MCU basada en SIP para videoconferencias seguras[1]

SIP based MCU for secure video conferences

Guillem Cabrera, Javier López, Xavier Calvo, Antoni Oller, Flaminio Minerva, Jesus Alcober

Resumen

La mayoría de las soluciones comerciales no incluyen la opción de asegurar las videoconferencias

Los sistemas actuales de videoconferencia IP presentan diferentes problemáticas que entorpecen la comunicación a través de ellos. La mayoría están basados en implementaciones propietarias y conllevan costes elevados de adquisición de los equipos. Además, la calidad de los recursos multimedia transferidos por estos sistemas acostumbra a ser bastante baja, cosa que afecta a la experiencia del usuario y a la percepción de presencia. Finalmente, la mayoría de las soluciones comerciales no incluyen la opción de asegurar las videoconferencias.

En este artículo se presenta una solución implementada basada en un servidor de videoconferencias en software de código abierto. Una capa de señalización basada en el protocolo SIP y un plano multimedia, basado en un replicador de paquetes RTP. Esta solución constituye una plataforma que permite disminuir los costes del sistema usando ordenadores personales con clientes SIP en software, permitiendo reducir drásticamente el coste de los equipos.

Finalmente, se propone un mecanismo de protección de los flujos multimedia usando Secure Real-time Transport Protocol (SRTP) y Multimedia Internet Keying (MIKEY) para el intercambio de claves.

Palabras clave: SIP, SDP, MCU, RTP, SRTP, MIKEY, seguridad, videoconferencia

Summary

En este artículo se presenta una solución implementada basada en un servidor de videoconferencias en software de código abierto

Current commercial IP based video conferencing systems present several problems that hold up the communication through them. Most of them are based on proprietary implementations with a high price. Moreover, the media quality offered by these systems is usually low, so the user experience and the presence feeling decrease.

Eventually, most of these solutions can not provide any security to secure the media in a conference.

In this article, it is presented an implementation based in an OpenSource software video conferencing server. This software includes a SIP signalling layer and a RTP packet reflector in the media layer. These solutions constitute a cheap video conferencing platform, since low-end computers using software SIP phones are used in the client side.

Finally, a media protection mechanism based in Secure Real-time Transport Protocol (SRTP) and Multimedia Internet Keying (MIKEY) as a key exchange protocol.

Keywords: SIP, SDP, MCU, RTP, SRTP, MIKEY, security, video conferencing

1 *Este trabajo ha sido parcialmente financiado por el Ministerio de Industria, Turismo y Comercio y por el CIDEM (Generalitat de Catalunya) en el marco del proyecto HDVIPER (www.hdviper.org) y por el MCyT (Ministerio de Ciencia y Tecnología del Gobierno de España) bajo el proyecto TSI2007-66637-C02-01.*

1. Introducción

En un mundo global, con colaboraciones entre empresas e instituciones de alrededor del mundo y con Internet implantada como una herramienta de uso cotidiano, las videoconferencias se han convertido en una herramienta básica para las comunicaciones entre distintas ubicaciones.

El uso de las videoconferencias representa un claro ejemplo del beneficio de la tecnología a sociedad. Su uso se extiende a entornos educativos (e-Learning), médicos (telemedicina) y de negocios, evitando la necesidad de desplazarse para el desarrollo de la actividad. Esta característica ha permitido reducir los viajes, lo que supone una contribución al cuidado del medio ambiente. Pensando a nivel empresarial, se ahorra tiempo, recursos y dinero a la vez que se mejora la eficiencia y la calidad de vida de los trabajadores. Además acelera la toma de decisiones. Por otro lado, la videoconferencia está permitiendo el acceso a expertos desde lugares remotos, como médicos especialistas o profesores de renombre.

Las videoconferencias se han convertido en una herramienta básica para las comunicaciones entre distintas ubicaciones

Si a estos beneficios le añadimos el rápido crecimiento de las tecnologías de Internet en la última década, nos encontramos con que grandes fabricantes comercializan equipos de fácil uso para el gran público que usa las redes IP para la transmisión de los datos.

