Network Working Group
Request for Comments: 4249
Category: Informational

B. Lilly

January 2006

Implementer-Friendly Specification of Message and MIME-Part Header Fields and Field Components

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Abstract

Implementation of generators and parsers of header fields requires certain information about those fields. Interoperability is most likely when all such information is explicitly provided by the technical specification of the fields. Lacking such explicit information, implementers may guess, and interoperability may suffer. This memo identifies information useful to implementers of header field generators and parsers.

Table of Contents

1. Introduction

Internet messages consist of a message header and a body [N1.STD11], [N2.RFC2822]. MIME content begins with a MIME-part header [N3.RFC2045], [N4.RFC2046]. Message headers and MIME-part headers consist of fields. While the Message Format and MIME specifications define their respective overall formats and some specific fields, they also have provision for extension fields. A number of extension fields have been specified, some more or less completely than others. Incomplete or imprecise specification has led to interoperability problems as implementers make assumptions in the absence of specifications. This memo identifies items of potential interest to implementers, and section 3 of this memo may serve as an informational guide for authors of specifications of extension fields and field components.

2. Scope

This memo is intended as a non-binding informational supplement to various specifications, guidelines, and procedures for specification of header fields [N1.STD11], [N2.RFC2822], [N3.RFC2045], [N4.RFC2046], [N5.BCP9], [N6.BCP90]. It does not absolve authors of header field specifications from compliance with any provisions of those or other specifications, guidelines, and procedures. It offers clarification and supplementary suggestions that will promote interoperability and may spare specification authors many questions regarding incomplete header field specifications.

3. Specification Items

3.1. Established Conventions

A number of conventions exist for naming and specifying header fields. It would be unwise and confusing to specify a field that conflicts with those conventions.

3.1.1. Standard Terminology

Terms related to the Internet Message Format are defined in [N2.RFC2822]. Authors specifying extension header fields should use the same terms in the same manner in order to provide clarity and avoid confusion. For example, a "header" [I1.FYI18], [N2.RFC2822] is comprised of "header fields", each of which has a "field name" and usually has a "field body". Each message may have multiple "headers", viz. a message header and MIME-part [N4.RFC2046] headers.

A message header has a Date header field (i.e., a field with field name "Date"). However, there is no "Date header"; use of such non-standard terms is likely to lead to confusion, possibly resulting in interoperability failures of implementations.

3.1.2. Naming Rules and Conventions

Several rules and conventions have been established for naming of header fields. Rules are codified in published RFCs; conventions reflect common use.

3.1.2.1. Naming Rules

Some RFCs define a particular prefix, reserving use of that prefix for specific purposes.

3.1.2.1.1. Content- prefix rule

This prefix must be used for all MIME extension fields and must not be used for fields that are not MIME extension fields [N3.RFC2045] (section 9).

3.1.2.1.2. Resent- prefix rule

Specified for certain standard fields as given in [N1.STD11] (also used by [N2.RFC2822], although not specified as a prefix therein). If a Resent- version of a field is applicable, an author should say so explicitly and should provide a comprehensive specification of any differences between the plain field and the Resent- version.

3.1.2.2. Naming Conventions

Some prefixes have developed as conventions. Although not formally specified as reserved prefixes, these conventions are or have been in use in multiple fields with common semantics for each prefix.

3.1.2.2.1. Accept- prefix convention

This prefix should be used for all extension fields intended for use in content negotiation [I2.RFC2616] and should not be used for fields that are not intended for such use. An example may be found in [I3.RFC3282].

3.1.2.2.2. List- prefix convention

Used to indicate information about mailing lists when a list expansion takes place. Examples of defined fields can be found in [I4.RFC2369] and [I5.RFC2919].

3.1.2.2.3. Illegal- prefix convention

This prefix provides a record of illegal content in a field when fields are transformed at a gateway [I6.RFC886].

3.1.2.2.4. Disposition-Notification- prefix convention

Specification of information used in conjunction with Message Disposition Notifications (MDNs) [I7.RFC3798].

3.1.2.2.5. Original- prefix convention

Used to reference some content from a related message. Examples include Original-Message-ID as used by [I8.RFC3297] and [I7.RFC3798], Original-Encoded-Information-Types [I9.RFC2156], Original-Envelope-ID [I10.RFC3464], and Original-Recipient [I7.RFC3798].

