

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: March 29, 2019

H. Zhang
P. Zuo
J. Yao
CNNIC
September 25, 2018

**New weighted resource record for traffic scheduling
draft-zhang-dnsop-weighted-address-records-00**

Abstract

Scheduling traffic in proportion to different nodes can optimize bandwidth utilization and improve user experience when the Internet traffic exceeds the bandwidth limit of a CDN cache node. This document defines AX and AAAAX RR types by adding weight field to existing A and AAAA resource records. A group of weighted records will be sorted in the DNS response according to the weight in AX and AAAAX records.

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 March 29, 2019.

Copyright Notice

Copyright (c) 2018 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

1. Introduction

Most of CDNs can bring a client to the closest cache node according to the source IP of the client nowadays. However, the rapid development of Internet raises new challenges. According to public information, Global IP traffic will increase nearly threefold over the next 5 years and the busy-hour Internet traffic is growing more rapidly than average traffic. For some CDNs, it becomes more frequent that peak traffic will exceed the bandwidth limit of their nodes and they need new solutions to optimize traffic scheduling.

For example, if the maximum bandwidth for a CDN node is 10G and the busy-hour traffic reaches 12G, it is necessary to schedule 10G traffic to this node and the extra 2G traffic to other adjacent nodes. Scheduling traffic in proportion to different nodes can optimize bandwidth utilization and improve user experience.

2. Possible solutions

In general, there are three ways to solve this problem:

(1) Get browsers to comply with [RFC 2782](#)

[RFC 2782](#) describes a DNS RR which specifies the location of the server for a specific protocol and provides a server selection mechanism. However, getting browsers to move is a long project. In practice, [RFC2782](#) is more suitable for service discovery but not widely adopted for dynamic web server selection.

(2) Serve up multiple copies of the same RR and expect that record shuffling will deliver a specific ratio

The multiple IP address listings is done in practice. However, it is meaningless for two records to ever have label, class, type and data all equal - servers should suppress such duplicates if encountered.(as described in [[RFC2181](#)]) . Therefore, this method is not recommended as some servers may remove duplicates.

(3) Get authoritative servers and resolvers to serve weighted A/AAAA records and short TTL

CDNs are actually using a similar method nowadays. Weighted A/AAAA records and short TTLs brings a dynamic server selection mechanism

that optimizes traffic scheduling. Compared to [RFC2782](#), this is a light-weight solution that can be widely adopted easily.

This document defines AX/AAAAX resource record to standardize the third method. By adding weight field in AX/AAAAX, a DNS server can serve address records with weighted frequency. The similar mechanism currently works well in many authoritative servers in practice but very few resolvers support it. Support from resolvers further optimizes scheduling but also increases complexity. If the resolver should support this mechanism is a problem for future discussion. For compatibility, the authoritative DNS server that has AX/AAAAX RRs must reply A/AAAA RRs to weight-RR-not-aware resolvers.

[3.](#) AX and AAAAX resource record

The AX RR presentation format is similar to that of A RR except the weight field. The weight is a 16 bit unsigned integer:

```
owner ttl class AX weight address
```

The AAAAX RR presentation format is similar to that of AAAA RR except the weight field. The weight is a 16 bit unsigned integer:

```
owner ttl class AAAAX weight address
```

[4.](#) Authoritative Server Behavior

[4.1.](#) Use of AX and AAAAX

When AX or AAAAX records are present at a DNS node and a query is received by an authoritative server for type AX or AAAAX, the authoritative server should: (1) sort AX or AAAAX RRs according to the weight; (2) place the AX or AAAAX RRs in the answer section.

When AX or AAAAX records are present at a DNS node and a query is received by an authoritative server for type A or AAAA, the authoritative server should: (1) remove weight field in AX or AAAAX; (2) sort A or AAAA RRs according to the weight in AX or AAAAX; (3) place the sorted A or AAAA RRs in the answer section;

[4.2.](#) example

Given the following zone:

```
$ORIGIN example.com.  
@ IN SOA example.com hostmaster.example.com 1 7200 3600 2419200 600  
@ IN NS ns1  
www IN AX 3 1.1.1.1.  
www IN AX 1 2.2.2.2
```

the weight filed in the above 2 AX RRs indicates 1.1.1.1 should account for 75% for example.com/A query and 2.2.2.2 account for 25%.

For the example.com/A queries:
Three quarters of responses should be:

```
;; QUESTION SECTION:  
www.example.com. IN A  
;; ANSWER SECTION:  
www.example.com. 60 IN A 1.1.1.1  
www.example.com. 60 IN A 2.2.2.2
```

one quarter of responses should be:

```
;; QUESTION SECTION:  
www.example.com. IN A  
;; ANSWER SECTION:  
www.example.com. 60 IN A 2.2.2.2  
www.example.com. 60 IN A 1.1.1.1
```

For the example.com/AX queries, the server would return:

```
;; QUESTION SECTION:  
www.example.com. IN A  
;; ANSWER SECTION:  
www.example.com. 60 IN AX 3 1.1.1.1  
www.example.com. 60 IN AX 1 2.2.2.2
```

4.3. DNSSEC consideration

If the zone that contains AX or AAAAX is DNSSEC-signed, for each AX or AAAAX, there must be a corresponding A or AAAA RR that has identical RDATA. The corresponding A or AAAA RR must be signed.

5. Recursive Server Behavior

Support from resolvers will further optimizes traffic scheduling but increases complexity at the same time. As the authoritative DNS server that has AX/AAAAX RRs must reply A/AAAA RRs for an A/AAAA query, this mechanism works well even if a resolver does not support it.

If a resolver does not support this mechanism, which currently accounts for the majority of cases, it never receives any AX/AAAAX RRs. It requires no extra implementation to support the weighted RRs mechanism.

If a resolver supports this mechanism, it should determine if to use the weighted RRs according to local configuration. Use of weighted address records in resolvers is an issue for further study.

6. Implementation Status

PowerDNS [<https://powerdns.com>] currently implements a similar authoritative-only feature using LUA records.

DNS service provider CloudXNS [<https://www.cloudxns.net>] currently implements a similar authoritative-only feature using AX record that contains a priority value.

CDN service provider UPYUN [<https://www.upyun.com>] currently adopts a similar mechanism to schedule traffic in busy hour.

7. IANA Considerations

IANA is requested to assign 2 DNS RR data type values for the AX and AAAAX RR type.

8. References

- [RFC2181] Elz, R. and R. Bush, "Clarifications to the DNS Specification", July 1997, <<https://tools.ietf.org/html/rfc2181>>.
- [RFC2782] Gulbrandsen, A., Vixie, P., and L. Esibov, "A DNS RR for specifying the location of services (DNS SRV)", February 2000, <<https://tools.ietf.org/html/rfc2782>>.

Authors' Addresses

Haikuo Zhang
CNNIC
4 South 4th Street, Zhongguancun, Haidian District
Beijing, Beijing 100190
China

Phone: +86 10 5881 3163
Email: zhanghaikuo@cnnic.cn

Peng Zuo
CNNIC
4 South 4th Street, Zhongguancun, Haidian District
Beijing, Beijing 100190
China

Phone: +86 10 5881 2629
Email: zuopeng@cnnic.cn

Jiankang yao
CNNIC
4 South 4th Street, Zhongguancun, Haidian District
Beijing, Beijing 100190
China

Phone: +86 10 5881 3007
Email: yaojiankang@cnnic.cn