

# 1

## Introduction

In an August 2011 interview, Federal Communications Commission (FCC) Chairman Julius Genachowski said that “It’s hard to imagine that airlines can send text messages if your flight is delayed, but you can’t send a text message to 911 in an emergency.” He continued “The unfortunate truth is that the capability of our emergency-response communications has not kept pace with commercial innovation, has not kept pace with what ordinary people now do every day with communications devices” [1].

Vice Presidents of the European Commission Neelie Kroes and Siim Kallas have decided to work together to ensure every European can access a 112 smart-phone application, in their own language. This announcement was made in February 2012 [2] on the European 112 day when surveys revealed that “74% of Europeans don’t know what emergency number to call when traveling in the European Union.”

In May 2012 Verizon Wireless announced that it will offer text-to-911 service nationwide to the public as the first US wireless carrier [3].

A lot has changed since the first standardization work on IP-based emergency services began more than 10 years ago. Today, the understanding of Internet applications and their potential has improved significantly. This book will help you to understand how we reached the point where we are today and will also show you a complicated ecosystem that still requires lots of deployment and education work.

We start this journey into the emergency services field with a short history, told from the viewpoint of one of the editors of this book (Hannes Tschofenig). We then provide a short introduction to the relevant part of the Internet Protocol that makes this work technically demanding, and conclude with an overview of the core building blocks for emergency services that will be discussed in the rest of the book.

### 1.1 History

In November 2004 Jon Peterson and I (HT) attended the 61st IETF Meeting in Washington, DC, to discuss the formation of a new Working Group in the IETF, in a Birds of a Feather (“BoF”) meeting [4]. The name of the proposed group was “Emergency Context Resolution

Jon: Is there interest in the room to work  
on the problem outlined in the proposed  
Charter?

hum ... seems broad interest!

Nobody opposed to create a Working Group.

**Figure 1.1** IETF 61 ECRIT BOF: Decision to form a Working Group.

with Internet Technologies (ECRIT).” Such BOF meetings require a fair amount of preparation to have a credible proposal for the audience to evaluate. However, the group does not expect a worked-out solution, but rather requires a critical mass of people to work on a topic and a good starting point. The BOF chairs, namely Jon and I, had worked on a Charter text proposal based on the technical material that was created by Brian Rosen and Henning Schulzrinne the year before this BOF. Speakers were also recruited to talk about requirements, the state of the art, and the design considerations. Jon Peterson was the area director of the Transport Area, and the ECRIT BOF fell under his supervision. As such, I could not have had a better co-chair for my first BOF, since Jon was experienced and the community knew him.

In any case, the audience was convinced that the IETF should start their work on IP-based emergency services standardization and the meeting minutes [5] reflect the decision, as shown in Figure 1.1. The Tao of IETF [4] describes the process of humming in the IETF as follows: “Sometimes consensus is determined by ‘humming’ – if you agree with a proposal, you hum when prompted by the chair. Most ‘hum’ questions come in two parts: you hum to the first part if you agree with the proposal, or you hum to the second part if you disagree with the proposal. Newcomers find it quite peculiar, but it works. It is up to the chair to decide when the Working Group has reached rough consensus.”

Not every BOF receives such support; sometimes there is a fierce fight if many oppose the formation of a new Working Group or argue to radically change the Charter.

Support from the community does not instantly lead to the formation of a new Working Group. Instead, the Internet Engineering Steering Group (IESG) reviews the outcome of the BOF [4]. However, this is uncontroversial and quickly finalized. At the 62nd IETF Meeting in Minneapolis, early March, the ECRIT Working Group met for the first time. The IESG had selected Marc Linsner and me to co-chair the group. I had not worked with Marc prior to that time, but we quickly got along very well. This was the starting point for a fair amount of work that followed over the subsequent years in the IETF.

Late 2005 and early 2006, after a number of ECRIT Working Group meetings, Marc and I started wondering why there was suddenly such a huge amount of activity in the standardization on emergency services. We heard about various efforts either starting or already ongoing in other standards developing organizations.

