Technical Considerations for Next-Generation 911

Comments in the matter of Framework for Next Generation 911 Deployment, PS Docket No. 10-255

Richard Barnes

Alissa Cooper

Hannes Tschofenig

Abstract

The IETF ECRIT and GEOPRIV working groups have developed standards for critical components of the next-generation, IP-based 911 system. This document reviews some of the basic principles of these technologies and responds to several questions in the FCC's recent Notice of Inquiry on the topic of NG911.

Clarification to the Original Filing

This comment is a clarification to the comments "Technical Considerations for Next-Generation 911"'[1]. The authors wish to make the following clarifications regarding the relationship between the comments submitted and the IETF[2] / IAB[3]. (The body of the report is unchanged from the previous filing.)

- The comments were filed by the authors in their role as individual contributors to the Internet standards process. They do not reflect the consensus output of any IETF working group or the IAB.
- The titles noted on the cover page of the report (ECRIT co-chair, GEOPRIV co-chair, IAB) are intended to demonstrate the individual authors' roles and experience within the Internet community, and do not indicate that comments by these authors reflect the consensus of the working groups or the IAB.
- The comments should be understood as a summary of IETF technologies related to NG911 from the perspective of experienced individual technical experts.
- The comments reference several documents that *do* reflect IETF consensus, in particular the references that are indicated as RFCs.

[1] http://fjallfoss.fcc.gov/ecfs/document/view?id=7021031915

[2] Internet Engineering Task Force; consensus procedures in RFC 2026

[3] Internet Architecture Board

About the Authors

Richard Barnes has served as co-chair of the Internet Engineering Task Force (IETF) working group on Geographic Location / Privacy (GEOPRIV) since November 2008, and the Emergency Context Resolution with Internet Technologies (ECRIT) working group since March 2010. He also co-chairs the Emergency Services Workshop meeting series. He is co-author of RFC 6155 as well as several Internet drafts, and the book "VoIP Emergency Calling: Foundations and Practice".

Alissa Cooper has served as co-chair of the GEOPRIV working group within the IETF since March 2009.

Hannes Tschofenig co-chaired the IETF ECRIT working group from 2005 to early 2010. For his work in the area of IP-based emergency services he received the 'Outstanding Vision for 112' award from the European Emergency Number Association (EENA) and later even became the co-chair of the EENA Next Generation 112 Technical Committee. He contributed to the technical specifications developed within the National Emergency Number Association (NENA), is a contributor to the work in the ECRIT as well as the IETF GEOPRIV working group, and co-chairs the Emergency Services Workshop series. In March 2010 he joined the Internet Architecture Board (IAB). He has published several articles on emergency services, most recently "Emergency Services for Internet Multimedia" in the December 2010 edition of the IP Protocol Journal.

Contents

1 Introduction

The ability for users to summon help in an emergency has long been a critical aspect of the telecommunications system, embodied in the United States by the decades-long tradition of 911 as a national, unified emergency number. Although more and more communications are conducted using Internet technologies (i.e., media based on the Internet Protocol, or IP), the 911 system will need to adapt to leverage the Internet as well.

One of the most challenging aspects of the NG911 transition will be the process of enabling the many thousands of Internet applications active today to communicate with 911 authorities. In his addendum to the NOI, Chairman Genachowski noted that with modern phones, "you can pretty much text anyone except a 9-1-1 call center." This comment is especially true for Internet applications: Users of these applications can communicate using voice, text, video, or even 3-D virtual reality, but except in a few very special cases, they cannot establish any sort of communication to a 9-1-1 call center.

In its role as the standards body for the Internet, the IETF has a long history of developing protocols that are central to the NG911 architecture. The ECRIT working group has developed a general architecture for enabling IP applications to discover and connect to emergency services [1]. The GEOPRIV working group has developed protocols that allow IP networks to inform end devices about their geolocation, a critical pre-requisite for emergency calling. The application-specific working groups in the IETF (for example, the SIPCORE working group) have developed extensions to support emergency calling as required.

The FCC recently issued a Notice of Inquiry (NOI) seeking comment on several aspects of this transition [2]. This document is intended to provide information on technical questions raised by the NOI, informed by the experience of the IETF in developing the critical enabling standards for NG911. The focus of our comments will thus be on the technical feasibility and architectural soundness of the various ideas discussed in the NOI, avoiding comment on matters of policy. In the remainder of this document, we first discuss some general technical principles that must undergird the NG911 architecture, then we apply these principles to the specific questions in the

NOI. These comments are largely a reflection of the many years of work of participants in the IETF's ECRIT and GEOPRIV working groups.

2 Terminology

In this document, we use the following terms:

- Network operator: An entity that manages a physical, link-layer, or IP network. For example, a telecommunications provider that leases optical circuits or a cellular ISP.
- Call-routing information: Information that specifies a destination for an emergency call based on information such as the caller's location and the type of emergency service desired.
- Internet application or calling application: A communications service that is carried over the Internet, including both applications running on end hosts and any network-based servers. (In the NOI, the entity providing this service is called an Application Service Provider or ASP.)
- Network-integrated application: An Internet application that relies on the underlying host being connected to the Internet through a particular access network. Usually a service provided by the operator of that network, such as a carrier-provided VoIP service.
- Internet-general application: An Internet application that works the same way regardless of how the underlying host is connected to the Internet.

