A CONFERENCE MECHANISM FOR THE SIMULTANEOUS TRANSMISSION OF VOICE AND MEDICAL INFORMATION USING SIP

Alexis Dorais-Joncas
Université de Sherbrooke
adorais@usherbrooke.ca

Wajdi Elleuch
Université de Sherbrooke
wajdi.elleuch@usherbrooke.ca

Alain C. Houle
Université de Sherbrooke
alain.houle@usherbrooke.ca

Abstract

New services based on Session Initiation Protocol (SIP) signaling, such as videoconferencing and presence-based applications, are emerging. New services mean new sets of requirements which, in turn, mean that existing architectures or protocols are not always adequate. A concrete example of such a new service is a system enabling an ambulance attendant to communicate with a number of physicians at the same time. In addition to regular voice transmission, the conference thus established would allow participants to study a patient's biometric data and share annotations directly. After formalizing the new set of requirements this service calls for, we turned to the most common mechanisms available and found that there are none perfectly suited to this new application. We therefore considered a less well-known mechanism, which is based on SIP and the implementation of a full mesh topology. This solution provides strong reliability, easy integration in future SIP user agents, powerful security capabilities and other essential features. ecgML, an innovative open XML-based format developed for the storage and transmission of electrocardiograms (ECGs), was analyzed and integrated into the proposed solution for the transmission of the patient's biometric data.

Keywords*: SIP conferencing, ECG transmission, full mesh, ecgML, decentralized communications.*

1. Introduction

When a person suffers from a heart problem, it is widely known that rapid intervention and treatment by a specialized medical team is crucial to maximize the victim's chances of survival and minimize after-effects. Furthermore, rapid intervention usually results in a shorter hospitalization.

Having a means of monitoring the patient's condition during transportation by ambulance and quickly establishing a multimedia conference between multiple medical caregivers would solve two issues at once while decreasing delays before treatment.

First, voice and medical data (e.g. electrocardiograms (ECG)) communication between the ambulance and the designated hospital would allow the emergency physician to establish a diagnosis before the patient's arrival, allowing better preparation before the patient's admission and triage. Second, that same emergency physician would be able to judge whether the patient's condition requires the attention of a cardiologist. In such a case, a voice conference could be established between the ambulance attendant, emergency physician and cardiologist, and the patient's biomedical data would be available to all three.

A multimedia conferencing system capable of monitoring the patient's health status during transportation can be designed using current technologies. Many technical issues, from low-level radio-communication considerations to high-level content adaptation and presence services, need to be addressed. In this project, our contribution to such a system is to develop a mechanism that allows voice and data exchange between multiple entities connected through an IP-based network.

In this article, you will be presented with an overview of the requirements and general characteristics our solution should meet. Current conferencing technologies will then be presented, followed by our system's design and a brief conclusion.

2. Requirements

We interviewed physicians and cardiologists to establish an extensive list of all the possible requirements and restrictions constraints our solution must comply with. We focused on the three most important ones, those with the potential to greatly affect the architecture of our solution: strong reliability, security and interoperability.

2.1. Reliability

Medical systems require the highest possible level of reliability and our solution is no exception: it is primarily intended to be used on the frontlines of a medical care service. Our objective is to provide a design that is simple and cost effective, and that offers a greater level of reliability than the one traditional solutions have to offer.

2.2. Security

Government laws and regulations are usually very strict when it comes to handling private medical data transmitted through computer networks. Canada is no exception. The Privacy Policy Division of Health Canada has the responsibility of examining how privacy and confidentiality can be safeguarded when patient health information circulates among health care providers and on sites on a "need-to-know" basis. A pan-Canadian framework [1] was proposed in 2005 and is awaiting provincial approval. It is vital to take such constraints into account when creating a communication system in a medical context, as some of the best technological

ideas may never be approved for use in a medical context unless strict security is ensured.

2.3. Interoperability

Health network structures are becoming more and more complex every day. New technologies are constantly emerging, which justifies more and more inter-networking between back-end data storage, monitoring units and other kinds of devices that are highly relevant in a medical center. Most of these devices implement proprietary communication protocols and data formats that either restrict the interoperability to only a few devices made by the same manufacturer or increase the cost of integrating different devices made by different manufacturers. This is where open standards becomes very interesting, both technologically and financially. They help avoid vendor lock-in, providing a wider range of options and preventing constant increase of operating and support fees imposed by vendors.

