High Definition videoconferencing: The future of collaboration in healthcare and education

Jesus ALCOBER[1], Guillem CABRERA[2], Xavier CALVO[3], Erik ELIASSON[4], Kristina GROTH[5], Piotr PAWAŁOWSKI[6]

[1, 2, 3]*Dep. d'Enginyeria Telemàtica, Universitat Politècnica de Catalunya / Fundació i2CAT, Carrer Esteve Terradas 7,Castelldefels, Barcelona, 08860, Spain*

[1]*Tel: +34 93 4137052, Email: jesus.alcober@upc.edu*

[2]*Tel: +34 93 4137059, Email: guillem.cabrera@i2cat.net*

[3]*Tel: +34 93 4137059, Email: xavier.calvo-brugal@upc.edu*

[4]*Royal Institute of Technology (KTH), Valhallavagen 79, Stockholm, 100 44, Sweden*

Tel: +46 8 7906000, Email: ere@kth.se

[5]*Clintec, Karolinska Institutet, Stockholm, 171 77, Sweden*

Tel: +46 707 573662, Fax: +46 8 58582340, Email: kicki@csc.kth.se

[6]*Poznań Supercomputing and Networking Center, ul. Noskowskiego 10, Poznań, 61-704, Poland*

Tel: +48 61 8582189, Fax: +48 61 8582151, Email: astagor@man.poznan.pl

Abstract: The rapidly developing video technology and Internet market combined with the need of collaboration over distance has led us to design and develop a new videoconferencing system using High Definition (HD) picture. Our aim was to create an open and scalable platform based on the Service Oriented Architecture (SOA), making it easier to create and deploy new services and extend the platform's functionality. Not only business but also healthcare, education and home users where in the focus of our project, therefore we have tried to identify and implement the most cost-effective hardware and software solutions. In this paper we describe the methodology, usage scenarios and user requirements in the area of healthcare and education. We give a thorough summary of technologies used in the media, control and network planes of the project. We finalize the paper with the description of the current progress as well as conclusions and future plans.

1. Introduction

Collaboration over distance has increased during the last decade due to the improvement of IP networks and broadband access to the Internet. Especially in the healthcare, where specialised knowledge and capabilities are concentrated in a small number of hospitals, videoconferencing technology plays an important role. It gives the possibility to save money and time spent on travelling between hospitals and to increase the knowledge dissemination between doctors. Universities and schools are also looking for means to enhance the learning process by introducing videoconferencing technologies. Collaboration between schools and partnership programs with universities help students from different environments to get fair access to knowledge. This is especially important in developing countries with low budget for education. However, the required videoconferencing equipment is usually expensive and available only for larger organisations. Therefore, its use is not widely spread within education and healthcare. Furthermore, today's commonly used Standard Definition (SD) picture is not detailed enough for many applications,

especially in healthcare where details from an online operation may be crucial, motivating the introduction of High Definition (HD) picture into videoconferencing, including high resolutions such as 1920x1080 or 1280x720 pixels, and sharp picture with great detail visibility. Thus, specialized equipment and CODECs to make the transportation of video streams over the today's Internet feasible is required. A survey of the marked showed that currently there are no HD videoconferencing platforms with a completely open architecture available on the market. Commercial solutions are mostly proprietary platforms with limited functionality. Interoperability between products of different manufacturers often faces problems in fields such as session signalling, encryption and authentication, media and CODEC negotiations or network and Quality of Service (QoS) parameters reservation.

In the HDVIPER[1] project , we are designing and developing an open HD videoconferencing platform, focusing on the limitations of today's HD solutions as well as needs and expectations of real users in business, healthcare, education and home environments. In this paper, we describe key requirements, usage scenarios as well as different problems encountered in HD videoconferencing and our suggested solutions. We focus on healthcare and education showing innovative technology, usage of established standards and explain the setups in which the HDVIPER platform will be tested. Furthermore, in Section 2 and 3, methods and usage scenarios are detailed in healthcare and education, respectively. In Section 4, the technology used in HDVIPER project is described. In Section 5, current results of the project are outlined and, finally, in Section 6, conclusions and business benefits are drawn.

