Framework for Emergency Calling Using Internet Multimedia

Abstract

   The IETF has standardized various aspects of placing emergency calls.
   This document describes how all of those component parts are used to
   support emergency calls from citizens and visitors to authorities.

Table of Contents

1. Introduction

Requesting help in an emergency using a communications device such as
a telephone (landline or mobile) is an accepted practice in many
parts of the world. As communications devices increasingly utilize
the Internet to interconnect and communicate, users will expect to
use such devices to request help. This document describes
establishment of a communications session by a user to a "Public
Safety Answering Point" (PSAP), that is, a call center established by
response agencies to accept emergency calls. Such citizen-/
visitor-to-authority calls can be distinguished from those that are
created by responders (authority-to-authority) using public
communications infrastructure often involving some kind of priority
access as defined in Emergency Telecommunications Service (ETS) in IP
Telephony [RFC4190]. They can also be distinguished from emergency
warning systems that are authority-to-citizen.

Supporting emergency calling requires cooperation by a number of
elements, their vendors, and service providers. This document
discusses how end devices and applications create emergency calls,
how access networks supply location for some of these devices, how
service providers assist the establishment and routing, and how PSAPs
receive calls from the Internet.

The emergency response community will have to upgrade their
facilities to support a wider range of communications services, but
cannot be expected to handle wide variations in device and service
capability. New devices and services are being made available that
could be used to make a request for help that are not traditional
telephones, and users are increasingly expecting to use them to place
emergency calls. However, many of the technical advantages of
Internet multimedia require re-thinking the traditional emergency
calling architecture. This challenge also offers an opportunity to
improve the operation of emergency calling technology, while
potentially lowering its cost and complexity.

It is beyond the scope of this document to enumerate and discuss all
the differences between traditional (Public Switched Telephone
Network) and IP-based telephony, but calling on the Internet is
characterized by:

o interleaving over the same infrastructure of a wider variety of
   services;

o separation of the access provider from the application provider;

o media other than voice (for example, video and text in several
   forms);

o   potential mobility of all end systems, including endpoints
    nominally thought of as fixed systems and not just those using
    radio access technology. For example, consider a wired phone
    connected to a router using a mobile data network such as
    Evolution Data Optimized (EV-DO) as an uplink.

This document focuses on how devices using the Internet can place
emergency calls and how PSAPs can handle Internet multimedia
emergency calls natively, rather than describing how circuit-switched
PSAPs can handle Voice over IP (VoIP) calls. In many cases, PSAPs
making the transition from circuit-switched interfaces to packet-
switched interfaces may be able to use some of the mechanisms
described here, in combination with gateways that translate packet-
switched calls into legacy interfaces, e.g., to continue to be able
to use existing call taker equipment. There are many legacy
telephone networks that will persist long after most systems have
been upgraded to IP origination and termination of emergency calls.
Many of these legacy systems route calls based on telephone numbers.
Gateways and conversions between existing systems and newer systems
defined by this document will be required. Since existing systems
are governed primarily by local government regulations and national
standards, the gateway and conversion details will be governed by
national standards and thus are out of scope for this document.

Existing emergency call systems are organized locally or nationally;
there are currently few international standards. However, the
Internet crosses national boundaries, and thus Internet standards are
required. To further complicate matters, VoIP endpoints can be
connected through tunneling mechanisms such as virtual private
networks (VPNs). Tunnels can obscure the identity of the actual
access network that knows the location. This significantly
complicates emergency calling, because the location of the caller and
the first element that routes emergency calls can be on different
continents, with different conventions and processes for handling of
emergency calls.

The IETF has historically not created national variants of its
standards. Thus, this document attempts to take into account best
practices that have evolved for circuit-switched PSAPs, but it makes
no assumptions on particular operating practices currently in use,
numbering schemes, or organizational structures.

This document discusses the use of the Session Initiation Protocol
(SIP) [RFC3261] by PSAPs and calling parties. While other inter-
domain call signaling protocols may be used for emergency calling,
SIP is ubiquitous and possesses the proper support of this use case.
Only protocols such as H.323, XMPP/Jingle, ISUP, and SIP are suitable
for inter-domain communications, ruling out Media Gateway Controller

protocols such as the Media Gateway Control Protocol (MGCP) or H.248/
Megaco. The latter protocols can be used by the enterprise or
carrier placing the call, but any such call would reach the PSAP
through a media gateway controller, similar to how inter-domain VoIP
calls would be placed. Other signaling protocols may also use
protocol translation to communicate with a SIP-enabled PSAP. Peer-
to-peer SIP (p2psip) is not considered in this document.

Existing emergency services rely exclusively on voice and
conventional text telephony ("TTY") media streams. However, more
choices of media offer additional ways to communicate and evaluate
the situation as well as to assist callers and call takers in making
and handling emergency calls, respectively. For example, instant
messaging and video could improve the ability to communicate and
evaluate the situation and to provide appropriate instruction prior
to arrival of emergency crews. Thus, the architecture described here
supports the creation of sessions of any media type, negotiated
between the caller and PSAP using existing SIP mechanisms [RFC3264].

This document focuses on the case in which all three steps in the
emergency calling process -- location configuration, call routing,
and call placement -- can be and are performed by the calling
endpoint, with the endpoint's Access Service Provider supporting the
process by providing location information. In this case, calls may
be routed via an application-layer Communications Service Provider
(e.g., a Voice Service Provider) but need not be. The underlying
protocols can also be used to support other models in which parts of
the process are delegated to the Communications Service Provider.
This document does not address in detail either these models or
interoperability issues between them and the model described here.

Since this document is a framework document, it does not include
normative behavior. [PHONEBCP] describes the best current practice
for this subject and contains normative language for devices as well
as access and calling network elements.

Supporting emergency calling does not require any specialized SIP
header fields, request methods, status codes, message bodies, or
event packages, but it does require that existing mechanisms be used
in certain specific ways, as described below. User agents (UAs)
unaware of the recommendations in this document may be able to place
emergency calls, but functionality may be impaired. For example, if
the UA does not implement the location mechanisms described, an
emergency call may not be routed to the correct PSAP, and if the
caller is unable to supply his exact location, dispatch of emergency
responders may be delayed. Suggested behavior for both endpoints and
servers is provided.

From the point of view of the PSAP, three essential elements
characterize an emergency call:

o The call is routed to the most appropriate PSAP, based primarily
   on the location of the caller.

o The PSAP must be able to automatically obtain the location of the
   caller with sufficient accuracy to dispatch a responder to help
   the caller.

o The PSAP must be able to re-establish a session to the caller if
   for any reason the original session is disrupted.

2. Terminology

This document uses terms from [RFC3261], [RFC5222], and [RFC5012].
In addition, the following terms are used:

Access network: The access network supplies IP packet service to an
   endpoint. Examples of access networks include digital subscriber
   lines (DSLs), cable modems, IEEE 802.11, WiMaX, enterprise local
   area networks, and cellular data networks.

Confidence: Confidence is an estimate indicating how sure the
   measuring system is that the actual location of the endpoint is
   within the bounds defined by the uncertainty value, expressed as a
   percentage. For example, a value of 90% indicates that the actual
   location is within the uncertainty nine times out of ten.

Dispatch location: The dispatch location is the location used for
   dispatching responders to the person in need of assistance. The
   dispatch location must be sufficiently precise to easily locate
   the caller; typically, it needs to be more accurate than the
   routing location.

Location configuration: During location configuration, an endpoint
   learns its physical location.

Location Configuration Protocol (LCP): A protocol used by an
   endpoint to learn its location.

Location conveyance: Location conveyance delivers location
   information to another element.

Location determination: Location determination finds where an
   endpoint is physically located. For example, the endpoint may
   contain a Global Navigation Satellite System (GNSS) receiver used
   to measure its own location or the location may be determined by a
   network administrator using a wiremap database.

Location Information Server (LIS): A Location Information Server
   stores location information for retrieval by an authorized entity.

Mobile device: A mobile device is a user agent that may change its
   physical location and possibly its network attachment point during
   an emergency call.

National Emergency Number Association (NENA): The National Emergency
   Number Association is an organization of professionals to "foster
   the technological advancement, availability and implementation of
   a universal emergency telephone number system in North America".
   It develops emergency calling specifications and procedures.

Nomadic device (user): A nomadic user agent is connected to the
   network temporarily, for relatively short durations, but does not
   move significantly during the emergency call. Examples include a
   laptop using an IEEE 802.11 hotspot or a desk IP phone that is
   moved occasionally from one cubicle to another.

Physical location: A physical location describes where a person or
   device is located in physical space, described by a coordinate
   system. It is distinguished from the network location, described
   by a network address.

Public Safety Answering Point (PSAP): A PSAP is a call center that
   answers emergency calls.

Routing location: The routing location of a device is used for
   routing an emergency call and may not be as precise as the
   dispatch location.

Stationary device: An stationary device is not mobile and is
   connected to the network at a fixed, long-term-stable physical
   location. Examples include home PCs or pay phones.

