

**SMTP extension for internationalized email address
draft-yao-ima-smtpext-00.txt**

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on March 30, 2006.

Copyright Notice

Copyright (C) The Internet Society (2005).

Abstract

The internationalized email address (IMA) should be solved in the transport-level, which is an architecturally desirable approach. This document specifies the use of SMTP extension for internationalized email address (IMA) delivery. And also mention about the backward compatible mechanism for downgrade procedure, as specified in an associated specification. The protocol proposed here is MTA-level solution which is feasible, architecturally more elegant, and not as difficult to deploy in relevant communities.

Table of Contents

1.	Introduction	3
1.1.	Role of this specification	3
1.2.	Proposal Context	3
1.3.	Terminology	3
2.	Mail Transport-level Protocol	4
2.1.	Framework for the Internationalization Extension	4
2.2.	The Address Internationalization Service Extension	4
2.3.	Extended Mailbox Address Syntax	5
2.4.	The ALT-ADDRESS parameter	5
2.5.	Additional ESMTP Changes and Clarifications	6
2.5.1.	The Initial SMTP Exchange	6
2.5.2.	Trace Fields	6
3.	Implementation Advice	6
4.	IANA Considerations	7
5.	Security considerations	7
6.	Acknowledgements	7
7.	References	7
7.1.	Normative References	7
7.2.	Informative References	8
	Author's Address	9
	Intellectual Property and Copyright Statements	10

1. Introduction

1.1. Role of this specification

An overview document [[IMA-overview](#)] specifies the requirements for, and components of, full internationalization of electronic mail. This document specifies an element of that work, specifically the definition of an SMTP extension [[RFC1869](#)] for internationalized email address (IMA) transport delivery.

1.2. Proposal Context

In order to use internationalized email addresses, we need to internationalize both the domain part and the local part of email address. Domain part of email addresses may be internationalized through IDNA [[RFC3490](#)]. But the local part of email address still remains as non-internationalized.

The syntax of Internet email addresses is restricted to a subset of 7-bit ASCII for the domain-part, with a less-restricted subset for the local-part. These restrictions are specified in [RFC 2821](#) [[RFC2821](#)]. To be able to deliver internationalized email through SMTP servers, we need to upgrade SMTP server to able to carry internationalized email address. Since older servers and the mail-reading clients and other systems that are downstream from them will not be prepared to handle these extended addresses, an SMTP extension is specified to identify and protect the addressing mechanism.

This specification describes a change to the email transport mechanism that permits internationalized addresses in both the envelope and header fields of messages. The context for the change is described in [[IMA-overview](#)] and the details of the header changes are described in [[IMA-utf8header](#)],

1.3. Terminology

The key words "MUST", "SHALL", "REQUIRED", "SHOULD", "RECOMMENDED", and "MAY" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

All specialized terms used in this specification are defined in the IMA overview [[IMA-overview](#)] or in [[RFC2821](#)] and [[RFC2822](#)].

This document is being discussed on the ima mailing list. See <https://www1.ietf.org/mailman/listinfo/ima> for information about subscribing. The list's archive is at <http://www1.ietf.org/mail-archive/web/ima/index.html>.

2. Mail Transport-level Protocol

2.1. Framework for the Internationalization Extension

The following service extension is defined:

1. The name of the SMTP service extension is "Internationalized eMail Address";
2. The EHLO keyword value associated with this extension is "IMA";
3. No parameter values are defined for this EHLO keyword value. In order to permit future (although unanticipated) extensions, the EHLO response MUST NOT contain any parameters for that keyword. If a parameter appears, the SMTP client that is conformant to this version of this specification MUST treat the ESMTP response as if the IMA keyword did not appear.
4. An optional parameter is added to the SMTP MAIL and RCPT commands. This parameter is named ALT-ADDRESS. It requires an argument as a substitute for the internationalized (UTF-8 coded) address, which is discussed in [\[IMA-downgrading\]](#). This all-ASCII address MAY incorporate the IDNA "punycode" form if the domain name is internationalized. No algorithmic transformation is specified for the local-part; in the general case, it may identify a completely separate mailbox from the one identified in the primary command argument. The domain part of the ALT-ADDRESS may either be the same one as in the primary address (or its punycode equivalent) or may be completely different.
5. No additional SMTP verbs are defined by this extension.
6. Servers offering this extension MUST provide support for, and announce, the 8BITMIME extension [\[RFC1652\]](#).