2. Methods and usage scenarios in healthcare

Within the healthcare setting qualitative methods for collecting requirements for the videoconference platform have been used. More specifically, different situations, in which videoconference technology is used today and where the need for it is high, have been studied. Surgeons in the conduct of their work, i.e., participating in multi-site decision meetings, in round meetings, and in operation theatres have been observed. Several surgeons have also formally and informally been interviewed about the situations observed.

The highly specialised medical care of serious diseases of the liver, pancreas, and gastro oesophagus was, in the beginning of the 21st century, centralised to the section of upper abdominal surgery (Gastro) at a university hospital in Sweden. Gastro has a close collaboration with several other units, such as pathology, radiology, oncology, and hepatology. One task at Gastro is the weekly multi-site video-mediated decision meetings with the focus on how to treat the patients. These meetings are held in rooms equipped with radiology workstations, video projectors and videoconference equipment. All remote participants can simultaneously see the same pictures shown by the radiologist using two projectors and each other on a large screen. Another task is surgery performed in the operation theatres, e.g., ERCP (Endoscopic Retrograde Cholangiopancreatography), used primarily to diagnose and treat conditions of the bile ducts, including gallstones, inflammatory strictures (scars), leaks (from trauma and surgery), and cancer. ERCP combines the use of X-rays and an endoscope, which is a long, flexible, lighted tube, and sometimes a spyglass. Through the endoscope, the physician can see the inside of the stomach and duodenum, and inject dyes into the ducts in the biliary tree and pancreas so they can be seen on X-rays. The spyglass is a finer instrument than the endoscope making it possible to, through the endoscope, go into more narrow paths in the body. A recording unit is used to record video snippets and take still images from either of the main equipment.

[1] This work has been partially funded by the Spanish Ministry of Industry, Tourism and Trade in the framework of the "Avanza I+D" programme (project number TSI-020400-2008-147) and by the Spanish Ministry of Science and Technology, under the project TSI2007-66637-C02-01, which is partially funded by FEDER.

Major requirements identified from the field data were: the technology must be secure in that all participants in a videoconference are identified, it needs functionality such as a possibility to transmit information objects (still images and video) from various devices together with audio-video streams from the conference room at the same time, to let the participants see when others interact with the information objects. Also the quality of the transmitted video, audio and information objects needs to be high, but may change during the videoconference depending on the current requirements.

More specifically, when connecting a surgeon in the ERCP room, through videoconference, with a surgeon located in another part of the building the following requirements were identified: high security of who participates in the videoconference and can see the transmitted pictures, both surgeons can see the same pictures without any loss of quality, the remote site must be able to control the camera and what pictures that are transmitted, no extra equipment should be needed when the operator talks to the remote person, the videoconference equipment used in the operation theatre should be small in size. A scenario from the ERCP room: 1) Pelle, a younger surgeon, is performing an examination in the ERCP room and needs help from a senior surgeon, Bill who is in his office, in another part of the hospital. 2) Pelle asks one of the nurses to call Bill and tell him to meet through the videoconference system. 3) Bill connects to the videoconference and can now see Pelle and Pelle can see Bill, they say hi to each other. 4) Pelle tells Bill about his problem seeing what is shown on the spyglass. 5) Bill listens and, at the same time, looks at the c-bow video. Bill sees exactly what Pelle can see on his screens in the ERCP room. 6) Bill switches one of the video windows to the spyglass. 7) Bill asks Pelle to move the spyglass a little bit further and explains what he thinks.

