Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 13, 2022

J. Peterson
Neustar
C. Wendt
Comcast

July 12, 2021

Connected Identity for STIR
draft-peterson-stir-rfc4916-update-04

Abstract

The SIP Identity header conveys cryptographic identity information about the originators of SIP requests. The Secure Telephone Identity Revisited (STIR) framework however provides no means for determining the identity of the called party in a traditional telephone calling scenario. This document updates prior guidance on the "connected identity" problem to reflect the changes to SIP Identity that accompanied STIR, and considers a revised problem space for connected identity as a means of detecting calls that have been retargeted to a party impersonating the intended destination, as well as the spoofing of mid-dialog or dialog-terminating events by intermediaries or third parties.

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 https://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 January 13, 2022.

Copyright Notice

Table of Contents

1. Introduction

The Session Initiation Protocol (SIP) [RFC3261] initiates sessions, and as a step in establishing sessions, it exchanges information about the parties at both ends of a session. Users review information about the calling party, for example, to determine whether to accept communications initiated by a SIP, in the same way that users of the telephone network assess "Caller ID" information before picking up calls. This information may sometimes be consumed by automata to make authorization decisions.

STIR [RFC8224] provides a cryptographic assurance of the identity of calling parties in order to prevent impersonation, which is a key enabler of unwanted robocalls, swatting, vishing, voicemail hacking, and similar attacks (see [RFC7340]). There also exists a related problem: the identity of the party who answers a call can differ from that of the initial called party for various innocuous reasons such as call forwarding, but in certain network environments, it is

possible for attackers to hijack the route of a called number and direct it to a resource controlled by the attacker. It can potentially be difficult to determine why a call reached a target other than the one originally intended, and whether the party ultimately reached by the call is one that the caller should trust. The lack of mutual authentication of parties moreover makes it possible for outside attackers to inject forged messages (e.g. BYE) into a SIP session.

The property of providing identity in the backwards direction of a call is here called "connected identity." Previous work on connected identity focused on fixing the core semantics of SIP. [RFC4916] allowed a mid-dialog request, such as an UPDATE [RFC3311], to convey identity in either direction within the context of an existing INVITE-initiated dialog. In an update to the original [RFC3261] behavior, [RFC4916] allowed that UPDATE to alter the From header field value for requests in the backwards direction: previously [RFC3261] required that the From header field values sent in requests in the backwards direction reflect the To header field value of the dialog-forming request, for various backwards-compatibility reasons. Under the original [RFC3261] rules, if Alice sent a dialog-forming request to Bob, then even if Bob's SIP service forwarded that dialog-forming request to Carol, Carol would still be required to put Bob's identity in the From header field value in any mid-dialog requests in the backwards direction.

One of the original motivating use cases for [RFC4916] was the use of connected identity with the SIP Identity [RFC4474] header field. While a mid-dialog request in the backwards direction (e.g. UPDATE) can be signed with Identity like any other SIP request, forwarded requests would not be signable without the ability to change the mid-dialog From header field value: Carol, say, would not be able to furnish a key to sign for Bob's identity, if Carol wanted to sign requests in the backwards direction. Carol would however be able to sign for her own identity in the From header field value, if mid-dialog requests in the backwards direction were permitted to vary from the original To header field value.

With the obsolence of [RFC4474] by [RFC8224], this specification updates [RFC4916] to reflect the changes to the SIP Identity header and the revised problem space of STIR. It also explores some new features that would be enabled by connected identity for STIR, including the use of connected identity to prevent route hijacking and to notify callers when an expected called party has successfully been reached. This document also addresses concerns about applying [RFC4916] connected identity to STIR discussed in the SIPBRANDY framework [RFC8862].

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Connected Identity Problem Statement for STIR

The STIR problem statement [RFC7340] enumerates robocalling, voicemail hacking, vishing, and swatting as problems with the modern telephone network that are enabled, or abetted, by impersonation: by the ability of a calling party to arbitrarily set the telephone number that will be rendered to end users to identify the caller.

