



UNIVERSIDAD DE OVIEDO

Departamento de Informática

**CARACTERIZACIÓN, EVALUACIÓN Y OPTIMIZACIÓN DE
SISTEMAS MULTIMEDIA INTERACTIVOS EN ENTORNOS
DE E-LEARNING SÍNCRONO**

Juan Carlos Granda Candás

Gijón, julio de 2008



Universidad
de Oviedo

Reservados todos los derechos
© El autor

Edita: Universidad de Oviedo
Biblioteca Universitaria, 2009
Colección Tesis Doctoral-TDR nº 43

ISBN 978-84-691-7822-5
D.L.: AS.05361-2008





UNIVERSIDAD DE OVIEDO

Departamento de Informática

Tesis doctoral

**CARACTERIZACIÓN, EVALUACIÓN Y OPTIMIZACIÓN DE
SISTEMAS MULTIMEDIA INTERACTIVOS EN ENTORNOS
DE E-LEARNING SÍNCRONO**

Juan Carlos Granda Candás

Directores:

Francisco José Suárez Alonso
Daniel Fernando García Martínez

Gijón, julio de 2008

a Marta

Agradecimientos

Muchas son las personas que me han prestado su apoyo durante el desarrollo de esta tesis doctoral. Con estas breves líneas les quiero mostrar mi más sincero agradecimiento empezando por mis directores de tesis, Francisco José Suárez y Daniel García, ya que con su ayuda y orientación me han guiado por el camino correcto hasta la meta que esta tesis representa.

En segundo lugar a mi hasta hace poco compañero de laboratorio, Christian Uría, con quien llegué a entablar una sincera amistad fruto de horas y horas de discusiones banales y otras no tanto, y cuya ayuda me ha resultado inestimable.

A mi novia Marta que ha soportado con estoica paciencia este momento y ha sido capaz de compartirme con esta amante tan posesiva que ha resultado ser la tesis. Sus correcciones de los manuscritos y sus sugerencias visceralmente sinceras me sirvieron para no perder el Norte con el pasar de los capítulos.

También debo agradecer su apoyo a todos los compañeros del área, con quienes las conversaciones en los cafés adquirirían niveles místicos por momentos y que me servían para recordarme que hay vida más allá de la tesis.

Quiero reconocer también la comprensión de mis amigos durante estos años en que la tesis ha sido mi excusa más reiterada.

Por último, la familia, que siempre está ahí pase lo que pase y que con sus preguntas del tipo “¿Cuánto te falta?” me han empujado y animado hasta escribir la última palabra de esta memoria.

A todos ellos, muchas gracias,

Juan Carlos Granda Candás

Resumen

Los sistemas multimedia interactivos han ido ganando en popularidad a la par que crecía la capacidad de proceso de los computadores personales. Uno de los campos en los que se utilizan con mayor éxito es en el aprendizaje asistido por computador o e-learning, que, dependiendo de si impone restricciones temporales al proceso de enseñanza-aprendizaje, se denomina e-learning síncrono o asíncrono. El primero supone la utilización de tecnologías multimedia para posibilitar una interacción en tiempo real entre instructor y alumnos.

Tradicionalmente, para llevar a cabo actividades de e-learning síncrono es muy común la utilización de herramientas inicialmente no diseñadas para tal fin, como herramientas de videoconferencia o de trabajo colaborativo. En algunos casos, en lugar de utilizarse una única herramienta, estas actividades se desarrollan mediante un conjunto de herramientas, cada una de ellas ofreciendo una funcionalidad específica. En el mejor de los casos, cuando estas herramientas se diseñan con el e-learning síncrono como su principal objetivo, suelen constituir complejas plataformas que involucran redes de servidores dedicados. En este aspecto, la presente tesis viene a identificar cuáles son aquellas características que debe presentar una herramienta multimedia cuya principal misión sea dar soporte a acciones de e-learning síncrono. Para ello, se realiza un amplio estudio de las herramientas para dar soporte a eventos de e-learning síncrono que existen en el ámbito comercial y de investigación. Todas ellas presentan funcionalidades similares, por lo que es posible extraer un mínimo de funcionalidad común que cualquier herramienta de este tipo debería satisfacer.