3. Methods and usage scenarios in education

Regarding videoconferencing in education, previous experiences on distant lectures of a professor living abroad were taken as a starting point. Additionally, several meetings with teachers and school authorities have been held. During such meetings, a survey was conducted to gather information about the participants' experiences with videoconferencing technologies, needs, expectations and foreseen problems. Based on this knowledge, a set of usage scenarios and requirements was prepared.

One of the basic usage scenarios of a videoconferencing system is a remote lecture. A teacher travelling or staying in a different place for some time may not have to cancel his classes. Using a videoconferencing system, he still can hold a planned lecture at his home university or school. In addition, many science subjects need to visualize processes during experimental lessons to explain ideas better. To perform such experiments can be expensive, and schools often lack the required infrastructure. Moreover, other kinds of experiments, due to their hazardous nature, can only be conducted by scientists in a specialized lab. The teacher or scientists need access not only to a videoconferencing system, but also to additional equipment such as multimedia projectors, blackboards or maps. The classroom should have a similar setup. The need of HD picture is important to increase the feeling of presence and the possibilities for a common ground in discussions. Students need to freely ask questions, read what is written on the blackboard or see other artefacts used, which can be difficult using low quality video and small screens.

In almost every school, there are small numbers of students with excellent grades and outstanding achievements. However, the level of education in school focuses on the average student and the outstanding ones often cannot develop as fast as they could. Some universities prepare weekend courses, but distance and travelling costs often make it impossible for students from small towns to attend. This problem could be solved by introducing videoconferencing technology and enabling students to join the university courses in a multipoint conference from their own schools. The multipoint setup gives the

opportunity not only to listen to the lecture and ask questions, but also to exchange ideas between students. As this scenario requires a videoconferencing system to be installed in every school, it is crucial to keep its price low.

A thorough analysis of the scenarios, experiences and surveys showed that the use of SD video images is a limitation for students and teachers. Students lack a feeling of being present in a lecture and miss details of explanations, which in turn leads to less attention to what the lecturer is saying. A solution could be to improve the quality by using HD images displayed on a large screen to show real-sized, high detailed videos and extra-tools, such as whiteboards or remote slide presentations. Other key issues, such as equipment cost, privacy and usability, were taken into account during the HDVIPER project.

4. Technology Description

Based on the collected data, many technical problems have been identified to be solved including media CODEC compatibility, signalling to establish and control both, direct calls, and multipoint conferences, supporting multiple simultaneous media streams, information security and procedures to establish reliable network connections.

In order to solve those problems, a new scalable videoconferencing platform called Snake was designed. The main idea was to define a common audio and video CODEC, common signalling strategies to establish media sessions, a dynamic reservation procedure for network resources and enhanced services to improve user experience. This new platform was designed following the Service Oriented Architecture (SOA) paradigm, aiming to improve the interoperability between different videoconferencing clients. The main aim was to define the minimal parameters that a Session Initiation Protocol (SIP) compatible videoconferencing client should fulfil to perform HD videoconferences. Thus, the platform assures that any videoconferencing agent can participate in virtual meetings, although only Snake labelled clients can use the extra services defined within the platform, as shown in Figure 1. Moreover, a minimal infrastructure to create and offer these new extra services to the final user was planned. These services communicate using WebServices (WS) interfaces and are deployed within an Enterprise Service Bus (ESB). On one hand, the choice of WS was motivated due to its wide spread use among the Internet services industry and the ease to rapidly create clients to use these services. On the other hand, the deployment within an ESB (see Figure 1) was selected because services stubs can interact between themselves using standard messaging protocols, and the addition of other services does not interfere with previously offered services and enables third party companies to create new added value services.

