Network Working Group	J. Abley
Internet-Draft	ICANN
Intended status: Informational	W. Maton
Expires: October 30, 2011	NRC-CNRC
	April 28, 2011

AS112 Nameserver Operations draft-ietf-dnsop-as112-ops-07

Abstract

Many sites connected to the Internet make use of IPv4 addresses that are not globally-unique. Examples are the addresses designated in RFC 1918 for private use within individual sites.

Devices in such environments may occasionally originate Domain Name System (DNS) queries (so-called "reverse lookups") corresponding to those private-use addresses. Since the addresses concerned have only local significance, it is good practice for site administrators to ensure that such queries are answered locally. However, it is not uncommon for such queries to follow the normal delegation path in the public DNS instead of being answered within the site.

It is not possible for public DNS servers to give useful answers to such queries. In addition, due to the wide deployment of private-use addresses and the continuing growth of the Internet, the volume of such queries is large and growing. The AS112 project aims to provide a distributed sink for such queries in order to reduce the load on the root and IN-ADDR.ARPA authoritative servers. The AS112 project is named after the Autonomous System Number (ASN) that was assigned to it. This document describes the steps required to install a new AS112 node, and offers advice relating to such a node's operation.

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 http://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 October 30, 2011.

Copyright Notice

Copyright (c) 2011 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 (http://trustee.ietf.org/licenseinfo) 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. This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

- *1. Introduction
- *2. AS112 DNS Service
- *2.1. Zones
- *2.2. <u>Nameservers</u>
- *3. <u>Installation of a New Node</u>
- *3.1. Useful Background Knowledge
- *3.2. Topological Location
- *3.3. Operating System and Host Considerations
- *3.4. Routing Software
- *3.5. DNS Software
- *3.6. Testing a Newly-Installed Node
- *4. Operations
- *4.1. Monitoring
- *4.2. Downtime
- *4.3. Statistics and Measurement

- *5. Communications
- *6. On the Future of AS112 Nodes
- *7. IANA Considerations
- *8. Security Considerations
- *9. Acknowledgements
- *10. References
- *10.1. <u>Normative References</u>
- *10.2. <u>Informative References</u>
- *Appendix A. <u>History</u>
- *Appendix B. Change History
- *Authors' Addresses

1. Introduction

Many sites connected to the Internet make use of IPv4 addresses that are not globally unique. Examples are the addresses designated in [RFC1918] for private use within individual sites.

Devices in such environments may occasionally originate Domain Name System (DNS) [RFC1034] queries (so-called "reverse lookups") corresponding to those private-use addresses. Since the addresses concerned have only local significance, it is good practice for site administrators to ensure that such queries are answered locally [I-D.ietf-dnsop-default-local-zones]. However, it is not uncommon for such queries to follow the normal delegation path in the public DNS instead of being answered within the site.

It is not possible for public DNS servers to give useful answers to such queries. In addition, due to the wide deployment of private-use addresses and the continuing growth of the Internet, the volume of such queries is large and growing. The AS112 project aims to provide a distributed sink for such queries in order to reduce the load on the root and IN-ADDR.ARPA authoritative servers.

The AS112 project encompasses a loosely coordinated collection of independently operated nameservers. Each nameserver functions as a single node in an AS112 <u>anycast cloud</u> [RFC4786], and is configured to answer authoritatively for a particular set of nominated zones. The AS112 project is named after the Autonomous System Number (ASN) that was assigned to it.

2. AS112 DNS Service

2.1. Zones

AS112 nameservers answer authoritatively for the following zones, corresponding to [RFC1918] private-use netblocks: [RFC5735]:

*10.IN-ADDR.ARPA

*16.172.IN-ADDR.ARPA, 17.172.IN-ADDR.ARPA, ..., 31.172.IN-ADDR.ARPA

*168.192.IN-ADDR.ARPA

and the following zone, corresponding to the "link local" netblock 169.254.0.0/16 listed in

*254.169.IN-ADDR.ARPA

To aid identification of AS112 anycast nodes, each node also answers authoritatively for the zone HOSTNAME.AS112.NET.

