ATOCA	H. Schulzrinne
Internet-Draft	Columbia University
Intended status: Informational	S. Norreys
Expires: April 27, 2012	BT Group
	B. Rosen
	NeuStar, Inc.
	H. Tschofenig
	Nokia Siemens Networks
	October 25, 2011

Requirements, Terminology and Framework for Exigent Communications
draft-ietf-atoca-requirements-02.txt

Abstract

Before, during and after emergency situations various agencies need to provide information to a group of persons or to the public within a geographical area. While many aspects of such systems are specific to national or local jurisdictions, emergencies span such boundaries and notifications need to reach visitors from other jurisdictions.

This document provides terminology, requirements and an architectural description for protocols exchanging alerts between IP-based end points.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

1.1. Classical Early Warning Situations

During large-scale emergencies, public safety authorities need to reliably communicate with citizens in the affected areas, to provide warnings, indicate whether citizens should evacuate and how, and to

dispel misinformation. Accurate information can reduce the impact of such emergencies.

Traditionally, emergency alerting has used church bells, sirens, loudspeakers, radio and television to warn citizens and to provide information. However, techniques, such as sirens and bells, provide limited information content; loud speakers cover only very small areas and are often hard to understand, even for those not hearing impaired or fluent in the local language. Radio and television offer larger information volume, but are hard to target geographically and do not work well to address the “walking wounded” or other pedestrians. Both are not suitable for warnings, as many of those needing the information will not be listening or watching at any given time, particularly during work/school and sleep hours.

This problem has been illustrated by the London underground bombing on July 7, 2006, as described in a government report [July2005]. The UK authorities could only use broadcast media and could not, for example, easily announce to the "walking wounded" where to assemble.

1.2. Exigent Communications

With the usage of the term 'Exigent Communications' this document aims to generalize the concept of conveying alerts to IP-based systems and at the same time to describe the actors that participate in the messaging communication. More precisely, exigent communications is defined as:

* Communication that requirs immediate action or remedy. Information about the reason for action and details about the steps that have to be taken are provided in the alert message.

* An alert message (or warning message) is a cautionary advice about something imminent (especially imminent danger or other unpleasantness). In the context of exigent communication such an alert message refers to a future, ongoing or past event as the signaling exchange itself may relate to different stages of the lifecycle of the event. The alert message itself, and not the signaling protocol that convey it, provides sufficient context about the specific state of the lifecycle the alert message refers to.

On a high level the communication occurs in two phases with the subscription phase sometimes being implicit:

Subscription:

In this step Recipients express their interest in receiving certain types of alerts. This step happens prior to the actual delivery of the alert. This expression of interest may be in form of an explicit communication step by having the Receiver send a subscribe message (potentially with an indication of the type of alerts they are interested in, the duration of the subscription and a number of other indicators). For example, parents may want to be alerted of emergencies affecting the school attended by their children and adult children may need to know about emergencies affecting elderly parents. The subscription step may,

however, also happen outside the Internet communication infrastructure and instead by the Recipient signing a contract and thereby agreeing to receive certain alerts. Additionally, certain subscriptions may happen without the Recipient's explicit consent and without the Receiver sending a subscription. For example, a Tsunami flood alert may be delievered to Recipients in case they are located in a specific geographical area.

It is important to note that a protocol interaction initiated by the Receiver may need to take place to subscribe to certain types of alerts. In some other cases the subscription does not require such interaction from the Receiver. Orthogonal to the need to have a protocol interaction is the question of opt-in vs. opt-out. This is a pure policy decision and largely outside the scope of a technical specification.

Alert Delivery:

In this step the alert message is distributed to one or multiple Receivers. The Receiver as a software module that presents the alert message to the Receipient. The alert encoding is accomplished via the Common Alerting Protocol (CAP) and such an alert message contains useful information needed for dealing with the imminent danger.

