dnsop
Internet-Draft
Intended status: Standards Track
Expires: July 8, 2016

W. Kumari
Google
Z. Yan
CNNIC
W. Hardaker
Parsons, Inc.

January 05, 2016

Returning multiple answers in DNS responses. draft-wkumari-dnsop-multiple-responses-02

Abstract

This document (re)introduces the ability to provide multiple answers in a DNS response.

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 8, 2016.

Copyright Notice

Copyright (c) 2016 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. Introduction 2
1.1. Requirements notation 3
2. Background 3
3. Terminology 3
4. Returning multiple answers 3
5. Additional records pseudo-RR 5
5.1. File Format 5
5.2. Wire Format 5
6. Signaling support 6
7. Stub-Resolver Considerations 6
8. Use of Additional information 6
9. IANA Considerations 6
10. Security Considerations 7
11. Acknowledgements 7
12. References 7
12.1. Normative References 7
12.2. Informative References 8
Appendix A. Changes / Author Notes. 8
Authors' Addresses 8

1. Introduction

Often the name being resolved in the DNS provides information about why the name is being resolved, allowing the nameserver to predict what other answers the client will soon query for. By providing multiple answers in the response, the authoritative name server operator can ensure that the recursive server that the client is using has all the answers in its cache.

For example, the name server operator of Example Widgets, Inc (example.com) knows that the example.com web page at www.example.com contains various other resources, including some images (served from images.example.com), some Cascading Style Sheets (served from css.example.com) and some JavaScript (data.example.com). A client attempting to resolve www.example.com is very likely to be a web browser rendering the page and so will need to also resolve all of the other names to obtain these other resources. Providing all of these answers in response to a query for www.example.com allows the recursive server to populate its cache and have all of the answers available when the client asks for them.

Other examples where this technique is useful include SMTP (including the mail server address when serving the MX record), SRV (providing the target information in addition to the SRV response) and TLSA (providing any TLSA records associated with a name).

This is purely an optimization - by providing all of other, related answers that the client is likely to need along with the answer that they requested, users get a better experience, iterative servers need to perform less queries, authoritative servers have to answer fewer queries, etc.

1.1. Requirements notation

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. Background

The existing DNS specifications allow for additional information to be included in the "additional" section of the DNS response, but in order to defeat cache poisoning attacks most implementations either ignore or don't trust additional information (other than for "glue"). For some more background, see [Ref.Bellovin], [RFC1034], [RFC2181].

Not trusting the information in the additional section was necessary because there was no way to authenticate it. If you queried for www.example.com and got back answers for www.invalid.com you couldn't tell if these were actually from invalid.com or if an attacker was trying to get bad information for invalid.com into your cache. In a world of ubiquitous DNSSEC deployment [Ed note: By the time this document is published, there *will* be ubiquitous DNSSEC :-)] the iterative server can validate the information and trust it.

3. Terminology

Additional records Additional records are records that the authoritative nameserver has included in the Additional section.

Primary query A Primary query (or primary question) is a QNAME that the name server operator would like to return additional answers for.

Supporting information Supporting information is the DNSSEC RRSIGs that prove the authenticity of the Additional records.

4. Returning multiple answers

The authoritative nameserver should include as many of the instructed Additional records and Supporting information as will fit in the response packet.

In order to include Additional records in a response, certain conditions need to be met. [Ed note: Some discussion on each rule is below]

1. Additional records MUST only be included when the primary name and each additional record are signed using DNSSEC "valid".

2. Additional records MUST only be served over TCP connections, or when DNS Cookies [ToDo: Ref] are in use. This is to mitigate Denial of Service reflection attacks.[1]

3. Additional records SHOULD be contained within the same zone as the primary name[2], or MAY be additionally be contained within a child zone for which the name server is authoritative for, assuming all DNSSEC validation records required to validate the child(ren) are included as well. Note that the DS record, and NS and glue records for a child zone may be returned even when no other additional data for the child will be included.

4. The DNSSEC supporting information necessary to perform validation on the records must be included. I.E., the RRSIGs required to validate the Additional record information must be included.

5. The authoritative nameserver SHOULD include as many of the additional records as will fit in the response. Each Additional record MUST have its matching Supporting information. Additional records MUST be inserted in the order specified in the Additional records list.

6. Operators SHOULD only include Additional answers that they expect a client to actually need. [3]

[Ed note 1: The above MAY be troll bait. I'm not really sure if this is a good idea or not - moving folk towards TCP is probably a good idea, and this is somewhat of an optional record type. Then again, special handing (TCP only) for a record would be unusual. Additional records could cause responses to become really large, but there are already enough large records that can be used for reflection attacks that we can just give up on the whole "keep responses as small as possible" ship.]

[Ed note 2: This is poorly worded. I mumbled about bailiwick, subdomains, etc but nothing I could come up with was better. Also, is this rule actually needed? I *think* it would be bad for .com servers to be able to include Additional records for www.foo.bar.baz.example.com, but perhaps <handwave>public-suffix-list?! This rule also makes it easier to decide what all DNSSEC information is required.]

[Ed note 3: This is not enforceable.]