3 General Context of IP Emergency Calling

As outlined in the ECRIT specifications [1] and elaborated in the NENA NG911 specifications [3], there are three basic steps in the establishment of an emergency call:

1. The calling application gains access to information about the caller's geolocation.
2. The calling application uses this geolocation information to determine the proper destination for the call.
3. The calling application directs the call to the discovered destination.

The main question for NG911 is thus what entities will play each of the four critical roles in these steps: The calling application, the geolocation provider, the location-to-service mapping provider, and the recipient of the call. Of these, there is an obvious answer only for the last.

The question of which entities are calling applications presents several complex challenges. We understand that one of the goals of NG911 is to enable emergency calling from some general Internet applications, such as Skype or Google Voice. These applications differ in a few important ways from the telephony services that can call 911 today. (Here, we understand the "application" to encompass both a program on a user device together with any necessary servers in the network.)

First, because the Internet is global, these applications are typically designed to work in the same way wherever they are used in the world. A VoIP application can place calls to other VoIP users in the same way, regardless of whether it is connected via DSL in the US or a 3G modem in Australia. (Indeed, this uniformity accounts for much of the popularity of VoIP.) There is thus a strong need for the core interfaces for NG911 to be globally consistent, in the sense that a calling application should be able to place a call using the same set of actions regardless of where it is.

Second, as the NOI correctly notes, Internet applications in general do not have any inherent knowledge of the caller's geolocation, or of where emergency calls should be routed. Internet applications can be broadly divided into two classes: "Network-integrated" applications that are limited to a specific network, but can benefit from network- and physical-layer information, and "Internet-general" applications that can be used anywhere, but need specific, interoperable interfaces to acquire information from the network. The former class covers things like ISP-provided VoIP services, while the latter covers "over-the-top" applications such as Skype, Google Talk, and

Facebook. Network-integrated applications will be able to benefit from a network operators' existing stores of location and call-routing information. Internet-general applications, however, will need to acquire geolocation and call-routing information from elsewhere, via standard interfaces.

Location information is clearly the sine qua non for emergency services. Unless a calling application has access to geolocation information, it will be unable to discover the proper destination for an emergency call – or even to tell that an emergency call has been placed, since calling numbers vary across the globe (for example, 112 vs. 911). There is a strong analogy to be made here to MVNOs, which can only place emergency calls to the extent that the underlying licensee has deployed E911 technologies. An Internet calling application can only place emergency calls to the extent that it can access geolocation and call routing information.

Techniques for determining the location of Internet devices fall into two broad categories: Those that exploit the physics of the device's connection to the network and those that do not. The former class covers techniques such as the use of wireless network signals or maps of wired networks; GPS is by far the most common example of the latter class. GPS and other network-external techniques can produce high-precision location information, but they require special hardware in the device (beyond the device's network interface hardware) and there are situations where they cannot provide location (e.g., GPS indoors). Network-based geolocation techniques re-use existing hardware, and typically work well where network-external techniques do not (e.g., WiFi-based positioning within a building).

So NG911 systems will deliver optimal reliability by ensuring that devices have access to network-based location information. Devices that have special hardware for positioning (e.g., phones with GPS chips) will be able to fall back to network-based location resources when they are not able to determine their location using this hardware. Of course, devices without special hardware (for example, laptops without GPS hardware) will have to rely exclusively on network-based positioning.

For Internet-connected callers, the authoritative network-based information about a caller's geolocation is held by the entities that operate the physical networks and IP networks overlaid on them, namely ISPs and their underlying access providers. We will refer to these organizations collectively as "network operators." By aggregating information about which node in a network has been assigned an IP address with information about the physical structure of a network, a network operator can provide location information that is usually very high-precision and high-accuracy, and always high-confidence.

Current Internet location-based services are draw their location information mostly from either generic "IP-geo" databases that provide roughly metro-level accuracy (e.g. MaxMind or Quova), or from services that attempt to reconstruct provider infrastructure based on observations (e.g., Google or Skyhook). Because they are operated by third parties, independent of the infrastructure, these location sources are inherently low-confidence, and typically not of high enough accuracy to ensure that emergency responders can be directed to the scene of the emergency.

While these third party sources could in principle be used to support emergency calling, it would clearly make the system much more robust if IP geolocation information could be provided to calling applications from network operators. (Of course, there are no technical barriers to this information being provided through intermediaries, as long as it is ultimately based on information from the network infrastructure.) The IETF GEOPRIV working group has defined a suite of protocols that allow an ISP to expose geolocation information to subscribers' devices [4][5][6][7].

By the same token, authoritative information about emergency call routing is typically maintained by 9-1-1 authorities, so the NG911 system will work best if these authorities provide routing information to calling applications. (Again, possibly through intermediaries.) The IETF ECRIT working group has defined the Location-to-Service Translation (LoST) protocol explicitly for this purpose [8].

Finally, we urge the Commission to keep in mind that, to paraphrase Gertrude Stein, "a network is a network is a network." The networks that comprise the Internet today take on many different forms, from wired residential broadband networks, to Ethernet-based corporate networks, to 4G wireless networks. As more types of networks are used to carry Internet traffic, the opportunities for innovation and interconnection increase, and a network operators' control of the edge decreases. While only a few years ago, only phones authorized by a telephone company could connect to a cellular network, today a 3G-to-WiFi gateway that can be bought at any consumer electronics store can connect any number of devices to that network, without the network knowing about the devices or vice versa. Likewise, more devices than ever are capable of using more than one different type of connection to access the Internet; some laptops are now capable of accessing the Internet over Ethernet, WiFi, and WiMAX.

