RFC 4871 DomainKeys Identified Mail (DKIM) Signatures -- Update

Abstract

   This document updates RFC 4871, "DomainKeys Identified Mail (DKIM)
   Signatures". Specifically, the document clarifies the nature, roles,
   and relationship of the two DKIM identifier tag values that are
   candidates for payload delivery to a receiving processing module.
   The Update is in the style of an Errata entry, albeit a rather long
   one.

Status of This Memo

   This document specifies an Internet standards track protocol for the
   Internet community, and requests discussion and suggestions for
   improvements. Please refer to the current edition of the "Internet
   Official Protocol Standards" (STD 1) for the standardization state
   and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Table of Contents

1. Introduction

   About the purpose for DKIM, [RFC4871] states:

      The ultimate goal of this framework is to permit a signing domain
      to assert responsibility for a message, thus protecting message
      signer identity...

   Hence, DKIM has a signer that produces a signed message, a verifier
   that confirms the signature, and an assessor that consumes the
   validated signing domain. So, the simple purpose of DKIM is to
   communicate an identifier to a receive-side assessor module. The
   identifier is in the form of a domain name that refers to a
   responsible identity. For DKIM to be interoperable and useful, the
   signer and assessor must share the same understanding of the details
   about the identifier.

   However, the RFC 4871 specification defines two, potentially
   different, identifiers that are carried in the DKIM-Signature: header
   field, d= and i=. Either might be delivered to a receiving
   processing module that consumes validated payload. The DKIM
   specification fails to clearly define which is the "payload" to be
   delivered to a consuming module, versus what is internal and merely
   in support of achieving payload delivery.

This currently leaves signers and assessors with the potential for
making different interpretations between the two identifiers and may
lead to interoperability problems. A signer could intend one to be
used for assessment, and have a different intent in setting the value
in the other. However the verifier might choose the wrong value to
deliver to the assessor, thereby producing an unintended (and
inaccurate) assessment.

This Update resolves that confusion. It defines additional, semantic
labels for the two values, clarifies their nature, and specifies
their relationship. More specifically, it clarifies that the
identifier intended for delivery to the assessor -- such as one that
consults a whitelist -- is the value of the "d=" tag. However, this
does not prohibit message filtering engines from using the "i=" tag,
or any other information in the message's header, for filtering
decisions.

For signers and verifiers that have been using the i= tag as the
primary value that is delivered to the assessor, a software change to
using the d= tag is intended.

So, this Update clarifies the formal interface to DKIM, after
signature verification has been performed. It distinguishes DKIM's
internal signing and verification activity, from its standardized
delivery of data to that interface.

The focus of the Update is on the portion of DKIM that is much like
an API definition. If DKIM is implemented as a software library for
use by others, it needs to define what outputs are provided, that is,
what data that an application developer who uses the library can
expect to obtain as a result of invoking DKIM on a message.

This Update defines the output of that library to include the yes/no
result of the verification and the "d=" value. In other words, it
says what (one) identifier was formally specified for use by the
signer and whether the use of that identifier has been validated.
For a particular library, other information can be provided at the
discretion of the library developer, since developers of assessors --
these are the consumers of the DKIM library -- well might want more
information than the standardized two pieces of information.
However, that standardized set is the minimum that is required to be
provided to a consuming module, in order to be able to claim that the
library is DKIM compliant.

This does not state what the implicit value of "i=" is, relative to
"d=". In this context, that fact is irrelevant.

Another example is the difference between the socket interface to TCP
versus the TCP protocol itself. There is the activity within the
protocol stack, and then there is the activity within in the software
libraries that are actually used.

NOTE: The text provided here updates [RFC4871]. Text appearing in
   the "Corrected Text:" replaces text in RFC 4871. Hence,
   references that appear in the "Original Text:" can be found in RFC
   4871, and are not duplicated 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 [RFC2119].

2. RFC 4871, Abstract

   Original Text:

      The ultimate goal of this framework is to permit a signing domain
      to assert responsibility for a message,

   Corrected Text:

      The ultimate goal of this framework is to permit a person, role or
      organization that owns the signing domain to assert responsibility
      for a message,

3. RFC 4871, Section 1, Introduction

   Original Text:

      ...permitting a signing domain to claim responsibility

   Corrected Text:

      permitting a person, role, or organization that owns the signing
      domain to claim responsibility

4. RFC 4871, Section 2.7, Identity

   Original Text:

      (None. New section. Additional text.)

Corrected Text:

   A person, role, or organization. In the context of DKIM, examples
   include author, author's organization, an ISP along the handling
   path, an independent trust assessment service, and a mailing list
   operator.