Early July 2006, the 3GPP CT1 and the IETF ECRIT Working Group organized a joint meeting on emergency calls. This meeting was organized by Hannu Hietalahti (3GPP TSG CT chairman at that time), Marc Linsner and me. The discussions were fruitful and we reached a better understanding of the architectural approaches the two organizations had in mind, even though it was clear that our views were far apart.

Although we did not have the complete picture of the entire standardization landscape, we wondered whether it would make sense to arrange a workshop to meet all those people working in other organizations. The goal of such a workshop would be to reach a better understanding

of what everyone else was working on. We hoped that the sharing of information would avoid overlapping and conflicting standards efforts. We were convinced that emergency services would have to interoperate on a global scale in order to avoid problems for those calling for help. Of course, it was not clear whether other organizations would be interested in discussing their work, and various aspects (such as policies regarding Intellectual Property Rights) could easily become meeting show-stoppers. We also knew it would be time-consuming to find our counterparts from other organizations, and to prepare the workshop.

We nevertheless gave it a try and started with the preparation of the first standards development organizations (SDOs) Emergency Services Coordination Workshop [6], which took place in New York hosted by Henning Schulzrinne at Columbia University on 5th and 6th October 2006. The workshop was organized by Marc Linsner (IETF ECRIT), Hannu Hietalahti and Atle Monrad (3GPP TSG CT), Henning Schulzrinne, and me. The agenda was simply a list of speakers from various organizations talking about their standardization efforts.

Our initial workshop was a big success: we managed to get many of the stakeholders to attend the meeting and to talk about their work. We suddenly had a much better idea of what everyone else was working on. It was also a lively workshop: we were interrupted by a fire alarm twice in the workshop building, and it was clear that the participants had very different views about where the standardization efforts should be going and what global interoperability meant for them. Although the architectural disconnections should not have been surprising, I was surprised again and again during the meeting. As one of the meeting organizers, I had to pause the meeting for some time because some of the participants had gotten into a huge argument, and I was worried that this would derail the entire workshop.

Despite the difficulties, we were convinced that further workshops were needed to improve the cooperation between the different standards bodies. After our second workshop in Washington, DC, we organized workshops in Europe and even went to Kuala Lumpur, Malaysia, in November 2009.

As we progressed our cooperation with other organizations, and regularly met with them, we realized that just hosting workshops was not enough. Marc, Henning, and I were active in different Working Groups in the National Emergency Number Association (NENA), where Brian Rosen led the work on NENA i3 – NENA’s vision for the long-term technical architecture. NENA had a large number of members from the emergency services community, and their feedback on and input to the standards process were crucial. Hence, it was obvious to many of the IETF ECRIT Working Group participants that they should also participate in the NENA work process, which was heavily driven by weekly or bi-weekly conference calls.

During our workshop in Brussels, we contacted the European counterpart of NENA, namely the European Emergency Number Association (EENA). EENA had a very different history than NENA and was initially focused only on lobbying for citizen rights. NENA, on the other hand, had been weaker on the lobbying side but was much stronger on the technical and operational side.

The emergency services communities,<sup>1</sup> which we also invited to our workshops, needed venues for discussions as well, and our highly technical discussions were not necessarily all that they were looking for. Many of them had questions about operating and funding

---

<sup>1</sup> In this context, an emergency services community refers to the public authorities in a specific country or part of a country who are responsible for maintaining the emergency services infrastructure, and their vendors and service providers who contribute the equipment and communications infrastructure.

new emergency services technology. We added tutorial tracks to our workshops to bring newcomers up to speed, but, for many, the tutorials were still too technical. We knew that in the long run we would not be able to organize our workshop series for this extended audience of standards professionals and the emergency services community. And that was not even the entire stakeholder community. We had contacted state and federal regulators, disability groups, Internet Service Providers, researchers, emergency services organizations, advisory bodies, technology providers, over-the-top application server providers as well as telecommunication operators.

Consequently, we started to work more closely with EENA and NENA. For example, in April 2011, we organized our emergency services coordination workshop with the EENA emergency services workshop in Budapest: we provided the technical track and EENA staff organized everything else, including the venue.

In February 2009, I received the “Outstanding Vision for 112” award from EENA for my work on IP-based emergency services. Later I was asked to co-chair EENA’s Next Generation 112 Technical Committee. Since I was familiar with NENA’s work style and had the technical background, I worked with Gary Machado and the rest of the EENA Advisory Board on changing the shape of EENA.

