Network Working Group
Internet-Draft
Updates: 1034, 2181 (if approved)
Intended status: Standards Track
Expires: May 5, 2017

K. Fujiwara
JPRS

November 01, 2016

Updating Resolver Algorithm
draft-fujiwara-dnsop-resolver-update-00

Abstract

Parent side NS RRSet and glue records are all information to access servers for child zone. However, they may be overwritten by child zone data (zone apex NS RRSet and other A/AAAA RRSets). The overwrite makes name resolution unstable and induces vulnerabilities. RFC 2181 section 5.4.1 specifies trustworthiness of DNS data. And it is deemed that that all cached data (authoritative data, non-authoritative data, referrals and glue records) are merged into one. Resolvers may answer non-authoritative data, referrals and glue records that should not be returned. This document proposes updating resolver algorithm that separates the cache to "authoritative data cache" and "delegation cache". The former is used to answer stub resolvers, and the latter is used to iterate zones.

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 May 5, 2017.

Copyright Notice

Table of Contents

1. Introduction

Resolver algorithm is defined in [RFC1034] and updated by [RFC2181]. The resolver algorithm seems to assume single cache that holds all RRSets from received responses. [RFC2181] Section 5.4.1 Ranking data specifies the trustworthiness order of RRSets. When a resolver receives higher trustworthy data, the cached data is replaced by the received data.

Parent side NS RRSet is very important because it creates new zone and specifies how to access name servers for the created zone described in Section 3.1. However, parent side NS RRSets and glue records have least trustworthiness. The parent side NS RRSets and

glue records are replaced by authoritative data if resolvers receive authoritative data described in Section 3.3.

The overwrite makes name resolution unstable and some vulnerabilities described in Section 4. And it may break requirements of resolvers' answers described in Section 3.2.

This document proposes updated resolver algorithm that separate authoritative data cache that is answered to stub resolvers and delegation cache that is used to iterate zones. Details are described in Section 5

2. Terminology

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 RFC 2119 [RFC2119].

Many of the specialized terms used in this specification are defined in DNS Terminology [RFC7719].

3. Traditional resolver algorithm

[RFC1034] defines "zone cut", "delegation", "referral", "glue records", "authoritative", and "resolver algorithm".

[RFC2181] clarified the resolver algorithm defined in [RFC1034].

3.1. Importance of parent side NS RRSet

[RFC1034], [RFC2181] and [RFC7719] defines zone "cut", "delegation", "referral", and parent side NS RRSet functions as follows.

"'cuts' in the name space can be made between any two adjacent nodes. After all cuts are made, each group of connected name space is a separate zone. The zone is said to be authoritative for all names in

the connected region." (Quoted from RFC 1034, Section 4.2)

"The RRs that describe cuts around the bottom of the zone are NS RRs that name the servers for the subzones. Since the cuts are between nodes, these RRs are NOT part of the authoritative data of the zone, and should be exactly the same as the corresponding RRs in the top node of the subzone." (Quoted from RFC 1034, section 4.2.1)

"That is, parent zones have all the information needed to access servers for their children zones." (Quoted from RFC 1034, section 4.2.1)

"Resolvers must be able to access at least one name server and use that name server's information to answer a query directly, or pursue the query using referrals to other name servers." (Quoted from RFC 1034, Section 2.4)

"Delegation is the process by which a separate zone is created in the name space beneath the apex of a given domain. Delegation happens when an NS RRset is added in the parent zone for the child origin." (Quoted from RFC 7719)

"This situation typically occurs when the glue address RRs have a smaller TTL than the NS RRs marking delegation," (Quoted from RFC 1035, Section 7.2)

"The existence of a zone cut is indicated in the parent zone by the existence of NS records specifying the origin of the child zone." (Quoted from RFC 2181, Section 6)

As described above, parent side NS RRSet makes a new zone. Parent side NS RRSet (referral) and glue records are all the information to access servers for the child zone.