Este trabajo se ha organizado de la siguiente manera. En el primer apartado se trata la situación actual de la videoconferencia y se analizan las limitaciones que presentan. En segundo lugar, se presenta la solución planteada, entrando en detalles con la señalización utilizada y el plano multimedia. Finalmente, se discute cómo asegurar este tipo de sesiones y qué estrategia se ha implementado en la propuesta presentada.

2. Estado actual

A día de hoy el uso de la videoconferencia está limitado a la comunicación entre sedes de grandes instituciones, ya que disponen de equipos compatibles, con redes propias y controladas. Estos detalles se unen a otras barreras como el precio, que limitan la plena expansión de la videoconferencia.

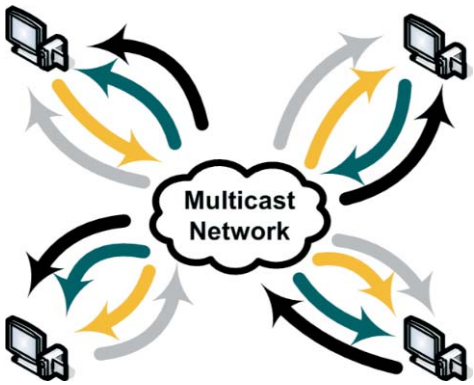
El coste de estos equipos suele ser elevado y dificulta su presencia en entornos educativos o pequeñas empresas

En primer lugar, la mayoría de sistemas están basados en implementaciones propietarias de los grandes fabricantes. Una gran parte de estas soluciones no son plenamente compatibles entre ellas, perdiendo muchas funcionalidades al establecer comunicaciones entre equipos de diferentes fabricantes. Añadir que el coste de estos equipos suele ser elevado y dificulta su presencia en entornos educativos o pequeñas empresas.

En segundo lugar, las videoconferencias tratan de recrear un encuentro físico, pero esto se suele romper por la escasa sensación de presencia que ofrecen la mayoría de soluciones, representando una barrera en la comunicación. Para mejorar la experiencia del usuario, las nuevas tecnologías audiovisuales ofrecen herramientas como multitud de canales de audio y el uso de vídeos de alta calidad, alta definición (High Definition) e incluso 3D que permiten tener un interlocutor a tamaño real, apreciar detalles de gesticulación y un sonido directivo que sumergen a los participantes en la reunión, olvidando la tecnología. Otros actos típicos en una reunión, como anotaciones en una pizarra o la lectura de un mismo documento se pueden solventar con la inclusión de recursos extra, como pizarras digitales o escritorios compartidos.

En tercer lugar, la falta de implantación de multicast en Internet representa un obstáculo para videoconferencias multipunto. En este escenario, los flujos multimedia generados por cada usuario deben ser entregados a todos los demás participantes en la reunión (Figura 1).

FIGURA 1. MULTICONFERENCIA EN RED MULTICAST



Finalmente, el bajo o nulo nivel de seguridad que soportan los sistemas actuales hace que no sean del agrado de las empresas e incluso de los médicos. Al usar Internet como canal de transporte de la información, cualquier observador podría ser capaz de recibir y reproducir los flujos multimedia correspondientes a una videoconferencia. Algunos sistemas incorporan mecanismos de seguridad, pero, de nuevo, solo son compatibles en combinación con sistemas del mismo fabricante.

El bajo o nulo nivel de seguridad que soportan los sistemas actuales hace que no sean del agrado de las empresas e incluso de los médicos

3. Propuesta Implementada

La solución propuesta [4] intenta mitigar algunos de los problemas descritos anteriormente. Consiste en un servidor de videoconferencias (Multipoint Conference Unit, MCU) en software de código abierto y diseñado arquitectónicamente en capas: una capa de señalización basada en SIP (Session Initiation Protocol) [1] y un replicador de paquetes RTP (Packet Reflector) [2] en el plano de media. Esta solución no precisa de un hardware específico, sino que está pensada para funcionar en máquinas de propósito general.