Today, sophisticated adversaries can redirect calls on the PSTN to destinations other than the intended called party. For some call centers, like those associated with financial institutions, healthcare, and emergency services, an attacker could hope to gain valuable information about people or to prevent some classes of important services. Moreover, on the Internet, the lack of any centralized or even federated routing system for telephone numbers has resulted in deployments where the routing of calls is arbitrary: calls to telephone numbers might be unceremoniously dumped on a PSTN gateway, they might be sent to a default intermediary that makes forwarding decisions based on a local flat file, various mechanisms like private ENUM [RFC6116] might be consulted, or routing might be determined in some other, domain-specific way. In short, there are numerous attack surfaces that an adversary could explore to attempt to redirect calls to a particular number to someplace other than the intended destination.

Another motivating use case for connected identity is mid-dialog requests, including BYE. The potential for an intermediary to generate a forged BYE in the backwards direction has always been built-in to the stateful dialog management of SIP. For example, there is a class of mobile fraud attacks ("short-stopping") that rely on intermediary networks making it appear as if a call has terminated to one side, while maintaining that the call is still active to the other, in order to create a billing discrepancy that could be pocketed by the intermediary. If BYE requests in both directions of a SIP dialog could be authenticated with STIR, just like dialog-forming requests, then another impersonation vector leading to fraud in the telephone network could be shut down.

There are however practical limits to what securing the signaling can achieve. [RFC4916] rightly observed that once a SIP call has been

answered, the called party can be replaced by a different party (with a different identity) due to call transfer, call park and retrieval, and so on. In some cases, due to the presence of a back-to-back user agent, it can be effectively impossible for the calling party to know that this has happened. The problem statement considered for STIR focuses solely on signaling, not whether media from the connected party should be rendered to the caller when a dialog has been established. This specification does not consider further any threats that arise from a substitution of media.

4. Connected Identity in Provisional Dialogs

[RFC4916] identified a means of sending Identity header field values in the backwards direction before a final response to a dialog has been received by the UAC. It relied on the use of the UPDATE method to send the connected identity in the backwards direction after the UAS had received and responded to a PRACK [RFC3262] from the UAC, which would in turn have been triggered by a provisional 1xx response sent earlier by the UAC. [RFC4916] permits the From header field of the UPDATE to change the address of record of the recipient: if the original INVITE had been sent with a To header field value of "sip:bob@example.com", the UAS in its UPDATE could set the From header field value to "sip:carol@example.com." For STIR, this is a very important property, as Carol might not even possess a credential that can legitimately sign for Bob.

Per [RFC3262], UAC's that require connected identity MUST thus send a Require header field with the option tag 100rel in INVITEs in addition to an Identity header field value containing a PASSporT. UAS's that support this mechanism will first send a Require header field with the option tag 100rel in 1xx class responses to INVITEs that they receive, along with the necessary RSeq header field. The UAC will send a PRACK when it receives the reliable 1xx response from the UAS; the UAS, upon receiving a PRACK, responds with a 200 OK. At this point, the terminating UA is free to send an UPDATE [RFC3311] request in the backwards direction to the originating UA. This update will contain an Identity header, with a PASSporT that signs for the connected identity in its "orig" claim, which typically corresponds to the From header field value of the UPDATE request. If the PASSporT is valid, the originating UA will respond with an OK, and may perform any behaviors associated with the updated identity (see Section 7). Even if connected identity is not required by the originator of an INVITE request, it can still be solicited if available by sending the 100rel option tag in a Supported header field when sending an INVITE with an Identity header, which will trigger the preceding flow if the UAS supports connected identity.

