Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: September 10, 2015

P. Koch
DENIC eG
M. Larson
Dyn, Inc.

March 9, 2015

Initializing a DNS Resolver with Priming Queries
draft-ietf-dnsop-resolver-priming-05

Abstract

This document describes the initial queries a DNS resolver is supposed to emit in order to initialize its cache with both a current NS RRSet for the root zone and the necessary address information.

1. Introduction

Domain Name System (DNS) resolvers need a starting point to resolve queries. [RFC1034], section 5.3.2, defines the SBELT structure in a full resolver as:

> a "safety belt" structure of the same form as SLIST, which is initialized from a configuration file, and lists servers which should be used when the resolver doesn't have any local information to guide name server selection. The match count will be -1 to indicate that no labels are known to match.

Section 5.3.3 of [RFC1034] adds

> the usual choice is two of the root servers and two of the servers for the host's domain

Today's practice generally seperates serving and resolving functionalities, so the servers "for the host's domain" might no longer be an appropriate choice, even if they were only intended to resolve "local" names, especially since the SBELT structure does not distinguish between local and global information. In addition, DNS server implementations have for a long time been seeded with not only two but an exhaustive list of the root servers' addresses. This list is either supplied as a configuration file (root "hints", an excerpt of the DNS root zone) or even compiled into the software.

The list of root name servers has been rather stable over the last fifteen years. After the last four servers had been added and moved to their final (network) destinations in 1997, there have been only five address changes affecting the L (twice), J, B, and D servers. Research is available for B [Mann2006] and J [BLKT2004], which shows that several months or even years after the change had become effective, traffic is still received on the old addresses. Therefore, it is important that resolvers be able to cope with change, even without relying upon configuration updates to be applied by their operator.

Work by the ICANN SSAC and RSSAC committees, [SSAC016] and [SSAC017], aiming at adding AAAA RRs for the root name servers' names, deals with priming queries and so does a draft on DNSSEC Trust Anchor maintenance [I-D.ietf-dnsop-dnssec-trust-anchor]. However, it turned out that despite having been practiced for a long time, priming queries have not yet been documented as an important resolver feature.

The following sections cover parameters of both the priming query and the response to be sent by a root name server.

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. Priming Queries

This document only deals with recursive name servers (recursive resolvers, resolvers) for the IN CLASS.

2.1. Parameters of a Priming Query

A priming query SHOULD use a QNAME of "." and a QTYPE of NS. The priming query MUST be sent over UDP (section 6.1.3.2 of [RFC1123]). The UDP source port SHOULD be randomly selected [RFC5452]. The RD bit MUST NOT be set.

The resolver SHOULD also use EDNS0 [RFC6891] and SHOULD announce and handle a reassembly size of at least 1024 octets [RFC3226]. This is to cover the size of a full priming response (see Section 3.3).

2.2. Repeating Priming Queries

A resolver SHOULD NOT originate a priming query more often than once per day (or whenever the resolver starts). It SHOULD adhere to the TTL values given in the priming response. To avoid amnesia, the resolver MAY proactively re-prime before the old root NS RRSet expires from the cache, but only after 75 percent of the NS RRSet's TTL (or of the A/AAAA RRSets' TTL, whichever is lower) have passed.

Should the priming query time out, the resolver SHOULD retry with a different target address.

2.3. Target Selection

A resolver MUST select the target for a priming query randomly from the list of addresses (IPv4 and IPv6) available in its SBELT structure and it MUST ensure that all targets are selected with equal probability even upon startup. For resending the priming query to a different server the random selection SHOULD also be used.

2.4. DNSSEC with Priming Queries

The resolver SHOULD NOT set the DNSSEC OK [RFC4033] bit.