La elección de SIP como protocolo de señalización no es casual

3.1. Capa de Señalización

La capa de señalización consiste en un agente SIP (Server Agent) que negocia los parámetros de la sesión multimedia con los clientes usando SDP (Session Description Protocol) [3] . La necesidad de seguir estándares para estos fines es vital para la interoperatibilidad entre diferentes sistemas y huir de implementaciones propietarias.

La elección de SIP como protocolo de señalización no es casual. El crecimiento experimentado en los últimos años por los sistemas de voz sobre IP (VoIP) ha generalizado el uso de los programas que utilizan SIP y SDP como protocolos de señalización (establecimiento, mantenimiento y cierre de sesiones multimedia). La gran mayoría de estos clientes soportan también vídeo, haciéndolos una herramienta ideal para su uso en videoconferencias. Este tipo de software es cada vez más corriente en las computadoras personales y está presente en la mayoría de instituciones y hogares. Esto hace que se disponga de un cliente de videoconferencias de bajo coste y al alcance de cualquier posible usuario.

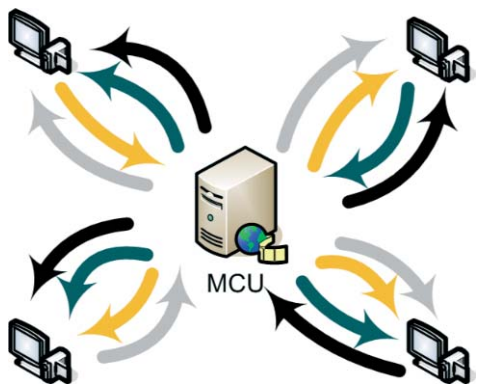
El servicio incorpora una gestión mínima de salas, que permite ofrecer la posibilidad de diversas videoconferencias simultáneas. Esta funcionalidad permite crear eventos y permitir o denegar el acceso de diferentes SIP URIs a ellas.

Para el desarrollo de la capa de señalización SIP se utilizaron las librerías de MiniSIP [5] , un softphone OpenSource. Además, su implementación basada en librerías permite construir nuevas herramientas SIP utilizando solo las partes necesarias del cliente SIP completo.

3.2. Capa de Media

El plano de media de la MCU se basa en un replicador de paquetes RTP. Esta herramienta permite la recepción de flujos de paquetes UDP (User Datagram Protocol) para ser reenviados a múltiples destinos, de manera que se definen unas rutas tal y como se hace en los equipos de enrutamiento IP. Su funcionamiento intenta emular el comportamiento de una red multicast sin que éste sea soportado (Figura 2).

FIGURA 2. REPLICADOR DE PAQUETES



Las MCUs convencionales suelen recibir los flujos multimedia (audio y vídeo generalmente) de los participantes y generan un nuevo flujo compuesto que es enviado a los interesados. En cambio, la propuesta de este trabajo es el uso de un replicador de paquetes. Los beneficios que aporta esta solución son comentados a continuación.

La propuesta de este trabajo es el uso de un replicador de paquetes

La primera y más importante es la independencia total del tipo de media o datos que manden los clientes, pues no hay ningún tipo de transcodificación o procesado de los datos. Este hecho permite que los clientes puedan mandar cualquier tipo de media y no solo los soportados por la MCU. Con esto, es posible realizar conferencias usando codificadores propios, vídeo 3D o incluso mandar tipos de media heterogéneos.

Otra ventaja clave es que al no haber ningún tipo de tratamiento sobre los flujos, el servicio puede ser prestado desde un equipo con relativamente baja capacidad computacional funcionando bajo Linux y sin necesidad de un equipo diseñado específicamente para ello, suponiendo un ahorro considerable de costes.

Finalmente, y ligado también con la ausencia de codificación, el retardo añadido por el uso de este sistema es mínimo, cosa que favorece el desarrollo de una comunicación fluída.

