Internet Engineering Task Force (IETF)
Request for Comments: 5825
Category: Experimental
ISSN: 2070-1721

K. Fujiwara
JPRS
B. Leiba
Huawei Technologies

April 2010

Displaying Downgraded Messages for Email Address Internationalization

Abstract

This document describes a method for displaying downgraded messages that originally contained internationalized email addresses or internationalized header fields.

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/rfc5825.

Copyright Notice

Copyright (c) 2010 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/license-info) 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.

Table of Contents

1. Introduction

The Email Address Internationalization (UTF8SMTP) extension document set [RFC4952] [RFC5336] [RFC5335] [RFC5337] expands Email address structure, syntax, and email header format. To avoid rejection of internationalized email messages, the downgrading mechanism [RFC5504] converts an internationalized message to a traditional email message when a server in the delivery path does not support the UTF8SMTP extension. The downgraded message is a traditional email message, except the message has "Downgraded-" header fields.

A perfect reverse-function of the downgrading does not exist because the encoding defined in [RFC2047] is not exactly reversible and "Received" header field downgrading may remove FOR clause information. The restoration of the downgrading should be done once at the final destination of the downgraded message such as Mail User Agents (MUAs) or IMAP servers. This document describes the restoration methods for displaying downgraded messages in MUAs.

2. Terminology

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 RFC 2119 [RFC2119].

Specialized terms used in this specification are defined in the EAI overview [RFC4952] or in [RFC5321], [RFC5322], or the MIME documents [RFC2045], [RFC2047], [RFC2183], and [RFC2231].

This document depends on [RFC5335] and [RFC5504]. Key words used in those documents are used in this document, too.

The term "MIME decode" is used for both "encoded-word" decoding defined by [RFC2047] and MIME parameter value decoding defined by [RFC2231].

3. Converting Downgraded Message Headers for Display

3.1. Considerations

The order of some header fields (such as "Resent-*" fields) is significant. The process of regenerating the original fields from the downgraded ones MUST NOT reorder the fields.

In order to regenerate a field from a specific downgraded header field, it's necessary to find the corresponding replacement in the current message. If the corresponding field cannot be found, the downgraded header field in question cannot be regenerated and used.

In any case where reconstruction of a particular downgraded header field fails, both header fields (the "downgraded-YYY" header field and the "YYY" header field) SHOULD be left in the message as they are. The MUA MAY choose to communicate the situation to the user (see the "Security Considerations" section).

3.2. The Process

A MUA MAY decode and regenerate the original header fields of the message (Mail Transport Agents (MTAs) and Mail Delivery Agents (MDAs) SHOULD NOT attempt to do this; it SHOULD be left to the MUA). This procedure can be used to approximately reverse the downgrade process, but it will not always construct the original header fields exactly.

Three types of downgraded header fields are described in Section 3 of [RFC5504]:

1. "Envelope Information Preservation Header Fields", described in RFC5504 Section 3.1 and in Section 3.2.1, below.

2. "Address Header Fields' Preservation Header Fields", described in RFC5504 Section 3.2 and in Section 3.2.2, below.

3. "Unknown Header Fields' Preservation Header Fields", described in RFC5504 Section 3.3 and in Section 3.2.3, below.

After processing downgraded header fields, decode all header fields, as described in [RFC2047] and [RFC2231].

3.2.1. No Reconstruction of the Envelope Information Preservation Header Fields

Envelope information preservation header fields are new fields that might have been added by the downgrade process. Because they do not represent fields that appeared in the original message, this process is not applicable to them.

3.2.2. Reconstructing the Address Header Fields' Preservation Header Fields

Reconstructing address header fields' preservation header fields is OPTIONAL, and a decision MAY be made on each field, individually. In particular, it might be less important to process the "Resent-*" header fields, so an implementation MAY choose to skip those.

To construct a displayable copy of a header field from one of these downgraded header fields, follow this procedure:

1. In an edit buffer, create a new header field:

   (a) For the field name, remove the "Downgraded-" prefix from the downgraded field name. For example, "Downgraded-From" becomes "From", and "Downgraded-Resent-To" becomes "Resent-To".

