Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 26, 2013

J. Abley
ICANN

June 24, 2013

A Summary of Various Mechanisms Deployed at L-Root for the Identification of Anycast Nodes
draft-jabley-dnsop-anycast-mapping-02

Abstract

Anycast is a deployment technique commonly employed for authoritative-only servers in the Domain Name System (DNS). L-Root, one of the thirteen root servers, is deployed in this fashion.

Various techniques have been used to map deployed anycast infrastructure externally, i.e. without reference to inside knowledge about where and how such infrastructure has been deployed. Motivations for performing such measurement exercises include operational troubleshooting and infrastructure risk assessment. In the specific case of L-Root, the ability to measure and map anycast infrastructure using the techniques mentioned in this document is provided for reasons of operational transparency.

This document describes all facilities deployed at L-Root to facilitate mapping of its infrastructure and serves as documentation for L-Root as a measurable service.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

The Domain Name System (DNS) is described in [RFC1034] and [RFC1035]. L-Root, one of the thirteen root servers, is deployed using anycast [RFC4786]; its service addresses, published in the A and AAAA Resource Record (RR) Sets for "L.ROOT-SERVERS.NET", are made available from a substantial number of semi-autonomous servers deployed throughout the Internet. A list of locations served by L-Root can be found at <http://www.root-servers.org>.

[Fan2013] describes a technique using open DNS resolvers to distribute mapping queries to the service addresses of authoritative-only servers. This technique relies upon the ability to acquire meaningful information about individual anycast nodes by means of an IN-class query. At the time the experiments described in that paper were conducted, such ability existed with AS112 servers [RFC6304] but not with any root server. Modifications were subsequently made to the infrastructure of the L-Root service to facilitate this technique.

This document describes all facilities currently provided at L-Root to aid node identification.

2. Conventions Used in this Document

This document contains several examples of commands typed at a Unix (or Unix-like) command line to illustrate use of the various mechanisms available to identify L-Root nodes. Such examples are presented in this document with lines typed by the user preceded by the "%" prompt character; a bare "%" character indicates the end of the output resulting from the command.

In some cases the output shown in examples is too long to be represented directly in the text. In those cases a backslash character ("\") is used to indicate continuation.

3. Naming Scheme for L-Root Nodes

Individual L-Root nodes have structured hostnames that are constructed as follows:

```
<IATA Code><NN>.L.ROOT-SERVERS.ORG
```

where

- <IATA Code> is chosen from the list of three-character airport codes published by the International Air Transport Association (IATA) in the IATA Airline Coding Directory [1]; and

- <NN> is a two-digit numeric code used to distinguish between two different locations in the vicinity of the same airport.

Where multiple airports exist in the vicinity of a single L-Root node, one is arbitrarily chosen.

More granular location data published for L-Root nodes (e.g. see Section 4.4) is derived from the location of the airport, not the actual location of the node.

4. Identification of L-Root Nodes

L-Root service is provided using a single IPv4 address (199.7.83.42) and a single IPv6 address (2001:500:3::42). It should be noted that it is preferable to refer to the service using its DNS name (L.ROOT-SERVERS.NET) rather than literal addresses, since addresses can change from time to time.

At the time of writing there are 273 separate name server elements ("nodes") deployed in 143 locations which together provide L-Root service. A DNS query sent to an L-Root service address will be routed towards exactly one of those nodes for processing, and the corresponding DNS response will be originated from the same node. Queries from different clients may be routed to different nodes.

The following sections provide a summary of all mechanisms provided by L-Root to allow a client to identify which L-Root node is being used.

Using HOSTNAME.BIND/CH/TXT (Section 4.2), ID.SERVER/CH/TXT (Section 4.3) or IDENTITY.L.ROOT-SERVERS.ORG/IN/TXT or .../IN/A (Section 4.4) to identify a node for the purposes of reporting a problem is frequently reasonable, but it should be acknowledged that there is potential for re-routing between successive queries: an observed problem might relate to one node, whilst a subsequent query using one of those three techniques could be answered by a different node. Use of the NSID option can obviate this possibility (see Section 4.1).

4.1. Use of NSID

L-Root supports the use of the Name Server Identifier (NSID) Option [RFC5001] to return the identity of an L-Root node along with the response to a DNS query. The NSID payload of such responses is the fully-qualified hostname of the responding L-Root node.