Discussion: Delegations in referral responses are not signed, so consequently the priming response is not validated, either. For that to work, the priming response would also have to be self-contained in that it would allow the resolver to not only validate the NS RRSet

(with the root DNSKEY RRSet and the root NS RRSet's signature), but also the A and AAAA RRSets. All this information cannot be guaranteed to be either present at the root name servers or fit into the priming reponse even with the largest feasible EDNS0 buffer size. In fact, in today's Internet, with the root name servers' names under "ROOT-SERVERS.NET.", this isn't even true for the top level domain involved. So, even though a poisoned priming response could drastically influence the resolver's operations, there is little a DNSSEC enhanced priming response could achieve without the whole validation chain. This would probably call for a different naming scheme (see section 6.1 of [I-D.koch-dns-glue-clarifications]).

3. Priming Responses

A root name server cannot distinguish a priming query from any other query for the root NS RRSet, except that QTYPE NS would not usually be part of the DNS resolution process.

3.1. Expected Properties of the Priming Response

The priming response can be expected to have an RCODE of NOERROR and the AA bit set. Also, there should be an NS RRSet in the answer section (since the NS RRSet originates from the root zone), an empty authority section (since the NS RRSet already appears in the answer section) and an additional section with A and AAAA RRSets for the root name servers pointed at by the NS RRSet. Resolver software SHOULD NOT expect a fixed number of 13 NS RRs, since "internal" root server setups in split DNS configurations might use a different number of servers. Resolver software SHOULD warn the operator about any change in the number or names of name servers or their addresses compared to the SBELT information.

3.2. Use of the Priming Response

A resolver MAY use the priming response as it would use any other data fed to its cache. However, it SHOULD NOT use the SBELT information directly in any responses it hands out.

3.3. Completeness of the Response

Assuming an upper bound of thirteen root name servers and one address each for IPv4 and IPv6, the combined size of all the A and AAAA RRSets is 13 * (16 + 28) == 572, independent of the naming scheme. Not even counting the NS RRSet, this value exceeds the original 512 octet payload limit.

For an EDNS response, a resolver SHOULD consider the address information found in the additional section complete for any

particular server that appears at all. In other words: if the additional section only has an A RRSet for a server, the resolver SHOULD assume that no AAAA RRSet exists. This is to avoid repeated unnecessary queries for names of name servers that do not or do not yet offer IPv6 service, or, in perspective, will have ceased IPv4 service.

If the resolver did not announce a reassembly size larger than 512 octets, this assumption is invalid. Simple re-issuing of the priming query does not help with those root name servers that respond with a fixed order of addresses in the additional section. Instead the resolver ought to issue direct queries for A and AAAA RRSets for the remaining names. In today's environment these RRSets would be authoritatively available from the root name servers.

4. Requirements for Root Name Servers and the Root Zone

The operational requirements for root name servers are described in [RFC2870]. This section specifies additional guidance for the configuration of and software deployed at the root name servers.

All DNS root name servers need to be able to provide for all addresses of all root name servers. This can easily achieved by keeping all root name server names in a single zone and by making all root name servers authoritative for that zone.

If the response packet does not provide for more than 512 octets due to lack of EDNS0 support, A RRSets SHOULD be given preference over AAAA RRSets when filling the additional section.

[[Issue 1: EDNS0 is used as an indication of AAAA understanding on the side of the client. What to do with payload sizes indicated by EDNS0 that are smaller than 1024, is open to discussion. At the time of writing, some root name servers will fill the additional section with all available A RRSets, only adding some AAAA RRSets, when queried over IPv4 without EDNS0. Other servers will deliver more AAAA RRSets, therefore withholding some A RRSets completely [RFC4472].]]

To ensure equal availability the A and AAAA RRSets for the root name servers' names SHOULD have identical TTL values at the authoritative source.

[[Issue 2: Do the TTLs for the root NS RRSet and address RRSets in the root and the ROOT-SERVERS.NET. zones need to be aligned? In real life responses, the address RRSet's TTL values vary by name server implementation. Is this diversity we can live with? Should the

authoritative source prevail? Is it therefore a protocol issue rather than an operational choice of parameters?]]

5. Security Considerations

This document deals with priming a DNS resolver's cache. The usual DNS caveats apply. Use of DNSSEC with priming queries is discussed in section 2.4.