The work in ECRIT took much, much longer than we expected: we added new Charter items to address new requirements and new technology; lots of work was also done in the IETF Geopriv Working Group on location and location protocols. When I was elected to the Internet Architecture Board (IAB) in 2010, I had to step down from my Working Group co-chair role to ensure that I had enough time for my IAB duties. Marc continued as a co-chair and was joined by Richard Barnes, who also co-chaired the IETF Geopriv Working Group with Alissa Cooper. In 2012 Richard decided to step down and Roger Marshall, who is a very active NENA member, took over his position.

At the time of writing, there are still various specifications in progress. Deployments are picking up, although in a way that we had never expected. We always thought that the innovation in communication protocols would also push emergency services forward, but instead application service providers were very reluctant to move into the emergency services space, largely because of liability and regulation fears. Instead, new developments in the communication space (e.g., instant messaging, Voice over IP, social networks) failed to lead to improvements in the ability to call for help. This reinforces the observation by FCC Chairman Julius Genachowski stated at the beginning of this chapter. On the side of emergency services authorities, however, we saw many changes. Authorities learned that Internet Protocols and the new off-the-shelf products led to lower capital investment and lower operational expenses. By focusing on future-proof technology, many of the investments were in IP-based systems rather than into legacy technology. IP also provided them with new flexibility that allowed many countries to reduce the number of Public Safety Answering Points (PSAPs),<sup>2</sup> leading to a more efficient emergency services organization.

Over all these years, I have met many people working on emergency services and tried to understand their views to better see the big picture of IP-based emergency services. Working with Henning Schulzrinne over the years, we both thought it would be valuable for others to

---

<sup>2</sup> RFC 5012 [7] defines a Public Safety Answering Point (PSAP) as “a facility where emergency calls are received under the responsibility of a public authority.” In simplified terms, a PSAP is a call center with call-takers waiting for incoming emergency calls from persons in need of help.

see that same big picture without having to spend 10 years or more in this field. We have asked those whom we have met in our workshops and in our standardization work to contribute their views on the emergency services to this book.

## 1.2 Overview

IP-based emergency services is best understood as an extension of existing communication architectures that are used for everyday communication between friends, coworkers, and businesses. Consequently, there are two parts that have evolved independently from each other, namely the communication architectures used by all of us (e.g., various forms of instant messaging tools, VoIP clients, social networks, voice integration into gaming platforms), and the emergency service networks with the call-takers as the other communication partners.

The drivers behind change for these two ecosystems are different. Day-to-day communication platforms are evolving to meet the needs of the users. The Internet has allowed companies and individuals to easily provide new services. Installing and operating a new Session Initiation Protocol (SIP) VoIP server or an Extensible Messaging and Presence Protocol (XMPP) server requires neither a huge time investment for a technically skilled person nor a huge financial investment. In most cases, all that is needed is a hosting provider that allows the software of the preferred communication software to be installed and run. Companies offering their services to millions of users on the Internet of course need a different level of sophistication and resource investment, but, as history has shown, even small startups can deploy exciting services that are accepted and used on a daily basis by many end customers. The speed of innovation is astonishing if we only look at the types of services offered via browser-based or downloadable applications, such as smart-phone applications. While there is lot of easily recognizable development happening in the communication section, these mechanisms have not been designed with emergency services in mind. In 2012, it is not possible in most parts of the world<sup>3</sup> to use instant messaging to start a real-time communication with call-takers working at a PSAP, to establish video calls, to transmit pictures, or to transmit highly accurate location information available in so many of today's smart phones.

There are many reasons for this development. Many communication service providers see emergency services support as a burden. This impression was partially created by the regulatory community, which is fragmented throughout the world, and the desire to aim for the perfect solution. For those who are mandated to provide emergency services support, there is a liability obligation that comes with it that is often very fuzzy. What service provider would want to spend the time and resources to monitor regulatory developments throughout the world (which are unfortunately not synchronized even within regions), and to invest in systems whose designs only suit a subset of the stakeholders? Without the strict obligation to offer emergency services functionality in our communication software, many emergency services authorities do not see the need to accept this new form of communication. Consequently, one would get the impression that no progress had been made by any of the emergency services authorities in upgrading their emergency services network and their PSAPs. This is interestingly enough

---

<sup>3</sup> Although it would not be too far fetched to say that it is not possible at all to communicate with PSAPs using modern communication technology. There are a few exceptions where emergency services authorities showed leadership and are trying to accept Internet multimedia communication, for example as part of pilot projects.

