Sistemas de videoconferencia: experiencias en la UPV/EHU

José Miguel-Alonso, Jesús Mª Aramberri Miranda, Javier Lasa Jáuregui, José E. Poza Sanz,

Departamento de Arquitectura y Tecnología de Computadores
Universidad del País Vasco
Paseo Manuel de Lardizábal, 1
20009 San Sebastián
{miguel, acparmij, acremulu, jose}@si.ehu.es

Resumen. Los sistemas de video-reunión (videoconferencia) están abriendo las puertas a nuevas formas de desarrollar la actividad docente universitaria. Hacemos en este artículo una revisión de los medios técnicos para desarrollar estas reuniones, y analizamos una experiencia concreta: la de la Universidad del País Vasco.

1 Introducción

Los sistemas de videoconferencia son aplicaciones *síncronas* (requieren la presencia simultánea de los comunicantes), pensadas para su utilización por *grupos de personas* y diseñadas para facilitar el *trabajo en común* entre los miembros de esos grupos. Aunque se hable de videoconferencia, siempre se entenderá que también se produce un intercambio de sonido, y quizá de otros tipos de información (datos). En cualquier caso, quizá sería más apropiado usar el término *video-reunión*. Un sistema de videoconferencia puede funcionar en diferentes entornos: escritorio a escritorio, escritorio a sala, sala a sala, etc., en función de a cuántas personas preste servicio el sistema en cada lugar participante. La videoconferencia se diferencia de la videofonía en que está diseñada para la comunicación en grupo (no para comunicación entre dos interlocutores[1]), y se diferencia de la videodistribución en que la comunicación es bidireccional.

Un sistema de videoconferencia tiene unas necesidades técnicas importantes. Se necesitan cámaras de TV que, si se usan en sala, tienen que ser capaces de captar a un grupo, no sólo a una persona (opciones: un objetivo gran angular o una cámara móvil). Durante la reunión también será necesario intercambiar documentos,

[1] Muchas veces se llama videoconferencia a una comunicación a dos bandas, utilizando aplicaciones como NetMeeting (Microsoft). Aunque estas aplicaciones tienen cierta capacidad para trabajar en grupo, realmente están optimizadas para su utilización por dos personas, por lo que entraríamos en el ámbito de lo que denominamos “videofonía”.

electrónicos o en papel. Dependiendo de cómo sean esos documentos, habrá que añadirle al sistema elementos como una cámara más (de mayor calidad) enfocada sobre la mesa, una cámara que enfoque una pizarra o pantalla de proyección, un sistema de facsímil (escáner más impresora), o un canal de datos para el intercambio de documentos ya digitales (hojas de cálculo, por ejemplo). Se requiere encontrar un compromiso entre calidad (resolución) y movimiento (cuadros por segundo): suele preferirse más movimiento a costa de menor calidad, excepto cuando se transmiten documentos. Los requisitos en cuanto al sonido son bastante exigentes, puesto que en una reunión tradicional buena parte de la información se transmite de forma vocal. No se admiten cortes, ruidos, ecos, acoples, retornos, etc., lo que exige una buena acústica en las salas, y equipamiento (micrófonos, amplificadores, altavoces, mesas de mezclas) de calidad.

Existen dos formas básicas de ofrecer videoconferencia, aunque las diferencias entre ambas cada vez son más difusas: sistemas basados en conmutación de circuitos, y sistemas basados en conmutación de paquetes:

- Si se usa la conmutación de circuitos, los elementos que forman el núcleo del sistema de videoconferencia se denominan video-codecs. Los video-codecs realizan funciones de codificación/decodificación (paso de sonido y vídeo analógico a digital, y viceversa), compresión/descompresión, acceso al sistema de transmisión, y control del sistema. Están conectados a una línea que ofrece una capacidad de transmisión constante y dedicada (normalmente, una línea RDSI). Inicialmente estaban pensados para comunicación entre salas especiales, aunque con el tiempo fueron apareciendo sistemas portables (*rollabout*) y de sobremesa. Se comercializan en forma de sistemas “llave en mano” que incluyen codificadores, compresores, acceso a la red y control en un sólo paquete. En general el sonido transmitido es de buena calidad, y se ofrecen múltiples mecanismos de control de cámaras. Es relativamente frecuente que sistemas de diferentes fabricantes no interoperen correctamente, aunque la aparición de diversas recomendaciones de la ITU-T ha mejorado considerablemente las cosas—más adelante analizaremos estos estándares.
- La videoconferencia basada en conmutación de paquetes trata de explotar, para dar soporte al trabajo en grupo, la capacidad de los PCs instalados en una organización, así como las redes a las que éstos se conectan (redes de conmutación de paquetes, típicamente basadas en los protocolos TCP/IP). Inicialmente los sistemas estaban ideados para PCs de sobremesa, aunque van apareciendo sistemas para salas, especialmente para comunicación asimétrica (el conferenciante está en su despacho y los asistentes en una sala, o bien se da una conferencia en una sala y algunos usuarios participan de forma remota desde sus PCs). Se ve, por lo tanto, cómo los sistemas de conmutación de circuitos y los de conmutación de paquetes van ocupando las mismas zonas del mercado. Los sistemas de captura de vídeo no suelen ser de alta calidad: el control de la cámara es limitado, y la resolución también. Algo similar ocurre con el sonido. Lo mejor de estos sistemas es la excelente integración de las herramientas de audio y vídeo con otras herramientas como las pizarras electrónicas y aplicaciones compartidas,

ofreciendo un entorno completo para las video-reuniones. Si el equipo utilizado lo permite, se pueden añadir funciones como grabación de sesiones (para su posterior reproducción), lo que da pie al uso de herramientas adicionales para la edición, manipulación y reproducción de secuencias previamente almacenadas, así como a la integración con servidores de vídeo. El sistema de videoconferencia puede *alimentar* al servidor de vídeo, o puede actuar como cliente usando secuencias grabadas. Como último comentario, decir que la mayor parte de los sistemas de videoconferencia sobre redes de paquetes son incompatibles entre sí, y que el proceso de estandarización avanza aquí bastante lentamente.