3.1.2.2.6. Reporting- prefix

Specifies a host that generated a type of report, such as those defined in [I7.RFC3798], [I10.RFC3464].

3.1.2.2.7. X400- prefix convention

Used in conversion from X.400 environments by gateways [I9.RFC2156].

3.1.2.2.8. Discarded-X400- prefix convention

Also used by gateways from X.400 [I9.RFC2156].

3.1.2.2.9. P1- prefix convention

Was used by X.400 gateways [I11.RFC987].

3.1.2.2.10. Delivery-Report-Content- prefix convention

Also used by legacy X.400 gateways [I11.RFC987].

3.2. Common Specification Items

Several items are specified for standard header fields; these items should also be specified for extension fields.

3.2.1. ABNF

[N1.STD11] is vague about where whitespace is permitted or required in header field syntax. [N2.RFC2822] addresses that issue by defining grammar productions such as FWS and CFWS, in conjunction with formal ABNF [N7.RFC4234] and in accordance with the necessity for specificity of such issues as noted in section 3.1 of [N7.RFC4234]. Extension field ABNF should clearly specify where comments, line folding, and whitespace are prohibited and permitted, and should use the [N2.RFC2822] grammar productions in ABNF for that purpose.

All ABNF must be carefully checked for ambiguities and to ensure that all productions resolve to some combination of terminal productions provided by a normative reference [N8.CKLIST] ("All ABNF must be checked"). [N7.RFC4234] provides several productions that may be useful. While use of suitable productions defined and in use is encouraged, specification authors are cautioned that some such productions have been amended by subsequently issued RFCs and/or by formal errata [I12.Errata].

Authors and designers should be careful not to mix syntax with disparate semantics within a single field. Examples of disparate semantics are [N2.RFC2822] comments (which use parentheses as delimiters), [I13.RFC2533] feature sets (which also use parentheses as delimiters, but not for comments), and [I14.RFC3986] Uniform Resource Identifiers (URIs), which permit parentheses in URI text.

It is sometimes necessary or desirable to define keywords as protocol elements in structured fields. Protocol elements should be case insensitive per the Internet Architecture [I15.RFC1958] (section 4.3). Keywords are typically registered by IANA; a specification using registered keywords must include an IANA Considerations section [N9.BCP26], [I16.RFC3692], and should indicate to readers of the specification precisely where IANA has set up the registry (authors

will need to coordinate this with IANA prior to publication as an RFC). In many cases, it will be desirable to make provision for extending the set of keywords; that may be done by specifying that the set may be extended by publication of an RFC, or a formal review and registration procedure may be specified (typically as a BCP RFC).

If keywords are defined, and if there is any chance that the set of keywords might be expanded, a registry should be established via IANA. If a registry is not established initially, there is a good chance that initially-defined keywords will not be registered or will subsequently be registered with different semantics (this has happened!).

Provision may be made for experimental or private-use keywords. These typically begin with a case-insensitive "x-" prefix. Note that [N10.BCP82] has specific considerations for use of experimental keywords.

If some field content is to be considered human-readable text, there must be provision for specifying language in accordance with [N11.BCP18] (section 4.2). Header fields typically use the mechanism specified in [I17.RFC2047] as amended by [I18.RFC2231] and [I12.Errata] for that purpose. Note, however, that that mechanism applies only to three specific cases: unstructured fields, an RFC 822 "word" in an RFC 822 "phrase", and comments in structured fields. Any internationalization considerations should be detailed in an Internationalization Considerations section of the specification as specified in [N11.BCP18] (section 6).

Some field bodies may include ABNF representing numerical values. Such ABNF, its comments, and supporting normative text should clearly indicate whether such a numerical value is decimal, octal, hexadecimal, etc.; whether or not leading and/or trailing zeroes are significant and/or permitted; and how any combinations of numeric fields are intended to be interpreted. For example, two numeric fields separated by a dot, exemplified by "001.100", "1.1", "1.075", and "1.75", might be interpreted in several ways, depending on factors such as those enumerated above.

While ABNF [N7.RFC4234] is used by [N2.RFC2822] and is mentioned above, alternate formal syntax formats may be used in specifications [I19.Syntax].