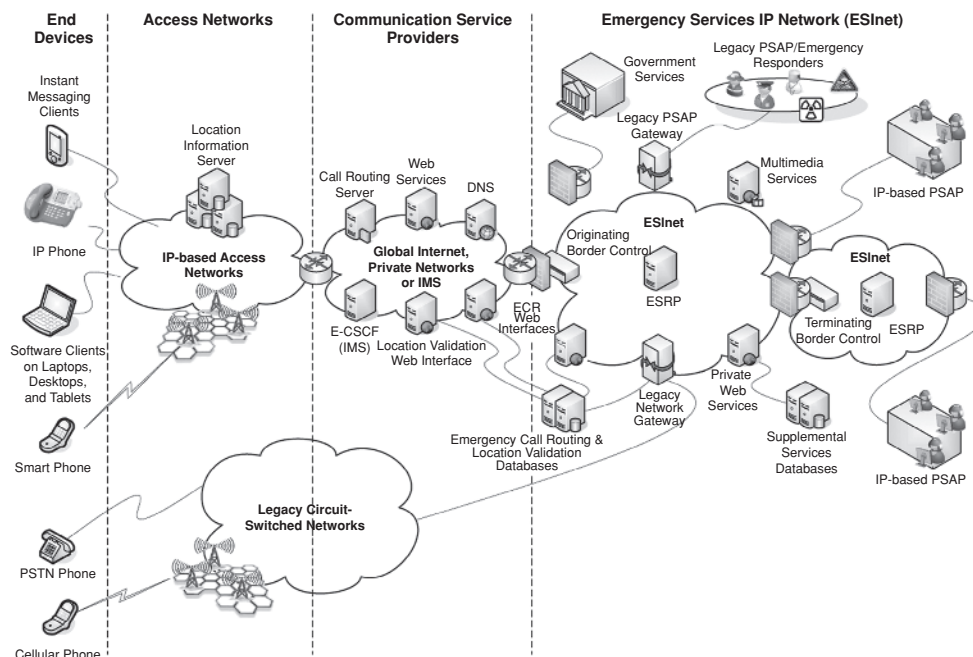
not correct either. In many countries, we can observe a shift to IP-based emergency services network deployments mainly due to four reasons:

1. lower costs due to the commodity nature and increased competition,
2. lower operational costs compared to the legacy (circuit-switched) infrastructure,
3. future-proofing since the rest of the entire industry is moving to IP as well, and
4. more functionality and better extensibility.

From a functionality point of view, this allows an emergency services network to connect PSAPs with each other and to re-route calls. Re-routing can happen for various reasons, including load balancing, better utilizing call-taker skills (e.g., language skills), re-routing incorrectly routed calls, conference bridging capabilities, and so on. Furthermore, multimedia data from service providers as well as from end devices can be received and passed on to other parties to improve the situational awareness.

In some countries, the transition to an all-IP-based emergency services network was also used to re-think the current emergency services organization. As a result, the existing processes and organizational structures were simplified, typically leading to a reduction in the number of PSAPs. This has been a trend in Europe for the last few years.

Figure 1.2 shows the vision of an IP-enabled emergency services architecture graphically. The left side of the picture shows end devices attached to different access networks and communication service providers. In the Internet many applications are provided by companies



**Figure 1.2** Communication architecture overview. (Courtesy of Guy Caron, ENP.)



independent from the access network operator, and therefore the separation between the access network and the communication service provider is indicated explicitly, whereas in the circuit-switched telephony world the application functionality was provided by the access network operator. The right side of Figure 1.2 depicts the IP-based emergency services network. This book covers technical specifications and developments that concern both sides. For communication architectures we focus on those systems that utilize Session Initiation Protocol (SIP), since most emergency services standardization efforts have focused on SIP.