Un aspecto clave en la videoconferencia es la forma de dar soporte a reuniones multi-parte. En el caso de conmutación de circuitos se necesita emplear un sistema central, denominado vídeo-concentrador o *unidad de control multipunto* (MCU) al cual se conectan los participantes. Es decir, el funcionamiento multi-parte está basado en múltiples conexiones punto a punto. Los vídeo-concentradores pueden ser unos conmutadores analógicos (lo que los hace independientes del sistema de videoconferencia) o digitales (normalmente asociados a un fabricante). Estos concentradores suelen imponer un número máximo de participantes (típicamente, ocho). Los operadores de telefonía suelen alquilarlos, aunque también pueden comprarse. Los vídeo-concentradores suelen activarse por voz: el canal de vídeo a retransmitir a todos los participantes en cada instante es el del interlocutor que en ese momento tiene la palabra.

En el caso de la videoconferencia sobre paquetes, la comunicación en grupo se puede basar en uno de dos modelos distintos: el uso de múltiples conexiones punto a punto, o el uso de técnicas multidestino (*multicast*). En el primer caso, un paso previo a una videoconferencia es la conexión de los participantes al concentrador, usando para ello canales de comunicación punto a punto tradicionales (por ejemplo, IP "convencional"). Un problema evidente es la posible existencia de varios canales circulando por un mismo segmento de la red, todos ellos transportando la misma información. Como ventaja podemos citar la posibilidad de incorporar sistemas para garantizar la confidencialidad: sólo aquellos sistemas autorizados pueden conectarse al concentrador.

La segunda alternativa aprovecha las capacidades multidestino de algunas redes de datos para optimizar el flujo de información, como es el caso de IP (IP *multicast*). Sin embargo, esta alternativa presenta como inconveniente la imposibilidad (actualmente) de establecer grupos cerrados[2]. Otro tema de importancia es el control del turno: no siempre es necesario transmitir flujos de datos desde aquellos lugares que no están aportando información. Los sistemas más primitivos ofrecen un control manual: los usuarios activan/desactivan manualmente los flujos de audio, vídeo, etc. En ocasiones esto es misión de un moderador. En algunos casos el sonido no es objeto de este control, ya que no genera un tráfico elevado (especialmente si se usan sistemas que suprimen el envío de datos durante los períodos de silencio).

[2] Se puede lograr una aproximación a los grupos cerrados, empleando técnicas criptográficas— lo que añade una considerable carga de trabajo al sistema.

Una ventaja clara de los sistemas apoyados en tecnologías multidestino es que el coste no crece de forma lineal con el número de participantes. Por otra parte, la proliferación de videoconferencias sobre una red corporativa puede exigir una actualización (mejora) de los elementos que la componen.

En cuanto a las redes de datos empleadas, la videoconferencia sobre circuitos se diseñó para trabajar sobre redes públicas de conmutación de circuitos, como la mencionada RDSI, o también sobre las líneas punto a punto alquiladas que constituyen el esqueleto de una red corporativa. Con el paso a la RDSI de banda ancha, basada en ATM (que también da servicio de emulación de circuitos) se puede trabajar de la misma manera, aunque con mucha mejor calidad (cercana a la de la TV).

Las LAN más extendidas hoy en día son, en general, suficientes para dar soporte a videoconferencias sobre paquetes, con una calidad media: se requieren aproximadamente 100 Kb/s para el vídeo, más 13-64 Kb/s para el sonido. Algunos de estos sistemas de videoconferencia explotan la capacidad multicast de las LAN, pero no siempre es así. Si pasamos a redes de paquetes de área amplia, recordemos que conviene usar aquellas que ofrecen multidestino, como es el caso de las redes basadas en IP.

El mercado de la videoconferencia está experimentando un fuerte crecimiento (se estima un incremento del 50% anual), especialmente en los sistemas de sobremesa. Existen varias compañías dedicadas en exclusiva a este segmento, y grandes empresas como Intel y Microsoft están entrando con fuerza en él.

2 Videoconferencia basada en estándares ITU-T

El estándar de más éxito en el ámbito de la videoconferencia es el H.320[3], que define una metodología para tranportar tráfico de videoconferencia sobre la red digital de servicios integrados (RDSI) con una calidad aceptable para comunicaciones de empresa. Recientemente han ido apareciendo otros estándares: H.323, H.321, H.324 and H.310, que indican cómo implementar videoconferencias con diferentes calidades sobre medios de transporte alternativos:

- H.321. Videoconferencia sobre ATM, calidad media-alta.
- H.323. Videoconferencia sobre IP/Ethernet, calidad media-baja.
- H.324. Videoconferencia sobre la red telefónica básica, calidad baja
- H.310. Videoconferencia MPEG-2 sobre ATM, calidad alta

El resultado de este esfuerzo de estandarización es una colección de especificaciones que cubren prácticamente todos los niveles de calidad, y que facilitan en gran medida la interoperabilidad entre sistemas de diferentes organizaciones.

[3] En realidad, la ITU-T no produce “estándares”, sino “recomendaciones”.