Uncertainty: Uncertainty is an estimate, expressed in a unit of
   length, indicating the diameter of a circle that contains the
   endpoint with the probability indicated by the confidence value.

3. Overview of How Emergency Calls Are Placed

   An emergency call can be distinguished (Section 5) from any other
   call by a unique service URN [RFC5031] that is placed in the call
   setup signaling when a home or visited emergency dial string is
   detected. Because emergency services are local to specific
   geographic regions, a caller obtains his location (Section 6) prior
   to making emergency calls. To get this location, either a form of
   measuring, for example, GNSS (Section 6.2.3) is deployed or the
   endpoint is configured (Section 6.5) with its location from the
   access network's Location Information Server (LIS) using a Location
   Configuration Protocol (LCP). The location is conveyed (Section 6.7)
   in the SIP signaling with the call. The call is routed (Section 8)
   based on location using the Location-to-Service Translation (LoST)
   protocol [RFC5222], which maps a location to a set of PSAP URIs.
   Each URI resolves to a PSAP or an Emergency Services Routing Proxy
   (ESRP) that serves as an incoming proxy for a group of PSAPs. The
   call arrives at the PSAP with the location included in the INVITE
   request.

   The following is a quick overview for a typical Ethernet-connected
   telephone using SIP signaling. It illustrates one set of choices for
   various options presented later in this document.

   o The phone "boots" and connects to its access network.

   o The phone gets location via a Location Configuration Protocol
      (LCP), for example, from the DHCP server in civic [RFC4776] and/or
      geo [RFC6225] forms, a HTTP-Enabled Location Delivery (HELD)
      server [RFC5985] or the first-level switch's Link-Layer Discovery
      Protocol (LLDP) server [LLDP].

   o The phone obtains the local emergency dial string(s) from the LoST
      [RFC5222] server for its current location. It also receives and
      caches the PSAP URI obtained from the LoST server.

   o Some time later, the user places an emergency call. The phone
      recognizes an emergency call from the dial strings and uses the
      "urn:service:sos" [RFC5031] URN to mark an emergency call.

   o It refreshes its location via DHCP and updates the PSAP's URI by
      querying the LoST mapping server with its location.

   o It puts its location in the SIP INVITE request in a Geolocation
      header [RFC6442] and forwards the call using its normal outbound
      call processing, which commonly involves an outbound proxy.

o The proxy recognizes the call as an emergency call and routes the
   call using normal SIP routing mechanisms to the URI specified.

o The call routing commonly traverses an incoming proxy server
   (ESRP) in the emergency services network. That proxy then routes
   the call to the PSAP.

o The call is established with the PSAP and mutually agreed upon
   media streams are created.

o The location of the caller is displayed to the call taker.

```
          Configuration Servers
   . . . . . . . . . . . . . . .
   .                             .
   .   +--------+    +----------+  .
   . +--------+ |  +----------+ |  .
   . | LIS    | |  | SIP      | |  .
   . |        |-+  | Registrar|-+  .
   . +--------+    +----------+    .
   .   ^                ^          .
   . . | . . . . . . | . . . . . .
       |             |
       |[M1][M4]     |[M2]
       |             |       +--------+
       |+-------------+      +--------+ |
       ||                    | LoST   | |
       ||+------------------->| Servers|-+
       |||      [M3][M5]     +--------+        +-------+
       |||                                     | PSAP2 |
       |||                                     +-------+
       |||
       |||   [M6]   +-------+ [M7]+------+ [M8]+-------+
     Alice ------>| Proxy |---->| ESRP |---->| PSAP1 |-----> Call Taker
                  +-------+     +------+     +-------+

                                        +-------+
                                        | PSAP3 |
                                        +-------+
```
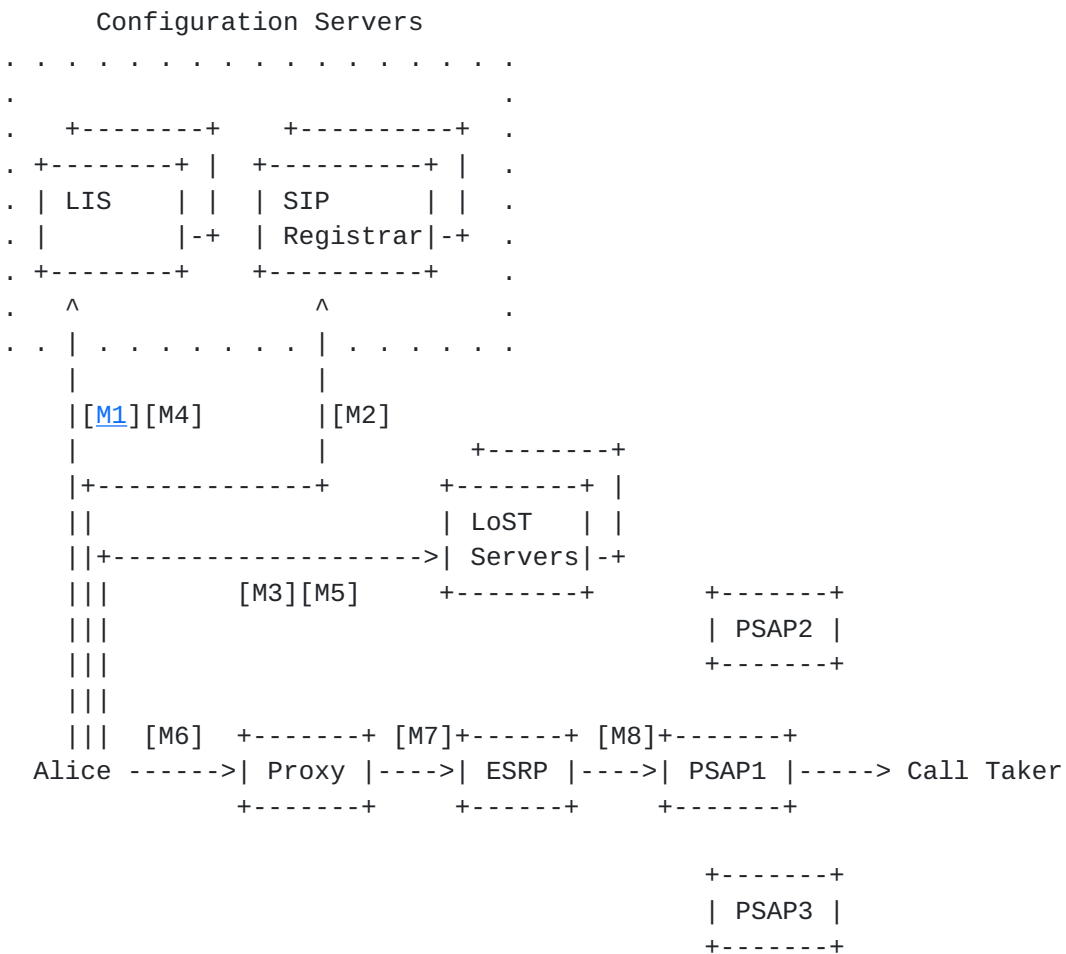
              Figure 1: Emergency Call Component Topology

The typical message flow for this example using Alice as the caller:

```
[M1] Alice -> LIS:  LCP Request(s) (ask for location)
     LIS -> Alice:  LCP Reply(s) (replies with location)
[M2] Alice -> Registrar: SIP REGISTER
     Registrar -> Alice: SIP 200 OK (REGISTER)
[M3] Alice -> LoST Server: Initial LoST Query (contains location)
     Lost Server -> Alice: Initial LoST Response (contains
                        PSAP-URI and dial string)
```

Some time later, Alice dials or otherwise initiates an emergency call:

```
[M4] Alice -> LIS:  LCP Request (updates location)
     LIS -> Alice:  LCP Reply (replies with location)
[M5] Alice -> LoST Server: Update LoST Query (contains location)
     Lost Server -> Alice: LoST Response (contains PSAP-URI)
[M6] Alice -> Outgoing Proxy: SIP INVITE (contains service URN,
                                   Location and PSAP URI)
[M7] Outgoing Proxy -> ESRP: SIP INVITE (contains service URN,
                                   Location and PSAP URI)
[M8] ESRP -> PSAP: SIP INVITE (contains service URN,
                                   Location and PSAP URI)
```

The 200 OK response is propagated back from the PSAP to Alice and the ACK response is propagated from Alice to the PSAP.

