Network Working Group
Internet-Draft
Intended Status: Proposed Standard
Updates: 3501

Arnt Gulbrandsen
August 2012

Simplified POP/IMAP Downgrading for Internationalized Email

draft-ietf-eai-simpledowngrade-06.txt

Status of this Memo

Abstract

This document specifies a method for IMAP and POP servers to serve internationalized messages to conventional clients. The specification is simple, easy to implement and provides only rudimentary results.

1. Overview

It may happen that a conventional IMAP or POP client opens a mailbox containing internationalized messages, or even attempt to read internationalized messages, for instance when a user has both internationalized and conventional MUAs.

Some operations cannot be performed by conventional clients. Most importantly, an internationalized message usually contains at least one internationalized address, so address-based operations are only rarely possible. This includes displaying the addresses, replying, and most types of address-based signature or security processing.

Still, the sender's name, the message subject, body text and attachments can easily be displayed, so a helpful IMAP/POP server may prefer to provide access to what it can rather than hide the message entirely.

This document specifies a way to present such messages to the client. It values simplicity of implementation over fidelity of representation, since implementing a high-fidelity downgrade algorithm is likely more work than implementing proper support for [RFC5721] and/or [RFC5738].

The server is assumed to be internationalized internally, and to store messages internationalized messages natively. When it needs to present an internationalized message to a conventional client, it synthesizes a conventional message containing most of the information and presents that (the "synthetic message").

2. Information preserved and lost

The synthetic message is intended to convey the most important information to the user. Where information is lost, the user should see the message as incomplete rather than modified.

The synthetic message is not intended to convey any information to the client software that would require or enable it to apply special handling to the message. Client authors who wish to handle internationalized messages are encouraged to implement [RFC5738].

Upper case in examples represents non-ASCII. example.com is a plain domain, EXAMPLE.com represents a non-ASCII domain in the .com top-level domain.

2.1 Email addresses

Each internationalized email address in the header fields listed below is replaced with an invalid email address whose display-name tells the user what happened.

The format of the display-name is explicitly unspecified. Anything which tells the user what happened is good. Anything which produces an email address which might belong to someone else is bad.

Given an internationalized address "Fred Foo <fred@EXAMPLE.com>", an implementation may choose to render it e.g. as these examples:

```
"fred@EXAMPLE.com" <invalid@internationalized-address.invalid>
Fred Foo <invalid@internationalized.invalid>
internationalized-address:;
fred:;
```

(The .invalid top-level domain is reserved by [RFC2606], therefore the first two examples are syntactically valid, but will never belong to anyone. Note that the display-name often will need [RFC2047] encoding.)

The affected header fields are Bcc, Cc, From, Reply-To, Resent-Bcc, Resent-Cc, Resent-From, Resent-Sender, Resent-To, Return-Path, Sender and To. Any addresses present in other header fields, Received for example, are not regarded as addresses by this specification.

2.2 MIME parameters

Any MIME parameter [RFC2045] (whether in the message header or a bodypart header) which cannot be presented as-is to the client is silently excised.

Given a field such as

```
Content-Disposition: attachment; filename=FOO
```

the field is presented as

```
Content-Disposition: attachment
```

2.3 "Subject"

If the Subject field cannot be presented as-is, the server presents a representation encoded as specified in [RFC2047].

2.4 Remaining header fields

Any header field which cannot be presented to the client even after the modifications in sections 2.1-2.3 is silently excised.

3. IMAP-specific details

IMAP allows clients to retrieve the message size without downloading it, using RFC822.SIZE, BODY.SIZE[] and so on. [RFC3501] requires that the returned size be exact.

This specification relaxes that requirement: When a conventional client requests size information for a message, the IMAP server is permitted to return size information for the internationalized message, even though the synthetic message's size differs.

When an IMAP server carries out downgrading as part of generating nFETCH responses, it reports which messages were synthesised using a response code and attendant UID set. This can be helpful to humans debugging the server and/or client.

```
C: a UID FETCH 1:* BODY.PEEK[HEADER.FIELDS(To From Cc)]
S: 1 FETCH (UID 65 [...]
S: 2 FETCH (UID 70 [...]
S: a OK [DOWNGRADED 70,105,108,109] Done
```

The message-set argument to DOWNGRADED contains UIDs.