2.2. The Address Internationalization Service Extension

An SMTP Server that announces this extension MUST be prepared to accept a UTF-8 string [\[RFC3629\]](#) in any position in which [RFC 2821](#) specifies that a "mailbox" may appear. That string must be parsed only as specified in [RFC 2821](#), i.e., by separating the mailbox into source route, local part and domain part, using only the characters colon (U+003A), comma (U+002C), and at-sign (U+0040) as specified there. Once isolated by this parsing process, the local part MUST be treated as opaque unless the SMTP Server is the final delivery MTA. Any domain names that are to be looked up in the DNS MUST be processed into punycode form as specified in IDNA [\[RFC3490\]](#) unless they are already in that form. Any domain names that are to be compared to local strings SHOULD be checked for validity and then MUST be compared as specified in IDNA.

An SMTP Client that receives the IMA extension keyword MAY transmit a mailbox name as an internationalized string in UTF-8 form. It MAY

Yao

Expires March 30, 2006

[Page 4]

transmit the domain part of that string in either punycode (derived from the IDNA process) or UTF-8 form. If it sends the domain in UTF-8 form, the original SMTP client SHOULD first verify that the string is valid for a domain name according to IDNA rules. As required by [RFC 2821](#), it MUST not attempt to parse, evaluate, or transform the local part in any way. If the IMA SMTP extension is not offered by the Server, the SMTP Client MUST not transmit an internationalized address. Instead, it MUST either return the message to the user as undeliverable or replace it. If it is replaced, the replacement MUST be either the ASCII-only address specified with the ALT-ADDRESS parameter or with an address obtained from another source that conforms to the syntax rules of [RFC 2821](#).

2.3. Extended Mailbox Address Syntax

[RFC 2821, section 4.1.2](#), defines the syntax of a mailbox as

Mailbox = Local-part "@" Domain

Local-part = Dot-string / Quoted-string
; MAY be case-sensitive

Dot-string = Atom *("." Atom)

Atom = 1*atext

Quoted-string = DQUOTE *qcontent DQUOTE

Domain = (sub-domain 1*("." sub-domain)) / address-literal
sub-domain = Let-dig [Ldh-str]

The key changes made by this specification are, informally, to

- o Change the definition of "sub-domain" to permit either the definition above or a UTF-8 string representing a DNS label that is conformant with IDNA [[RFC3490](#)]. That label MUST NOT contain the characters "@" or ".", even though those characters can normally be inserted into a DNS label.
- o Change the definition of "Atom" to permit either the definition above or a UTF-8 string. That string MUST NOT contain any of the ASCII characters (either graphics or controls) that are not permitted in "atext"; it is otherwise unrestricted.

2.4. The ALT-ADDRESS parameter

If the IMA extension is offered, the syntax of the SMTP MAIL and RCPT

commands is extended to support the optional "ALT-ADDRESS" parameter, which is specified in an associated document [[IMA-downgrading](#)].

2.5. Additional ESMTP Changes and Clarifications

The mail transport process involves addresses ("mailboxes") and domain names in contexts in addition to the MAIL and RCPT commands and extended alternatives to them. In general, the rule is that, when [RFC 2821](#) specifies a mailbox, this document expects UTF-8 to be used for the entire string; when [RFC 2821](#) specifies a domain name, the name should be in punycode form if its raw form is non-ASCII.

The following subsections list and discuss all of the relevant cases.

Support and use of this extension requires support for 8BITMIME.