   (b) For the field value, decode the MIME-encoded value of the downgraded field according to [RFC2047].

2. Apply "Email Header Fields Downgrading", defined in Section 5 of [RFC5504], to the field in the edit buffer. The process generates two header fields, one is ASCII header field and the other is the Address Header Fields' Preservation Header Field. Put the generated ASCII header field into comparison buffer 1.

3. Canonicalize the header field in the comparison buffer 1:

   1. Unfold all header field continuation lines as described in [RFC5322].

2. Ensure that there is one space character before and one after the <mailbox-list> separator ",". If a space character is missing, insert one.

3. Ensure that there is one space character before and one after each <comment>. If a space character is missing, insert one.

4. Decode each <encoded-word> whose charset is "UTF-8".

5. Convert all sequences of one or more WSP characters to a single space character. WSP characters here include those before and after a line-folding boundary.

6. Delete all WSP characters at the end of each unfolded header field value.

7. Delete any WSP characters remaining before and after the colon separating the header field name from the header field value, retaining the colon separator.

4. Locate the first instance of the corresponding field in the message's headers.

5. Canonicalize the located field as in step 3, and put the result into comparison buffer 2.

6. Compare the header field in comparison buffer 1 with the header field in comparison buffer 2. If they match, go to step 8.

7. Locate the next instance of the corresponding field in the message's headers. If one is found, go to step 5. If none is found, stop: you cannot use this downgraded field because you can't find its replacement in the message.

8. Replace the located header field with the one in the edit buffer. You MUST NOT reorder the header fields when you do this; it's important to replace the field in the same place. Remove the target downgraded header field in the message header.

3.2.3. The Unknown Header Fields' Preservation Header Fields

The unknown header fields' preservation header fields SHOULD be left as they are unless the MUA has special knowledge of a particular field. An MUA with such knowledge MAY use the procedure similar to the procedure in Section 3.2.2, above, for those fields about which it knows. (Note that the whitespace canonicalization rule might not be applicable to some header fields.)

4. Security Considerations

While information in any email header should usually be treated with some suspicion, current email systems commonly employ various mechanisms and protocols to make the information more trustworthy. For example, an organization's boundary MTA can modify "From" lines so that messages arriving from outside the organization are easily distinguishable from internal emails. As a result of that rewriting, the "From" header field might not match the "Downgraded-From" header field.

A MUA MAY emphasize bogus or broken address header fields' preservation header fields found in step 7 of Section 3.2.2.

Hiding the information from the actual header fields when using the "Downgraded-" header fields does not cause loss of information if generating MIME-decoded header fields in step 1 of Section 3.2.2 and the comparison done in step 7 are successful. To ensure that no information is lost, a MUA SHOULD have a function that uses the actual message that was received (with/without MIME decoding) to render the message.

We have focused, here, on issues with displaying downgraded messages. For more discussion of downgraded and internationalized messages in general, see the "Security Considerations" section in [RFC5504] and [RFC4952].

5. Acknowledgements

This document was separated from [RFC5504]. Both documents were developed in the EAI WG. Significant comments and suggestions were received from John Klensin, Harald Alvestrand, Chris Newman, Randall Gellens, Charles Lindsey, Marcos Sanz, Alexey Melnikov, Pasi Eronen, Frank Ellermann, Edward Lewis, S. Moonesamy, and JET members.

6. References

6.1. Normative References

[RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

[RFC2047] Moore, K., "MIME (Multipurpose Internet Mail Extensions) Part Three: Message Header Extensions for Non-ASCII Text", RFC 2047, November 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2183] Troost, R., Dorner, S., and K. Moore, "Communicating Presentation Information in Internet Messages: The Content-Disposition Header Field", RFC 2183, August 1997.

[RFC2231] Freed, N. and K. Moore, "MIME Parameter Value and Encoded Word Extensions:
Character Sets, Languages, and Continuations", RFC 2231, November 1997.

[RFC4952] Klensin, J. and Y. Ko, "Overview and Framework for Internationalized Email", RFC 4952, July 2007.

[RFC5322] Resnick, P., Ed., "Internet Message Format", RFC 5322, October 2008.

