

Internet Engineering Task Force
INTERNET-DRAFT
Expires: July 15, 2004

Motonori Nakamura
Kyoto University
Jun-ichiro itojun Hagino
IIJ Research Laboratory
January 15, 2004

SMTP operational experience in mixed IPv4/IPv6 environements
draft-ietf-ngtrans-ipv6-smtp-requirement-08.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with 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.''

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

Distribution of this memo is unlimited.

The internet-draft will expire in 6 months. The date of expiration will be July 15, 2004.

Abstract

This document talks about SMTP operational experiences in IPv4/v6 dual stack environments. As IPv6-capable SMTP servers are deployed, it has become apparent that certain configurations are necessary in IPv6-capable MX DNS records for stable dual-stack (IPv4 and IPv6) SMTP operation. This document clarifies the problems that exist in the transition period between IPv4 SMTP and IPv6 SMTP. It also defines operational requirements for stable IPv4/v6 SMTP operation.

This document does not define any new protocol.

1. Introduction

Deliveries of mail messages to the final mail drop is not always done by direct IP communication with submitter and final receiver, and there may be some intermediate hosts to relay the messages. So it is difficult to know at message submission (also at receiver side) that all intermediate relay hosts are properly configured. It is not so easy to configure all the system with consistency since mail message delivery system is rather complex on DNS setting than other Internet services. For the transition state from IPv4 to IPv6, both IPv4 and IPv6 interoperability should be kept more carefully.

There are several technologies defined for the transition from IPv4 to IPv6. This document concentrates on SMTP issues in a dual-stack environment, e.g. delivery of emails from IPv6-only MTA to IPv4-only MTA is outside of the scope of the document.

IPv6-only MTA to IPv4-only MTA case could use help from IPv6-to-IPv4 translators such as [Hagino, 2001], however, there are no special SMTP considerations for translators needed; If there is SMTP traffic from an IPv6 MTA to an IPv4 MTA over an IPv6-to-IPv4 translator, the IPv4 MTA will consider this as a normal IPv4 SMTP traffic. Protocols like IDENT [St.Johns, 1993] as well as SMTP HELO/EHLO may require special consideration when translators are used.

The following sections explain how to make IPv4 SMTP and IPv6 SMTP coexist in a dual-stack environment.

This document does not discuss the problems encountered when the sending MTA and the receiving MTA have no common protocol (e.g. the sending MTA is IPv4-only while the receiving MTA is IPv6-only). Such a situation should be resolved by making either side dual-stack or by making either side use a protocol translator.

2. Basic DNS resource record definitions for mail routing

Mail messages on the Internet are delivered based on domain name system generally. MX RRs are looked up to know destination hosts associated with domain part of a mail address. Similar to the way RFC's for IPv6 DNS lookup [Thomson, 1995] use IN class for both IPv4 and IPv6, IN MX records will be used for both IPv4 and IPv6 on mail message routing, hosts which have IPv6 transport and want to be delivered with the IPv6 transport must define IPv6 IP addresses for the host name as well as IPv4 IP addresses.

A MX RR has two data, a preference value and the name of destination host. IP addresses for the destination host are also looked up to make

SMTP transport [Partridge, 1986] . In IPv4 environment, IPv4 IP addresses are defined with A RRs.

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 2]

For example, IPv6 only site may have the following DNS definitions:

```
example.org.          IN MX   1  mx1.example.org.
                     IN MX  10  mx10.example.org.
mx1.example.org.      IN AAAA  3ffe:501:ffff::1
mx10.example.org.     IN AAAA  3ffe:501:ffff::2
```

In transition period from IPv4 to IPv6, there are many IPv4 sites, and such sites will not have mail interoperability with IPv6 only sites. For the transition period, all mail domains should have MX records such that MX targets with IPv4 and IPv6 addresses exist, e.g, for example:

```
example.org.          IN MX   1  mx1.example.org.
                     IN MX  10  mx10.example.org.
mx1.example.org.      IN AAAA  3ffe:501:ffff::1
                     IN A     192.0.2.1
mx10.example.org.     IN AAAA  3ffe:501:ffff::2
                     IN A     192.0.2.2
```

But, every host may not support dual stack operation, some host entries may have only IPv4 or IPv6 RRs:

```
example.org.          IN MX   1  mx1.example.org.
                     IN MX  10  mx10.example.org.
mx1.example.org.      IN AAAA  3ffe:501:ffff::1
mx10.example.org.     IN A     192.0.2.1
```

In the following sections, how sender side operates with IPv4/IPv6 combined RR definitions ([section 3](#)), and how receiver side should define RRs to keep interoperability with both IPv4 and IPv6 Internet ([section 4](#)) are considered.