3. General characteristics

Further interviews allowed us to extract the general characteristics of a communication system that would adequately solve the problem described in the introduction.

- A conference between the ambulance attendant and the relevant caregivers will typically involve two to five people, as shown in Figure 1.
- A conference needs to be created and fully established the moment it is needed (*ad hoc*). Therefore, it is hard to foresee and manage the number of simultaneous conferences that may be in use at any given time.
- No conference control mechanism is needed, apart from letting any user invite another one (dial-out invitation). In no case is participant exclusion or conference termination expected to be required..
- No sharing of exclusive resources is expected, thus eliminating the need for floor control [2].
- Transmission of the patient's demographic information, ECG, medical history and other relevant information must be made in an open, cross-platform format.

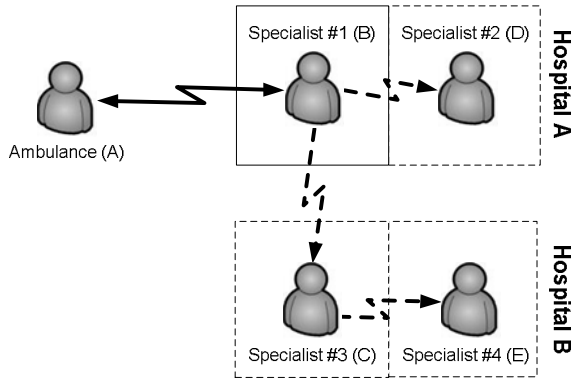

Figure 1 : People involved in a typical medical conference

4. Existing conference mechanisms

SIP conferencing is a very promising technology that aims to go beyond conventional audio and video conferencing. Among the key functionalities are [3]:

- The ability to invite a participant to a conference without having to enter a phone number or pass code.
- The easy display of the basic conference information on any SIP-capable device.
- The ability to support multiple signalling and mixing modes.
- The possibility for a user agent to communicate its "in conference" status to another, a caller for example, allowing the latter to inform the user of the situation and to take specific actions based on that information.

In order to design the best solution based on the requirements and general characteristics presented in section 2 and 3, we evaluated and compared the existing conferencing and mixing technologies available using baseline SIP [4].

4.1. End system mixing

In this conference model, a user agent (A) involved in a two-way call with user agent B can invite a third party (C) to create a conference. The user agent A then turns into a focus [5], meaning it becomes the central point of the conference. The focus is always in charge of the signalling and, usually, of the media mixing, although other mixing mechanisms are available. This centralization hides participants from each other, except for the focus of course. This model has the benefit of allowing conferencing between user agents that are conference-unaware, reducing the complexity of the non-focus user agents (focuses must include the added complexity of managing the conference signalling and the media mixing). A major drawback of this model is that as soon as the focus (A) leaves the conference, there is no way for B and C to stay connected, due to their conference-unawareness. Furthermore, conference-unaware user agents will not be able to benefit from many of the advanced SIP conference capabilities, some of which were mentioned at the beginning of the current section.

4.2. Centralized mixing

This model implies the presence of a central conference server (CS). The server can provide many distinct ways of creating and maintaining a conference. A dial-in CS allows a conference to be planned in advance, where conference users send an INVITE to the CS in order to be integrated into the conference. A dial-out CS acts inversely, as it is the only entity that can invite other users. *Ad hoc* centralized conferences are also possible with the use of a CS: a two-way call is transferred over to a dial-in CS and the third user is invited to dial into the conference using a REFER [6] request.

A CS is quite simple to use for planned conferences. It also provides a certain quality of service (QoS), since a user may reserve a certain amount of resources in advance. However,

having a central entity managing the conferences means that if the CS goes down, so do all the existing conferences. Moreover, while moving a standard two-way call to a *n*-way conference is possible, it is a procedure that is both unintuitive and complex.

4.3. Centralized signalling, distributed media

This is similar to the centralized mixing model presented in section 4.2, except for the way media distribution is handled. Media mixing and distribution is no longer the responsibility of the CS: the media is sent directly to every participant. Standard multicast or multi-unicast (referred to as Decentralized Multipoint Conference in H.323 [7]) can be used for data transmission. This has the advantage of reducing the processing load on the conference controller, but it increases the user agent complexity slightly.