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

[RFC5504] Fujiwara, K. and Y. Yoneya, "Downgrading Mechanism for Email Address Internationalization", RFC 5504, March 2009.

6.2. Informative References

[RFC5321] Klensin, J., "Simple Mail Transfer Protocol", RFC 5321, October 2008.

[RFC5336] Yao, J. and W. Mao, "SMTP Extension for Internationalized Email Addresses", RFC 5336, September 2008.

[RFC5337] Newman, C. and A. Melnikov, "Internationalized Delivery Status and Disposition Notifications", RFC 5337, September 2008.

Appendix A. Examples

This section shows an example of displaying a downgraded message. First, an example of the original UTF8SMTP message and its downgraded message are shown. The example comes from "Example 1" of [RFC5504] and three header fields, "Unknown-Field", "Resent-From", and "Resent-To", are added. The example UTF8SMTP message is shown in Figure 1.

```
Message-Id: MESSAGE_ID
Mime-Version: 1.0
Content-Type: text/plain; charset="UTF-8"
Content-Transfer-Encoding: 8bit
Subject: NON-ASCII-SUBJECT
Unknown-Field: NON-ASCII-Unknown
From: DISPLAY-local <NON-ASCII-local@example.com
 <ASCII-local@example.com>>
To: DISPLAY-remote1 <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Cc: DISPLAY-remote2 <NON-ASCII-remote2@example.org>
Resent-From: DISPLAY-remote1 <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Resent-To: DISPLAY-reto <NON-ASCII-reto@example.net
 <ASCII-reto@example.net>>
Date: DATE

MAIL_BODY
```

Figure 1: Original message

A delivered downgraded message is shown in Figure 2. A Return-Path header will be added by the final destination MTA. Some "Received" header fields may be added.

```
Return-Path: <ASCII-local@example.com>
Received: ...
Downgraded-Mail-From: =?UTF-8?Q?<NON-ASCII-local@example.com_?=
 =?UTF-8?Q?<ASCII-local@example.com>>?=
Downgraded-Rcpt-To: =?UTF-8?Q?<NON-ASCII-remote1@example.net_?=
 =?UTF-8?Q?<ASCII-remote1@example.net>>?=
Message-Id: MESSAGE_ID
Mime-Version: 1.0
Content-Type: text/plain; charset="UTF-8"
Content-Transfer-Encoding: 8bit
Subject: =?UTF-8?Q?NON-ASCII-SUBJECT?=
Downgraded-Unknown-Field: =?UTF-8?Q?NON-ASCII-Unknown?=
From: =?UTF-8?Q?DISPLAY-local?= <ASCII-local@example.com>
Downgraded-From: =?UTF-8?Q?DISPLAY-local_<NON-ASCII-local@example.com_?=
 =?UTF-8?Q?<ASCII-local@example.com>>?=
To: =?UTF-8?Q?DISPLAY-remote1?= <ASCII-remote1@example.net>
Downgraded-To: =?UTF-8?Q?DISPLAY-remote1_?=
 =?UTF-8?Q?<NON-ASCII-remote1@example.net_<ASCII-remote1@example.net>>?=
Cc: =?UTF-8?Q?DISPLAY-remote2?= Internationalized address
 =?UTF-8?Q?NON-ASCII-remote2@example.org?= removed:;
Downgraded-Cc: =?UTF-8?Q?DISPLAY-remote2_?=
 =?UTF-8?Q?<NON-ASCII-remote2@example.org>?=
Resent-From: =?UTF-8?Q?DISPLAY-remote1?= <ASCII-remote1@example.net>
Downgraded-Resent-From: =?UTF-8?Q?DISPLAY-remote1_?=
 =?UTF-8?Q?<NON-ASCII-remote1@example.net_<ASCII-remote1@example.net>>?=
Resent-To: =?UTF-8?Q?DISPLAY-reto?= <ASCII-reto@example.net>
Downgraded-Resent-To: =?UTF-8?Q?DISPLAY-reto_?=
 =?UTF-8?Q?<NON-ASCII-reto@example.net_<ASCII-reto@example.net>>?=
Date: DATE

MAIL_BODY
```

