Network Working Group
Request for Comments: 2919
Category: Standards Track

R. Chandhok
G. Wenger
QUALCOMM, Inc.

March 2001

List-Id: A Structured Field and Namespace for the Identification of Mailing Lists

Status of this Memo

Copyright Notice

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 management headers, it has become even more important to provide a unique identifier for a mailing list regardless of the particular host that serves as the list processor at any given time.

The List-Id header provides a standard location for such an identifier. In addition, a namespace for 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, list servers can make it easier for mail clients to provide automated tools for users to perform list functions. The list identifier can serve as a key to make many automated processing tasks easier, and hence more widely available.

1. Introduction

Internet mailing lists have evolved into fairly sophisticated forums for group communication and collaboration; however, corresponding changes in the underlying infrastructure have lagged behind. Recent

proposals like [RFC2369] have expanded the functionality that the MUA can provide by providing more information in each message sent by the mailing list distribution software.

Actually 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 list server host, the list processing software, or the submission policy of the list changes the submission address itself can change. This causes great difficulty for automated processing and filtering.

In order to further automate (and make more accurate) the processing a software agent can do, there needs to be some unique identifier to use as an identifier for the 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.

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 RFC 2119.

2. The List Identifier Syntax

The list identifier will, in most cases, appear like a host name in a domain of the list owner. In other words, the domain name system is used to delegate namespace authority for list identifiers just as it has been used to distribute that authority for other internet resources.

Using the domain name system as a basis for the list identifier namespace is intended to leverage an existing authority structure into a new area of application. By using the domain name system to delegate list identifier namespace authority, it becomes instantly clear who has the right to create a particular list identifier, and separates the list identifier from any particular delivery host or mechanism. Only the rights-holder of a domain or subdomain has the authority to create list identifiers in the namespace of that domain. For example, only the rights-holder to the "acm.org" domain has the authority to create list identifiers in "acm.org" domain.

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 namespace 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 namespace, 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 list does not have the authority to create a list identifier in a domain-based namespace, they may create unmanaged list identifiers in the special unmanaged domain "localhost". This would apply to personal users, or users unable to afford domain name registration fees.

The syntax for a list identifier in ABNF [RFC2234] follows:

```
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    = "localhost"

domain-name = dot-atom-text
```

Where:

dot-atom-text is defined in [DRUMS]

"localhost" is a reserved domain name is 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 "localhost" is not valid for the domain-name rule.

3. The List-Id Header Field

This document presents a header field which will provide an identifier for an e-mail distribution list. This header SHOULD be included on all messages distributed by the list (including command responses to individual users), and on other messages where the message clearly applies to this particular distinct list. There MUST be no more than one of each field present in any given message.

This field MUST only be generated by mailing list software, not end users.

The contents of the List-Id header mostly consist of angle-bracket ('<', '>') enclosed identifier, with internal whitespace being ignored. MTAs MUST NOT insert whitespace within the brackets, but client applications should treat any such whitespace, that might be inserted by poorly behaved MTAs, as characters to ignore.

The list header fields are subject to the encoding and character restrictions for mail headers as described in [RFC822].

The List-Id header MAY optionally include a description by including it as a "phrase" [DRUMS] 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 will simply consist of extracting the list identifier from between the delimiting angle brackets.

The syntax of the List-Id header follows:

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

where phrase and CRLF are as defined in [DRUMS]. Unlike most headers in [RFC822], the List-Id header 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.

Examples:

```
List-Id: List Header Mailing List <list-header.nisto.com>
List-Id: <commonspace-users.list-id.within.com>
List-Id: "Lena's Personal Joke List"
         <lenas-jokes.da39efc25c530ad145d41b86f7420c3b.021999.localhost>
List-Id: "An internal CMU List" <0Jks9449.list-id.cmu.edu>
List-Id: <da39efc25c530ad145d41b86f7420c3b.052000.localhost>
```

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.) A MUA may not recognize the change to the list identifier because the MUA SHOULD treat 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 list

changes. On the other hand, transitioning from an informal unmanaged-list-id-namespace to a domain namespace is an acceptable reason to change the list identifier. Also if the focus of the list changes sufficiently the administrator may wish to retire the previous list and its associated identifier to start a new list reflecting the new focus.