That is, resolvers SHOULD NOT use child side NS RRSet to iterate zones.

3.2. Recommendation of resolver's answer

RFC 1034 describes resolver's answer as follows.

"The ideal answer is one from a server authoritative for the query which either gives the required data or a name error. The data is passed back to the user and entered in the cache for future use if its TTL is greater than zero." (Quoted from RFC 1034, Section 5.3.3)

"The simplest mode for the client is recursive, since in this mode the name server acts in the role of a resolver and returns either an error or the answer, but never referrals." (Quoted from RFC 1034, Section 4.3.1)

Recently, most of full-service resolver implementations answer only authoritative data to stub resolvers.

As described above, recommendation of resolver's answer is "answer only authoritative data." It does not break existing standards.

3.3. Traditional resolver algorithm

Resolver algorithm is defined in [RFC1034] Section 5.3.3. Resolvers cache all RRSets during the iterations in their cache. When resolvers receive new data, they will update their cache. The update is explained using an example resolution of "www.example.com/A" in "example.com" zone as follows.

When a resolver iterates "www.example.com/A" query, then one of root servers responds "com" NS RRSet (referral) with glue records, one of "com" servers responds "example.com" NS RRSet with glue records, and one of "example.com" servers responds "www.example.com" A RRSet.

As a result, the resolver caches all RRSets during the iterations in its cache.

After then, a resolver receives "example.com/NS" query, it retrieves "example.com" NS authoritative data defined in zone apex of "example.com" and it overwrites "example.com/NS" in the resolver's cache as "trustworthiness" rule of [RFC2181].

If the parent side "example.com" NS RRSet and the child side

"example.com" NS RRSet are different, next resolution result will be changed because the resolver will send "www.example.com/A" query to name servers that are specified by "example.com" NS RRSet defined in zone apex. Glue records in the cache are also overwritten by authoritative data, and then, IP addresses of name servers that the resolver send to will be changed.

The other case, if one of "example.com" name servers responds "www.example.com" A RRSet with "example.com" NS RRSet in authority section (several existing authoritative server implementations perform this) , "example.com" NS RRSet from "com" TLD servers (referral) is overwritten by "example.com" NS data attached in the authoritative answer from child zone.

The overwrite is specified by [RFC2181] Section 5.4.1 Ranking data. "Referrals" is the ranking 7: "Data from the authority section of a non-authoritative answer". And "example.com" NS RRSet attached in the authoritative answer is the ranking 4: "Data from the authority section of an authoritative answer".

4. Problem Statement

[RFC1034] section 4.2.1 states that "the parent side NS RRSet should be exactly the same as the corresponding RRs in the top node of the subzone".

However, people sets different NS RRSets with mistakes, or intentionally. Name server configuration changes will make the differences because the changes take time.

If the zone data of name server(s) specified by referrals and specified by zone apex NS RRSet are different, name resolution becomes unstable. The cache overwrite of NS RRSet may break "Referrals and glue records are information to access servers for child zones" specified by [RFC1034] section 4.2.1.

The overwrite by zone apex NS RRSet induced security vulnerabilities. In 2012, "Ghost Domain Names: Revoked Yet Still Resolvable" [DUAN2012GHOST] was reported. The attack uses updates of NS RRSet attached in authoritative answer. Assume a resolver caches and uses zone apex NS RRSet, and the parent side NS RRSet is updated or

removed. The resolver send queries to name servers specified by zone apex NS RRSet and update NS RRSet by the NS RRSet attached in the authority section of the answer. Parent side NS RRSet specifies the existence of delegation, however, resolvers may not check the existence of the parent side NS RRSet and the domain name will remain in the resolvers.

DNS software vendors fixed the problem to restrict the TTL of NS RRSet not to exceed the cached TTL value of old NS RRSet when replacing it.

5. Updating of Resolver Algorithm

5.1. Recommendations to resolvers

Resolvers MUST answer one of the following results: required data, name error, or empty (NODATA) from a server that is authoritative for the query, other name resolution errors (SERVFAIL, REFUSED), or no answer.