Spoofing a response to a priming query can be used to redirect all queries originating from a victim resolver. Therefore, any difference between the inital SBELT list and the priming response SHOULD be brought to the operators' attention. There is also a chance that the random target selection chooses the address of a retired root name server. Operational measures to prevent reuse of these addresses are out of the scope of this document.

6. IANA Considerations

This document does not propose any new IANA registry nor does it ask for any allocation from an existing IANA registry.

However, this document deals with requirements for the root zone and root server operations.

[[Issue3: related to issue 2 - any recommendation on the "." NS RRSet TTL or the TTLs of the respective A and/or AAAA RRSets might go here.]]

7. References

7.1. Normative References

[RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, November 1987.

[RFC1123] Braden, R., "Requirements for Internet Hosts - Application and Support", STD 3, RFC 1123, October 1989.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3226] Gudmundsson, O., "DNSSEC and IPv6 A6 aware server/resolver message size requirements", RFC 3226, December 2001.

[RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, March 2005.

[RFC5452] Hubert, A. and R. van Mook, "Measures for Making DNS More Resilient against Forged Answers", RFC 5452, January 2009.

[RFC6891] Damas, J., Graff, M., and P. Vixie, "Extension Mechanisms for DNS (EDNS(0))", STD 75, RFC 6891, April 2013.

7.2. Informative References

[BLKT2004]
Barber, P., Larson, M., Kosters, M., and P. Toscano, "Life and Times of J-Root", NANOG 32, October 2004.

[I-D.ietf-dnsop-dnssec-trust-anchor]
Larson, M. and O. Gudmundsson, "DNSSEC Trust Anchor Configuration and Maintenance", draft-ietf-dnsop-dnssec-trust-anchor-04 (work in progress), October 2010.

[I-D.koch-dns-glue-clarifications]
Koch, P., "DNS Glue RR Survey and Terminology Clarification", draft-koch-dns-glue-clarifications-04 (work in progress), July 2010.

[Mann2006]
Manning, B., "persistent queries and phantom nameservers", WIDE/CAIDA Workshop , October 2006.

[RFC2870] Bush, R., Karrenberg, D., Kosters, M., and R. Plzak, "Root Name Server Operational Requirements", BCP 40, RFC 2870, June 2000.

[RFC4472] Durand, A., Ihren, J., and P. Savola, "Operational Considerations and Issues with IPv6 DNS", RFC 4472, April 2006.

[RFC4697] Larson, M. and P. Barber, "Observed DNS Resolution Misbehavior", BCP 123, RFC 4697, October 2006.

[SSAC016] ICANN Security and Stability Advisory Committee, "Testing Firewalls for IPv6 and EDNS0 Support", SSAC 016, January 2007.

[SSAC017] ICANN Security and Stability Advisory Committee, "Testing Recursive Name Servers for IPv6 and EDNS0 Support", SSAC 017, February 2007.

Appendix A. Document Revision History

This section is to be removed should the draft be published.

$Id: draft-ietf-dnsop-resolver-priming.xml,v 1.8 2015/03/09 20:22:52 pk Exp $

A.1. -05 WG Document

Revived. Minor edits for readability - Warren.

A.2. -04 WG Document

Revived. Updated EDNS0 reference. Minor edits for clarity.

A.3. -03 WG Document

Revived. Resolved most open issues [[]] as per previous WG discussion. Minor edits on history and wording. All root servers authoritative for ROOT-SERVERS.NET.

A.4. -02 WG Document

Revived. Changed use of DNSSEC OK in the priming query as per the WG discussion.

A.5. -01 WG Document

Revived with minor edits. Open issues marked [[]].

A.6. -00 WG Document

Reposted as WG document with minor edits.

Added re-primimg proposal and A/AAAA TTL considerations.

A.7. Initial Document

First draft

Authors' Addresses

Peter Koch
DENIC eG
Kaiserstrasse 75-77
Frankfurt 60329
DE

Phone: +49 69 27235 0
Email: pk@DENIC.DE

Matt Larson
Dyn, Inc.
150 Dow St
Manchester, NH 03101
USA

Email: mlarson@dyn.com