To enable emergency calling in such an environment, interoperable, universal standards that apply across all types of IP networks are more important than ever. For the laptop connected over WiFi to a gateway to the 3G network, the laptop should not have to know what type of network the gateway is using – in no small part because there is no practical way for it to detect this information with current technologies. Likewise, it only increases the complexity and decreases the reliability of the NG911 system if a device with multiple interfaces has to execute different procedures depending on which interface it is using to connect to the Internet. Instead, the interfaces and procedures that a calling device uses to make an emergency call need to be uniform across all the different ways that it can access the Internet.

4 NOI Responses

In this section, we discuss the implications of the above context with respect to the particular questions of the NOI, by section of the NOI.

4.1 NG911 Capabilities and Applications (Section IV.A)

Interfaces that are used by calling applications must be standard across all NG911 deployments, and to the extent possible, these interfaces should be consistent across the Internet, regardless of national or operational context. The basic Internet standards for the required protocol interfaces for geolocation and call routing are laid out in the IETF ECRIT and GEOPRIV specifications. (Call signaling between the user device and any application servers need not be standardized; only the interface to deliver calls to the emergency services network.) Network-integrated applications (for example, IMS-based carrier-provided services) may not require all of these interfaces, but Internet-general applications require full, standard interfaces from which they can obtain geolocation and call routing information.

The above paragraph does not imply that the same set of services must be offered by every NG911 deployment. Modern signaling protocols are designed so that the two endpoints can negotiate support for specific types of media, for example, video or real-time text. The critical issue is that the high-level protocols over which services are requested and delivered are standardized (for example, the Session Initiation Protocol (SIP) and the Real-Time Protocol (RTP)). The ECRIT emergency calling framework describes a standard profile of SIP for this purpose, and recommends a baseline set of capabilities.

Due to the digital nature of NG911, which allows translation between protocols, a baseline set of capabilities can actually be very robust with respect to technological change. For example, the contents of many different types of instant message – from SIP to XMPP to social-network messaging – can all be translated to a common protocol (for example, SIP) by an appropriate gateway. Indeed, this is common in deployed systems: The MSN and Yahoo chat services, for example, exchange messages via XMPP. However, there may ultimately be services that are different enough that they cannot be translated in this way. The goal of initial NG911 standardization should be to establish a robust set of baseline service interfaces (for example, voice, video, instant messaging, real-time text), then evolve these interfaces as necessary.

4.2 Primary vs. Secondary Usage of Media Types (Section IV.A.2)

There is not necessarily a distinction to be made between some forms of "text-based messaging" (SMS, IM) and others (email, social-network messages). On many modern devices, all of these can be entered and received by users with roughly equal speed. Live video, while it is a "conversational" medium, may be appropriately classed as secondary, since it may not be available in many circumstances, for example, due to device limitations or network congestion.

The question in Paragraph 40 asks what media types service providers “will” support. In fact, for each media type above, there are multiple providers that already offer service (in some cases, hundreds of providers). So the main question here is what PSAPs should support, which will determine which of these service providers will be able to connect to them.

It may facilitate deployment for all PSAPs to accept some common baseline set of media types (for example, voice, IM, real-time text), with others being optional (for example, video, photos, telemetry). Such requirements would provide a clear baseline for application vendors to target, while allowing innovation and competition around advanced services.

The question in Paragraph 40 about privacy is covered by our response to Section IV.D.4 of the NoI below. There is a fair amount of variation in the charging models to which current Internet applications are adapted. Some applications accommodate differential charging depending on such attributes as the source and destination of the call, the prime example being VoIP services that are interconnected with the PSTN. Other applications, rely on the standard tariffing model for Internet access, in which each edge subscriber pays an ISP a set fee for general access to the Internet, which allows the application to send calls anywhere in the world with no additional cost to the user. This latter class accounts for many VoIP applications that are provided to users free of charge.

In considering charging and tariffing models for NG911, the Commission should consider how these models compare to typical tariffing models for Internet applications. Differences in cost models could either encourage or discourage ASPs to participate in the NG911 system. A system in which implementing NG911 in VoIP products causes a net reduction in interconnection costs would create an incentive for ASPs to provide NG911 services. However, a system that imposes new interconnection costs for ASPs that have none today (the latter class above) would discourage those ASPs from enabling NG911 in their products.

4.3 SMS for Emergency Communication (Section IV.A.3)

It should be noted that in the context of NG911, support for SMS and other forms of carrier-based text messaging is a legacy support issue, since these are not an Internet services. SMS and other forms of legacy text messaging do not use Internet protocols and are constrained in various ways (see also [27]).

In terms of end-user experience, SMS is not significantly different from other instant-messaging systems that lack guaranteed delivery (e.g. , SIP MESSAGE over UDP). Indeed, many modern SMS user interfaces group messages into "sessions" in order to more closely resemble IM interfaces. Given this similarity and the overall trend in the industry to transition to IP-based communication in general, there may be benefits to directly implementing IP-based text messaging to PSAPs rather than trying to support legacy technologies like SMS. For a transition period, carriers that do not offer even basic Internet access to enable IP-based messaging could employ gateways to translate SMS messages into appropriate IP-based protocols.