Usually, the updated Identity reflects a changed to the From header field value. But in many operating environments, the From header field value does not contain the identity of the caller that has been asserted by the network, which is instead conveyed by the P-Asserted-Identity header field [RFC3325]. The contents of PAID are not used for dialog matching, and so in environments where PAID is used, it can be altered more dynamically. However, in order for the connected identity and a PASSporT to be conveyed in the backwards direction, a provisional dialog still needs to be established, and an UPDATE sent: in this case, it will be the UDPATE that contains the connected identity in its P-Asserted-Identity header field value, and that value will be signed by the PASSporT in its "orig" claim.

5. Connected Identity in Mid-Dialog and Dialog-Terminating Requests

The use of the connected identity mechanism here specified is not limited to provisional dialog requests. Once a dialog has been established with connected identity, any re-INVITEs from either the originating and terminating side, as well as any BYE requests, MUST contain Identity headers with valid PASSporTs based on the current To/From header field values for the dialog. This prevents third-parties from spoofing any mid-dialog requests in order to redirect media or similarly interfere with communications, as well as preventing denial of service teardowns by attackers.

Theoretically, any SIP requests in a dialog could be signed in this fashion, though it is unclear how valuable it would be for some (e.g. OPTIONS). Requests with specialized payloads such as INFO or MESSAGE, however, would require additional specification for how integrity protection for their bodies could be implemented. Some work has been done toward that for MESSAGE (see [I-D.peterson-stir-messaging]. This specification thus does not mandate PASSporTs for any requests sent in a dialog other than INVITE, UPDATE, and BYE.

It might seem tempting to require that, if an INVITE has been with an Identity header containing a PASSporT, any CANCEL request received for the dialog initiated by that INVITE must also contain an Identity header with a PASSporT. However, CANCEL requests can also sent be sent by stateful proxy servers engaged in parallel forking; for example, when branches need to be canceled because a final response has been received from a UAS. It is however REQUIRED by this specification that if a UAC sends a CANCEL for its own PASSporT-protected INVITE request, that it include an Identity header with a valid PASSporT in the CANCEL. UAS policy will have to determine the instances where it will accept unsigned CANCEL requests for a dialog initiated with a signed INVITE.

6. Interaction with Divert PASSporT

Many of the use cases that motivate connected identity involve retargeting: when a call acquires a new target (in its Request-URI) during transit, the terminating side needs a way to express to the originating side which destination the INVITE reached. In STIR, the "div PASSporT type [RFC8946] was created to securely record when a call was retargeted from one destination to another. Those "div" PASSporTs can be consumed on the terminating side by verification services to determine that a call has reached its eventual destination for the right reasons.

As specified in [RFC8946], the only way those diversion PASSporTs will be seen by the calling party is if redirection is used (SIP 3XX responses) instead of retargeting; while some network policies may want to conceal service logic from the originating party, sending redirections in the backwards direction is the only current defined way for secure indications of redirection to be revealed to the calling party. That in turn would allow the calling user agent to have a strong assurance that legitimate entities in the call path caused the request to reach a party that the caller did not anticipate.

This specification introduces another alternative. When per Section 4 the terminating side sends an UPDATE with an Identity header containing a PASSporT for its current From (or PAID) header field value, it MAY include in Identity header field values any "div" PASSporTs it received in the INVITE that initiated this provisional dialog. These "div" PASSporTs will enable the originating side to receive a secure assurance that the call is being fielded by the proper recipient per the routing of the call. Note however that "div" is not universally supported, and thus calls may be retargeted with generating a "div" PASSporT, so sending those PASSporTs in the backwards direction cannot be mandated. Also note that this will potentially reveal service logic to the called party.

7. Authorization Policy for Callers

In a traditional telephone call, the called party receives an alerting signal and can make a decision about whether or not to pick up a phone. They may have access to displayed information, like "Caller ID", to help them arrive at an authorization decision. The situation is more complicated for callers, however: callers typically expect to be connected to the proper destination and are often holding telephones in a position that would not enable them to see displayed information, if any were available for them to review--and moreover, their most direct response to a security breach would be to hang up the call they were in the middle of placing.