Resolvers iterate queries using referrals with corresponding glue records to other name servers. If referrals contain out-of-bailiwick name server names, resolvers need to resolve address records of out-of-bailiwick name servers.

Resolvers MUST NOT use glue records and referrals except iterating delegations. Resolvers MUST NOT use zone apex NS RRSets to iterate.

5.2. Update to resolver algorithm

This document update RFC 1034 Section 5.3.3. Algorithm as follows.

Separate the cache into "authoritative data cache" and "delegation cache". Pre-load root hint information (root NS RRSet and root server addresses) into the delegation cache.

"Step 4.a." is changed as "if the response answers the question or contains a name error, cache the data into authoritative cache as well as returning it back to the client".

"Step 4.b." is changed as if the response contains a better delegation to other servers, cache the delegation information into delegation cache, and go to step 2".

The cache in "Step 1" is the authoritative data cache.

The cache used in "step 2" is the delegation cache. "Set up their addresses using local data" is replaced as "Set up their addresses using the delegation cache".

As a result, RFC 2181 Section 5.4.1 Ranking data becomes useless because the overwrite will not happen. Pre-loaded zone files (or zones retrieved from zone transfer) are treated as answers from authoritative servers. They are treated as static authoritative data, referrals, and glue records. Referrals and glue records in pre-loaded zone files MUST NOT be answered to stub resolvers. They MUST be used to iterate name servers only.

Root zone is special because it is not delegated. Root hint and priming are exceptions because priming replaces pre-configured root hint by root zone apex NS RRSet (authoritative data).

5.3. Characteristics of the update

This update does not change resolver algorithm described in RFC 1034 section 5.3.3, except updates of referrals. It separates authoritative data (possible to answer) and referrals (used to iterate DNS tree). It does not require no special ordering (e.g. trustworthiness and ranking data). It offers more stability of name resolution because the results of traditional name resolution will flap if NS RRSets between the parent and the child are different.

This algorithm is similar to traditional algorithm when the cache is empty.

The update does not effect to DNSSEC [RFC4033] [RFC4034] [RFC4035] because DNSSEC validates authoritative data and does not validate referrals.

The update does not effect to DNS Query Name Minimisation [RFC7816]

because answers from authoritative servers don't change. Delegation cache and authoritative data cache separation will need small implementation changes.

5.4. Issues of the update

This update makes impossible to control of TTL value of NS RRSet by the child zone owner. However, overwrite of the referral does not occur always and TTL control may increase queries to authoritative servers.

6. Implementation Ideas

Some implementers already implemented similar algorithm to their products.

7. IANA Considerations

{#:ianacons}

This document has no IANA actions.

8. Security Considerations

9. Acknowledgments

10. Change History

11. References

11.1. Normative References

[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>.

[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, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc4035>.

[RFC7719] Hoffman, P., Sullivan, A., and K. Fujiwara, "DNS Terminology", RFC 7719, DOI 10.17487/RFC7719, December 2015, <http://www.rfc-editor.org/info/rfc7719>.

11.2. Informative References

[DUAN2012GHOST] D, H., Jiang, J., Liang, J., Li, K., Li, J., and J. Wu, "Ghost domain names:Revoked yet still resolvable", 2012, <The 19th Annual Network & Distributed System Security Symposium (NDSS 2012)>.

[RFC7816] Bortzmeyer, S., "DNS Query Name Minimisation to Improve Privacy", RFC 7816, DOI 10.17487/RFC7816, March 2016, <http://www.rfc-editor.org/info/rfc7816>.

Author's Address

Kazunori Fujiwara
Japan Registry Services Co., Ltd.
Chiyoda First Bldg. East 13F, 3-8-1 Nishi-Kanda
Chiyoda-ku, Tokyo 101-0065
Japan

Phone: +81 3 5215 8451
Email: fujiwara@jprs.co.jp