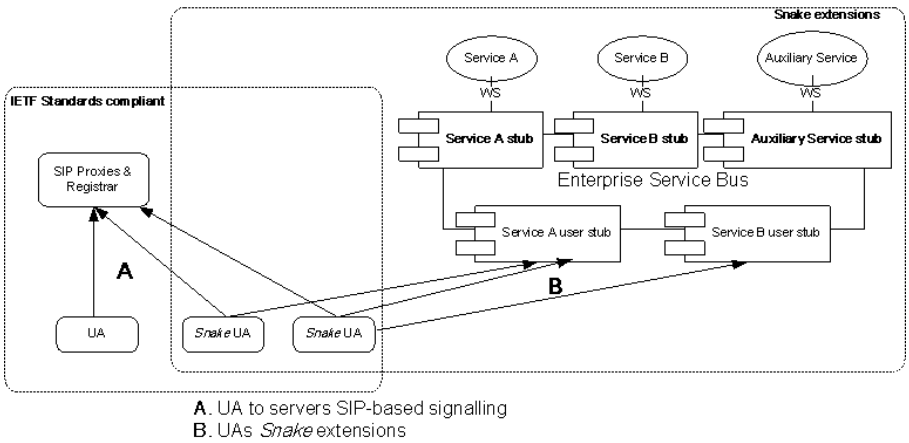

Figure 1: User agent relations and Snake services deployment

In the media plane, the most important characteristic of the Snake platform is the use of HD video streams for videoconferencing, which combined with large screens and augmented detail visibility offers the user the best possible quality, thereby improving the experience by perceiving gestures and body language.

From the survey of commercial products available in the market, it was concluded that most of HD capable videoconferencing products implement proprietary media CODECs. This becomes a big interoperability issue mainly between different manufacturer agents or even between different versions of the same device.

With the goal to solve the interoperability issues, a common Application Programming Interface (API) to capture, encode, packetize, transmit, depacketize, decode and visualize images was defined (see Figure 2). The goal was to make all clients fulfil this coding interface, thereby becoming compatible, as the only change needed is in the coding/decoding steps for each specific video CODEC.

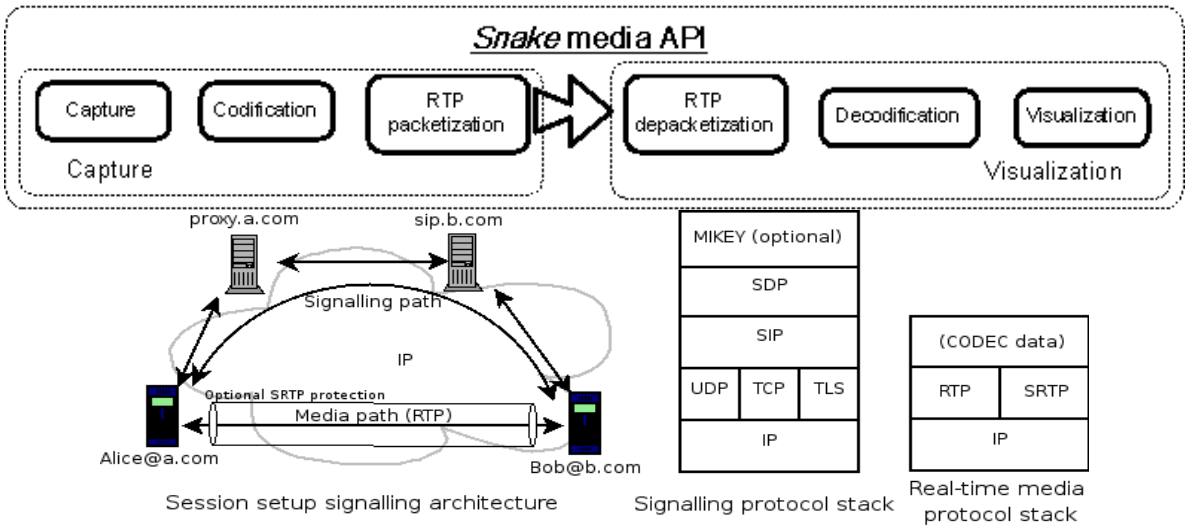

Figure 2: Snake structures

Another fact deducted from our user surveys was the need of interaction improvement through videoconferencing systems. To achieve this, media should be transmitted and displayed with as minimum delay as possible, trying to reduce it between the source and the destination of the information. Furthermore, a trade off between the codification and decodification time, due to its algorithmic complexity and the computing process, and the resulting bandwidth usage must be found to enable better end user interaction.

