-sent.

(4) Re-assembling fragments requires buffer resources, which opens up denial of service attacks.

Instead, it is possible to use TCP for large responses, but this is undesirable, as TCP imposes significant overhead and state that may be vulnerable to denial of service attack.

Nearly all current DNS traffic is carried by UDP with a maximum size of 512 bytes, and relying on TCP is a risk for the deployment of DNSSEC.

A particular problem occurs with DNS proxies, which often truncate responses at 512 bytes. In this case, TCP does not help, and it is impossible to retrieve responses through the proxy.

Therefore an EDNS option [RFC2181] to allow large DNS responses to be sent using small UDP packets is proposed.

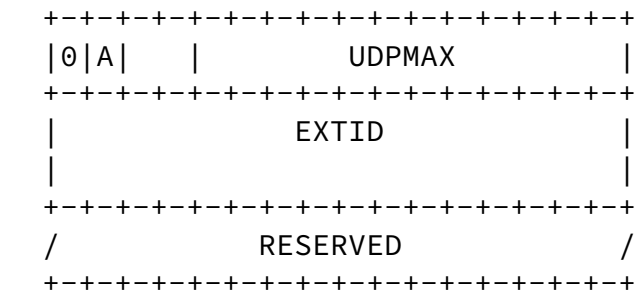
The option includes an authentication mechanism that prevents blind spoofing of the response, provided IP fragmentation does not occur.

2. Protocol

Reserved areas and undefined bits must be set to zero length / zero by the sender and must be ignored by the receiver.

2.1 Initial request

The client signals support in it's initial request by including an EDNS Page option with option data :



where :

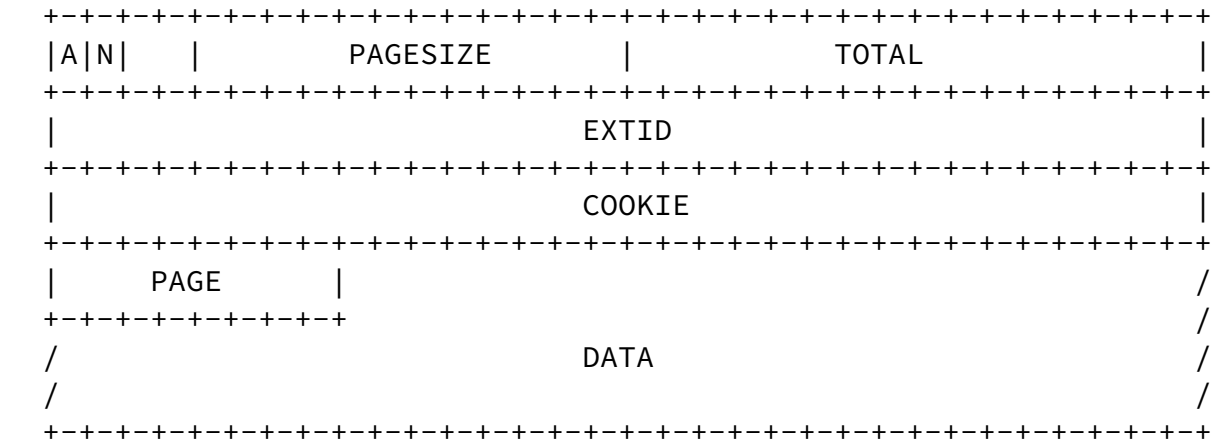
A is a 1 bit field set to request that the server send all pages immediately. It must not be used with proxy servers that do not support it, except for discovery. The server may decline the request.

UDPMAX is a 12 bit field that limits the UDP payload of response packets. Commonly set to 512, as proxies often limit responses to the RFC 1035 UDP limit. The minimum value is 512.

EXTID is a 32 bit field used to validate the response, preventing blind spoofing.

2.2 Server response

The server responds with an EDNS Page option. The server does not send a copy of the question. The Page option data is :



where :

A is a 1 bit field set to indicate that all pages have been sent.

N is a 1 bit field set to indicate that the cookie is omitted, and follow-up requests are not possible.

PAGESIZE is a 12 bit field, the size of the pages into which the full response is divided, chosen so that the UDP payload does not exceed UDPMAX from the initial request. Servers may also limit the UDP payload for other reasons, for example to mitigate an amplification attack, or to avoid IP

fragmentation.

TOTAL is a 16 bit field, the size in bytes of the whole response.

EXTID is a copy of the EXTID from the request. The client must check that the value is as expected.