It should also be noted that some new forms of emergency services communications, like instant messaging, may require end user training to create awareness about proper usage. For example, when instant messaging support was launched in some provinces in Spain, users were required to register. This registration interaction was used as an opportunity to educate users, inform them about liability issues, and discourage misuse.

As noted in the NOI, the primary challenge is the correlation of multiple SMS messages into a logical "conversation" so that they can all be delivered to the same call taker. The document cited in reference 70 of the NOI [9] was presented to the ECRIT working group in March 2010; there was not consensus in the group at that time to make an Internet standard on this topic. Part of this decision was driven by the fact that SMS is not an Internet service, so the work of converting SMS to an IP-based messaging system might be better done in other standards organization. If there remains a gap in current standards, however, the issue could be raised again in ECRIT in light of stronger requirements.

4.4 Transport Mechanisms in an NG911 Environment (Section IV.B.1)

NG911 systems are fundamentally based on IP, and should use Internet technologies to the greatest extent possible. Encouraging consolidation on a single network architecture (IP vs. a mixture of IP and legacy) will dramatically reduce the complexity of the overall NG911 system, increasing the reliability and scalability of the system overall.

NG911 systems should also be based on open standards wherever possible, to reduce cost and increase robustness for the overall system. The ability of the Internet to maintain high levels of stability and reliability over decades of operation is based in large part to the fact that it is entirely based on open standards. This openness allows a broad community to obtain the technical expertise necessary to contribute to the Internet, and encourages open discussion of operational issues, to ensure that problems in the network are observed and corrected. The ultimate impact of open standards is thus to make the overall system more reliable and scalable.

Indeed, one of the critical challenges for NG911 deployment will be encouraging and enabling new players to join the 911 ecosystem. These new players – namely, Internet applications – have grown up in the Internet environment, where open standards are the norm. Building an NG911 system that relies on legacy or proprietary technologies will discourage these applications from taking part in the system, reducing the effectiveness of the NG911 transition.

We acknowledge that there is a significant deployed base of legacy systems. These systems are best used in support of IP-based NG911 technologies rather than as first-class parts of any architecture. Some systems can be re-used directly. For example, location systems developed for E911 in cellular networks can be re-used as Location Information Servers in an NG911 context simply by adding the appropriate Internet-standard interface [4][10]. Other systems, for example circuit-based emergency call routing systems, may not be useful at all in the context of NG911. Legacy PSAPs and calling applications should be supported at the edge of the network, according to the NENA standards for Legacy PSAP Gateways and Legacy Network Gateways.

Finally, any new infrastructure deployed in support of NG911 must be built to be part of the Internet of the future. It is therefore a requirement that any new infrastructure support version 6 of the Internet Protocol (IPv6). While commercial networks are still in the process of enabling support for IPv6, a global transition is in progress, so any infrastructure that supports only IPv4 is already becoming obsolete. This focus on IPv6 is consistent with the US CIO's requirements that the government communications infrastructure transition to IPv6 over the course of the next few years [11].

4.5 NG911 Participants (Section IV.B.2)

The technical barriers to the implementation of NG911 on consumer devices are fairly low. Most of the difficult processing tasks (for example, HTTP queries, XML parsing) can be handled via APIs provided by operating systems. Ecritdroid, an open-source program that implements the majority of the ECRIT emergency calling system for the Android mobile phone operating system required only around 900 lines of Java source code. [12]. Thus, it is within the realm of technical possibility for NG911 to be enabled on many classes of devices with suitable user interfaces, including both general-purposes devices such as desktop PCs and mobile phones as well as more special-purpose devices such as gaming consoles.

A device's ability to make emergency calls is always conditional on its ability to gain access to information about its geographical location. This may vary depending on its physical situation (if it is using GPS) or on support by the underlying network for GEOPRIV technologies. In certain devices, access may also be contingent on the presence of certain software applications. In the example above, an Android phone with the Ecritdroid app is capable of NG911, but the generic Android device (without the app) is not.

Nonetheless, it can be empirically verified that a device or application service is capable of placing an emergency call at any given time. The ECRIT architecture includes an explicit mechanism for testing emergency calls (see Section 15 of [13]), in which all call processing is done in the same way as for an emergency call, but the call is delivered to an automatic response system instead of to a call taker. If this mechanism is supported in NG911, then devices will be able to tell dynamically whether they are capable of emergency calling in a given situation.

With regard to the questions in Paragraph 53 about the expansion of 911 requirements to hot spot providers and other additional participants: The ability of a calling application to place NG911 calls will always be predicated on the ability of the application to obtain geolocation and call-routing information, and thus ultimately on the ability of the underlying network to provide geolocation information. So it is not feasible for a calling application to ensure unilaterally that it can make emergency calls in all cases and on all networks.

The technical feasibility of non-traditional networks providing geolocation will vary significantly depending on the type of network. In some cases, it will be possible for geolocation functions can be outsourced to a third party, such as the upstream ISP that provides connectivity for the non-traditional network. Such outsourcing does require some reconfiguration of the local (non-traditional) network, and many deployed non-traditional networks use simple consumer network devices that lack support for these configuration changes and are difficult to upgrade. So while it may be technically possible for many non-traditional networks to support emergency calling functions (i.e., to provide geolocation), there will be an extended period of transition.