Note that an alert receiver software modules may not necessarily only be executed on end devices humans typically carry around, such as mobile phones, Internet tablets, or laptops. Instead, alerts may well be directly sent to displays in subway stations, or electronic bill boards. Furthermore, a software module that obtains an alert may not necessarily need to interact with a human (as the Recipient) but may instead use it as input to another process to trigger automated behaviors, such as closing vents during a chemical spill or activating sirens or other warning systems in commercial buildings.

Finally, a few introductory words about the scope of this writeup: This document provides terminology, requirements and an architectural description. Note that the requirements focus on the communication protocols for subscription and alert delivery rather than on the content of the alert message itself. With the usage of CAP these alert message content requirements are delegated to the authors and originators of alerts.

2. Terminology

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119], with the important qualification that, unless otherwise stated, these terms apply to the design of a protocol conveying warning messages, not its implementation or application.

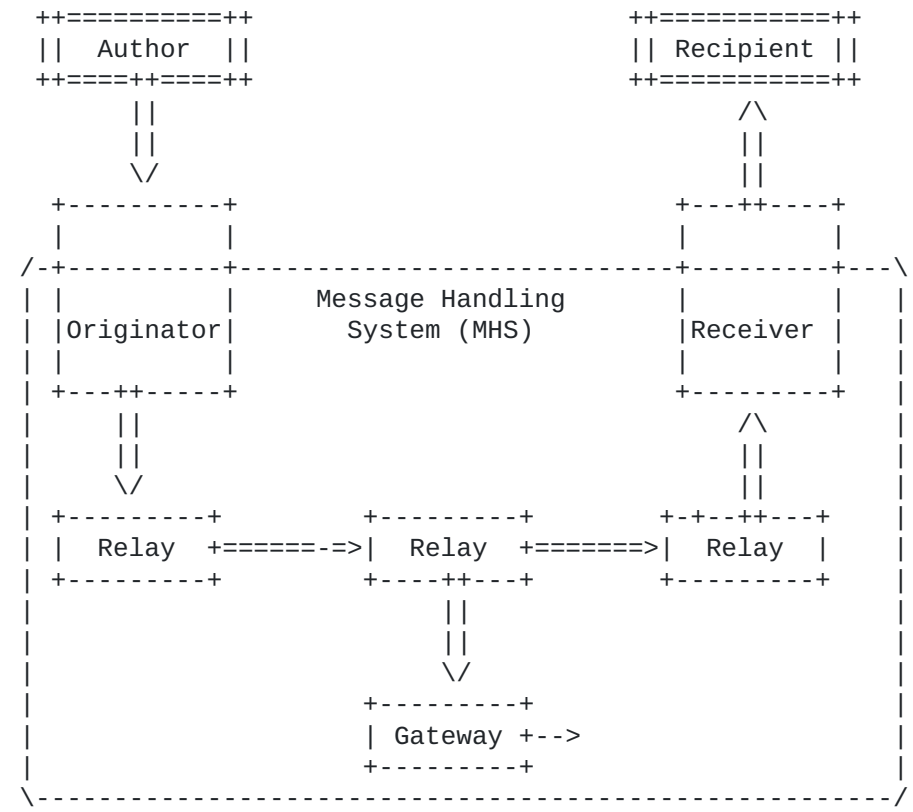
Alert messages are typically produced by humans and consumed by users, Authors and Recipients in our system, are the sources and sinks of alert messages.

The Author is a human responsible for creating the content of the alert message, and to make a decision about the intended recipients, even though the exact list of recipients may be unknown to the Author at the time of writing the alert message. Instead, the recipients may, for example, be described in terms of a geographical region, or recipients with interest in a specific alert type.

The Recipient is a consumer of the delivered alert message. It is a human reading the alert message.

From the user's perspective, all alert message transfer activities are performed by a monolithic Message Handling Service (MHS), even though the actual service can be provided by many independent organizations. The Message Handling Service (MHS) performs a single end-to-end transfer of warning messages on behalf of the Author to reach the Recipients.