See <u>Section 3.5</u> for the recommended contents of all these zones. It is possible that other zones corresponding to private-use infrastructure will be delegated to AS112 servers in the future. A list of zones for which AS112 servers answer authoritatively is maintained at http://www.as112.net/.

2.2. Nameservers

The zones listed in <u>Section 2.1</u> are delegated to the two nameservers BLACKHOLE-1.IANA.ORG (192.175.48.6) and BLACKHOLE-2.IANA.ORG (192.175.48.42).

Additionally, the server PRISONER.IANA.ORG (192.175.48.1) is listed in the MNAME field of the SOA records of the IN-ADDR.ARPA zones served by AS112 nameservers. PRISONER.IANA.ORG receives mainly dynamic update queries.

The addresses of all these nameservers are covered by the single IPv4 prefix 192.175.48.0/24.

3. Installation of a New Node

3.1. Useful Background Knowledge

Installation of an AS112 node is relatively straightforward. However, experience in the following general areas may prove useful:

*inter-domain routing with BGP [RFC4271];

*DNS authoritative server operations;

*anycast [RFC4786] distribution of DNS services.

3.2. Topological Location

AS112 nodes may be located anywhere on the Internet. For nodes that are intended to provide a public service to the Internet community (as opposed to private use), it may well be advantageous to choose a location that is easily (and cheaply) reachable by multiple providers, such as an Internet exchange point.

AS112 nodes may advertise their service prefix to BGP peers for local use (analogous to a conventional peering relationship between two providers) or for global use (analogous to a customer relationship with one or more providers).

It is good operational practice to notify the community of users that may fall within the reach of a new AS112 node before it is installed. At an Internet Exchange, local mailing lists usually exist to facilitate such announcements. For nodes that are intended to be globally reachable, coordination with other AS112 operators is highly recommended. See also Section 5.

3.3. Operating System and Host Considerations

Examples in this document are based on UNIX and UNIX-like operating systems, but other operating systems exist which are suitable for use in construction of an AS112 node.

The chosen platform should include support for either cloned loopback interfaces, or the capability to bind multiple addresses to a single loopback interface. The addresses of the nameservers listed in Section2.2 will be configured on these interfaces in order that the DNS software can respond to queries properly.

A host that is configured to act as an AS112 anycast node should be dedicated to that purpose, and should not be used to simultaneously provide other services.

System startup scripts should be arranged such that the various AS112-related components start automatically following a system reboot. The order in which interfaces are configured and software components started should be arranged such that routing software startup follows DNS software startup, and DNS software startup follows loopback interface configuration.

Wrapper scripts or other arrangements should be employed to ensure that the anycast service prefix for AS112 is not advertised while either the anycast addresses are not configured, or while the DNS software is not running.

3.4. Routing Software

AS112 nodes signal the availability of AS112 nameservers to the Internet using BGP [RFC4271]: each AS112 node is a BGP speaker, and

announces the prefix 192.175.48.0/24 to the Internet with origin AS 112 (see also Section 2.2).

The examples in this document are based on the <u>Quagga Routing Suite</u>, but other software packages exist which also provide suitable BGP support for AS112 nodes.

The "bgpd.conf" file is used by Quagga's bgpd daemon, which provides BGP protocol support. The router id in this example is 203.0.113.1; the AS112 node peers with external peers 192.0.2.1 and 192.0.2.2. Note the local AS number 112, and the origination of the prefix 192.175.48.0/24.

```
! bgpd.conf
!
hostname as112-bgpd
password <something>
enable password <supersomething>
!
! Note that all AS112 nodes use the local Autonomous System
! Number 112, and originate the IPv4 prefix 192.175.48.0/24.
! All other addresses shown below are illustrative, and
! actual numbers will depend on local circumstances.
!
router bgp 112
bgp router-id 203.0.113.1
network 192.175.48.0
neighbor 192.0.2.1 remote-as 64496
neighbor 192.0.2.2 remote-as 64497
neighbor 192.0.2.2 remote-as 64497
neighbor 192.0.2.2 next-hop-self
```