While this specification will not prescribe any user experience associated with placing a call, it assumes that callers might have some way to a set an authorization posture that will result in the right thing happening when the connected identity is not expected. This is analogous to a situation where SRTP negotiation fails because the keys exchanges at the media layer do not match fingerprints exchanged at the signaling layer: when a user requests confidentiality services, and they are unavailable, media should not be exchanged. Thus we assume that users have a way in their interface to require this criticality, on a per-call basis, or perhaps on a per-destination basis, that would cause their user agent to send the INVITE with a Require for 100rel. Similarly, users will not always place calls where the connected identity is crucial--but when they do, they should have a way to tell their devices that the call should not be completed if it arrives at an unexpected party.

Ultimately, authorization policy for called parties is difficult to set, as calls can end up at unexpected places for legitimate reasons. Some work has been done to make sure that secure diversion works with STIR, as described in Section 6.

8. Pre-Association with Destinations

Any connected identity mechanism will work best if the user knows before initiating a call that connected identity is supported by the destination side. Not every institution that a user wants to connect to securely will support STIR and connected identity out of the gate.

The user interface of modern smartphones support an address book from which users select telephone numbers to dial. Even when dialing a number manually, the interface frequently checks the address book and will display to users any provisioned name for the target of the call if one exists. Similarly, when clicking on a telephone number viewed on a web page, or similar service, smartphones often prompt users approve the access to the outbound dialer. These sorts of decision points, when the user is still interacting with the user interface, provide an opportunity to form a pre-association with the destination, and potentially even to exchange STIR PASSporTs in order to validate whether or not the expected destination can be reached securely. Again, this is probably most meaningful for contacting financial, government, or emergency services, for cases where reaching an unintended destination may have serious consequences.

The establishment of media-less dialogs has long been specified as a component of third-party call control in SIP [RFC3375], in which an INVITE is sent with no SDP. Similar media-less dialogs have been proposed for certain automata per [RFC5552]. In the STIR context, a media-less dialog is established by sending an INVITE with an

Identity header but no SDP. STIR-aware UAS's that support this specification, upon receiving an INVITE with no SDP, carrying a PASSporT with a 100rel in the Require header field value, SHOULD follow the mechanism described in Section 4 to send a provisional response and then an UPDATE carrying a PASSporT in the backwards direction. The PASSporT received in the backwards direction could be rendered to the originating user to help them decide if they want to place the call.

[More TBD. In some environments, that may require the use of mechanisms defined by [RFC8816].]

9. Connected Identity and Conferencing

The establishment of connected identity when there are more than two parties in a session will depend on the conferencing mechanism used.

[More TBD.]

10. Examples

[TBD: Revise RFC4916 examples to show new Identity header present in UPDATE and in a backwards-direction BYE.]

11. Updates to RFC4916

[TBD - ways that UPDATEs in the backwards direction can carry additional information in support of the above]

In general, the guidance of RFC4916 remains valid for RFC8224.

The deprecation of the Identity-Info header has a number of implications for RFC4916; all of the protocol examples need to be updated to reflect that.

12. Acknowledgments

We would like to thank YOU for your contributions to this specification.

13. IANA Considerations

This memo includes no request to IANA.

14. Security Considerations

TBD.

15. References

15.1. Informative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <https://www.rfc-editor.org/info/rfc3261>.

[RFC3262] Rosenberg, J. and H. Schulzrinne, "Reliability of Provisional Responses in Session Initiation Protocol (SIP)", RFC 3262, DOI 10.17487/RFC3262, June 2002, <https://www.rfc-editor.org/info/rfc3262>.

[RFC3311] Rosenberg, J., "The Session Initiation Protocol (SIP) UPDATE Method", RFC 3311, DOI 10.17487/RFC3311, October 2002, <https://www.rfc-editor.org/info/rfc3311>.