5. Additional records pseudo-RR

To allow the authoritative nameserver operator to configure the name server with the additional records to serve when it receives a query to a label, we introduce the Additional Resource Record (RR).

5.1. File Format

The format of the Additional RR is:

label ADD "label,type; label,type; label,type; ..."

For example, if the operator of example.com would like to also return A record answers for images.example.com, css.html.example.com and both an A and AAAA for data.example.com when queried for www.example.com he would enter:

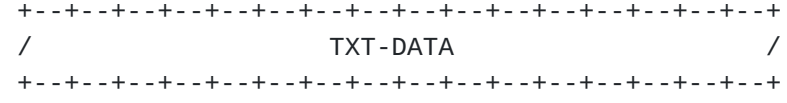
www ADD "images,A; css.html,A; data,A; data,AAA;"

The entries in the ADD list are ordered. An authoritative nameserver SHOULD insert the records in the order listed when filling the response packet. This is to allow the operator to express a preference in case all the records to not fit. The TTL of the records added to the Additional section are MUST be the same as if queried directly.

In some cases the operator might not know what all additional records clients need. For example, the owner of www.example.com may have outsourced his DNS operations to a third party. DNS operators may be able to mine their query logs, and see that, in a large majority of cases, a recursive server asks for foo.example.com and then very soon after asks for bar.example.com, and so may decide to optimize this by opportunistically returning bar when queried for foo. This functionality could also be included in the authoritative name server software itself, but discussions of these re outside the scope of this document.

5.2. Wire Format

The wire format of the Additional RR is the same as the wire format for a TXT RR:



Where TXT-DATA is one or more charter-strings.

The Additional RR has RR type TBD [RFC Editor: insert the IANA assigned value and delete this note]

6. Signaling support

Iterative nameservers that support Additional records signal this by setting the OPT record's PL ("plus") bit (bit NN [TBD: assigned by IANA] in the EDNS0 extension header to 1.

7. Stub-Resolver Considerations

No modifications need to be made to stub-resolvers to get the predominate benefit of this protocol, since the majority of the speed gain will take place between the validating recursive resolver and the authoritative name server. However, stub resolvers may wish to query directly for the Additional RR if it wants to pre-query for data that will likely be needed in the process of supporting its application.

8. Use of Additional information

When receiving Additional information, an iterative server follows certain rules:

1. Additional records MUST be validated before being used.

2. Additional records SHOULD be annotated in the cache as having been received as Additional records.

3. Additional records SHOULD have lower priority in the cache than answers received because they were requested. This is to help evict Additional records from the cache first, and help stop cache filling attacks.

4. Iterative servers MAY choose to ignore Additional records for any reason, including CPU or cache space concerns, phase of the moon, etc. It may choose to only accept all, some or none of the Additional records.

9. IANA Considerations

This document contains the following IANA assignment requirements:

1. The PL bit discussed in Section 6 needs to be allocated.

10. Security Considerations

Additional records will make DNS responses even larger than they are currently, leading to more large records that can be used for DNS reflection attacks. We mitigate this by only serving these over TCP.

A malicious authoritative server could include a large number of Additional records (and associated DNSSEC information) and attempt to DoS the recursive by making it do lots of DNSSEC validation. I don't view this as a very serious threat (CPU for validation is cheap compared to bandwidth), but we mitigate this by allowing the iterative to ignore Additional records whenever it wants.

By requiring the ALL of the Additional records are signed, and all necessary DNSSEC information for validation be included we avoid cache poisoning (I hope :-))

11. Acknowledgements

The authors wish to thank some folk.

12. References

12.1. Normative References

[Ref.Bellovin] Bellovin, S., "Using the Domain Name System for System Break-Ins", 1995, <https://www.cs.columbia.edu/~smb/papers/dnshack.pdf>.

[RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, DOI 10.17487/RFC1034, November 1987, <http://www.rfc-editor.org/info/rfc1034>.

[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>.

[RFC2181] Elz, R. and R. Bush, "Clarifications to the DNS Specification", RFC 2181, DOI 10.17487/RFC2181, July 1997, <http://www.rfc-editor.org/info/rfc2181>.

[RFC5395] Eastlake 3rd, D., "Domain Name System (DNS) IANA Considerations", RFC 5395, DOI 10.17487/RFC5395, November 2008, <http://www.rfc-editor.org/info/rfc5395>.

12.2. Informative References

[I-D.ietf-sidr-iana-objects]
Manderson, T., Vegoda, L., and S. Kent, "RPKI Objects issued by IANA", draft-ietf-sidr-iana-objects-03 (work in progress), May 2011.

Appendix A. Changes / Author Notes.

[RFC Editor: Please remove this section before publication]

From -00 to -01.

- Nothing changed in the template!

Authors' Addresses

Warren Kumari
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
US

Email: warren@kumari.net

Zhiwei Yan
CNNIC
No.4 South 4th Street, Zhongguancun
Beijing 100190
P. R. China

Email: yanzhiwei@cnnic.cn

Wes Hardaker
Parsons, Inc.
P.O. Box 382
Davis, CA 95617
US

Email: ietf@hardakers.net