DNSOP
Internet-Draft
Updates: 1034 (if approved)
Intended status: Standards Track
Expires: 27 January 2022

M. Andrews
ISC
S. Huque
Salesforce
P. Wouters
Aiven
D. Wessels
Verisign

26 July 2021

Glue In DNS Referral Responses Is Not Optional
draft-ietf-dnsop-glue-is-not-optional-02

Abstract

The DNS uses 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 glue records in referrals. If message size constraints prevent the inclusion of all glue records in a UDP response, the server MUST set the TC flag to inform the client that the response is incomplete, and that the client SHOULD use TCP to retrieve the full 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 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 27 January 2022.

Table of Contents

1. Introduction

The Domain Name System (DNS) [RFC1034], [RFC1035] uses glue records to allow iterative clients to find the addresses of nameservers that are contained within a delegated zone. 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 glue records in referrals. If message size constraints prevent the inclusion of all glue records in a UDP response, the server MUST set the TC (Truncated) flag to inform the client that the response is incomplete, and that the client SHOULD use TCP to retrieve the full response. This document clarifies that expectation.

DNS responses sometimes contain optional data in the additional section. 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].

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. Glue record example

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       192.0.1.1
ns2.foo.test.            86400   IN A       192.0.1.2
```

Referral responses from "test" for "foo.test" must include the glue records in the additional section (and set TC=1 if they do not fit):

```
;; 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       192.0.1.1
ns2.foo.test.           86400        IN      A       192.0.1.2
```

2.1. Missing glue

While not common, real life examples of servers that fail to set TC=1

when glue records are available, exist and they do cause resolution failures.

The example below from June 2020 shows a case where none of the glue records, present in the zone, fitted into the available space and TC=1 was not set in the response. While this example shows an DNSSEC [RFC4033], [RFC4034], [RFC4035] referral response, this behaviour has also been seen with plain DNS responses as well. The records have been truncated for display purposes. Note that at the time of this writing, this configuration has been corrected and the response correctly sets the TC=1 flag.


```
% dig +norec +dnssec +bufsize=512 +ignore @a.gov-servers.net \
       rh202ns2.355.dhhs.gov

; <<>> DiG 9.15.4 <<>> +norec +dnssec +bufsize +ignore \
       @a.gov-servers.net rh202ns2.355.dhhs.gov
; (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.dhhs.gov.          IN A

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

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 addresses are available into the additional section, using glue RRs if the addresses are not available from authoritative data or the cache. If all glue RRs do not fit, set TC=1 in the header. Go to step 4."

4. Sibling Glue

Sibling glue are glue records that are not contained in the delegated zone itself, but in another delegated zone from the same parent. In many cases, these are not strictly required for resolution, since the resolver can make follow-on queries to the same 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.

This document clarifies that sibling glue (being part of all available glue records) MUST be returned in referral responses, and that the requirement to set TC=1 applies to sibling glue that cannot fit in the response too.

4.1. Sibling Glue example

Here the delegating zone "test" contains 2 sub-delegations for the subzones "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       192.0.1.1
ns2.bar.test.               86400   IN A       192.0.1.2

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

Referral responses from "test" for "foo.test" must include the sibling glue (and set TC=1 if they do not fit):

```
;; 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       192.0.1.1
ns2.bar.test.           86400        IN      A       192.0.1.2
```


5. Promoted (or orphaned) glue

When a zone is deleted but the parent notices that its NS glue records are required for other zones, it MAY opt to take these (now orphaned) glue records into its own zone to ensure that other zones depending on this glue are not broken. Technically, these address records are no longer glue records, but authoritative data of the parent zone, and should be added to the DNS response similarly to regular glue records.

6. Security Considerations

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

7. IANA Considerations

There are no actions for IANA.

8. 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>.

9. Informative References

[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