3.2.2. Minimum and Maximum Instances of Fields per Header

Some fields are mandatory, others are optional. It may make sense to permit multiple instances of a field in a given header; in other cases, at most a single instance is sensible. [N2.RFC2822] specifies

a minimum and maximum count per header for each standard field in a message; specification authors should likewise specify minimum and maximum counts for extension fields.

3.2.3. Categorization

[N2.RFC2822] defines categories of header fields (e.g., trace fields, address fields). Such categories have implications for processing and handling of fields. A specification author should indicate any applicable categories.

3.3. Semantics

In addition to specifying syntax of a field, a specification document should indicate the semantics of each field. Such semantics are composed of several aspects:

3.3.1. Producers, Modifiers, and Consumers

Some fields are intended for end-to-end communication between author or sender and recipient; such fields should not be generated or altered by intermediaries in the transmission chain [I20.Arch]. Other fields comprise trace information that is added during transport. Authors should clearly specify who may generate a field, who may modify it in transit, who should interpret such a field, and who is prohibited from interpreting or modifying the field.

3.3.2. What's it all about?

When introducing a new field or modifying an existing field, an author should present a clear description of what problem or situation is being addressed by the extension or change.

3.3.3. Context

The permitted types of headers in which the field may appear should be specified. Some fields might only be appropriate in a message header, some might appear in MIME-part headers [N4.RFC2046] as well as message headers, still others might appear in specialized MIME media types.

3.4. Overall Considerations

Several factors should be specified regarding how a field interacts with the Internet at large, with the applications for which it is intended, and in interacting with other applications.

3.4.1. Security

Every specification is supposed to include a carefully-considered Security Considerations section [N12.RFC2223] (section 9), [I21.BCP72].

3.4.2. Backward Compatibility

There is a large deployed base of applications that use header fields. Implementations that comprise that deployed base may change very slowly. It is therefore critically important to consider and specify the impact of a new or revised field or field component on that deployed base. A new field, or extensions to the syntax of an existing field or field component, might not be recognizable to deployed implementations. Depending on the care with which the authors of an extension have considered such backward compatibility, such an extension might, for example:

a. Cause a deployed implementation to simply ignore the field in its entirety. That is not a problem provided that it is a new field and that there is no assumption that such deployed implementations will do otherwise.

b. Cause a deployed implementation to behave differently from how it would behave in the absence of the proposed change, in ways that are not intended by the proposal. That is a failure of the proposal to remain backward compatible with the deployed base of implementations.

There are many subtleties and variations that may come into play. Authors should very carefully consider backward compatibility when devising extensions, and should clearly describe all known compatibility issues.

3.4.3. Compatibility With Legacy Content

Content is sometimes archived for various reasons. It is sometimes necessary or desirable to access archived content, with the semantics of that archived content unchanged. It is therefore important that lack of presence of an extension field or field component should not be construed (by an extension specification) as conferring new semantics on a message or piece of MIME content that lacks that field or field component. Any such semantics should be explicitly specified.

3.4.4. Interaction With Established Mechanisms

Header fields are handled specially by gateways under various circumstances, e.g., message fragmentation and reassembly [N4.RFC2046]. If special treatment is required for a header field under such circumstances, it should be clearly specified by the author of the specification. [I7.RFC3798] is an example of how this might be handled (however, because that specification requires deployed RFC 2046-conforming implementations to be modified, it is not strictly backward compatible).

4. Acknowledgements

The author would like to acknowledge the helpful comments provided by members of the ietf-822 mailing list. In particular, Peter Koch and Keith Moore have made useful comments.

5. Security Considerations

No new security considerations are addressed by this memo. The memo reinforces the need for careful consideration and specification of security issues.

6. Internationalization Considerations

This memo does not directly have internationalization considerations; however, it reminds specification authors of the need to consider internationalization of textual field components.

7. IANA Considerations

While no specific action is required of IANA in regard to this memo, it does note that some coordination between IANA and specification authors who do require IANA to set up registries is at least desirable, if not a necessity. IANA should also closely coordinate with the RFC Editor so that registries are set up and properly referenced at the time of publication of an RFC that refers to such a registry. IANA is also encouraged to work closely with authors and the RFC Editor to ensure that descriptions of registries maintained by IANA are accurate and meaningful.

