Internet Draft   D. MacQuigg
Document: draft-macquigg-authent-declare-00g
Expires: November 2005   25 May 2005

Email Sender's Declaration of Identity

Status of this Memo

Abstract

A key item that must be standardized to allow interoperation of different email authentication methods is the ID declaration. Current authentication methods assume that one or another of the existing fields in a mail transfer can be used as the Identity to be verified. Since there is no way to tell which field, if any, the sender is prepared to authenticate, extra DNS queries must be made, in the worst-case, testing all possibilities just to find no authentication is offered at all. This draft proposes a neutral syntax that can be used by all methods.

Conventions used in this document

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].

Table of Contents

Changes

5/16 SMTP Service Extension - Added statement in section 4 and example in IANA Considerations section.

5/18 Clarifying the relationship between the declared ID and the IDs assumed by various methods – section 4.

5/18 Added example to clarify how a standard ID declaration avoids DNS "hunting" – New section 5.

5/22 Added SRS example at end of section 4 and explanations as to why neither SUBMITTER nor SRS are suitable as a universal ID.

5/22 Added paragraphs to section 1, clarifying the need for a standard.

5/22 Added paragraph to section 2 stating the semantics of the new ID name.

5/23 Moved fundamental requirements from section 4 to section 2.

5/24 Added ABNF syntax in section 6.

5/24 Updated IANA Considerations.

5/25 Clarify limitations of alternative IDs (SUBMITTER and SRS).

The current working copy of this draft and a summary of mailing-list discussions can be found at http://purl.net/macquigg/email

1. The Need for an Email Identity Declaration

A fundamental requirement for all email authentication methods is that the sender must declare, or at least reveal, its Identity to the receiver. Unfortunately, there is no agreement on how this should be done. Some believe firmly that it should be done in the EHLO command at the start of each session, others insist that it should be done in the MAIL FROM command with each email. Still others think the true Identity should be extracted from one or another of the email headers that the recipient actually sees. Adding to the confusion is the fact that each of these identities may legitimately differ from the Identity that is to be authenticated, and may differ in having extra "subdomain" labels that are not easily separated from the Identity to be checked.

The fundamental problem with the use of existing identities is that none of them were intended for the purpose of email authentication. Changing current standards and practice is difficult. Adding new syntax, as done by [SUBMITTER] and [SRS] avoids this problem, but each of these proposals addresses only the needs of one method. We need a standard that will work with all methods.

Each of the methods is now going its own way, with no thought as to how one will communicate with another as an email is forwarded across the Internet. A receiver must try all possible methods, by "hunting" for DNS records at various locations. The most costly DNS hunts will be for the typical randomly-generated spammer name that offers no authentication.

We need a clear and simple standard that will allow any receiver to know exactly what Identity is authorizing a transfer, regardless of which authentication method is used, and to treat with suspicion any sender that does not offer a reputable ID.

We need a clear and simple standard that will allow any sender to declare, regardless of what other identities may be found in an email, "I am <ID>, and I take responsibility for this transfer." We need to change the culture of irresponsibility in the current operations of Public Mail Servers. An email ID should become the equivalent of a "broadcast license", and owners of those IDs should not tolerate their abuse. Standardizing an ID declaration will help achieve this change.

2. A Possible Compromise

The proposed ID Declaration syntax is designed to fit in with a future Inter-Operability Protocol for all methods. See [draft-macquigg-authent-IOP] for one such protocol. The relevant fundamental requirements from that protocol are:

1) This protocol must not favor any one authentication method over another. It must allow an arbitrary number of Forwarders using different methods to work together in the same authenticated transfer.

2) Each Sending Mail Transfer Agent (MTA) in an IP-authenticated transfer must declare, in the SMTP session, the Identity responsible for the transfer.

One way to standardize the Identity declaration is to use a new field, independent of existing fields, and not constrained by any pre-existing semantics.

```
   EHLO  mailserver7.bigforwarder.com
   ID  bigforwarder.com
   MAIL FROM:<bob@sales.some-company.com>
```

The ID command provides a domain name independent of other names in the envelope and headers. It should be a short, memorable name to enhance its value as a Public Mail Server identity. There are three semantics associated with this new name.

