

IETF DNSOPS working group
Internet draft
Category: Work-in-progress

E.Hardie
Equinix, Inc
June 1999

[draft-hardie-dnsop-shared-root-server-00.txt](#)

Distributing Root Name Servers via Shared Unicast Addresses

Status of this memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC 2026](#).

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/ietf/lid-abstracts.txt>

To view the list Internet-Draft Shadow Directories, see <http://www.ietf.org/shadow.html>.

Copyright Notice

Copyright (C) The Internet Society 1999. All Rights Reserved.

Abstract

This memo describes a set of practices designed to enable the distribution of a root DNS server to multiple geographically and topologically distinct network locations. These practices presume that a single entity remains administratively and operationally responsible for each of the distributed servers.

The primary motivation for the development of these practices is to increase the availability of root servers. The current root servers already provide a highly distributed mesh, but the concentration of servers in U.S.-based networks limits their availability for users outside North America. These practices should enable the

distribution of root servers to areas not historically well-served by the current mesh without disrupting the operation of the DNS.

Discussion of methods like those outlined below date back at least to the December 1996 meeting of the IEPG. Recent discussions have taken place on the dnsop@cafax.se mailing list, and an internet draft "Root Name Servers Sharing Administratively Scoped Shared Unicast Addresses" was distributed there by Masataka Ohta.

[1. Architecture](#)

[1.1 Server Requirements](#)

In addition to meeting the host requirements for root servers listed in [1], each of the hosts should be configured with two network interfaces. One of the network interfaces should use the shared unicast address associated with the root name server. The other interface, referred to as the AS-internal interface below, should use a distinct address specific to that host. The host should respond to DNS queries only on the shared-unicast interface. The host should use the AS-internal interface and address for all mesh coordination.

[1.2 Zone file delivery](#)

In order to minimize the risk of man-in-the-middle attacks, zone files should be delivered to the AS-internal interface of the servers participating in the mesh. Secure file transfer methods and strong authentication should be used for all transfers.

[1.3 Synchronization](#)

As noted below in [section 3.2](#), lack of synchronization among servers could create problems for users of this service. In order to minimize the risk, switch-overs from one data set to another data set should be coordinated. The use of synchronized clocks on the participating hosts and set times for switch-overs provides a basic level of coordination. The full coordination process would involve transferring new data, checking for full receipt of data on all participating hosts, setting switch-over times for all participating hosts, and instituting a failure process to ensure that hosts which did not succeed in switching over ceased to respond to incoming queries.

[1.4 Server Placement](#)

Though the geographic diversity of server placement helps reduce the effects of service disruptions due to local problems, it is diversity of placement in the network topology which is the driving force behind these distribution practices. Server placement should

emphasize that diversity. Ideally, servers should be placed topologically near the points at which the operator exchanges routes and traffic with other networks.

[1.5](#) Routing

The organization administering the mesh of servers sharing a unicast address must have an autonomous system number and speak BGP to its peers. To those peers, the organization announces a route to the network containing the shared-unicast address of the root name server. The organization's border routers must then deliver the traffic destined for the root name server to the nearest instantiation. To avoid internal routing difficulties, a static route to that network is recommended. Routing to the AS-internal interfaces for the servers can use the normal routing methods for the administering organization, but care should be taken that traffic for the AS-internal interfaces does not leak onto the internal networks.

[Appendix A](#). contains an ASCII diagram of a simple implementation of this system. In it, the odd numbered routers deliver traffic to the shared-unicast interface network and filter traffic from the AS-internal network; the even numbered routers deliver traffic to the AS-internal network and filter traffic from the shared-unicast network. These are depicted as separate routers for the ease this gives in explanation, but they could easily be separate interfaces on the same router. Similarly, a local NTP source is depicted for synchronization, but the level of synchronization needed would not require that source to be either local or a stratum one NTP server.

[2](#). Administration

[2.1](#) Points of Contact

A single point of contact for reporting problems is crucial to the correct administration of this system. If an external user of the system needs to report a problem related to the service, there must be no ambiguity about whom to contact. If internal monitoring does not indicate a problem, the contact may, of course, need to work with the external user to identify which server generated the error.

[3](#). Security Considerations

As a core piece of internet infrastructure, the root servers are a common target of attack. The practices outlined here increase the risk of certain kinds of attack and reduce the risk of others.

[3.1](#) Increased Risks

As a first principal, it should be recognized that the architecture outlined in this document increases the number of physical servers acting as roots, which increases the possibility that a server mis-configuration will occur which allows for a security breach. If the mechanism used to distribute zone files among the servers is not well secured, a man-in-the-middle attack could result in the injection of false information. Digital signatures will alleviate this risk, but encrypted transport and tight access lists are a necessary adjunct to them.

A fundamental risk in the distribution of data using the methods outlined above is that the servers in the mesh will fall out of synch with one another. The use of ntp to provide a synchronized time for switch-over eliminates some aspects of this problem, but mechanisms to handle failure during the switchover are required. In particular, a server which cannot make the switchover must not roll-back to a previous version; it must cease to respond to queries so that other root servers are queried.

3.2 Decreased Risks

The increase in number of physical servers reduces, however, the likelihood that a denial-of-service attack will take out a significant portion of the DNS infrastructure. The increase in servers also reduces the effect of machine crashes, fiber cuts, and localized disasters by reducing the number of users dependent on on a specific machine.

4. Full copyright statement

Copyright (C) The Internet Society 1999. All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be

revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

5. Acknowledgements

Masataka Ohta, Bill Manning, Randy Bush, Chris Yarnell, Ray Plzak, Mark Andrews, Robert Elz, Geoff Houston, Bill Norton, and Akira Kato all provided input and commentary on this work.

[6]. References

1 "Root Name Server Operational Requirements", Randy Bush.
<ftp://ftp.ietf.org/internet-drafts/draft-bush-dnsop-root-opreq-00.txt>

7. Editor's address

Edward (Ted) Hardie
Equinix, Inc.
901 Marshall St.
Redwood City, CA 94063
hardie@equinix.com
Tel: 1.650.817.2226
Fax: 1.650.298.0420

Appendix A.