2.5.1. The Initial SMTP Exchange

When an SMTP or ESMTP connection is opened, the server sends a "banner" response consisting of the 220 reply code and some information. The client then sends the EHLO command. Since the client cannot know whether the server supports internationalized addresses until after it receives the response from EHLO, any domain names that appear in this dialogue, or in responses to EHLO, must be in hostname form, i.e., internationalized ones must be in punycode form.

2.5.2. Trace Fields

Internationalized domain names in Received fields should be transmitted in punycode form. Addresses in "for" clauses need further examination and might be treated differently depending on [[IMA-utf8header](#)]. The reasoning in the introductory portion of [[IMA-overview](#)] strongly suggests that these addresses be in UTF-8 form, rather than some specialized encoding.

3. Implementation Advice

In the absence of this extension, SMTP clients and servers are constrained to using only those addresses permitted by [RFC 2821](#). The local parts of those addresses may be made up of any ASCII characters, although certain of them must be quoted as specified there. It is notable in an internationalization context that there is a long history on some systems of using overstruck ASCII characters (a character, a backspace, and another character) within a quoted string to approximate non-ASCII characters. This form of internationalization should be phased out as this extension becomes

widely deployed but backward-compatibility considerations require that it continue to be supported.

4. IANA Considerations

IANA is requested to add "IMA" to the SMTP extensions registry with the entry pointing to this specification for its definition.

5. Security considerations

See the extended security considerations discussion in [[IMA-overview](#)]

6. Acknowledgements

Much of the text in the initial version of this document was derived or copied from [[Klensin-emailaddr](#)] with the permission of the author. Significant comments and suggestions were received from Xiaodong LEE, Nai-Wen Hsu, Yangwoo KO, Yoshiro YONEYA, and other members of the JET team and were incorporated into the document.

7. References

7.1. Normative References

[ASCII] American National Standards Institute (formerly United States of America Standards Institute), "USA Code for Information Interchange", ANSI X3.4-1968, 1968.

ANSI X3.4-1968 has been replaced by newer versions with slight modifications, but the 1968 version remains definitive for the Internet.

[IMA-overview]

Klensin, J. and Y. Ko, "Overview and Framework of Internationalized Email Address Delivery", [draft-klensin-ima-framework-00](#) (work in progress), October 2005.

[IMA-utf8header]

Klensin, J. and J. Yeh, "Transmission of Email Headers in UTF-8 Encoding", [draft-yeh-utf8headers-00](#) (work in progress), October 2005.

[RFC1652] Klensin, J., Freed, N., Rose, M., Stefferud, E., and D.

- Crocker, "SMTP Service Extension for 8bit-MIMEtransport", [RFC 1652](#), July 1994.
- [RFC1869] Klensin, J., Freed, N., Rose, M., Stefferud, E., and D. Crocker, "SMTP Service Extensions", STD 10, [RFC 1869](#), November 1995.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2821] Klensin, J., "Simple Mail Transfer Protocol", [RFC 2821](#), April 2001.
- [RFC2822] Resnick, P., "Internet Message Format", [RFC 2822](#), April 2001.
- [RFC3490] Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications (IDNA)", [RFC 3490](#), March 2003.
- [RFC3492] Costello, A., "Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)", [RFC 3492](#), March 2003.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", [RFC 3629](#), November 2003.

7.2. Informative References

- [IMA-downgrading]
YONEYA, Y. and K. Fujiwara, "Downgrade Mechanism for Internationalized Email Address (IMA)", [draft-yoneya-ima-downgrade-00](#) (work in progress), October 2005.
- [Klensin-emailaddr]
Klensin, J., "Internationalization of Email Addresses", [draft-klensin-emailaddr-i18n-03](#) (work in progress), July 2005.
- [RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", [RFC 2045](#), November 1996.

Author's Address

Jiankang YAO (editor)
CNNIC
No.4 South 4th Street, Zhongguancun
Beijing

Phone: +86 10 58813007

Email: yaojk@cnnic.cn

Intellectual Property Statement

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 procedures with respect to rights in RFC 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.

Disclaimer of Validity

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.

Copyright Statement

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

Acknowledgment

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

