Workgroup: Network Working Group
Internet-Draft:
draft-klensin-email-for-clause-00
Published: 24 July 2022
Intended Status: Informational
Expires: 25 January 2023
Authors: J.C. Klensin

Issues with the SMTP/IMF 'for' Clause and Remedies

Abstract

The "for" clause of the "Received:" header field goes back to the first widely deployed version of SMTP (RFC 821). However SMTP also allows multiple-recipient messages to be transmitted in a single mail transaction. The combination may, in some cases, lead to undesirable disclosure of information, including disclosing mail addresses that were intended to be kept hidden from other recipients. In the process of working on revisions to RFC 5321 and developing a new Applicability Statement in the EMAILCORE WG, there have been attempts to fix the problems by find-tuning text about actions and warnings. This document is an attempt to explore the issues in somewhat more depth for members of the community who are, or should be, participating in the WG.

Status and Audience

This document is intended for discussion during the scheduled 2022-07-26 meeting of the EMAILCORE WG. It is being posted in Internet-Draft form rather than simply to the WG mailing list because it addresses issues outside the WG's scope that might be of interest to the community in the future. It is not expected to evolve into a Standards Track document in it current form and may be abandoned after IETF 114. In conjunction with those characteristics, it is written very informally.

The current version assumes familiarity with the current versions of the specs being developed in that WG [rfc5321bis][rfc5322bis] [email-as].

Status of This Memo

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 25 January 2023.

Table of Contents

1. Introduction

The "for" clause of the "Received:" header field goes back to the first normative version of SMTP (usually known just as "RFC 821" [RFC0821]). However SMTP also allows multiple-recipient messages to

be transmitted in a single mail transaction. The combination may, in some cases, lead to undesirable disclosure of information, including disclosing mail addresses that were intended to be kept hidden from other recipients. In the process of working on revisions to RFC 5321 [rfc5321bis] and developing a new Applicability Statement [email-as] in the EMAILCORE WG [EMAILCORE-wg], there have been attempts to fix the problems by find-tuning text about actions and warnings. This document is an attempt to explore the issues in somewhat more depth for members of the community who are, or should be, participating in the WG.

Because some of the changes it suggests as possibilities might change the syntax of the "for" clause, syntax currently defined in the Internet Message Format (IMF, Mail Headers) document [rfc5322bis], that document or its successors could, in principle, be affected as well.

While it may suggest paths that might lead to normative specifications, this document is not intended to contain any normative language and any text that appears to be normative is a result of haste in writing and should not be interpreted that way.

2. Issues and the Problem

2.1. History

Section 4.1.2 of the Internet Standard SMTP protocol [RFC0821], defines the "for" clause as an optional part of the <time-stamp-line> (aka the "Received:" header field, later known as a "trace field"). It allows only one argument, a <path>, which, with later modifications, became a mailbox name. It also allowed only one "for" clause in a "Received:" header field. However, it also allows multiple RCPT commands in a mail transaction. It does not specify when "for" should or should not be supplied nor which mailbox name should be used (i.e., which RCPT command is more important than the others). We can look at that today and say "whatever were they (or was he) thinking?", but, a month short of forty years later, that question would be a bit late. For most of those forty years, or at least the first half of them, the ambiguity was not seen as causing harm.

Skipping over some intermediate documents, the current spec in general use, RFC 5321 [RFC5321], is somewhat more specific. Its section 4.4 includes:

> If the FOR clause appears, it MUST contain exactly one <path> entry, even when multiple RCPT commands have been given. Multiple <path>s raise some security issues and have been deprecated, see Section 7.2.

And then Section 7.2 discusses "blind copies" and the possible advantages of sending such copies as separate mail transactions with only a single RCPT command each. It also reinforces the principle that there should be "no more than one mailbox" in the "for" clause of a "Received" mail header field and urges that other information not supply more than one address either. Its description avoids telling implementers what they should do, only what they should avoid.

2.2. Status as of draft-ietf-emailcore-rfc5321bis-11 ("rfc5321bis" below) and draft-ietf-emailcore-as-05 (the "Applicability Statement" or "A/S" below)

In an attempt to better deal with the problem, the last sentence in the current version of rfc5321bis [rfc5321bis] contains the sentence:

> Also, the optional FOR clause should be supplied with caution or not at all when multiple recipients are involved lest it inadvertently disclose the identities of 'blind copy' recipients to others.

That still does not say very much and its meaning in various edge cases may be subject to interpretation. That is uncomfortable, but it may be hard to do much better without getting severely tangled up in those edge cases (see the next Section). The WG has also reached consensus that whatever detailed explanation of the "for" clause is needed (presumably including discussion of the edge cases) will be included in the Applicability Statement [email-as] rather than in rfc5321bis, but there is no clear agreement yet about exactly what that other document should say.

3. Edge Cases, Elephants in the Room, and Other Technical Complications

While numbered, the issues below are not in any particular order and may not be completely separable.