As described in earlier paragraphs, various communication architectures exist today and Figure 1.2 just refers to them as originating networks. Three core features must be provided by such an communication architecture in order to interwork smoothly with an IP-based emergency services network:

1. ability to identify an emergency call;
2. ability to communicate location and/or a location reference; and
3. ability to convey multimedia content.

We will describe these three core building blocks in more detail in section 1.3. Some of the communication architectures in use today provide support for this functionality, such as the SIP-based IP Multimedia Subsystem (IMS) architecture, and the SIP-based VoIP architecture. At the time of writing, there are other communication architectures that do not yet support emergency services, such as the Real-Time Communication in WEB-browsers [8] and the Extensible Messaging and Presence Protocol (XMPP) [9]. There are also many non-standardized and proprietary communication architectures, such as Skype, and many smart-phone applications that do not support emergency services.

You may wonder why there is not just one communication architecture used by everyone. There are probably many reasons, but the history and background of the people who did the work often had a huge influence on the direction of the standardization work. There are also different business models that motivate the work in different standards developing organizations. We will describe one such differentiator in terms of the chosen design assumptions, which will also explain the developments described in other sections of this book.

The Internet architecture follows a layered design, which is a common design pattern for communication systems in general, and allows the replacement of different components (such as a new radio technology, or new applications) while only impacting the neighboring layers. In student textbooks, the Internet architecture has the link layer, Internet layer, transport layer, and the applications layer. In the real world, the layering is much more complex, and it is sometimes hard to assign specific protocols to specific layers (since protocols can be used in a very flexible way and tunneled inside other protocols). The responsibility of providing implementations and deployments of certain layers is also distributed to different parties. In many deployments, the provider of physical connectivity (e.g., a wire) and the provider offering connectivity to the Internet are different companies. Furthermore, those companies offering Internet connectivity are very often different from those offering application layer services. This separation of functionality is a consequence of the end-to-end principle rather than the layering alone. The end-to-end principle states that end-to-end functions can best be realized by end-to-end protocols [10]. While the end-to-end principle still leaves room for discussion and interpretation, the fact is that many application deployments today happen independently of those who provide Internet access. Needless to say, that those who provide

Internet access would like to provide applications as well, or benefit from the deployment of the applications in some way, very much like they did in the past with the Plain Old Telephone Service (POTS). Consequently, a design that assumes that the Internet access provider also offers application services (as is done with the IMS architecture) is different from a design that assumes a separation between the two parties. Such fundamental design assumptions lead to different communication architectures and consequently also to different designs for the emergency services system (even though it is possible to reuse the same building blocks).

The differences between XMPP-based, Skype-based, and over-the-top SIP-based VoIP deployments are, on the other hand, less dramatic. All three assume independence from the Internet access provider, but they are different in the protocol choice. SIP is an IETF standard and is today widely used for voice traffic, Skype software clients use a proprietary protocol that only relies on SIP for interconnecting with other non-Skype-based systems, and XMPP is also an IETF protocol that has found widespread usage for instant messaging.

The term “smart-phone app(lication)” just refers to an implementation of some protocol. It does not indicate whether a standardized or a proprietary protocol has been implemented. As such, a smart-phone app may be an implementation of any of the protocols above, even an SIP-based VoIP client.

## 1.3 Building Blocks

### 1.3.1 Recognizing Emergency Calls

In the early days of Public Switched Telephone Network (PSTN)-based emergency calling, callers would dial a local number for the fire or police department. It was recognized in the 1960s that trying to find this number in an emergency caused unacceptable delays. Thus, most countries have been introducing single nationwide emergency numbers, such as 9-1-1 in North America and 1-1-2 in all European Union countries. This became even more important as mobile devices started to supplant landline phones. As can be seen from the introduction of 1-1-2 in Europe, this education effort takes many years, and the old emergency numbers are therefore still in use today as well (in addition to the European-wide 1-1-2 number).

In many countries, different types of emergency services, such as police or mountain rescue, are still identified by separate numbers. Unfortunately, there are more than 60 different emergency numbers in use worldwide, many of which also have non-emergency uses in other countries, so that simply storing the list of numbers in all devices is not feasible. Furthermore, hotels, university campuses, and larger enterprises often use dial prefixes, so that an emergency caller may have to dial 0-1-1-2 to reach the fire department.