Figure 1 shows the relationships among transfer participants. Transfers typically entail one or more Relays. However, direct delivery from the Originator to Receiver is possible.



Legend: === and || lines indicate primary (possibly indirect) transfers or roles

2.1. Originator

The Originator ensures that a warning message is valid for transfer and then submits it to a Relay. A message is valid if it conforms to both communication and warning message encapsulation standards and

local operational policies. The Originator can simply review the message for conformance and reject it if it finds errors, or it can create some or all of the necessary information.

The Originator serves the Author and can be the same entity in absence of a human crafting alert messages.

The Originator also performs any post-submission, Author-related administrative tasks associated with message transfer and delivery. Notably, these tasks pertain to sending error and delivery notices, and enforcing local policies. The Author creates the message, but the Originator handles any transmission issues with it.

2.2. Relay

The Relay performs MHS-level transfer-service routing and store-and-forward, by transmitting or retransmitting the message to its Recipients. The Relay may add history information (e.g., as available with SIP History Info [RFC4244]) or security related protection (e.g., as available with SIP Identity [RFC4474]) but does not modify the envelope information or the message content semantics.

A Message Handling System (MHS) network consists of a set of Relays. This MHS network is above any underlying packet-switching network that might be used and below any Gateways.

2.3. Gateway

A Gateway connects heterogeneous communication infrastructures and its purpose is to emulate a Relay and the closer it comes to this, the better. A Gateway needs the ability to modify message content.

Differences between the different communication systems can be as small as minor syntax variations, but they usually encompass significant, semantic distinctions. Hence, the Relay function in a Gateway presents a significant design challenge, if the resulting performance is to be seen as nearly seamless. The challenge is to ensure end-to-end communication between the communication services, despite differences in their syntax and semantics.

2.4. Receiver

The Receiver performs final delivery and is typically responsible for ensuring that the appropriate user interface rendering is executed to interact with the Recipient.

3. Alert Delivery Architecture

Section 1 describes the basic two steps that are involved with the alert message handling, namely subscription and alert delivery. From an architectural point of view there are, however, a few variations possible depending on the characteristics of the subscription process and the style of message delivery. This section offers more details on the communication architecture and highlights the necessary standardization actions.

3.1. Point-to-Point Alert Delivery

We start our description with the so-called "school closed" example where school authorities send alerts to all parents to notify theim about the fact that their children cannot attend school. Parents subscribe to these events when their children start attending the school and unsubscribe when they are finished with a particular school. The subscription procedures establishes some form of group communication by requiring an initial registration procedure. Typically, alert messages stay within the closed group and are not shared with others and alert message delivery is point-to-point with whatever communication protocol is most suitable. This also means that the alerts reach those who have subscribed rather than those who are in the vicinity of the school. The number of Recipients is typically rather small, in the order of hundreds to several thousands.

A variation of the "school closed" example is an explicit subscription model where no closed group pattern exists. Consider a traveler who would like to receive weather alerts about a specific geographical region. He may have to manually search for how to subscribe to alerts for the desired region, potentially looking a different subscription points for different types of alerts. As an automated version of this procedure some form of discovery may help to find these subscription servers. The approach described in [I-D.rosen-ecrit-lost-early-warning] is one possible way to discovery such alert subscription servers. The number of alert message Recipients is larger than in the previous example but will typically stay below the millions.

These alert delivery examples are supported by a number of standardized communication protocols, such as SIP, XMPP, eMail, or RSS feeds. The standardized distribution of alert messages is, however, a lower priority.