Note that DOWNGRADED does not necessarily mention all the internationalized messages in the mailbox. In the example above, we know that UID 65 does not contain internationalized addresses in From, To and Cc. It may contain an internationalized Subject, etc.

4. POP-specific details

The number of lines specified in the TOP command (see [RFC1939]) refers to the synthetic message. The message size reported by e.g. LIST may refer to either the internationalized or the synthetic message.

5. Security Considerations

If the internationalized message uses any sort of signature, the synthetic message's signature almost certainly is invalid. This is a necessary limitation of displaying internationalized messages in conventional clients, since the client does not support internationalized addresses.

If any excised information is significant, then that information does not arrive at the recipient. Notably, the message-id, in-reference-to and/or references fields may be excised, which might cause a lack of context when the recipient reads the message.

Some POP/IMAP clients delete the original message and use only the what they downloaded, Fetchmail is one well-known example. This may lead to permament loss of information.

Other clients cache messages for a very long time, even across client upgrades, such as the stock Android client. When such a client is internationalized, care must be taken so that it will not use an old synthetic message from its cache rather than retrieve the real message from the server.

6. Acknowledgements

Claudio Allocchio, Ned Freed, Kazunori Fujiwara, Ted Hardie, John Klensin, Barry Leiba, John Levine, Alexey Melnikov, Chris Newman, Joseph Yee and the originator of rule 12 in [RFC1925] helped with this document.

7. IANA Considerations

The IANA is requested to add DOWNGRADED to the IMAP response code registry.

(RFC editor: Please edit the previous paragraph suitably once the IANA has added the code. The registry is the one specified in RFC 5530.)

8. Normative References

[RFC1939] Myers, J and M. Rose, "Post Office Protocol - Version 3", RFC 1939, May 1996.

[RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail

Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

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

[RFC2606] Eastlake, D. and A. Panitz, "Reserved Top Level DNS Names", BCP 32, RFC 2606, June 1999.

[RFC3501] Crispin, "Internet Message Access Protocol - Version 4rev1", RFC 3501, June 2003.

9. Informative References

[RFC1925] Callon, R., "Fundamental Truths of Networking", RFC 1925, April 1996.

[RFC5721] Gellens, R., and C. Newman, "POP3 Support for UTF-8", RFC 5721, February 2010.

[RFC5738] Resnick, P. and C. Newman, "IMAP Support for UTF-8", RFC 5738, March 2010.

10. Author's Address

Arnt Gulbrandsen
Schweppermannstr. 8
D-81671 Muenchen
Germany

Fax: +49 89 4502 9758

Email: arnt@gulbrandsen.priv.no

(RFC Editor: Please delete everything after this point)

Open Issues

Should Kazunori Fujiwara's downgrade document also mention DOWNGRADED?

RFC Editor: IF 5721 and/or 5738 have been superseded by new RFCs at this time, please change the references to those RFCs throughout this document. Well, except in the previous sentence. I'm such a pedant.

RFC Editor: I do not know the difference between that and which. Will and shall outnumber me too. Please fix all that. Thank you.

Changes since -00

Added a rule to handle Subject

Removed the sentence about unknown:;

Terminology fixes

Changes since -01

Nits from Joseph Yee.

Clarified the address rendering and added non-.invalid examples, based on suggestions from Kazunori Fujiwara.

Many changes from Barry Leiba: Simplified and better terminology, reformatted examples, more references, etc.

Specified POP TOP. A bit of a no-op specification.

Mention BODY.SIZE[] as well as RFC822.SIZE. Wave hands so BODY.SIZE[1] sneaks past.

http://rant.gulbrandsen.priv.no/good-bad-rfc fwiw

Changes since -02

Added the DOWNGRADED response code, since both Barry and Alexey wants it.

Changes since -03

Added/changed text in response to appsdir reviews from Ted Hardie and Claudio Allocchio.

Changes since -04

Closed two open issues; the interest in them was clearly negligible.

"Updates: 3501" because of the SIZE relaxation.

Security considerations about download-and-delete and long-term caching.

Bring on the WGLC!

Changes since -05

Text changes from John Klensin