5. Uniqueness of List Identifiers

This proposal seeks to leverage the existing administrative process already in place for domain name allocation. In particular, we exploit the fact that domain name ownership creates a namespace that by definition can be used to create unique identifiers within the domain.

In addition, there must be a mechanism 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 "within.com" are generated in the subdomain of "list-id.within.com".

List-IDs not ending with ".localhost" MUST be globally unique in reference to all other mailing lists.

List owners wishing to use the special "localhost" namespace for their list identifier SHOULD use the month and year (in the form MMYYYY) that they create the list identifier as a "subdomain" of the "localhost" namespace. In addition, some portion of the list identifier MUST be a randomly generated string. List owners generating such identifiers should refer to [MSGID] for further suggestions on generating a unique identifier, and [RFC1750] for suggestions on generating random numbers. In particular, list identifiers that have a random component SHOULD contain 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.localhost> and <da39efc25c530ad145d41b86f7420c3b.051998.localhost> conform to these guidelines, while <lenas-jokes.021999.localhost> and

<mylist.localhost> do not. A particular list owner with several lists MAY choose to use the same random number subdomain when generating list identifiers for each of the lists.

List-IDs ending with ".localhost" are not guaranteed to be globally unique.

6. Operations on List Identifiers

There is only one operation defined for list identifiers, that of case insensitive equality (See Section 3.4.7., CASE INDEPENDENCE [RFC822]). The sole use of a list identifier is to identify a mailing list, and the sole use of the List-Id header is to mark a particular message as belonging to that list. The comparison operation MUST ignore any part of the List-Id header outside of the angle brackets, the MUA MAY choose to 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. If a mailing list processor encounters a List-Id header field from any unexpected source it SHOULD NOT pass it through to the list. This implies that mailing list processors may have to be updated to properly support List-Ids for nested lists.

8. Security Considerations

There are very few new security concerns generated with this proposal. Message headers are an existing standard, designed to easily accommodate new types. There may be concern with headers being forged, but this problem is inherent in Internet e-mail, not specific to the header described in this document. Further, the implications are relatively harmless.

As mentioned above, mail list processors SHOULD NOT allow any user-originated List-Id 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. Acknowledgements

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

Grant Neufeld <grant@acm.org> focused much of the early discussion, and thus was essential in the creation of this document.

References

[LISTHEADER] "List-Header" Mail list. list-header@list.nisto.com <http://www.nisto.com/listspec/mail/> <http://www.nisto.com/listspec/>

[LISTMOM] "ListMom-Talk" Mail list. listmom-talk@skyweyr.com <http://cgi.skyweyr.com/ListMom.Home>

[MSGID] J. Zawinski, M. Curtin, "Recommendations for generating Message IDs", Work in Progress.

[RFC822] Crocker, D., "Standard for the Format of ARPA Internet Text Messages", RFC 822, August 1982.

[RFC1750] Eastlake, D., Crocker S. and J. Schiller, "Randomness Recommendations for Security", RFC 1750, December 1994.

[RFC2234] Crocker, D. and P. Overell. "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, November 1997.

[RFC2369] Neufeld G. and J. Baer, "The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields", RFC 2369, July 1998.

[RFC2606] Eastlake, 3rd, D., and S. Panitz. "Reserved Top Level DNS Names", BCP 32, RFC 2606, June 1999.

[RFC2822] Resnick, P., Editor, "Internet Message Format Standard", STD 11, RFC 2822, March 2001.

Authors' Addresses

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

Full Copyright Statement

Acknowledgement

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