Internet Engineering Task Force
INTERNET-DRAFT
draft-schulzrinne-sipping-emergency-req-00.ps

SIPPING
Schulzrinne
Columbia U.

February 21, 2003
Expires: August 2003

Requirements for Session Initiation Protocol (SIP)-based Emergency Calls

Status of this Memo

Copyright Notice

Abstract

This document enumerates requirements for emergency calls in VoIP and general Internet multimedia systems. We divide the requirements into "last-mile" and "end-to-end". Last-mile solutions only exchange the emergency call center's circuit-switched access by an IP-based system. The requirements for end-to-end IP-based emergency calling address functional and security issues for determining the correct emergency address, for identifying the appropriate emergency call center and for identifying the caller and its location. While we focus on systems that employ the Session Initiation Protocol (SIP), many of the requirements also apply to other environments, such as those using H.248/Megaco or H.323.

Contents

1 Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [1].

2 Introduction

Users of telephone-like services expect to be able to call for emergency help, such as police, the fire department or an ambulance, regardless of where they are, what (if any) service provider they are using and what kind of device they are using. Unfortunately, the mechanisms for emergency calls that have evolved in the public circuit-switched telephone network (PSTN) are not quite appropriate for evolving IP-based voice and real-time multimedia communications. This document outlines some of the requirements that end systems and network elements such as SIP proxies need to satisfy in order to provide emergency call services that offer at least the same functionality as existing PSTN services, while hopefully making emergency calling more robust, cheaper to implement and multimedia-capable.

In the future, users of other real-time and near real-time services may also expect to be able to summon emergency help. For example, instant messaging (IM) users may want to use such services. IM is particularly helpful for hearing-disabled users [3] and in cases where bandwidth is scarce. For lack of a better term, we will use the term "caller" or "emergency caller" to refer to the person placing an emergency call or sending an emergency IM.

The emergency calls described in this document differ from the emergency telecommunications service (ETS) described in [4]. In ETS, relatively small numbers of emergency workers need to maintain communication even when parts of the infrastructure are destroyed or disabled. Emergency calls, on the other hand, are placed by civilians to call for emergency services such as fire, ambulance and police services. Thus, these two services are complementary.

3 Definitions

Emergency call center (ECC): An emergency call center (ECC) receives emergency calls within a specific geographic area and dispatches emergency services, such as fire, police and rescue services. An ECC may also serve as a backup for another ECC and, in backup mode, dispatch emergency services outside of its normal service region. In the United States and Canada, ECCs are called Public Safety Access Points (PSAPs).

Internet Protocol ECC (IECC): An Internet protocol emergency call center (IECC) is an ECC that uses Internet protocols, such as SIP for call signaling, RTP for media delivery, to receive emergency calls.

Call taker: A call taker is an agent, typically a government employee, at the ECC that accepts calls and may dispatch emergency help. (Sometimes the functions of call taking and dispatching are handled by different groups of people, but these divisions of labor are not generally visible to the outside and thus do not concern us here.)

Basic emergency service: Basic emergency service allows a user to reach an ECC serving its current location, but the ECC may not be able to determine the identity or geographic location of the caller (except by having the call taker ask the caller).

Enhanced emergency service: Enhanced emergency services add the ability to identify the caller identity and/or caller location to basic emergency services. (Sometimes, only the caller location may be known, e.g., from a public access point that is not owned by an individual.)

Last-mile emergency service: In last-mile emergency service, the caller uses the existing PSTN infrastructure to place an emergency call. Only the path from the “selective router” to the ECC uses IP-based communications. The call may well be placed from a VoIP device, but is assumed to enter the PSTN very close to the location of the caller. The use of Internet protocols is invisible to the caller.

End-to-end emergency service: In end-to-end emergency service, the caller and ECC both use Internet protocols end-to-end.

4 Last-Mile Access

In last-mile access, an ECC replaces an analog (CAMA) or digital (ISDN) trunk with packet-based access, typically over one or more high-speed access lines such as DSL or leased lines. The packet-based access terminates in the “selective router” that normally hands off calls to the ECC. Thus, the ECC becomes an EICC, but no larger scale infrastructure changes are required.

