Email Authentication for Internationalized Mail

Abstract

   Sender Policy Framework (SPF) (RFC 7208), DomainKeys Identified Mail
   (DKIM) (RFC 6376), and Domain-based Message Authentication,
   Reporting, and Conformance (DMARC) (RFC 7489) enable a domain owner
   to publish email authentication and policy information in the DNS.
   In internationalized email, domain names can occur both as U-labels
   and A-labels. This specification updates the SPF, DKIM, and DMARC
   specifications to clarify which form of internationalized domain
   names to use in those specifications.

Status of This Memo

   This is an Internet Standards Track document.

   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). Further information on
   Internet Standards is available in Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8616.

Table of Contents

1. Introduction

   SPF [RFC7208], DKIM [RFC6376], and DMARC [RFC7489] enable a domain
   owner to publish email authentication and policy information in the
   DNS. SPF primarily publishes information about what host addresses
   are authorized to send mail for a domain. DKIM places cryptographic
   signatures on email messages, with the validation keys published in
   the DNS. DMARC publishes policy information related to the domain in
   the From: header field of email messages.

   In conventional email, all domain names are ASCII in all contexts, so
   there is no question about the representation of the domain names.
   All internationalized domain names are represented as A-labels
   [RFC5890] in message header fields, SMTP sessions, and the DNS.

   Internationalized mail [RFC6530] (generally called "EAI" for Email
   Address Internationalization) allows U-labels in SMTP sessions
   [RFC6531] and message header fields [RFC6532].

   Every U-label is equivalent to an A-label, so in principle, the
   choice of label format will not cause ambiguities. But in practice,
   consistent use of label formats will make it more likely that code
   for mail senders and receivers interoperates.

   Internationalized mail also allows UTF-8-encoded Unicode characters
   in the local parts of mailbox names, which were historically only
   ASCII.

2. Definitions

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
   "OPTIONAL" in this document are to be interpreted as described in
   BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all
   capitals, as shown here.

The term "IDN", for Internationalized Domain Name, refers to a domain
name containing either U-labels or A-labels.

Since DMARC is not currently a Standards Track protocol, this
specification offers advice rather than requirements for DMARC.

3. General Principles

In headers in EAI mail messages, domain names that were restricted to
ASCII can be U-labels, and mailbox local parts can be UTF-8. Header
field names and other text intended primarily to be interpreted by
computers rather than read by people remains ASCII.

Strings stored in DNS records remain ASCII since there is no way to
tell whether a client retrieving a DNS record expects an EAI or an
ASCII result. When a domain name found in a mail header field
includes U-labels, those labels are translated to A-labels before
being looked up in the DNS, as described in [RFC5891].

4. SPF and Internationalized Mail

SPF [RFC7208] uses two identities from the SMTP session: the host
name in the EHLO command and the domain in the address in the MAIL
FROM command. Since the EHLO command precedes the server response
that tells whether the server supports the SMTPUTF8 extension, an IDN
host name MUST be represented as A-labels. An IDN in MAIL FROM can
be either U-labels or A-labels.

All U-labels MUST be converted to A-labels before being used for an
SPF validation. This includes both the labels in the name used for
the original DNS lookup, described in Section 3 of [RFC7208], and
those used in the macro expansion of domain-spec, described in
Section 7. Section 4.3 of [RFC7208] states that all IDNs in an SPF
DNS record MUST be A-labels; this rule is unchanged since any SPF
record can be used to authorize either EAI or conventional mail.

SPF macros %{s} and %{l} expand the local part of the sender's
mailbox. If the local part contains non-ASCII characters, terms that
include %{s} or %{l} do not match anything, because non-ASCII local
parts cannot be used as the DNS labels the macros are intended to
match. Since these macros are rarely used, this is unlikely to be an
issue in practice.

5. DKIM and Internationalized Mail

DKIM [RFC6376] specifies a mail header field that contains a
cryptographic message signature and a DNS record that contains the
validation key.

Section 2.11 of [RFC6376] defines dkim-quoted-printable. Its
definition is modified in messages with internationalized header
fields so that non-ASCII UTF-8 characters need not be quoted. The
ABNF [RFC5234] for dkim-safe-char in those messages is replaced by
the following, adding non-ASCII UTF-8 characters from [RFC3629]:

```
dkim-safe-char          =  %x21-3A / %x3C / %x3E-7E /
                                UTF8-2 / UTF8-3 / UTF8-4
                        ; '!' - ':', '<', '>' - '~', non-ASCII

UTF8-2                  = <Defined in Section 4 of RFC 3629>

UTF8-3                  = <Defined in Section 4 of RFC 3629>

UTF8-4                  = <Defined in Section 4 of RFC 3629>
```

Section 3.5 of [RFC6376] states that IDNs in the d=, i=, and s= tags
of a DKIM-Signature header field MUST be encoded as A-labels. This
rule is relaxed only for internationalized message header fields
[RFC6532], so IDNs SHOULD be represented as U-labels. This provides
improved consistency with other header fields. (A-labels remain
valid to allow a transition from older software.) The set of
allowable characters in the local part of an i= tag is extended in
the same fashion as local parts of email addresses as described in
Section 3.2 of [RFC6532]. When computing or verifying the hash in a
DKIM signature as described in Section 3.7 of [RFC6376], the hash
MUST use the domain name in the format it occurs in the header field.

