DNS load balancing
draft-sonoda-dnsop-dnslb-03

Abstract

This document defines a new DNS load balance function that is able to transfer information in zone transfer and not need online signing. DNS base load balance is popular technology. It provides weight base response and location base response. It have become an indispensable part of traffic engineering. However, DNS base load balance can’t transfer load balance information in zone transfer and need online singing because it is not standardized. This document defines a new DNS resource record called "LB" and new EDNS option bit called "LS". LB RR provides the balancing information weight, location and target domain name. LS bit provides the change response mechanism in name servers.

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 December 25, 2018.

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.
1. Introduction

1.1. Legacy DNS load balancing

Legacy DNS load balancing provides traffic engineering. It uses a special authoritative name server. Its response is one address record that is dynamic changes using network location and weight. It's used for large traffic WEB site domain name that is important domain name. Important domain name should be secure.

But legacy DNS load balancing is not secure. Because Legacy DNS load balancing can't send zone data by zone transfer. It's mean very difficult to use multi service providers. That means weak for DDoS Attack. If zone is signed, All name servers require private key for dynamic signing because response is dynamic changes. Distributing private key is not secure, It is increased risk of leakage private key.

1.2. New DNS load balancing

New DNS load balancing concept is that Authoritative name server uses "LB" RR to provide load balancing and Full resolver use "LB" RR to select target name and respond with the target name.

"LB" RR defines load balancing settings that is network location and weight and target domain name. Network location is string that meaningful name of network. For example Country code (ex. JP), subdivision code (ex. US-CA) and Autonomous System Number (ex. AS65536). Full resolver use resolver configuration to select network location and refine "LB" RRs. Weight is integer value. Full resolver use weight ratio to select target name. Target name is pointer to address record. It is same "ANAME" [I-D.ietf-dnsop-aname].

2. Mechanism

1. Stub resolver sends A or AAAA query to full resolver.

2. Full Resolver sends A or AAAA query with LS bit to authoritative server.

4. Full Resolver selects target domain name using location and weight.

5. Full Resolver resolv target domain name.

6. Full Resolver response target domain name and LB RRSet to stub resolver.

3. The LB Resource Record

The LB RR has mnemonic LB. LB RR define load balancing information.

LB format below.

<owner> <ttl> <class> LB <weight> <location> <target>

The format is not class-sensitive. All fields are required.

<weight> field is a 2 octets, 1 or more natural number.

<location> field is a "<character-string>" [RFC1035].

<target> field is a "<domain-name>" [RFC1035]. It has A or AAAA RR or CNAME RR or DNAME RR.

3.1. Define location

<location> ::= "*" | <continental code> | <country code> | <subdivision code> | <asn code> | <private code>

<region code> ::= "AF" | "NA" | "AS" | "EU" | "NA" | "OC" | "SA"
<country code> ::= ISO 3166-1 alpha-2 Country code.
<subdivision code> ::= ISO 3166-2 Codes for the representation of names of countries and their subdivisions.
<asn code> ::= "AS" <asn> [ ":" <asn>]
<private code> ::= "+" <let>
<ans> ::= <nonzero digit> <digit>
<nonzero digit> ::= any one of the ten digits 1 through 9
<digit> ::= any one of the ten digits 0 through 9
<let> ::= any one of the 26 alphabetic characters A through Z in upper case or any one of the ten digits 0 through 9.

3.2. Record example

example.jp. 3600 IN LB 1 * www.example.com. ; for any region
example.jp. 3600 IN LB 1 AS as.example.com. ; for ASIA region
example.jp. 3600 IN LB 1 JP jpl.example.jp. ; for JP region, weight 1
example.jp. 3600 IN LB 3 JP jp2.example.jp. ; for JP region, weight 3
example.jp. 3600 IN LB 1 JP-13 tokyo.example.jp. ; for tokyo region
example.jp. 3600 IN LB 1 AS2496 as65536.example.jp. ; for AS65536
example.jp. 3600 IN LB 1 AS2496:1 as65536.example.jp. ; for AS65536
example.jp. 3600 IN LB 1 +BEER beer.example.jp. ; private use

4. The LB Support Flag

Defines a new "EDNS Header Flags" [RFC6891] call LB Support Flag(LS) using full resolver sends LB RR supported to authoritative server.

LS bit provides change response mechanism in authoritative name server. If LS bit is flagged, Authoritative name server can response LB RR for A,AAAA query.

5. Authoritative name server Behavior

When authoritative name server receives a query of type A or AAAA with LS bit and LB record is present at a SNAME, The authoritative server returns the LB RRSet in the answer section with LS bit.

5.1. Example of authoritative name server behavior

Example zone data:

example.jp. 3600 IN SOA ( ns1.example.com.
    postmaster.example.com.
     1
     3600
     900
     1814400
     900 )
example.jp. 3600 IN NS ns1.example.com.
example.jp. 3600 IN NS ns2.example.com.
example.jp. 3600 IN A 198.51.100.1
example.jp. 3600 IN LB 1 * www.example.com.
example.jp. 3600 IN LB 1 JP jpl.example.com.
example.jp. 3600 IN LB 3 JP jp2.example.com.