Por otra parte, las grandes corporaciones empresariales han descubierto hace tiempo la importancia de la capacitación continua de sus recursos humanos. En este sentido, el e-learning síncrono aparece como una opción más que viene a completar el abanico de posibilidades para el desarrollo de acciones formativas. En la presente tesis se desarrolla un prototipo de plataforma de e-learning síncrono apropiado para el departamento de recursos humanos de una empresa multinacional exponiendo los posibles diseños que pueden plantearse, siguiendo los requisitos funcionales extraídos de la previa caracterización de las herramientas de e-learning síncrono.

Se hace especial hincapié en la definición del modelo de interacción entre el instructor y los alumnos que la plataforma posibilita, es decir, qué medios serán unidireccionales desde el instructor a los alumnos y cuáles serán multidireccionales para soportar el trabajo colaborativo. Esto tiene gran implicación en cómo los datos deben transportarse y, en definitiva, en la arquitectura global de la plataforma y las restricciones que impone sobre la red subyacente. Concretamente, la plataforma utilizará tráfico multicast siempre que sea posible, por ejemplo dentro de una red corporativa. Para el caso de redes que no permitan este tipo de transporte, se implementa un servidor que emula el comportamiento de una red multicast. Gracias a este servidor, es posible la implantación de arquitecturas multi-servidor altamente escalables. Sobre este tráfico multicast, se utilizan sesiones RTP independientes en el transporte de todos los medios, sean continuos o no. En este último caso, se enumeran las ventajas que supone la utilización de RTP para un medio de tipo no continuo y cómo la plataforma saca provecho de ello.

El prototipo final es el resultado de un proceso iterativo consistente en el diseño e implementación de sucesivos prototipos para evaluar el comportamiento en función de las decisiones de diseño tomadas. En base a estas evaluaciones, se proponen diferentes optimizaciones orientadas a mejorar el comportamiento del prototipo en la siguiente iteración. Dichas optimizaciones se podrán aplicar a cualquier herramienta multimedia de e-learning síncrono.

Asimismo, se presta gran atención a la simplicidad de la plataforma resultante. Por un lado,

por el perfil de alumnos a los que va dirigido: empleados con escasos conocimientos informáticos; y, por otro, para que el mantenimiento de toda la plataforma sea mínimo, no necesitando personal especializado para su gestión.

En ocasiones, es necesario asumir una reducción deliberada de funcionalidad para satisfacer requisitos como la operación con muy bajo ancho de banda. Por tanto, la plataforma desarrollada implementa el mínimo de funcionalidad necesaria para dar soporte a actividades de e-learning síncrono en el ámbito de la capacitación de los recursos humanos en grandes corporaciones utilizando la mínima cantidad de recursos, de tal forma que resulta aplicable a casi cualquier escenario posible de red.

Por último, en esta tesis también se analizan los resultados de la utilización de la plataforma en diferentes actividades de e-learning síncrono reales; resultados tales como la satisfacción de los usuarios que han utilizado la plataforma para asistir a clases virtual. Su grado de satisfacción ha sido medido a través de encuestas al final de cada una de las sesiones.

Abstract

Interactive multimedia systems have grown in popularity as throughput of personal computers improves. Computer-assisted learning or e-learning is a field in which these systems are used successfully. If these systems impose temporal restrictions on the learning process they are called synchronous e-learning systems, otherwise they are called asynchronous e-learning systems. The former implies the use of multimedia technologies to enable real-time interaction between instructor and students.

Traditionally, tools such as videoconference or computer-supported collaborative tools are used in order to carry out synchronous e-learning activities. However, these tools are not designed with the aim of e-learning. In some cases, rather than a single tool, multiple tools are used during these activities, each of them providing a specific functionality. Even though some tools are designed with the aim of providing support to e-learning activities, they usually constitute complex platforms with many dedicated servers. This PhD thesis identifies the characteristics to be accomplished by a tool in order to provide support to the development of synchronous e-learning activities. Thus, an extensive survey of the tools used to promote synchronous e-learning session is carried out, which covers market tools and others developed in the investigation environment. All of them offer similar functionalities, so this PhD thesis analyzes the set of common functionality that every tool must satisfy to be appropriate for synchronous e-learning.

Moreover, large business corporations have discovered the importance of synchronous e-learning applied to the continuous training of human resources. Synchronous e-learning is now another option to promote training. In this PhD thesis a synchronous e-learning prototype platform is developed to provide support to e-training activities in the scope of a large international company. All possible designs following the requirements extracted from the previous characterization of synchronous e-learning tools are discussed.

