             STIR Out of Band Architecture and Use Cases
                  draft-rescorla-stir-fallback-02.txt

Abstract

   The PASSporT format defines a token that can be carried by signaling
   protocols, including SIP, to cryptographically attest the identify of
   callers. Not all telephone calls use Internet signaling protocols,
   however, and some calls use them for only part of their signaling
   path. This document describes use cases that require the delivery of
   PASSporT objects outside of the signaling path, and defines
   architectures and semantics to provide this functionality.

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 December 16, 2017.

Copyright Notice

include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as
described in the Simplified BSD License.

Table of Contents

1. Introduction

   The STIR problem statement [RFC7340] describes widespread problems
   enabled by impersonation in the telephone network, including illegal
   robocalling, voicemail hacking, and swatting. As telephone services
   are increasingly migrating onto the Internet, and using Voice over IP
   (VoIP) protocols such as SIP [RFC3261], it is necessary for these
   protocols to support stronger identity mechanisms to prevent
   impersonation. For example, [I-D.ietf-stir-rfc4474bis] defines an
   Identity header of SIP requests capable of carrying a PASSporT
   [I-D.ietf-stir-passport] object in SIP as a means to
   cryptographically attest that the originator of a telephone call is
   authorized to use the calling party number (or, for native SIP cases,
   SIP URI) associated with the originator of the call. of the request.

Not all telephone calls use SIP today, however; and even those that do use SIP do not always carry SIP signaling end-to-end. Most calls from telephone numbers still traverse the Public Switched Telephone Network (PSTN) at some point. Broadly, calls fall into one of three categories:

1. One or both of the endpoints is actually a PSTN endpoint.

2. Both of the endpoints are non-PSTN (SIP, Jingle, ...) but the call transits the PSTN at some point.

3. Non-PSTN calls which do not transit the PSTN at all (such as native SIP end-to-end calls).

The first two categories represent the majority of telephone calls associated with problems like illegal robocalling: many robocalls today originate on the Internet but terminate at PSTN endpoints. However, the core network elements that operate the PSTN are legacy devices that are unlikely to be upgradable at this point to support an in-band authentication system. As such, those devices largely cannot be modified to pass signatures originating on the Internet--or indeed any inband signaling data--intact. Even if fields for tunneling arbtirary data can be found in traditional PSTN signaling, in some cases legacy elements would strip the signatures from those fields; in others, they might damage them to the point where they cannot be verified. For those first two categories above, any in-band authentication scheme does not seem practical in the current environment.

But while the core network of the PSTN remains fixed, the endpoints of the telephone network are becoming increasingly programmable and sophisticated. Landline "plain old telephone service" deployments, especially in the developed world, are shrinking, and increasingly being replaced by three classes of intelligent devices: smart phones, IP PBXs, and terminal adapters. All three are general purpose computers, and typically all three have Internet access as well as access to the PSTN. Additionally, various kinds of gateways increasingly front for legacy equipment. All of this provides a potential avenue for building an authentication system that implements stronger identity while leaving PSTN systems intact.

This capability also provides an ideal transitional technology while in-band STIR adoption is ramping up. It permits early adopters to use the technology even when intervening network elements are not yet STIR-aware, and through various kinds of gateways it may allow providers with a significant PSTN investment to still secure their calls with STIR.

This specification therefore builds on the PASSporT
[I-D.ietf-stir-passport] mechanism and the work of
[I-D.ietf-stir-rfc4474bis] to define a way that a PASSporT object
created in the originating network of a call can reach the
terminating network even when it cannot be carried end-to-end in-band
in the call signaling. This relies on a new service defined in this
document that permits the PASSporT object to be stored during call
processing and retrieved for verification purposes.

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 RFC
2119 [RFC2119].