When non-traditional networks do not support geolocation technologies, connected devices will not be able to use NG911 services unless they have access to alternative location mechanisms. Devices with special location hardware (e.g., GPS) can use that hardware, provided that is enabled and functional; other devices would have to rely on lower-quality sources of location information such as the third-party services mentioned in the General Principles section above.

4.6 Interoperability and Standards (Section IV.B.3)

The Internet is a global network that works in the same way everywhere, regardless of national boundaries. Internet applications are likewise global. Indeed, most Internet applications have no concept of the geographical location of a user. They are thus incapable for the most part of tailoring their behavior based on a user's current location. The interfaces that PSAPs and networks present to user devices need to be consistent across the Internet (and thus across the globe). The IETF is the standards body for the Internet, and the IETF ECRIT working group has developed the core protocols that comprise NG911. Other standards groups, such as 3GPP, ATIS, and NENA have created more specific architectures based on these protocols. Existing standards should be re-used to the greatest extent possible. Should further standards development related to NG911 be necessary, it should be done in coordination with the IETF.

Within that constraint, however, there is room for national variation in the internals of the NG911 system. These variations can be accommodated in cases where interfaces are purely among local entities (for example, between 911 authorities and ISPs) and where the underlying protocols are indifferent to specific choices (for example, image formats to be carried in a SIP or HTTP transaction). NENA has had a strong track record of solid technical standards within these classes, and good coordination with the IETF on issues where there is more general impact.

4.7 PSAP Functions in an NG911 Environment (Section IV.B.4)

Technologies for distributed PSAPs are already in place, and are in full production use in several emergency calling deployments in Europe. For example, in 2004, the Niederösterreich province of Austria united 84 previously disconnected areas into a single virtual PSAP. In this implementation and others like it, the virtual PSAP receives calls from the PSTN, translates them to SIP, and distributes them to a human call taker station over IP. In an NG911 context, the only major change to this emergency services network would be to remove the translation step and allow calls to arrive directly over SIP. (Of course, on the caller-facing side, there is still a need to provide geolocation and call-routing information to calling applications.)

While none of the three infrastructure components proposed in Paragraph 57 (a LoST "forest guide", public-key cryptography certificates, and a national emergency network) are technically necessary, they could significantly reduce the complexity of implementation and transition of NG911. Without a national forest guide (1), there would be a need for state or regional forest guides and for these entities to share information with each other in order to facilitate national roaming. Without a national certification authority (2), there would be a need for PSAPs and holders of sensitive information to negotiate trust relationships more locally, adding complexity and increasing the risk of unintended authentication failures. A national emergency network (3) would be of the least technical utility, since all IP networks will carry call-related traffic in more or less the same way. A dedicated network, though, could help reduce latency between PSAPs. If connectivity to this network were properly controlled, it could act as a secure enclave for emergency-related services, allowing easier data sharing among emergency response entities.

4.8 Other Specialized NG911 Applications (Section IV.C)

In the context of Internet calling and NG911, there is not a significant technical difference between human-initiated calls and device initiated calls – both are ultimately mediated by end devices. The main difference is that for device-initiated calls, the media transferred will originate from sensors rather than from a human, and there will be less of a need for media to flow back from a PSAP to the caller. The IETF ECRIT working group is developing a simplified framework for device-initiated calls, based on the Common Alerting Protocol (CAP) [14].

There is no technical difference between IP calling to 9-1-1 and other N11 numbers. At a technological level, all of these numbers can take advantage of the same geolocation and call routing resources as NG911. The ECRIT architecture anticipates this multiple use by creating an extensible system of identifiers for services, the so-called “service URNs” [15]. The only difference between a 911 call and a call to an N11 number is which service URN is used, an emergency URN from the “urn:service:sos” class, or a non-emergency URN such as those in the ”urn:service:counseling” class.

Paragraph 61 lists a number of auxiliary data types that could be provided to PSAPs: the caller’s medical history, a description of the caller’s residence or business location, building floor plans, information about hazardous materials, and information about building occupants with special needs. In all of these cases, the fundamental challenge is identifying and locating data. Only two types of identifiers are guaranteed to come with an emergency call: The application-layer identity of the caller (for example, a SIP URI or Skype handle) and network-layer identity of the calling device (i.e., its IP address). Neither of these is necessarily useful for obtaining additional data. Application-layer identities do not necessarily relate to anything else outside of the calling application. Network-layer identities do not always uniquely identify the calling device, for example, due to carrier-grade NAT.

It would be inappropriate and ineffective to automatically expect calling applications to supply additional data. Most calling applications simply lack access to any sort of rich information about their users. For cases where a SIP-based calling application does have access to additional data, the IETF ECRIT working group is developing a mechanism that allows the calling application to add a pointer to that information to an emergency call [16].

It could be helpful to establish standard ways for PSAPs to retrieve additional data based on some standard identifiers, especially when these identifiers are used consistently across many calling applications (e.g., IP addresses and SIP URIs). When this interaction is between entities in the same jurisdiction (e.g., ISPs and PSAPs only), there is not necessarily a technical need for these standards to be consistent internationally, since all data and interactions would be local. It would thus be appropriate for such standards to be developed by a national body such as NENA. In cases where information is to be collected over the broader Internet, there is more of a need for global consistency. This could arise, for example, with an ASP or health care provider that could be in another country. It would thus be appropriate to develop standards for these sorts of interactions in a global forum such as the IETF.

