List-Id: A Structured Field and Namespace for the Identification of Mailing Lists
draft-moonesamy-rfc2919bis-04

Abstract

   Software that handles electronic mailing list messages (servers and
   user agents) needs a way to reliably identify messages that belong to
   a particular mailing list. With the advent of list header fields, it
   has become even more important to provide a unique identifier for a
   mailing list regardless of the particular host that runs the mailing
   list manager at any given time.

   The List-Id header field provides a standard location for such an
   identifier. In addition, a namespace for mailing list identifiers
   based on fully qualified domain names is described. This namespace
   is intended to guarantee uniqueness for list owners who require it,
   while allowing for a less rigorous namespace for experimental and
   personal use.

   By including the List-Id field, mailing list managers can make it
   easier for mail user agents to provide automated tools for users to
   perform mailing list functions. The mailing list identifier can
   serve as a key for automated processing tasks

Status of this Memo

   This Internet-Draft will expire on August 23, 2012.

Copyright Notice

Table of Contents

1. Introduction

   Internet mailing lists have evolved into fairly sophisticated forums
   for group communication and collaboration. draft-moonesamy-rfc2369bis
   [ID.moonesamy-rfc2369bis] expands the functionality that the Mail
   User Agent (MUA) can provide by providing more information in each
   message sent by the mailing list manager (MLM).

   Implementing such functionality in the MUA depends on the ability to
   accurately identify messages as belonging to a particular mailing
   list. The problem then becomes what attribute or property to use to
   identify a mailing list. The most likely candidate is the submission
   address of the mailing list itself. Unfortunately, when the server
   where the mailing list manager is hosted or the submission policy of
   the list changes the submission address itself can change. This
   affects automated processing and filtering of messages from mailing
   lists.

   In order to further automate (and make more accurate) the processing
   a software agent can do, RFC 2919 [RFC2919] specified how to generate
   a unique identifier for a mailing list. This identifier can be
   simply used for string matching in a filter, or it can be used in
   more sophisticated systems to uniquely identify messages as belonging
   to a particular mailing list independent of the particular host
   delivering the actual messages. This identifier can also act as a
   key into a database of mailing lists.

   This document obsoletes RFC 2919 [RFC2919].

1.1. 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 [RFC2119].

1.2. Syntax Notation

   This specification uses the Augmented Backus-Naur Form (ABNF)
   [RFC5234] notation for the formal definitions of the syntax of list
   header fields.

1.3. Note

   This Internet-Draft can be discussed on the apps-discuss@ietf.org
   mailing list. [RFC-Editor: please remove this paragraph]

[2](#). List Identifier Syntax

   The List Identifier will, in most cases, appear like a host name in a
   domain of the list owner. Using the domain name space as a basis for
   the List Identifier namespace, it is intended to leverage an existing
   name space structure to generate a unique identifier. It also allows
   for the List Identifier to be kept separate from any particular
   delivery host or mechanism.

   While it is perfectly acceptable for a List Identifier to be
   completely independent of the domain name of the host machine
   servicing the mailing list, the owner of a mailing list MUST NOT
   generate List Identifiers in any domain name space for which they do
   not have authority. For example, a mailing list hosting service may
   choose to assign List Identifiers in their own domain-based name
   space, or they may allow their clients (the list owners) to provide
   List Identifiers in a namespace for which the owner has authority.

   If the owner of the mailing list does not have the authority to
   create a List Identifier in a domain-based name space, they may
   create unmanaged List Identifiers in the special unmanaged domain
   "invalid".

   Syntax:

      list-id = list-label "." list-id-namespace

      list-label = dot-atom-text

      list-id-namespace = domain-name / unmanaged-list-id-namespace

      unmanaged-list-id-namespace = "invalid"

      domain-name = dot-atom-text

   Where:

      dot-atom-text is defined in [RFC5322]

      "invalid" is a reserved domain name defined in [RFC2606]

   In addition, a List Identifier (list-id) MUST NOT be longer than 255
   octets in length, for future compatibility. It should be noted that
   "invalid" is not valid for the domain-name rule.

3. **List-Id Header Field**

   This document specifies a List-Id header field which will provide an
   identifier for a mailing list. This header field should be included
   on all messages distributed by the mailing list manager (including
   command responses to individual users), and on other messages where
   the message clearly applies to this particular distinct mailing list.
   There MUST be not be more than one of List-Id header field in any
   given message.