3. Operating Environments

This section describes the environments in which the proposed
mechanism is intended to operate. In the simplest setting, Alice is
calling Bob through some set of gateways and/or the PSTN. Both Alice
and Bob have smart devices which can be modified, but they do not
have a clear connection between them: Alice cannot inject any data
into signaling which Bob can read, with the exception of the asserted
destination and origination E.164 numbers. The calling party number
might originate from her own device or from the network. These
numbers are effectively the only data that can be used for
coordination between the endpoints.

```
                                  +---------+
                                 /           \
                            +---              +---+
      +----------+         /                       \      +----------+
      |          |        |                         |     |          |
      |   Alice  |<------>|        Gateways         |<----->|    Bob   |
      | (caller) |        |         and/or          |     | (callee) |
      +----------+        |          PSTN           |     +----------+
                           \                       /
                            +---              +---+
                              \               /
                               +---------+
```
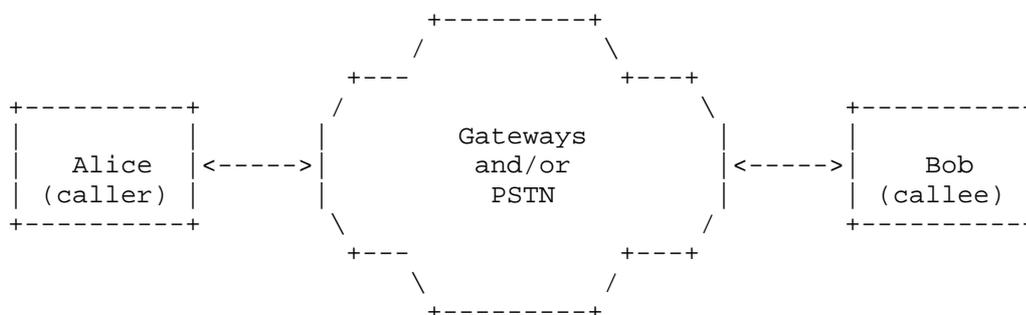
In a more complicated setting, Alice and/or Bob may not have a smart
or programmable device, but one or both of them are behind a STIR-
aware gateway that can participate in out-of-band coordination, as
shown below:

```
                         +---------+
                       /             \
                 +---                   +---+
   +----------+  +--+  /                    \  +--+  +----------+
   |          |  |  |  |                     |  |  |  |          |
   |  Alice   |<-|GW|->|      Gateways       |<-|GW|->|   Bob    |
   | (caller) |  |  |  |      and/or         |  |  |  | (callee) |
   +----------+  +--+  \       PSTN          /  +--+  +----------+
                 +---                   +---+
                       \             /
                         +---------+
```
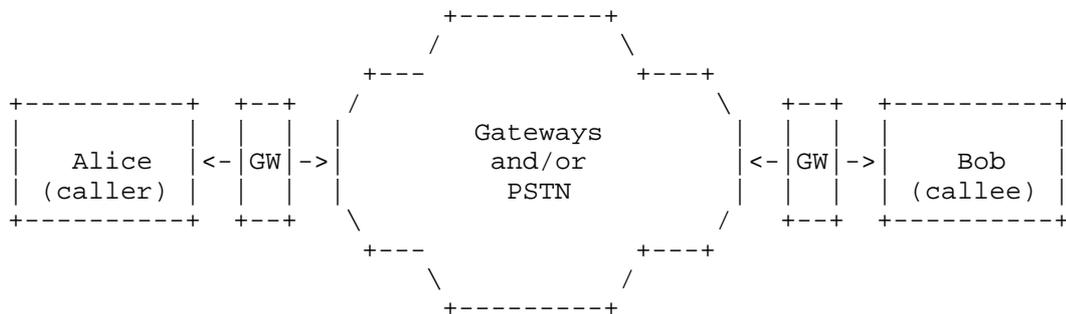
   In such a case, Alice might have an analog connection to her gateway/
   switch which is responsible for her identity. Similarly, the gateway
   would verify Alice's identity, generate the right calling party
   number information and provide that number to Bob using ordinary POTS
   mechanisms.

4. Dataflows

   Because in these operating environments endpoints cannot pass
   cryptographic information to one another directly through signaling,
   any solution must involve some rendezvous mechanism to allow
   endpoints to communicate. We call this rendezvous service a "call
   placement service" (CPS), a service where a record of call placement,
   in this case a PASSporT, can be stored for future retrieval. In
   principle this service could communicate any information, but
   minimally we expect it to include a full-form PASSporT that attests
   the caller, callee, and the time of the call. The callee can use the
   existence of a PASSporT for a given incoming call as rough validation
   of the asserted origin of that call. (See Section 9.1 for
   limitations of this design.)

   There are roughly two plausible dataflow architectures for the CPS:

      The callee registers with the CPS. When the caller wishes to
      place a call to the callee, it sends the PASSporT to the CPS,
      which immediately forwards it to the callee.

      The caller stores the PASSporT with the CPS at the time of call
      placement. When the callee receives the call, it contacts the CPS
      and retrieves the PASSporT.

