Workgroup: DNSOP
Internet-Draft:
draft-ietf-dnsop-glue-is-not-optional-04
Updates: 1034 (if approved)
Published: 8 February 2022
Intended Status: Standards Track
Expires: 12 August 2022
Authors: M. Andrews (ISC), S. Huque (Salesforce), P. Wouters (Aiven), D. Wessels (Verisign)

DNS Referral Glue Requirements

Abstract

The DNS uses referral glue records to allow iterative clients to find the addresses of nameservers that are contained within a delegated zone. Authoritative Servers are expected to return all available referral glue records in a referral response. If message size constraints prevent the inclusion of all in-domain referral glue records, the server MUST set the TC flag to inform the client that the response is incomplete, and that the client SHOULD use another transport to retrieve the full response. This document updates RFC 1034 to clarify correct server behavior.

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 https://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 12 August 2022.

Copyright Notice

Copyright (c) 2022 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 (https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction

The Domain Name System (DNS) [RFC1034], [RFC1035] uses referral glue records to allow iterative clients to find the addresses of nameservers that are contained within a delegated zone. Referral glue records are added to the parent zone as part of the delegation process and returned in referral responses, otherwise a resolver following the referral has no way of finding these addresses. Authoritative servers are expected to return all available in-domain referral glue records in a referral response. If message size constraints prevent the inclusion of all in-domain glue records over the chosen transport, the server MUST set the TC (Truncated) flag to inform the client that the response is incomplete, and that the client SHOULD use another transport retrieve the full response. This document clarifies that expectation.

DNS responses sometimes contain optional data in the additional section. In-domain referral glue records, however, are not optional. Several other protocol extensions, when used, are also not optional. This includes TSIG [RFC2845], OPT [RFC6891], and SIG(0) [RFC2931].

At the time of this writing, referral glue is the only type of glue defined for the DNS. Referral glue records are always addresses (A or AAAA records) of a delegation's authoritative name servers. New work underway in the IETF may lead to definitions for other types of glue data, with requirements that differ from referral glue. This document only describes requirements for referral glue. Unless stated otherwise, "glue" in the remainder of this document always means "referral glue."

Note that this document only clarifies requirements of name server software implementations. It does not place any requirements on data placed in DNS zones or registries.

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 [RFC2119].

2. Types of Referral Glue

This section describes different types of glue that may be found in DNS referral responses. Note that the type of glue depends on the QNAME. A particular record can be in-domain glue for one response and sibling glue for another.

2.1. In-Domain Referral Glue

The following is a simple example of glue records present in the delegating zone "test" for the child zone "foo.test". The nameservers for foo.test (ns1.foo.test and ns2.foo.test) are both below the delegation point. They are configured as glue records in the "test" zone:

```
foo.test.                86400   IN NS      ns1.foo.test.
foo.test.                86400   IN NS      ns2.foo.test.
ns1.foo.test.            86400   IN A       192.0.2.1
ns2.foo.test.            86400   IN AAAA    2001:db8::2:2
```

A referral response from "test" for "foo.test" with in-domain glue looks like this:

```
;; QUESTION SECTION:
;www.foo.test.       IN      A

;; AUTHORITY SECTION:
foo.test.               86400        IN      NS      ns1.foo.test.
foo.test.               86400        IN      NS      ns2.foo.test.

;; ADDITIONAL SECTION:
ns1.foo.test.           86400        IN      A       192.0.2.1
ns2.foo.test.           86400        IN      AAAA    2001:db8::2:2
```

2.2. Sibling Referral Glue

Sibling glue are glue records that are not contained in the delegated zone itself, but in another zone delegated from the same parent. In many cases, these are not strictly required for resolution, since the resolver can make follow-on queries to the other zone to resolve the nameserver addresses after following the referral to the sibling zone. However, most nameserver implementations today provide them as an optimization to obviate the need for extra traffic from iterative resolvers.

Here the delegating zone "test" contains two delegations for the child zones "bar.test" and "foo.test":

```
bar.test.                   86400   IN NS      ns1.bar.test.
bar.test.                   86400   IN NS      ns2.bar.test.
ns1.bar.test.               86400   IN A       192.0.2.1
ns2.bar.test.               86400   IN AAAA    2001:db8::2:2

foo.test.                   86400   IN NS      ns1.bar.test.
foo.test.                   86400   IN NS      ns2.bar.test.
```

A referral response from "test" for "foo.test" with sibling glue looks like this:

```
;; QUESTION SECTION:
;www.foo.test.       IN      A

;; AUTHORITY SECTION:
foo.test.               86400        IN      NS      ns1.bar.test.
foo.test.               86400        IN      NS      ns2.bar.test.

;; ADDITIONAL SECTION:
ns1.bar.test.           86400        IN      A       192.0.2.1
ns2.bar.test.           86400        IN      AAAA    2001:db8::2:2
```

2.3. Cyclic Sibling Referral Glue

The use of sibling glue can introduce cyclic dependencies. This happens when one domain specifies name servers from a sibling domain, and vice versa. This type of cyclic dependency can only be broken when the delegating name server includes the sibling glue in a referral response.