The final design is proposed with special emphasis on the definition of the interaction model allowed between instructor and student. It must resolve which multimedia data will flow in a one-way fashion from the instructor to the students, and which data will flow in a bidirectional way to enable collaborative work. This fact has enormous consequences in the way data is delivered and the restrictions that the global platform imposes on the underlying network. Specifically, the prototype platform will use multicast delivery whenever possible, for example in the company network. If multicast delivery is not available, a relay server developed within this PhD thesis could be used to emulate the behavior of a multicast network. With this server, highly-scalable multi-server architectures are possible. All multimedia data is organized in multiple RTP sessions according to its media type, and conveyed using multicast delivery whether or not they are continuous media (audio or video). The advantages of using RTP with non-continuous media are listed and how the prototype platform takes advantage of them.

The final prototype is the result of an iterative process of continuous design and implementation to estimate the behavior of the platform depending on the design decisions. Based on these evaluations, multiple optimizations are proposed, which aim to improve the global performance of the prototype. These optimizations can be applied to any synchronous e-learning tool.

Furthermore, the resulting prototype platform must be simple to use. This is a consequence of two factors. Firstly, the users of the prototype platform will be employees with little knowledge of computing. Secondly, the maintenance of the whole prototype platform should be minimum, so specialized staff is not needed to manage it.

Sometimes, it is necessary to assume a deliberate functionality reduction in order to satisfy

network bandwidth requirements, especially when little bandwidth is available. Therefore, the prototype platform offers the minimal functionality to promote synchronous e-learning activities in the scope of training in large enterprises, using as few resources as possible. Thus, the prototype could operate in any network scenario.

Finally, in this PhD thesis the results of the use of the prototype in real synchronous e-learning situations are documented and discussed. Questionnaires are used to estimate the satisfaction of participants attending e-learning sessions.

Índice general

Índice de figuras	XIII
Índice de tablas	XV
Acrónimos	XVII
1. Introducción	1
1.1. Los sistemas multimedia interactivos	1
1.1.1. Clasificación de los sistemas multimedia interactivos	2
1.1.2. Diseño de sistemas multimedia interactivos	4
1.1.3. Ámbitos de aplicación	6
1.2. El e-learning	9
1.2.1. Modalidades de enseñanza	12
1.2.2. Clasificación de los sistemas de e-learning	13
1.2.3. Beneficios del e-learning síncrono	18
1.3. Motivación	19
1.4. Objetivos	21
1.4.1. Plan de trabajo	21
1.5. Organización de la documentación	21
2. Herramientas de e-learning síncrono	23
2.1. Introducción	23
2.2. Herramientas de e-learning síncrono comerciales	24
2.2.1. WebEx Training Center	24
2.2.2. Adobe Acrobat Connect Professional	26
2.2.3. Microsoft Office Live Meeting	29
2.2.4. Elluminate Live!	31
2.2.5. Saba Centra Live	32
2.2.6. GoToMeeting	33
2.2.7. Otras herramientas de interés	33
2.2.8. Comparativa	36
2.3. Herramientas de investigación	36
2.3.1. Conclusiones	40
2.4. Caracterización de una herramienta de e-learning síncrono	40
2.4.1. Funcionalidad de una herramienta de e-learning síncrono	40
2.4.2. Requisitos mínimos de una herramienta de e-learning síncrono	46
2.4.3. Interfaz de usuario de una herramienta de e-learning síncrono	49
3. Análisis de tecnologías para el e-learning síncrono	57
3.1. Introducción	57
3.2. Transporte de la información multimedia	58
3.2.1. Comportamiento de las redes IP	58
3.2.2. IP Multicast	65