1) It may be used for accreditation and reputation.
2) It may be used to specify the location for authentication records.
3) It may be used, after authentication, as a bounce address for complaints and challenges relating to spam.

One advantage of this syntax is that the sender's ID is explicitly declared, not just assumed from existing information. Not only will this remove the current uncertainty as to which ID the sender intends to use, but false information here is evidence of a serious problem, not just a forgivable error in passing on existing information (a long-standing problem with email). This will greatly reduce the administrative burden in deciding whether to trust a sender. It will also allow an immediate reject when a declared ID has no authentication record.

Another advantage is that there is no "hunting" for DNS records at various locations and multiple levels of a deep subdomain tree. The ID should provide the exact location where at least the first authentication record will be found. The first record should specify what methods are used, and thereby avoid the hunt. See the example in section 5.

Most reputable Public Mail Servers will chose their top domain name as their ID, but it can be any name under DNS control. This could be a domain set up specifically to authorize mail servers, or it could be some other organization's ID. The latter should be allowed but discouraged, since any miscommunication over the use of someone else's ID could result in authentication failures, suspicion of forgery, and loss of reputation by the owner of the ID.

Although the ID command may be repeated, to provide a different ID with every message, senders should organize their messages so that only one ID command is issued for many subsequent MAIL commands. This will minimize the number of DNS queries made by the receiving MTA.

Senders with a large organization, and a desire to decentralize their mail system management, should still consider putting their authorization records under their topmost domain name. Consolidating the records for ten busy subdomains should reduce DNS queries by a factor of ten.

3. Levels of Compliance

During the early days of email authentication, it may be useful to rate Public Mail Servers as to their level of compliance with authentication standards. This will encourage all servers to provide at least minimum security, and allow mail receivers to put special trust in servers that provide the highest levels. One possible scheme having three levels is described in [draft-macquigg-authent-IOP]. The proposed ID syntax will satisfy level one, and this is all that is needed for domains that do not forward emails from other domains.

Level 1) Servers that will declare their ID, and provide a DNS record for that ID to authorize that server.

4. Relations with Existing and Proposed Standards or Practice

The proposed syntax will require an SMTP service extension for a new ID command. See section 7, IANA Considerations and [RFC-2821].

MTA software will need to be enhanced and deployed at sites that provide email authentication. To minimize upgrade efforts these changes should be bundled with the upgrade to enable authentication.

Each authentication method should consider what it will do if the declared ID differs from the default ID that is used by their method. The options are:
  a) Ignore the default. The ID declaration over-rides.
  b) Ignore the declared ID except to find the initial DNS record and determine what methods are available. Then use the default ID, start with a fresh query for DNS records at that ID.
  c) Do a cross-check, then proceed with the declared ID. e.g. The default ID must be a subdomain of the declared ID.

The proper procedure depends on what the requirements of the particular method are. If they are simply to verify that the ID authorizes the transfer, option 'a' will be the quickest. If additional requirements are important, options 'b' or 'c' may be necessary. Additional requirements may include such things as matching between header fields and the authorizing Identity, or existence of a particular DNS record structure for the sending MTA.

The problem of introducing a new identity into the SMTP session has been addressed before. See [SUBMITTER] for one alternative.

```
   MAIL FROM:<bob@sales.some-company.com> SUBMITTER=bigforwarder.com
```

The proposed SUBMITTER parameter for the MAIL FROM command is intended to provide header information (the "PRA" address) in the SMTP commands. The limitation to PRA makes it inapplicable as a universal ID declaration.

See [SRS] for another alternative. The MAIL FROM command is rewritten so that it contains both the original return path before any forwarding and a new return path for the current hop.

```
   MAIL FROM:<bob#sales.some-company.com@bounce.bigforwarder.com>
```

The limitation to defining a new return path makes SRS inapplicable as a universal ID declaration.