Section 3.4.2 of [RFC6376] describes relaxed header canonicalization.
Its first step converts all header field names from uppercase to
lowercase. Field names are restricted to printable ASCII (see
[RFC5322], Section 3.6.8), so this case conversion remains ASCII case
conversion.

DKIM key records, described in Section 3.6.1 of [RFC6376], do not
contain domain names, so there is no change to their specification.

6. DMARC and Internationalized Mail

DMARC [RFC7489] defines a policy language that domain owners can
specify for the domain of the address in an RFC5322.From header
field.

Section 6.6.1 of [RFC7489] specifies, somewhat imprecisely, how IDNs
in the RFC5322.From address domain are to be handled. That section
is updated to say that all U-labels in the domain are converted to
A-labels before further processing. Section 7.1 of [RFC7489] is

similarly updated to say that all U-labels in domains being handled
are converted to A-labels before further processing.

DMARC policy records, described in Sections 6.3 and 7.1 of [RFC7489],
can contain email addresses in the "rua" and "ruf" tags. Since a
policy record can be used for both internationalized and conventional
mail, those addresses still have to be conventional addresses, not
internationalized addresses.

7. IANA Considerations

This document has no IANA actions.

8. Security Considerations

Email is subject to a vast range of threats and abuses. This
document attempts to slightly mitigate some of them but does not, as
far as the author knows, add any new ones. The updates to SPF, DKIM,
and DMARC are intended to allow the respective specifications to work
as reliably on internationalized mail as they do on ASCII mail, so
that applications that use them, such as some kinds of mail filters
that catch spam and phish, can work more reliably on
internationalized mail.

9. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <https://www.rfc-editor.org/info/rfc2119>.

   [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO
              10646", STD 63, RFC 3629, DOI 10.17487/RFC3629, November
              2003, <https://www.rfc-editor.org/info/rfc3629>.

   [RFC5234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax
              Specifications: ABNF", STD 68, RFC 5234,
              DOI 10.17487/RFC5234, January 2008,
              <https://www.rfc-editor.org/info/rfc5234>.

   [RFC5322] Resnick, P., Ed., "Internet Message Format", RFC 5322,
              DOI 10.17487/RFC5322, October 2008,
              <https://www.rfc-editor.org/info/rfc5322>.

   [RFC5890] Klensin, J., "Internationalized Domain Names for
              Applications (IDNA): Definitions and Document Framework",
              RFC 5890, DOI 10.17487/RFC5890, August 2010,
              <https://www.rfc-editor.org/info/rfc5890>.

   [RFC5891] Klensin, J., "Internationalized Domain Names in
              Applications (IDNA): Protocol", RFC 5891,
              DOI 10.17487/RFC5891, August 2010,
              <https://www.rfc-editor.org/info/rfc5891>.

   [RFC6376] Crocker, D., Ed., Hansen, T., Ed., and M. Kucherawy, Ed.,
              "DomainKeys Identified Mail (DKIM) Signatures", STD 76,
              RFC 6376, DOI 10.17487/RFC6376, September 2011,
              <https://www.rfc-editor.org/info/rfc6376>.

   [RFC6530] Klensin, J. and Y. Ko, "Overview and Framework for
              Internationalized Email", RFC 6530, DOI 10.17487/RFC6530,
              February 2012, <https://www.rfc-editor.org/info/rfc6530>.

   [RFC6531] Yao, J. and W. Mao, "SMTP Extension for Internationalized
              Email", RFC 6531, DOI 10.17487/RFC6531, February 2012,
              <https://www.rfc-editor.org/info/rfc6531>.

   [RFC6532] Yang, A., Steele, S., and N. Freed, "Internationalized
              Email Headers", RFC 6532, DOI 10.17487/RFC6532, February
              2012, <https://www.rfc-editor.org/info/rfc6532>.

   [RFC7208] Kitterman, S., "Sender Policy Framework (SPF) for
              Authorizing Use of Domains in Email, Version 1", RFC 7208,
              DOI 10.17487/RFC7208, April 2014,
              <https://www.rfc-editor.org/info/rfc7208>.

   [RFC7489] Kucherawy, M., Ed. and E. Zwicky, Ed., "Domain-based
              Message Authentication, Reporting, and Conformance
              (DMARC)", RFC 7489, DOI 10.17487/RFC7489, March 2015,
              <https://www.rfc-editor.org/info/rfc7489>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

Author's Address

   John Levine
   Taughannock Networks
   PO Box 727
   Trumansburg, NY 14886
   United States of America

   Email: standards@taugh.com
   URI: http://jl.ly