   While the first architecture is roughly isomorphic to current VoIP
   protocols, it shares their drawbacks. Specifically, the callee must
   maintain a full-time connection to the CPS to serve as a notification
   channel. This comes with the usual networking costs to the callee
   and is especially problematic for mobile endpoints. Indeed, if the

endpoints had the capabilities to implement such an architecture,
they could surely just use SIP or some other protocol to set up a
secure session; even if the media were going through the traditional
PSTN, a "shadow" SIP session could convey the PASSporT. Thus, we
focus on the second architecture in which the PSTN incoming call
serves as the notification channel and the callee can then contact
the CPS to retrieve the PASSporT.

5. Use Cases

   The following are the motivating use cases for this mechanism. Bear
   in mind that just as in [I-D.ietf-stir-rfc4474bis] there may be
   multiple Identity headers in a single SIP INVITE, so there may be
   multiple PASSporTs in this out-of-band mechanism associated with a
   single call. For example, a SIP user agent might create a PASSporT
   for a call with an end user credential, and as the call exits the
   originating administrative domain the network authentication service
   might create its own PASSporT for the same call. As such, these use
   cases may overlap in the processing of a single call.

5.1. Case 1: VoIP to PSTN Call

   A call originates in the SIP world in a STIR-aware administrative
   domain. The local authentication service for that administrative
   domain creates a PASSporT which is carried in band in the call per
   [I-D.ietf-stir-rfc4474bis]. The call is routed out of the
   originating administrative domain and reaches a gateway to the PSTN.
   Eventually, the call will terminate on a mobile smartphone that
   supports this out-of-band mechanism.

   In this use case, the originating authentication service can store
   the PASSporT with the appropriate CPS for the target telephone number
   as a fallback in case SIP signaling will not reach end-to-end. When
   the destination mobile smartphone receives the call over the PSTN, it
   consults the CPS and discovers a PASSporT from the originating
   telephone number waiting for it. It uses this PASSporT to verify the
   calling party number.

5.2. Case 2: Two Smart PSTN endpoints

   A call originates with an enterprise PBX that has both Internet
   access and a built-in gateway to the PSTN. It will immediately drop
   its call to the PSTN, but before it does, it provisions a PASSporT on
   the CPS associated with the target telephone number.

   After normal PSTN routing, the call lands on a smart mobile handset
   that supports the STIR out-of-band mechanism. It queries the
   appropriate CPS over the Internet to determine if a call has been

placed to it by a STIR-aware device. It finds the PASSporT
provisioned by the enterprise PBX and uses it to verify the calling
party number.

5.3. Case 3: PSTN to VoIP Call

A call originates with an enterprise PBX that has both Internet
access and a built-in gateway to the PSTN. It will immediate drop
the call to the PSTN, but before it does, it provisions a PASSporT
with the CPS associated with the target telephone number. However,
it turns out that the call will eventually route through the PSTN to
an Internet gateway, which will translate this into a SIP call and
deliver it to an administrative domain with a STIR verification
service.

In this case, there are two subcases for how the PASSporT might be
retrieved. In subcase 1, the Internet gateway that receives the call
from the PSTN could query the appropriate CPS to determine if the
original caller created and provisioned a PASSporT for this call. If
so, it can retrieve the PASSporT and, when it creates a SIP INVITE
for this call, add a corresponding Identity header per
[I-D.ietf-stir-rfc4474bis]. When the SIP INVITE reaches the
destination administrative domain, it will be able to verify the
PASSporT normally. Note that to avoid discrepancies with the Date
header field value, only full-form PASSporT should be used for this
purpose. In subcase 2, the gateway does not retrieve the PASSporT
itself, but instead the verification service at the destination
administrative domain does so. Subcase 1 would perhaps be valuable
for deployments where the destination administrative domain supports
in-band STIR but not out-of-band STIR.

5.4. Case 4: Gateway Out-of-band

A call originates in the SIP world in a STIR-aware administrative
domain. The local authentication service for that administrative
domain creates a PASSporT which is carried in band in the call per
[I-D.ietf-stir-rfc4474bis]. The call is routed out of the
originating administrative domain and eventually reaches a gateway to
the PSTN.

