                 Operational Criteria for Root Name Servers

Status of this Memo

Abstract

   This document specifies the operational requirements of root name
   servers, including host hardware capacities, name server software
   revisions, network connectivity, and physical environment.

1 - Rationale and Scope

   1.1. Historically, the name servers responsible for the root (".")
   zone have also been responsible for all international top-level
   domains (iTLD's, for example: COM, EDU, INT, ARPA). These name
   servers have been operated by a cadre of highly capable volunteers,
   and their administration has been loosely coordinated by the NIC
   (first SRI-NIC and now InterNIC). Ultimate responsibility for the
   correct operation of these servers and for the content of the DNS
   zones they served has always rested with the IANA.

   1.2. As described in [Postel96], many new TLD's may be created
   shortly. Servers for all new and existing iTLD's will be subject to
   the operational requirements given in [Postel96]. The set of servers
   for the root (".") zone is likely to become disjoint from the set of
   servers for any TLD or group of TLD's, including those maintained by
   the InterNIC.

1.3. In spite of the similarities in operational requirements between
the servers for the iTLD's and the servers for the root (".") zone,
they are in fact different server sets with different administrators
and slightly different operational requirements. It is likely that
many contry code tld servers will have even more divergent
operational requirements. That said, the requirements set down in
this document could be successfully applied to any name server
(whether root, top level, or any other level), but may be more
draconian than necessary for servers other than those of the root
(".") zone.

    Disclaimer: The selection of name server locations and
                  administrators, and the procedures for addressing
                  noncompliance with these stated operational
                  requirements, are outside the scope of this document.

    Definition: For the purpose of this document, the term "zone master"
                  shall be used to designate the administrative owner of
                  the content of a zone. This person is expected to have
                  final responsibility for the selection and correct
                  operation of all of the zone's servers. For the root
                  (".") zone, this is the IANA.

2 - Operational Requirements

2.1. Name server software. The zone master shall initially and
periodically choose a name server package to run on all of the zone's
servers. It is expected that the BIND server will be used, at least
initially, and that new versions or other servers will be specified
from time to time.

    Rationale: This requirement is based on the wide and free
                availability of BIND's source code, and the active
                analysis and development it constantly receives from
                several members of the IETF.

Name server software upgrades will be specified and scheduled by the
zone master, and must occur on all of a zone's servers within a
specified 96 hour window.

    Rationale: In some cases it has proven necessary to "cold start" a
                zone's servers in order to clear out oscillating bad
                data. By forcing all software upgrades to happen at
                about the same time, it will be possible to coordinate
                a software change with a zone content change.

2.2. UDP checksums. UDP checksums must be generated when sending
datagrams, and verified when receiving them.

   Rationale: Some vendors turn off UDP checksums for performance
                reasons, citing the presence of MAC-level frame checks
                (CRC, for example) as "strong enough." This has been
                a disaster in actual practice.

2.3. Dedicated host. A name server host should have no other
function, and no login accounts other than for system or network
administrators. No other network protocols should be served by a
name server host (e.g., SMTP, NNTP, FTP, et al). If login is
permitted from other than the system console, then the login service
must be by encrypted channel (e.g., Kerberized and encrypted
rlogin/telnet, the secure shell (SSH), or an equivilent).

   Rationale: Each additional service performed by a host makes it
                less reliable and potentially less secure, as well as
                complicating fault isolation procedures. While name
                service does not consume very much in the way of system
                resources, it is thought best that a host do a few
                things well rather than many things poorly.

2.4. Clock synchronization. A name server host should synchronize
its clock using the NTP protocol (currnet version) with
authentication. At least two NTP servers should be used. As an
exception to section 2.3 above, a name server host can be an NTP
server as well.

   Rationale: For distributed fault isolation reasons, synchronized
                time stamps in system event logs are quite helpful.
                NTP is easily spoofed by UDP blast attacks, thus the
                requirement for authentication between the name server
                host and its NTP servers. A name server host is
                allowed to be an NTP server because it has been
                observed that a single host running both name service
                and stratum 1 NTP is still quite reliable and secure.

2.5. Network interfaces. Name servers must send UDP responses with
an IP source address (and UDP source port number) equal to the IP
destination address (and UDP destination port number) of the request.
Also, a name server might have multiple real interfaces, but only one
will be advertised in the zone's NS RRset and associated glue A RRs.
The advertised address should be that of the "best" interface on the
host, in terms of network performance and reliability to the largest
number of destinations.

       Rationale: While not required by [RFC1035], many extant DNS
                   implementations require the source address and port of
                   a reply to match the destination address and port to
                   which the request was sent. The number of advertised
                   addresses is limited to one (1) so that DNS delegation
                   responses containing this name server can be as short
                   as possible.

   2.6. Physical environment. A name server host must be located in a
   secure space such as a locked computer room or a data center with
   restricted access. The power supply should be redundant, using
   batteries, generators or some other means to protect against utility
   power failures. Network connectivity should be redundant, so that a
   single wide area line failure cannot completely isolate the name
   server host from the rest of the network.