3. SMTP sender algorithm in a dual-stack environment

In a dual-stack environment MX records for a domain resemble the following:

```
example.org.          IN MX   1  mx1.example.org.
                     IN MX  10  mx10.example.org.
mx1.example.org.      IN A     192.0.2.1          ; dual-stack
                     IN AAAA  3ffe:501:ffff::1
mx10.example.org.     IN AAAA  3ffe:501:ffff::2 ; IPv6 only
```

For a single MX record there are many possible final states, including: (a) one or more A records for the IPv4 destination, (b) one or more AAAA records for the IPv6 destination, (c) a mixture of A and AAAA records. Because multiple MX records may be defined using different preference

values, multiple addresses based on multiple MX's must be traversed.
Domains without MX records and failure recovery cases must be handled properly as well.

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 3]

The algorithm for an SMTP sender is basically the same as that for an IPv4-only sender, but it now includes AAAA lookups of MX records for SMTP-over-IPv6 delivery. IPv4/v6 dual stack destinations should be treated just like multihomed destinations as described in [RFC2821](#) [Klensin, 2001] [section 5](#). When there is no reachable destination address record found (for example, the sender MTA is IPv4 only and there are no A records available) the case should be treated just like MX records without address records, and deliveries should never fail because of no known address if other addresses are available related to other MX records.

```

; if the sender MTA is IPv4 only, email delivery to a.example.org
; should fail with the same error as deliveries to b.example.org.
a.example.org.          IN MX      1  mx1.a.example.org.
mx1.a.example.org.      IN AAAA    3ffe:501:ffff::1 ; IPv6 only
b.example.org.          IN MX      1  mx1.b.example.org.
mx1.b.example.org.      IN HINFO   "NO ADDRESS RECORDS"
```

An algorithm for a dual-stack SMTP sender is as follows:

- (1) Lookup the MX record for the destination domain. If a CNAME record is returned, go to the top of step (1) with replacing the destination domain by the query's result. If any MX records are returned, go to step (2) with the query's result (explicit MX). If NO_DATA (i.e. empty answer with NOERROR(0) RCODE) is returned, there is no MX record but the name is valid. Assume that there is a record like "name. IN MX 0 name." (implicit MX) and go to step (3). If HOST_NOT_FOUND (i.e. empty answer with NXDOMAIN(3) RCODE) is returned, there is no such domain. Raise a permanent email delivery failure. Finish.
- (2) Compare each host name in MX records with the name of sending host. If there is a record which has the same name, drop MX records which have equal to or larger than preference value of the matched MX record (including itself). If multiple MX records remain, sort the MX records in ascending order based on their preference values. Loop over steps (3) to (9) on each host name in MX records in a sequence. If no MX records remain, the sending host must be the primary MX host. Other routing rule should be applied. Finish.
- (3) If the sending MTA has IPv4 capability, lookup the A record. Keep the resulting address until step (5).
- (4) If the sending MTA has IPv6 capability, lookup the AAAA record.

NOTE: IPv6 addresses for hosts defined by MX records may be informed in additional information section of DNS querie's result

as well as IPv4 addresses. If there is no additional address information for the MX hosts, separate queries for A or AAAA records should be sent. There is no way to query A and AAAA records at once in current DNS implementation.

- (5) If there is no A or AAAA record present, try the next MX record (go to step (3)). Note that the next MX record could have the same preference.

NOTE: If one or more address records are found, some MTA implementation may sort addresses based on the implementation's preference of A or AAAA records. To encourage the transition from IPv4 SMTP to IPv6 SMTP, AAAA records should take precedence. But this type of sorting is optional.