> Last-mile access is motivated by cost and call setup considerations. It may be cheaper to use IP-based technology for the access link and ECC-internal communications. Also, many existing (US) PSAPs use analog technology, so-called CAMA trunks, to receive emergency calls. These trunks, originally designed for operator positions, can pulse out the ten or 20-digit (for wireless) caller’s number, but as dialed digits. Thus, they add several seconds of call setup delay. This can be particularly disconcerting since it affects the time until the call taker can pick up the call. IP-based communications, using, for example, SIP as a call signaling protocol, can effectively eliminate this extra caller identification delay. (Additional delays are caused by the often very low speed access to the mapping database that maps caller identity to geographic location.) Finally, since pending calls do not consume access network resources, such systems may be more robust in the face of overload.

M1: Coexistence: Due to the investment required, not all ECCs will convert to IP-based access at the same time. Thus, emergency calls MUST work in a network where some ECCs use existing (analog) technology, some ISDN, others IP. In particular, existing back-up relationships between ECCs must continue to work.

M2: Call setup delay: The call setup delay MUST NOT be no larger than for existing analog trunks and SHOULD be significantly smaller.

M3: Call identification: Signaling from the PSTN switch must be able to convey both ten and 20-digit caller identities (ANI – automatic number identification) used in North America and other digit strings used elsewhere.

M4: Call transfer: Call takers MUST be able to transfer active sessions to other call takers within the same ECC and to other ECCs, even those not using Internet

M5: Conferencing: Occasionally, supervisors, translators or other specialists need to participate in an emergency call. Thus, it MUST be possible to add one or more parties, not necessarily located in the IECC, to any emergency call at any time.

M6: Monitoring and recording: In many jurisdictions, both sides of all emergency calls are automatically recorded as potential legal evidence. Thus, it MUST be possible to record and timestamp all signaling and media from all successful, failed and aborted calls.

M7: Transition to end-to-end: Protocols and architecture SHOULD be chosen so that a last-mile IECC can receive emergency calls placed by IP endpoints without major system changes or hardware upgrades.

M8: Authentication of incoming calls: The IECC MUST be able to ascertain that the calls it receives are indeed originating from the selective router.

M9: Authentication of the IECC: The selective router MUST be able to be assured that the calls it places reach the desired IECC rather than an impostor.

M10: Confidentiality: Call signaling and media streams MUST be protected against unauthorized disclosure to third parties.

M11: Robustness: An IECC SHOULD be able to automatically route all incoming calls to another backup IECC, even if the access link(s) to the primary IECC are inoperative. Any such redirection MUST be authenticated.

M12: Overflow handling: An IECC SHOULD be able to automatically route calls to another IECC if the (expected) waiting time exceeds a configured threshold.

5 End-to-End IP-Based Emergency Calls

End-to-end emergency calls originate on an Internet device, traverse IP networks and terminate on an IP-capable ECC (IECC).

As noted, emergency calls need to be identified as such (Section 5.1) and be routed to the appropriate emergency call center (Section 5.2). The ECC needs to determine who (Section 5.3) placed the call from where (Section 5.4). Emergency calls may not be subject to access restrictions placed on non-emergency calls. Also, some call features may interfere with emergency calls, particularly if triggerd accidentally (Section 6).

5.1 Emergency Address

The emergency address is used by the emergency caller to declare a call to be an emergency call and to guide the call to an ECC. The emergency address could a be “sip”, “sips” or “tel” URI, or some other, yet-to-be-defined URI scheme.

A1: Universal: Each device and all network elements MUST recognize one or more global emergency call identifiers, regardless of the location of the device, the service provider used (if any) or other factors.

