dnsop
Internet-Draft
Updates: 1035 (if approved)
Intended status: Standards Track
Expires: 26 July 2024

T. April
P. paek
ISC
R. Weber
Akamai Technologies
D. Lawrence
Salesforce

23 January 2024

Extensible Delegation for DNS
draft-dnsop-deleg-00

Abstract

A delegation in the Domain Name System (DNS) is a mechanism that enables efficient and distributed management of the DNS namespace. It involves delegating authority over subdomains to specific DNS servers via NS records, allowing for a hierarchical structure and distributing the responsibility for maintaining DNS records.

An NS record contains the hostname of the nameserver for the delegated namespace. Any facilities of that nameserver must be discovered through other mechanisms. This document proposes a new extensible DNS record type, DELEG, which contains additional information about the delegated namespace and the capabilities of authoritative nameservers for the delegated namespace.

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 26 July 2024.

Copyright Notice

Copyright (c) 2024 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

In the Domain Name System [STD13], subdomains within the domain name hierarchy are indicated by delegations to servers which are authoritative for their portion of the namespace. The DNS records that do this, called NS records, contain hostnames of nameservers, which resolve to addresses. No other information is available to the resolver. It is limited to connect to the authoritative servers over UDP and TCP port 53.

This limitation is a barrier for efficient introduction of new DNS technology. New features come with additional overhead as they are constrained by the intersection of resolver and nameserver functionality. New functionality could be discovered insecurely by trial and error, or negotiated after first connection, which is costly and unsafe.

The proposed DELEG record type remedies this problem by providing extensible parameters to indicate capabilities that a resolver may use for the delegated authority, for example that it should be contacted using a transport mechanism other than DNS over UDP or TCP on port 53.

DELEG records are served with NS and DS records in the Authority section of DNS delegation type responses. Standard behavior of legacy DNS resolvers is to ignore the DELEG type and continue to rely on NS and DS records (see compliance testing described in Appendix A). Resolvers that do understand DELEG and its associated parameters can efficiently switch to the new mechanism.

The DELEG record leverages the Service Binding (SVCB) record format defined in [RFC9460], using a subset of the already defined service parameters.

DELEG can use AliasMode, inherited from SVCB, to insert a level of indirection to ease the operational maintenance of multiple zones by the same servers. For example, an operator can have numerous customer domains all aliased to nameserver sets whose operational characteristics can be easily updated without intervention from the customers. Most notably, this provides a method for addressing the long-standing problem operators have with maintaining DS records on behalf of their customers. This type of facility will be handled in separate documents.

1.1. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Terminology regarding the Domain Name System comes from [BCP219], with addition terms defined here:

* legacy name servers: An authoritative server that does not support the DELEG record.

* legacy resolvers: A resolver that does not support the DELEG record.

1.2. Motivation for DELEG

* There is no secure way to signal capabilities or new features of an authoritative server, such as authenticated DNS-over-TLS. A resolver must resort to trial-and-error methods that can potentially fall victim to downgrade attacks.

* Delegation point NS records and glue address records are, by design, not DNSSEC signed. This presents a leap of faith. Spoofed delegation point NS records can be detected eventually if the delegated domain was signed, but only after traffic was sent to the (potentially) spoofed endpoint.

* The Registry, Registrar, Registrant (RRR) model has no formally defined role for DNS operators. Consequently, registrants are the channel between DNS operators and registries/registrars on purely operational elements, such as adding NS records, modify DS records when rolling keys, etc. Deleg’s AliasMode allows the registrants

to delegate these facilities to a DNS Operator.

1.3. Introductory Examples

To introduce DELEG record, this example shows the authority section of a DNS response that delegates a subdomain to another nameserver.

```
example.com.  86400  IN DELEG  1 ns1.example.com. (
                ipv4hint=192.0.2.1 ipv6hint=2001:DB8::1 )
example.com.  86400  IN NS     ns1.example.com.
ns1.example.com.    86400   IN  A  192.0.2.1
ns1.example.com     86400   IN  AAAA    2001:DB8::1
```

In this example, the authoritative nameserver is delegating using the same parameters as regular DNS, but the delegation as well as the glue can be signed.