3.2. Multicast/Broadcast Alert Delivery

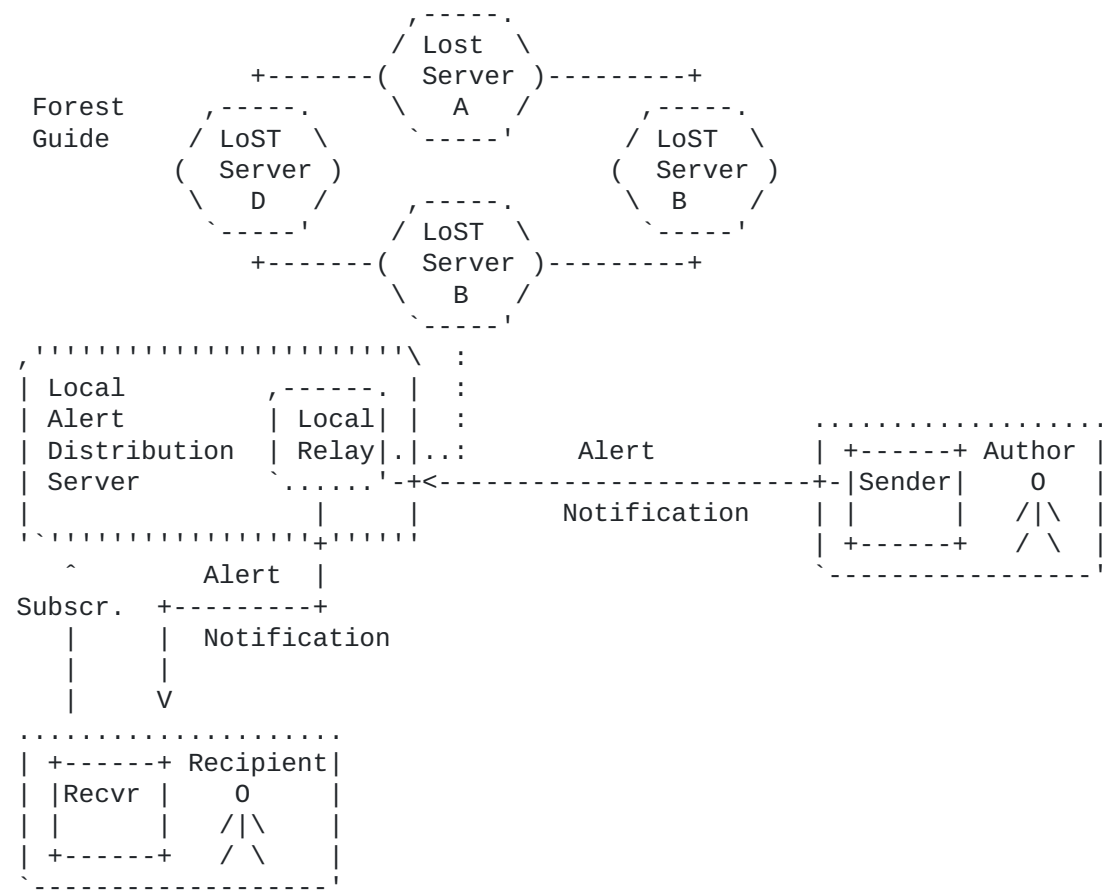
With the next category we move to a scenario where large number of Recipients shall be notified but the subscription itself is implicit, as it is the case when persons are within a specific region that can easily be reached by making use of broadcast link layer technologies. The placement of the actors from Figure 1 is thereby important. An Originator distributes the alert message to Relays within the geographically affected area. Those Relays are located within Internet Service Providers so that multicast and broadcast communication protocols can be utilized for efficient distribution to a large number of Recipients within the affected area. When the alert message delivery has to be accomplished at the networking layer then various requirements, such as the ability to traves NATs and firewalls, have to be met by such a protocol. The number of alert message Recipients is very large, potentially in the millions.

As a variation of the previously described model consider an alert distribution that uses a multicast network layer distribution mechanism but subscription to the alerts is explicit. Figure 2 shows the architecture. The LoST Forest Guide ensures that there is a way for Receivers to discovery local alert subscription servers very

much in the same way as LoST is used for citizen-to-authority emergency services, see [RFC5582]. The individual LoST servers know about the authoritative LoST servers for their region and redirect discovery requests in case they cannot return a cached response. The result of the LoST lookup is a URI to an local alert distribution server.

