Internet Engineering Task Force (IETF)
Request for Comments: 5983
Category: Experimental
ISSN: 2070-1721

R. Gellens
Qualcomm

October 2010

Mailing Lists and Internationalized Email Addresses

Abstract

This document describes considerations for mailing lists with the introduction of internationalized email addresses.

This document makes some specific recommendations on how mailing lists should act in various situations.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for examination, experimental implementation, and evaluation.

This document defines an Experimental Protocol for the Internet community. This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc5983.

Table of Contents

1. Introduction

This document describes considerations for mailing lists with the introduction of internationalized email addresses [RFC5335].

Mailing lists are an important part of email usage and collaborative communications. The introduction of internationalized email addresses affects mailing lists in three main areas: (1) transport (receiving and sending messages), (2) message headers of received and retransmitted messages, and (3) mailing list operational policies.

A mailing list is a mechanism whereby a message may be distributed to multiple recipients by sending to one address. An agent (typically not a human being) at that single address receives the message and then causes the message to be redistributed to a list of recipients. This agent sets the envelope return address of the redistributed message to a different address from that of the original message. Using a different envelope return address (reverse-path) directs error (and other automatically generated) messages to an error

handling address associated with the mailing list. (This avoids having error and other automatic messages go to the original sender, who typically doesn't control the list and hence can't do anything about them.)

In most cases, the mailing list agent redistributes a received message to its subscribers as a new message, that is, conceptually it uses message submission [submission] (as did the sender of the original message). The exception, where the mailing list is not a separate agent that receives and redistributes messages in separate transactions, but is instead an expansion step within an SMTP transaction where one local address expands to multiple local or non-local addresses, is out of scope for this document.

Some mailing lists alter message header fields, while others do not. A number of standardized list-related header fields have been defined, and many lists add one or more of these header fields. Separate from these standardized list-specific header fields, and despite a history of interoperability problems from doing so, some lists alter or add header fields in an attempt to control where replies are sent. Such lists typically add or replace the "Reply-To" field and some add or replace the "Sender" field. Poorly behaved lists may alter or replace other fields, including "From".

Among these list-specific header fields are those specified in RFC 2369 ("The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields") [List-*] and RFC 2919 ("List-Id: A Structured Field and Namespace for the Identification of Mailing Lists") [List-ID]. For more information, see Section 5.

While the mail transport protocol does not differ between regular email recipients and mailing list recipients, lists have special considerations with internationalized email addresses because they retransmit messages composed by other agents to potentially many recipients.

There are considerations for internationalized email addresses in the envelope as well as in header fields of redistributed messages. In particular, an internationalized message cannot be downgraded unless all envelope addresses are available in ASCII (that is, each address either is ASCII or has an alt-address [UTF8SMTP]).

With mailing lists, there are two different types of considerations: first, the purely technical ones involving message handling, error cases, downgrades, and the like; and second, those that arise from the fact that humans use mailing lists to communicate. As an example of the first, mailing lists might choose to reject all messages from internationalized addresses that lack an alt-address, or even all

internationalized messages that cannot be downgraded. As an example of the second, a user who sends a message to a list often is unaware of the list membership. In particular, the user often doesn't know if the members are UTF-8 mail users or not, and often neither the original sender nor the recipients personally know each other. As a consequence of this, remedies that may be readily available for a missed email in one-to-one communications might not be appropriate when dealing with mailing lists. For example, if a user sends a message that is undeliverable, normally the telephone, instant messaging, or other forms of communication are available to obtain a working address. With mailing lists, the users may not have any recourse. Of course, with mailing lists, the original sender usually does not know if the message was successfully received by any list members or if it was undeliverable to some.

Conceptually, a mailing list's internationalization can be divided into three capabilities: First, does it have a UTF-8 submission or return-path address? Second, does it accept subscriptions to UTF-8 addresses? And third, does it accept [UTF8SMTP] messages? This is explored in Section 4.

A brief discussion on a few additional considerations for mailing list operation is in Section 6.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [KEYWORDS].

3. Scenarios Involving Mailing Lists

Generally (and exclusively within the scope of this document), an original message is sent to a mailing list as a completely separate and independent transaction from the mailing list agent sending the retransmitted message to one or more list recipients. In both cases, the message might have only one recipient, or might have multiple recipients. That is, the original message might be sent to additional recipients as well as the mailing list agent, and the mailing list might choose to send the retransmitted message to each list recipient in a separate message submission [submission] transaction, or it might choose to include multiple recipients per transaction. (Often, mailing lists are constructed to work in cooperation with, rather than include the functionality of, a message submission server [submission], and hence the list transmits to a single submission server one copy of the retransmitted message, with

all list recipients specified in the SMTP envelope. The submission server then decides which recipients to include in which transaction.)

The retransmitted message sent by the mailing list to its subscribers might need to be downgraded [EAI-Downgrade]. In order for a downgrade to be possible, the return path set by the mailing list agent must be an ASCII address or have an alt-address [UTF8SMTP] specified. In addition, the recipient addresses need to have ASCII addresses available. It may be advisable for mailing list operators to pre-obtain an alt-address for all its internationalized member addresses.