With regard to the interaction between disaster recovery and emergency services, as discussed in Paragraph 62: The transition to the use of Internet technologies for 911 will enable the 911 system to benefit from the large body of knowledge that has been developed around making Internet systems robust. To take just one example, even though the Haitian earthquake in 2010 destroyed most of the country's telecommunications infrastructure, Haitian websites under the .ht domain remained reachable because that domain was redundantly distributed across the world [17].

The same techniques can be applied to make the NG911 infrastructure very robust to local failures. For example, since NG911 endpoints have an inherent requirement to dynamically discover geolocation information (for example, via LIS discovery [5]) and call destinations (via LoST [8]), these functions can be dynamically assigned to backup facilities in the case that primary facilities are impaired.

4.9 Confidentiality and Privacy Concerns (Section IV.D.4)

The Commission notes that "the NG911 network may be only one part of a much larger system that will be shared with government, private sector, and other public safety entities" (Paragraph 74), and questions whether the evolution from the single-purpose legacy 911 system to shared-use NG911 infrastructure will create new privacy concerns. The GEOPRIV suite of protocols were designed to help ensure the privacy of location information even as it moves through shared-use networks and hosts. The privacy features built into GEOPRIV protocols allow flexible use of location information for emergency services while proscribing the distribution and use of location information for other purposes.

A central feature of the GEOPRIV architecture is that location information is always bound to privacy rules to ensure that entities that receive location are informed of how they may use it [18][19]. In the simplest case, the rules convey directives about further distribution and retention of location information; for example, a user who passes his location as part of a VoIP call or web request might set rules directing the recipient not to redistribute the location or to retain it for longer than 24 hours. For conveyance of location information that is unrelated to emergency calling, the rules are conveyed explicitly together with the user's location information.

In the case of emergency services, the privacy rules conveyed to a PSAP may be implicit, or a PSAP that receives privacy rules as part of a GEOPRIV location object may be required to ignore the rules in order to respond to an emergency. But because explicit rules can be included by default in most cases, they can be used to govern non-emergency uses of location information that may be included as part of emergency calls. For example, if the introduction of NG911 infrastructure allows a PSAP to share data with other public safety entities for analysis or review, the privacy rules can be used as directives governing those data exchanges. The conveyance of the rules in general effectively creates a presumption of privacy despite the fact that it may not always be possible to honor the rules during emergencies.

Unlike the PSTN, IP-based networks allow advanced and granular privacy functionality to be built into communications services and applications. While the move to NG911 may expand possibilities for data sharing, it should also be viewed as an opportunity to incorporate greater privacy protection for location and other sensitive information. Building mechanisms like GEOPRIV privacy rules into applications and services that support NG911 is one way to seize that opportunity.

4.10 Location Capabilities (Section IV.D.5)

It is a technical requirement for NG911 that network operators provide geolocation information to calling applications through standard protocol interfaces. Location information must be sourced from network operators to ensure the highest possible confidence in its correctness and timeliness. The use of standard interfaces is important because it allows an application to place emergency calls in the same way, no matter how it is connected to the Internet.

Some current network operators have deployed location services that are accessible using standard protocols defined by legacy telecommunications standards groups, such as the ETSI Parlay/X protocols [20] or the Open Mobile Alliance MLP protocol [21]. These protocols, however, are not suitable for NG911 because they are not general to the Internet (they are bound to specific types of access networks). For example, MLP contains several fields that can hold values relevant to cellular networks (for example, identifiers and measurements), but not those relevant to other networks, such as cable or DSL.

Thus, while these legacy services can be very valuable as sources of location information, the interface presented to calling applications must use a protocol that is designed to work across the Internet. The IETF GEOPRIV working group has defined a set of standard interfaces that networks can use to advertise geolocation information services [7][6][5] and that end devices can use to query these services [4]. In contrast to protocols that are predicated on the use of a specific layer-2 network, the GEOPRIV technologies have been designed from the start for use by any type of endpoint (stationary, nomadic, or mobile), in any IP network, including both wired networks (for example, cable, DSL, Ethernet, fibre-optic) and wireless networks (for example, WiFi, WiMAX, 3G, 4G). They are therefore suitable as a basis for a national NG911 system that applies to all Internet-based emergency calls.

It should also be noted that network operators do not necessarily need to provide the most precise geolocation at all times and to any requestor. Location information should always accurately represent the location of the caller, and authorities, such as call takers and first responders, need to have access to the most precise information possible. However, for purposes of call routing, the calling application only needs access to location information that is precise enough to identify the correct destination for an emergency call, in some cases city- or county-level precision. The IETF ECRIT working group has developed detailed criteria for location precision and algorithms for simplifying location delivery [22].

While not strictly necessary, the creation of a certification entity to allow digital signatures of location for emergency calling could be a helpful step in creating a consistent way for PSAPs and other emergency entities to authenticate that a location object has not been modified. It should be noted, however, that signing location is not a universal solution for location-related attacks. For example, it can be difficult for a PSAP to verify that a location object identifies a particular endpoint (due to the separation between the network and the application layer), so even with signed location, it will still be possible in some cases for two entities to “swap” locations. Detailed security considerations related to the use of location information for NG911 are currently ongoing in the ECRIT working group [23].

4.11 Network and Data Security Concerns (Section IV.D.6)

NG911 systems will face two broad classes of security threats: Those that face any IP-based system, and those that are specific to NG911 as an application.