2.1 H.320

La recomendación H.320 ha sido clave para posibilidar la interoperación, usando la RDSI, entre equipos de videoconferencia de diferentes fabricantes. Esta recomendación define aspectos importantes de una videoconferencia, pero muchas veces se la define cono un “estándar-paraguas”, porque incluye multitud de referencias a otras recomendaciones que deben emplearse en diferentes aspectos de la videoconferencia, como son la codificación del sonido, la codificación de la imagen, el multiplexado de la información multimedia, y el control del sistema.

La recomendación básica para la codificación del sonido es la G.711 (sonido PCM con un consumo de 64 Kb/s), aunque se permiten otras con diferentes calidades y consumos de ancho de banda: G.722 (que ofrece más calidad que G.711 con los mismos 64 Kb/s) y G.728 (calidad similar a G.711, pero usando sólo 16 Kb/s)

Para la codificación y compresión del vídeo se especifica la recomendación H.261, que puede trabajar con un ancho de banda que oscila entre los 64 Kb/s y los 2 MB/s, en múltiplos de 64 Kb/s.

En lo que se refiere a conferencia de datos (transferencia de ficheros, pizarra compartida, compartición de aplicaciones, conversación mediante texto, etc.), H.320 incluye una referencia a la recomendación T.120, que a su vez incluye referencias a otra colección de recomendaciones para facetas concretas de la conferencia de datos. En el apéndice 1 podemos encontrar una lista de algunas de estas recomendaciones.

2.2 H.321

Esta recomendación define cómo adaptar los terminales H.320 para que trabajen en entornos RDSI de banda ancha—es decir, ATM—sin perder la capacidad de interoperar con sistemas H.320. Por esta razón, sigue especificando el uso de los estándares G.711 para el sonido y H.261 para el vídeo. La ineroperabilidad impuesta impide que un sistema basado en H.321 aproveche al máximo el potencial de ATM. Por una parte, se impone que el ancho de banda máximo a utilizar para el vídeo sea 2 MB/s (el límite de H.261). También se impone el uso del AAL1 (servicio CBR—tasa de bits constante), lo que es una limitación, porque un sistema de videoconferencia podría aprovechar de forma óptima un servicio VBR (tasa de bits variable), dadas las características del tráfico que genera.

2.3 H.310

En 1996 la ITU-T aprobó la recomendación H.310, que define sistemas y terminales para llevar a cabo comunicaciones audiovisuales sobre ATM—incluyendo videoconferencias, pero también otros métodos de comunicación. Se definen terminales unidireccionales y bidireccionales, y también se clasifican los terminales en función del nivel de adaptación empleado (AAL1 o AAL5).

Aunque H.310 especifica H.321 como una posible forma de trabajar, define también un modo ATM nativo. Se mantienen las especificaciones H.261 y G.711 como codificaciones básicas (por motivos de compatibilidad), pero se introducen otras nuevas, en particular MPEG. Esta especificación incluye mecanismos de codificación y compresión tanto para audio como para vídeo, con una excelente calidad.

La arquitectura de H.310 es bastante compleja; incluye cinco pilas de protocolos, con múltiples niveles (recomendaciones) para cada pila:

- Pila de señalización fuera de banda
- Pila de señalización en banda H.245—señalización de canales y gestión de las capacidades de los terminales
- Pila para la interoperación H.320/H.321 utilizando multiplexado multimedia H.221
- Pila nativa H.310 utilizando multiplexado multimedia H.222.1/H.222.0
- Pila T.120 para aplicaciones de datos.

2.4 H.323

Esta recomendación define los sistemas de videoconferencia sobre redes locales e Internet. De hecho, la especificación es aplicable a cualquier red de conmutación de paquetes, independientemente del medio físico de transmisión que se emplee. Especifica los ya conocidos formatos H.261, G.711 y T.120 para el vídeo y el audio y los datos, por motivos de compatibilidad (entre otros). Como formatos alternativos, se citan el H.263 para vídeo, y los G.722, G.723 y G.728 para audio.

El intercambio de información de manera fiable (por ejemplo, los datos y la información de control) se realiza usando protocolos como TCP. Sin embargo, la transmisión de paquetes de sonido y vídeo se realiza con protocolos no fiables, como UDP, para evitar la sobrecarga asociada a las retransmisiones en caso de error.

Puesto que es posible transmitir tráfico IP sobre ATM, esta recomendación también es aplicable a la RDSI de banda ancha. El uso de ATM es totalmente invisible al sistema de videoconferencia. Sin embargo, esta forma de trabajar es poco eficiente: sería mejor utilizar directamente el AAL5 para transportar los canales de sonido y vídeo. Al fin y al cabo, AAL5 da un servicio de conmutación de paquetes. Además, el uso nativo de ATM AAL5 puede permitir la selección de calidad de servicio en los circuitos virtuales.

3 Mbone y la videoconferencia en Internet

Desde hace ya bastantes años se ha visto que Internet puede convertirse en una excelente plataforma para la realización de video-reuniones. El comportamiento “best effort” de IP no es el más indicado para las aplicaciones isócronas, pero conforme la tecnología avanza (las velocidades de transmisión aumentan, mejoran los algoritmos de compresión, se desarrollan protocolos que minimizan el impacto de las variaciones en el retardo, etc.) Internet y en general las redes TCP/IP están siendo seriamente consideradas como alternativa viable (y económica) a la RDSI.

Una de las ventajas de las redes IP es el soporte nativo para comunicación multicast. Mientras que la RDSI funciona sólo en modo punto a punto, y para tener una auténtica reunión (con más de dos interlocutores) se precisa de una unidad de control multipunto, en IP es posible formar grupos constituidos por varias máquinas, y representados mediante una dirección IP especial (una dirección multidestino). De esta forma, la infraestructura de la red se encarga de que un datagrama enviado a esa dirección llegue a todos los miembros del grupo.