Here the delegating zone "test" contains two delegations for the child zones "bar.test" and "foo.test", and each use name servers under the other:

```
bar.test.                 86400   IN NS      ns1.foo.test.
bar.test.                 86400   IN NS      ns2.foo.test.
ns1.bar.test.             86400   IN A       192.0.2.1
ns2.bar.test.             86400   IN AAAA    2001:db8::2:2

foo.test.                 86400   IN NS      ns1.bar.test.
foo.test.                 86400   IN NS      ns2.bar.test.
ns1.foo.test.             86400   IN A       192.0.2.3
ns2.foo.test.             86400   IN AAAA    2001:db8::2:4
```

A referral response from "test" for "bar.test" with sibling glue looks like this:

```
;; QUESTION SECTION:
;www.bar.test.       IN      A

;; AUTHORITY SECTION:
bar.test.               86400        IN      NS      ns1.foo.test.
bar.test.               86400        IN      NS      ns2.foo.test.

;; ADDITIONAL SECTION:
ns1.foo.test.           86400        IN      A       192.0.2.3
ns2.foo.test.           86400        IN      AAAA    2001:db8::2:4
```

In late 2021 the authors analyzed zone file data available from ICANN's Centralized Zone Data Service [CZDS] and found 222 out of approximately 209,000,000 total delegations that had only sibling NS RRs in a cyclic dependency as above.

2.4. Missing Referral Glue

An example of missing glue is included here, even though it can not be considered as a type of glue. While not common, real examples of responses that lack required glue, and with TC=0, have been shown to occur and cause resolution failures.

The example below is based on a response observed in June 2020. The names have been altered to fall under documentation domains. It

shows a case where none of the glue records present in the zone fit into the available space of the UDP response, and the TC flag was not set. While this example shows a referral with DNSSEC records [RFC4033], [RFC4034], [RFC4035], this behavior has been seen with plain DNS responses as well. Some records have been truncated for display purposes. Note that at the time of this writing, the servers originally responsible for this example have been updated and now correctly set the TC flag.

```
% dig +norec +dnssec +bufsize=512 +ignore @ns.example.net \
       rh202ns2.355.foo.example

; <<>> DiG 9.15.4 <<>> +norec +dnssec +bufsize +ignore \
       @ns.example.net rh202ns2.355.foo.example
; (2 servers found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 8798
;; flags: qr; QUERY: 1, ANSWER: 0, AUTHORITY: 9, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
;; QUESTION SECTION:
;rh202ns2.355.foo.example.         IN A

;; AUTHORITY SECTION:
foo.example.          86400   IN NS      rh120ns2.368.foo.example.
foo.example.          86400   IN NS      rh202ns2.355.foo.example.
foo.example.          86400   IN NS      rh120ns1.368.foo.example.
foo.example.          86400   IN NS      rh202ns1.355.foo.example.
foo.example.          3600    IN DS      51937 8 1 ...
foo.example.          3600    IN DS      635 8 2 ...
foo.example.          3600    IN DS      51937 8 2 ...
foo.example.          3600    IN DS      635 8 1 ...
foo.example.          3600    IN RRSIG   DS 8 2 3600 ...
```

3. Requirements

This section describes updated requirements for including glue in referral responses.

3.1. In-Domain Referral Glue

This document clarifies that when a name server generates a referral response, it MUST include all available in-domain glue records in the additional section, or MUST set TC=1 if constrained by message size.

At the time of writing, most iterative clients send initial queries over UDP and retry over TCP upon receiving a response with the TC

flag set. UDP responses are generally limited to between 1232 and 4096 bytes, due to values commonly used for the EDNS0 UDP Message Size field [RFC6891], [FLAGDAY2020]. TCP responses are limited to 65,536 bytes.

3.2. Sibling Referral Glue

This document clarifies that when a name server generates a referral response, it SHOULD include all available glue records in the additional section. If after adding all in-domain glue records, not all sibling glue records fit due to message size constraints, the name server is NOT REQUIRED to set TC=1.