In this case, the originating authentication service does not support
the out-of-band mechanism, so instead the gateway to the PSTN
extracts the PASSporT from the SIP request and provisions it to the
CPS. (When the call reaches the gateway to the PSTN, the gateway
might first check the CPS to see if a PASSporT object had already
been provisioned for this call, and only provision a PASSporT if none
is present).

Ultimately, the call may terminate on the PSTN, or be routed back to
the IP world. In the former case, perhaps the destination endpoints
queries the CPS to retrieve the PASSporT provisioned by the first
gateway. Or if the call ultimately returns to the IP world, it might
be the gateway from the PSTN back to the Internet that retrieves the
PASSporT from the CPS and attaches it to the new SIP INVITE it
creates, or it might be the terminating administrative domain's
verification service that checks the CPS when an INVITE arrives with
no Identity header field. Either way the PASSporT can survive the
gap in SIP coverage caused by the PSTN leg of the call.

6. Authorization for Storing and Retrieving PASSporTs

   The use cases show a variety of entities accessing the CPS to store
   and retrieve PASSporTs. The question of how the CPS authorizes the
   storage and retrieval of PASSporT is thus a key design decision in
   the architecture.

   The STIR architecture assumes that service providers and in some
   cases end user devices will have credentials suitable for attesting
   authority over telephone numbers per [I-D.ietf-stir-certificates].
   These credentials provide the most obvious way that a CPS can
   authorize the storage and retrieval of PASSporTs. However, as use
   cases 3 and 4 in Section 5 show, it may sometimes make sense for the
   entity storing or retrieving PASSporTs to be an intermediary rather
   than a device associated with either the originating or terminating
   side of a call, and those intermediaries often would not have access
   to STIR credentials covering the telephone numbers in question.

   It is an explicit design goal of this mechanism to minimize the
   potential privacy exposure of using a CPS. Ideally, the out-of-band
   mechanism should not result in a worse privacy situation than in-band
   [I-D.ietf-stir-rfc4474bis] STIR: for in-band, we might say that a SIP
   entity is authorized to receive a PASSporT if it is an intermediate
   or final target of the routing of a SIP request. As the originator
   of a call cannot necessarily predict the routing path a call will
   follow, an out-of-band mechanism could conceivably even improve on
   the privacy story. As a first step, transport-level security can
   provide confidentiality from eavesdroppers for both the storage and
   retrieval of PASSporTs.

6.1. Storage

   For authorizing the storage of PASSporTs, the architecture can permit
   some flexibility. A CPS could adopt a policy where it will store any
   valid PASSporT - that is, the CPS could act as a limited verification
   service and validate the PASSporT, only storing it if the timestamp
   and signature are valid. In that case, it would not matter whether

the CPS received a PASSporT from the authentication service that
created it or from an intermediary gateway downstream in the routing
path as in case 4: so long as the PASSporT is valid, it would be
stored.

6.2. Retrieval

For retrieval of PASSporTs, the story is a bit more complicated.
Beyond using transport-level security when storing and retrieving
PASSporTs, the architecture must include some way to constrain access
to the PASSporTs stored at a CPS. How those constraints should
operate depends on the semantics of the request used to retrieve
PASSporTs. A retrieval request could have one of the following three
semantics:

   a) Are there any current PASSporTs for calls originating from
   1.111.111.1111?

   b) Are there any current PASSporTs for calls destined to
   2.222.222.2222?

   c) Are there any current PASSporTs for calls originating from
   1.111.111.1111 and destined to 2.222.222.2222?

Each of these three semantics results in very different properties
for the architecture. If a CPS permitted just anyone to ask for all
PASSporTs that happen to exist for current calls to or from a given
telephone number, that would be an unacceptable privacy situation.
Although on the surface semantic (c) may seem sufficiently strict, a
particular adversary might only be interested in learning when one
specific party calls another, and there are certainly cases in which
that could pose a significant security risk. While a CPS could
eventually refuse to answer repeated requests from a single device
that is obviously polling to collect the state of calls in progress,
more sophisticated adversaries could outwit any attempt to do source
filtering on requests at the CPS.

The semantics of (a) or (b) vs. (c) could be very significant when
the originating and destination numbers are for call centers or
similar organizations that send or receive a vast amount of calls for
a single number. In a case where many thousands of people are trying
to call a number where tickets have just gone on sale, for example,
it might be difficult using semantics (b) to sift through all of the
call setup attempts in progress to find a PASSporT matching any
particular call. A more narrow semantic like (c) would make it far
easier.