[RFC3325] Jennings, C., Peterson, J., and M. Watson, "Private Extensions to the Session Initiation Protocol (SIP) for Asserted Identity within Trusted Networks", RFC 3325, DOI 10.17487/RFC3325, November 2002, <https://www.rfc-editor.org/info/rfc3325>.

[RFC3375] Hollenbeck, S., "Generic Registry-Registrar Protocol Requirements", RFC 3375, DOI 10.17487/RFC3375, September 2002, <https://www.rfc-editor.org/info/rfc3375>.

[RFC4474] Peterson, J. and C. Jennings, "Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)", RFC 4474, DOI 10.17487/RFC4474, August 2006, <https://www.rfc-editor.org/info/rfc4474>.

[RFC4916] Elwell, J., "Connected Identity in the Session Initiation Protocol (SIP)", RFC 4916, DOI 10.17487/RFC4916, June 2007, <https://www.rfc-editor.org/info/rfc4916>.

[RFC7159] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", RFC 7159, DOI 10.17487/RFC7159, March 2014, <https://www.rfc-editor.org/info/rfc7159>.

[RFC7340] Peterson, J., Schulzrinne, H., and H. Tschofenig, "Secure Telephone Identity Problem Statement and Requirements", RFC 7340, DOI 10.17487/RFC7340, September 2014, <https://www.rfc-editor.org/info/rfc7340>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8224] Peterson, J., Jennings, C., Rescorla, E., and C. Wendt, "Authenticated Identity Management in the Session Initiation Protocol (SIP)", RFC 8224, DOI 10.17487/RFC8224, February 2018, <https://www.rfc-editor.org/info/rfc8224>.

[RFC8396] Peterson, J. and T. McGarry, "Managing, Ordering, Distributing, Exposing, and Registering Telephone Numbers (MODERN): Problem Statement, Use Cases, and Framework", RFC 8396, DOI 10.17487/RFC8396, July 2018, <https://www.rfc-editor.org/info/rfc8396>.

[RFC8816] Rescorla, E. and J. Peterson, "Secure Telephone Identity Revisited (STIR) Out-of-Band Architecture and Use Cases", RFC 8816, DOI 10.17487/RFC8816, February 2021, <https://www.rfc-editor.org/info/rfc8816>.

[RFC8862] Peterson, J., Barnes, R., and R. Housley, "Best Practices for Securing RTP Media Signaled with SIP", BCP 228, RFC 8862, DOI 10.17487/RFC8862, January 2021, <https://www.rfc-editor.org/info/rfc8862>.

[RFC8946] Peterson, J., "Personal Assertion Token (PASSporT) Extension for Diverted Calls", RFC 8946, DOI 10.17487/RFC8946, February 2021, <https://www.rfc-editor.org/info/rfc8946>.

15.2. Informative References

[I-D.peterson-stir-messaging] Peterson, J. and C. Wendt, "Messaging Use Cases and Extensions for STIR", draft-peterson-stir-messaging-01 (work in progress), February 2021.

[RFC5552] Burke, D. and M. Scott, "SIP Interface to VoiceXML Media Services", RFC 5552, DOI 10.17487/RFC5552, May 2009, <https://www.rfc-editor.org/info/rfc5552>.

[RFC6116] Bradner, S., Conroy, L., and K. Fujiwara, "The E.164 to Uniform Resource Identifiers (URI) Dynamic Delegation Discovery System (DDDS) Application (ENUM)", RFC 6116, DOI 10.17487/RFC6116, March 2011, <https://www.rfc-editor.org/info/rfc6116>.

Authors' Addresses

Jon Peterson
Neustar, Inc.
1800 Sutter St Suite 570
Concord, CA 94520
US

Email: jon.peterson@team.neustar

Chris Wendt
Comcast
One Comcast Center
Philadelphia, PA 19103
USA

Email: chris-ietf@chriswendt.net