El funcionamiento de IP multicast exige que la tecnología subyacente lo soporte. Las redes locales lo hacen de forma nativa. Para que dos "islas multicast" se conecten, se necesita que los encaminadores reconozcan la comunicación multicast o, alternativamente, establecer un "túnel" entre ambas islas—que no es otra cosa que un sistema para encapsular datagramas multicast dentro de datagramas normales y encaminarlos de la forma convencional. En cada extremo del túnel se sitúa un *mrouter*, o encaminador multidestino.

La red Mbone (Multicast Backbone) consiste en una colección de islas multicast unidas mediante túneles, establecidos a través de Internet. Se dice que es una red virtual, porque no tiene infraestructura propia. Si una organización quiere integrarse en Mbone, lo único que requiere (a nivel técnico) es establecer un túnel con una organización que ya esté integrada. En estos momentos la cobertura de Mbone es prácticamente global.

Una vez que se cuenta con la infraestructura, se necesitan aplicaciones para sacarle provecho. Las aplicaciones desarrolladas para Mbone permiten la realización de reuniones virtuales con una gran riqueza de medios, aunque quizá con una deficiencia importante: se trabaja con múltiples aplicaciones separadas, sin integración entre ellas. Diferentes grupos de investigación en el mundo (incluyendo el nuestro en la UPV/EHU) están trabajando en esta integración.

Alrededor de Mbone se han desarrollado aplicaciones para la transmisión de vídeo (vic), para la transmisión de audio (vat, rat), para edición cooperativa de textos (nte), para realización de encuestas, etc. Merece la pena destacar sdr, el directorio de sesiones, quizá la aplicación clave de Mbone. Sdr permite a los usuarios anunciar los eventos que se van a difundir por la red, y consultar los que están en marcha o están planificados. Parte de la información asociada a un evento es la colección de flujos de información que lo componen (normalmente sonido+vídeo+pizarra); pues bien, sdr es capaz de abrir automáticamente las aplicaciones requeridas, para que al usuario le resulte sencillo incorporarse a la sesión.

Por supuesto, las aplicaciones de Mbone no funcionan sólo sobre esta red. En realidad, cualquier red con soporte multicast (una red local, una colección de redes locales unidas mediante encaminadores multidestino) puede utilizar estas aplicaciones, que en la mayor parte de los casos son de libre distribución.

4 La situación de la UPV/EHU

En esta sección vamos a ver qué características tiene la UPV/EHU para que la implantación de lo que llamamos un campus virtual: un sistema para la asistencia a cursos ofrecidos en un campus desde cualquier otro. También veremos con qué infraestructura técnica se cuenta para poder abordar el proyecto.

4.1 Estructura de la UPV/EHU

La Universidad del País Vasco / Euskal Herriko Unibertsitatea es la única universidad pública de la Comunidad Autónoma del País Vasco. Los grandes números dicen que en la UPV/EHU hay unos 3.300 profesores, 65.000 alumnos, y tres campus principales, además de una serie de centros fuera de los campus. Se trata de una universidad tradicional, orientada a la enseñanza "presencial".

Una de las directrices organizativas de la UPV/EHU es evitar, en la medida de lo posible, la duplicación de su oferta docente en los diferentes campus. El resultado de esta decisión es que la oferta de titulaciones se dispersa por la geografía de la CAPV, y muchos alumnos tienen que optar por:

1. elegir una universidad privada que funcione en su territorio,
2. trasladar su lugar de residencia hasta el lugar de estudio,
3. elegir una titulación diferente a la deseada, o
4. realizar todos los días el trayecto entre su domicilio y el centro de estudios.

Las dos primeras opciones plantean problemas serios, fundamentalmente desde un punto de vista económico. Algunos alumnos no acceden a las universidades privadas porque cobran tasas más elevadas que las de la pública, por no haber sido admitidos en ellas, o por motivos ideológicos. En lo referente al traslado de domicilio, la UPV/EHU no dispone de una oferta amplia de residencias o colegios mayores, y el acceso a viviendas en alquiler es especialmente costoso en la CAPV. La tercera opción resulta bastante habitual, aunque su motivación suele venir de las limitaciones de matrícula en los centros con más demanda.

El resultado de todo esto es que un elevado número de alumnos recorren todos los días distancias de cientos de Km. En este contexto, la incorporación de un campus virtual a la UPV/EHU puede ser especialmente útil.

Por otra parte, los nuevos planes de estudio que se han implantado (o se están implantando) en todas las universidades españolas permiten a los alumnos la realización de asignaturas optativas de libre elección. Disponer del campus virtual permitirá a los alumnos de la UPV/EHU acceder a asignaturas de otros centros.

Otra razón más para la implementación de estas tecnologías es la optimización de recursos humanos. En ocasiones resulta difícil alcanzar la "masa crítica" para la realización de ciertas actividades, como por ejemplo la apertura de una asignatura optativa o la impartición de un curso de doctorado. La existencia del campus virtual facilita sobremanera la consecución de un número apropiado de participantes en una actividad. Incluso es posible llegar a acuerdos con otras Universidades para realización de actividades conjuntas.

Por último, y desde un punto de vista de gestión, es importante señalar que la actual estructura distribuida de la UPV/EHU exige frecuentes reuniones de trabajo correspondientes a tareas diferentes de la docencia: equipo rectoral, comisiones de contratación, comisión de investigación, etc. Como en cualquier empresa, la disponibilidad de sistemas de video-reunión supone un importante ahorro en costes directos (kilometraje, dietas) e indirectos (tiempo empleado en el transporte).