The "zebra.conf" file is required to provide integration between protocol daemons (bgpd, in this case) and the kernel.

```
! zebra.conf
!
hostname as112
password <something>
enable password <supersomething>
!
interface lo
!
interface eth0
```

3.5. DNS Software

Although the queries received by AS112 nodes are definitively misdirected, it is important that they be answered in a manner that is

accurate and consistent. For this reason AS112 nodes operate as <u>fully-functional</u> and <u>standards-compliant DNS authoritative servers</u> [RFC1034], and hence require DNS software.

Suitable choices of free DNS software for AS112 nodes include, but are not limited to:

*ISC BIND9

*NLnet Labs' NSD

Examples in this document are based on ISC BIND9. The following is a sample BIND9 "named.conf" file for a dedicated AS112 server. Note that the nameserver is configured to act as an authoritative-only server (i.e. recursion is disabled). The nameserver is also configured to listen on the various AS112 anycast nameserver addresses, as well as its local addresses.

```
// named.conf
// global options
options {
 listen-on {
                // localhost
   127.0.0.1;
 // the following address is node-dependent, and should be set to
 // something appropriate for the new AS112 node
                      // local address (globally-unique, unicast)
    203.0.113.1;
 // the following addresses correspond to AS112 addresses, and
  // are the same for all AS112 nodes
   192.175.48.1;
                    // prisoner.iana.org (anycast)
                    // blackhole-1.iana.org (anycast)
   192.175.48.6;
   192.175.48.42;
                     // blackhole-2.iana.org (anycast)
  };
  directory "/var/named";
  recursion no;
                      // authoritative-only server
  query-source address *;
};
// log queries, so that when people call us about unexpected
// answers to queries they didn't realise they had sent, we
// have something to talk about. Note that activating this
// has the potential to create high CPU load and consume
// enormous amounts of disk space.
logging {
 channel "querylog" {
   file "/var/log/query.log" versions 2 size 500m;
   print-time yes;
 };
  category queries { querylog; };
};
// RFC 1918
zone "10.in-addr.arpa" { type master; file "db.empty"; };
zone "16.172.in-addr.arpa" { type master; file "db.empty"; };
zone "17.172.in-addr.arpa" { type master; file "db.empty"; };
zone "18.172.in-addr.arpa" { type master; file "db.empty"; };
zone "19.172.in-addr.arpa" { type master; file "db.empty"; };
zone "20.172.in-addr.arpa" { type master; file "db.empty"; };
zone "21.172.in-addr.arpa" { type master; file "db.empty"; };
zone "22.172.in-addr.arpa" { type master; file "db.empty"; };
zone "23.172.in-addr.arpa" { type master; file "db.empty"; };
```

```
zone "24.172.in-addr.arpa" { type master; file "db.empty"; };
zone "25.172.in-addr.arpa" { type master; file "db.empty"; };
zone "26.172.in-addr.arpa" { type master; file "db.empty"; };
zone "27.172.in-addr.arpa" { type master; file "db.empty"; };
zone "28.172.in-addr.arpa" { type master; file "db.empty"; };
zone "29.172.in-addr.arpa" { type master; file "db.empty"; };
zone "30.172.in-addr.arpa" { type master; file "db.empty"; };
zone "31.172.in-addr.arpa" { type master; file "db.empty"; };
zone "168.192.in-addr.arpa" { type master; file "db.empty"; };
// RFC 5735
zone "254.169.in-addr.arpa" { type master; file "db.empty"; };
// also answer authoritatively for the HOSTNAME.AS112.NET zone,
// which contains data of operational relevance
zone "hostname.as112.net" {
  type master;
  file "db.hostname.as112.net";
};
The "db.empty" file follows, below. This is the source data used to
populate all the IN-ADDR.ARPA zones listed in Section 2.1. Note that
the RNAME specified in the SOA record corresponds to hostmaster@root-
servers.org, a suitable e-mail address for receiving technical queries
about these zones.
; db.empty
; Empty zone for AS112 server.
$TTL
        1W
 IN SOA prisoner.iana.org. hostmaster.root-servers.org. (
                                       ; serial number
                                       ; refresh
                               1M
                                       ; retry
                                       ; expire
                               1W
                                      ; negative caching TTL
       NS
              blackhole-1.iana.org.
       NS
              blackhole-2.iana.org.
; There should be no other resource records included in this zone.
; Records that relate to RFC 1918-numbered resources within the
; site hosting this AS112 node should not be hosted on this
; nameserver.
```

