Indicating Email Handling States in Trace Fields
draft-ietf-appsawg-received-state-01

Abstract

   This memo registers a trace field clause for use in indicating
   transitions between handling queues or processing states, including
   enacting inter- and intra-host message transitions. This might
   include message quarantining, mailing list moderation, timed
   delivery, queueing for further analysis, content conversion, or other
   similar causes, as well as optionally identifying normal handling
   queues.

Status of This Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at http://datatracker.ietf.org/drafts/current/.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   This Internet-Draft will expire on November 19, 2012.

Table of Contents

1. Introduction

   [SMTP] defines the content of email message trace fields, commonly
   the "Received" field. These are typically used to record an audit
   trail of the path a message follows from origin to destination, with
   one such field added each time a message moves from one host to the
   next.

   Section 3.7.2 of that memo mentions that "the most important use of
   of Received: lines is for debugging mail faults [...]".

   There are some cases where there may be large time gaps between trace
   fields. Though this might be caused by transient communication
   issues, they might also be caused by policy decisions or special
   processing regarding the content of the message, authorization of
   some identity on the message, or transitions between major software
   components. Common examples include message quarantines (filters
   that delay relaying or delivery of a message pending manual operator
   action), pending content analysis, or mailing list servers that
   impose moderation rules (mailing list owner action required regarding
   mail from authors not subscribed to those lists).

   This memo registers a new optional clause that can be used in trace
   fields to indicate that a message entered such a special processing
   queue or state for some period. This allows inspection of the trace
   information to reveal that the cause for a time gap in trace fields
   was an imposed delay rather than one caused by transient technical
   difficulties.

2. Key Words

   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 [KEYWORDS].

3. New Trace Clause

   This memo creates a new trace field clause, called "state", which can
   be used to indicate the nature of a delay imposed on relaying of a
   message toward its recipient(s). It is followed by a single keyword
   that provides that detail. A Mail Transfer Agent (MTA) or other
   handling agent that determines a message has entered a state other
   than normal queueing of messages for relaying or delivery MAY
   generate a trace field including one of these clauses. That is, the
   presence of this clause on a trace field is an indication of the
   entry of the message into that state; a later trace field added would
   indicate its departure from that state.

Appropriate use of this mechanism does not include associating meta-
data with the message, such as categorizing the message (e.g., the
notions of "is spam" or "was 8-bit, converted to 7-bit").

The following state keywords are defined in this document; extensions
may define other registered keywords (see Section 6.2):

auth: The message entered a queue pending authentication of some
   identifier in the message.

content: The message entered a queue pending content analysis, such
   as scanning for spam or viruses.

convert: The message entered a queue pending content conversion.

moderation: The message entered a hold pending mailing list
   moderator action.

normal: The message is not in an administrative hold and is queued
   for or is being handed off to the next handling agent (which may
   be local delivery). This is the default interpretation when no
   "state" clause is present.

other: The message entered a hold or queue for reasons not covered
   by other keywords in this list, and not for transient technology
   issues.

outbound: The message entered a queue for outbound relaying. This
   is typically the last case added for a single host, and the next
   Received field is expected to be added by some other host.

quarantine: The message entered a hold in an isolation queue pending
   operator action for local policy reasons.

timed: The message entered a hold in order to meet a requested
   delivery window, such as is defined in [FUTURERELEASE].

The ABNF for this clause:

```
    State = CFWS "state" FWS queue-state-keyword *( "/" 1*value )

    queue-state-keyword = ( reg-state-keyword / unreg-state-keyword )

    reg-state-keyword = ( "auth" / "content" / "convert" /
                          "moderation" / "normal" / "other" /
                          "outbound" / "quarantine" / "timed" /
                          additional-state-keyword )

    additional-state-keyword = unstructured
                            ; see "IANA Considerations" below

    unreg-state-keyword = unstructured
                        ; from [MAIL]
```

"FWS" and "CFWS" are defined in [MAIL]; "value" is defined in [MIME].

A transfer agent making use of this extension MAY also include header field comments to provide additional information.

Use of this clause by transfer agents is OPTIONAL.

4. Discussion

Handling agents are not expected to implement or support all of these. Indeed, recording trace information for all of the states described above could make the header of a message inordinately large. Rather, an agent is encouraged to apply state annotations when a message enters a handling queue where substantial delay is possible, and especially when a handoff has occurred between two different, independent agents.

For example, an MTA receiving a message, doing message authentication, scanning for viruses and spam, and then putting it in an outbound queue could add four Received fields denoting each of these states. However, where they are all done as part of a single system process, in a single pass, doing so would be considered unusual (and extremely verbose). This method SHOULD NOT be applied except when doing detailed analysis of a single component to identify performance issues with those steps.

Rather, an agent that wishes to make a state annotation SHOULD add only a single Received field including such annotation, thus indicating (a) the time of completion of its handling of the message via the date portion of the field, and (b) the final disposition of that message relative to that agent. For example, an MTA receiving a message that performs various checks on the message before immediately handing it off to a Mailing List Manager (MLM) would only

record a "normal" state, assuming it passes those checks. The MLM
would then evaluate the message and record its own state once it
decides what the next step will be for the handling of that message.

5. Granularity

The degree of granularity -- and therefore the degree of verbosity --
recorded through the use of this additional trace clause is likely to
vary depending on circumstances. It will typically be the case that
use of this clause will be limited to "unusual" transitions, such as
when a message requires additional scrutiny or other processing, or
needs to be quarantined.