4.2 La red de servicios telemáticos de la UPV/EHU

Desde el punto de vista de la tecnología, nuestra universidad abordó hace ya tiempo, y con recursos abundantes, la instalación de una red de servicios telemáticos. En los años 1990-1993, y por iniciativa de la SPRI[4], se instaló una red de comunicaciones avanzadas. Esta red corporativa, basada en Ethernet y con una red troncal de tecnologías FDDI y Frame Relay, ha servido también para familiarizar a un elevado colectivo de docentes e investigadores con las aplicaciones telemáticas.

Recientemente esta infraestructura ha experimentado una importante remodelación, especialmente en su red troncal. Aunque los equipos de los usuarios siguen estando basados en Ethernet a 10 Mb/s, la red troncal ha pasado a usar tecnología ATM, con enlaces entre campus a 34 Mb/s.

5 La solución instalada

Partiendo de la situación descrita en la sección anterior, el rectorado de la UPV/EHU decidió abordar el proyecto del campus virtual partiendo de estos principios básicos:

1. Utilización desde salas preparadas al efecto.
2. Sistema basado en la red corporativa ya implantada (es decir, sin nuevos costes en infraestructura de comunicaciones.
3. Calidad comparable a la de la TV de difusión.

En el momento en el que se inició el proyecto (principios de 1998) el mercado de productos de videoconferencia sobre ATM no estaba suficientemente maduro. Esto quiere decir que no existía una gama significativa de productos, ni propietarios ni estándar (acordes a las recomendaciones de la ITU-T antes mencionadas). Por diversas razones, se optó por una solución propietaria de Fore, empresa que había suministrado los equipos para la red troncal.

Realmente, la solución adoptada está a mitad de camino entre un sistema de TV por circuito cerrado y un sistema de vídeo-reunión, aunque ciertamente cumple

[4] Sociedad para la Promoción y Reconversión Industrial, S.A., una sociedad pública dependiente del gobierno de la Comunidad Autónoma del País Vasco.

adecuadamente sus funciones. Las piezas clave son las aulas de videoconferencia, que de momento son cuatro, ubicadas en estos centros: ETSII e IT en Bilbao, Facultad de Informática en San Sebastián, Facultad de Ciencias en Leioa y Facultad de Farmacia en Vitoria (Figura 1). El número de aulas crecerá en el futuro: ya está prevista una más en el Edificio Aulario de San Sebastián. También se prevé la instalación de una sala adicional en cada campus, para la realización de reuniones de carácter no docente.

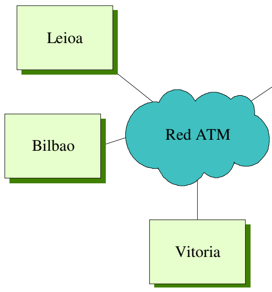


Figura 1. Distribución de las aulas de videoconferencia.

Todas las aulas actuales constan de equipamiento similar: sistemas de captación de vídeo (cámaras, ordenadores, etc.), sistemas de visualización de vídeo (un cañón de proyección y varios monitores), sistemas de captación de sonido, altavoces, mesa de mezclas, matrices de conmutación, etc. La disposición e interconexión de los equipos es la misma en todas las salas, con el fin de poder establecer procedimientos comunes a todas ellas, y de resolver de manera más sencilla y coherente los problemas que puedan surgir. Se trata de pequeños estudios de TV capaces de generar y reproducir sonido y vídeo de calidad semi-profesional (en el apéndice 2 se muestran fotografías de la sala de la Facultad de Informática).

Las piezas clave en cada aula son dos equipos de la casa Fore: un StreamRunner AVA-300 (codificador/compresor que recoge una señal de audio/vídeo y la envía por un circuito ATM) y un StreamRunner ATV-300 (que realiza la operación inversa). La figura 2 muestra un esquema simplificado del aula. Como elementos que generan vídeo[5] tenemos:

- Una cámara que enfoca al conferenciante
- Una o varias cámaras que enfocan al público, es decir, a los alumnos que están en la sala
- Una cámara para proyección de documentos y transparencias
- Un “proyector” de diapositivas
- Un computador personal con salida directa de vídeo
- Un conversor de VGA a vídeo compuesto que permite obtener una señal de vídeo compuesto a partir de la salida de monitor de cualquier computador (por ejemplo, un PC o un Macintosh portátil)

[5] Hablamos de vídeo analógico: vídeo compuesto en formato PAL.

– Un grabador/reproductor de vídeo (VCR)
– Un ATV
– Tomas auxiliares para “inyectar” cualquier entrada de vídeo

En cuanto al sonido se cuenta entre otras fuentes de entrada con diversos micrófonos (conferenciante, ambiente), el computador personal, el VCR, y varias auxiliares.

Las señales de vídeo se dirigen todas ellas hacia una matriz de conmutación que permite asociar entradas con salidas. Es posible seleccionar en todo momento qué señal de vídeo se dirige al AVA. Del mismo modo, es posible seleccionar qué señal se envía al proyector que ven los asistentes a la sala (normalmente, lo generado por el ATV), al monitor que tiene el profesor, al VCR, o a cualquier otro “consumidor” de vídeo.

El sonido es objeto tanto de conmutación como de mezcla, para lo cual se dispone de una mesa de mezclas. Lo habitual es enviar una suma de todas las señales generadas en la sala a través del AVA, y sacar por la megafonía sólo lo recibido por el ATV.