Once a Receiver had discovered a local alert distribution server it sends a subscribe message to it (with additional information about the type of alert it is interested in). As a response, it will receive information about the security credential the relay is going to use for subsequent alert delivery.

When an Author creates an alert for distribution the affected region will be indicated and so the alert will be sent to a Relay within the realm of the local alert distribution server and a notification will be sent to all the subscribed Receivers. The local Relay and the local alert subscription server will therefore cooperate in the handling of the alerts.



4. Requirements

4.1. Requirements for the Discovery of an Alert Distribution Server

Req-D1:

The protocol solution MUST allow a receiver to discover a local alert distribution server, as discussed in Section 3 and shown in Figure 2.

4.2. Requirements for Alert Subscription

The requirements listed below refer to the alert subscription phase.

Req-S1:

The protocol solution MUST allow a potential Recipient to indicate the language used by alert messages.

Req-S2:

The protocol solution MUST allow a potential Recipient to express the geographical area it wants to receive alerts about.

Req-S3:

The protocol solution MUST allow a potential Recipient to indicate preferences about the type of alerts it wants to receive.

Req-S4:

The protocol solution MUST allow a potential Recipient to express preference for certain media types. The support for different media types depends on the content of the warning message but also impacts the communication protocol. This functionality is, for example, useful for hearing and vision impaired persons.

4.3. Requirements for Alert Message Delivery

The requirements listed below refer to the delivery of alerts for two types of alerts communication patterns, namely point-to-point communication and broadcast communication. We separate the requirements for these two communication protocols.

4.3.1. Point-to-Point Alert Delivery

Req-P1:

The protocol solution MUST build on existing communication protocols and support the delivery of alert messages. Examples of such protocols are SIP, and XMPP.

Req-P2:

The protocol solution MUST allow targeting notifications to specific subscribers.

4.3.2. Broadcast Alert Message Delivery

Req-B1:

The protocol solution MUST Leverage lower-layer multi-/broadcast technologies. This implies non-TCP transport and congestion control being considered.

Req-B2:

The protocol solution MUST allow delivery of messages simultaneously to a large audience.

Req-B3:

The protocol solution MUST work in realistic network environments with firewalls and NATs. This typically requires a registration procedure and regular fresh messages to ensure that state at firealls and NATs is kept alive.

5. IANA Considerations

This document does not require actions by IANA.

6. Security Considerations

Figure 1 shows the actors for delivering an alert message assuming that a prior subscription has taken place already. The desired security properties of an MHS for conveying alerts will depend on the number of administrative domains involved. Each administrative domain can have vastly different operating policies and trust-based decision-making. One obvious example is the distinction between alert messages that are exchanged within an closed group (such as alert messages received by parents affecting the school attended by their children) and alert messages that are exchanged between independent organizations (e.g., in case of large scale disasters). The rules for handling both types of communication architectures tend to be quite different. That difference requires defining the boundaries of each.

Operation of communication systems that are used to convey alert messages are typically carried out by different providers (or operators). Since each be in operated in an independent administrative domain it is useful to consider administrative domain boundaries in the description to facilitate discussion about designs, policies and operations that need to distinguish between internal issues and external entities. Most significant is that the entities communicating across administrative boundaries typically have the added burden of enforcing organizational policies concerning external communications. For example, routing alerts between administrative domains can create requirements, such as needing to route alert messages between organizational partners over specially trusted paths.

The communication interactions are subject to the policies of that domain, which cover concerns such as these:

*Reliability

*Access control

*Accountability

*Content evaluation, adaptation, and modification

Many communication system make the distinction of administrative domains since they impact the requirements on security solutions. However, with the distribution of alert messages a number of additional security threats need to be addressed. Due to the nature of alerts it is quite likely that end device implementations will offer user interface enhancements to get the Recipients attention whenever an alert arrives, which is an attractive property for adversaries to exploit. Below we list the most important threats any solution will have to deal with.