5. RFC 4871, Section 2.8, Identifier

   Original Text:

   (None. New section. Additional text.)

   Corrected Text:

   A label that refers to an identity.

6. RFC 4871, Section 2.9, Signing Domain Identifier (SDID)

   Original Text:

   (None. New section. Additional text.)

   Corrected Text:

   A single domain name that is the mandatory payload output of DKIM
   and that refers to the identity claiming responsibility for
   introduction of a message into the mail stream. For DKIM
   processing, the name has only basic domain name semantics; any
   possible owner-specific semantics are outside the scope of DKIM.
   It is specified in Section 3.5.

7. RFC 4871, Section 2.10, Agent or User Identifier (AUID)

   Original Text:

   (None. New section. Additional text.)

   Corrected Text:

   A single identifier that refers to the agent or user on behalf of
   whom the Signing Domain Identifier (SDID) has taken
   responsibility. The AUID comprises a domain name and an optional
   <Local-part>. The domain name is the same as that used for the
   SDID or is a sub-domain of it. For DKIM processing, the domain
   name portion of the AUID has only basic domain name semantics; any
   possible owner-specific semantics are outside the scope of DKIM.
   It is specified in Section 3.5.

8. RFC 4871, Section 2.11, Identity Assessor

    Original Text:

       (None. New section. Additional text.)

    Corrected Text:

       A module that consumes DKIM's mandatory payload, which is the
       responsible Signing Domain Identifier (SDID). The module is
       dedicated to the assessment of the delivered identifier. Other
       DKIM (and non-DKIM) values can also be delivered to this module as
       well as to a more general message evaluation filtering engine.
       However, this additional activity is outside the scope of the DKIM
       signature specification.

9. RFC 4871, Section 3.5, The DKIM-Signature Header Field

    Original Text:

    d= The domain of the signing entity (plain-text; REQUIRED). This is
        the domain that will be queried for the public key. This domain
        MUST be the same as or a parent domain of the "i=" tag (the
        signing identity, as described below), or it MUST meet the
        requirements for parent domain signing described in Section 3.8.
        When presented with a signature that does not meet these
        requirement, verifiers MUST consider the signature invalid.

        Internationalized domain names MUST be encoded as described in
        [RFC3490].

        ABNF:

           sig-d-tag   = %x64 [FWS] "=" [FWS] domain-name
           domain-name   = sub-domain 1*("." sub-domain)
                   ; from RFC 2821 Domain,
                     but excluding address-literal

    Corrected Text:

       d=

          Specifies the SDID claiming responsibility for an introduction
          of a message into the mail stream (plain-text; REQUIRED).
          Hence, the SDID value is used to form the query for the public
          key. The SDID MUST correspond to a valid DNS name under which
          the DKIM key record is published. The conventions and
          semantics used by a signer to create and use a specific SDID

       are outside the scope of the DKIM Signing specification, as is
       any use of those conventions and semantics. When presented
       with a signature that does not meet these requirements,
       verifiers MUST consider the signature invalid.

       Internationalized domain names MUST be encoded as described in
       [RFC3490].

       ABNF:

          sig-d-tag   = %x64 [FWS] "=" [FWS] domain-name
         domain-name = sub-domain 1*("." sub-domain)
                         ; from RFC 5321 Domain,
                           but excluding address-literal

10. RFC 4871, Section 3.5, The DKIM-Signature Header Field

    Original Text:

    i= Identity of the user or agent (e.g., a mailing list manager) on
        behalf of which this message is signed (dkim-quoted-printable;
        OPTIONAL, default is an empty Local-part followed by an "@"
        followed by the domain from the "d=" tag). The syntax is a
        standard email address where the Local-part MAY be omitted. The
        domain part of the address MUST be the same as or a subdomain of
        the value of the "d=" tag.

        Internationalized domain names MUST be converted using the steps
        listed in Section 4 of [RFC3490] using the "ToASCII" function.

        ABNF:

           sig-i-tag = %x69 [FWS] "=" [FWS]
                        [Local-part] "@" domain-name

        INFORMATIVE NOTE: The Local-part of the "i=" tag is optional
        because in some cases a signer may not be able to establish a
        verified individual identity. In such cases, the signer may
        wish to assert that although it is willing to go as far as
        signing for the domain, it is unable or unwilling to commit
        to an individual user name within their domain. It can do so
        by including the domain part but not the Local-part of the
        identity.