In the case where a member or non-member with an internationalized email address is sending to a mailing list, no alt-address [UTF8SMTP] is specified, and a downgrade is required, the message cannot be delivered. To protect against this, a UTF8SMTP-aware [UTF8SMTP] mailing list might prefer to reject submissions from internationalized email addresses that lack an alt-address.

(Note that this situation is not unique to mailing lists. Mail relays that are UTF8SMTP-aware will potentially encounter the same situation.) Further discussions are included in Section 6 of this document.

4. Capabilities and Requirements

There are three primary internationalization capabilities of mailing lists: First, does it have a UTF-8 submission or return-path address? Second, does it allow subscriptions from UTF-8 addresses? And third, does it accept [UTF8SMTP] messages?

In theory, any list can support any combination of these. In practice, only some offer any benefit. For example, neither allowing UTF-8 addresses to subscribe, nor accepting UTF8SMTP messages, makes much sense without the other (an all-ASCII address might or might not be capable of receiving UTF8SMTP messages, but a UTF-8 address of necessity needs to accept UTF8SMTP messages). Likewise, there is no real benefit to a list in using a UTF-8 submission address unless it also accepts UTF8SMTP messages and permits UTF-8 addresses to subscribe.

However, requirements for lists can be discussed separately for each of the three capabilities.

1. If the list uses a UTF-8 submission or return-path address, it SHOULD specify an alt-address [UTF8SMTP] for it. Clearly, it needs to sit behind a UTF8SMTP-enabled final-delivery SMTP server

[UTF8SMTP] and delivery agent. Likewise, if a list uses a UTF-8 return-path address, then its Message Submission Agent (MSA) [submission] needs to support UTF8SMTP.

The list's return-path address is usually separate from its submission address (so that delivery reports and other automatically generated messages are not sent to the submission address). For reliability in receiving delivery status notifications, a list MAY choose to use an all-ASCII return path even if it uses a UTF-8 submission address. If the list does use a UTF-8 return path, it MUST specify an alt-address [UTF8SMTP] (or else there is a high risk of being unable to receive non-delivery reports).

There are also implications for the List-* header fields (see below).

2. If it allows UTF-8 addresses to subscribe, it MAY require an alt-address [UTF8SMTP] to be specified for each UTF-8 subscriber.

   Naturally, if it permits UTF-8 addresses to subscribe, it needs a mechanism to accept subscription requests from such addresses (preferably specified in the form <utf8@utf8 <ascii@ascii>>). In order to send email to its subscribers who have UTF-8 addresses, its MSA needs to support [UTF8SMTP].

3. If it accepts UTF8SMTP messages, the Message Transfer Agents (MTAs) and Mail Delivery Agent (MDA) in its inbound path need to support UTF8SMTP.

5. List Header Fields

A number of header fields, specifically for mailing lists, have been introduced in RFCs 2369 and 2919. For example, these include:

```
List-Id: List Header Mailing List <list-header.example.com>
List-Help: <mailto:list@example.com?subject=help>
List-Unsubscribe: <mailto:list@example.com?subject=unsubscribe>
List-Subscribe: <mailto:list@example.com?subject=subscribe>
List-Post: <mailto:list@example.com>
List-Owner: <mailto:listmom@example.com> (Contact Person for Help)
List-Archive: <mailto:archive@example.com?subject=index%20list>
```

As described in RFC 2369, "The contents of the list header fields mostly consist of angle-bracket ('<', '>') enclosed URLs, with internal whitespace being ignored" [List-*]. For List-ID, RFC 2919 specifies that, "The list identifier will, in most cases, appear like a host name in a domain of the list owner" [List-ID].

Except for the List-ID header field, these mailing list header fields contain URLs [RFC3986]. The most common schemes are generally HTTP, HTTPS, mailto, and FTP. Schemes that permit both URI and Internationalized Resource Identifier (IRI) [IRI] forms should use the URI-encoded form described in [IRI]. Future work may extend these header fields or define replacements to directly support non-encoded UTF-8 in IRIs (for example, [mailto-bis]), but in the absence of such extension or replacement, non-ASCII characters can only appear within when encoded as ASCII. Note that discussion on whether internationalized domain names should be percent encoded or puny coded is ongoing; see [IRI-bis].

Even without these header fields being extended to support UTF-8, some special provisions may be helpful when downgrading. In particular, if a List-* header field contains a UTF-8 mailto (even encoded in ASCII) followed by an ASCII mailto, it may be advisable not only to copy and preserve the original header field as usual (ENCAPSULATION method of [EAI-Downgrade]), but also to edit the header field to remove the UTF-8 address. Otherwise, a client might run into trouble if the decoded mailto results in a non-ASCII address.

When mailing lists use a UTF-8 form of a List-* header field, an ASCII form SHOULD also be used. These header fields are vital to good operations and use of mailing lists; caution is called for when considering how to form and use these header fields in a non-ASCII environment.