Sometimes the more narrow semantics of (c) can pose an obstacle to
acquiring the right PASSporT, for example in call forwarding cases
where retargeting of the request has occurred. Even using semantic
(b) would be problematic if the PASSporT stored by the originating
authentication service had a different original "dest". Mechanisms
have been proposed for STIR to patch this by creating PASSporTs that
record the diversion (see [I-D.peterson-passport-divert]), and
potentially a CPS could store these additional PASSporT objects and
supply them through the retrieval interface.

If we assume that the party retrieving PASSporTs from the CPS has a
STIR credential attesting authority over the terminating number, then
two more attractive mechanisms become possible: using authentication
and encryption. Note however that in some use cases, like case 3
subcase 1 above, the retrieving party is an intermediary who would
not have access to the necessary credentials. However, this might
argue that subcase 1 should be disallowed for security reasons, and
only subcase 2 should be permitted.

6.2.1. Authentication

For any of the three proposed retrieval semantics, a CPS could
authenticate a request to retrieve PASSporTs and only release
PASSporTs that have a destination that matches the credential
provided by the requestor. Per semantic (b), if a smart endpoint has
a credential for 2.222.222.2222, it could send a request to the CPS
signed with that credential to retrieve any PASSporTs for calls in
progress to 2.222.222.2222. In this case, (a) and (c) have very
similar semantics: when the requestor asks for (a), effectively they
would receive only those PASSporTs coming from 1.111.111.1111 that
are destined to 2.222.222.2222 - though perhaps in cases where the
call had been forwarded, a CPS aware of the situation could
understand that the new destination should be authorized to see the
original PASSporT.

On balance, an approach along the lines of requiring authenticating
requests with semantic (a) appears attractive as a direction for out-
of-band.

6.2.2. Encryption

Some of the privacy risks on the retrieval side could potentially be
mitigated with encryption. If all PASSporTs stored at a CPS were
encrypted with a key belonging to the intended destination, then
potentially the CPS could allow almost anyone to download PASSporTs
using semantics (a) or (b) without much fear of compromising private
information about calls in progress - provided that the CPS always
provided at least one encrypted blob in response to a request, even

if there was no call in progress. It would also prevent the CPS
itself from learning the contents of PASSporTs, and thus metadata
about calls in progress, which would make the CPS a less attractive
target for pervasive monitoring (see [RFC7258]). However, encrypting
PASSporTs faces some substantial difficulties.

First, this requires the entity that stores the PASSporT to have
access to a public key associated with the intended called party to
be used to encrypt the PASSporT. Discovering this key would require
some new service that does not exist today; depending on how the CPS
is architected, however, some kind of key store or repository could
be implemented adjacent to it, and perhaps even incorporated into its
operation. This key discovery problem is compounded by the fact that
there can potentially be multiple entities that have authority over a
telephone number: a carrier, a reseller, an enterprise, and an end
user might all have credentials permitting them to attest that they
are allowed to originate calls from a number, say. PASSporTs might
need to be encrypted with multiple keys in the hopes that one will be
decipherable by the relying party.

Second, in call forwarding cases, the difficulties in managing the
relationship between PASSporTs with the diversion extension
[I-D.peterson-passport-divert] become more serious. The originating
authentication service would encrypt the PASSporT with the public key
of the intended destination, but when a call is forwarded, it may go
to a destination that does not possess the corresponding private key.
It would require special behavior on the part of the retargeting
entity, and probably the CPS as well, to accommodate encrypted
PASSporTs that show a secure chain of diversion.

Another side effect of encrypting PASSporTs before storing them is
that the CPS can no longer validate the PASSporTs since it cannot in
fact read them. However, a CPS needs to know enough about PASSporTs
so that it can respond to requests to retrieve them, whichever
semantics are used - which means the CPS will always process some
amount of metadata (even if some sort of hash function is used to
index PASSporTs). Unless the storer of PASSporTs is authenticated,
it may be possible for attackers to inject bogus PASSporTs into the
system. Note however that merely injecting a bogus PASSporT into a
CPS will not allow attackers to impersonate parties. That is because
verification services trust a PASSporT based its own internal
signature, not based on where the verification service found it.
This is orthogonal to the current question of how the CPS authorizes
an endpoint to acquire a PASSporT; though of course spamming a CPS
with large numbers of bogus PASSporTs could cause a denial of service
or similar problems with retrieval of PASSporTs.