With the introduction of smart phones, new user interface designs emerged as well. For this reason, some devices may use dedicated emergency calling buttons or similar user interface elements to initiate an emergency call. Such mechanisms need to be carefully designed so that they are not accidentally triggered, for example, when the device is in a pocket.

Instead of conveying the actual dial string in a protocol message once the user has entered it, a symbolic representation is used instead. This allows unambiguous emergency call identification and automatic treatment of calls. The mechanism used for this emergency call marking uses the Uniform Resource Names (URNs) defined in RFC 5031 [11], such as urn:service.sos.



### *1.3.2 Obtaining and Conveying Location Information*

Location information is needed by emergency services for three reasons: routing the call to the right PSAP, dispatching first responders (e.g., policemen), and determining the emergency service dial strings that are supported in a specific area. It is clear that the location has to be automatic for the first and third purposes, but experience has shown that automated, highly accurate location information is vital to dispatching as well, rather than relying on the caller to report his or her location to the call-taker. This increases accuracy and avoids dispatch delays when the caller is unable to provide location information due to language barriers, lack of familiarity with his or her surroundings, stress, physical or mental impairment. For this reason, automatic location retrieval for emergency calls is a mandatory requirement in nearly all countries in the world.

Location information for emergency purposes comes in two representations: geospatial (also called geodetic), that is, longitude and latitude; and civic, that is, street addresses similar to postal addresses. Particularly for indoor location, vertical information (floors) is also very useful. Civic locations are most useful for fixed Internet access, including wireless hotspots, and are often preferable for specifying indoor locations; while geodetic location is frequently used for cell phones. However, with the advent of femto- and pico-cells, civic location is both possible and probably preferable since accurate geodetic information can be very hard to acquire indoors.

The requirements for location accuracy differ between routing and dispatch. For call routing, city- or even county-level accuracy is often sufficient, depending on how large the PSAP service areas are. First responders, however, benefit greatly when they can pinpoint the caller to a particular building or, better yet, apartment or office for indoor locations, and an outdoor area of at most a few hundred meters outdoors. This avoids having to search multiple buildings, for example, for medical emergencies.

Various protocol mechanisms have been developed to obtain location information from Location Servers operated by different entities, if the calling device does not yet have location information available. Chapter 2 discusses various location protocols that have been developed to allow the calling device, the PSAP, or some other proxy on behalf of them to request location information.

### *1.3.3 Routing Emergency Calls*

Once an emergency call is recognized, the call needs to be routed to the appropriate PSAP. Each PSAP is only responsible for a limited geographic region, its service region. In addition to the geographical region, different PSAPs often only provide their service for specific services, such as police, fire, ambulance, and so on. There is a wide range of different deployment models used throughout the world, and a description of the different PSAP models for Europe can be found elsewhere [12, 13].

The number of PSAPs serving a country varies quite a bit: Sweden, for example, has 18 PSAPs; and the United States has approximately 6200 (at the time of writing). Therefore, there is roughly one PSAP per 500,000 inhabitants in Sweden, and one per 50,000 in the United States. As all-IP infrastructure is rolled out, smaller PSAPs may be consolidated into regional PSAPs.

Emergency calls are primarily routed based on the emergency caller's location information. Routing may also take place in multiple stages, taking into account factors such as the number of available call-takers, the load situation of a PSAP, or the language capabilities of the call-takers. We expect that dynamic routing will be the predominant mechanism in the near future. In order to perform the initial location-based routing step closer to the PSAP, information about the service boundaries of PSAPs need to be known to end devices or to VoIP Service Provider (VSPs). This information has somehow to be exchanged and shared. Today, Excel sheets are used [14] to exchange phone numbers associated with PSAP service boundaries. A call-taker then has to manually analyze the situation, look up the appropriate phone number, and re-route the call. This may seem a reasonable initial step for the small number of transnational emergency calls that are misrouted in today's emergency services system, but it is not a future-proof approach for a next-generation IP-based emergency services system. The inability of emergency centers to quickly re-route calls due to an overload situation in a mass incident has already been shown in various incidents today.