Appendix A. Disclaimers

This document has exactly one (1) author.

In spite of the fact that the author's given name may also be the surname of other individuals, and the fact that the author's surname may also be a given name for some females, the author is, and has always been, male.

The presence of "/SHE", "their", and "authors" (plural) in the boilerplate sections of this document is irrelevant. The author of this document is not responsible for the boilerplate text.

Comments regarding the silliness, lack of accuracy, and lack of precision of the boilerplate text should be directed to the IESG, not to the author.

Normative References

[N1.STD11] Crocker, D., "Standard for the format of ARPA Internet text messages", STD 11, RFC 822, August 1982.

[N2.RFC2822] Resnick, P., "Internet Message Format", RFC 2822, April 2001.

[N3.RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

[N4.RFC2046] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", RFC 2046, November 1996.

[N5.BCP9] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

[N6.BCP90] Klyne, G., Nottingham, M., and J. Mogul, "Registration Procedures for Message Header Fields", BCP 90, RFC 3864, September 2004.

[N7.RFC4234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 4234, October 2005.

[N8.CKLIST] "Checklist for Internet-Drafts (IDs) submitted for RFC publication", http://www.ietf.org/ID-Checklist.html.

[N9.BCP26] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[N10.BCP82] Narten, T., "Assigning Experimental and Testing Numbers Considered Useful", BCP 82, RFC 3692, January 2004.

[N11.BCP18] Alvestrand, H., "IETF Policy on Character Sets and Languages", BCP 18, RFC 2277, January 1998.

[N12.RFC2223] Postel, J. and J. Reynolds, "Instructions to RFC Authors", RFC 2223, October 1997.

Informative References

[I1.FYI18] Malkin, G., "Internet Users' Glossary", FYI 18, RFC 1983, August 1996.

[I2.RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[I3.RFC3282] Alvestrand, H., "Content Language Headers", RFC 3282, May 2002.

[I4.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.

[I5.RFC2919] Chandhok, R. and G. Wenger, "List-Id: A Structured Field and Namespace for the Identification of Mailing Lists", RFC 2919, March 2001.

[I6.RFC886] Rose, M., "Proposed standard for message header munging", RFC 886, December 1983.

[I7.RFC3798] Hansen, T. and G. Vaudreuil, "Message Disposition Notification", RFC 3798, May 2004.

[I8.RFC3297] Klyne, G., Iwazaki, R., and D. Crocker, "Content Negotiation for Messaging Services based on Email", RFC 3297, July 2002.

[I9.RFC2156] Kille, S., "MIXER (Mime Internet X.400 Enhanced Relay): Mapping between X.400 and RFC 822/MIME", RFC 2156, January 1998.

[I10.RFC3464] Moore, K. and G. Vaudreuil, "An Extensible Message Format for Delivery Status Notifications", RFC 3464, January 2003.

[I11.RFC987] Kille, S., "Mapping between X.400 and RFC 822", RFC 987, June 1986.

[I12.Errata] RFC-Editor errata page, http://www.rfc-editor.org/errata.html

[I13.RFC2533] Klyne, G., "A Syntax for Describing Media Feature Sets", RFC 2533, March 1999.

[I14.RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[I15.RFC1958] Carpenter, B., "Architectural Principles of the Internet", RFC 1958, June 1996.

[I16.RFC3692] Narten, T., "Assigning Experimental and Testing Numbers Considered Useful", BCP 82, RFC 3692, January 2004.

[I17.RFC2047] Moore, K., "MIME (Multipurpose Internet Mail Extensions) Part Three: Message Header Extensions for Non-ASCII Text", RFC 2047, November 1996.

[I18.RFC2231] Freed, N. and K. Moore, "MIME Parameter Value and Encoded Word Extensions: Character Sets, Languages, and Continuations", RFC 2231, November 1997.

[I19.Syntax] Carpenter, B., "Syntax for format definitions", http://ietf.org/IESG/STATEMENTS/syntax-format-def.txt

[I20.Arch] Crocker, D., "Internet Mail Architecture", Work in Progress, March 2005.

[I21.BCP72] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.

Author's Address

Bruce Lilly

EMail: blilly@erols.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).