Las características listadas del replicador de paquetes permiten salvar los puntos críticos de toda conferencia

Las características listadas del replicador de paquetes (independencia de codec, no transcodificación, bajo coste y bajo retardo) permiten salvar los puntos críticos de toda conferencia. No tratar los datos implica que el retardo añadido es mínimo y ayuda a mantener la interactividad. Por otro lado, implica que se puede enviar cualquier tipo de codec, limitación que las implementaciones actuales presentan al tener que transcodificar la media. Esto no les permite transmitir altas calidades y la percepción del usuario se ve reducida. En la solución propuesta, la MCU no representa un impedimento para que emisor y receptor puedan intercambiar flujos de vídeo de muy alta calidad, al ser capaz de gestionar un gran volumen de datos. Finalmente, el uso de equipos no específicos lo convierten en una buena solución para PYMES, permitiéndoles tener un servicio sin un equipo dedicado y sin una gran inversión.

Los resultados de las pruebas iniciales realizadas sobre este sistema nos permiten afirmar que es capaz de gestionar unos 15 flujos de 20Mbps, gestionando de forma global más de 300Mbps entre tráfico

entrante y saliente. Estos números representan un incremento de unos 2 órdenes de magnitud (en términos de ancho de banda gestionado) respecto a los sistemas de multiconferencia actuales.

La implementación del servidor de videoconferencias se ha realizado íntegramente en C++, pues es un lenguaje lo suficientemente flexible para el desarrollo de aplicaciones complejas y a la vez ofrece un buen rendimiento.

3.3. Caso de uso

Para unirse a una multiconferencia usando el servicio ofrecido por la MCU, un cliente debería llamar a la SIP URI de la MCU. Además, para indicar la reunión a la que quiere unirse, debería añadir un parámetro indicando el nombre de la sala. La MCU negociaría los parámetros de la sesión multimedia con SDP (Figura 3) y, en caso de llegar a un acuerdo, añadiría el usuario a la reunión.

La implementación del servidor de videoconferencias se ha realizado íntegramente en C++

En este momento la MCU guarda la información (direcciones IP, puertos y tipos de media) correspondiente a la sesión multimedia establecida con la SIP URI alice@i2cat.net.

Cuando otro usuario (bob@i2cat.net) accede a la misma sala, la MCU negocia con él otra vez los parámetros de la sesión y los almacena (Figura 4).

Para unirse a una multiconferencia usando el servicio ofrecido por la MCU, un cliente debería llamar a la SIP URI de la MCU

En este momento, la capa de señalización se comunicaría con el plano de media, añadiendo las rutas necesarias para hacer llegar los flujos de media enviados por alice@i2cat.net a bob@i2cat.net y los enviados por bob@i2cat.net a alice@i2cat.net (Figura 5).

Si un tercer participante entrara a la sala, el diálogo SIP con la MCU sería el mismo que en los casos anteriores. En la comunicación con el plano de media, los flujos ya existentes serían reenviados hacia el nuevo participante y los flujos del nuevo participante se reenviarían a los demás participantes de la multiconferencia.