Figure 2: Message Flow

Figure 1 shows emergency call component topology and the text above shows call establishment. These include the following components:

o Alice - the user of a UA that places the emergency call.

o Configuration servers - Servers providing Alice's UA its IP
   address and other configuration information, perhaps including
   location-by-value or location-by-reference. Configuration servers
   also may include a SIP registrar for Alice's UA. Most SIP UAs
   will register, so it will be a common scenario for UAs that make
   emergency calls to be registered with such a server in the
   originating calling network. In most cases, a UA would have to
   register in order for the PSAP to be able to call it back after an
   emergency call has been completed. All the configuration messages
   are labeled M1 through M3, but could easily require more than
   three messages to complete.

o LoST server - Processes the LoST request for location plus a
   service URN to a PSAP-URI, either for an initial request from a UA
   or an in-call routing by the proxy server in the originating
   network, or possibly by an ESRP.

o ESRP - Emergency Services Routing Proxy, a SIP proxy server that
   is the incoming call proxy in the emergency services domain. The
   ESRP makes further routing decisions (e.g., based on PSAP state
   and the location of the caller) to choose the actual PSAP that
   handles the call. In some jurisdictions, this may involve another
   LoST query.

o PSAP - Emergency calls are answered at a Public Safety Answering
   Point, a call center.

Generally, Alice's UA either has location configured manually, has an
integral location measurement mechanism, or runs an LCP [M1] to
obtain location from the access (broadband) network. Then, Alice's
UA will most likely register [M2] with a SIP registrar. This allows
her to be contacted by other SIP entities. Next, her UA will perform
an initial LoST query [M3] to learn a URI for use if the LoST query
fails during an emergency call or to use to test the emergency call
mechanism. The LoST response contains the dial string for emergency
calls appropriate for the location provided.

At some time after her device has booted, Alice initiates an
emergency call. She may do this by dialing an emergency dial string
valid for her current ("local") location or for her "home" location.

The UA recognizes either dial string. The UA attempts to refresh its
location [M4], and with that location, to refresh the LoST mapping
[M5], in order to get the most accurate information to use for
routing the call. If the location request or the LoST request fails,
or takes too long, the UA uses values it has cached.

The UA creates a SIP INVITE [M6] request that includes the location.
[RFC6442] defines a SIP Geolocation header that contains either a
location-by-reference URI or a [RFC3986] "cid:" URL indicating where
in the message body the location-by-value is.

The INVITE message is routed to the ESRP [M7], which is the first
inbound proxy for the emergency services domain. This message is
then routed by the ESRP towards the most appropriate PSAP for Alice's
location [M8], as determined by the location and other information.

A proxy in the PSAP chooses an available call taker and extends the
call to its UA.

The 200 OK response to the INVITE request traverses the path in
reverse, from call taker UA to PSAP proxy to ESRP to originating
network proxy to Alice's UA. The ACK request completes the call
setup and the emergency call is established, allowing the PSAP call
taker to talk to Alice about Alice's emergency.

4. Which Devices and Services Should Support Emergency Calls?

Current PSAPs support voice calls and real-time text calls placed
through PSTN facilities or systems connected to the PSTN. However,
future PSAPs will support Internet connectivity and a wider range of
media types and provide higher functionality. In general, if a user
could reasonably expect to be able to place a call for help with the
device, then the device or service should support emergency calling.
Certainly, any device or service that looks like and works like a
telephone (wired or mobile) should support emergency calling, but
increasingly, users have expectations that other devices and services
should work.

Devices that create media sessions and exchange audio, video, and/or
text and that have the capability to establish sessions to a wide
variety of addresses and communicate over private IP networks or the
Internet should support emergency calls.

Traditionally, enterprise support of emergency calling is provided by
the telephony service provider to the enterprise. In some more
recent systems, the enterprise Private Branch Exchange (PBX) assists
emergency calling by providing more fine-grained location in larger
enterprises. In the future, the enterprise may provide the
connection to emergency services itself, not relying on the telephony
service provider.

5. Identifying an Emergency Call

Using the PSTN, emergency help can often be summoned by dialing a
nationally designated, widely known number, regardless of where the
telephone was purchased. The appropriate number is determined by the
infrastructure to which the telephone is connected. However, this
number differs between localities, even though it is often the same
for a country or region, as it is in many countries in the European
Union. In some countries, there is only one uniform digit sequence
that is used for all types of emergencies. In others, there are
several sequences that are specific to the type of responder needed,
e.g., one for police, another for fire. For end systems, on the
other hand, it is desirable to have a universal identifier,
independent of location, to allow the automated inclusion of location

information and to allow the device and other entities in the call
path to perform appropriate processing within the signaling protocol
in an emergency call setup.

Since no such universal identifier existed, the overall emergency
calling architecture described here defines common emergency call
URNs [RFC5031]. When all emergency services use a single number, the
URN is "urn:service:sos". Users are not expected to "dial" an
emergency URN. Rather, appropriate emergency dial strings are
translated to corresponding service URNs, carried in the Request-URI
of the INVITE request. Such translation is best done by the
endpoint, because, among other reasons, emergency calls convey
location in the signaling but non-emergency calls normally do not.
If the device recognizes the emergency call, it can include location,
if known. A signaling intermediary (proxy server) can also recognize
emergency dial strings if the endpoint fails to do so.

For devices that are mobile or nomadic, an issue arises of whether
the home or visited dial strings should be used. Many users would
prefer that their home dialing sequences work no matter where they
are. However, local laws and regulations may require that the
visited dialing sequence(s) work. Therefore, the visited dial string
must work. Devices may have a way to be configured or learn home
dial strings.

LoST [RFC5222] provides the mechanism for obtaining the dialing
sequences for a given location. LoST servers must return dial
strings for emergency services. If the endpoint does not support the
translation of dial strings to service URNs, the dialing sequence
from the endpoint to its proxy is represented as a dial string
[RFC4967] and the outgoing proxy must recognize the dial string and
translate it to the equivalent service URN. To determine the local
emergency dial string, the proxy needs the location of the endpoint.
This may be difficult in situations where the user can roam or be
nomadic. Endpoint recognition of emergency dial strings is therefore
preferred. If a service provider is unable to guarantee that it can
correctly determine local emergency dial strings, wherever its
subscribers may be, then it is required that the endpoint do the
recognition.

Note: The emergency call practitioners consider it undesirable to
have a single-button emergency call user interface element. These
mechanisms tend to result in a very high rate of false or accidental
emergency calls. In order to minimize this issue, practitioners
recommend that devices should only initiate emergency calls based on
entry of specific emergency call dial strings. Speed dial mechanisms
may effectively create single-button emergency call invocation and
practitioners recommend they not be permitted.

6. **Location and Its Role in an Emergency Call**

   Location is central to the operation of emergency services. Location
   is used for two purposes in emergency call handling: routing of the
   call and dispatch of responders. It is frequently the case that the
   callers reporting an emergency are unable to provide a unique, valid
   location themselves. For this reason, location provided by the
   endpoint or the access network is needed. For practical reasons,
   each PSAP generally handles only calls for a certain geographic area,
   with overload arrangements between PSAPs to handle each others'
   calls. Other calls that reach it by accident must be manually
   re-routed (transferred) to the more appropriate PSAP, increasing call
   handling delay and the chance for errors. The area covered by each
   PSAP differs by jurisdiction, where some countries have only a small
   number of PSAPs, while others decentralize PSAP responsibilities to
   the level of counties or municipalities.

   In most cases, PSAPs cover at least a city or town, but there are
   some areas where PSAP coverage areas follow old telephone rate center
   boundaries and may straddle more than one city. Irregular boundaries
   are common, often due to historical reasons. Routing must be done
   based on actual PSAP service boundaries -- the closest PSAP, or the
   PSAP that serves the nominal city name provided in the location, may
   not be the correct PSAP.

   Accuracy of routing location is a complex subject. Calls must be
   routed quickly, but accurately, and location determination is often a
   time/accuracy trade-off, especially with mobile devices or self-
   measuring mechanisms. If a more accurate routing location is not
   available, it is considered acceptable to base a routing decision on
   an accuracy equal to the area of one sector of a mobile cell site.

   Routing to the most appropriate PSAP is always based on the location
   of the caller, despite the fact that some emergency calls are placed
   on behalf of someone else, and the location of the incident is
   sometimes not the location of the caller. In some cases, there are
   other factors that enter into the choice of the PSAP that gets the
   call, such as time of day, caller media requests, language
   preference, and call load. However, location of the caller is the
   primary input to the routing decision.

   Many mechanisms used to locate a caller have a relatively long "cold
   start" time. To get a location accurate enough for dispatch may take
   as much as 30 seconds. This is too long to wait for emergencies.
   Accordingly, it is common, especially in mobile systems, to use a
   coarse location, for example, the cell site and sector serving the
   call, for call-routing purposes, and then to update the location when

   a more precise value is known prior to dispatch. In this document,
   we use "routing location" and "dispatch location" when the
   distinction matters.

   Accuracy of dispatch location is sometimes determined by local
   regulation, and is constrained by available technology. The actual
   requirement is more stringent than available technology can deliver:
   It is required that a device making an emergency call close to the
   "demising" or separation wall between two apartments in a high-rise
   apartment building report location with sufficient accuracy to
   determine on what side of the wall it is. This implies perhaps a 3
   cm accuracy requirement. As of the date of this memo, assisted GNSS
   uncertainty in mobile phones with 95% confidence cannot be relied
   upon to be less than hundreds of meters. As technology advances, the
   accuracy requirements for location will need to be tightened. Wired
   systems using wire-tracing mechanisms can provide location to a wall
   jack in specific room on a floor in a building, and may even specify
   a cubicle or even smaller resolution. As this discussion
   illustrates, emergency call systems demand the most stringent
   location accuracy available.