3.2.3.	Efectos del protocolo de transporte	67
3.2.4.	Real-time Transport Protocol	68
3.2.5.	Calidad de servicio	72
3.2.6.	Protocolos de señalización	74
3.3.	Procesamiento de la información multimedia	85
3.3.1.	Audio y vídeo	86
3.3.2.	Pizarra virtual	90
3.3.3.	Mensajería y control de presencia	94
3.4.	Control del uso de la palabra	94
3.4.1.	Protocolo BFCP	95
4.	Diseño de un prototipo de herramienta: evaluación y optimización	97
4.1.	Introducción	97
4.1.1.	Modelo de interacción	98
4.2.	Arquitectura de red	104
4.2.1.	Evaluación de alternativas y optimización	107
4.3.	Servidor relay RTP	116
4.3.1.	Rutas estáticas	117
4.3.2.	Rutas dinámicas	118
4.3.3.	Network Address Translation	118
4.4.	Diseño de alto nivel de la herramienta	120
4.5.	Procesamiento de medios	121
4.5.1.	Evaluación de alternativas y optimización	121
4.6.	Procesamiento de medios: audio y vídeo	122
4.6.1.	Audio	124
4.6.2.	Vídeo	124
4.6.3.	Evaluación de alternativas y optimización	125
4.7.	Procesamiento de medios: mensajería instantánea y control de presencia	126
4.7.1.	Evaluación de alternativas y optimización	127
4.8.	Procesamiento de medios: pizarra virtual	129
4.8.1.	Evaluación de alternativas y optimización	132
4.8.2.	Presentación de contenidos	134
4.8.3.	Anotaciones	140
4.8.4.	Puntero virtual	143
4.9.	Control de la sesión	144
4.9.1.	SIP	146
4.9.2.	H.323	147
4.9.3.	SAP	148
4.9.4.	SDP	148
4.9.5.	Evaluación de alternativas y optimización	149
4.10.	Interfaz de usuario	151
4.10.1.	Menú principal	151
4.10.2.	Barra de estado	154
4.10.3.	Ventanas auxiliares	155
4.10.4.	Área de presentación	156
5.	Valoración del funcionamiento	161
5.1.	Introducción	161
5.1.1.	Métricas cualitativas para la evaluación de medios continuos	162
5.2.	Pruebas de laboratorio	163
5.2.1.	Prueba internacional	164

5.2.2. Relay RTP	168
5.3. Docencia en tercer ciclo y acción tutorial	170
5.3.1. Viabilidad	171
5.3.2. Metodología	172
5.3.3. Planificación	173
5.3.4. Resultados esperados	174
5.4. Resultados obtenidos	175
6. Conclusiones y trabajo futuro	179
6.1. Objetivos alcanzados	179
6.2. Líneas de investigación futuras	180
A. Red corporativa de ArcelorMittal	183
A.1. Red corporativa	183
A.1.1. Red MPLS internacional	184
A.1.2. Redes MPLS nacionales	186
A.1.3. Backbone Avilés - Veriña	187
A.1.4. Calidad de Servicio (QoS)	188
B. Encuestas	189
B.1. Encuestas para los alumnos	189
Bibliografía	193

Índice de figuras

1.1. Sistema multimedia aislado	3
1.2. Sistema multimedia de comunicación entre iguales	3
1.3. Sistema multimedia cliente/servidor	3
1.4. Sistema multimedia de presentación	4
1.5. Sistema multimedia de conferencia	4
1.6. Modelo de ciclo de vida del software en cascada	5
1.7. Detalle de la página web de Nokia	8
1.8. Second Life aplicado a la enseñanza	8
1.9. Ventana principal de Google Earth	9
1.10. Modalidades de enseñanza	12
1.11. Aspecto de un curso desarrollado sobre Moodle	15
1.12. Aspecto de un curso desarrollado sobre WebCT	16
2.1. Cuotas de mercado de las herramientas comerciales	25
2.2. Pantalla principal de WebEx Training Center	27
2.3. Funcionalidad para compartir un aplicación en WebEx Training Center	28
2.4. Interfaz de usuario del instructor en WebEx Training Center	29
2.5. Pantalla principal de Acrobat Connect Professional	30
2.6. Pantalla principal de Live Meeting	31
2.7. Planteamiento de preguntas de los alumnos al instructor	48
2.8. Ventana principal de diferentes aplicaciones de e-learning síncrono	51
2.9. Menu principal y barra de herramientas de WebConference	52
2.10. Barra de estado de WebEx Training Center	52
2.11. Área de presentación de contenidos de WebEx Training Center	53
2.12. Ventanas auxiliares de varias aplicaciones de e-learning síncrono	54
2.13. Ventanas auxiliares en WebEx Training Center	55
3.1. Pilas de protocolos multimedia	58
3.2. Funciones de cada protocolo de la pila multimedia	59
3.3. Modelo del reloj de arena de una red IP	60
3.4. Ratio de pérdidas de paquetes en Internet	61
3.5. Distribución de frecuencia de las pérdidas de paquetes en ráfagas	61
3.6. Calidad percibida en función de la latencia de red en audioconferencias	63
3.7. RTT en Internet	63
3.8. Efectos del jitter de la red sobre las aplicaciones multimedia	64
3.9. Transmisión de datos en una red IP Multicast	65
3.10. Elementos en una red H.323	77
3.11. Elementos en una red SIP	81
3.12. Establecimiento de una sesión con SIP	82
3.13. Funcionamiento de un proxy SIP	83
3.14. Tablet PC	93
4.1. Traslación de las clases presenciales a un entorno en línea	99