The signalling layer, to establish media sessions in videoconferences, was designed using Session Initiation Protocol (SIP) and Session Description Protocol (SDP) . SIP was chosen since it is the standardized protocol in Voice over IP (VoIP) environments and it is able to set up any kind of conference session. Moreover, already existing SIP infrastructure can be reused. In combination with SIP, SDP becomes a powerful mechanism to negotiate media session parameters, such as ports, media types and video CODECs.

In business and healthcare scenarios, sensitive information with legal requirements to handled it securely, is often discussed. To solve the privacy requirement, and encrypt the sessions, we have specified how Snake applications can implement strong security (end-to-end mutual authentication and encryption with perfect forward secrecy), by using the IETF standardized protocols Multimedia Internet KEYing (MIKEY) and Secure Real-time Transport Protocol (SRTP) . MIKEY lets applications use passwords or a Public Key Infrastructure (PKI) to exchange encryption keys, which are then used by SRTP to encrypt all audio and video media sent over the network. Figure 2 shows the signalling architecture and the protocols used to secure the HD conferences.

Videoconference's data is sensible traffic in the network plane, mainly in terms of needed bandwidth and maximum admitted delay, jitter and packet loss. Video and audio streams are sensitive to jitter, and a high delay or packet loss has a direct effect on the user experience. Additionally, related to high quality video and audio streams, the required bandwidth is high and the acceptable jitter is low. This traffic is modelled as a restrictive traffic pattern and requires high quality network access lines. These are often expensive and their utilization is low, since average generated traffic is far from their capacity.

Nowadays, usually the only way to fulfil the traffic requirements of the media transmitted in video conferences is to hire a fixed access line with the Quality of Service (QoS) fixed by the used videoconferencing system. This is, basically, due to the lack of dynamic network reservation mechanisms available to the final user and is a problem for network operators, since they need much more resources to offer a static service than in a resource reallocation scenario. Within the Snake platform a Bandwidth on Demand (BoD) service has been planned to be a dynamic network reservation service. It will be used as a transparent network resources reservation mechanism offered directly to the end user.

Several network reservation systems are available today, so the BoD service should be compatible with as many as possible. To achieve this, the service was designed as a public interface interacting with connectors to the different available systems. Therefore, the service is ready to be extended to new reservation systems with no need for the user to know about them (see Figure 3). If the BoD service user interface is integrated within a videoconference client, it could work as a transparent service when a meeting is started, making the needed reservations on the network plane without any actions by the user.