Note that users may experience resolution failures for domains with only sibling glue when a name servers chooses to omit them in a referral response. As described in Section 2.3, such domains are rare.

3.3. Updates to RFC 1034

Replace

"Copy the NS RRs for the subzone into the authority section of the reply. Put whatever addresses are available into the additional section, using glue RRs if the addresses are not available from authoritative data or the cache. Go to step 4."

with

"Copy the NS RRs for the subzone into the authority section of the reply. Put whatever NS addresses are available into the additional section, using glue RRs if the addresses are not available from authoritative data or the cache. If all in-domain glue RRs do not fit, set TC=1 in the header. Go to step 4."

4. Security Considerations

This document clarifies correct DNS server behavior and does not introduce any changes or new security considerations.

5. Operational Considerations

At the time of this writing, the behavior of most DNS server implementations is to set the TC flag only if none of the available glue records fit in a response over UDP transport. The updated requirements in this document might lead to an increase in the fraction of UDP responses with the TC flag set, and consequently an increase in the number of queries to over TCP transport.

6. IANA Considerations

There are no actions for IANA.

7. Acknowledgements

The authors wish to thank Joe Abley, Brian Dickson, Geoff Huston, Jared Mauch, George Michaelson, Benno Overeinder, John R Levine, Shinta Sato, Puneet Sood, Ralf Weber, Tim Wicinski, Suzanne Woolf, and other members of the DNSOP working group for their input.

8. Changes

RFC Editor: Please remove this section before publication.

This section lists substantial changes to the document as it is being worked on.

From -01 to -02:

- *Clarified that "servers" means "authoritative servers".
- *Clarified that "available glue" means "all available glue".
- *Updated examples and placed before RFC 1034 update.

From -02 to -03:

- *Clarified scope to focus only on name server responses, and not zone/registry data.
- *Reorganized with section 2 as Types of Glue and section 3 as Requirements.
- *Removed any discussion of promoted / orphan glue.
- *Use appropriate documentation addresses and domain names.
- *Added Sibling Cyclic Glue example.

From -03 to -04:

- *Use "referral glue" on the assumption that other types of glue may be defined in the future.
- *Added Operational Considerations section.
- *Note many current implementations set TC=1 only when no glue RRs fit. New requirements may lead to more truncation and TCP.

*Sibling glue can be optional. Only require TC=1 when all in-domain glue RRs don't fit.

*Avoid talking about requirements for UDP/TCP specifically, and talk more generically about message size constraints regardless of transport.

9. Normative References

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

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <https://www.rfc-editor.org/info/rfc1035>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

10. Informative References

[CZDS] ICANN, "Centralized Zone Data Service", , January 2022, <https://czds.icann.org/>.

[FLAGDAY2020] Various DNS software and service providers, "DNS Flag Day 2020", , October 2020, <https://dnsflagday.net/2020/>.

[RFC2845] Vixie, P., Gudmundsson, O., Eastlake 3rd, D., and B. Wellington, "Secret Key Transaction Authentication for DNS (TSIG)", RFC 2845, DOI 10.17487/RFC2845, May 2000, <https://www.rfc-editor.org/info/rfc2845>.

[RFC2931] Eastlake 3rd, D., "DNS Request and Transaction Signatures (SIG(0)s)", RFC 2931, DOI 10.17487/RFC2931, September 2000, <https://www.rfc-editor.org/info/rfc2931>.

[RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, DOI 10.17487/RFC4033, March 2005, <https://www.rfc-editor.org/info/rfc4033>.

[RFC4034] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Resource Records for the DNS Security Extensions", RFC 4034, DOI 10.17487/RFC4034, March 2005, <https://www.rfc-editor.org/info/rfc4034>.

[RFC4035] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Protocol Modifications for the DNS Security Extensions", RFC 4035, DOI 10.17487/RFC4035, March 2005, <https://www.rfc-editor.org/info/rfc4035>.

[RFC6891] Damas, J., Graff, M., and P. Vixie, "Extension Mechanisms for DNS (EDNS(0))", STD 75, RFC 6891, DOI 10.17487/RFC6891, April 2013, <https://www.rfc-editor.org/info/rfc6891>.

Authors' Addresses

M. Andrews
ISC

Email: marka@isc.org

Shumon Huque
Salesforce

Email: shuque@gmail.com

Paul Wouters
Aiven

Email: paul.wouters@aiven.io

Duane Wessels
Verisign

Email: dwessels@verisign.com