Like in SVCB, DELEG also offer the ability to use the Alias form of delegation. The example below shows an example where example.com is being delegated with a DELEG AliasMode record which can then be further resolved using standard SVCB to locate the actual parameters.

```
example.com.  86400  IN DELEG 0   config2.example.net.
example.com.  86400  IN NS     ns2.example.net.
```

The example.net authoritative server may return the following SVCB records in response to a query as directed by the above records.

```
config2.example.net 3600    IN SVCB . (
                ipv4hint=192.0.2.54,192.0.2.56
                ipv6hint=2001:db8:2423::3,2001:db8:2423::4 )
```

The above records indicate to the client that the actual configuration for the example.com zone can be found at config2.example.net

Later sections of this document will go into more detail on the resolution process using these records.

1.4. Goal of the DELEG record

The primary goal of the DELEG records is to provide zone owners a method to signal capabilities to clients how to connect and validate a subdomain. This method coexists with NS records in the same zone.

The DELEG record is authoritative in the parent zone and, if signed, has to be signed with the key of the parent zone. The target of an alias record is an SVCB record that exists and can be signed in the zone it is pointed at, including the child zone.

1.5. DNSSEC is RECOMMENDED

While DNSSEC is RECOMMENDED, unsigned DELEG records may be retrieved in a secure way from trusted, Privacy-enabling DNS servers using encrypted transports.

FOR DISCUSSION: This will lead to cyclical dependencies. A DELEG record can introduce a secure way to communicate with trusted, Privacy-enabling DNS servers. For that, it needs to be DNSSEC signed.

1.5.1. Preventing downgrade attacks

A flag in the DNSKEY record is used as a backwards compatible, secure signal to indicate to a resolver that DELEG records are present or that there is an authenticated denial of a DELEG record. Legacy resolvers will ignore this flag and use the DNSKEY as is.

Without this secure signal an on-path adversary can remove DELEG records and its RRsig from a response and effectively downgrade this to a legacy DNSSEC signed response.

1.6. Facilities

The DELEG record is extensible in such a way that future innovations in the domain name system, such as new methods of secure transport, message encoding, error reporting, etc, does not depend on a re-design of the DNS.

2. DELEG Record Type

The SVCB record allows for two types of records, the AliasMode and the ServiceMode. The DELEG record takes advantage of both and each will be described below in depth. The wire format of and the registry for the DELEG record is the same as SVCB record defined in [RFC9460]

2.1. Difference between the records

This document uses two different resource record types. Both records have the same functionality, with the difference between them being that the DELEG record MUST only be used at a delegation point, while the SVCB is used as a normal resource record and does not indicate that the label is being delegated. For example, take the following DELEG record:

Zone com.:
example.com. 86400 IN DELEG 1 config2.example.net.

When a client receives the above record, the resolver should send queries for any name under example.com to the nameserver at config2.example.net unless further delegated. By contrast, when presented with the records below:

Zone com.:
example.com. 86400 IN DELEG 0 config3.example.org.

Zone example.org.:
config3.example.org. 86400 IN SVCB 1 . (ipv4hint=192.0.2.54,192.0.2.56
ipv6hint=2001:db8:2423::3,2001:db8:2423::4)

A resolver trying to resolve a name under example.com would get the first record above from the parent authoritative server, .COM, indicating that the SVCB records found at config3.example.org should be used to locate the authoritative nameservers of example.com, and other parameters.

The primary difference between the two records is that the DELEG record means that anything under the record label should be queried at the delegated server while the SVCB record is just for redirection purposes, and any names under the record's label should still be resolved using the same server unless otherwise delegated.

It should be noted that both DELEG and SVCB records may exist for the

same label, but they will be in different zones. Below is an example of this:

```
Zone com.:
example.com.  86400  IN DELEG 0   c1.example.org.

Zone example.org.:
c1.example.org.  86400  IN DELEG  1   config3.example.net. (
                            ipv6hint=2001:db8:2423::3 )
c1.example.org.  86400  IN NS test.c1.example.org.
test.c1.example.org. 600 IN A 192.0.2.1

Zone c1.example.org:
c1.example.org.  86400  IN SVCB 1   config2.example.net. (
                     ipv6hint=2001:db8:4567::4  )
c1.example.org.  86400  IN NS test.c1.example.org.
test.c1.example.org. 600 IN A 192.0.2.1
```

In the above case, the DELEG record for c1.example.org would only be used when trying to resolve names at or below c1.example.org. This is why when an AliasMode DELEG or SVCB record is encountered, the resolver MUST query for the SVCB record associated with the given name.