The most commonly used URI schemes in List-* header fields tend to be HTTP and mailto. The current specification for mailto does not permit unencoded UTF-8 characters, although work has been proposed to extend or more likely replace mailto in order to permit this. For mailto URIs, a separate consideration is how to include an alternate ASCII address (alt-address) [UTF8SMTP] for a UTF-8 address. Note that the existing ability to specify multiple URLs within each List-* header field provides one solution.

[List-*] says:

> A list of multiple, alternate, URLs MAY be specified by a comma-separated list of angle-bracket enclosed URLs. The URLs have order of preference from left to right. The client application should use the left most protocol that it supports, or knows how to access by a separate application.

When a UTF-8 mailto is used in a List-* header field, an alt-address [UTF8SMTP], if available, SHOULD be supplied.

The List-ID header field provides an opaque value that uniquely identifies a list. The intent is that the value of this header field remain constant, even if the machine or system used to operate or host the list changes. This header field is often used in various filters and tests, such as client-side filters, Sieve filters, and so forth. Such filters and tests may not properly compare a non-ASCII value that has been encoded into ASCII. In addition to these comparison considerations, it is generally desirable that this header field contain something meaningful that users can type in. However, ASCII encodings of non-ASCII characters are unlikely to be meaningful to users or easy for them to accurately type.

6. Further Discussion

While mailing lists do not create a significant additional burden to the deployment of internationalized email address functionalities, there are some specific areas that need to be considered when the operator of a mailing list or of a final delivery MTA that serves a mailing list upgrades to internationalized mail.

Mailing lists face additional complexity since they redistribute messages composed by other agents. Hence, they may be asked to accept a message with non-ASCII header fields composed by a UTF8SMTP-aware user agent [UTF8SMTP] and redistribute it to UTF-8 mail and all-ASCII mail users via systems that are not UTF8SMTP-aware.

1. Obtaining Downgrade Information -- for a mailing list, or mail relay server for that matter, which is UTF8SMTP-aware, receiving mail from an internationalized email address, the alt-address [UTF8SMTP] is not required from the sending MTA for the transport to be complete. When the mailing list then retransmits the message to its subscribers, it may encounter paths where a downgrade is needed (if a relay or final MSA does not supports UTF8SMTP). In order to mitigate this situation, the mailing list might perhaps decide to reject all incoming mail from an internationalized email address that lacks an alt-address. However, note that in general, downgrades are not expected to be the normal case.

2. Downgrading Considerations for mailto URLs -- UTF-8 addresses in mailto links in List-* header fields will be easier to downgrade if they contain an alt-address [UTF8SMTP].

7. Security Considerations

Because use of both a UTF-8 address and an alt-address for the same entity introduces a potential ambiguity regarding the identity of list subscribers and message senders, implementers are advised to

carefully handle authorization decisions regarding subscriptions, sender filters, and other common list administration features. For example, a binding between a UTF-8 address and an ASCII alt-address can be used by an attacker to deny another person admission to an Email Address Internationalization (EAI) mailing list.

Other relevant security considerations are discussed in the Framework document [EAI-Framework].

8. Acknowledgments

Edmon Chung of Afilias wrote the original version of this document. Thanks to Harald Alvestrand for his extensive comments. Ted Hardie contributed helpful text on IRIs. Last-Call comments from S. Moonesamy and Amanda Baber, plus shepherd review by Pete Resnick, improved the document.

9. References

9.1. Normative References

[EAI-Framework] Klensin, J. and Y. Ko, "Overview and Framework for Internationalized Email", RFC 4952, July 2007.

[KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[List-*] Neufeld, G. and J. Baer, "The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields", RFC 2369, July 1998.

[List-ID] Chandhok, R. and G. Wenger, "List-Id: A Structured Field and Namespace for the Identification of Mailing Lists", RFC 2919, March 2001.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[RFC5335] Abel, Y., Ed., "Internationalized Email Headers", RFC 5335, September 2008.

[submission] Gellens, R. and J. Klensin, "Message Submission for Mail", RFC 4409, April 2006.

[UTF8SMTP] Yao, J., Ed., and W. Mao, Ed., "SMTP Extension for Internationalized Email Addresses", RFC 5336, September 2008.

9.2. Informative References

[EAI-Downgrade]
Fujiwara, K., Ed., and Y. Yoneya, Ed., "Downgrading Mechanism for Email Address Internationalization", RFC 5504, March 2009.

[IRI] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, January 2005.

[IRI-bis] Duerst, M., Suignard, M., and L. Masinter, "Internationalized Resource Identifiers (IRIs)", Work in Progress, July 2010.

[mailto-bis]
Duerst, M., Masinter, L., and J. Zawinski, "The 'mailto' URI Scheme", Work in Progress, May 2010.

Author's Address

Randall Gellens
QUALCOMM Incorporated
5775 Morehouse Drive
San Diego, CA 92121
rg+ietf@qualcomm.com