FIGURA 3. DIÁLOGO SIP Y SDP

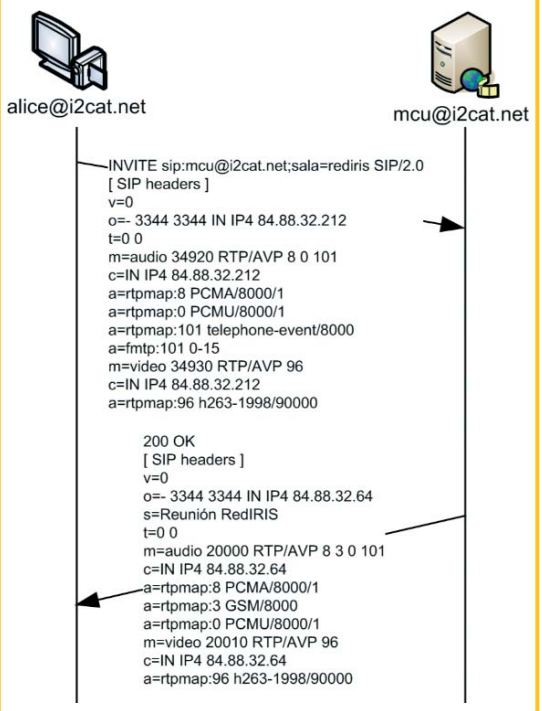


FIGURA 4. DIÁLOGO SIP Y SDP

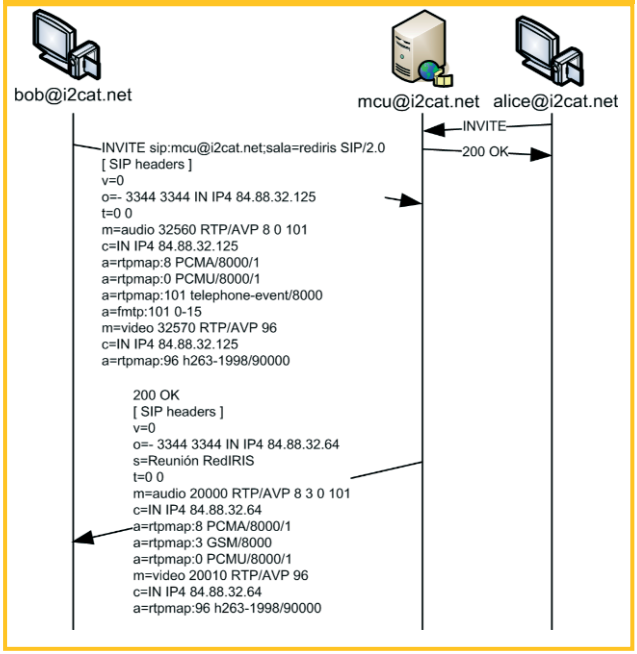
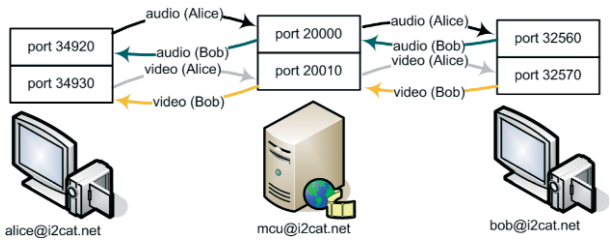


FIGURA 5. FLUJOS REENVIADOS EN EL REPLICADOR



Destacar que en este último caso los clientes SIP recibirían varios flujos de cada tipo de media (2 de audio y 2 de vídeo) en el mismo puerto, con lo que deberían ser capaces de diferenciar los flujos multimedia a partir del campo SSRC de los flujos RTP para poder reproducirlos.

3.4. Multiconferencias seguras

Para solventar los problemas de seguridad derivados del uso de redes públicas para la transmisión de los flujos multimedia correspondientes a una multiconferencia, se aprovechan las posibilidades que ofrece el protocolo Secure Real-time Transport Protocol (SRTP) [6] .

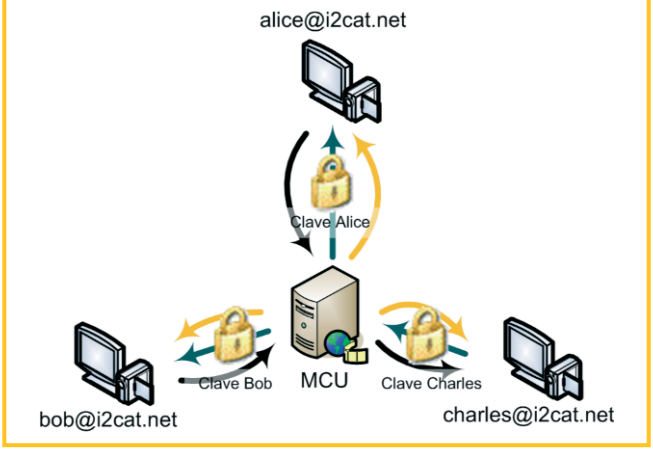
El protocolo SRTP permite autenticar y cifrar flujos de datos RTP, ofreciendo confidencialidad, integridad y autenticación

Este protocolo permite autenticar y cifrar flujos de datos RTP, ofreciendo confidencialidad, integridad y autenticación. Aún así, este protocolo necesita una clave que debe ser intercambiada entre los 2 extremos de la comunicación. Para este intercambio, SRTP delega en mecanismos externos de intercambio de claves.