The "db.hostname.as112.net" file follows, below. This zone contains various resource records that provide operational data to users for troubleshooting or measurement purposes, and should be edited to suit local circumstances. Note that the response to the query "HOSTNAME.AS112.NET IN TXT" should fit within a 512 octet DNS/UDP datagram: i.e. it should be available over UDP transport without requiring EDNS0 support.

The optional <u>LOC record</u> [RFC1876] included in the zone apex provides information about the geospatial location of the node.

```
; db.hostname.as112.net
$TTL
        1W
        SOA
                server.example.net. admin.example.net. (
                                         ; serial number
                        1
                                         ; refresh
                        1W
                                        ; retry
                        1M
                                        ; expire
                        1W
                                        ; negative caching TTL
                        1W )
        NS
                blackhole-2.iana.org.
                blackhole-1.iana.org.
        NS
                "Name of Facility or similar" "City, Country"
        TXT
                "See http://www.as112.net/ for more information."
        TXT
        LOC
                45 25 0.000 N 75 42 0.000 W 80.00m 1m 10000m 10m
```

3.6. Testing a Newly-Installed Node

The BIND9 tool "dig" can be used to retrieve the TXT resource records associated with the domain "HOSTNAME.AS112.NET", directed at one of the AS112 anycast nameserver addresses. Continuing the example from above, the response received should indicate the identity of the AS112 node that responded to the query. See Section 3.5 for more details about the resource records associated with "HOSTNAME.AS112.NET".

```
% dig @prisoner.iana.org hostname.as112.net txt +short +norec "Name of Facility or similar" "City, Country" "See http://www.as112.net/ for more information." %
```

If the response received indicates a different node is being used, then there is probably a routing problem to solve. If there is no response received at all, there might be host or nameserver problem. Judicious use of tools such as traceroute, and consultation of BGP looking glasses might be useful in troubleshooting.

Note that an appropriate set of tests for a new server will include queries sent from many different places within the expected service area of the node, using both UDP and TCP transport, and exercising all three AS112 anycast nameserver addresses.

4. Operations

4.1. Monitoring

AS112 nodes should be monitored to ensure they are functioning correctly, just as with any other production service. An AS112 node that stops answering queries correctly can cause failures and timeouts in unexpected places and can lead to failures in dependent systems that can be difficult to troubleshoot.

4.2. Downtime

An AS112 node that needs to go off-line (e.g. for planned maintenance or as part of the diagnosis of some problem) should stop advertising the AS112 service prefix to its BGP peers. This can be done by shutting down the routing software on the node altogether or by causing the routing system to withdraw the route.

Withdrawing the service prefix is important in order to avoid blackholing query traffic in the event that the DNS software on the node is not functioning normally.

4.3. Statistics and Measurement

Use of the AS112 node should be measured in order to track long-term trends, identify anomalous conditions, and to ensure that the configuration of the AS112 node is sufficient to handle the query load. Examples of free monitoring tools that might be useful to operators of AS112 nodes include, but are not limited to:

- *bindgraph
- *dnstop
- *DSC

5. Communications

It is good operational practice to notify the community of users that may fall within the reach of a new AS112 node before it is installed. At Internet Exchanges, local mailing lists usually exist to facilitate such announcements.