Con la ayuda de una consola de gestión se pueden configurar los circuitos ATM para la realización de comunicaciones entre dos sedes, o para la comunicación entre múltiples sedes. Los equipos ATV pueden realizar funciones de mezcla para mostrar en una misma pantalla varias imágenes de pequeño tamaño y ver así desde una sede lo que está ocurriendo en todas las demás. El número máximo de señales de vídeo que se pueden visualizar simultáneamente es cuatro, lo que limita el ámbito de las reuniones a cuatro sedes (aunque se pueden tener más señales y usar el software de control para seleccionar aquellas que se deseen visualizar).

La situación con el audio es más difícil: los ATV no pueden reproducir más que una señal de audio, lo que obliga a que, cuando se realiza una reunión entre más de dos sedes, la conexión de audio no sea todos-con-todos, sino que cada sede establece un único circuito de sonido (en realidad dos, uno de ida y otro de vuelta) con un equipo que hace de mezclador de flujos de sonido[6]. Así, de forma efectiva todas las sedes reciben una única señal con la combinación de las señales de audio generadas por todas las salas.

[6] Este equipo es en realidad un ATV con un *firmware* modificado.

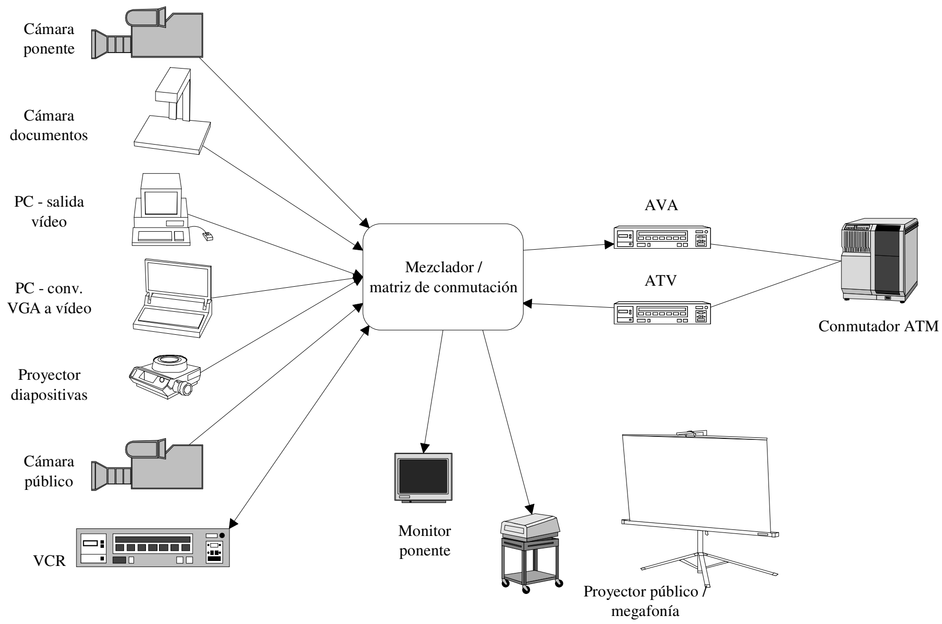


Figura 2. Esquema simplificado del equipamiento de cada aula.

El sistema se controla desde una consola de gestión: una estación de trabajo directamente conectada a la red ATM en la que se ejecuta el software SVA de Fore. Este software incluye (entre otros) los siguientes componentes:

- **Trader**. Gestor global de una video-reunión. Es el que coordina al resto de los procesos.
- **Svamgr**. Gestiona los equipos AVA y ATV, e implementa los sistemas de señalización UNI 3.0 o UNI 3.1 en nombre de estos dispositivos.
- **Svapatch**. Gestiona la conexión entre fuentes (AVAs) y sumideros (ATVs) de corrientes de audio/vídeo.
- **Svc-rtds**. Ofrece un interfaz gráfico de usuario para editar, visualizar y almacenar corrientes de audio/vídeo generadas por un AVA en la estación de trabajo directamente conectada a la red ATM.

En lo que se refiere a las características de los dispositivos AVA/ATV, resumimos algunas en las tablas 1 y 2.

Tabla 1. Características básicas del Fore StreamRunner AVA-300

Característica	Especificación
Interfaz ATM	155 Mb/s MMF (otros disponibles).
Formatos de vídeo	PAL y NTSC.
Entradas de vídeo	6 conectores RCA que pueden configurarse como 6 entradas de vídeo compuesto ó 3 entradas S-vídeo.
Vídeo digital	Hasta 4 corrientes de vídeo digital (seleccionadas de entre las 6 entradas). Formatos sin compresión de 24/16/8 bits por píxel (color) y 8 bpp (b/n). Formato con compresión M-JPEG, con factor de compresión seleccionable por software.
Entradas de sonido	6 conectores RCA, para tres señales estéreo.
Audio digital	Una corriente de sonido digital. Formatos PCM de 8/16 bits, estéreo/mono. Frecuencia de muestreo seleccionable entre 5 kHz y 48 kHz.
Control del ritmo de emisión de celdas ATM	Seleccionable por software, para asegurar que se cumple con el tráfico contratado.
Protocolos ATM	UNI 3.0 y UNI 3.1, realizados por el software SVA.

Tabla 2. Características básicas del Fore StreamRunner ATV-300

Característica	Especificación
Interfaz ATM	155 Mb/s MMF (otros disponibles).
Formatos de vídeo	PAL y NTSC.
Vídeo digital	Descompresión concurrente de hasta 4 corrientes de vídeo digital generadas por dispositivos ATV.
Salidas de vídeo	Una salida de vídeo (duplicada), como vídeo compuesto o S-vídeo, formada por combinaciones de las corrientes de vídeo recibidas.
Audio digital	Recepción de una corriente de sonido digital generada por un dispositivo ATV.
Salida de sonido	2 conectores RCA, para una señal estéreo.
Protocolos ATM	UNI 3.0 y UNI 3.1, realizados por el software SVA.
Configuración	Por software (SVA) y mediante mando a distancia por infrarrojos.

