Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: 5 January 2026

K. Fujiwara
JPRS

4 July 2025

Upper limit values for DNS
draft-fujiwara-dnsop-dns-upper-limit-values-03

Abstract

DNS was designed to have as few hard upper limits as possible to allow for future extensibility. However, the lack of a clear upper limit leads to vulnerabilities, and several attack methods have been reported. This document proposes reasonable upper-limit values for DNS protocols as a Best Current Practice.

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 5 January 2026.

Table of Contents

1. Introduction

There are parameters in the DNS protocol that do not have clear upper limits. For example, the number of alias levels using CNAME Resource records and the number of resource records in an RRSet.

If a protocol is implemented without considering the upper limit, it may become vulnerable to DoS attacks, and several attack methods have been proposed.

This document introduces some upper limits to the DNS protocol as a best current practice, to allow secure use of the DNS while preserving the flexibility that the DNS protocol was designed to provide.

2. 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.

The specialized terms used in this document are defined in DNS Terminology [RFC9499].

Gluelessness is the term used to describe delegation without glue.

3. Background

DNS was designed to have as few upper limits as possible to allow for future extensibility described in the paper "Development of the Domain Name System" [Mockapetris1988].

If a protocol is implemented without considering the upper limit, it may become vulnerable to DoS attacks.

[RFC5358] describes DNS Reflector Attacks and how to prevent the use of default configured recursive nameservers. Simply preparing a large RRSet can increase the amplification factor of DNS Reflector Attacks. Countermeasures for recursive resolvers were described in [RFC5358], however, countermeasures for authoritative servers have not been standardized as an RFC and are implemented in various software as DNS Response Rate Limiting.

In recent years, DNS vulnerabilities research have been actively progressed and many vulnerabilities have been made public. Each time a vulnerability is discovered, upper limits on the execution time, number of attempts, and size are added to DNS software implementations.

CNAME aliases are widely used; however, when there are multiple levels of CNAME aliases, full-service resolvers have to redo the name resolution, which increases the load. And more, each stub resolver that receives a response containing multiple CNAME aliases must find the final A, AAAA Resource record that corresponds to the CNAME in each application. Unbound and BIND 9 introduced ’max-query-restarts’ parameter, and the default is 11. (Hard limit on the number of times Unbound is allowed to restart a query upon encountering a CNAME record.)

Currently, many web services that use domain names require that a TXT record containing a token of their choosing be placed at the zone apex to verify the registrant domain name. A large enterprise set 74 TXT records at its zone apex.

CVE-2024-1737, "BIND’s database will be slow if a very large number of RRs exist at the same name" was reported, and BIND 9.18.28 implemented the limit. BIND 9 introduced ’max-records-per-type’ parameter that limits the number of resource records in an RRSet, and the default is 100.

KeyTrap [KeyTrap] is a vulnerability caused by the fact that there is no upper limit on the number of DNSKEY, DS, or RRSIG resource records. Unbound introduced the maximum number of RRSIG validations for an RRset (MAX_VALIDATE_RRSIGS) as 8, and the maximum allowed digest match failures per DS, for DNSKEYs with the same properties (MAX_DS_MATCH_FAILURES) as 4.

4. Problem Statement

DNS was designed to have as few upper limits as possible to allow for future extensibility. However, in reality, DoS attacks using large responses and the use of excessively long CNAME chains has been reported, and resource-wasting attacks using many DNSKEY/DS/RRSIG combinations have been reported.

To mitigate these attacks, it is possible to set realistic upper limits on some parameters, and to cause a name resolution error if those limits are exceeded.

In order to avoid breaking existing mechanisms, widely used values should not be treated as errors.

Secondary authoritative servers simply provide copies from the primary servers, so they should not cause errors due to the upper limit.

Primary authoritative servers are expected to report errors when they read zone files, or receive dynamic updates that contains RRsets that exceed the upper limit.

Full-service resolvers may prevent malfunction by treating a name resolution error as responses from authoritative servers that exceed the upper limit.

Since restrictions due to datagram size are possible, it seems reasonable to propose an upper limit on data size, etc., as a best current practice.

Best Current Practice documents should allow for values that are currently in widespread use. However, apparent anomalies may be excluded.