   In Internet emergency calling, where the endpoint is located is
   determined using a variety of measurement or wire-tracing methods.
   Endpoints may be configured with their own location by the access
   network. In some circumstances, a proxy server may insert location
   into the signaling on behalf of the endpoint. The location is mapped
   to the URI to send the call to, and the location is conveyed to the
   PSAP (and other elements) in the signaling. The terms
   "determination", "configuration", "mapping", and "conveyance" are
   used for specific aspects of location handling in IETF protocols.
   Likewise, we employ Location Configuration Protocols, Location
   Mapping Protocols, and Location Conveyance Protocols for these
   functions.

   This document provides guidance for generic network configurations
   with respect to location. It is recognized that unique issues may
   exist in some network deployments. The IETF will continue to
   investigate these unique situations and provide further guidance, if
   warranted, in future documents.

6.1. Types of Location Information

   Location can be specified in several ways:

   Civic: Civic location information describes the location of a person
      or object by a street address that corresponds to a building or
      other structure. Civic location may include more fine-grained
      location information such as floor, room, and cubicle. Civic
      information comes in two forms:

         "Jurisdictional" refers to a civic location using actual
         political subdivisions, especially for the community name.

         "Postal" refers to a civic location for mail delivery. The
         name of the post office sometimes does not correspond to the
         community name and a postal address may contain post office
         boxes or street addresses that do not correspond to an actual
         building. Postal addresses are generally unsuitable for
         emergency call dispatch because the post office conventions
         (for community name, for example) do not match those known by
         the responders. The fact that they are unique can sometimes be
         exploited to provide a mapping between a postal address and a
         civic address suitable to which to dispatch a responder. In
         IETF location protocols, there is an element (Postal Community
         Name) that can be included in a location to provide the post
         office name as well as the actual jurisdictional community
         name. There is also an element for a postal code. There is no
         other accommodation for postal addresses in these protocols.

   Geospatial (geo): Geospatial addresses contain longitude, latitude,
      and altitude information based on an understood datum and earth
      shape model (datum). While there have been many datums developed
      over time, most modern systems are using or moving towards the
      WGS84 [WGS84] datum.

   Cell tower/sector: Cell tower/sector is often used for identifying
      the location of a mobile handset, especially for routing of
      emergency calls. Cell tower and sectors identify the cell tower
      and the antenna sector that a mobile device is currently using.
      Traditionally, the tower location is represented as a point chosen
      to be within a certain PSAP service boundary that agrees to take
      calls originating from that tower/sector, and routing decisions
      are made on that point. Cell tower/sector information could also
      be represented as an irregularly shaped polygon of geospatial
      coordinates reflecting the likely geospatial location of the
      mobile device. Whatever representation is used must route
      correctly in the LoST database, where "correct" is determined by
      local PSAP management.

In IETF protocols, both civic and geospatial forms are supported.
The civic forms include both postal and jurisdictional fields. A
cell tower/sector can be represented as a geo point or polygon or
civic location. Other forms of location representation must be
mapped into either a geo or civic value for use in emergency calls.

For emergency call purposes, conversion of location information from
civic to geo or vice versa prior to conveyance is not desirable. The
location should be sent in the form it was determined. Conversion
between geo and civic requires a database. Where PSAPs need to
convert from whatever form they received to another for responder
purposes, they have a suitable database. However, if a conversion is
done before the PSAP's, and the database used is not exactly the same
one the PSAP uses, the double conversion has a high probability of
introducing an error.

6.2. Location Determination

As noted above, location information can be entered by the user or
installer of a device ("manual configuration"), measured by the end
system, be delivered to the end system by some protocol or measured
by a third party, and be inserted into the call signaling.

In some cases, an entity may have multiple sources of location
information, possibly some that are partially contradictory. This is
particularly likely if the location information is determined both by
the end system and a third party. Although self-measured location
(e.g., GNSS) is attractive, location information provided by the
access network could be much more accurate, and more reliable in some
environments such as high-rise buildings in dense urban areas.

The closer an entity is to the source of location, the more likely it
is able to determine which location is most appropriate for a
particular purpose when there is more than one location determination
for a given endpoint. In emergency calling, the PSAP is the least
likely to be able to appropriately choose which location to use when
multiple conflicting locations are presented to it. While all
available locations can be sent towards the PSAP, the order of the
locations should be the sender's best attempt to guide the recipient
of which one(s) to use.

6.2.1. User-Entered Location Information

Location information can be maintained by the end user or the
installer of an endpoint in the endpoint itself, or in a database.

Location information routinely provided by end users is almost always
less reliable than measured or wire database information, as users
may mistype location information or may enter civic address
information that does not correspond to a recognized (i.e., valid,
see Section 6.10) address. Users can forget to change the data when
the location of a device changes.

However, there are always a small number of cases where the automated
mechanisms used by the access network to determine location fail to
accurately reflect the actual location of the endpoint. For example,
the user may deploy his own WAN behind an access network, effectively
removing an endpoint some distance from the access network's notion
of its location. To handle these exceptional cases, there must be
some mechanism provided to manually provision a location for an
endpoint by the user or by the access network on behalf of a user.
The use of the mechanism introduces the possibility of users falsely
declaring themselves to be somewhere they are not. However, this is
generally not a problem in practice. Commonly, if an emergency
caller insists that he is at a location different from what any
automatic location determination system reports he is, responders
will always be sent to the user's self-declared location.

6.2.2. Access Network "Wire Database" Location Information

Location information can be maintained by the access network,
relating some form of identifier for the end subscriber or device to
a location database ("wire database"). In enterprise LANs, wiremap
databases map Ethernet switch ports to building locations. In DSL
installations, the local telephone carrier maintains a mapping of
wire-pairs to subscriber addresses.

Accuracy of location historically has been to a street-address level.
However, this is not sufficient for larger structures. The Presence
Information Data Format (PIDF) Location Object [RFC4119], extended by
[RFC5139] and [RFC5491], permits interior building, floor, and room
and even finer specification of location within a street address.
When possible, interior location should be supported.

The threshold for when interior location is needed is approximately
650 square meters. This value is derived from US fire brigade
recommendations of spacing of alarm pull stations. However, interior
space layout, construction materials, and other factors should be
considered.

Even for IEEE 802.11 wireless access points, wire databases may
provide sufficient location resolution. The location of the access
point as determined by the wiremap may be supplied as the location
for each of the clients of the access point. However, this may not

   be true for larger-scale systems such as IEEE 802.16 (WiMAX) and IEEE
   802.22 that typically have larger cells than those of IEEE 802.11.
   The civic location of an IEEE 802.16 base station may be of little
   use to emergency personnel, since the endpoint could be several
   kilometers away from the base station.

   Wire databases are likely to be the most promising solution for
   residential users where a service provider knows the customer's
   service address. The service provider can then perform address
   validation (see Section 6.10), similar to the current system in some
   jurisdictions.

6.2.3. End System Measured Location Information

   Global Positioning System (GPS) and similar Global Navigation
   Satellite Systems (e.g., GLONAS and Galileo) receivers may be
   embedded directly in the end device. GNSS produces relatively high
   precision location fixes in open-sky conditions, but the technology
   still faces several challenges in terms of performance (time-to-fix
   and time-to-first-fix), as well as obtaining successful location
   fixes within shielded structures, or underground. It also requires
   all devices to be equipped with the appropriate GNSS capability.
   Many mobile devices require using some kind of "assist", that may be
   operated by the access network (A-GPS) or by a government (WAAS). A
   device may be able to use multiple sources of assist data.

   The GNSS satellites are active continuously; thus, location will
   always be available as long as the device can "see" enough
   satellites. However, mobile devices may not be able to afford the
   power levels required to keep the measuring system active. In such
   circumstances, when location is needed, the device has to start up
   the measurement mechanism. Typically, this takes tens of seconds,
   far too long to wait to be able to route an emergency call. For this
   reason, devices that have end system measured location mechanisms but
   need a cold start period lasting more than a couple seconds need
   another way to get a routing location. Typically, this would be a
   location associated with a radio link (cell tower/sector).

6.2.4. Network Measured Location Information

   The access network may locate end devices. Techniques include
   various forms of triangulation. Elements in the network
   infrastructure triangulate end systems based on signal strength,
   angle of arrival or time of arrival. Common mechanisms deployed
   include the following:

o Time Difference Of Arrival - TDOA

o Uplink Time Difference Of Arrival - U-TDOA

o Angle of Arrival - AOA

o Radio Frequency (RF) fingerprinting

o Advanced Forward Link Trilateration - AFLT

o Enhanced Forward Link Trilateration - EFLT

Sometimes multiple mechanisms are combined, for example A-GPS with
AFLT.