6 Integración de sistemas

Para integrar dos sistemas de videoconferencia distintos es necesario emplear una pasarela conectada a ambos entornos. Las pasarelas más sencillas son analógicas: la salida de sonido y vídeo analógicos de un sistema se utilizan como entradas del otro. Esto añade retardos, y puede afectar bastante a la calidad de extremo a extremo, debido a las distorsiones introducidas por las múltiples conversiones analógico/digital y digital/analógico. La ventaja es la independencia de las tecnologías (especialmente de los sistemas de codificación y compresión) empleadas en los dos sistemas conectados.

Una alternativa es usar pasarelas digitales. Si todos nuestros sistemas están basados en recomendaciones de la ITU-T, y nos limitamos al conjunto de codificaciones comunes para audio/vídeo/datos (G.711, H.261, T.120), la convivencia es posible, y el resultado final es de calidad. Desafortunadamente, esta circunstancia es poco habitual, bien porque se usan productos propietarios, bien porque no se encuentra un subconjunto de protocolos comunes, o bien porque la adhesión a las recomendaciones internacionales no es completa.

En la UPV/EHU, el grupo de investigación en Redes de Computadores se ha propuesto la instalación en el aula de videoconferencia de la Facultad de Informática de un conjunto de pasarelas para interconectar al sistema “principal” (el basado en ATM) con otros alternativos—Mbone, H.320, H.323.

Una estación de trabajo SGI con capacidad de entrada/salida de sonido y vídeo analógico nos sirve para la interconexión con Mbone. La señal recibida en la sala por el ATV (que puede ser una combinación de la generada por todas las salas participantes) se puede inyectar en una sesión de Mbone. La señal recibida desde Mbone y recibida en la estación se puede considerar como una señal de audio/vídeo analógico adicional, y como tal ser seleccionada para su envío por el ATV.

En breve se adquirirá un equipo de la compañía PictureTel para poder interoperar con sistemas de videoconferencia H.320. De nuevo, se usarán señales de sonido y vídeo analógicos como “lengua franca”. El método de interoperación será similar al descrito para Mbone.

La interoperación con sistemas basados en el estándar H.323 ya es posible desde cualquier sede, puesto que en todas ellas se dispone de PCs conectados a Internet mediante un adaptador Ethernet 10BaseT. Estos PCs están dotados de capturadoras de sonido y vídeo (tarjetas de sonido compatibles SoundBlaster, tarjetas de captura de vídeo Miro PC-TV) y generan tanto vídeo digital (VGA) como vídeo analógico (vídeo compuesto). Se está buscando una configuración hardware/software apropiada para que este equipo pueda hacer también las veces de pasarela con Mbone, eliminando así la necesidad de la estación de trabajo antes mencionada.

7 Conclusiones

Las tecnologías multimedia para apoyo a las aplicaciones síncronas de trabajo en grupo han llegado a un punto de madurez tal que las hacen viables para su utilización en entornos empresariales y educativos. Precisamente en este segundo campo es donde grupos de investigación de muchas universidades, entre ellas la UPV/EHU, están centrando sus intereses. La instalación de redes de altas prestaciones y salas con equipamiento audiovisual están permitiendo la configuración de lo que aquí hemos denominado campus virtuales, cuyo objetivo es eliminar la barrera de la distancia a la hora de asistir a cursos de interés. Las características de la UPV/EHU hacen que esta solución sea interesante para “consumo” propio, pero el horizonte es mucho más amplio si pensamos que estas tecnologías pueden permitir unas formas de colaboración entre universidades extraordinariamente valiosa, con un enorme ahorro económico tanto para las instituciones educativas como para los alumnos.

Para que este escenario se convierta en una realidad, aún queda bastante camino por andar. En concreto, se requiere que la red que une las universidades (es decir, RedIRIS o su sucesora) asuma el reto de la interconexión a alta velocidad: la red no debe limitarse ya a su utilización para transferencia de correo electrónico y navegación web. También es necesario que se avance en el campo de la estandarización: la solución de la UPV/EHU puede resultar muy satisfactoria para su uso interno, pero sin duda resultará incompatible con las que instalen otras instituciones.

Para que la oferta educativa sea más completa, la infraestructura para clases presenciales debería complementarse con otra para clases no presenciales, es decir, con materiales para autoformación, accesibles a través de una interfaz web (o similar). Estos repositorios de información podrían enriquecerse considerablemente con grabaciones (digitales) de clases previamente impartidas por videoconferencia. Del mismo modo, un profesor que imparta una clase presencial podría tener al alcance de sus manos un rico conjunto de materiales educativos (incluyendo vídeos previamente grabados) para mostrárselos a sus alumnos. Como podemos ver, los sistemas para formación a distancia presenciales y no presenciales no sólo no resultan excluyentes, sino que se complementan extraordinariamente bien.

Por último, hacer mención a un importante reto que va más allá de la instalación de infraestructuras: la utilización adecuada de las mismas. ¿Es aplicable el modelo de clase tradicional al nuevo entorno? Probablemente, no. Es necesario desarrollar nuevas metodologías para impartición de clases y colaboración entre grupos dispersos.