Figure 2: Downgraded message

Figure 3 shows the MIME-decoded message of Figure 2. The recipient can read the original "From", "To", "Cc", "Resent-From", "Resent-To" and "Unknown-Field" header fields as "Downgraded-From", "Downgraded-To", "Downgraded-Cc", "Downgraded-Resent-From", "Downgraded-Resent-To", and "Downgraded-Unknown-Field" header fields.

```
Return-Path: <ASCII-local@example.com>
Received: ...
Downgraded-Mail-From: <NON-ASCII-local@example.com
 <ASCII-local@example.com>>
Downgraded-Rcpt-To: <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Message-Id: MESSAGE_ID
Mime-Version: 1.0
Content-Type: text/plain; charset="UTF-8"
Content-Transfer-Encoding: 8bit
Subject: NON-ASCII-SUBJECT
Downgraded-Unknown-Field: NON-ASCII-Unknown
From: DISPLAY-local <ASCII-local@example.com>
Downgraded-From: DISPLAY-local <NON-ASCII-local@example.com
 <ASCII-local@example.com>>
To: DISPLAY-remote1 <ASCII-remote1@example.net>
Downgraded-To: DISPLAY-remote1 <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Cc: DISPLAY-remote2 Internationalized address
 NON-ASCII-remote2@example.org removed:;
Downgraded-Cc: DISPLAY-remote2 <NON-ASCII-remote2@example.org>
Resent-From: DISPLAY-remote1 <ASCII-remote1@example.net>
Downgraded-Resent-From: DISPLAY-remote1
 <NON-ASCII-remote1@example.net <ASCII-remote1@example.net>>
Resent-To: DISPLAY-reto <ASCII-reto@example.net>
Downgraded-Resent-To: DISPLAY-reto
 <NON-ASCII-reto@example.net <ASCII-reto@example.net>>
Date: DATE

MAIL_BODY
```

Figure 3: MIME-decoded message

A.1. Displaying Example

This example shows how to display the message in Figure 2, above, using the process defined in Section 3. For simplicity, we will show the reconstruction of all the applicable fields at once.

Selecting all Downgraded-* fields gives this:

```
Downgraded-Mail-From: =?UTF-8?Q?<NON-ASCII-local@example.com_?=
 =?UTF-8?Q?<ASCII-local@example.com>>?=
Downgraded-Rcpt-To: =?UTF-8?Q?<NON-ASCII-remote1@example.net_?=
 =?UTF-8?Q?<ASCII-remote1@example.net>>?=
Downgraded-Unknown-Field: =?UTF-8?Q?NON-ASCII-Unknown?=
Downgraded-From: =?UTF-8?Q?DISPLAY-local_<NON-ASCII-local@example.com_?=
 =?UTF-8?Q?<ASCII-local@example.com>>?=
Downgraded-To: =?UTF-8?Q?DISPLAY-remote1_?=
 =?UTF-8?Q?<NON-ASCII-remote1@example.net_<ASCII-remote1@example.net>>?=
Downgraded-Cc: =?UTF-8?Q?DISPLAY-remote2_?=
 =?UTF-8?Q?<NON-ASCII-remote2@example.org>?=
Downgraded-Resent-From: =?UTF-8?Q?DISPLAY-remote1_?=
 =?UTF-8?Q?<NON-ASCII-remote1@example.net_<ASCII-remote1@example.net>>?=
Downgraded-Resent-To: =?UTF-8?Q?DISPLAY-reto_?=
 =?UTF-8?Q?<NON-ASCII-reto@example.net_<ASCII-reto@example.net>>?=
```