6.3. Who Adds Location, Endpoint, or Proxy?

The IETF emergency call architecture prefers endpoints to learn their
location and supply it on the call. Where devices do not support
location, proxy servers may have to add location to emergency calls.
Some calling networks have relationships with all access networks the
device may be connected to, and that may allow the proxy to
accurately determine the location of the endpoint. However, NATs and
other middleboxes often make it impossible to determine a reference
identifier the access network could provide to a LIS to determine the
location of the device. System designers are discouraged from
relying on proxies to add location. The technique may be useful in
some limited circumstances as devices are upgraded to meet the
requirements of this document, or where relationships between access
networks and calling networks are feasible and can be relied upon to
get accurate location.

Proxy insertion of location complicates dial-string recognition. As
noted in Section 6, local dial strings depend on the location of the
caller. If the device does not know its own location, it cannot use
the LoST service to learn the local emergency dial strings. The
calling network must provide another way for the device to learn the
local dial string, and update it when the user moves to a location
where the dial string(s) change, or do the dial-string determination
itself.

6.4. Location and References to Location

Location information may be expressed as the actual civic or
geospatial value but can be transmitted as by value, i.e., wholly
contained within the signaling message, or by reference, i.e., as a
URI pointing to the value residing on a remote node waiting to be
dereferenced.

When location is transmitted by value, the location information is
available to entities in the call path. On the other hand, location
objects can be large and only represent a single snapshot of the
device's location. Location references are small and can be used to
represent a time-varying location, but the added complexity of the
dereference step introduces a risk that location will not be
available to parties that need it if the dereference transaction were
to fail.

6.5. End System Location Configuration

Unless a user agent has access to provisioned or locally measured
location information, it must obtain it from the access network.
There are several Location Configuration Protocols (LCPs) that can be
used for this purpose including DHCP, HELD, and LLDP:

   DHCP can deliver civic [RFC4776] or geospatial [RFC6225]
   information. User agents need to support both formats. Note that
   a user agent can use DHCP, via the DHCP REQUEST or INFORM
   messages, even if it uses other means to acquire its IP address.

   HELD [RFC5985] can deliver a civic or geo location object, by
   value or by reference, via a Layer 7 protocol. The query
   typically uses the IP address of the requester as an identifier
   and returns the location value or reference associated with that
   identifier. HELD is typically carried in HTTP.

   Link-Layer Discovery Protocol [LLDP] with Media Endpoint Device
   (MED) extensions [LLDP-MED] can be used to deliver location
   information directly from the Layer 2 network infrastructure and
   also supports both civic and geo formats identical in format to
   DHCP methods.

Each LCP has limitations in the kinds of networks that can reasonably
support it. For this reason, it is not possible to choose a single
mandatory-to-deploy LCP. For endpoints with common network
connections, such as an Ethernet jack or a WiFi connection, location
determination could easily fail unless every network supported every
protocol, or alternatively, every device supported every protocol.
For this reason, a mandatory-to-implement list of LCPs is established
in [PHONEBCP]. Every endpoint that could be used to place emergency
calls must implement all of the protocols on the list. Every access
network must deploy at least one of them. Since it is the
variability of the networks that prevent a single protocol from being
acceptable, it must be the endpoints that implement all of them, and
to accommodate a wide range of devices, networks must deploy at least
one of them.

Often, network operators and device designers believe that they have
a simpler environment and some other network specific mechanism can
be used to provide location. Unfortunately, it is very rare to
actually be able to limit the range of devices that may be connected
to a network. For example, existing mobile networks are being used
to support routers and LANs behind the WAN connection of a wireless
data network, with Ethernet connected phones connected to that. It
is possible that the access network could support a protocol not on
the list and require every handset in that network to use that
protocol for emergency calls. However, the Ethernet-connected phone
will not be able to acquire location, and the user of the phone is
unlikely to be dissuaded from placing an emergency call on that
phone. The widespread availability of gateways, routers, and other
network-broadening devices means that indirectly connected endpoints
are possible on nearly every network. Network operators and vendors
are cautioned that shortcuts to meeting this requirement are seldom
successful.

Location for non-mobile devices is normally expected to be acquired
at network attachment time and retained by the device. It should be
refreshed when the cached value expires. For example, if DHCP is the
acquisition protocol, refresh of location may occur when the IP
address lease is renewed. At the time of an emergency call, the
location should be refreshed, with the retained location used if the
location acquisition does not immediately return a value. Mobile
devices may determine location at network attachment time and
periodically thereafter as a backup in case location determination at
the time of call does not work. Mobile device location may be
refreshed when a Time-to-Live (TTL) expires or the device moves
beyond some boundaries (as provided by [RFC5222]). Normally, mobile
devices will acquire their location at call time for use in an
emergency call routing. See Section 6.8 for a further discussion on
location updates for dispatch location.

There are many examples of endpoints that are user agent applications
running on a more general purpose device, such as a personal
computer. On some systems, Layer 2 protocols like DHCP and LLDP may
not be directly accessible to applications. It is desirable for an
operating system to have an API that provides the location of the
device for use by any application, especially those supporting
emergency calls.

6.6. When Location Should Be Configured

Devices should get routing location immediately after obtaining local
network configuration information. The presence of NAT and VPN
tunnels (that assign new IP addresses to communications) can obscure
identifiers used by LCPs to determine location, especially for HELD.

In some cases, such as residential NAT devices, the NAT is placed
between the endpoint and the access network demarcation point and
thus the IP address seen by the access network is the right
identifier for location of the residence. However, in many
enterprise environments, VPN tunnels can obscure the actual IP
address. Some VPN mechanisms can be bypassed so that a query to the
LCP can be designated to go through the direct IP path, using the
correct IP address, and not through the tunnel. In other cases, no
bypass is possible, but location can be configured before the VPN is
established. Of course, LCPs that use Layer 2 mechanisms (DHCP
location options and LLDP-MED) are usually immune from such problems
because they do not use the IP address as the identifier for the
device seeking location.

It is desirable that routing location information be periodically
refreshed. A LIS supporting a million subscribers each refreshing
once per day would need to support a query rate of 1,000,000 / (24 *
60 * 60) = 12 queries per second. For networks with mobile devices,
much higher refresh rates could be expected.

It is desirable for routing location information to be requested
immediately before placing an emergency call. However, if there is
any significant delay in getting more recent location, the call
should be placed with the most recent location information the device
has. In mobile handsets, routing is often accomplished with the cell
site and sector of the tower serving the call, because it can take
many seconds to start up the location determination mechanism and
obtain an accurate location.

There is a trade-off between the time it takes to get a routing
location and the accuracy (technically, confidence and uncertainty)
obtained. Routing an emergency call quickly is required. However,
if location can be substantially improved by waiting a short time
(e.g., for some sort of "quick (location) fix"), it is preferable to
wait. Three seconds, the current nominal time for a quick fix, is a
very long time add to post-dial delay. NENA recommends [NENAi3TRD]
that IP-based systems complete calls in two seconds from last dial
press to ring at the PSAP.

6.7. Conveying Location

When an emergency call is placed, the endpoint should include
location in the call signaling. This is referred to as "conveyance"
to distinguish it from "configuration". In SIP, the location
information is conveyed following the procedures in [RFC6442]. Since

the form of the location information obtained by the acquisition
protocol may not be the same as the conveyance protocol uses (PIDF-LO
[RFC4119]), mapping by the endpoint from the LCP form to PIDF may be
required.

6.8. Location Updates

As discussed above, it may take some time for some measurement
mechanisms to get a location accurate enough for dispatch, and a
routing location with less accuracy may be provided to get the call
established quickly. The PSAP needs the dispatch location before it
sends the call to the responder. This requires an update of the
location. In addition, the location of some mobile callers, e.g., in
a vehicle or aircraft, can change significantly during the emergency
call.

A PSAP has no way to request an update of a location provided by
value. If the User Agent Client (UAC) gets new location information,
it must signal the PSAP using a new INVITE or an UPDATE transaction
with a new Geolocation header field to supply the new location.

With the wide variation in determination mechanisms, the PSAP does
not know when accurate location may be available. The preferred
mechanism is that the LIS notifies the PSAP when an accurate location
is available rather than requiring a poll operation from the PSAP to
the LIS. The SIP Presence subscription [RFC3856] provides a suitable
mechanism.

When using a HELD dereference, the PSAP must specify the value
"emergencyDispatch" for the ResponseTime parameter. Since,
typically, the LIS is local relative to the PSAP, the LIS can be
aware of the update requirements of the PSAP.

6.9. Multiple Locations

Getting multiple locations all purported to describe the location of
the caller is confusing to all, and should be avoided. Handling
multiple locations at the point where a PIDF is created is discussed
in [RFC5491]. Conflicting location information is particularly
harmful if different routes (PSAPs) result from LoST queries for the
multiple locations. When they occur anyway, the general guidance is
that the entity earliest in the chain generally has more knowledge
than later elements to make an intelligent decision, especially about
which location will be used for routing. It is permissible to send
multiple locations towards the PSAP, but the element that chooses the
route must select exactly one location to use with LoST.