5. System design

5.1. Design approach

Based on the many benefits SIP conferencing can provide, we looked for an appropriate SIP-based conference mechanism that would satisfy every requirement exposed in section 3. It is clear that none of the more traditional models described in section 4 are fully satisfactory. Instead, we chose a fully-connected model (figure 2), which is an experimental design that is categorized as a loosely coupled" model, as opposed to the traditional tightly-coupled models. Tightly-coupled models necessarily include a central point of control of the conference, whereas in a loosely-coupled model, both signalling and media transmission are fully distributed. Of course, a mechanism insuring the consistency of the conference representation among participants is needed. Such a mechanism is described by J. Lennox and H. Schulzrinne in [8].

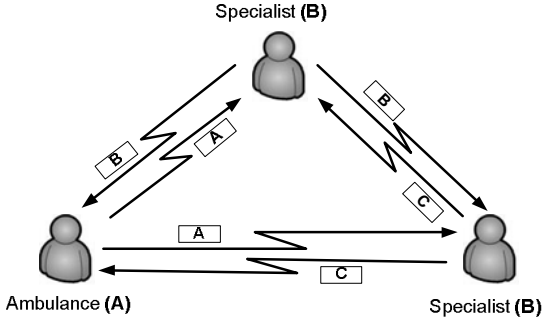

Figure 2: Fully-connected conference

Our approach offers many benefits:

- It eliminates the problem of a single point of failure, greatly increasing the system's reliability.
- It allows end-to-end integrity verification and encryption.
- Most of the signalling can be used with baseline SIP, thus making integration of the mechanism into existing SIP-enabled devices easier.

These benefits offer an almost perfect solution to the requirements presented in section 2.

5.2. System limitations

However, this solution does not come without limitation. The most important one is regarding bandwidth usage. In a *n*-way fully-distributed conference, each participant must maintain *n*-1 signalling and multimedia streams. The protocol used to maintain the mesh permanently fully-connected also generates a certain amount of bandwidth whenever a modification to the conference is detected (participant arrival/departure).

Processing power can also be an issue. For most of the popular voice codecs, encoding is more processing-intensive than decoding [8]. Each participant must have enough processing power to decode simultaneously at the most *n*-1 media streams, but will only need to encode one stream per outgoing codec. It is said "at the most" since typically, not every member of the conference will be speaking at the same time. If a silence-suppression technique is applied, the number of streams to be decoded by an end point is reduced to exactly the number of participants that are transmitting an audio signal.

5.3. Medical data design

In addition to the voice transmission, the conference must also carry the patient's biometric data. Typically, ECGs will regularly be sent by the ambulance to the conference. Caregiver will have the ability to annotate the ECGs and send back the annotations to the conference.

We examined the most popular formats for storage and transmission of medical data (i.e. HL7, DICOM, SCP-ECG) and concluded that most are either proprietary or overwhelmingly complex to support. Such formats make interoperability between different devices much difficult than it should be. Searching and data mining for information is also very inefficient, when at all possible, leading to multiple data duplication and increased operating costs.

This is where a new, innovative ECG file format is introduced. Based on the XML technology [9], ecgML [10] is system-, application- and format-independent and is meant only to carry basic patient's information and raw ECG content. Presentation and display are strictly separated from the ecgML format, maximising information re-usability and facilitating data mining. Content separation also allows the use of on-the-fly XSLT transformations to transparently convert ECG's into various other formats such as MPEG (audio), MatLab (text) and PNG (image).

Many XML parsers exist on a variety of platforms and can be used to efficiently parse ecgML documents. XML also provides an easy-to-follow hierarchical data structure, distinguishing ecgML from proprietary, binary flat file formats.

5.4. System messaging design

The simultaneous transmission of voice and medical data is done within the SIP sessions resulting from the conference establishment.

Full-mesh conference signalling

The aim of the message exchange is to allow any party to communicate in a conference without a central point of control. The deployed protocol allows parties to join and leave the conference at any time, and ensures that all members of the conference are always informed of new members [8]. New headers were added to the basic SIP messages in order to implement the protocol. For example, the Conference-Member header allows user agents to maintain and update their view of the conference participants. Figure 3 shows the messaging sequence diagram of a typical conference establishment.