(1) In the Internet most of us deal with most of the time, connectivity is very good and most of the email is handled by a few large providers. Many of those providers have their own infrastructure and are not dependent on the SMTP relay mechanisms defined in RFC 821 and subsequent specifications. We have strongly discouraged so-called "open relays" (relay MTAs that have no administrative relationship with either the originating or delivery systems), but that does not mean that mail necessarily moves from a collection of originating systems (starting with an MUA and sometimes a submission server) controlled by one entity to a collection of receiving ones (including the one that makes "final delivery" to either

a mailbox, a gateway into some other environment, or the equivalent). For example, a destination system might contract with an independent third party provider to relay mail as a so-called "backup MX". Unless specified by the contract or in other ways, that does not make that third party part of the administrative domain of the recipient, at least as we usually think of those terms. Changing SMTP to ignore or exclude those other cases would, at least in my opinion, make the specification more confusing and less generally applicable and the Internet worse.

(2) With regard to that "for" clause, it may be worth remembering that SMTP has no mechanism for specifying or determining which RCPT command is more important than others (for example, RFC 821 could easily have made the first one, or the first one that was accepted, primary, but did not). Whatever heuristics we might invent notwithstanding, in many cases, the only computer entities that really know what mailbox should be exposed in the "for" clause are the MUA and maybe the Submission system. We given them permission to do things "ordinary" MTAs are prohibited from doing and even require that they do such things in order to ensure conceptual message integrity and conformance with standards. Both are, fwiw, out of scope for EMAILCORE and we have never required a simple first-hop SMTP server to conform to Submission server rules or allowed it the same flexibility.

(3) We do allow an MTA to take a message that comes to it with multiple RCPT commands, break it apart, and send out separate messages in separate mail transactions. That is essentially required when the RCPT commands it receives contain different domains that resolve to different MX record collections (or at least the preferred server for the next hop), but it is clearly allowed for other cases, and, in rfc5321bis section 7.2, we encourage it for cases involving "bcc"s where the server knows what is going on. But there may be cases where MTAs might want to recombine things as well as splitting them apart.

As an extreme example, suppose there were a relay at the boundary between the high-speed and well-connected part of the Internet and a part of the network that involves intermittent connections, long delays, and expensive bandwidth (if an example is needed, think Mars or something further out). That situation implies that it should have, and carefully curate, mail queues, at least somewhat organized around the availability of particular destinations (or appropriate intermediaries in their direction). So it receives two messages (different mail transactions in the same or different

SMTP sessions), each of which has one RCPT command, and the arguments to those commands have the same target domains. Given what we seem to be telling the servers in the path up to that point, it is entirely reasonable for the trace fields in those messages to contain "for" clauses that identify the (only) recipient address. However, because bandwidth and transaction time are important for that relay, it checks the Message-IDs, discovers that they are identical, and then maybe goes on to verify that the message bodies are bit-for-bit identical. There might be other cases, but the most obvious one is a message that left a Submission server with multiple recipients (RCPT commands in the same mail transaction) but was broken into separate mail transactions along the line. Ignoring the widely-abused rule against MTAs looking at headers to figure out what to do (see rfc5321bis Section 3.6), I don't think there is anything in rfc5321bis (or 5321 itself) that prohibits it from re-combining those messages and sending them out in a single mail transaction with multiple RCPT commands. I can imagine some interesting choices if the messages took different paths reaching it where recombining might do a real job on signatures over the message headers, but that is getting far afield. Now, if we discourage inserting a "for" clause when one receives mail transactions with multiple RCPT commands, no more "for" clauses are going to get added, at least prior to the delivery server. But the privacy damage is already done because earlier servers in the path have already inserted "for" clauses that, with a little bad luck, could disclose bcc recipients.

(4) As far as I can tell, the only MTAs who may actually have the information needed to "understand" what should should go into the "for" clause (or if it should be provided) are the submission server and the delivery MTA. The submission server at least knows that it is a submission server. In many cases, the delivery MTA might not actually know whether it is in that role or just thinks it is.

Snark: If, by now, readers are not either suffering from bad headaches or muttering things about cans of worms, this document may be failing in its intent.

4. Some Options

Note: Before considering the subsections below (and variations on them) in context with rfc5321bis, remember or review the very limited type of changes that can be made, relative to prior specs, in bringing a document to Internet Standard [RFC2026] [RFC6410]. Some of what follows might be consistent with those requirements; some others clearly are not. Also note that at least one of them is

clearly out of scope for EMAILCORE even if it required a syntax change in rfc5321bis/rfc5322bis (and it is not clear whether the syntax change would be allowed while going to Internet Standard).

4.1. Time to Give Up

We could conclude that, because of under-specification going back at least to RFC 821 and the complexities of the modern world (including but certainly not limited to privacy considerations) the "for" clause has outlived its usefulness. Deprecating it would require an explanation in rfc5321bis but, otherwise, would just require removing some syntax and assorted clumsy explanations.

One of several problems with this approach is that I gather there are people out there who think it is useful (that includes myself on some days). Probably there are enough of them that some implementations would ignore the prohibition and then we would have no guidance at all (I don't see how the Applicability Statement could include an extensive discussion of a feature that rfc5321bis had discarded, presumably with its own Deprecated, Obsolete, or "NOT RECOMMENDED" text).