Guidelines for dealing with multiple locations are also given in
[RFC5222]. If a UA gets multiple locations, it must choose the one
to use for routing, but it may send all of the locations it has in
the signaling. If a proxy is inserting location and has multiple
locations, it must choose exactly one to use for routing and send it
as well as any other locations it has that correspond to this UA.

The UA or proxy should have the ability to understand how and from
whom it learned its location, and should include this information in
the location objects that are sent to the PSAP. That labeling
provides the call taker with information to make decisions upon, as
well as guidance for, what to ask the caller and what to tell the
responders.

Endpoints or proxies may be tempted to send multiple versions of the
same location. For example a database may be used to "geocode" or
"reverse geocode", that is, convert from civic to geo or vice versa.
It is very problematic to use derived locations in emergency calls.
The PSAP and the responders have very accurate databases that they
use to convert most commonly from a reported geo to a civic suitable
for dispatching responders. If one database is used to convert from,
say, civic to geo, and another converts from geo to civic, errors
will often occur where the databases are slightly different. Errors
of even a single house number are serious as it may lead first
responders to the wrong building. Derived locations should be marked
with a "derived" method token [RFC4119]. If an entity gets a
location that has a measured or other original method, and another
with a derived method, it must use the original value for the
emergency call.

6.10. Location Validation

Validation, in this context, means that there is a mapping from the
address to a PSAP and that the PSAP understands how to direct
responders to the location. It is recommended that location be
validated prior to a device placing an actual emergency call; some
jurisdictions require that this be done.

Determining whether an address is valid can be difficult. There are,
for example, many cases of two names for the same street, or two
streets with the same name but different "suffixes" (Avenue, Street,
Circle) in a city. In some countries, the current system provides
validation. For example, in the United States of America, the Master
Street Address Guide (MSAG) records all valid street addresses and is
used to ensure that the service addresses in phone billing records
correspond to valid emergency service street addresses. Validation
is normally only a concern for civic addresses, although there could

   be some determination that a given geo is within at least one PSAP
   service boundary; that is, a "valid" geo is one where there is a
   mapping in the LoST server.

   LoST [RFC5222] includes a location validation function. Validation
   is normally performed when a location is entered into a Location
   Information Server. It should be confirmed periodically, because the
   mapping database undergoes slow change and locations that previously
   validated may eventually fail validation. Endpoints may wish to
   validate locations they receive from the access network, and will
   need to validate manually entered locations. Proxies that insert
   location may wish to validate locations they receive from a LIS.
   When the test functions (Section 15) are invoked, the location used
   should be validated.

   When validation fails, the location given should not be used for an
   emergency call, unless no other valid location is available. Bad
   location is better than no location. If validation is completed when
   location is first loaded into a LIS, any problems can be found and
   fixed before devices could get the bad location. Failure of
   validation arises because an error is made in determining the
   location, although occasionally the LoST database is not up to date
   or has faulty information. In either case, the problem must be
   identified and should be corrected before using the location.

6.11. Default Location

   Occasionally, the access network cannot determine the actual location
   of the caller. In these cases, it must supply a default location.
   The default location should be as accurate as the network can
   determine. For example, in a cable network, a default location for
   each Cable Modem Termination System (CMTS), with a representative
   location for all cable modems served by that CMTS could be provided
   if the network is unable to resolve the subscriber to anything more
   precise than the CMTS. Default locations must be marked as such so
   that the PSAP knows that the location is not accurate.

6.12. Location Format Conversion

   The endpoint is responsible for mapping any form of location it
   receives from an LCP into PIDF-LO form if the LCP did not directly
   return a PIDF-LO.

7. LIS and LoST Discovery

   Endpoints must be able to discover a LIS, if the HELD protocol is
   used and a LoST server. DHCP options are defined for this purpose,
   namely [RFC5986] and [RFC5223].

Until such DHCP records are widely available, it may be necessary for
the service provider to provision a LoST server address in the
device. The endpoint can also do a DNS SRV query to find a LoST
server. In any environment, more than one of these mechanisms may
yield a LoST server, and they may be different. The recommended
priority is DHCP first, provisioned value second, and DNS SRV query
in the SIP domain third.

8. Routing the Call to the PSAP

Emergency calls are routed based on one or more of the following
criteria expressed in the call setup request (INVITE):

Location: Since each PSAP serves a limited geographic region and
   transferring existing calls delays the emergency response, calls
   need to be routed to the most appropriate PSAP. In this
   architecture, emergency call setup requests contain location
   information, expressed in civic or geospatial coordinates, that
   allows such routing.

Type of emergency service: In some jurisdictions, emergency calls
   for specific emergency services such as fire, police, ambulance,
   or mountain rescue are directed to just those emergency-specific
   PSAPs. This mechanism is supported by marking emergency calls
   with the proper service identifier [RFC5031]. Even in single-
   number jurisdictions, not all services are dispatched by PSAPs and
   may need alternate URNs to route calls to the appropriate call
   center.

Media capabilities of caller: In some cases, emergency call centers
   for specific caller media preferences, such as typed text or
   video, are separate from PSAPs serving voice calls. ESRPs are
   expected to be able to provide routing based on media. Also, even
   if media capability does not affect the selection of the PSAP,
   there may be call takers within the PSAP that are specifically
   trained, e.g., in real-time text or sign language communications,
   where routing within the PSAP based on the media offer would be
   provided.

Providing a URL to route emergency calls by location and by type of
service is the primary function LoST [RFC5222] provides. LoST
accepts a query with location (by-value) in either civic or geo form,
plus a service identifier, and returns a URI (or set of URIs) to
which to route the call. Normal SIP [RFC3261] routing functions are
used to resolve the URI to a next-hop destination.

The endpoint can complete the LoST mapping from its location at boot
time, and periodically thereafter. It should attempt to obtain a
"fresh" location, and from that a current mapping when it places an
emergency call. If accessing either its location acquisition or its
mapping functions fail, it should use its cached value. The call
would follow its normal outbound call processing.

Determining when the device leaves the area provided by the LoST
service can tax small mobile devices. For this reason, the LoST
server should return a simple (small number of points) polygon for
geospatial location. This can be a simple enclosing rectangle of the
PSAP service area when the reported point is not near an edge, or a
smaller polygonal edge section when the reported location is near an
edge. Civic location is uncommon for mobile devices, but reporting
that the same mapping is good within a community name, or even a
street, may be very helpful for WiFi connected devices that roam and
obtain civic location from the access point to which they are
connected.

Networks that support devices that do not implement LoST mapping
themselves may need the outbound proxy do the mapping. If the
endpoint recognized the call was an emergency call, provided the
correct service URN and/or included location on the call in a
Geolocation header, a proxy server could easily accomplish the
mapping.

However, if the endpoint did not recognize the call was an emergency
call, and thus did not include location, the proxy's task is more
difficult. It is often difficult for the calling network to
accurately determine the endpoint's location. The endpoint may have
its own location, but would not normally include it on the call
signaling unless it knew it was an emergency call. There is no
mechanism provided in [RFC6442] for a proxy to request the endpoint
supply its location, because that would open the endpoint to an
attack by any proxy on the path to get it to reveal location. The
proxy can attempt to redirect a call to the service URN, which, if
the device recognizes the significance, would include location in the
redirected call from the device. All network elements should detect
emergency calls and supply default location and/or routing if it is
not already present.

The LoST server would normally be provided by the local emergency
authorities, although the access network or calling network might run
its own server using data provided by the emergency authorities.
Some enterprises may have local responders and call centers, and
could operate their own LoST server, providing URIs to in-house
"PSAPs". Local regulations might limit the ability of enterprises to
direct emergency calls to in-house services.

The ESRP, which is a normal SIP proxy server in the signaling path of
the call, may use a variety of PSAP state information, the location
of the caller, and other criteria to route onward the call to the
PSAP. In order for the ESRP to route on media choice, the initial
INVITE request has to supply an SDP offer.

9. Signaling of Emergency Calls

9.1. Use of TLS

Best current practice for SIP user agents [RFC4504] including
handling of audio, video, and real-time text [RFC4103] should be
applied. As discussed above, location is carried in all emergency
calls in the call signaling. Since emergency calls carry privacy-
sensitive information, they are subject to the requirements for
geospatial protocols [RFC3693]. In particular, signaling information
should be carried in Transport Layer Security (TLS), i.e., in 'sips'
mode with a ciphersuite that includes strong encryption, such as AES.
There are exceptions in [RFC3693] for emergency calls. For example,
local policy may dictate that location is sent with an emergency call
even if the user's policy would otherwise prohibit that.
Nevertheless, protection from eavesdropping of location by encryption
should be provided.