> SIP is not specific to one country or service provider and devices are likely to be used across national or service provider boundaries. Since services such as disabling mandatory authentication for emergency calls (S1) requires the cooperation of outbound proxies, the outbound proxy has to be able to recognize the emergency address and be assured that it will be routed as an emergency call. Thus, a simple declaration on a random URI that it is an emergency call will likely lead to fraud and possibly attacks on the network infrastructure. A universal address also makes it possible to create user interface elements that are correctly configured without user intervention. UA features could be made to work without such an identifier, but the user interface would then have to provide an unambiguous way to declare a particular call an emergency call.

A2: Local: Since many countries have already deployed national emergency numbers, such as 911 in North America and 112 in large parts of Europe, UAs, proxies and call routers MUST recognize local emergency numbers. In addition, they SHOULD recognize emergency numbers that are found elsewhere.

> The latter requirement is meant to help travelers that may not know the local emergency number and instinctively dial the number they are used to from home. However, it is unlikely that all systems could be programmed to recognize any emergency number used anywhere as some of these numbers are used for non-emergency purposes, in particular extensions and service numbers.

A3: Recognizable: Emergency calls MUST be recognizable by user agents, proxies and other network elements. To prevent fraud, an address identified as an emergency number for call features or authentication override MUST also cause routing to an ECC.

A5: Minimal configuration: Any local emergency numbers SHOULD be configured automatically, without user intervention.

> A new UA "unofficially imported" into an organization from elsewhere should have the same emergency capabilities as one officially installed.

A6: Secure configuration: Devices SHOULD be assured of the correctness of the local emergency numbers that are automatically configured.

> If we assume a fixed, global emergency service identifier that requires no configuration and only configure local "traditional" emergency numbers, users are not likely to suddenly dial some random number if a rogue configuration server introduces this as an additional emergency number. The ability to override all locally configured emergency number is of more concern.

A7: Testable: A user SHOULD be able to test whether a particular address reaches emergency help, without actually causing emergency help to be dispatched or consuming ECC call taker resources.

5.2 Identifying the Appropriate Emergency Call Center

From the previous section, we take the requirement of a single (or small number of) emergency addresses which are independent of the caller's location. However, since for reasons of robustness, jurisdiction and local knowledge, ECCs only serve a small region, having the call reach the correct ECC is crucial. While an ECC may be able to transfer an errant call, any such transfer is likely to add tens of seconds to call setup latency and is prone to errors. (In the United States, there are about 5,000 PSAPs.)

There appear to be two basic architectures for translating an emergency address into the correct IECC. We refer to these as caller-based and mediated. In *caller-based resolution*, the caller's UA consults a directory and determines the correct IECC based on its location. For *mediated resolution*, a SIP (outbound) proxy or redirect server performs this function. Note that the latter case includes the architecture where the call is effectively routed to a copy of the database, rather than having some non-SIP protocol query the database. (It appears undesirable to have either the UA or every outbound proxy server contain a copy of the location-to-ECC mapping since this table changes frequently.)

The problem is harder than for traditional web or email services. There, the originator knows which entity it wants to reach, identified by the email address or HTTP URL. However, the emergency caller only dialed an emergency address. Depending on the location, any of several ten thousand destinations could be valid. In addition, the caller probably does not care which specific ECC answers the call, but rather that it be an accredited ECC, e.g., one run by the local government authorities. (Many ECCs are run by private entities. For example, universities and corporations with large campuses often have their own emergency response centers.)

I1: Correct IECC: The system MUST reach the correct IECC regardless of the location of the caller. In particular, the location determination should not be fooled by the location of IP telephony gateways or dial-in lines into a corporate LAN (and dispatch emergency help to the gateway or campus, rather than the caller), multi-site LANs and similar arrangements.

I2: Choice of IECCs: The system SHOULD offer the emergency caller a choice as to whether he wants to reach a local private emergency response center, e.g., on a corporate campus, or the government-run emergency call center responsible for his current location.

> This choice is often, but not always, provided today. For example, in some cases, the local campus emergency center is reachable by a different number or 9-911 reaches the external ECC, while 911 reaches campus security.