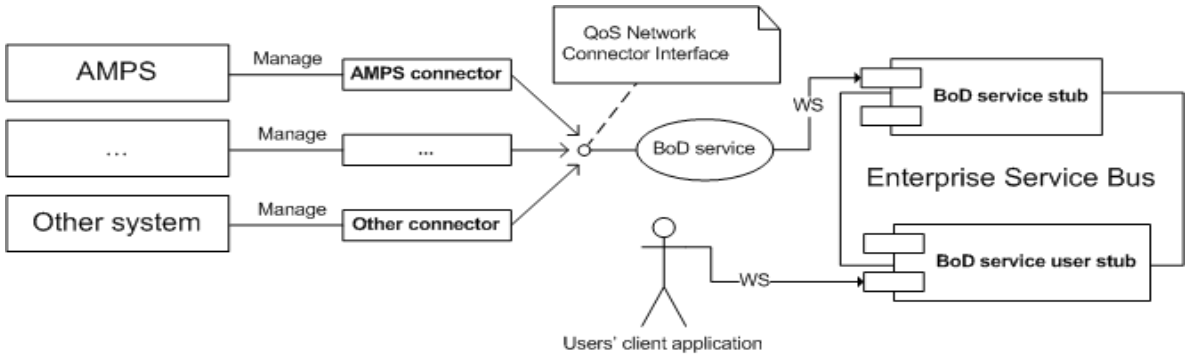

Figure 3: Network resources reservation service

As stated in the usage scenarios section, multipoint videoconferences were planned as part of the Snake platform. However, a major technical problem arose. In a multipoint conference, all participants should receive media streams from all other parties, which makes necessary to have a protocol that signals all participant media streams. Although several approaches can be taken here, in HDVIPER , the use of a Multipoint Control Unit (MCU), a central point that receives all media streams and delivers it to other parties, was chosen. The proposed solution within HDVIPER was to create a software-based MCU from an RTP packet reflector and a SDP capable SIP Server Agent. The use of an RTP packet reflector made the service ready to work with any RTP stream, with no effect over the media carried inside. This was very interesting for videoconferences, since any media CODEC supported by the clients could be used and the added delay by the central element was very low, having no effect on the user perception. In order to build a full MCU service, an SDP capable SIP Server Agent was developed. This control layer acted as a SIP User Agent, negotiating with clients the session establishment and their parameters. From these parameters, it is possible to configure the RTP Packet reflector to receive media streams from all parties and to forward them to all other participants in the conference. SDP was a valid option for a signal protocol for media streams, but some net procedures were required.

During the development phase many unpredicted issues appeared. Differences in SIP protocol implementation between various clients, performance issues with single threaded applications and operating systems network buffer overloads had to be solved. These experience and lessons learned will make future developments much easier.

5. Current results

The HDVIPER project is currently in the final stage of the integration phase. All main implementation tasks have been finished and tests among partners are being performed.

After an analysis over several options, H.264 was pointed out as the most suitable video CODEC, mainly due to its flexibility in terms of compression level and strategies. The x264 open source CODEC implementation was taken as a starting point. As a second implementation, the H.264 CODEC from Intel Integrated Performance Primitives library has been also integrated. Both CODECs deliver similar performance taking advantage of multiple core processors. For the audio streams, an open source implementation was chosen as well, Speex , which is a speech oriented audio CODEC that provides high quality audio streams using relatively low bandwidth. In terms of security, no open source implementation of the key exchange protocol (MIKEY) existed, and a LGPL licensed version, which any conferencing application can use to add strong authentication and encryption of the audio and video streams, have been released. In addition, the following basic services were implemented and integrated in the reference platform. The Bandwidth on Demand (BoD) service enables the Snake platform to make bandwidth reservation using various QoS systems. As a reference implementation, a connector to the Advance Multi-domain Provisioning System (AMPS) was designed and developed. The AMPS was developed within the GEANT2 project, and allows reservations for Premium IP bandwidth, being effective across a chain of participating domains. AMPS is a solution that may allow to make bandwidth reservations just as a conference starts. The Presence Agent (PA) is a service that informs the user about the status of people from its contact list. The PA shows if a user is online and provides an interface for status messages as well. When a user client registers or unregisters using SIP, it is informed by the SIP Registrar and forwards this information to other users. The Services Manager (SM) is a registry of all services deployed within an instance of the platform. It provides a list of services and their WS interface addresses for the user application to connect. This makes sense in an open platform, since new services can be deployed at any time and user agents need to create their clients to access to them. The Call-back service enables the ESB to send information to the user application bypassing problem caused by firewalls or NATs. It uses the Long-polling technique in which, like in the standard polling technique, the user application sends a request for information to the ESB and, if the platform has some, it returns it immediately. However, if the platform does not have any information for the user ready, it holds the request and makes the user application wait until new information is available.