7. Solution Architecture

   In this section, we discuss a strawman architecture for providing the
   service described in the previous sections. This discussion is
   deliberately sketchy, focusing on broad concepts and skipping over
   details. The intent here is merely to provide a rough concept, not a
   complete solution.

7.1. Credentials and Phone Numbers

   We start from the premise of the STIR problem statement [RFC7340]
   that phone numbers can be associated with credentials which can be
   used to attest ownership of numbers. For purposes of exposition, we
   will assume that ownership is associated with the endpoint (e.g., a
   smartphone) but it might well be associated with a provider or
   gateway acting for the endpoint instead. It might be the case that
   multiple entities are able to act for a given number, provided that
   they have the appropriate authority. [I-D.ietf-stir-certificates]
   describes a credentials system suitable for this purpose; the
   question of how an entity is determined to have control of a given
   number is out of scope for the current document.

7.2. Solution Architecture

   An overview of the basic calling and verification process is shown
   below. In this diagram, we assume that Alice has the number
   +1.111.111.1111 and Bob has the number +2.222.222.2222.

```
Alice                        Call Placement Service                    Bob
--------------------------------------------------------------------------
Store PASSporT ---------------->

Call from 1.111.111.1111 ------------------------------------------------->


                                 <-  Authenticate as 1.222.222.2222  ---->

                                    <-------------- Retrieve call record
                                                         from 1.111.111.1111?

                                    (1.222.222.2222,1.111.111.1111) -->

                                                    [Ring phone with callerid
                                                         = 1.111.111.1111]
```
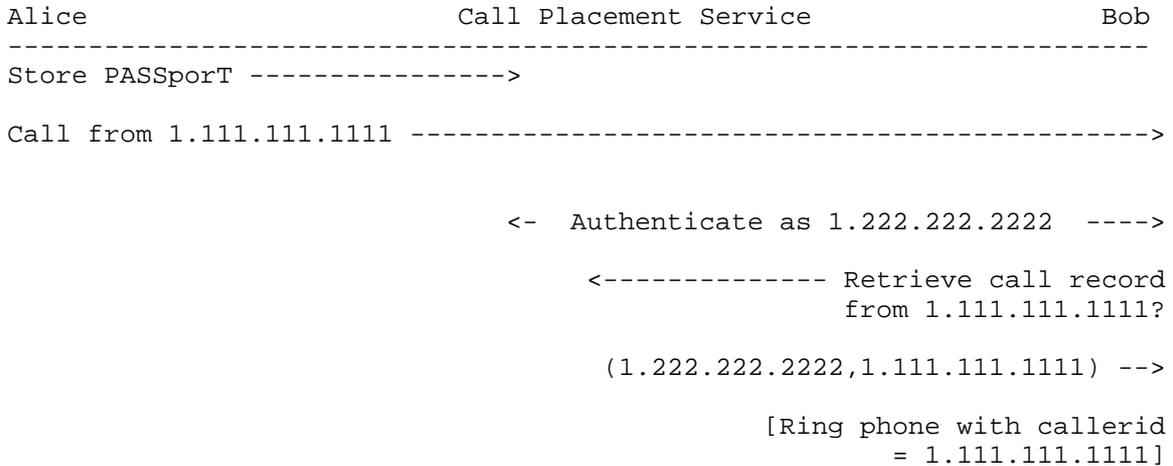
   When Alice wishes to make a call to Bob, she contacts the CPS and
   stores a PASSporT on the CPS. The CPS validates the PASSporT before

indexing it so that it can be acquired with a request from Bob's number

Once Alice has stored the PASSporT, she then places the call to Bob as usual. At this point, Bob's phone would usually ring and display Alice's number (+1.111.111.1111), which is informed by the existing PSTN mechanisms for relying a calling party number (i.e., the CIN field of the IAM). Instead, Bob's phone transparently contacts the CPS, authenticates itself, and requests any current PASSporTs for calls from Alice. The CPS responds with any such PASSporTs (assuming they exist). If such a PASSpoRT exists, and the verification service in Bob's phone validates it, then Bob's phone can then present the calling party number information as valid. Otherwise, the call is unverifiable. Note that this does not necessarily mean that the call is bogus; because we expect incremental deployment many legitimate calls will be unverifiable.