Originator Impersonation:

An attacker could then conceivably attempt to impersonate the Originator of an alert message. This threat is particularly applicable to those deployment environments where authorization decisions are based on the identity of the Originator.

Alert Message Forgery:

An attacker could forge or alter an alert message in order to convey custom messages to Recipients to get their immediate attention.

Replay:

An attacker could obtain previously distributed alert messages and to replay them at a later time in the hope that Recipients could be tricked into believing they are fresh.

Unauthorized Distribution:

When a Receiver receives an alert message it has to determine whether the Author distributing the alert messages is genuine to avoid accepting messages that are injected by malicious entities with the potential desire to at least get the immediate attention of the Recipient.

Amplification Attack:

An attacker may use the Message Handling System to inject a single alert message for distribution that may then be instantly turned into potentially millions of alert messages for distribution.

One important security challenge is related to authorization. When an alert message arrives at the Receiver then certain security checks may need to be performed to ensure that the alert message meets certain criteria. The final consumer of the alert message is, however, the Recipient - a human. From a security point of view the work split between the Recipient and the Receiver for making the authorization decision is important, particularly when an alert message is rejected due to a failed security verfication by the Receiver. False positives may be fatal but accepting every alert message lowers the trustworthiness in the overall system.

7. Acknowledgments

This document re-uses text from [RFC5598]. The authors would like to thank Dave Crocker for his work.

The authors would like to thank Martin Thomson, Carl Reed, Leopold Murhammer, and Tony Rutkowski for their comments.

At IETF#79 the following persons provided feedback leading to changes in this document: Keith Drage, Scott Bradner, Ken Carberg, Keeping Li, Martin Thomson, Igor Faynberg, Mark Wood, Peter Saint-Andre.

8. References

8.1. Normative References

[RFC2119]	Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
[RFC5598]	Crocker, D., "Internet Mail Architecture", RFC 5598, July 2009.

8.2. Informative References

[RFC4244]	Barnes, M., "An Extension to the Session Initiation Protocol (SIP) for Request History Information", RFC 4244, November 2005.
[RFC4474]	Peterson, J. and C. Jennings, "Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)", RFC 4474, August 2006.
[RFC5582]	Schulzrinne, H., "Location-to-URL Mapping Architecture and Framework", RFC 5582, September 2009.
[I-D.rosen-ecrit-lost-early-warning]	Rosen, B, Schulzrinne, H and H Tschofenig, "A Uniform Resource Name (URN) for Early Warning Emergency Services and Location-to-Service Translation (LoST) Protocol Usage", Internet-Draft draft-rosen-ecrit-lost-early-warning-01, July 2009.
[July2005]	, , "Report of the 7 July Review Committee, ISBN 1 85261 878 7", (PDF document), http://www.london.gov.uk/assembly/reports/7july/report.pdf, June 2006.

Authors' Addresses

Henning Schulzrinne Schulzrinne Columbia University Department of Computer Science 450 Computer Science Building New York, NY 10027 US Phone: +1 212 939 7004 EMail: hgs+ecrit@cs.columbia.edu URI: http://www.cs.columbia.edu

Steve Norreys Norreys BT Group 1 London Road Brentwood, Essex CM14 4QP UK Phone: +44 1277 32 32 20 EMail: steve.norreys@bt.com

Brian Rosen Rosen NeuStar, Inc. 470 Conrad Dr Mars, PA 16046 US EMail: br@brianrosen.net

Hannes Tschhofenig Tschofenig Nokia Siemens Networks Linnoitustie 6 Espoo, 02600 Finland Phone: +358 (50) 4871445 EMail: Hannes.Tschofenig@gmx.net URI: http://www.tschofenig.priv.at