4.2.	Modelo de interacción 1-N entre instructor y alumnos	100
4.3.	Modelo de interacción N-1 entre instructor y alumnos	100
4.4.	Modelo de interacción N-N entre instructor y alumnos	101
4.5.	Transporte de un medio mediante una o varias sesiones RTP	106
4.6.	Posibles alternativas de distribución de los datos	106
4.7.	Transporte distribuido utilizando técnicas multicast	108
4.8.	Transporte distribuido utilizando técnicas multi-unicast	110
4.9.	Transporte centralizado utilizando técnicas multi-unicast	111
4.10.	Transporte distribuido utilizando técnicas híbridas	112
4.11.	Transporte centralizado utilizando técnicas híbridas	114
4.12.	Arquitectura de la plataforma dentro de la Universidad de Oviedo	115
4.13.	Relay RTP para la traducción de flujos de un ámbito multicast a otro unicast	116
4.14.	Utilización de varios relay RTP para la traducción de flujos	117
4.15.	Comunicación entre dos equipos a través de un NAT	119
4.16.	Modelo arquitectónico de la herramienta prototipo	120
4.17.	Grafo DirectShow para la transmisión mediante RTP de audio y vídeo	124
4.18.	Grafo DirectShow para la reproducción de flujos RTP de audio y vídeo	124
4.19.	Características deseables de los codecs de audio y vídeo	125
4.20.	Mensajería instantánea y control de presencia con RTP	128
4.21.	Organización en capas de un área de trabajo	129
4.22.	Renderización de un área de trabajo de la pizarra virtual	131
4.23.	Latencia durante la unión a la clase de un nuevo participante	132
4.24.	Mapeo de las estructuras de un área de trabajo en la red	135
4.25.	Implementación de la presentación de contenidos con RTP	138
4.26.	Implementación de la capa de anotaciones con RTP	141
4.27.	Datos que se transportan del puntero virtual	143
4.28.	Conversión de unidades de espacio de trabajo a unidades de transporte	144
4.29.	Muestreo del puntero local y su representación en un equipo remoto	145
4.30.	Negociación para la unión a la clase virtual	145
4.31.	Unión a la clase virtual utilizando SIP	147
4.32.	Interfaz de usuario del prototipo	152
4.33.	Menú principal del prototipo	153
4.34.	Posibles disposiciones de la interfaz del prototipo	155
4.35.	Barra de herramientas principal	157
4.36.	Barra de anotaciones	158
5.1.	Recursos consumidos durante una clase virtual	165
5.2.	Consumo de ancho de banda del audio y el vídeo del instructor	166
5.3.	Consumo de ancho de banda de los mensajes de texto de un participante	166
5.4.	Consumo de ancho de banda de la pizarra virtual	167
5.5.	Consumo de ancho de banda de las anotaciones de dos usuarios	167
5.6.	Consumo de ancho de banda de los punteros virtuales de dos usuarios	168
5.7.	Recursos consumidos por el relay RTP durante una clase virtual	170
5.8.	Tareas del proyecto de innovación docente e-pSylon	175
A.1.	Red MPLS global de ArcelorMittal en Europa	184
A.2.	Red MPLS de ArcelorMittal en España	186
A.3.	Backbone entre las factorías de Avilés y Veriña	187
A.4.	Modelo simplificado del backbone entre las factorías de Avilés y Veriña	188