I3: Assuring IECC identity: The emergency caller SHOULD be able to determine conclusively that he has reached an accredited emergency call center.

> This requirement is meant to address the threat that a rogue, possibly criminal, entity pretends to accept emergency calls.

Implementations SHOULD allow callers to proceed, with appropriate warnings or user confirmations, if the identity of the destination IECC cannot be verified.

> Verification can fail for any number of reasons, such as lack of a common certificate chain, especially when traveling, call forwarding, or the expiration of certificates. Accreditation, e.g., in the case of corporate or university campuses, may not exist.

I4: Traceable resolution: Particularly for mediated resolution, the caller SHOULD be able to definitively and securely determine who provided the resolution answer.

I5: Assuring directory identity: The querier (UA or server) MUST be able to assure that it is querying the intended directory.

I6: Query response integrity: The querier MUST be able to be confident that the query or response has not been tampered with.

I7: Assuring update integrity: Any update mechanism for the directory MUST ensure that only authorized users can change directory information. An audit trail MUST be provided.

I8: Call setup latency: The directory lookup SHOULD add minimal delay to the call setup. Since outbound proxies will likely be asked to resolve the same geographic coordinates repeatedly, a suitable time-limited caching mechanism SHOULD be supported (see also ”Ix”).

I9: Multiple directories: A UA or proxy SHOULD be able to use multiple different directories to resolve the emergency address. We do not assume that a single directory has worldwide or even nationwide coverage.

> This allows competing or regional data sources.

I10: Referral: All directories SHOULD refer out-of-area queries to an appropriate default or region-specific directory.

> This requirement alleviates the potential for misconfigurations to cause calls to fail, particularly for caller-based queries.

I9: Multiple protocols: It MAY be useful if directories support multiple query protocols, such as SIP (for proxying), LDAP, a SOAP-based query and others. A mandatory-to-implement protocol

> It appears likely that the resolution mechanism will be needed by a variety of session protocols and user applications.

I11: Robustness: The resolution mechanism MUST allow to deploy systems that are robust in the face of partial network and directory server failures. Caching MAY be used to mitigate temporary unavailability of directories or network connectivity.

I12: Incrementally deployable: An Internet-based emergency call system MUST be able to deployed incrementally. In the initial stages of deployment, an emergency call may not reach the optimal ECC.

5.3 Identifying the Caller

Enhanced emergency call systems provide the ECC with the identity and location of the caller. In PSTN-based systems, the identity is represented by the number of the terminal the call is placed from. In a SIP-based system, we have two distinct identities, namely the address of the terminal (Contact header field) and the identity (name and/or AOR) of the person using the terminal. Depending on the circumstances, only one of them may be available. For example, from a public terminal (“Internet payphone”), only the Contact address may be useful.

In most jurisdictions, callers do not have a choice as to whether they want to reveal their location or identity; such disclosure is typically mandated by law. Emergency numbers are generally not meant for anonymous tips. [TBD: Are there any exceptions?]

C1: Identity: The system SHOULD allow to identify both the caller’s identity and his or her terminal address.

C2: Privacy override: The end system MUST be able to automatically detect that a call is an emergency call so that it can override any privacy settings that conflict with emergency calling. (Whether this override can be configured by the user or is considered a condition of service is considered a legal matter, not a protocol issue.)

Since emergency calls are often placed by children, by people using somebody else's end system or by people in panic, any configuration should be automated rather than relying on user interaction at the time of the call. Delaying a call until the user discovers that they have to answer some screen prompt or deal with a voice prompt in an unfamiliar language is likely to lead to large call setup delays or call failures. This does not preclude that end systems can allow, on a call-by-call basis, to configure special call parameters.

5.4 Identifying the Caller Location

This section supplements the requirements outlined in [2]. Thus, the requirements enumerated there are not repeated here. In general, we can distinguish two modes of operation: *direct* and *indirect* location provision. In *direct* location provision, the calling end system knows its own location and can convey this location to the ECC. In an indirect system, the caller is identified by a permanent or temporary identifier, which the ECC then uses to map the caller to a current location. (In the current North American enhanced emergency calling system, the landline terminal phone number is mapped to a location using the so-called ALI database. For wireless phones, a temporary identifier is created and then mapped to the location information.)