Referencias

ITU-T Recommendation H.310: Broadband audio-visual communications systems and terminal equipment (1996)

ITU-T Recommendation H.320: Narrow-Band Visual Telephone Systems And Terminal Equipment (1993/1996)

ITU-T Recommendation H.321: Adaptation of H.320 Visual Telephone Terminals to B-ISDN Environments (1996)

ITU-T Recommendation H.323: Visual Telephone Systems and Equipment for Local Area Networks which provide a Non-Guaranteed Quality of Service (1996)

Bellido, L.: A survey on standards for videoconferencing over broadband networks. 1st LEVERAGE conference on broadband communications in education and training, Education in the communication. University of Cambridge 7-8 January 1998. Disponible en http://greco.dit.upm.es/~leverage/conf1/bellido.htm.

Aramberri, J., Abal, F., Gamboa, M., Lasa, J., Miguel, J.: Un proyecto de teleformación: la Facultad Virtual. Novática n. 134 (1998) 54–58.

Fluckiger, F. Understanding networked multimedia: applications and technology. Prentice-Hall, 1995. ISBN 0-13-190992-4.

Team Solutions, Inc. Standards and Terminology. Documento electrónico en http://www.teamsolutions.co.uk/tsindex.html.

Fore Systems, Inc. StreamRunner™ AVA/ATV User's Manual, software version 5.0.x.

Apéndice 1: Lista de estándares y recomendaciones

Vídeo:

- H.261 – codec de vídeo para servicios audiovisuales a *p* x 64 Kb/s.
- H.263 – codec de vídeo para canales de comunicación de menos de 64 Kb/s.
- MPEG – codec para audio y vídeo con alto grado de compresión utilizado para grabación de "clips" (MPEG-1) y para emisión de TV digital (MPEG-2). La compresión tiene en cuenta redundancias intra-cuadro e inter-cuadros.
- M-JPEG – codec para vídeo consistente en la emisión de cuadros completos, cada un comprimido usando JPEG; no se eliminan redundancias inter-cuadros.

Audio:

- G.711 – Codificación PCM (Pulse Code Modulation) de sonido calidad teléfono (3 KHz), tal como se usa en un canal B (64 Kb/s) de la RDSI.
- G.722 – Codificación de sonido de 7 KHz de ancho de banda usando 64Kb/s.
- G.723.1 – Codificación para voz, usando 6'4 ó 5'3 Kb/s.
- G.728 – Codificación para sonido con predicción lineal y bajo retardo, usando 16 Kb/s.

Datos:

- T.120 – Protocolos de datos para conferencia multimedia.
- T.120 Annex C – Perfiles "ligeros" para la arquitectura T.120.
- T.121 – Plantilla genérica de aplicación.
- T.122 – Servicio de comunicación multipunto – definición del servicio
- T.123 – Pilas de protocolos específicas de diferentes redes, para su uso en conferencias multimedia.
- T.124 – Control genérico de la conferencia.
- T.125 – especificación del protocolo para el servicio de comunicación multipunto.
- T.126 – Protocolo para imágenes fijas y anotación (pizarra electrónica).
- T.127 – Protocolo para transferencia multipunto de ficheros binarios.
- T.128 – Compartición de aplicaciones en entornos multipunto.
- T.134 – Entidad de aplicación para conversación mediante texto (*chat*).
- T.135 – Transacciones entre el usuario y el sistema de reservas, en conferencias T.120.

Tamaños de ventanas de vídeo

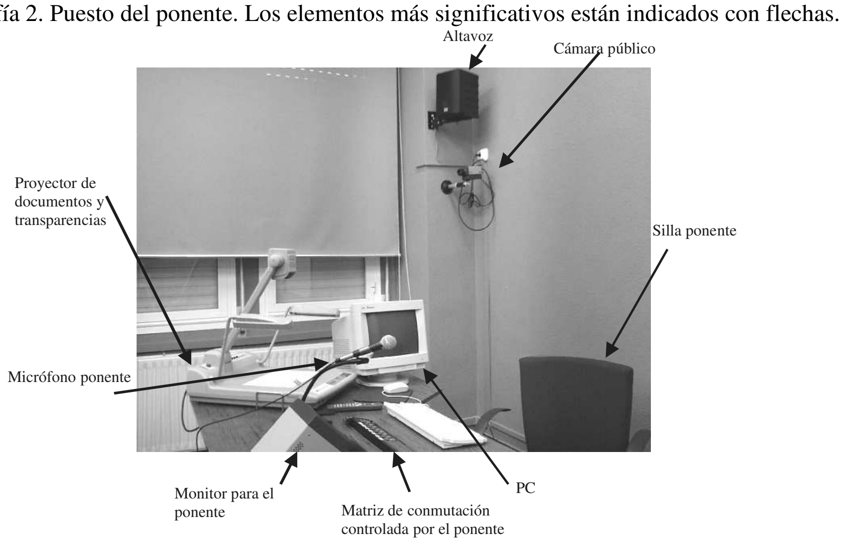
- NTSC – National Television Standards Committee, utilizado en los EE.UU, Canadá y Japón. Equivale a 640 x 480 píxeles.
- PAL – Phase Alternation by Line, utilizado en Europa (excepto Francia), África y oriente medio. Equivale 768 x 576 píxeles.
- SECAM – Sequentielle Couleur Avec Memoire, usado en Francia y Rusia. Equivale 768 x 576 píxeles.
- CIF – Common Interchange Format; de uso opcional en H.261. Especifica una ventana de 352 x 288 píxeles.
- QCIF – Quarter Common Interchange Format; estándar para H.261. Especifica una ventana de 176 x 144 píxeles.

Apéndice 2: Fotografías

Fotografías del aula de videoconferencia de la Facultad de Informática de la UPV/EHU.

Fotografía 1. Panorámica de la sala. Mesas de los asistentes y cámara (con control remoto) para captar al ponente.

Fotografía 2. Puesto del ponente. Los elementos más significativos están indicados con flechas.



Fotografía 3. Puesto de control.