Every system that is connected to the Internet is inherently faced by a suite of possible threats, ranging from viruses to denial-of-service attacks. Fortunately, over many years of operating such systems, the Internet community has developed a large body of experience in mitigating these threats, and a commercial marketplace that offers many solutions. To protect against these general threats, NG911 systems should apply best practices from the IT security industry.

In addition to best practices that are common to enterprise networks (for example, the use of firewalls), NG911 networks should consider best practices from the ISP and content-provider community for ensuring reliability and continuity of service. Central points of failure should be avoided (for example, PSAPs should consider having more than one connection to the network), and critical assets such as DNS servers should be duplicated across multiple physical locations.

The NG911 system itself, as an application running over the Internet, has some security challenges of its own. Each of the participants in the NG911 calling process puts something at risk, including callers and PSAPs, but there are also mitigations to these risks. (For a detailed treatment of security issues, see [24].)

Callers entrust their safety to the proper functioning of the NG911 system. They do not expect to be denied emergency services or to be directed to a false PSAP. The ECRIT and GEOPRIV protocols that undergird the NG911 system include security mechanisms that enable calling applications to authenticate and encrypt NG911 information in order to ensure that this information is safe from tampering and observation by third parties.

By accepting calls over IP, PSAPs place themselves at a much higher risk of false calls, and of calls without associated location or caller identity information, than they were faced with in the PSTN. This is in large part a cost of enabling consumers to have greater access to 911: rather than receiving calls from a fixed set of local carriers, PSAPs may have to accept calls from anywhere in the world, since even local callers can use calling applications developed and hosted in other parts of the world. Even filtering calls based on the originating IP address is risky, since callers that are physically close to the PSAP may be connected to the Internet via encrypted tunnels to foreign networks (such as VPNs). PSAPs should keep these considerations in mind as they set access control policies, and consider "softer" access controls, such as call ranking and ordering, as a complement to "harder" access controls.

The transition to NG911 should make secure communications easier among NG911 entities, including PSAPs, first responders, and related entities such as hospitals. Using Internet technologies for these communications will simplify the process of connecting different organizations and services, making it a process of setting up trust and authorization relationships rather than any sort of physical interconnection. Every IETF protocol is required to have strong security mechanisms [25], so by using standard Internet technologies, NG911 systems will be able to benefit from these mechanisms and the strengthening and refinement they have undergone through years of use in the general Internet.

4.12 Unidentified Caller Access to NG911 (Section IV.D.8)

The problem of enabling access for all callers who can physically connect to a network is somewhat more complicated for NG911 than for PSTN-based emergency calling systems. While PSTN service is provided as an integrated service, so that access to network connectivity and voice service are controlled as a unitary decision, in the Internet, these two concepts are very much independent of one another. Thus, there can arise situations where a user would be permitted to use a voice service, but cannot connect to any available access network, and vice versa, a connected user may not have authorization to place calls through a particular provider.

The latter case, where the user at least has Internet access, is more straightforward to address, since no third-party "voice provider" is technically necessary for an NG911 call; all that is technically required is the proper software on the caller's device. The destination PSAP or emergency network must also have an open-access policy, in the sense that it will accept calls that come from any provider. Indeed, it is generally very risky in an NG911 environment to have a more restrictive policy (i.e., one that requires calls to come from certain origins), since NG911 will make the set of call origins both more global and more dynamic than it is today.

The case where a user is not authorized to access the Internet is much more challenging. Many access controls of this type are applied at the link layer, before the user has even been assigned an IP address. The technologies involved are thus not subject to IETF standardization, requiring solutions to be developed by the individual standards organizations responsible for the different link-layer technologies (for example, IEEE, 3GPP, WiMAX Forum). There are, however, some common technologies (for example, the Extensible Authentication Protocol, EAP) that are used across several link-layer technologies, and could serve as the basis for at least a partial solution.

The IETF ECRIT working group is developing a document that discusses these problems in more detail, and some proposed solutions [26]. The document contains a discussion of "emergency only" credentials, but there is no current consensus on how these credentials should be implemented. As mentioned above, the underlying challenge is that different types of networks use different security mechanisms, which require different types of credentials. So an authority issuing "emergency-only credentials" may have to issue them differently for different network types (for example, 3G, WiMAX, WiFi) and different security mechanisms used by these networks (for example, 802.1X, WEP, and WPA for WiFi alone).

There is a secondary question of what a network should be expected to provide for "emergency only" access. In particular, it is not technically feasible for a network to distinguish emergency call traffic from other traffic. For example, if a device is configured to only send Internet traffic through a corporate VPN, the local network would be unable to tell whether this device is watching online movies or making a video call to a PSAP – to block the former would block the latter as well, negating the benefit of emergency-only access. So the best that can be done is to carefully monitor the use of emergency-only Internet access, apply heuristic rules to detect and follow up on possible abuse.

There is a clear trade-off between requirements for authentication and/or authorization and requirements for callers' access to NG911. If NG911 is to provide access to the full collection of calling applications (i.e., ASPs) that US callers make use of, then requirements for authentication and authorization will need to be very low. The set of ASPs is both very global, since a caller in the US can use an ASP from anywhere in the world, and very dynamic, since the technical and business barriers to establishing a new ASP are very low. It may nonetheless be appropriate for NG911 systems not to trust all ASPs equally (for example, giving preference to calls from known ASPs), although such preferences should clearly be managed with care to avoid denying NG911 access to users of less-trusted ASPs.