It is desirable to determine the upper limit values by conducting extensive measurements on the Internet and excluding obvious errors or malicious errors, and to set the value that has the least impact.

5. Possible upper limits

5.1. Possible upper limit items

* Packet size
* Number of Resource Records in a RRSet
* Number of NS Resource Records in a delegation
* Number of DS Resource Records in a delegation
* Number of glue RRs in a delegation
* Number of DNSKEY Resource Records in a DNSKEY RRSet
* Number of RRSIG RRs for each name and type
* Number of levels of gluelessness delegations
* Number of alias levels using CNAME Resource records

5.2. Packet size

There were comments that there are size limitations even if there are no precise upper limits are set.

The DNS packet format has an upper limit of 65535 octets, so an RRset cannot exceed that size. Also, the upper limit size of a single resource record is 65535 octets minus DNS header size because RDLENGTH is 16 bits.

Section 4.2.1 UDP usage of [RFC1035] limits the UDP message size to 512.

The size of a DNS response that can be sent using unfragmented UDP is about 1400 octets. [RFC9715]

The size 65535 is large, and attackers use this upper limit to carry out resource-wasting attacks.

5.3. Number of Resource Records in a RRSet

Since there are 13 root name servers and 13 name servers for com and net TLDs, the maximum number of NS RR in an NS RRSet should be larger than or equal to 13.

Since there are 13 name servers for root, com, net and they have both IPv4 and IPv6 addresses, 26 glue records in a delegation should be allowed.

Although we could not find a clear explanation, the upper limit value of the number of name server names that can be registered for gTLDs and ccTLDs is 13. Therefore, the practical upper limit for the number of name servers is 13.

In recent years, there have been cases where many TXT resource records have been set at the zone apex. Many services seem to request their designated authentication tokens written as TXT records at the zone apex to verify domain name registrants. However, there seem to be cases where authentication tokens are only added, as there seems to be no procedure for deleting them once they have been set. It is necessary to standardize and deploy [I-D.ietf-dnsop-domain-verification-techniques], and to write TXT records not at the zone apex, but application-specific undercore prefix labels.

The number of zone apex TXT records will have to remain as is until [I-D.ietf-dnsop-domain-verification-techniques] will be standardized, and the upper limit of the number of resource records in an RRSet may be 100, as the BIND 9 limits by default.

5.4. Number of alias levels using CNAME/DNAME

Many resolver implementations can resolve over 10 CNAME aliases. Unbound and BIND 9 introduced ’max-query-restarts’ parameter, and the default is 11.

Each application interprets complex CNAME chains. To avoid complexity in applications, it is recommended to use as few CNAME chains as possible.

A maximum of 9 CNAME chains were observed for the top 1 million popular domain names and domain names observed by a university’s resolvers. Therefore, a realistic upper limit for CNAME chains is thought to be 9 or 11.

5.5. Number of RRSIGs/DNSKEYs/DSs in a RRSet

KeyTrap [KeyTrap] is a vulnerability caused by the fact that there is no upper limit on the number of DNSKEY, DS, or RRSIG resource records. If there were upper limits on these, the damage could be mitigated.

Unbound introduced the maximum number of RRSIG validations for an RRset (MAX_VALIDATE_RRSIGS) as 8.

There are no good values to base the number of DS resource records or DNSKEY resource records, but having 8 DS records and 8 DNSKEY resource records will be sufficient to handle key rollovers and multi-signers.

5.6. Number of delegation levels of gluelessness delegations

[RFC9471] states that all in-domain glue records are attached to the delegation response. Therefore, using in-domain name server names for DNS delegation minimizes name resolution costs.

Unrelated (or, rarely sibling) name server names are used/required for DNS hosting services.

However, using unrelated name server names increases the name resolution costs and may increase the likelihood of name resolution errors.

For some domain names, there are multiple layers of dependence on unrelated name server names when resolving the name. If information on unrelated name server names is not in the cache, the recursive resolver must perform A/AAAA name resolution for the unrelated name server names. Frequent use of unrelated name server names can cause unstable name resolution.

Furthermore, there are cases where cyclic dependencies in delegation occur, settings that depend on sibling glue, and cases where the sibling glue disappears or some name servers stop responding, making it impossible to resolve names.