7.3. Security Analysis

The primary attack we seek to prevent is an attacker convincing the callee that a given call is from some other caller C. There are two scenarios to be concerned with:

   The attacker wishes to impersonate a target when no call from that target is in progress.

   The attacker wishes to substitute himself for an existing call setup as described in Section 7.4.

If an attacker can inject fake PASSporT into the CPS or in the communication from the CPS to the callee, he can mount either attack. As PASSporTs should be digitally signed by an appropriate authority for the number and verified by the callee (see Section 7.1), this should not arise in ordinary operations. For privacy and robustness reasons, using TLS on the originating side when storing the PASSporT at the CPS is recommended.

The entire system depends on the security of the credential infrastructure. If the authentication credentials for a given number are compromised, then an attacker can impersonate calls from that number. However, that is no different from in-band [I-D.ietf-stir-rfc4474bis] STIR.

7.4. Substitution Attacks

All that receipt of the PASSporT from the CPS proves to the called party is that Alice is trying to call Bob (or at least was as of very recently) - it does not prove that any particular incoming call is

from Alice. Consider the scenario in which we have a service which
provides an automatic callback to a user-provided number. In that
case, the attacker can try to arrange for a false caller-id value, as
shown below:

```
Attacker              Callback Service              CPS              Bob
------------------------------------------------------------------------
Place call to Bob ---------->

                              Store PASSporT for
                              CS:Bob -------------->

Call from CS (forged caller-id info)  -------------------------------->

                              Call from CS --------------------------> X


                                              <----- Retrieve PASSporT
                                                        for CS:Bob

                      PASSporT for CS:Bob -------------------------->

                                              [Ring phone with callerid = CS]
```

In order to mount this attack, the attacker contacts the Callback
Service (CS) and provides it with Bob's number. This causes the CS
to initiate a call to Bob. As before, the CS contacts the CPS to
insert an appropriate PASSporT and then initiates a call to Bob.
Because it is a valid CS injecting the PASSporT, none of the security
checks mentioned above help. However, the attacker simultaneously
initiates a call to Bob using forged caller-id information
corresponding to the CS. If he wins the race with the CS, then Bob's
phone will attempt to verify the attacker's call (and succeed since
they are indistinguishable) and the CS's call will go to busy/voice
mail/call waiting. Note: in a SIP environment, the callee might
notice that there were multiple INVITEs and thus detect this attack.

8. Call Placement Service Discovery

   In order for the two ends of the out-of-band dataflow to coordinate,
   they must agree on a way to discover a CPS and retrieve PASSporT
   objects from it based solely on the rendezvous information available:
   the calling party number and the called number. There are a number
   of potential service discovery mechanisms that could be used for this
   purpose. The means of service discovery may vary by use case.

   Although the discussion above is written in terms of a single CPS,
   having a significant fraction of all telephone calls result in

storing and retrieving PASSporTs at a single monolithic CPS has
obvious scaling problems, and would as well allow the CPS to gather
metadata about a very wide set of callers and callees. These issues
can be alleviated by operational models with a federated CPS; any
service discovery mechanism for out-of-band STIR should enable
federation of the CPS function.

Some service discovery possibilities under consideration include the
following:

   If a credential lookup service is already available, the CPS
   location can also be recorded in the callee's credentials; an
   extension to [I-D.ietf-stir-certificates] could for example
   provide a link to the location of the CPS where PASSporTs should
   be stored for a destination.

   There exist a number of common directory systems that might be
   used to translate telephone numbers into the URIs of a CPS. ENUM
   [RFC6116] is commonly implemented, though no "golden root" central
   ENUM administration exists that could be easily reused today to
   help the endpoints discover a common CPS. Other protocols
   associated with queries for telephone numbers, such as the TeRI
   [I-D.peterson-modern-teri] protocol, could also serve for this
   application.

   Another possibility is to use a single distributed service for
   this function. VIPR [I-D.rosenberg-dispatch-vipr-overview]
   proposed a RELOAD [RFC6940] usage for telephone numbers to help
   direct calls to enterprises on the Internet. It would be possible
   to describe a similar RELOAD usage to identify the CPS where calls
   for a particular telephone number should be stored. One advantage
   that the STIR architecture has over VIPR is that it assumes a
   credential system that proves authority over telephone numbers;
   those credentials could be used to determine whether or not a CPS
   could legitimately claim to be the proper store for a given
   telephone number.

Future versions of this specification will identify suitable service
discovery mechanisms for out-of-band STIR.