For nodes that are intended to be globally reachable, coordination with other AS112 operators is especially recommended. The mailing list mailto:as112-ops@lists.dns-oarc.net is operated for this purpose.

Information pertinent to AS112 operations is maintained at http://www.as112.net/.

Information about an AS112 node should also be published within the DNS, within the "HOSTNAME.AS112.NET" zone. See Section 3.5 for more details.

6. On the Future of AS112 Nodes

It is recommended practice for the operators of recursive nameservers to answer queries for zones served by AS112 nodes locally, such that queries never have an opportunity to reach AS112 servers [I-D.ietf-dnsop-default-local-zones]. Operational experience with AS112 nodes does not currently indicate an observable trend towards compliance with those recommendations, however.

It is expected that some DNS software vendors will include default configuration that will implement measures such as those described in [I-D.ietf-dnsop-default-local-zones]. If such software is widely deployed, it is reasonable to assume that the query load received by AS112 nodes will decrease; however, it is safe to assume that the query load will not decrease to zero, and consequently that AS112 nodes will continue to provide a useful service for the foreseeable future. There may be a requirement in the future for AS112 nodes to answer for their current set of zones over IPv6 transport. Such a requirement would necessitate the assignment of a corresponding IPv6 netblock for use as an anycast service prefix.

There may be a requirement in the future for AS112 nodes to serve additional zones, or to stop serving particular zones that are currently served. Such changes would be widely announced in operational forums, and published at http://www.as112.net/.

7. IANA Considerations

The AS112 nameservers are all named under the domain IANA.ORG (see <u>Section 2.2</u>). However, the anycast infrastructure itself is operated by a loosely-coordinated, diverse mix of organisations across the Internet, and is not an IANA function.

The autonomous system number 112 and the IPv4 prefix 192.175.48.0/24 were assigned by ARIN.

This document makes no request of the IANA.

8. Security Considerations

Hosts should never normally send queries to AS112 servers; queries relating to private-use addresses should be answered locally within a site. Hosts that send queries to AS112 servers may well leak information relating to private infrastructure to the public network, and this could present a security risk. This risk is orthogonal to the presence or absence of authoritative servers for these zones in the public DNS infrastructure, however.

Queries that are answered by AS112 servers are usually unintentional; it follows that the responses from AS112 servers are usually unexpected. Unexpected inbound traffic can trigger intrusion detection systems or alerts by firewalls. Operators of AS112 servers should be prepared to be contacted by operators of remote infrastructure who believe their security has been violated. Advice to those who mistakenly believe that responses from AS112 nodes constitutes an attack on their infrastructure can be found in [I-D.ietf-dnsop-as112-under-attack-help-help].

The deployment of AS112 nodes is very loosely coordinated compared to other services distributed using anycast. The malicious compromise of an AS112 node and subversion of the data served by the node is hence more difficult to detect due to the lack of central management. Since it is conceivable that changing the responses to queries received by AS112 nodes might influence the behaviour of the hosts sending the queries, such a compromise might be used as an attack vector against private infrastructure.

Operators of AS112 should take appropriate measures to ensure that AS112 nodes are appropriately protected from compromise, such as would normally be employed for production nameserver or network infrastructure. The guidance provided for root nameservers in [RFC2870] may be instructive.

The zones hosted by AS112 servers are not signed with DNSSEC [RFC4033]. Given the distributed and loosely-coordinated structure of the AS112 service, the zones concerned could only be signed if the private key material used was effectively public, obviating any security benefit resulting from the use of those keys.

9. Acknowledgements

The authors wish to acknowledge the assistance of Bill Manning, John Brown, Marco D'Itri, Daniele Arena, Stephane Bortzmeyer, Frank Habicht, Chris Thompson, Peter Losher, Peter Koch and Alfred Hoenes in the preparation of this document.

