Network Working Group
Request for Comments: 3974
Category: Informational

M. Nakamura
Kyoto University
J. Hagino
IIJ Research Laboratory

January 2005

SMTP Operational Experience in Mixed IPv4/v6 Environments

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2005).

IESG Note:

The content of this RFC was at one time considered by the IETF, and therefore it may resemble a current IETF work in progress or a published IETF work. This RFC is not a candidate for any level of Internet Standard. The IETF disclaims any knowledge of the fitness of this RFC for any purpose, and in particular notes that the decision to publish is not based on IETF review for such things as security, congestion control, or inappropriate interaction with deployed protocols. The RFC Editor has chosen to publish this document at its discretion. Readers of this RFC should exercise caution in evaluating its value for implementation and deployment.

This document contains a specific interpretation of the applicability of the MX processing algorithm in RFC 2821, Section 5, to dual-stack environments. Implementors are cautioned that they must reference RFC 2821 for the full algorithm; this document is not to be considered a full restatement of RFC 2821, and, in case of ambiguity, RFC 2821 is authoritative.

Abstract

This document discusses SMTP operational experiences in IPv4/v6 dual stack environments. As IPv6-capable SMTP servers are deployed, it has become apparent that certain configurations of MX records are necessary for stable dual-stack (IPv4 and IPv6) SMTP operation. This document clarifies the existing problems 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

Delivery of mail messages to the final mail drop is not always done by direct IP communication between the submitter and final receiver, and there may be some intermediate hosts that 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 easy to configure all systems consistently since the DNS configuration used by mail message delivery systems is more complex than other Internet services. During the transition period from IPv4 to IPv6, more care should be applied to IPv4/v6 interoperability.

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 of MX records are necessary for stable dual-stack (IPv4 and IPv6) SMTP operation.

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 can be resolved by making either side dual-stack or by making either side use a protocol translator (see Appendix A on issues with protocol translator).

2. Basic DNS Resource Record Definitions for Mail Routing

Mail messages on the Internet are typically delivered based on the Domain Name System [Mockapetris]. MX RRs are looked up in DNS to retrieve the names of hosts running MTAs associated with the domain part of the mail address. DNS lookup uses IN class for both IPv4 and IPv6, and similarly IN MX records will be used for mail routing for both IPv4 and IPv6. Hosts which have IPv6 connectivity and also want to have the mails delivered using IPv6 must define IPv6 addresses for the host name as well as IPv4 addresses [Thomson].

An MX RR has two parameters, a preference value and the name of destination host. The name of the destination host will be used to look up an IP address to initiate an SMTP connection [Partridge].

For example, an 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 2001:db8:ffff::1
mx10.example.org.       IN AAAA 2001:db8:ffff::2
```

In the transition period from IPv4 to IPv6, there are many IPv4-only 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.,

```
example.org.            IN MX   1  mx1.example.org.
                        IN MX   10 mx10.example.org.
mx1.example.org.        IN AAAA 2001:db8:ffff::1
                        IN A    192.0.2.1
mx10.example.org.       IN AAAA 2001:db8:ffff::2
                        IN A    192.0.2.2
```

But, not every MX target may support dual-stack operation. Some host entries may have only A RRs or AAAA RRs:

```
example.org.            IN MX   1  mx1.example.org.
                        IN MX   10 mx10.example.org.
mx1.example.org.        IN AAAA 2001:db8:ffff::1
mx10.example.org.       IN A    192.0.2.1
```

The following sections discuss how the sender side should operate with IPv4/v6 combined RRs (section 3), and how the receiver should define RRs to maintain interoperability between IPv4 and IPv6 networks (section 4).

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 2001:db8:ffff::1
mx10.example.org.       IN AAAA 2001:db8:ffff::2 ; IPv6-only
```

For a single MX record, there are multiple 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 must be traversed based on multiple MXs. Domains without MX records and failure recovery cases must be handled properly as well.

The algorithm for a dual-stack 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 RFC 2821 [Klensin], section 5. When there is no destination address record found (i.e., 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 fail.

```
; 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 2001:db8:ffff::1 ; IPv6-only
b.example.org.          IN MX   1  mx1.b.example.org. ; no address
```

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 NODATA (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. If SERVFAIL is returned, retry after a certain period of time.

(2) Compare each host name in MX records with the names of the sending host. If there is match, drop MX records which have an equal or larger value than the lowest-preference matching 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 rules should be applied. Finish.