Diferentes estudios [7] [8] demuestran que el Multimedia Internet KEYing (MIKEY) [9] representa la mejor opción para la negociación de claves SRTP a través de SIP y SDP, pues consigue autenticar a los clientes y acordar una clave maestra en un solo intercambio de mensajes. A partir de esta clave maestra, SRTP deriva las claves de autenticación y de cifrado que serán usadas para asegurar los flujos RTP.

En la implementación descrita, los clientes intercambian la clave maestra con la MCU. Los flujos SRTP son desencriptados en el servidor de videoconferencias y son encriptadas de nuevo con la clave acordada con los destinatarios de dicho flujo (Figura 6).

FIGURA 6. FLUJOS SRTP



Para el desarrollo de la capa de señalización se usaron las librerías que forman MiniSIP

Tal y como se comenta anteriormente, para el desarrollo de la capa de señalización se usaron las librerías que forman MiniSIP. Éstas constituyen la única implementación del protocolo MIKEY, cosa que nos permitió hacer uso de éste en nuestro sistema.

4.Conclusiones

Este artículo propone un nuevo modelo para gestionar un entorno de comunicaciones multipunto seguro. Este esquema proporciona ventajas frente a las soluciones actuales del mercado, al utilizar SIP como protocolo de señalización y permitiendo la utilización de cualquier tipo de codificación.

Por otro lado, cambia la visión de aparatos de videoconferencia caros, rígidos y dedicados hacia una videoconferencia accesible desde cualquier lugar, para todo tipo de organizaciones y con funcionalidades de seguridad incorporadas.

Por tanto, es una herramienta ideal para el trabajo colaborativo entre empresas y usuarios, salvando la distancia que les separa y sin necesidad de grandes inversiones.

Referencias

[1] *J. Rosenberg y otros. SIP: Session Initiation Protocol. Internet Engineering Task Force, RFC 3261, 2002.*

[2] A. Ríos, A. Oller, X. Miguélez, J. López, J. Alcober. Multiconferencia con Video de Alta Calidad usando SIP. A: Jornadas Telecom I+D 2007. TelecomI+D, 2007.

[3] M. Handley; V. Jacobson; C. Perkins. SDP: Session Description Protocol. Internet Engineering Task Force, RFC 4566, 2006.

[4] G. Cabrera, E. Eliasson. Secure High Definition Video Conferencing. i2Cat & Kungliga Tekniska Högskolan, 2008.

[5] MiniSIP (http://www.minisip.org) 06-10-2008

[6] M. Baugher y otros. The Secure Real-time Transport Protocol (SRTP). Internet Engineering Task Force, RFC 3711, 2004.

[7] J. Bilien, Key Agreement for secure Voice over IP, Kungliga Tekniska Högskolan, 2003.

[8] J. Orrblad, Alternatives to MIKEY/SRTP to secure VoIP, Kungliga Tekniska Högskolan, 2005.

[9] J. Arkko y otros. MIKEY: Multimedia Internet KEYing. Internet Engineering Task Force, RFC 3830, 2004.

Xavier Calvo
(xavier.calvo-brugal@upc.edu)
Antoni Oller
(antoni.oller@upc.edu)
Jesús Alcober
(jesus.alcober@upc.edu)

Departamento de Ingeniería Telemática,
Universitat Politècnica de Catalunya

Guillem Cabrera
(guillem.cabrera@i2cat.net)
Javier López
(javi.lopez@i2cat.net)
Flaminio Minerva
(flaminio.minerva}@i2cat.net)

Fundació i2cat