Allowing emergency calling by zero-balance customers may be conceptually easy, but there are some subtleties. ASPs may not always be able to recognize which calls are emergency calls, since PSAPs cannot necessarily be distinguished from other call destinations based on the information in call signaling.

As discussed above, technologies are not sufficiently mature at this point to support unauthenticated layer-2 access, but this is in part due to disagreement in standards bodies over the will of regulatory bodies. So while a hard requirement may not be appropriate at this point, an indication that such a capability would be desirable could help orient the standards process. It should also be noted that requiring unauthenticated access to 911 without the proper technologies for attributing calls would likely lead to many hoax calls, due to the inability of law enforcement to identify and prosecute these callers

4.13 International Issues (Section IV.D.9)

If the NG911 systems follow the ECRIT architecture (for example, as profiled in the NENA i3 specification), then international roaming will supported without any further technology. In particular, in order to allow devices from abroad to access NG911 services, ISPs need to provide geolocation over standard interfaces (as specified in the GEOPRIV RFCs discussed in the Location Capabilities section above), and the LoST infrastructure will need to be available to these devices as well. The calling applications used by foreign devices will also need to support the ECRIT architecture.

The main requirement for PSAPs is that they be liberal in the sources from which they receive calls: They should accept calls from international as well as domestic calling applications / ASPs. There is actually a purely domestic reason for having an open-access policy, since as the 911 ecosystem opens up to include both large LECs and small Internet start-ups, it will likely not be possible to keep a master list of US carriers in order to determine whether the origin of a call is domestic or foreign. And of course, as mentioned above, VoIP calls need have no carrier involved at all.

References

[1] Rosen, B., Schulzrinne, H., Polk, J., and A. Newton, "Framework for Emergency Calling using Internet Multimedia," draft-ietf-ecrit-framework-12 (work in progress), October 2010

[2] Federal Communications Commission, "Framework for Next Generation 911 Deployment," PS Docket No. 10-255, Notice of Inquiry, December 2010

[3] NENA, "Functional & Interface Standards for NG9-1-1," December 2007

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

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

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

[7] Polk, J., Schnizlein, J., and M. Linsner, "Dynamic Host Configuration Protocol Option for Coordinate-based Location Configuration Information," RFC 3825, July 2004

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

[9] Kim, J., Song, W., Schulzrinne, H., Boni, P., and M. Armstrong, "Emergency Text Messaging using SIP MESSAGE," draft-kim-ecrit-text-00 (work in progress), November 2009

[10] Winterbottom, J., Thomson, M., Tschofenig, H., and R. Barnes, "Use of Device Identity in HTTP-Enabled Location Delivery (HELD)," draft-ietf-geopriv-held-identity (work in progress), March 2011

[11] Kundra, V., "Transition to IPv6," September 2010, `http://www.cio.gov/Documents/IPv6MemoFINAL.pdf`

[12] Barnes, R., "Ecritdroid," May 2010, `http://ecritdroid.googlecode.com/`

[13] Rosen, B. and J. Polk, "Best Current Practice for Communications Services in support of Emergency Calling," draft-ietf-ecrit-phonebcp-16 (work in progress), October 2010

[14] Rosen, B., Schulzrinne, H., and H. Tschofenig, "Common Alerting Protocol (CAP) based Data-Only Emergency Alerts using the Session Initiation Protocol (SIP)," draft-ietf-ecrit-data-only-ea-01 (work in progress), October 2010

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

[16] Rosen, B. and H. Tschofenig, "Additional Data related to an Emergency Call," draft-ietf-ecrit-additional-data-00 (work in progress), September 2010

[17] Raj Upadhaya, G., ".ht: Recovering DNS from the Quake," March 2010, `http://meetings.apnic.net/29/program/lightning-talks`

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

[19] Barnes, R., Lepinski, M., Cooper, A., Morris, J., Tschofenig, H., and H. Schulzrinne, "An Architecture for Location and Location Privacy

in Internet Applications," draft-ietf-geopriv-arch-03 (work in progress), October 2010

[20] ETSI, "Open Service Access (OSA); Parlay X Web Services; Part 9: Terminal Location," November 2007

[21] Open Mobile Alliance, "Mobile Location Protocol (MLP)," March 2004

[22] Barnes, R. and M. Lepinski, "Using Imprecise Location for Emergency Context Resolution," draft-ietf-ecrit-rough-loc-03 (work in progress), August 2010

[23] Tschofenig, H., Schulzrinne, H., and B. Aboba, "Trustworthy Location Information," draft-ietf-ecrit-trustworthy-location-01 (work in progress), October 2010

[24] Barnes, R. and K. Wolf, "VoIP Emergency Calling: Foundations and Practice," January 2011

[25] Schiller, J., "Strong Security Requirements for Internet Engineering Task Force Standard Protocols," BCP 61, RFC 3365, August 2002

[26] Schulzrinne, H., McCann, S., Bajko, G., Tschofenig, H., and D. Kroeselberg, "Extensions to the Emergency Services Architecture for dealing with Unauthenticated and Unauthorized Devices," draft-ietf-ecrit-unauthenticated-access-01 (work in progress), October 2010

[27] 4G Americas, "Texting to 9-1-1: Examining the Design and Limitations of SMS," http://www.4gamericas.org/documents/SMS%20to%20911%20White%20Paper%20Final%20October%202010.pdf, October 2010