Somewhat greater granularity might also include transitions of
administrative responsibility, such as between an Mail Transfer Agent
(MTA) operator and a Mailing List Manager (MLM) operator. This could
be further enhanced to note some transitions that are interesting
only when other transitions have occurred, such as noting entry to
the outbound queue only when the message is originating from an
"interesting" source, like an MLM, since an MLM can introduce
significant delay and it could be useful to know when it completed
its processing, as distinct from the subsequent processing by the
originating MTA. In circumstances needing very fine-grained trace
information, fields might be created to note all of these
"significant" network architecture transitions.

One should note, however, when choosing higher levels of granularity,
that the Received fields present on a message could be counted by
MTAs when trying to decide whether or not a message routing loop is
in effect. A message with an abundance of these might cause an
incorrect determination that the message is in a delivery loop,
causing it to be removed from the mail stream. See Section 6.3 of
[SMTP] for further discussion.

6. IANA Considerations

6.1. Mail Parameters Additional-registered-clauses Sub-Registry

This memo adds to the "Additional-registered-clauses" sub-registry of
the "Mail Parameters" registry, created by [SMTP], the following
entry:

Clause name: state

Description: Indicates entry into a special queue state

   Syntax Summary: state <state-name>

   Reference: [this memo]

6.2. Mail Parameters Registered-states Sub-Registry

   The "Mail Parameters" registry at IANA is updated by the creation of
   the "Registered-states" sub-registry to contain valid state keywords
   for use with this specification. Updates to this registry are
   governed by the Specification Required rules of [IANA].
   Registrations must include the following entries:

   Name: The name of the state keyword being defined or updated

   Description: A brief description of the keyword's meaning

   Specification: The specification document that defines the queue
      state being registered

   Use: One of "current" (the state keyword is in current use),
      "deprecated" (the state keyword is in use but not recommended for
      new implementations), or "historic" (the state keyword is no
      longer in substantial current use).

   The initial registration set is as follows:

Name	Description	Specification	Use
auth	Held for message authentication	[this memo]	current
content	Held for message content analysis	[this memo]	current
convert	Held for message content conversion	[this memo]	current
moderation	Held for list moderation	[this memo]	current
normal	Message is not being held other than to accommodate typical relaying delays	[this memo]	current
other	Held for causes not covered by other registered state keywords	[this memo]	current
outbound	Message placed in outbound queue	[this memo]	current
quarantine	Held for operator action due to content analysis or local policy	[this memo]	current
timed	Held to accommodate a specific requested delivery window	[this memo]	current

7. Security Considerations

   The use of this trace information can reveal hints as to local policy
   that was in effect at the time of message handling.

8. References

8.1. Normative References

   [IANA]   Narten, T. and H. Alvestrand, "Guidelines for
                   Writing an IANA Considerations Section in RFCs",
                   BCP 26, RFC 5226, May 2008.

   [KEYWORDS]   Bradner, S., "Key words for use in RFCs to Indicate

Requirement Levels", BCP 14, RFC 2119, March 1997.

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

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

[SMTP]   Klensin, J., "Simple Mail Transfer Protocol",
                RFC 5321, October 2008.

8.2. Informative References

[FUTURERELEASE] White, G. and G. Vaudreuil, "SMTP Submission Service
                Extension for Future Message Release", RFC 4865,
                May 2007.

Appendix A. Trace Field Examples

This section includes a sample of the new trace field clause in use.

A.1. Typical Delivery Without Obvious Delays

Typical message delivery

```
    Received: from newyork.example.com
            (newyork.example.com [192.0.2.250])
        by mail-router.example.net (8.11.6/8.11.6)
            with ESMTP id i7PK0sH7021929
            for <recipient@example.net>;
        Fri, Feb 15 2002 17:19:22 -0800
    Received: from internal.example.com
            (internal.example.com [192.168.0.1])
        by newyork.example.com (8.11.6/8.11.6)
            with ESMTP id i9MKZCRd064134
            for <recipient@example.net>;
        Fri, Feb 15 2002 17:19:08 -0800
```

Example 1: Typical message delivery with no appreciable handling
delays; only Received fields shown

A.2. **Delivery With Moderation**

   Message delivery after moderation

```
        Received: from newyork.example.com
                  (newyork.example.com [192.0.2.250])
             by mail-router.example.net (8.11.6/8.11.6)
                  with ESMTP id i7PK0sH7021929
                  for <recipient@example.net>;
             Fri, Feb 15 2002 18:33:29 -0800
        Received: from internal.example.com
                  (internal.example.com [192.168.0.1])
             by newyork.example.com (8.11.6/8.11.6)
                  with ESMTP id i9MKZCRd064134
                  for <secret-list@example.com>
                  state moderation (sender not subscribed);
             Fri, Feb 15 2002 17:19:08 -0800
```

   Example 2: Message held for moderation; only Received fields shown

   The message passed from internal.example.com to newyork.example.com
   intended for a mailing list hosted at the latter. For list
   administrative reasons, the message is held there for moderation. It
   is finally released over an hour later and passed to the next host.
   A comment after the state expression indicates the actual cause for
   the administrative hold.

Appendix B. **Acknowledgements**

   The authors wish to acknowledge the following for their reviews and
   constructive criticisms of this proposal: Tony Finch, Ned Freed, Carl
   S. Gutenkunst, John Levine, Bill McQuillan, Alexey Melnikov, Robert
   A. Rosenberg, Hector Santos, Rolf Sonneveld, and Mykyta Yevstifeyev.

Authors' Addresses

   D. Crocker
   Brandenburg InternetWorking
   675 Spruce Dr.
   Sunnyvale 94086
   USA

   Phone: +1.408.246.8253
   EMail: dcrocker@bbiw.net
   URI: http://bbiw.net

   Murray S. Kucherawy
   Cloudmark, Inc.
   128 King St., 2nd Floor
   San Francisco, CA 94107
   US

   Phone: +1 415 946 3800
   EMail: msk@cloudmark.com