        INFORMATIVE DISCUSSION: This document does not require the value
        of the "i=" tag to match the identity in any message header
        fields. This is considered to be a verifier policy issue.
        Constraints between the value of the "i=" tag and other

identities in other header fields seek to apply basic
authentication into the semantics of trust associated with a
role such as content author. Trust is a broad and complex
topic and trust mechanisms are subject to highly creative
attacks. The real-world efficacy of
bindings between the "i=" value and other identities is not
well established, nor is its vulnerability to subversion by
an attacker. Hence reliance on the use of these options
should be strictly limited. In particular, it is not at all
clear to what extent a typical end-user recipient can rely on
any assurances that might be made by successful use of the
"i=" options.

Corrected Text:

i=

    The Agent or User Identifier (AUID) on behalf of which the SDID
    is taking responsibility (dkim-quoted-printable; OPTIONAL,
    default is an empty Local-part followed by an "@" followed by
    the domain from the "d=" tag).

    The syntax is a standard email address where the Local-part MAY
    be omitted. The domain part of the address MUST be the same
    as, or a subdomain of the value of, the "d=" tag.

    Internationalized domain names MUST be converted using the
    steps listed in Section 4 of [RFC3490] using the "ToASCII"
    function.

    ABNF:

       sig-i-tag = %x69 [FWS] "=" [FWS]
                    [Local-part] "@" domain-name

    The AUID is specified as having the same syntax as an email
    address, but is not required to have the same semantics.
    Notably, the domain name is not required to be registered in
    the DNS -- so it might not resolve in a query -- and the Local-
    part MAY be drawn from a namespace that does not contain the
    user's mailbox. The details of the structure and semantics for
    the namespace are determined by the Signer. Any knowledge or
    use of those details by verifiers or assessors is outside the
    scope of the DKIM Signing specification. The Signer MAY choose
    to use the same namespace for its AUIDs as its users' email
    addresses or MAY choose other means of representing its users.
    However, the signer SHOULD use the same AUID for each message
    intended to be evaluated as being within the same sphere of

responsibility, if it wishes to offer receivers the option of
using the AUID as a stable identifier that is finer grained
than the SDID.

INFORMATIVE NOTE: The Local-part of the "i=" tag is optional
because, in some cases, a signer may not be able to establish a
verified individual identity. In such cases, the signer might
wish to assert that although it is willing to go as far as
signing for the domain, it is unable or unwilling to commit to
an individual user name within their domain. It can do so by
including the domain part but not the Local-part of the
identity.

11. RFC 4871, Section 3.8, Signing by Parent Domains

Original Text:

e.g., a key record for the domain example.com can be used to
verify messages where the signing identity ("i=" tag of the
signature) is sub.example.com, or even sub1.sub2.example.com. In
order to limit the capability of such keys when this is not
intended, the "s" flag may be set in the "t=" tag of the key
record to constrain the validity of the record to exactly the
domain of the signing identity. If the referenced key record
contains the "s" flag as part of the "t=" tag, the domain of the
signing identity ("i=" flag) MUST be the same as that of the d=
domain. If this flag is absent, the domain of the signing
identity MUST be the same as, or a subdomain of, the d= domain.

Corrected Text:

...for example, a key record for the domain example.com can be
used to verify messages where the AUID ("i=" tag of the signature)
is sub.example.com, or even sub1.sub2.example.com. In order to
limit the capability of such keys when this is not intended, the
"s" flag MAY be set in the "t=" tag of the key record, to
constrain the validity of the domain of the AUID. If the
referenced key record contains the "s" flag as part of the "t="
tag, the domain of the AUID ("i=" flag) MUST be the same as that
of the SDID (d=) domain. If this flag is absent, the domain of
the AUID MUST be the same as, or a subdomain of, the SDID.

12. RFC 4871, Section 3.9, Relationship between SDID and AUID

   Original Text: (None. New section. Additional text.)

   Corrected Text:

      DKIM's primary task is to communicate from the Signer to a
      recipient-side Identity Assessor a single Signing Domain
      Identifier (SDID) that refers to a responsible identity. DKIM MAY
      optionally provide a single responsible Agent or User Identifier
      (AUID).

      Hence, DKIM's mandatory output to a receive-side Identity Assessor
      is a single domain name. Within the scope of its use as DKIM
      output, the name has only basic domain name semantics; any
      possible owner-specific semantics are outside the scope of DKIM.
      That is, within its role as a DKIM identifier, additional
      semantics cannot be assumed by an Identity Assessor.

      A receive-side DKIM verifier MUST communicate the Signing Domain
      Identifier (d=) to a consuming Identity Assessor module and MAY
      communicate the Agent or User Identifier (i=) if present.

