

Root Name Servers with Inter Domain Anycast Addresses

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

Abstract

This memo describes an operational guideline for millions of name servers to share an interdomain anycast address.

It enables people operate as many root name servers as they want and still make them traceable.

1. Motivation

DNS root servers are the essential component of the Internet that all the ISPs in the world want to run several root servers.

To satisfy them, we need to have thousands or millions of root servers.

However, because of the restriction on DNS message size over UDP, the number of unicast IP addresses of root servers is severely limited.

Thus, it is necessary to increase the number of root servers by

INTERNET DRAFT

Anycast Root Name Servers

November 2002

assigning an IP address to a lot of root servers.

Even if DNS were designed to allow a lot of root servers, it is difficult for DNS clients to choose the best (with regard to availability, credibility of zone content, delay, domain policy etc.) root servers among so many root servers. It is not practical to ping millions of servers to find which has the smallest RTT.

This memo proposes a mechanism of policy based selection of a root server sharing an IP address (anycast IP address) with other root servers and discusses operational issues related to it.

Because the selection is policy based, domain administrators have some control over the selection of the best root server among root servers sharing an IP address.

Note that operations similar to that described in this memo are possible today locally without global coordination by any operator who may be irritated by the lack of his control on (sufficiently many) root servers, which may be a source of some operational problems. This memo is an attempt to document the way to solve the problem in a least harmful manner.

Similar operation described in this memo may be applicable to gTLD or other global servers but it is outside the scope of this memo.

2. Suggested Operation

As is demonstrated by proliferated private use addresses, it is easy to set up routers to let unicast addresses have local scopes. It is also easy to let the unicast addresses have nested local scopes. The important difference between the addresses for private use and root servers is in their semantics that the root servers sharing an address also share the globally unique semantics of the address. The root servers may share a globally unique DNS host name, too.

A possible problem of such addresses is that the shared addresses can not be used for global communication.

So, the root name servers with the anycast addresses must have additional globally unique unicast address (or addresses), which may be used for global communication such as zone transfer.

The other possible problem of such addresses is that the shared addresses are not managed by a single entity that the mapping from the shared address of a root server to some operational entity is impossible.

However, if a router adjacent to (or near) the root server has a globally unique address, it is possible to map from the global address to an operational entity, which is expected to be operating the root server. That is, tools like "traceroute" work to uniquely identify the operational entity of the root servers sharing a anycast address.

To be compatible with the current practice that a single address belong to a single AS, each anycast address is assigned its own AS number. There will be multiple ASes of the AS number containing the same address ranges.

ASes, still, can be identified by adjacent ASes. For example, network operators may choose their favorite root server based on the AS numbers of the next hop ASes with, for example, AS path and BGP policy.

It is required that operators of an AS adjacent to the root servers' AS be fully responsible to the operation of the root servers. If a root server's AS is adjacent to multiple ASes, operators of all the ASes must be fully responsible to the operation of the root server. Thus, if there is a routing problem related to a root server, operators of the next hop AS(es) should be contacted.

To avoid complex routing tricks, globally unique unicast address(es) of the root name servers must be taken from the AS(es) adjacent to the root server's AS. Then, in a likely simple case that the root server's AS consists of a single host, which acts as a name server and a BGP router, the host should peer with adjacent AS(es) through an interface(s) address(es) of which belongs to the adjacent AS(es). If the root server's AS has more complex structure, special IGP arrangement of globally unique unicast address(es) is necessary in the AS and at the border router(s) of the adjacent AS(es). The border router(s) must accept IGP information advertised from the root server's AS.

3. Assignment

When a server with an anycast address fails but a route to it is still available, there is no way for clients use other servers with the same anycast address, anycast does not improve availability of servers so much.

So, even with anycast addresses, there should be multiple root servers.

However, as anycast solves the problem of load concentration, we don't need so many anycast IP addresses,

M. Ohta

Expires on May 1, 2003

[Page 3]

INTERNET DRAFT

Anycast Root Name Servers

November 2002

We should have at least 3 and at most 7 anycast addresses for root servers.

4. Ongoing Experiment

An experiment is ongoing with a block of IP addresses 192.83.230/24 in AS 4128 and DNS servers sharing an IP address of 192.83.230.1.

Domains served by the DNS servers are not root domain but are: real-internet.org. and psg.com.

If one is interested in joining (or just watching) the test, send a mail containing a single line of:

```
subscribe aroot
```

to

```
majordomo@ops.ietf.org
```

the mailing list is located at:

```
aroot@ops.ietf.org
```

Archive of the list is available at
<ftp://ops.ietf.org:/pub/lists/aroot>.

If there is something wrong with route information from AS 4128, the author of the memo or the mailing list for the test may be contacted.

However, for direct contact to the source of the problem, the contact person of an AS next to AS 4128 in the AS path of problematic route information should be contacted.

5. Security Considerations

This memo describes just an operational guideline with no protocol change. As such, the guideline does not introduce any security issues of the protocol level.

As the route forgery to the root servers can be implemented today without this memo by anyone including local intruders, the guideline does not introduce any security issues of the operational level, either.

A guideline to track down and verify a route or an AS path to a valid or a forged root server is described in [section 2](#).

M. Ohta

Expires on May 1, 2003

[Page 4]

INTERNET DRAFT

Anycast Root Name Servers

November 2002

Furthermore, if an ISP or a site operate its own anycast root server, hosts of the ISP or the site using the root server is protected from external forged route.

In addition, if a lot of local root servers share an anycast address, it reduce the effect of distributed denial of service attack on the anycast address.

6. Author's Address

Masataka Ohta
Graduate School of Information Science and Engineering
Tokyo Institute of Technology
2-12-1, O-okayama, Meguro-ku
Tokyo 152-8552, JAPAN

Phone: +81-3-5734-3299

Fax: +81-3-5734-3299

EEmail: mohta@necom830.hpcl.titech.ac.jp