- (6) For each of the addresses, loop over steps (7) to (9).
- (7) Try to make a TCP connection to the destination's SMTP port (25). The client needs to follow timeouts documented in [RFC2821 section 4.5.3.2](#). If successful, go to step (9).
- (8) If unsuccessful and there is another available address, try the next available address. Go to step (7). If all addresses are not reachable and if a list of MX records is being traversed, try the next MX record (go to step (3)). If there is no list of MX records, or if the end of the list of MX records has been reached, raise a temporary email delivery failure. Finish.
- (9) Attempt to deliver the e-mail over the connection established, as specified in [RFC2821](#) [Klensin, 2001] . If a transient failure condition reported, try the next MX record (go to step (3)). If an error condition reported, raise a permanent email delivery error, and further MX records are not tried. Finish. If successful, SMTP delivery has succeeded. Finish.

[4.](#) MX configuration in the recipient domain

[4.1.](#) Ensuring reachability for both protocol versions

If a site has dual-stack reachability, the site SHOULD configure both A and AAAA records for its MX hosts (NOTE: MX hosts can be outside of the site). This will help both IPv4 and IPv6 senders to reach the site efficiently.

[4.2.](#) Reachability between the primary and secondary MX

When registering MX records in a DNS database in a dual-stack environment, reachability between MX hosts must be considered carefully. Suppose all inbound email is to be gathered at the primary MX host, "mx1.example.org.":

```
example.org.      IN MX    1    mx1.example.org.
```

```
IN MX    10  mx10.example.org.  
IN MX   100 mx100.example.org.
```

If "mx1.example.org" is an IPv6-only node, and the others are IPv4-only

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 5]

nodes, there is no reachability between the primary MX host and the other MX hosts. When email reaches one of the lower MX hosts, it cannot be relayed to the primary MX host based on MX preferencing mechanism, therefore mx1.example.org will not be able to collect all the emails (unless there is another transport mechanism(s) between lower-preference MX hosts and mx1.example.org).

```
; This configuration is troublesome.
; No secondary MX can reach mx1.example.org.
example.org.    IN MX    1    mx1.example.org.    ; IPv6 only
                IN MX    10   mx10.example.org.   ; IPv4 only
                IN MX    100  mx100.example.org.  ; IPv4 only
```

The easiest possible configuration is to configure the primary MX host as a dual-stack node. By doing so, secondary MX hosts will have no problem reaching the primary MX host.

```
; This configuration works well.
; The secondary MX hosts are able to relay email to the primary MX host
; without any problems.
example.org.    IN MX    1    mx1.example.org.    ; dual-stack
                IN MX    10   mx10.example.org.   ; IPv4 only
                IN MX    100  mx100.example.org.  ; IPv6 only
```

It may not be needed that the primary MX host and lower MX hosts reach directly one another with IPv4 or IPv6 transport. For example, it is possible to establish a routing path with UUCP or an IPv4/v6 translator. It is also possible to drop messages into single mailbox with shared storage using NFS or something else offered by a dual-stack server. It is receiver site's matter that all messages delivered to each MX hosts must be reached to recipient's mail drop. In such cases, dual-stack MX host may not be listed in the MX list.

5. Operational experience

Many of the existing IPv6-ready MTA's appear to work in the way documented in [section 3](#).

There were, however, cases where IPv6-ready MTA's were confused by broken DNS servers. When attempting to obtain a canonical hostname, some broken name servers return SERVFAIL (RCODE 2), a temporary failure, on AAAA record lookups. Upon this temporary failure, the email is queued for a later attempt. In the interest of IPv4/v6 interoperability, these broken DNS servers should be fixed. A draft by Yasuhiro Morishita [Morishita, 2003] has more detail on misconfigured/misbehaving DNS servers and their bad sideeffects.

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 6]

6. Open issues

- o How should scoped addresses (i.e. link-local addresses) in email addresses be interpreted on MTA's? We suggest prohibiting the use of IPv6 address literals in source routes (Scoped addresses should not appear on the global DNS database).
- o The document should really be integrated into [RFC2821](#) [Klensin, 2001] (i.e. [RFC2821](#) should talk about IPv6 cases).