The NSID option allows the identification of a node sending a specific, requested response to the client. This is of particular use if (for example) there is a desire to identify unequivocally what node is responding with a particularly troublesome response; the output of the diagnostic tool dig with NSID requested provides the problem response with the node identification, and its output in that case could form the basis of a useful trouble report.

NSID is specified as an EDNS0 option [RFC2671]. Clients that do not support EDNS0 signalling (or depend on other systems that do not support EDNS0) may find this mechanism unavailable.

The NSID option can be specified using the widely-used diagnostic tool "dig" using the "+nsid" option, as shown below. Note that long lines have been truncated for the purposes of this document ("\" at the end of a line indicates continuation).

```
% dig -4 @L.ROOT-SERVERS.NET . SOA +nsid \
  +norec +noall +comments
; <<>> DiG 9.6.-ESV-R3 <<>> -4 @L.ROOT-SERVERS.NET . SOA +nsid \
  +norec +noall +comments
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 14913
;; flags: qr aa; QUERY: 1, ANSWER: 1, AUTHORITY: 13, ADDITIONAL: 23

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
; NSID: 79 74 7a 30 31 2e 6c 2e 72 6f 6f 74 2d 73 65 72 76 65 72 73 \
  2e 6f 72 67  (y) (t) (z) (0) (1) (.) (l) (.) (r) (o) (o) (t) (-) \
  (s) (e) (r) (v) (e) (r) (s) (.) (o) (r) (g)
%


% dig -6 @L.ROOT-SERVERS.NET . SOA +nsid \
  +norec +noall +comments
; <<>> DiG 9.6.-ESV-R3 <<>> -6 @L.ROOT-SERVERS.NET . SOA +nsid \
  +norec +noall +comments
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 33374
;; flags: qr aa; QUERY: 1, ANSWER: 1, AUTHORITY: 13, ADDITIONAL: 23

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
; NSID: 79 74 7a 30 31 2e 6c 2e 72 6f 6f 74 2d 73 65 72 76 65 72 73 \
  2e 6f 72 67  (y) (t) (z) (0) (1) (.) (l) (.) (r) (o) (o) (t) (-) \
  (s) (e) (r) (v) (e) (r) (s) (.) (o) (r) (g)
%
```

4.2. Use of HOSTNAME.BIND/CH/TXT

L-Root supports the use of HOSTNAME.BIND/CH/TXT queries to return the identity of an L-Root node. The TXT RDATA returned is the fully-qualified hostname of the responding L-Root node.

The HOSTNAME.BIND/CH/TXT convention is described in [RFC4892].

```
% dig -4 @L.ROOT-SERVERS.NET HOSTNAME.BIND CH TXT +short
"ytz01.l.root-servers.org"
%

% dig -6 @L.ROOT-SERVERS.NET HOSTNAME.BIND CH TXT +short
"ytz01.l.root-servers.org"
%
```

4.3. Use of ID.SERVER/CH/TXT

L-Root supports the use of ID.SERVER/CH/TXT queries to return the identity of an L-Root node. The TXT RDATA returned is the fully-qualified hostname of the responding L-Root node.

ID.SERVER/CH/TXT functions identically (apart from the QNAME) to HOSTNAME.BIND/CH/TXT, as discussed in Section 4.2. The discussion there relating to the possibility of re-routing between successive queries also follows for ID.SERVER/CH/TXT.

The ID.SERVER/CH/TXT convention is described in [RFC4892].

```
% dig -4 @L.ROOT-SERVERS.NET ID.SERVER CH TXT +short
"ytz01.l.root-servers.org"
%

% dig -6 @L.ROOT-SERVERS.NET ID.SERVER CH TXT +short
"ytz01.l.root-servers.org"
%
```

4.4. Use of IDENTITY.L.ROOT-SERVERS.ORG/IN/TXT and .../IN/A

The operator of L-Root has distributed a separate DNS service in parallel with L-Root, operating on precisely the same set of nodes. Measurements of this separate service should give results which are representative of L-Root. Further discussion of this service can be found in Section 5.