2.2. AliasMode Record Type

In order to take full advantage of the AliasMode of DELEG and SVCB, the parent, child, and resolver must support these records. When supported, the use of the AliasMode will allow zone owners to delegate their zones to another operator with a single record in the parent. If a resolver were to encounter an AliasMode DELEG or SVCB record, it would then resolve the name in the TargetName of the original record using SVCB RR type to receive either another AliasMode record or a ServiceMode SVCB record.

For example, if the name www.example.com was being resolved, the .com zone may issue a referral by returning the following record:

```
example.com.    86400    IN  DELEG     0   config1.example.net.
```

The above record would indicate to the resolver that in order to obtain the authoritative nameserver records for example.com, the resolver should resolve the RR type SVCB for the name config1.example.net.

2.2.1. Multiple Service Providers

Some zone owners may wish to use multiple providers to serve their zone, in which case multiple DELEG AliasMode records can be used. In the event that multiple DELEG AliasMode records are encountered, the resolver SHOULD treat those as a union the same way this is done with NS records, picking one at random for the first lookup and eventually discovering the others. How exactly DNS questions are directed and split between configuration sets is implementation specific:

```
example.com.    86400    IN  DELEG     0   config1.example.net.
example.com.    86400    IN  DELEG     0   config1.example.org.
```

[DRAFT NOTE: SVCB says that there "SHOULD only have a single RR". This ignores that but keeps the randomization part. Section 2.4.2 of SVCB]

2.2.2. Loop Prevention

The TargetName of an SVCB or DELEG record MAY be the owner of a CNAME record. Resolvers MUST follow CNAMEs as well as further alias SVCB records as normal, but MUST not allow more then 4 total lookups per delegation, with the first one being the DELEG referral and then 3 SVCB/CNAME lookups maximal.

Special care should be taken by both the zone owner and the delegated zone operator to ensure that a lookup loop is not created by having two AliasMode records rely on each other to serve the zone. Doing so may result in a resolution loop, and likely a denial of service. The mechanism on following CNAME and SVCB alias above should prevent exhaustion of server resources. If a resolution can not be found after 4 lookups the server should reply with a SERVFAIL error code.

2.3. Deployment Considerations

The DELEG and SVCB records are intended to replace the NS record while also adding additional functionality in order to support additional transports for the DNS. Below are discussions of considerations for deployment.

2.3.1. AliasMode and ServiceMode in the Parent

Both the AliasMode and ServiceMode records can be returned for the DELEG record from the parent. This is different from the SCVB [RFC9460] specification and only applies for the DELEG RRSet in the parent.

2.3.2. Rollout

When introduced, the DELEG and SVCB records might not initially be supported by the DNS root or TLD operators. Zone owners may place these records into their zones before the zones above them have done so. However, doing so is only useful for further delegations down the tree as an SVCB record at the zone apex alone does not indicate a new delegation type. The only way to discover new delegations is with the DELEG record at the parent.

2.3.3. Availability

If a zone operator removes all NS records before DELEG and SVCB records are implemented by all clients, the availability of their zones will be impacted for the clients that are using non-supporting resolvers. In some cases, this may be a desired quality, but should be noted by zone owners and operators.

2.4. Response Size Considerations

For latency-conscious zones, the overall packet size of the delegation records from a parent zone to child zone should be taken into account when configuring the NS, DELEG and SVCB records. Resolvers that wish to receive DELEG and SVCB records in response SHOULD advertise and support a buffer size that is as large as possible, to allow the authoritative server to respond without truncating whenever possible.

3. Implementation

This document introduces the concept of signaling capabilities to clients on how to connect and validate a subdomain. This section

details the implementation specifics of the DELEG record for various DNS components.

3.1. Including DELEG RRs in a Zone

A DELEG RRset MAY be present at a delegation point. The DELEG RRset MAY contain multiple records. DELEG RRsets MUST NOT appear at a zone's apex.

A DELEG RRset MAY be present with or without NS or DS RRsets at the delegation point.

Construction of a DELEG RR requires knowledge which implies communication between the operators of the child and parent zones. This communication is an operational matter not covered by this document.