Índice de tablas

1.1. Comparativa entre e-learning asíncrono y síncrono	13
1.2. Ejemplos de sistemas de e-learning asíncronos y síncronos	14
1.3. Gastos que deben afrontarse en función de la modalidad de enseñanza	20
2.1. Comparativa de las plataformas de e-learning síncrono comerciales	37
2.2. Requisitos mínimos para una herramienta de e-training	48
3.1. Recomendaciones incluidas en la especificación H.323	78
3.2. Codecs disponibles para la audioconferencia	88
4.1. Utilización de los distintos medios como fuente de realimentación	102
5.1. Escala de valores de la métrica MOS	162
5.2. Características técnicas de los equipos de instructor y alumno	164
5.3. Características técnicas de relay RTP	169
A.1. Latencias de red medidas dentro de la red de ArcelorMittal	185

Acrónimos

- API** *Application Programming Interface*. 33, 35, 39, 86
- ASP** *Application Service Provider*. 25–27, 29, 31, 32, 34–36
- ATM** *Asynchronous Transfer Mode*. 76
- BBS** *Bulletin Board System*. 10
- BFCP** *Binary Floor Control Protocol*. 95, 96
- CAL** *Computer Aided Learning*. 6
- CELP** *Code Excited Linear Prediction*. 89
- CGI** *Common Gateway Interface*. 79
- CNAME** *Canonical NAME*. 70
- CPL** *Call Processing Language*. 83
- CSCL** *Computer-Supported Collaborative Learning*. 11
- DHCP** *Dynamic Host Configuration Protocol*. 82
- DNS** *Domain Name System*. 80–82
- EPSS** *Electronic Performance Support System*. 12, 13
- FEC** *Forward Error Correction*. 130, 131, 134, 139, 142, 163
- HCI** *Human Computer Interaction*. 5
- HTTP** *Hypertext Transfer Protocol*. 74, 75, 79, 81, 84, 85
- IANA** *Internet Assigned Numbers Authority*. 66, 75, 79
- ICANN** *Internet Corporation for Assigned Names and Numbers*. 66, 75
- IEEE** *Institute of Electrical and Electronics Engineers*. 82
- IETF** *Internet Engineering Task Force*. 57, 58, 68, 70, 71, 73, 74, 76, 78, 79, 84, 94, 95, 128, 143
- IGMP** *Internet Group Management Protocol*. 66
- IIS** *Internet Information Server*. 32
- ILS** *Internet Locator Server*. 34
- IM** *Instant Messaging*. 23

- IP** *Internet Protocol*. 3, 17, 57–60, 62, 65–68, 70, 72, 76, 79, 82, 84, 88–90, 92, 105, 118, 119, 139, 148
- IPTV** *Internet Protocol TV*. 7, 149, 180
- IRC** *Internet Relay Chat*. 94, 127
- ITU-T** Sector de Normalización de las Telecomunicaciones de la Unión Internacional de Telecomunicaciones. 57, 58, 70, 74–76, 84, 88–90, 94, 127, 129, 141, 162, 163
- LAN** *Local Area Network*. 66
- LCMS** *Learning Content Management System*. 15
- LDAP** *Lightweight Directory Access Protocol*. 79
- LMS** *Learning Management System*. 15
- MCU** *Multipoint Control Unit*. 38, 76, 77, 147
- MG** *Media Gateway*. 84
- MGC** *Media Gateway Controller*. 84
- MIME** *Multipurpose Internet Mail Extensions*. 79, 84, 85
- MOS** *Mean Opinion Score*. 162, 173, 176
- MPLS** *Multi-Protocol Label Switching*. 183–188
- MTU** *Maximum Transfer Unit*. 64
- NAT** *Network Address Translation*. 118, 119
- PCM** *Pulse Code Modulation*. 77, 87–89
- PDA** *Personal Digital Assistant*. 10, 93
- PDF** *Portable Document Format*. 91, 156, 172
- QoS** Calidad de Servicio. 73, 74, 76, 105
- RADIUS** *Remote Authentication Dial In User Service*. 79
- RAS** Registro, Admisión y Estado. 78
- RDSI** Red Digital de Servicios Integrados. 38, 58, 74, 76, 90, 95
- RFC** *Request For Comments*. 71
- RSVP** *Resource ReSerVation Protocol*. 73, 74
- RTCP** *Real-time Transport Control Protocol*. 69–71, 79, 105, 118, 119, 128, 131, 132, 138, 141–143
- RTP** *Real-time Transport Protocol*. 57, 58, 68–72, 74, 77–79, 85, 88, 94, 104–106, 116–119, 121–123, 125–128, 130–135, 139–147, 150, 164, 169, 172