   This field MUST only be generated by mailing list managers, not MUAs.

   The contents of the List-Id header field mostly consist of angle-
   bracket ('<', '>') enclosed identifier, with internal whitespace
   being ignored. Mailing List Managers MUST NOT insert whitespace
   within the brackets. Client applications SHOULD treat any such
   whitespace, that might be inserted by poorly behaved mailing list
   managers, as characters to ignore.

   The List-Id header field is subject to the encoding and character
   restrictions for mail headers as described in [RFC5322].

   The List-Id header field MAY optionally include a description by
   including it as a "phrase" [RFC5322] before the angle-bracketed List
   Identifier. The MUA MAY choose to use this description in its user
   interface; however, any MUA that intends to make use of the
   description should be prepared to properly parse and decode any
   encoded strings or other legal phrase components. For many MUAs the
   parsing of the List-Id header field will simply consist of extracting
   the List Identifier from between the delimiting angle brackets.

   Syntax:

      list-id-header = "List-Id:" [phrase] "<" list-id ">" CRLF

   where phrase and CRLF are as defined in [RFC5322]. Unlike most
   headers fields in RFC 5322 [RFC5322], the List-Id header field does
   not allow free insertion of whitespace and comments around tokens.
   Any descriptive text MUST be presented in the optional phrase
   component of the header field.

   Examples:

      List-Id: List Header Mailing List <list-name.example.com>

      List-Id: <commonspace-users.list-id.example.com>

      List-Id: "Lena's Personal Joke List" <lenas-
         jokes.da39efc25c530ad145d41b86f7420c3b.021999.invalid>

      List-Id: <da39efc25c530ad145d41b86f7420c3b.052000.invalid>

[4](). **Persistence of List Identifiers**

   Although the List Identifier may be changed by the mailing list
   administrator, this is not desirable. (Note that there is no
   disadvantage to changing the description portion of the List-Id
   header field.) A MUA would not recognize the change to the List
   Identifier because a MUA treats a different List Identifier as a
   different list. As such the mailing list administrator should avoid
   changing the List Identifier even when the host serving the mailing
   list changes. On the other hand, transitioning from an informal
   unmanaged-list-id-namespace to a domain name space is an acceptable
   reason to change the List Identifier. Also if the focus of the
   mailing list changes sufficiently the administrator may wish to
   retire the previous mailing list and its associated identifier to
   start a new mailing list reflecting the new focus.

[5](). **Uniqueness of List Identifiers**

   This proposal uses a namespace that by definition can be used to
   create unique identifiers within the domain.

   There is a need for identification of mailing lists that are
   administrated by some entity without administrative access to a
   domain. In this case, general heuristics can be given to reduce the
   chance of collision, but it cannot be guaranteed. If a list owner
   requires a guarantee, they are free to register a domain name under
   their control.

   It is suggested, but not required, that List Identifiers be created
   under a subdomain of "list-id" within any given domain. This can
   help to reduce internal conflicts between the administrators of the

subdomains of large organizations. For example, List Identifiers at
"example.com" are generated in the subdomain of "list-
id.example.com".

A List Identifier not ending with ".invalid" MUST be globally unique
in reference to all other mailing lists. List Identifiers ending
with ".invalid" are not guaranteed to be globally unique.

A List Identifier using the special "invalid" namespace SHOULD
include the month and year (in the form MYYYY), i.e. the List
Identifier is a "subdomain" of the "invalid" namespace. In addition,
some portion of the List Identifier MUST be a randomly generated
string, e.g. the random [RFC4086] component contains a hex encoding
of 128 bits of randomness (resulting in 32 hex characters) as part of
the List Identifier.

Thus, List Identifiers such as <lenas-
jokes.da39efc25c530ad145d41b86f7420c3b.021999.invalid> and
<da39efc25c530ad145d41b86f7420c3b.051998.invalid> follow these
guidelines, while <lenas-jokes.021999.invalid> and <mylist.invalid>
do not.

6. Operations on List Identifier

There is only one operation defined for List Identifiers, that of
case insensitive equality.

The sole use of a List Identifier is to identify a mailing list, and
the sole use of the List-Id header field is to mark a particular
message as belonging to that list. The comparison operation MUST
ignore any part of the List-Id header field outside of the angle
brackets, the MUA may inform the user if the descriptive name of a
mailing list changes.