7. Security considerations

It could be problematic if the route-addr email address format is used across multiple scope zones. MTA's would need to reject email with improper route-addr email address formats.

References

Hagino, 2001.

Jun-ichiro Hagino and Hal Snyder, "IPv6 multihoming support at site exit routers" in [RFC3178](#) (October 2001). <ftp://ftp.isi.edu/in-notes/rfc3178.txt>.

St.Johns, 1993.

M. St.Johns, "Identification Protocol" in [RFC1413](#) (January 1993). <ftp://ftp.isi.edu/in-notes/rfc1413.txt>.

Thomson, 1995.

S. Thomson and C. Huitema, "DNS Extensions to support IP version 6" in [RFC1886](#) (December 1995). <ftp://ftp.isi.edu/in-notes/rfc1886.txt>.

Partridge, 1986.

C. Partridge, "Mail routing and the domain system" in [RFC974](#) (January 1986). <ftp://ftp.isi.edu/in-notes/rfc974.txt>.

Klensin, 2001.

J. Klensin, Editor, "Simple Mail Transfer Protocol" in [RFC2821](#) (April 2001). <ftp://ftp.isi.edu/in-notes/rfc2821.txt>.

Morishita, 2003.

Y. Morishita and T. Jinmei, "Common Misbehavior against DNS Queries for IPv6 Addresses" in [draft-morishita-dnsop-misbehavior-against-aaaa-00.txt](#) (June 2003). work in progress material.

Change history

00 -> 01

Corrected the email address notation for source-routed emails,
based on a comment from Gregory Neil Shapiro.

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 7]

[01](#) -> [02](#)

Change a reference to refer to [RFC2822](#), not 822. Used "example.org", not "sample.org". These changes were based on comments from Arnt Gulbrandsen. Added an ``Operational experiences'' section. Clarified the case where an MX record points to a CNAME record, based on comments from Mohsen Souissi.

[02](#) -> [03](#)

In some cases, IPv6-ready MTA's are troubled by incorrect DNS server responses for AAAA queries. This change was based on comments from Gregory Neil Shapiro.

[03](#) -> [04](#)

Grammar cleanups by JJ Behrens. More text on the delivery error cases.

[04](#) -> [05](#)

Change title, suggested by Alain Durand. Limit the scope of the document to dual stack environment (interoperation of IPv6-only cloud and IPv4-only cloud is out of scope).

[05](#) -> [06](#)

Section on summary of IPv4 MX operation is deleted (Replaced by Introduction). Clarify on CNAME chain. Cleanups on sender's algorithm. Suggested by Patrik Faltstrom.

[06](#) -> [07](#)

Site local address is being obsoleted in IPv6 wg, so remove reference to site-locals. Reflect comments from John C Klensin: fixes to sending rules, correct route-addr issues. Reflect comments from Michael A. Patton: HELO on connection via translator. Reflect comments from Robert Elz.

[07](#) -> [08](#)

Refer a draft by Yasuhiro Morishita.

Acknowledgements

This draft was written based on discussions with Japanese IPv6 users and help from the WIDE research group. Here is a (probably incomplete) list of people who contributed to the draft: Gregory Neil Shapiro, Arnt Gulbrandsen, Mohsen Souissi, JJ Behrens, John C Klensin, Michael A. Patton, and Robert Elz.

Author's address

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 8]

^L

DRAFT

SMTP in dual stack environments

January 2004

Motonori NAKAMURA

Center for Information and Multimedia Studies, Kyoto University

Yoshida-nihonmatsu-cho, Sakyo, Kyoto 606-8501, JAPAN

Tel: +81-75-753-9063

Fax: +81-75-753-9056

Email: motonori@media.kyoto-u.ac.jp

Jun-ichiro itojun HAGINO

Research Laboratory, Internet Initiative Japan Inc.

1-105, Kanda Jinbo-cho,

Chiyoda-ku, Tokyo 101-0051, JAPAN

Tel: +81-3-5205-6464

Fax: +81-3-5205-6466

Email: itojun@iijlab.net

NAKAMURA, HAGINO

Expires: July 15, 2004

[Page 9]

^L