It is unacceptable to have an emergency call fail to complete because
a TLS connection was not created for any reason. Thus, the call
should be attempted with TLS, but if the TLS session establishment
fails, the call should be automatically retried without TLS.
[RFC5630] recommends that to achieve this effect, the target specify
a sip URI, but use TLS on the outbound connection. An element that
receives a request over a TLS connection should attempt to create a
TLS connection to the next hop.

In many cases, persistent TLS connections can be maintained between
elements to minimize the time needed to establish them [RFC5626]. In
other circumstances, use of session resumption [RFC5077] is
recommended. IPsec [RFC4301] is an acceptable alternative to TLS
when used with an equivalent crypto suite.

Location may be used for routing by multiple proxy servers on the
path. Confidentiality mechanisms such as Secure/Multipurpose
Internet Mail Extensions (S/MIME) encryption of SIP signaling
[RFC3261] cannot be used because they obscure location. Only hop-by-
hop mechanisms such as TLS should be used. Implementing location
conveyance in SIP mandates inclusion of TLS support.

9.2. **SIP Signaling Requirements for User Agents**

   SIP UAs that recognize local dial strings, insert location, and
   perform emergency call routing will create SIP INVITE messages with
   the service URN in the Request-URI, the LoST-determined URI for the
   PSAP in a Route header, and the location in a Geolocation header.
   The INVITE request must also have appropriate callback identifiers
   (in Contact and From headers). To enable media-sensitive routing,
   the call should include a Session Description Protocol (SDP) offer
   [RFC3264].

   SIP caller preferences [RFC3841] can be used to signal how the PSAP
   should handle the call. For example, a language preference expressed
   in an Accept-Language header may be used as a hint to cause the PSAP
   to route the call to a call taker who speaks the requested language.
   SIP caller preferences may also be used to indicate a need to invoke
   a relay service for communication with people with disabilities in
   the call.

9.3. **SIP Signaling Requirements for Proxy Servers**

   At least one SIP proxy server in the path of an emergency call must
   be able to assist UAs that are unable to provide any of the location-
   based routing steps and recognition of dial strings. A proxy can
   recognize the lack of location awareness by the lack of a Geolocation
   header. It can recognize the lack of dial-string recognition by the
   presence of the local emergency call dial string in the From header
   without the service URN being present. They should obtain the
   location of the endpoint if possible, and use a default location if
   they cannot, inserting it in a Geolocation header. They should query
   LoST with the location and put the resulting URI in a route, with the
   appropriate service URN in the Request-URI. In any event, they are
   also expected to provide information for the caller using SIP
   Identity or P-Asserted-Identity. It is often a regulatory matter
   whether calls normally marked as anonymous are passed as anonymous
   when they are emergency calls. Proxies must conform to the local
   regulation or practice.

10. **Call Backs**

   The call taker must be able to reach the emergency caller if the
   original call is disconnected. In traditional emergency calls,
   wireline and wireless emergency calls include a callback identifier
   for this purpose. There are two kinds of call backs. When a call is
   dropped, or the call taker realizes that some important information
   is needed that it doesn't have, it must call back the device that
   placed the emergency call. The PSAP, or a responder, may need to
   call back the caller much later, and for that purpose, it wants a

normal SIP address of record (AOR). In SIP systems, the caller must
include a Contact header field in an emergency call containing a
globally routable URI, possibly a Globally Routable User Agent URI
(GRUU) [RFC5627]. This identifier would be used to initiate
callbacks immediately by the call taker if, for example, the call is
prematurely dropped. A concern arises with back-to-back user agents
(B2BUAs) that manipulate Contact headers. Such B2BUAs should always
include a Contact header that routes to the same device.

In addition, a callback identifier as an address of record (AoR) must
be included either as the URI in the From header field [RFC3261]
verified by SIP Identity [RFC4474] or as a network-asserted URI
[RFC3325]. If the latter, the PSAP will need to establish a suitable
spec(t) with the proxies that send it emergency calls. This
identifier would be used to initiate a callback at a later time and
may reach the caller, not necessarily on the same device (and at the
same location) as the original emergency call as per normal SIP
rules. It is often a regulatory matter whether calls normally marked
as anonymous are passed as anonymous when they are emergency calls.
Proxies must conform to the local regulation or practice.

11. Mid-Call Behavior

Some PSAPs often include dispatchers, responders, or specialists on a
call. Some responders' dispatchers are not located in the primary
PSAP, the call may have to be transferred to another PSAP. Most
often, this will be an attended transfer, or a bridged transfer.
Therefore, a PSAP may need to a REFER request [RFC3515] a call to a
bridge for conferencing. Devices that normally involve the user in
transfer operations should consider the effect of such interactions
when a stressed user places an emergency call. Requiring user
interface manipulation during such events may not be desirable.
Relay services for communication with people with disabilities may be
included in the call with the bridge. The UA should be prepared to
have the call transferred (usually attended, but possibly blind) per
[RFC5359].

12. Call Termination

It is undesirable for the caller to terminate an emergency call. A
PSAP terminates a call using the normal SIP call termination
procedures, i.e., with a BYE request.

13. Disabling of Features

   Certain features that can be invoked while a normal call is active
   are not permitted when the call is an emergency call. Services such
   as call waiting, call transfer, three-way calling, and hold should be
   disabled.

   Certain features such as call forwarding can interfere with calls
   from a PSAP and should be disabled. There is no way to reliably
   determine a PSAP call back. A UA may be able to determine a PSAP
   call back by examining the domain of incoming calls after placing an
   emergency call and comparing that to the domain of the answering PSAP
   from the emergency call. Any call from the same domain and directed
   to the supplied Contact header or AoR after an emergency call should
   be accepted as a callback from the PSAP if it occurs within a
   reasonable time after an emergency call was placed.

14. Media

   PSAPs should always accept RTP media streams [RFC3550].
   Traditionally, voice has been the only media stream accepted by
   PSAPs. In some countries, text, in the form of Baudot codes or
   similar tone encoded signaling within a voiceband is accepted ("TTY")
   for persons who have hearing disabilities. Using SIP signaling
   includes the capability to negotiate media. Normal SIP offer/answer
   [RFC3264] negotiations should be used to agree on the media streams
   to be used. PSAPs should accept real-time text [RFC4103]. All PSAPs
   should accept G.711 A-law (and mu-law in North America) encoded voice
   as described in [RFC3551]. Newer text forms are rapidly appearing,
   with instant messaging now very common, PSAPs should accept IM with
   at least "pager-mode" MESSAGE request [RFC3428] as well as Message
   Session Relay Protocol [RFC4975]. Video media in emergency calling
   is required to support Video Relay Service (sign language
   interpretation) as well as modern video phones.

   It is desirable for media to be kept secure by the use of Secure RTP
   [RFC3711], using DTLS [RFC5764] for keying.

15. Testing

   Since the emergency calling architecture consists of a number of
   pieces operated by independent entities, it is important to be able
   to test whether an emergency call is likely to succeed without
   actually occupying the human resources at a PSAP. Both signaling and
   media paths need to be tested since NATs and firewalls may allow the
   session setup request to reach the PSAP, while preventing the
   exchange of media.

[PHONEBCP] includes a description of an automated test procedure that
validates routing, signaling, and media path continuity. This test
should be used within some random interval after boot time, and
whenever the device location changes enough that a new PSAP mapping
is returned by the LoST server.

The PSAP needs to be able to control frequency and duration of the
test, and since the process could be abused, it may temporarily or
permanently suspend its operation.

There is a concern associated with testing during a so-called
"avalanche-restart" event where, for example, a large power outage
affects a large number of endpoints, that, when power is restored,
all attempt to reboot and, possibly, test. Devices need to randomize
their initiation of a boot time test to avoid the problem.

16. Security Considerations

Security considerations for emergency calling have been documented in
[RFC5069] and [RFC6280].

This document suggests that security (TLS or IPsec) be used hop-by-
hop on a SIP call to protect location information, identity, and
other privacy-sensitive call data. It also suggests that if the
attempt to create a security association fails, the call be retried
without the security. It is more important to get an emergency call
through than to protect the data; indeed, in many jurisdictions
privacy is explicitly waived when making emergency calls. Placing a
call without security may reveal user information, including
location. The alternative, failing the call if security cannot be
established, is considered unacceptable.

17. Acknowledgments

This document was created from "Emergency Services for Internet
Telephony Systems" (Schulzrinne, 2004) together with sections from
"Emergency Context Routing of Internet Technologies Architecture
Considerations" (Polk, 2006).

Design Team members participating in this document creation include
Martin Dolly, Stu Goldman, Ted Hardie, Marc Linsner, Roger Marshall,
Shida Schubert, Tom Taylor, and Hannes Tschofenig. Further comments
and input were provided by Richard Barnes, Barbara Stark, and James
Winterbottom.

18. Informative References

   [LLDP]   IEEE, "IEEE802.1ab Station and Media Access Control",
                 December 2004.