3.1.1. Signing DELEG RRs

A DELEG RRset MUST be DNSSEC signed if the zone is signed.

If a signed zone contains DELEG records, the zone MUST be signed with a DNSKEY that has the DELEG flag set.

3.2. Authoritative Name Servers

3.2.1. Including DELEG RRs in a Response

If a DELEG RRset is present at the delegation point, the name server MUST return both the DELEG RRset and its associated RRSIG RR in the Authority section along with the DS RRset and its associated RRSIG RR and the NS RRset.

If no DELEG RRset is present at the delegation point, and the zone was signed with a DNSKEY that has the DELEG flag set, the name server MUST return the NSEC or NSEC3 RR that proves that the DELEG RRset is not present including its associated RRSIG RR along with the DS RRset and its associated RRSIG RR if present and the NS RRset, if present.

Including these DELEG, DS, NSEC or NSEC3, and RRSIG RRs increases the size of referral messages. If space does not permit inclusion of these records, including glue address records, the name server MUST set the TC bit on the response.

3.2.2. Responding to Queries for Type DELEG

DELEG records, when present, are included in referrals. When a parent and child are served from the same authoritative server, this referral will not be sent because the authoritative server will respond with information from the child zone. In that case, the resolver may query for type DELEG.

The DELEG resource record type is unusual in that it appears only on the parent zone's side of a zone cut. For example, the DELEG RRset for the delegation of "foo.example" is part of the "example" zone rather than in the "foo.example" zone. This requires special processing rules for both name servers and resolvers because the name server for the child zone is authoritative for the name at the zone cut by the normal DNS rules, but the child zone does not contain the DELEG RRset.

A DELEG-aware resolver sends queries to the parent zone when looking

for a DELEG RR at a delegation point. However, special rules are necessary to avoid confusing legacy resolvers which might become involved in processing such a query (for example, in a network configuration that forces a DELEG-aware resolver to channel its queries through a legacy recursive name server). The rest of this section describes how a DELEG-aware name server processes DELEG queries in order to avoid this problem.

The need for special processing by a DELEG-aware name server only arises when all the following conditions are met:

* The name server has received a query for the DELEG RRset at a zone cut.

* The name server is authoritative for the child zone.

* The name server is not authoritative for the parent zone.

* The name server does not offer recursion.

In all other cases, the name server either has some way of obtaining the DELEG RRset or could not have been expected to have the DELEG RRset, so the name server can return either the DELEG RRset or an error response according to the normal processing rules.

If all the above conditions are met, however, the name server is authoritative for the domain name being searching for, but cannot supply the requested RRset. In this case, the name server MUST return an authoritative "no data" response showing that the DELEG RRset does not exist in the child zone's apex.

3.2.3. Priority of DELEG over NS and Glue Address records

DELEG-aware resolvers SHOULD prioritize the information in DELEG records over NS and glue address records.

4. Privacy Considerations

All of the information handled or transmitted by this protocol is public information published in the DNS.

5. Security Considerations

TODO: Fill this section out

5.1. Availability of Zones Without NS

5.2. Resolution Procedure

An example of a simplified DNS interaction after priming. This is a query for www.example.com type AAAA with DELEG-aware com and example.com authoritative servers.

* Ask www.example.com qtype AAAA to a.root-servers.net the answer is: Answer section: (empty) Authority section: com. 172800 IN NS a.gtld-servers.net. Additional section: a.gtld-servers.net. 172800 IN AAAA 2001:db8:a83e::2:30

* Ask www.example.com qtype AAAA to a.gtld-servers.net the answer is: Answer section: (empty) Authority section: example.com. 172800 IN NS ns1.example.com. example.com. 172800 IN DELEG 1 config1.example.com. (ipv6hint=2001:db8:440:1:1f::24)

Additional section: ns1.example.com. 172800 IN AAAA 2001:db8:322c::35:42

* Ask www.example.com qtype AAAA to config1.example.com (2001:db8:1:1f::24) the answer is: Answer section: www.example.com. 3600 IN AAAA 2001:db8:a0:322c::2 Authority section: (empty) Additional section: (empty)

TODO: more resolution examples (e.g out of bailiwick)

5.2.1. Failures when DELEG Delegation is Present

When a delegation using DELEG to a child is present, the resolver MUST use it and SERVFAIL if none of the configurations provided work.