The fully-qualified DNS name IDENTITY.L.ROOT-SERVERS.ORG (note the use of ORG, not NET) has associated TXT and A RR Sets which are unique to the responding node. Clients are hence able to issue queries for IDENTITY.L.ROOT-SERVERS.ORG/IN/A and IDENTITY.L.ROOT-SERVERS.ORG/IN/TXT and use the results both to identify individual nodes and to distinguish between responses generated by different nodes.

The TXT record returned in the response to such queries is structured as follows:

1. The fully-qualified host name of the node responding to the query;

2. The city in which the node is located;

3. The region in which the node is located;

4. The economy in which the node is located; and

5. The ICANN region in which the node is located.

The A record returned in the response to such queries is guaranteed to be unique to the responding node.

Since in this case identity data is published using IN-class resource records, it is not necessary to send queries directly towards L-Root in order to obtain results. Responses can be obtained through recursive servers, the responses in those cases being the identity of L-Root as observed through the recursive server used rather than the "closest" L-Root node to the client. This facilitates some degree of remote troubleshooting, since a query for IDENTITY.L.ROOT-SERVERS.ORG/IN/TXT or .../IN/A directed a remote recursive resolver can help illustrate which L-Root node is being used by that server (or was used when the cache was populated).

A related caching effect is that responses to IDENTITY.L.ROOT-SERVERS.ORG/IN/A and IDENTITY.L.ROOT-SERVERS.ORG/IN/TXT may be cached at different times, and may hence persist in a cache for overlapping periods of time. One possible visible effect is that the responses to .../IN/A and .../IN/TXT as presented from a cache may appear to be incoherent (i.e. refer to different nodes) despite queries against of the cache happening (near) simultaneously. Caches may also discard the published TTLs in responses from the authoritative server and replace them with longer TTLs, as a matter of local policy. Interpretation of responses for these queries from caches should therefore be carried out with these possible effects in mind.

It has been observed that IDENTITY.L.ROOT-SERVERS.ORG/IN/A queries offer a useful mechanism for troubleshooting DNS problems with non-technical users, since such users can often be walked through the process of looking up an A record (e.g. as a side effect of utilities such as ping) far easier than they can be instructed on how use DNS-specific tools such as dig.

```
  % dig IDENTITY.L.ROOT-SERVERS.ORG TXT +short
  "ytz01.l.root-servers.org" "Toronto" "Ontario" "Canada" "NorthAmerica"
  %

  % dig IDENTITY.L.ROOT-SERVERS.ORG A +short
  67.215.199.91
  %
```

4.5. Use of NODES.L.ROOT-SERVERS.ORG/IN/TXT

The fully-qualified DNS name NODES.L.ROOT-SERVERS.ORG (note again the use of ORG, not NET) provides multiple TXT RRs, one per node, and represents the effective concatenation of all possible responses to the query IDENTITY.L.ROOT-SERVERS.ORG/IN/TXT.

Note that in the example below we have forced dig to send the query over TCP, since we expect the response to be too large for UDP transport to accommodate. Note also that the list shown is truncated for clarity, and can be expected to change from time to time as new L-Root nodes are provisioned and old ones decommissioned.

```
% dig NODES.L.ROOT-SERVERS.ORG TXT +short +tcp | head -10
"abj01.l.root-servers.org" "Abidjan" "" "Cote d'Ivoire" "Africa"
"abj02.l.root-servers.org" "Abidjan" "" "Cote d'Ivoire" "Africa"
"akl01.l.root-servers.org" "Mangere" "" "New Zealand" "AsiaPacific"
"akl41.l.root-servers.org" "Mangere" "" "New Zealand" "AsiaPacific"
"akl42.l.root-servers.org" "Mangere" "" "New Zealand" "AsiaPacific"
"akl43.l.root-servers.org" "Mangere" "" "New Zealand" "AsiaPacific"
"akl44.l.root-servers.org" "Mangere" "" "New Zealand" "AsiaPacific"
"ams01.l.root-servers.org" "Haarlemmermeer" "" "Netherlands" "Europe"
"anc01.l.root-servers.org" "Anchorage" "Alaska" "United States" \
  "NorthAmerica"
%
```

5. Provisioning of IDENTITY.L.ROOT-SERVERS.ORG

Individual L-Root nodes run a dedicated, separate authority-only DNS server process which serves the IDENTITY.L.ROOT-SERVERS.ORG zone. The contents of that zone are unique to every node, and hence each responding node will generate a node-specific response.