7. Supporting Nested Lists

A list that is a sublist for another list in a nested mailing list
hierarchy MUST NOT modify the List-Id header field; however, this
will only be possible when the nested mailing list is aware of the
relationship between it and its "parent" mailing lists.

8. Security Considerations

There are very few new security concerns generated with this
proposal. Message headers fields are an existing standard, designed

to easily accommodate new types. There may be concern with multiple
header fields being inserted or headers fields being forged, but
these are problems inherent in Internet mail, not specific to this
specification. If a mailing list manager encounters List-Id header
fields from any unexpected source it SHOULD NOT pass them through to
the mailing list.

As mentioned above, mail list managers should not allow any user-
originated List-Id header fields to pass through to their lists, lest
they confuse the user and have the potential to create security
problems.

On the client side, a forged List Identifier may break automated
processing. The List Identifier (in its current form) should not be
used as an indication of the authenticity of the message.

9. IANA Considerations

The List-Id reference in the Permanent Message Header Field Names
registry should be updated to point to this document.

10. Acknowledgements

The numerous participants of the List-Header and ListMom-Talk mailing
lists contributed much to the formation and structure of this
document.

Grant Neufeld focused much of the early discussion, and thus was
essential in the creation of this document.

Murray S. Kucherawy provided insightful comments.

11. References

11.1. Normative References

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

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax
           Specifications: ABNF", STD 68, RFC 5234, January 2008.

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

11.2. Informative References

[I-D.moonesamy-rfc2369bis]
            Moonesamy, S., Neufeld, G., and J. Baer, "The Use of URIs
            as Meta-Syntax for Core Mail List Commands and their
            Transport through Message Header Fields",
            draft-moonesamy-rfc2369bis-03 (work in progress),
            January 2012.

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

[RFC2919]   Chandhok, R. and G. Wenger, "List-Id: A Structured Field
            and Namespace for the Identification of Mailing Lists",
            RFC 2919, March 2001.

[RFC4086]   Eastlake, D., Schiller, J., and S. Crocker, "Randomness
            Requirements for Security", BCP 106, RFC 4086, June 2005.

Appendix A. Subject Tags

A popular feature of some MLMs is the "tagging" of the Subject header
field by prefixing the header field's contents with the name of the
mailing list, e.g. "[example]" for a list called "example". The
difference between a List Identifier and a subject tag is that while
a List Identifier is unique, a subject tag is usually the short form
of a mailing list's name. Such a namespace is inherently flat,
unmanaged and thus non-unique.

Appendix B. Changes from RFC 2919

This appendix contains a list of changes between this document and
RFC 2919.

o   Removed text about the domain name system and domain ownership in
    Section 2

o   "localhost" replaced with "invalid"

o   Replaced MTAs with mailing list managers in the sentence: "MTAs
    MUST NOT insert whitespace within the brackets" in Section 3

o   Case independence in Section 6 changed to case insensitivity for
    ASCII

o Added a paragraph in the appendix about subject tags

o Updated references

o Editorial changes

Appendix C. **Change Log**

[RFC-Editor: please remove this section]

C.1. **Changes between between version -03 and version -04**

o Removed List-Sequence header field as it does not fit it

o Reverted recommended to MUST NOT in Section 2 as it is a
   requirement in RFC 2919

C.2. **Changes between between version -02 and version -03**

o Added Ravinder Chandhok as co-author

o Added Geoffrey Wenger as co-author

C.3. **Changes between between version -01 and version -02**

o Added List-Sequence header field

o Removed text about about case-insensitive string comparison

o Removed recommendation in Section 4 about how the MUA SHOULD treat
   a different List Identifier

C.4. **Changes between between version -00 and version -01**

o Added informative reference to I-D.moonesamy-rfc2369bis

Authors' Addresses

   S. Moonesamy (editor)
   76, Ylang Ylang Avenue
   Quatre Bornes
   Mauritius

   Email: sm+ietf@elandsys.com

Ravinder Chandhok
QUALCOMM, Inc.
5775 Morehouse Drive
San Diego, CA 92121
USA

Email: chandhok@qualcomm.com

Geoffrey Wenger
QUALCOMM, Inc.
5775 Morehouse Drive
San Diego, CA 92121
USA

Email: gwenger@qualcomm.com