All these software components have been integrated and a wide range of tests and trails is being performed using the internal project test-bed. In a second step, test-beds will be deployed in hospitals and schools to perform trails with real users. The goal is to evaluate the open platform concept and encourage other videoconferencing manufacturers and software developers to adapt their agents, to be considered Snake compatible. Some public demos have been presented including ones at the Celtic Event in Paris in March 2009 and the Research for Industry fairs in Poznań in June 2009.

6. Conclusions and Business Benefits

Our analysis of user needs and usage scenarios has provided an overview of the technical and non-technical problems faced while designing a HD videoconferencing platform. Our

study shows that people working in healthcare and education need an affordable HD videoconferencing platform enhancing collaborative work. Our solution is a secure, scalable and expandable system that can provide services tailored for specific user needs.

We believe that future tests and demos performed in hospitals, universities and schools will give us the opportunity to encourage others to use HD videoconferencing technology over the Internet and promote the role of Europe as a leading technological player. It provides business benefits in healthcare and education as a cost efficient solution with better quality of the videoconference equipment enabling new healthcare and education scenarios. Reducing travelling between, for example, hospitals also has a positive environmental effect. Moreover, the fact of being an open and scalable platform allows third party organisations to deploy new services on it with no big efforts. This becomes an opportunity for SME to offer services within a big platform, since they have no need to deploy the whole system, and for final users, who will have the opportunity to use these new features.

Not only schools and hospitals are interested in the results of the HDVIPER project. In addition, other European project expressed interest for cooperation. The integration of the HDVIPER BoD service with the AMPS is a proof of concept for both of them. On one hand, the BoD service has a dynamic QoS system to use and, on the other hand, the AMPS has now a specific application that is in need of the innovative QoS approach it delivers. The Feasibility Study for African-European Research and Education Network Interconnection (FEAST) project is a study contracted by the European Commission (EC) to assess the feasibility of the AfricaConnect project, which works with interconnecting African research and education networks, and connecting them to the European GÉANT network. The representatives of the FEAST project expressed interest in having the HDVIPER platform as the HD videoconferencing tool used in their study.

Finally, HDVIPER is the first project to propose a new innovative approach in the videoconferencing market, offering business opportunities for service companies and showing that three different videoconferencing solutions can be interoperable and use common services through the new platform.

References

[1] *High Definition Videoconferencing over IP Environment (HDVIPER)*. Available on-line at [http://www.hdviper.org]. Last visited on 23-04-2009.

[2] J. Rosenberg *et al*. *SIP: Session Initiation Protocol*. Internet Engineering Task Force, RFC 3261, June 2002.

[3] M. Handley *et al*. *SDP: Session Description Protocol.* Internet Engineering Task Force, RFC 4566, July 2006.

[4] J. Arkko *et al*. *MIKEY: Multimedia Internet KEYing.* Internet Engineering Task Force, RFC 3830, August 2004.

[5] M. Baugher *et al*. *The Secure Real-time Transport Protocol.* Internet Engineering Task Force, RFC 3711, August 2004.

[6] *x264 - a free h264/avc encoder*. Available on-line at [http://www.videolan.org/developers/x264.html]. Last visited on 24-04-2009.

[7] *Intel Integrated Performance Primitives*. Available on-line at [http://software.intel.com/en-us/intel-ipp]. Last visited on 23-06-2009.

[8] *Speex: A Free Codec For Free*. Available on-line at [http://speex.org]. Last visited on 24-04-2009.

[9] *Advance Multi-domain Provisioning System, Geant2*. Available on-line at [http://www.geant2.net/server/show/nav.00d008009002]. Last visited on 23-04-2009.

[10] *Push Technology, Long Polling, Wikipedia*. Available on-line at [http://en.wikipedia.org/wiki/Push_technology#Long_polling]. Last visited on 23-04-2009.

[11] *FEAST Project*. Available on-line at [http://www.feast-project.org]. Last visited on 26-04-2009.