- RTSP** *Real Time Streaming Protocol*. 84, 85
- RTT** *Round-Trip Time*. 62, 63
- SAP** *Session Announcement Protocol*. 78, 80, 84, 85, 148–150
- SCTP** *Stream Control Transmission Protocol*. 68, 76
- SDK** *Software Development Kit*. 28, 31, 34, 35, 86, 93, 141
- SDP** *Session Description Protocol*. 75, 83–85, 96, 121, 146–150
- SIMPLE** *Session Initiation Protocol for Instant Messaging and Presence Leveraging Extensions*. 94, 127
- SIP** *Session Initiation Protocol*. 34, 58, 68, 74–76, 78–85, 94–96, 105, 127, 128, 146–150
- SMI** *Sistema Multimedia Interactivo*. 1–9, 57, 90, 163
- SMTP** *Simple Mail Transfer Protocol*. 75, 81
- SPT** *Shortest Path Tree*. 66, 67
- SRTP** *Secure Real-time Transport Protocol*. 71
- SSRC** *Synchronization SouRCe*. 70–72, 105, 135, 142
- SVG** *Scalable Vector Graphics*. 91, 137, 139, 156
- TCP** *Transmission Control Protocol*. 58, 59, 62, 67, 68, 70, 75, 76, 78, 94, 95, 118, 180
- TDT** *Televisión Digital Terrestre*. 180
- TPV** *Terminal Punto de Venta*. 7
- TRIP** *Telephony Routing over IP*. 79
- UA** *User Agent*. 79
- UAC** *User Agent Client*. 79, 81
- UAS** *User Agent Server*. 79
- UDP** *User Datagram Protocol*. 58, 62, 67–70, 75, 76, 78, 88, 93, 105, 118, 119, 130, 138
- URI** *Uniform Resource Identifier*. 138, 140, 146
- USDLA** *United States Distance Learning Association*. 11
- VoIP** *Voice over IP*. 25, 27, 30, 31, 34, 35, 42, 62, 69, 70, 74, 82, 87–89, 149, 180, 188
- XML** *Extensible Markup Language*. 83, 91, 94, 127
- XMPP** *Extensible Messaging and Presence Protocol*. 34, 94, 127, 128

Capítulo 1

Introducción

En este capítulo se presentan las motivaciones que han desencadenado el desarrollo de la presente tesis, así como los objetivos planteados al comienzo de la misma. Además, se introduce el concepto de sistema multimedia interactivo y cómo una de sus principales aplicaciones es la enseñanza a distancia. Ahondando en esto, se realiza una pequeña introducción al e-learning y se detallan sus modalidades síncrona y asíncrona, exponiendo y contrastando las características de cada una. Finalmente, se describe cuál es la organización de la documentación en esta memoria.

1.1. Los sistemas multimedia interactivos

La idea de Sistema Multimedia Interactivo (SMI) aparece por primera vez en un artículo de Vannevar Bush en el año 1945 [Bush, 1945]. Bush, que por aquel entonces era director de la Oficina de Desarrollo e Investigación Científica y asesor del presidente de los Estados Unidos, plantea la necesidad de reorientar los esfuerzos científicos tras un periodo de guerra. Una de las líneas de investigación que sugiere implicaba el desarrollo de un dispositivo llamado *memex* para el almacenamiento y consulta de información de una forma muy flexible y eficaz. Mediante este dispositivo:

«[...] un individuo almacena sus libros, registros y comunicaciones y cualquier cosa mecanizable de tal forma que pueda ser consultada con extremada velocidad y flexibilidad. Es un amplio suplemento a su memoria».

De esta forma, Bush se anticipaba a lo que más tarde constituiría el SMI por excelencia, la *World Wide Web*.

Pero no es hasta la década de los ochenta y noventa cuando, con el aumento en el rendimiento de los computadores personales, aparecieron los primeros equipos sobre los que era posible implementar aplicaciones multimedia. Esto, junto con la invención de dispositivos como el disco compacto y el desarrollo de la tecnología de audio y vídeo, posibilitó la evolución del PC multimedia, ayudada por la maduración de las interfaces gráficas y los sistemas operativos.

En la actualidad, se pueden proponer varias definiciones de sistema multimedia dependiendo del autor consultado. En [Dastbaz, 2002] se denomina *multimedia* a:

«[...] un nuevo entorno integrado que es capaz de procesar información en una variedad de formatos como audio, vídeo, animaciones, gráficos y texto; no posibles anteriormente».