9. To Do

Section 4 provides a broad sketch of an approach. In this section,
we consider some areas for additional work. Readers can feel free to
skip this section, as it is not necessary to get the flavor of the
document.

9.1. Credential Lookup

   In order to encrypt a PASSporT (see Section 6.2.2), the caller needs
   access to the callee's credentials (specifically their public key).
   This requires some sort of directory/lookup system. This document
   does not specify any particular scheme, but a list of requirements
   would be something like:

   Obviously, if there is a single central database and the caller and
   callee each contact it in real time to determine the other's
   credentials, then this represents a real privacy risk, as the central
   database learns about each call. A number of mechanisms are
   potentially available to mitigate this:

      Have endpoints pre-fetch credentials for potential counterparties
      (e.g., their address book or the entire database).

      Have caching servers in the user's network that proxy their
      fetches and thus conceal the relationship between the user and the
      credentials they are fetching.

   Clearly, there is a privacy/timeliness tradeoff in that getting
   really up-to-date knowledge about credential validity requires
   contacting the credential directory in real-time (e.g., via OCSP).
   This is somewhat mitigated for the caller's credentials in that he
   can get short-term credentials right before placing a call which only
   reveals his calling rate, but not who he is calling. Alternately,
   the CPS can verify the caller's credentials via OCSP, though of
   course this requires the callee to trust the CPS's verification.
   This approach does not work as well for the callee's credentials, but
   the risk there is more modest since an attacker would need to both
   have the callee's credentials and regularly poll the database for
   every potential caller.

   We consider the exact best point in the tradeoff space to be an open
   issue.

10. Acknowledgments

   The ideas in this document come out of discussions with Richard
   Barnes and Cullen Jennings.

11. IANA Considerations

   This memo includes no request to IANA.

12. Security Considerations

   This entire document is about security, but the detailed security
   properties depend on having a single concrete scheme to analyze.

13. Informative References

   [I-D.ietf-stir-certificates]
            Peterson, J. and S. Turner, "Secure Telephone Identity
            Credentials: Certificates", draft-ietf-stir-
            certificates-14 (work in progress), May 2017.

   [I-D.ietf-stir-passport]
            Wendt, C. and J. Peterson, "Personal Assertion Token
            (PASSporT)", draft-ietf-stir-passport-11 (work in
            progress), February 2017.

   [I-D.ietf-stir-rfc4474bis]
            Peterson, J., Jennings, C., Rescorla, E., and C. Wendt,
            "Authenticated Identity Management in the Session
            Initiation Protocol (SIP)", draft-ietf-stir-rfc4474bis-16
            (work in progress), February 2017.

   [I-D.peterson-modern-teri]
            Peterson, J., "An Architecture and Information Model for
            Telephone-Related Information (TeRI)", draft-peterson-
            modern-teri-02 (work in progress), October 2016.

   [I-D.peterson-passport-divert]
            Peterson, J., "PASSporT Extension for Diverted Calls",
            draft-peterson-passport-divert-01 (work in progress), June
            2017.

   [I-D.rosenberg-dispatch-vipr-overview]
            Rosenberg, J., Jennings, C., and M. Petit-Huguenin,
            "Verification Involving PSTN Reachability: Requirements
            and Architecture Overview", draft-rosenberg-dispatch-vipr-
            overview-04 (work in progress), October 2010.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
            Requirement Levels", BCP 14, RFC 2119,
            DOI 10.17487/RFC2119, March 1997,
            <http://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,
              <http://www.rfc-editor.org/info/rfc3261>.

   [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,
              <http://www.rfc-editor.org/info/rfc6116>.

   [RFC6940] Jennings, C., Lowekamp, B., Ed., Rescorla, E., Baset, S.,
              and H. Schulzrinne, "REsource LOcation And Discovery
              (RELOAD) Base Protocol", RFC 6940, DOI 10.17487/RFC6940,
              January 2014, <http://www.rfc-editor.org/info/rfc6940>.

   [RFC7258] Farrell, S. and H. Tschofenig, "Pervasive Monitoring Is an
              Attack", BCP 188, RFC 7258, DOI 10.17487/RFC7258, May
              2014, <http://www.rfc-editor.org/info/rfc7258>.

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

Authors' Addresses

   Eric Rescorla
   Mozilla

   Email: ekr@rtfm.com

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

   Email: jon.peterson@neustar.biz