5. Example Using the ID

Here is a typical SMTP session using the ID command. C is the client (sender). S is the server (receiver).

```
   C: EHLO mailserver7.bigforwarder.com
   S: 250-host.com, welcome
   S: 250-SIZE ETRN
   S: 250-AUTH LOGIN ID
   S: 250 HELP
   C: ID bigforwarder.com
   S: 250 ... Sender validation pending. Continue.
   C: MAIL FROM:<bob@sales.some-company.com>
   S: 250 Ok
```

Without the ID command, you will waste a bunch of DNS queries and possibly conclude this sender offers no authentication. For each possible Identity (mailserver7.bigforwarder.com, bigforwarder.com, sales.some-company.com, some-company.com) you need to search every possible location for DNS records (<Identity>, _client._smtp.<Identity>, ...), and we still haven't searched all the header identities. This is what we mean by DNS "hunting" – searching for records that may not exist.

With the ID command, the receiving MTA does a DNS query for a TXT record at a standard location, like _AUTH.bigforwarder.com.mail.net The query returns a record that specifies exactly what methods are supported by the owner of the Identity. If the method parameters all fit in the first record, no further queries are necessary. If the parameters don't all fit, you will at least know exactly where to look for the rest.

6. Formal Syntax

The following syntax specification uses the Augmented Backus-Naur Form (ABNF) as described in [RFC-2234].
The ID command can occur any time in an SMTP session except during data transfer. The specified Identity remains in effect until the end of the session, or another ID command. Clients MUST NOT send an ID command unless that keyword is offered in the server's EHLO response.

```
   ID-command      = "ID" 1*SP Domain 1*SP options CRLF
   Domain          = (sub-domain 1*("." sub-domain))
   sub-domain      = Let-dig [Ldh-str]
   Let-dig         = ALPHA / DIGIT
   Ldh-str         = *( ALPHA / DIGIT / "-" ) Let-dig
   options         = 1*(%d0-9 / %d11-12 / %d14-127)
                   ; string of any characters other than CR or LF

   ALPHA           =  %x41-5A / %x61-7A    ; A-Z / a-z
   DIGIT           =  %x30-39              ; 0-9
   SP              =  %x20   ; space
   CRLF            =  CR LF
   CR              =  %x0D   ; carriage return
   LF              =  %x0A   ; linefeed
```

The domain name used as an Identity, has the same syntax as the domain name in the EHLO command. Options are not defined, but are included here to allow future extensions to the ID command.

Security Considerations

ID strings are easily faked, the same as any other envelope or header parameters. Security depends entirely on the authentication method. Until the ID is authenticated, it should not be trusted.

IANA Considerations

The proposed syntax will require an SMTP service extension with the following addition to the Mail Parameters Registry.

Keywords	Description	Reference
ID	Sender's Declared Identity	[RFC....]

There are no additional parameters needing registration.

Normative References

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

[RFC-2234], Crocker, D. and Overell, P.(Editors), "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, Internet Mail Consortium and Demon Internet Ltd., November 1997

[RFC-2821], Klensin, J., "Simple Mail Transfer Protocol", April 2001

Informative References

[SUBMITTER] draft-katz-submitter-01, Allman, E., and Katz, H., "SMTP Service Extension for Indicating the Responsible Submitter of an E-mail Message", (work in progress) May 2005

[draft-macquigg-authent-IOP], MacQuigg, D., "Email Authentication Inter-Operability Protocol", (work in progress) May 2005, http://purl.net/net/macquigg/email

[SRS] draft-mengwong-sender-rewrite-01, Wong, M.,"Sender Rewriting Scheme", (expired) http://www.libsrs2.org/

Author's Addresses

David R. MacQuigg, PhD
9320 East Mikelyn Lane
Tucson Arizona 85710 USA
Phone: 520-721-4583
Email: david_macquigg a-t yahoo.com
URL: http://purl.net/macquigg/

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgments

The author thanks all those who participated in discussions with him on the ietf and spf mailing lists.

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