Incoming query with LS bit:
query: qtype = example.jp. qtype=A, LS=1

Response for include LS:

query: qtype = example.jp. qtype=A
response: LS=1
answer:
  example.jp. 3600 IN LB 1 * www.example.com.
  example.jp. 3600 IN LB 1 JP jp1.example.com.
  example.jp. 3600 IN LB 3 JP jp2.example.com.
authority:
  example.jp. 3600 IN NS ns1.example.com.
  example.jp. 3600 IN NS ns2.example.com.

Incoming query without LS bit (normal query):

query: qtype = example.jp. qtype=A, LS=0

Response for not include LS:

query: qtype = example.jp. qtype=A
response: LS=0
answer:
  example.jp. 3600 IN A 198.51.100.1
authority:
  example.jp. 3600 IN NS ns1.example.com.
  example.jp. 3600 IN NS ns2.example.com.

6. Full Service Resolver Behavior

When a full resolver sends a query of type A or AAAA with LS bit and receives a response with a LB RRset in the answer section with LS bit, Full resolver MUST re-resolv the either LB <target> of type "STYPE" [RFC1034]. <target> be selected by <location> and <weight>.

6.1. Location selection

Location selection needs full resolver or stub resolver location information. Full resolver location information with priority value SHOULD gets from config file. Full resolver MUST setting '*' location with lowest priority. Full resolver select location that match either LB RR <location> and highest priority resolver location. If all LB RRs don’t match all resolver locations, resolver selects a location randomly.
6.2. Weight selection

Full resolver selects <target> using <weight> from LB RR whose location matches.

6.3. Example

1. Full resolver location is configured "JP-13" "JP" "AS" ".*

2. Stub resolver query coming:
   
   query: qtype = example.jp. qtype=A

3. Full resolver send query:
   
   query: qtype = example.jp. qtype=A, LS=1

3. If response include LB:
   
   query: qtype = example.jp. qtype=A
   response: LS=1
   answer:
   
   example.jp. 3600 IN LB 1 * www.example.com.
   example.jp. 3600 IN LB 1 JP jpl.example.com.
   example.jp. 3600 IN LB 3 JP jpl2.example.com.
   authority:
   
   example.jp. 3600 IN NS ns1.example.com.
   example.jp. 3600 IN NS ns2.example.com.

4. select LB RR that’s location include resolver location
   
   example.jp. 3600 IN LB 1 JP jpl.example.com.
   example.jp. 3600 IN LB 3 JP jpl2.example.com.

5. Select one LB RR using <weight>:

   example.jp. 3600 IN LB 3 JP jpl2.example.com.

5. Name resolution <target>:

   query: qtype = jpl2.example.com. qtype=A, LS=1
   response:
   
   answer:
   
   jpl2.example.com. 3600 IN A 192.0.2.2
   authority:
   
   example.com. 3600 IN NS ns1.example.com.
   example.com. 3600 IN NS ns2.example.com.
6. Make response message:

```
query: qtype = example.jp. qtype=A, LS=1
response:
  answer:
  example.jp. 3600 IN LB 1 * www.example.com.
  example.jp. 3600 IN LB 1 JP jp1.example.com.
  example.jp. 3600 IN LB 3 JP jp2.example.com.
  jp2.example.com. 3600 IN A 192.0.2.2
  authority:
  example.jp. 3600 IN NS ns1.example.com.
  example.jp. 3600 IN NS ns2.example.com.
```

7. Stub resolver uses 192.0.2.2

7. Implementation Status

7.1. Authoritative name server

NSD proof of concept patches can be seen at
https://github.com/mimuret/nsd-with-ls

KnotDNS proof of concept patches can be seen at
https://github.com/mimuret/knotd-with-ls

7.2. Full resolver name server

Knot-resolver proof of concept patches can be seen at
https://github.com/mimuret/knot-resolver-with-ls

7.3. Stub resolver behavior

Stub resolvers status below.

<table>
<thead>
<tr>
<th>Implementation</th>
<th>gethostbyname*</th>
<th>getaddrinfo</th>
<th>AF_INET</th>
<th>AF_INET6</th>
<th>AF_UNSPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>macOS(10.13)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Windows 10</td>
<td>X</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Linux(glic)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>Solaris 11</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
8. IANA Considerations

IANA is requested to assign a DNS RR data type value for the LB RR type under the "Resource Record (RR) TYPES" sub-registry and a EDNS Header Flag value for the LB Support Flag under the "EDNS Header Flags (16 bits)" sub-registry under the "Domain Name System (DNS) Parameters" registry.

9. Security Considerations

Both authoritative server and resolvers that implement LB SHOUD carefully check for loops.

10. Normative References

[I-D.ietf-dnsop-aname]


Author’s Address

Manabu Sonoda (editor)
Internet Initiative Japan Inc.

Email: manabu-s@iij.ad.jp