6. IANA Considerations

DELEG will use the SVCB IANA registry definitions in section 14.3 of [RFC9460].

The IANA has assigned a bit in the DNSKEY flags field (see Section 7 of [RFC4034] for the DELEG bit (N).

7. References

7.1. Normative References

[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/rfc/rfc4034>.

[RFC9460] Schwartz, B., Bishop, M., and E. Nygren, "Service Binding and Parameter Specification via the DNS (SVCB and HTTPS Resource Records)", RFC 9460, DOI 10.17487/RFC9460, November 2023, <https://www.rfc-editor.org/rfc/rfc9460>.

[STD13] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, November 1987.

Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

<https://www.rfc-editor.org/info/std13>

7.2. Informative References

[BCP219] Hoffman, P., Sullivan, A., and K. Fujiwara, "DNS Terminology", BCP 219, RFC 8499, January 2019.

<https://www.rfc-editor.org/info/bcp219>

[I-D.tapril-ns2]
April, T., "Parameterized Nameserver Delegation with NS2 and NS2T", Work in Progress, Internet-Draft, draft-tapril-ns2-01, 13 July 2020, <https://datatracker.ietf.org/doc/html/draft-tapril-ns2-01>.

[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/rfc/rfc2119>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

Appendix A. Legacy Test Results

In December 2023, Roy Arends and Shumon Huque tested two distinct sets of requirements that would enable the approach taken in this document.

* legacy resolvers ignore unknown record types in the authority section of referrals.

* legacy resolvers ignore an unknown key flag in a DNSKEY.

Various recent implmentations were tested (BIND, Akamai Cacheserve, Unbound, PowerDNS Recursor and Knot) in addition to various public resolver services (Cloudflare, Google, Packet Clearing House). All possible variations of delegations were tested, and there were no issues. Further details about the specific testing methodology, please see test-plan.

Appendix B. Acknowledgments {:unnumbered}

This document is heavily based on past work done by Tim April in [I-D.tapril-ns2] and thus extends the thanks to the people helping on this which are: John Levine, Erik Nygren, Jon Reed, Ben Kaduk, Mashooq Muhaimen, Jason Moreau, Jerrod Wiesman, Billy Tiemann, Gordon Marx and Brian Wellington.

Appendix C. TODO

RFC EDITOR: PLEASE REMOVE THE THIS SECTION PRIOR TO PUBLICATION.

* Write a security considerations section

* worked out resolution example including alias form delegation

Appendix D. Change Log

RFC EDITOR: PLEASE REMOVE THE THIS SECTION PRIOR TO PUBLICATION.

0123456789012345678901234567890123456789012345678901234567890123456789012345678901234567891

Contributors

Christian Elmerot
Cloudflare
Email: christian@elmerot.se

Edward Lewis
ICANN
Email: edward.lewis@icann.org

Roy Arends
ICANN
Email: roy.arends@icann.org

Shumon Huque
Salesforce
Email: shuque@gmail.com

Klaus Darilion
nic.at
Email: klaus.darilion@nic.at

Libor Peltan
CZ.nic
Email: libor.peltan@nic.cz

Vladimír unát
CZ.nic
Email: vladimir.cunat@nic.cz

Shane Kerr
NS1
Email: shane@time-travellers.org

David Blacka
Verisign
Email: davidb@verisign.com

George Michaelson
APNIC
Email: ggm@algebras.org

Ben Schwartz
Meta
Email: bemasc@meta.com

Jan Velák
NS1
Email: jvcelak@ns1.com

Peter van Dijk
PowerDNS
Email: peter.van.dijk@powerdns.com

Philip Homburg
NLnet Labs
Email: philip@nlnetlabs.nl

Erik Nygren
Akamai Technologies
Email: erik+ietf@nygren.org

Vandan Adhvaryu

Team Internet
Email: vandan@adhvaryu.uk

Manu Bretelle
Meta
Email: chantr4@gmail.com

Authors' Addresses

Tim April
Email: ietf@tapril.net

Petr paek
ISC
Email: pspacek@isc.org

Ralf Weber
Akamai Technologies
Email: rweber@akamai.com

David C Lawrence
Salesforce
Email: tale@dd.org