10. References

10.1. Normative References

[RFC1034]	Mockapetris, P., " <u>Domain names - concepts and</u> <u>facilities</u> ", STD 13, RFC 1034, November 1987.	
[RFC1918]	Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G. and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.	
[RFC2870]	Bush, R., Karrenberg, D., Kosters, M. and R. Plzak, "Root Name Server Operational Requirements", BCP 40, RFC 2870, June 2000.	
[RFC4033]		

	Arends, R., Austein, R., Larson, M., Massey, D. and S.		
	Rose, "DNS Security Introduction and Requirements", RFC		
	4033, March 2005.		
[RFC4271]	Rekhter, Y., Li, T. and S. Hares, " <u>A Border Gateway</u>		
	<u>Protocol 4 (BGP-4)</u> ", RFC 4271, January 2006.		
[RFC4786]	Abley, J. and K. Lindqvist, "Operation of Anycast		
	Services", BCP 126, RFC 4786, December 2006.		

10.2. Informative References

[RFC1876]	Davis, C., Vixie, P., Goodwin, T. and I. Dickinson, "A Means for Expressing Location Information in the Domain Name System", RFC 1876, January 1996.
[RFC5735]	Cotton, M. and L. Vegoda, " <u>Special Use IPv4</u> <u>Addresses</u> ", BCP 153, RFC 5735, January 2010.
<pre>[I-D.ietf-dnsop- default-local- zones]</pre>	Andrews, M, " <u>Locally-served DNS Zones</u> ", Internet-Draft draft-ietf-dnsop-default-local- zones-15, March 2011.
[I-D.ietf-dnsop- as112-under- attack-help- help]	Abley, J and W Maton, "I'm Being Attacked by PRISONER.IANA.ORG!", Internet-Draft draft-ietf-dnsop-as112-under-attack-help-help-06, April 2011.

Appendix A. History

Widespread use of the private address blocks listed in [RFC1918] followed that document's publication in 1996.

The idea of off-loading IN-ADDR.ARPA queries relating to [RFC1918] addresses from the root nameservers was first proposed by Bill Manning and John Brown.

The use of anycast for distributing authoritative DNS service for [RFC1918] IN-ADDR.ARPA zones was subsequently proposed at a private meeting of root server operators.

ARIN provided an IPv4 prefix for the anycast service, and also the autonomous system number 112 for use in originating that prefix. This assignment gave the project its name.

In 2002, the first AS112 anycast nodes were deployed.

The use of anycast nameservers in the AS112 project contributed to the operational experience of anycast DNS services, and can be seen as a precursor to the anycast distribution of other authoritative DNS servers in subsequent years (e.g. various root servers).

<u>Appendix B.</u> Change History

This section to be removed prior to publication.

00 Initial draft, circulated as draft-jabley-as112-ops-00 and reviewed at the DNSOP working group meeting at IETF 66.

- 00
- Document adoped by the DNSOP working group and renamed accordingly.
- 01 Input from reviewers of DNSOP and others, some cosmetic tweaks.
- **02** Version bump as request by DNSOP chairs. Added missing IANA Considerations section. Updated author's addresses. Make http://www.as112.net/ URL consistent.
- 03 Fix BLACKHOLE-2.IANA.ORG IP address.
- **04** Bump version number. Refresh references. Add reference to BIRD. Minor wordsmithing.
- 05 Updated following review from Peter Koch.
- 06 Updated following review from Alfred Hoenes.

Authors' Addresses

Joe Abley Abley ICANN 4676 Admiralty Way, Suite 330 Marina del Rey, CA 90292 US Phone: +1 519 670 9327 EMail: joe.abley@icann.org

William F. Maton Sotomayor Maton National Research Council of Canada 1200 Montreal Road Ottawa, ON K1A 0R6 Canada Phone: +1 613 993 0880 EMail: wmaton@ryouko.imsb.nrc.ca