(3) If the sending MTA has IPv4 capability, lookup the A records. Keep the resulting addresses until step (5).

(4) If the sending MTA has IPv6 capability, lookup the AAAA records.

NOTE: IPv6 addresses for hosts defined by MX records may be informed in an additional information section of the DNS queries' 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 and no 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, an 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. The sorting may only reorder addresses from MX records of the same preference. RFC 2821 section 5 paragraph 4 suggests randomization of destination addresses. Randomization should only happen among A records, and among AAAA records (do not mix A and AAAA records).

(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 RFC 2821 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 email over the connection established, as specified in RFC 2821. If a transient failure condition is reported, try the next MX record (go to step (3)). If an error condition is reported, raise a permanent email delivery error, and do not try further MX records. 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 in reaching 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 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 necessary for the primary MX host and lower MX hosts to directly reach 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 a single mailbox with shared storage using NFS or something else offered by a dual-stack server. It is the receiver site's responsibility that all messages delivered to MX hosts arrive at the recipient's mail drop. In such cases, a 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 document by Yasuhiro Morishita [Morishita] has more detail on misconfigured/misbehaving DNS servers and their negative side effects.

6. Open Issues

- 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 destination specification.

- A future specification of SMTP (revision of RFC 2821) should be updated to include IPv6 concerns presented in this memo, such as (1) the additional query of AAAA RRs where A RRs and/or MX RRs are suggested, and (2) the ordering between IPv6 destination and IPv4 destination.

7. Security Considerations

It could be problematic if the route-addr email address format [Crocker] (or "obs-route" address format in [Resnick]) is used across multiple scope zones. MTAs would need to reject email with route-addr email address formats that cross scope zone borders.

Appendix A. Considerations on Translators

IPv6-only MTA to IPv4-only MTA cases could use help from IPv6-to-IPv4 translators such as [Hagino]. Normally 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 normal IPv4 SMTP traffic.

Protocols like IDENT [St.Johns] may require special consideration when translators are used. Also, there are MTAs which perform strict checks on the SMTP HELO/EHLO "domain" parameter (perform reverse/forward DNS lookups and see if the "domain" really associates to the SMTP client's IP address). In such a case, we need a special consideration when translators will be used (for instance, override "domain" parameter by translator's FQDN/address).

Even without a translator, it seems that there are some MTA implementations in the wild which send IPv6 address literals in a HELO/EHLO message (like "HELO [IPv6:blah]"), even when it is using IPv4 transport, or vice versa. If the SMTP peer is IPv4-only, it won't understand the "[IPv6:blah]" syntax and mails won't go out of the (broken) MTA. These implementations have to be corrected.

Normative References

[Mockapetris] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[Thomson] Thomson, S., Huitema, C., Ksinant, V., and M. Souissi, "DNS Extensions to Support IP Version 6", RFC 3596, October 2003.

[Partridge] Partridge, C., "Mail routing and the domain system", STD 10, RFC 974, January 1986.

[Klensin] Klensin, J., "Simple Mail Transfer Protocol", RFC 2821, April 2001.

[Crocker] Crocker, D., "Standard for the format of ARPA Internet text messages", STD 11, RFC 822, August 1982.

[Resnick] Resnick, P., "Internet Message Format", RFC 2822, April 2001.

[Hagino] Hagino, J. and H. Snyder, "IPv6 Multihoming Support at Site Exit Routers", RFC 3178, October 2001.

[St.Johns] Johns, M. St., "Identification Protocol", RFC 1413, February 1993.

Informative References

[Morishita] Morishita, Y. and T. Jinmei, "Common Misbehavior against DNS Queries for IPv6 Addresses", Work in Progress, June 2003.

Acknowledgements

This document 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 document: Gregory Neil Shapiro, Arnt Gulbrandsen, Mohsen Souissi, JJ Behrens, John C Klensin, Michael A. Patton, Robert Elz, Dean Strik, Pekka Savola, and Rob Austein.

Authors' Addresses

Motonori NAKAMURA
Academic Center for Computing and Media Studies, Kyoto University
Yoshida-honmachi, Sakyo, Kyoto 606-8501, JAPAN

Fax: +81-75-753-7450
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

Phone: +81-3-5205-6464
Fax: +81-3-5205-6466
EMail: itojun@iijlab.net

Full Copyright Statement

Copyright (C) The Internet Society (2005).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and at www.rfc-editor.org, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM 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.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the ISOC's procedures with respect to rights in ISOC Documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.