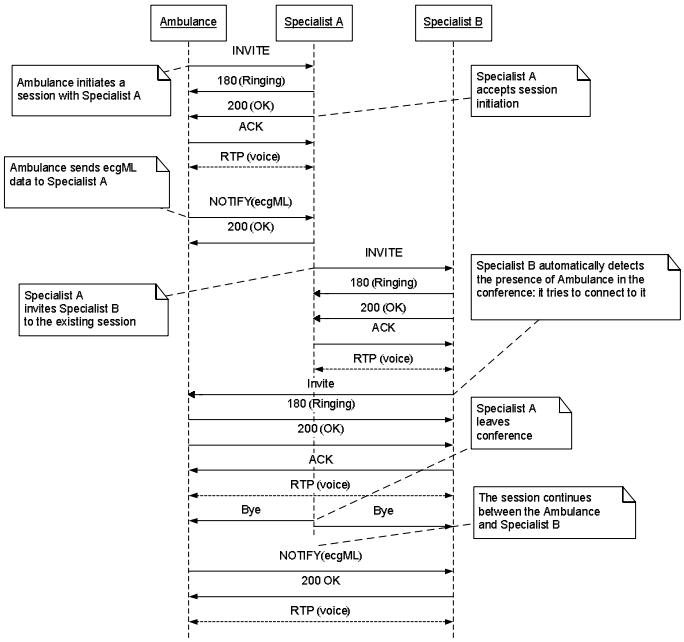

Figure 3: Overview of a typical messaging sequence diagram

Transmission of ecgML

Once the conference is established, the medical data can be propagated to every participant. ecgML data is transmitted by encapsulating the data into a SIP NOTIFY [11] payload. One NOTIFY packet is sent to every member of the conference.

6. Conclusions

A system allowing a specialist to monitor a patient's condition during transportation by ambulance and quickly establish a multimedia conference between multiple medical caregivers must meet some important requirements. Strong reliability, information security and interoperability were considered for this project.

Based on those requirements, existing SIP conferencing models were carefully examined. Analysis showed that a novel mechanism based on a full-mesh topology was the best approach possible.

The multimedia conference must carry voice and data transmission. Voice is transmitted using standard VoIP protocols and the patient's biometric data, stored in the ecgML file format, is encapsulated in the SIP's NOTIFY method's payload. ecgML is an XML-based file format designed to store medical information, including electrocardiogram data..

The full-mesh topology is an excellent solution to the problem described in this article, although at the cost of higher bandwidth usage and processing load on conference user agents.

Future research needs to be done in order to obtain an efficient wireless link between the ambulance fleets and health networks, as well as to integrate a presence service allowing caregivers to contact the most appropriate resources available at a given time and location.

Acknowledgements

The authors would like to thank Dr. Guy Bisson, Dr. Serge Lepage and Fakher Ben Salah for fruitful discussions. They also acknowledge the contribution of H. Wang and F. Azuaje for sharing their ecgML viewer under an open source license.

References

[1] Pan-Canadian Health Information Privacy and Confidentiality Framework, Health and the Information Highway Division, Health Canada, January 27, 2005

[2] G. Camarillo, J. Ot, and K. Drage, "*The binary Floor Control Protocol (BFCP),*" *Internet draft, IETF*, August 2004. Work in progress.

[3] H. Sinnreich, "*SIP Beyond VoIP,*" Von Publishing, New-York, 2005, 333 p.

[4] J. Rosenberg, "*A Framework for Conferencing with the Session Initiation Protocol (SIP),*" *IETF*, RFC 4353, February 2006

[5] A. Johnston, O. Levin, "*SIP Call Control – Conferencing for User Agents,*" *Internet draft, IETF*, December 2005. Work in progress.

[6] R. Sparks, "*The SIP Refer Method*," *IETF*, RFC 3515, April 2003.

[7] ITU-T Recommendation H.323 (2000), Packet-Based Multimedia Communications Systems.

[8] J. Lennox, H. Schulzrinne, "*A Protocol for Reliable Decentralized Conferencing,*" *Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video*, pp. 72-81, 2003

[9] Extensible Markup language (XML) [http://www.w3.org/xml/]

[10] H. Wang, F. Azuaje, B. Jung, N. Black, "*A markup language for electrocardiogram data acquisition and analysis (ecgML),*" *BMC Medical Informatics and Decision Making*, May 2003

[11] A. B. Roach, "*Session Initiation Protocol (SIP)-Specific Event Notification,*" *IETF*, RFC 3265, June 2002.