   [LLDP-MED]   ANSI/TIA, "Link Layer Discovery Protocol - Media
                 Endpoint Discovery", TIA Standard, TIA-1057, April 2006.

   [NENAi3TRD]   NENA, "08-751 v1 - i3 Technical Requirements (Long Term
                 Definition)", 2006.

   [PHONEBCP]   Rosen, B. and J. Polk, "Best Current Practice for
                 Communications Services in support of Emergency
                 Calling", Work in Progress, September 2011.

   [RFC3261]   Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston,
                 A., Peterson, J., Sparks, R., Handley, M., and E.
                 Schooler, "SIP: Session Initiation Protocol", RFC 3261,
                 June 2002.

   [RFC3264]   Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model
                 with Session Description Protocol (SDP)", RFC 3264,
                 June 2002.

   [RFC3325]   Jennings, C., Peterson, J., and M. Watson, "Private
                 Extensions to the Session Initiation Protocol (SIP) for
                 Asserted Identity within Trusted Networks", RFC 3325,
                 November 2002.

   [RFC3428]   Campbell, B., Rosenberg, J., Schulzrinne, H., Huitema,
                 C., and D. Gurle, "Session Initiation Protocol (SIP)
                 Extension for Instant Messaging", RFC 3428,
                 December 2002.

   [RFC3515]   Sparks, R., "The Session Initiation Protocol (SIP) Refer
                 Method", RFC 3515, April 2003.

   [RFC3550]   Schulzrinne, H., Casner, S., Frederick, R., and V.
                 Jacobson, "RTP: A Transport Protocol for Real-Time
                 Applications", STD 64, RFC 3550, July 2003.

   [RFC3551]   Schulzrinne, H. and S. Casner, "RTP Profile for Audio
                 and Video Conferences with Minimal Control", STD 65,
                 RFC 3551, July 2003.

   [RFC3693]   Cuellar, J., Morris, J., Mulligan, D., Peterson, J., and
                 J. Polk, "Geopriv Requirements", RFC 3693,
                 February 2004.

   [RFC3711]   Baugher, M., McGrew, D., Naslund, M., Carrara, E., and
                K. Norrman, "The Secure Real-time Transport Protocol
                (SRTP)", RFC 3711, March 2004.

   [RFC3841]   Rosenberg, J., Schulzrinne, H., and P. Kyzivat, "Caller
                Preferences for the Session Initiation Protocol (SIP)",
                RFC 3841, August 2004.

   [RFC3856]   Rosenberg, J., "A Presence Event Package for the Session
                Initiation Protocol (SIP)", RFC 3856, August 2004.

   [RFC3986]   Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform
                Resource Identifier (URI): Generic Syntax", STD 66,
                RFC 3986, January 2005.

   [RFC4103]   Hellstrom, G. and P. Jones, "RTP Payload for Text
                Conversation", RFC 4103, June 2005.

   [RFC4119]   Peterson, J., "A Presence-based GEOPRIV Location Object
                Format", RFC 4119, December 2005.

   [RFC4190]   Carlberg, K., Brown, I., and C. Beard, "Framework for
                Supporting Emergency Telecommunications Service (ETS) in
                IP Telephony", RFC 4190, November 2005.

   [RFC4301]   Kent, S. and K. Seo, "Security Architecture for the
                Internet Protocol", RFC 4301, December 2005.

   [RFC4474]   Peterson, J. and C. Jennings, "Enhancements for
                Authenticated Identity Management in the Session
                Initiation Protocol (SIP)", RFC 4474, August 2006.

   [RFC4504]   Sinnreich, H., Lass, S., and C. Stredicke, "SIP
                Telephony Device Requirements and Configuration",
                RFC 4504, May 2006.

   [RFC4776]   Schulzrinne, H., "Dynamic Host Configuration Protocol
                (DHCPv4 and DHCPv6) Option for Civic Addresses
                Configuration Information", RFC 4776, November 2006.

   [RFC4967]   Rosen, B., "Dial String Parameter for the Session
                Initiation Protocol Uniform Resource Identifier",
                RFC 4967, July 2007.

   [RFC4975]   Campbell, B., Mahy, R., and C. Jennings, "The Message
                Session Relay Protocol (MSRP)", RFC 4975,
                September 2007.

[RFC5012]   Schulzrinne, H. and R. Marshall, "Requirements for
             Emergency Context Resolution with Internet
             Technologies", [RFC 5012](#), January 2008.

[RFC5031]   Schulzrinne, H., "A Uniform Resource Name (URN) for
             Emergency and Other Well-Known Services", [RFC 5031](#),
             January 2008.

[RFC5069]   Taylor, T., Tschofenig, H., Schulzrinne, H., and M.
             Shanmugam, "Security Threats and Requirements for
             Emergency Call Marking and Mapping", [RFC 5069](#),
             January 2008.

[RFC5077]   Salowey, J., Zhou, H., Eronen, P., and H. Tschofenig,
             "Transport Layer Security (TLS) Session Resumption
             without Server-Side State", [RFC 5077](#), January 2008.

[RFC5139]   Thomson, M. and J. Winterbottom, "Revised Civic Location
             Format for Presence Information Data Format Location
             Object (PIDF-LO)", [RFC 5139](#), February 2008.

[RFC5222]   Hardie, T., Newton, A., Schulzrinne, H., and H.
             Tschofenig, "LoST: A Location-to-Service Translation
             Protocol", [RFC 5222](#), August 2008.

[RFC5223]   Schulzrinne, H., Polk, J., and H. Tschofenig,
             "Discovering Location-to-Service Translation (LoST)
             Servers Using the Dynamic Host Configuration Protocol
             (DHCP)", [RFC 5223](#), August 2008.

[RFC5359]   Johnston, A., Sparks, R., Cunningham, C., Donovan, S.,
             and K. Summers, "Session Initiation Protocol Service
             Examples", [BCP 144](#), [RFC 5359](#), October 2008.

[RFC5491]   Winterbottom, J., Thomson, M., and H. Tschofenig,
             "GEOPRIV Presence Information Data Format Location
             Object (PIDF-LO) Usage Clarification, Considerations,
             and Recommendations", [RFC 5491](#), March 2009.

[RFC5626]   Jennings, C., Mahy, R., and F. Audet, "Managing Client-
             Initiated Connections in the Session Initiation Protocol
             (SIP)", [RFC 5626](#), October 2009.

[RFC5627]   Rosenberg, J., "Obtaining and Using Globally Routable
             User Agent URIs (GRUUs) in the Session Initiation
             Protocol (SIP)", [RFC 5627](#), October 2009.

   [RFC5630]   Audet, F., "The Use of the SIPS URI Scheme in the
                 Session Initiation Protocol (SIP)", [RFC 5630](#),
                 October 2009.

   [RFC5764]   McGrew, D. and E. Rescorla, "Datagram Transport Layer
                 Security (DTLS) Extension to Establish Keys for the
                 Secure Real-time Transport Protocol (SRTP)", [RFC 5764](#),
                 May 2010.

   [RFC5985]   Barnes, M., "HTTP-Enabled Location Delivery (HELD)",
                 [RFC 5985](#), September 2010.

   [RFC5986]   Thomson, M. and J. Winterbottom, "Discovering the Local
                 Location Information Server (LIS)", [RFC 5986](#),
                 September 2010.

   [RFC6225]   Polk, J., Linsner, M., Thomson, M., and B. Aboba,
                 "Dynamic Host Configuration Protocol Options for
                 Coordinate-Based Location Configuration Information",
                 [RFC 6225](#), July 2011.

   [RFC6280]   Barnes, R., Lepinski, M., Cooper, A., Morris, J.,
                 Tschofenig, H., and H. Schulzrinne, "An Architecture for
                 Location and Location Privacy in Internet Applications",
                 [BCP 160](#), [RFC 6280](#), July 2011.

   [RFC6442]   Polk, J., Rosen, B., and J. Peterson, "Location
                 Conveyance for the Session Initiation Protocol",
                 [RFC 6442](#), December 2011.

   [WGS84]   NIMA, "NGA: DoD World Geodetic System 1984, Its
                 Definition and Relationships with Local Geodetic
                 Systems", Technical Report TR8350.2, Third Edition,
                 July 1997.

Authors' Addresses

    Brian Rosen
    NeuStar, Inc.
    470 Conrad Dr
    Mars, PA 16046
    USA

    EMail: br@brianrosen.net

    Henning Schulzrinne
    Columbia University
    Department of Computer Science
    450 Computer Science Building
    New York, NY 10027
    USA

    Phone: +1 212 939 7042
    EMail: hgs@cs.columbia.edu
    URI: http://www.cs.columbia.edu

    James Polk
    Cisco Systems
    3913 Treemont Circle
    Colleyville, Texas 76034
    USA

    Phone: +1-817-271-3552
    EMail: jmpolk@cisco.com

    Andrew Newton
    TranTech/MediaSolv
    4900 Seminary Road
    Alexandria, VA 22311
    USA

    Phone: +1 703 845 0656
    EMail: andy@hxr.us