COOKIE is a 32 bit field, used in follow-up requests.

PAGE is an 8 bit field.

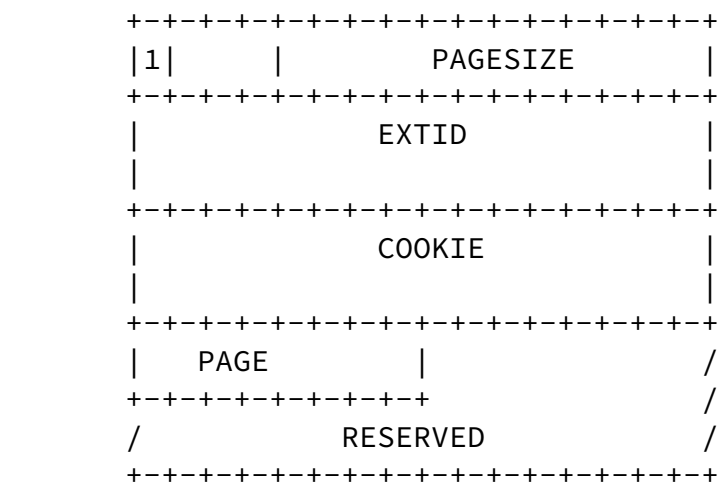
DATA is a variable length field containg the page data.

The client allocates an assembly buffer of TOTAL bytes, and copies DATA into it, at offset PAGE x PAGESIZE.

2.3 Follow-up request

If the A bit of the response is zero, the client sends a follow-up request for each page it has not yet received. The client should also send follow-up requests if an expected response is not received after a timeout period due to packet loss.

A follow-up request is identical to the initial request, except that the EDNS page option data is as follows:



where :

PAGESIZE is a copy of PAGESIZE in the response.

EXTID is a copy of EXTID in the initial request.

COOKIE is a copy of COOKIE from the response that identifies a read-only representation of the full response on the server, possibly in conjunction with the Question. The cookie has a lifetime of 5 seconds. After this time has elapsed, a SERVERFAIL error response may be generated.

PAGE identifies the required page.

When the client has received all of the pages, the complete assembled response is then processed normally.

Follow-up requests may be sent in parallel.

3. Compatibility

Servers are not required to support the EDNS Page option, however support is encouraged.

Authoritative servers that do not support the EDNS page option can expect a higher level of TCP traffic.

Authoritative servers need not support cookies. Initial requests to authoritative servers should normally set the A flag. However, cookie support is encouraged, as it allows dropped packets to be retried without re-sending the whole response.

DNSSEC aware recursive servers need to support cookies if they may be accessed via proxy servers that truncate responses at 512 bytes.

DNSSEC validating stub resolvers need to use the EDNS Page option if they may be deployed behind proxy servers that truncate responses at 512 bytes.

Firewalls may not allow multiple responses through, and servers should detect this possibility, and disable multiple responses, if the firewall cannot be re-configured.

4. Security Considerations

The EXTID may expose internal state to an attacker who controls a name server. It is essential that a cryptographically strong source of random numbers be used to generate the secret key. This must be seeded from data that cannot be guessed by an attacker, such as thermal noise or other random physical fluctuations.

Clients must verify the EXTID in each response.

Fragmented responses are vulnerable to blind spoofing, therefore fragmented responses should be avoided if possible.

If the response does not have a Page Option, to avoid a potential

spoofing downgrade attack the client may send an additional query, or adopt other measures to prevent blind spoofing that are outside the scope of this document.

To limit the effectiveness of amplification attacks on third parties, servers should make every effort to limit the maximum number of packets that are sent in response to a single query.

Suggested techniques include:

- Declining requests to send all pages when QTYPE=ANY.

- Not sending the NS RRset when QTYPE=DNSKEY.

- Limiting additional section processing so that it does not contribute to the maximum response.

- Checking that UDPMAX in the initial request is at least 512.

5. IANA Considerations

The EDNS TYPE code for Page Option.

6. Acknowledgments

Mark Andrews, Alex Bligh, Robert Elz, Douglas Otis, Wouter Wijngaards, Nicholas Weaver were each instrumental in creating and refining this specification.

7. Informative References

[RFC2181] P. Vixie, "Extension Mechanisms for DNS (EDNS0)", RFC 2181, August 1999.

Author's Address

George Barwood
33 Sandpiper Close
Gloucester
GL2 4LZ
United Kingdom

Phone: +44 452 722670
EMail: george.barwood@blueyonder.co.uk