   2.7. Network security. The system and network administrators should
   educate themselves about potential threats, and stay current on CERT
   bulletins regarding network breakins. The system staff should
   periodically audit the name server host's activity logs and be able
   to detect breakins during or after the fact.

   2.8. Host performance. As of the time of this writing, a name server
   must be able to answer 1,200 UDP transactions per second with less
   than 5 milliseconds of average latency. Because the network is still
   growing at a high rate, the ability to grow to 2,000 transactions per
   second and still support a 5 millisecond latency is highly desirable.
   Note that this requirement affects both the host and the network
   infrastructure to which that host is attached.

   2.9. Response time. The administrators responsible for a name server
   will respond to e-mail trouble reports within 24 hours. Personnel
   issues such as vacations and illness will cause responsibilities to
   be delegated and/or reassigned rather than ignored. After hours
   telephone numbers must be made available to the zone master for
   nonpublished use in emergencies. An escalation contact name, e-mail
   address, and telephone number will also be made available to the zone
   master in the event of nonresponse through the normal channel.

   2.10. Zone transfer access control. The name server shall be
   configured so that outbound zone transfers are permitted only to
   destinations on the server's local networks, and to whichever
   networks the zone master designates for remote debugging purposes.

        Rationale: Zone transfers can present a significant load on a name
                    server, especially if several transfers are started
                    simultaneously against the same server. There is no
                    operational reason to allow anyone outside the name
                    server's and zone's administrators to transfer the
                    entire zone.

   2.11. Zone transfer protocol. DNS AXFR shall be used in preference
   to FTP or any other non-DNS transfer protocol. DNS NOTIFY (see
   [NOTIFY]) and DNS IXFR (see [IXFR]) shall be supported and enabled
   when available.

        Rationale: Historically, the common implementations of DNS
                    (a.k.a., BIND) did not support zone transfer of the
                    root (".") zone due to programming errors. Thus, FTP
                    was used. In the future, DNS implementations which do
                    not support zone transfer of all zones will not be
                    considered suitable for use as root name servers. The
                    benefits of [IXFR] and [NOTIFY] should be obvious.

   2.12. Recursion shall be disabled for queries.

        Rationale: Recursion is a major source of cache pollution, and can
                    be a major drain on name server performance. An
                    organization's recursive DNS needs should be served by
                    some other host than its root name server(s). An
                    exception is made for missing glue since it's possible
                    that glue needed for some delegations will not be
                    within or beneath any zone for which the server is
                    authoritative. Such glue must be fetched via
                    recursive lookups to other servers.

   2.13. Outages shall be reported. All outages, scheduled or not,
   shall be reported to the zone master via e-mail. If an outage is
   unscheduled or if an outage is scheduled less than 24 hours in
   advance, then an additional notification of the zone master shall be
   made via telephone. Extended or repeated outages may beget special
   handling by the zone master.

   2.14. Inverse name lookups. The PTR RR associated with a server's
   primary interface address (that is, the address shown in in the
   zone's delegation) shall have its target specified by the zone
   master.

        Rationale: Since each organization has local control of their
                    network's PTR RRs, and since it is necessary for the
                    correct operation of some software that the forward and
                    reverse lookups have symmetrical results, it is left
                    up to the zone master to select the name for each
                    authority server's primary address.

3 - Possible Selection Criteria

    3.1. Host population. A server's location on the network should be
    such that it has a low IP hop count to a high number of end hosts.
    Duplication of service should be avoided, such that any given set of
    end hosts needs to have a low IP hop count to at most one authority
    server for any given zone.

    3.2. Infrastructure diversity. A server's location on the network
    should be such that most failures capable of isolating it from a
    large number of end hosts are diverse from the failures capable of
    similarly isolating other authority servers for the same zone(s).

4 - Security Considerations

    See section 2.7.

5 - References

    [RFC1035]
        Mockapetris, P., "Domain Names - Implementation and Specification",
        STD 13, RFC 1035, USC/Information Sciences Institute, November
        1987.

    [Postel96]
        Postel, J., "New Registries and the Delegation of International Top
        Level Domains", Work in Progress.

    [IXFR]
        Ohta, M., "Incremental Zone Transfer", RFC 1995, August 1996.

    [NOTIFY]
        Vixie, P., "A Mechanism for Prompt Notification of Zone Changes",
        RFC 1996, August 1996.

6 - Acknowledgements

    Constructive comments have been received from: Jon Postel, Michael
    Patton, Andrew Partan, Michael Dillon, Don Mitchell Steven Doyle,
    Owen DeLong and other members of the internet community.

7 - Authors' Addresses

     Bill Manning
     USC/ISI
     4676 Admiralty Way
     Marina del Rey, CA 90292

     Phone: +1 310 822 1511
     EMail: bmanning@isi.edu

     Paul Vixie
     Internet Software Consortium
     Star Route Box 159A
     Woodside, CA 94062

     Phone: +1 415 747 0204
     EMail: paul@vix.com