Figure 4: Downgraded header fields

Two of the fields, "Downgraded-Mail-From" and "Downgraded-Rcpt-To", are envelope information preservation header fields, and will not be reconstructed. One field, "Downgraded-Unknown-Field", is an unknown header fields' preservation header field and will also not be reconstructed. That leaves the address header fields' preservation header fields to be reconstructed.

```
Downgraded-From: =?UTF-8?Q?DISPLAY-local_<NON-ASCII-local@example.com_?=
 =?UTF-8?Q?<ASCII-local@example.com>>?=
Downgraded-To: =?UTF-8?Q?DISPLAY-remote1_?=
 =?UTF-8?Q?<NON-ASCII-remote1@example.net_<ASCII-remote1@example.net>>?=
Downgraded-Cc: =?UTF-8?Q?DISPLAY-remote2_?=
 =?UTF-8?Q?<NON-ASCII-remote2@example.org>?=
Downgraded-Resent-From: =?UTF-8?Q?DISPLAY-remote1_?=
 =?UTF-8?Q?<NON-ASCII-remote1@example.net_<ASCII-remote1@example.net>>?=
Downgraded-Resent-To: =?UTF-8?Q?DISPLAY-reto_?=
 =?UTF-8?Q?<NON-ASCII-reto@example.net_<ASCII-reto@example.net>>?=
```

Figure 5: Header fields for the reconstruction

Now, perform step 1 to the downgraded header fields shown in Figure 5 and create an edit buffer.

```
From: DISPLAY-local <NON-ASCII-local@example.com
 <ASCII-local@example.com>>
To: DISPLAY-remote1 <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Cc: DISPLAY-remote2 <NON-ASCII-remote2@example.org>
Resent-From: DISPLAY-remote1
 <NON-ASCII-remote1@example.net <ASCII-remote1@example.net>>
Resent-To: DISPLAY-reto
 <NON-ASCII-reto@example.net <ASCII-reto@example.net>>
```

Figure 6: edit buffer: Output of step 1

Apply "Email Header Fields Downgrading" to the "edit buffer". It generates downgraded ASCII header fields and the address header fields' preservation header fields. The latter fields are the same as the downgraded header fields. Put the former fields into "comparison buffer 1".

```
From:DISPLAY-local <ASCII-local@example.com>
To:DISPLAY-remote1 <ASCII-remote1@example.net>
Cc:DISPLAY-remote2 Internationalized address
 NON-ASCII-remote2@example.org removed:;
Resent-From:DISPLAY-remote1 <ASCII-remote1@example.net>
Resent-To:DISPLAY-reto <ASCII-reto@example.net>
```

Figure 7: comparison buffer 1: Output of step 3

Perform steps 4 to 6, comparison, for each header field. Five header fields, "From", "To", "Cc", "Resent-From" and "Resent-To" fields will match, and we will proceed to step 8. (Step 7, iteration, does not apply in this example.

Perform step 8, replacing all applicable fields, without changing the order. Then, do MIME decoding on everything, for display.

```
Return-Path: <ASCII-local@example.com>
Received: ...
Downgraded-Mail-From: <NON-ASCII-local@example.com
 <ASCII-local@example.com>>
Downgraded-Rcpt-To: <NON-ASCII-remote1@example.net>
 <ASCII-remote1@example.net>
Message-Id: MESSAGE_ID
Mime-Version: 1.0
Content-Type: text/plain; charset="UTF-8"
Content-Transfer-Encoding: 8bit
Subject: NON-ASCII-SUBJECT
Downgraded-Unknown-Field: NON-ASCII-Unknown
From: DISPLAY-local <NON-ASCII-local@example.com
 <ASCII-local@example.com>>
To: DISPLAY-remote1 <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Cc: DISPLAY-remote2 <NON-ASCII-remote2@example.org>
Resent-From: DISPLAY-remote1 <NON-ASCII-remote1@example.net
 <ASCII-remote1@example.net>>
Resent-To: DISPLAY-reto <NON-ASCII-reto@example.net
 <ASCII-reto@example.net>>
Date: DATE
```

Figure 8: The final result

As a result, in this simple example, some original header fields are now displayed in their original form. Differences between Figure 1 and Figure 8 are Return-Path, Downgraded-Mail-From, Downgraded-Rcpt-To, and Downgraded-Unknown-Field.

Authors' Addresses

Kazunori Fujiwara
Japan Registry Services Co., Ltd.
Chiyoda First Bldg. East 13F, 3-8-1 Nishi-Kanda
Chiyoda-ku, Tokyo 101-0065
Japan

Phone: +81-3-5215-8451
EMail: fujiwara@jprs.co.jp

Barry Leiba
Huawei Technologies

Phone: +1 646 827 0648
EMail: barryleiba@computer.org
URI: http://internetmessagingtechnology.org/