(This is somewhat similar to terminal-based and network-based location services in wireless "911' services. However, even in direct location provision, the terminal may well acquire the location information from a third party, e.g., a wireless location beacon or a DHCP server.)

L1: Multiple location providers: For indirect locations, ECCs MUST be able to access different location providers. The location provider may be tied to the service provider or may be independent of the service provider.

This requirement avoids that all users have to rely on a single location provider. This requirement is hard to avoid if there are no traditional national application-layer service providers.

L2: Civil and geographic: Where possible, both civil (street address) and geographic (longitude/latitude) information SHOULD be provided.

While geographic information can usually be translated into civil coordinates, some coordinates, such as building numbers and floors, are more easily provided as civil coordinates since they do not require a detailed surveying operation. For direct location determination, it may also be easier for the user to check civil coordinates for correctness.

6 Call Setup and Call Features

S1: Authentication override: In many jurisdictions, emergency calls can be placed by any device, regardless of whether it has subscribed for service. Similarly, outbound proxies and other call filtering elements MUST be able to be configured so that they allow unauthenticated emergency calls.

S2: Mid-call features: The end system MUST be able to recognize an emergency call and allow configuration so that certain call features are not triggered accidentally. For example, it may be inappropriate to transfer the ECC or put it on hold. An end system MAY make it more difficult to disconnect an on-going emergency call or accept other incoming calls while in an emergency call.

Call transfer initiated by the emergency caller is likely only to be a problem if a PSTN gateway or B2BUa is in the call path. It is not clear how much effort should be expended on preventing intentional, as opposed to accidental, disconnection, since callers can typically find physical-layer means to terminate the call.

S3: Testable: Users SHOULD be able to test the ability to place an emergency call without actually invoking an emergency response.

This capability is unfortunately missing from the current PSTN.

7 Security Considerations

Confidentiality, integrity and authentication are core requirements for multiple aspects of emergency calling. Threats exist at the infrastructure and individual call level. Security threats are identified throughout this document.

An adversary could corrupt call information or ECC resolution to cause emergency calls to fail subtly, without the caller necessarily noticing. This can be done on a call-by-call basis or by corrupting elements that perform the resolution, including the directory described in Section 5.2, Internet routing tables or DNS. (Obviously, there are typically other ways to make emergency calls fail completely, an approach phone-wire cutting burglars have practiced for years. However, the ability to spoof an ECC requires physical access to the PSTN cable plant, while this may not be required in the IP case.)

Here, we do not consider attacks on the emergency call infrastructure itself. The techniques for dealing with such attacks are likely to be similar as those for protecting other network infrastructure, although the stakes may well be higher.

8 References

Normative References

[1] S. Bradner, "Key words for use in rfcs to indicate requirement levels," RFC 2119, Internet Engineering Task Force, Mar. 1997.

[2] J. Cuellar, J. Morris, and D. Mulligan, "Geopriv requirements," internet draft, Internet Engineering Task Force, Jan. 2003. Work in progress.

Informative References

[3] N. Charlton, M. Gasson, G. Gybels, M. Spanner, and A. van Wijk, "User requirements for the session initiation protocol (SIP) in support of deaf, hard of hearing and speech-impaired individuals," RFC 3351, Internet Engineering Task Force, Aug. 2002.

[4] H. C. Folts, C. Beard, and K. Carlberg, "Requirements for emergency telecommunication capabilities in the Internet," internet draft, Internet Engineering Task Force, Oct. 2002. Work in progress.

9 Acknowledgments

Your name here.

10 Authors' Addresses

Henning Schulzrinne
Dept. of Computer Science
Columbia University
1214 Amsterdam Avenue
New York, NY 10027
USA
electronic mail: schulzrinne@cs.columbia.edu