      To the extent that a receiver attempts to intuit any structured
      semantics for either of the identifiers, this is a heuristic
      function that is outside the scope of DKIM's specification and
      semantics. Hence, it is relegated to a higher-level service, such
      as a delivery handling filter that integrates a variety of inputs
      and performs heuristic analysis of them.

      INFORMATIVE DISCUSSION: This document does not require the value
      of the SDID or AUID to match the identifier in any other message
      header field. This requirement is, instead, an assessor policy
      issue. The purpose of such a linkage would be to authenticate the
      value in that other header field. This, in turn, is the basis for
      applying a trust assessment based on the identifier value. Trust
      is a broad and complex topic and trust mechanisms are subject to
      highly creative attacks. The real-world efficacy of any but the
      most basic bindings between the SDID or AUID and other identities
      is not well established, nor is its vulnerability to subversion by
      an attacker. Hence, reliance on the use of such bindings should
      be strictly limited. In particular, it is not at all clear to
      what extent a typical end-user recipient can rely on any
      assurances that might be made by successful use of the SDID or
      AUID.

13. RFC 4871, Section 6.3, Interpret Results/Apply Local Policy

     Original Text:

        It is beyond the scope of this specification to describe what
        actions a verifier system should make, but an authenticated email
        presents an opportunity to a receiving system that unauthenticated
        email cannot. Specifically, an authenticated email creates a
        predictable identifier by which other decisions can reliably be
        managed, such as trust and reputation. Conversely,
        unauthenticated email lacks a reliable identifier that can be used
        to assign trust and reputation.

     Corrected Text:

        It is beyond the scope of this specification to describe what
        actions an Identity Assessor can make, but mail carrying a
        validated SDID presents an opportunity to an Identity Assessor
        that unauthenticated email does not. Specifically, an
        authenticated email creates a predictable identifier by which
        other decisions can reliably be managed, such as trust and
        reputation.

14. RFC 4871, Section 6.3, Interpret Results/Apply Local Policy

     Original Text:

        Once the signature has been verified, that information MUST be
        conveyed to higher-level systems (such as explicit allow/
        whitelists and reputation systems) and/or to the end user. If the
        message is signed on behalf of any address other than that in the
        From: header field, the mail system SHOULD take pains to ensure
        that the actual signing identity is clear to the reader.

     Corrected Text:

        Once the signature has been verified, that information MUST be
        conveyed to the Identity Assessor (such as an explicit allow/
        whitelist and reputation system) and/or to the end user. If the
        SDID is not the same as the address in the From: header field, the
        mail system SHOULD take pains to ensure that the actual SDID is
        clear to the reader.

15. RFC 4871, Appendix D, MUA Considerations

   Original Text:

      The tendency is to have the MUA highlight the address associated
      with this signing identity in some way, in an attempt to show the
      user the address from which the mail was sent.

   Corrected Text:

      The tendency is to have the MUA highlight the SDID, in an attempt
      to show the user the identity that is claiming responsibility for
      the message.

16. Security Considerations

   This Update clarifies core details about DKIM's payload. As such, it
   affects interoperability, semantic characterization, and the
   expectations for the identifiers carried with a DKIM signature.
   Clarification of these details is likely to limit misinterpretation
   of DKIM's semantics. Since DKIM is fundamentally a security
   protocol, this should improve its security characteristics.

17. Normative References

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

   [RFC3490] Faltstrom, P., Hoffman, P., and A. Costello,
              "Internationalizing Domain Names in Applications (IDNA)",
              RFC 3490, March 2003.

   [RFC4871] Allman, E., Callas, J., Delany, M., Libbey, M., Fenton,
              J., and M. Thomas, "DomainKeys Identified Mail (DKIM)
              Signatures", RFC 4871, May 2007.

Appendix A. ABNF Fragments

   This appendix contains the full set of corrected ABNF fragments
   defined in this document.

```
        sig-d-tag   = %x64 [FWS] "=" [FWS] domain-name
      domain-name = sub-domain 1*("." sub-domain)
                    ; from RFC 5321 Domain,
                      but excluding address-literal

       sig-i-tag =  %x69 [FWS] "=" [FWS]
                    [ Local-part ] "@" domain-name
```

Appendix B. Acknowledgements

   This document was initially formulated by an ad hoc design team,
   comprising: Jon Callas, D. Crocker, J. D. Falk, Michael Hammer, Tony
   Hansen, Murray Kucherawy, John Levine, Jeff Macdonald, Ellen Siegel,
   and Wietse Venema. The final version of the document was developed
   through vigorous discussion in the IETF DKIM working group.

Author's Address

   D. Crocker (editor)
   Brandenburg InternetWorking

   Phone: +1.408.246.8253
   EMail: dcrocker@bbiw.net