4.2. Deprecate It in the Applicability Statement Instead

Leave the text in rfc5321bis alone, plus or minus minor tuning (see Section 4.4 below) and push this entirely onto the Applicability Statement, including (presumably in a battle to be fought later) the possibility of a careful explanation of why the "for" clause has become more trouble than it is worth and is NOT RECOMMENDED regardless of what rfc5321bis appears to say.

Same comments as above about people who may think the feature is useful. Also, slipping on my hat as nominal co-author of the A/S for a moment, I really hope we don't dump such baggage there.

4.3. Prohibit "for" in SMTP Without Discarding It In Message Submission

Modify the Submission spec [RFC6409] to include more information about when a Submission server should, or should not, provide a "for" clause (even if nominally in the "Received:" header field associated with the message it received) and what should be in it. Ideally, allow some additional/special syntax to distinguish between "just did not include one of those" (since providing one probably cannot be required) and "'for' clause deliberately omitted".

Then leave the syntax (or include the new syntax) in rfc5321bis but indicate that SMTP MTAs SHOULD NOT (or maybe even MUST NOT) insert the "for" clause. That eliminates the insertion of a "for" clause by the delivery server, but maybe "Apparently-to:" (despite being

explicitly deprecated in RFC 5321 (and, so far, rfc5321bis)) and/or "Delivered-to:" [RFC9228] or some successor are better solutions to whatever the actual requirement is at that point.

Possibly the existing text in rfc5321bis that would be left after removing a good deal of handwaving ought to be modified as well. Or not.

4.4. Fine-tuning Example

In an offlist conversation, Alexey proposed changing the current text in rfc5321bis (see Section 2.2 above) to approximately:

> Also, the optional FOR clause should not be supplied when the same message body is sent, in the same mail transaction, to multiple recipients in order not to inadvertently disclose the identities of "blind copy" recipients to others.

While this does not cover all of the cases above, it may still be an improvement by virtue of being (at least apparently) less ambiguous. On the other hand, it does not cover all cases and hence does not completely protect against "for" clauses that inadvertently disclose information.

So:

1. Independent of any of the other options (other than those that would remove the subject text entirely), does the WG believe that, on balance, this (or something like it) would be an improvement?

2. Is it good enough or does it, or that section, or the sections that refer to it, need additional tuning?

5. Acknowledgments

Thanks to Alexey Melnikov, EMAILCORE co-chair, for finally getting me to the point where it became obvious (at least to me) that this document was needed.

6. IANA Considerations

RFC Editor: Please remove this section before publication.

This memo includes no requests to or actions for IANA.

7. Security Considerations

Once we decide what to do with the notes above, it will be possible to describe their security implications. On the other hand, there is

a sense in which the entire conversation about the "for" clause is about privacy and prevention of unwanted disclosures of information.

8. References

8.1. Consolidated References (temporary)

[email-as] Klensin, J.C., Murchison, K., and E. Sam, "Applicability Statement for IETF Core Email Protocols", 23 May 2022, <draft-ietf-emailcore-as-05>.

[EMAILCORE-wg] IETF, "Revision of core Email specifications (emailcore)", 2022, <https://datatracker.ietf.org/wg/emailcore/about/>.

[RFC0821] Postel, J., "Simple Mail Transfer Protocol", RFC 821, DOI 10.17487/RFC0821, STD 10, August 1982, <https://www.rfc-editor.org/info/rfc821>.

[RFC2026] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, DOI 10.17487/RFC2026, October 1996, <https://www.rfc-editor.org/info/rfc2026>.

[RFC5321] Klensin, J., "Simple Mail Transfer Protocol", RFC 5321, DOI 10.17487/RFC5321, October 2008, <https://www.rfc-editor.org/info/rfc5321>.

[rfc5321bis] Klensin, J.C., "Simple Mail Transfer Protocol", 9 July 2022, <draft-ietf-emailcore-rfc5321bis-12>.

[rfc5322bis] Resnick, P., "Internet Message Format", 4 April 2022, <draft-ietf-emailcore-rfc5322bis-03>.

[RFC6409] Gellens, R. and J. Klensin, "Message Submission for Mail", RFC 6409, DOI 10.17487/RFC6409, STD 72, November 2011, <https://www.rfc-editor.org/info/rfc6409>.

[RFC6410] Housley, R., Crocker, D., and E. Burger, "Reducing the Standards Track to Two Maturity Levels", BCP 9, RFC 6410, DOI 10.17487/RFC6410, October 2011, <https://www.rfc-editor.org/info/rfc6410>.

[RFC9228] Crocker, D., Ed., "Delivered-To Email Header Field", DOI 10.17487/RFC9228, RFC 9228, April 2022, <https://www.rfc-editor.org/info/rfc9228>.

Author's Address

John C Klensin
1770 Massachusetts Ave, Ste 322

Cambridge, MA 02140
United States of America

Email: john-ietf@jck.com