The contents of the IDENTITY.L.ROOT-SERVERS.ORG zone are hence deliberately incoherent, the apparent zone contents depending on the node responding to the corresponding query.

The IDENTITY.L.ROOT-SERVERS.ORG zone is delegated to the single name server BEACON.L.ROOT-SERVERS.ORG, numbered on IPv4 and IPv6 addresses that are covered by the same routing advertisements that cover the L-Root service addresses. Reachability of BEACON.L.ROOT-SERVERS.ORG is hence well-aligned with the reachability of L.ROOT-SERVERS.NET, and hence measurement of the IDENTITY service ought to give similar results to measurement of the L-Root service.

It is considered best practice always to delegate a DNS zone to more than one name server; however, as described, the IDENTITY.L.ROOT-SERVERS.ORG zone is delegated to just one server. Ordinarily this would present a risk of failure if that single server is not available; however, given the purpose of the delegation in this case and that the expected mitigation of a failure in a single node is the routing of a query to a different node, delegation to a single server in this particular use-case is effective.

The L.ROOT-SERVERS.ORG zone is signed using DNSSEC, and hence secure responses for BEACON.L.ROOT-SERVERS.ORG and NODES.L.ROOT-SERVERS.ORG are available. IDENTITY.L.ROOT-SERVERS.ORG is an insecure delegation from the L.ROOT-SERVERS.ORG zone, however, following the operational preference to serve static data from each node for that zone, and the disinclination to distribute key materials and zone signing machinery to every node.

6. Security Considerations

Some operators of anycast services choose not to disclose locations (or even numbers) of nodes, citing security concerns. The operator of L-Root considers that none of the published information described in this document is truly secret, since any service element which provides service to the Internet can be can never truly be obscured from view. Given that location information can be found regardless of any conscious, deliberate disclosure, and since easy access to this information has diagnostic value, the operator of L-Root has adopted a policy of operational transparency.

The information presented in this document presents no new threat to the Internet.

7. IANA Considerations

This document makes no request of the IANA.

8. Acknowledgements

The aspects of the L-Root service that were deployed to facilitate IN-class mapping were discussed and implemented as part of an informal collaboration with Xun Fan, John Heidemann and Ramesh Govidan, whose contributions are acknowledged.

Helpful reviews and comments from Gaurab Upadhaya, Hugo Salgado, Brian Dixon, Bob Harold and Paul Hoffman on earlier versions of this document were very much appreciated.

9. References

9.1. Normative References

[RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, November 1987.

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[RFC2671] Vixie, P., "Extension Mechanisms for DNS (EDNS0)", RFC 2671, August 1999.

[RFC4786] Abley, J. and K. Lindqvist, "Operation of Anycast Services", BCP 126, RFC 4786, December 2006.

[RFC4892] Woolf, S. and D. Conrad, "Requirements for a Mechanism Identifying a Name Server Instance", RFC 4892, June 2007.

[RFC5001] Austein, R., "DNS Name Server Identifier (NSID) Option", RFC 5001, August 2007.

9.2. Informative References

[Fan2013] Fan, X., Heidemann, J., and R. Govidan, "Evaluating Anycast in the Domain Name System", Proceedings of the IEEE Infocom Turin, Italy, April 2013.

[RFC6304] Abley, J. and W. Maton, "AS112 Nameserver Operations", RFC 6304, July 2011.

URIs

[1] <http://www.iata.org/publications/Pages/coding.aspx>

Appendix A. Editorial Notes

This section (and sub-sections) to be removed prior to publication.

A.1. Change History

00 Initial idea, circulated for the purposes of entertainment.

01 Added some commentary of use-cases of NSID vs various/CH/TXT. Moved discussion of IN-class queries from the NODES section to the IDENTITY section. Added a note about DNSSEC for IDENTITY, NODES. Updated acknowledgements section.

02 Clarified re-routing impact on HOSTNAME.BIND, ID.SERVER, LOCATION.L queries vs. NSID as not just applying to HOSTNAME.BIND. Fixed typos and absurd malapropisms. Cleaned up prompts in command-line examples and added text to clarify how such examples should be interpreted.

Author's Address

Joe Abley
ICANN
12025 Waterfront Drive
Suite 300
Los Angeles, CA 90094-2536
USA

Phone: +1 519 670 9327
Email: joe.abley@icann.org