[Tsuname2021] pointed out attacks and countermeasures that use increased load due to cyclic dependencies. Many cyclic delegations are likely due to misconfigurations.

[djbdns] allows three levels of gluelessness.

To avoid complex name resolutions and misconfigurations, it is better to avoid using unrelated name server names as much as possible.

Unrelated name server names SHOULD be hosted by a domain name with at least one in-domain name server name. In other words, DNS providers SHOULD have at least one in-domain nameserver for their domain names.

6. Recommendation of DNS upper limit values

Name	upper limit
number of RRs in an RRSet	100
number of NS RRs in a delegation	13
number of glue RRs in a delegation	26
number of DS RRs in a delegation	8
number of DNSKEY RRs in an RRSet	8
number of RRSIG RRs for each name and type	8
number of CNAME/DNAME chains	9
number of levels of gluelessness delegations	3

Table 1

DNS software is expected to make these items configurable parameters that operators can control.

Recursive resolvers SHOULD respond with a name resolution error (Server Failure) with Extended DNS Errors [RFC8914] if it receives a response from an authoritative server that exceeds the upper limit value.

Primary authoritative servers SHOULD be in an error state if they find RRSets that exceed the hard limits when they load zone files, receive zone data by zone transfers, or receive DNS Updates.

7. IANA Considerations

This document requests no IANA actions.

8. Security Considerations

9. References

9.1. Normative References

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <https://www.rfc-editor.org/rfc/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/rfc/rfc2119>.

[RFC5358] Damas, J. and F. Neves, "Preventing Use of Recursive Nameservers in Reflector Attacks", BCP 140, RFC 5358, DOI 10.17487/RFC5358, October 2008, <https://www.rfc-editor.org/rfc/rfc5358>.

[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>.

[RFC8914] Kumari, W., Hunt, E., Arends, R., Hardaker, W., and D. Lawrence, "Extended DNS Errors", RFC 8914, DOI 10.17487/RFC8914, October 2020, <https://www.rfc-editor.org/rfc/rfc8914>.

[RFC9471] Andrews, M., Huque, S., Wouters, P., and D. Wessels, "DNS Glue Requirements in Referral Responses", RFC 9471, DOI 10.17487/RFC9471, September 2023, <https://www.rfc-editor.org/rfc/rfc9471>.

[RFC9499] Hoffman, P. and K. Fujiwara, "DNS Terminology", BCP 219, RFC 9499, DOI 10.17487/RFC9499, March 2024, <https://www.rfc-editor.org/rfc/rfc9499>.

9.2. Informative References

[djbdns] Bernstein, D. J., "djbdns: Notes on the Domain Name System", n.d., <https://cr.yp.to/djbdns/notes.html>.

[I-D.ietf-dnsop-domain-verification-techniques] Sahib, S. K., Huque, S., Wouters, P., Nygren, E., and T. Wicinski, "Domain Control Validation using DNS", Work in Progress, Internet-Draft, draft-ietf-dnsop-domain-verification-techniques-07, 3 March 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-dnsop-domain-verification-techniques-07>.

[KeyTrap] Elias Heftrig, Haya Schulmann, Niklas Vogel, and Michael Waidner, "The KeyTrap Denial-of-Service Algorithmic Complexity Attacks on DNS", 2024.

[Mockapetris1988]
Paul V Mockapetris and Kevin J Dunlap, "Development of the domain name system", the Proceedings of SIGCOMM 1988 , August 1988.

[RFC9715] Fujiwara, K. and P. Vixie, "IP Fragmentation Avoidance in DNS over UDP", RFC 9715, DOI 10.17487/RFC9715, January 2025, <https://www.rfc-editor.org/rfc/rfc9715>.

[Tsuname2021]
Moura, G. M., Sebastian Castro, John S Heidemann, and Wes Hardaker, "TsuNAME: exploiting misconfiguration and vulnerability to DDoS DNS", IMC ’21: Proceedings of the 21st ACM Internet Measurement Conference , 2021.

Author’s Address

Kazunori Fujiwara
Japan Registry Services Co., Ltd.
Japan
Email: fujiwara@wide.ad.jp