En esta definición se resaltan, por un lado la variedad de formatos de los que se hace uso en los sistemas multimedia, y por otro, el carácter novedoso por la imposibilidad tecnológica de integrar todos esos medios tiempo atrás.

Una definición alternativa es propuesta en [Elsom-Cook, 2001]:

«Multimedia es la combinación de varios canales de comunicación dentro de una experiencia coordinada para la que no existe un lenguaje natural que integre todos los canales».

Esta definición plantea dos aspectos: uno, la necesaria utilización de varios canales de comunicación, y el otro, cómo la conjunción de diferentes canales se percibe como un único medio. Este segundo aspecto nos conduce a una tercera definición enunciada en [Reddi *et al.*, 2003]:

«[...] la multimedia puede ser definida como la integración de varios medios (audio, vídeo, gráficos, texto animaciones, etc.) en un único, sinérgico y simbiótico todo que redunde en mayores beneficios para el usuario final que los que proporcionan cada uno de ellos individualmente».

Aquí, el autor resalta los beneficios que produce al usuario el uso combinado de varios medios respecto al que se obtendría con su utilización de forma individual. En este sentido, la definición de [Chen, 2004] viene a incidir en el beneficio que supone la combinación de medios:

«La verdadera multimedia es la combinación de diferentes elementos (medios, modalidades, tecnologías o aplicaciones) que proporcionan una más completa experiencia como efecto de esa combinación».

Ya por último, en la definición propuesta por [Phillips, 1997] se introduce el término *interactiva*, no presente en las anteriores, para caracterizar un SMI:

«El término “multimedia interactiva” es una expresión para describir la nueva ola de software que principalmente se ocupa del suministro de información. La componente “multimedia” está caracterizada por la presencia de texto, imágenes, sonido animaciones y vídeo; estando varios o todos organizados en algún programa coherente. La componente “interactiva” referencia al proceso de facultar al usuario con el control del entorno mediante un computador».

Otros términos que suelen utilizarse de forma incorrecta para referirse a las tecnologías multimedia son el hipertexto y la hipermedia. No conviene confundirlos pues se refieren a distintos conceptos. El término *hipertexto* fue acuñado por primera vez en [Nelson, 1965] como:

«[...] una combinación de texto en lenguaje natural con la capacidad de los computadores para saltar a otras partes del texto de forma interactiva, o mostrar dinámicamente [...] texto no lineal [...] que no puede ser impreso de forma adecuada en una página convencional».

Por contra, aunque también acuñado por el mismo autor en [Nelson, 1970], el término *hipermedia* se refiere a aquellos sistemas de hipertexto que son capaces de enlazar conjuntamente algo más que información textual, por ejemplo audio y vídeo.

1.1.1. Clasificación de los sistemas multimedia interactivos

Los SMI pueden clasificarse en función de multitud de criterios. Una clasificación inicial es posible en base a los componentes que los forman o la función que cumplen.

Si atendemos a los diferentes elementos que componen un sistema multimedia, éste puede clasificarse en tres grupos diferentes: sistemas multimedia aislados, de comunicación entre iguales o cliente/servidor.

Un *sistema multimedia aislado* es aquél que se desarrolla sobre un equipo autónomo sin comunicación de ningún tipo con otro sistema informático. Habitualmente, un único usuario puede

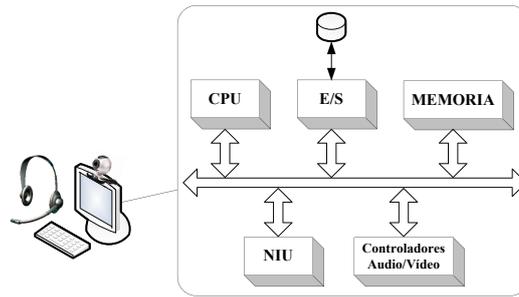


Figura 1.1: Sistema multimedia aislado

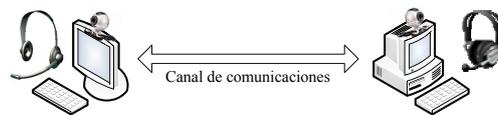


Figura 1.2: Sistema multimedia de comunicación entre iguales

interactuar en un determinado instante con un sistema multimedia de este tipo. Tal es el caso de los CD-ROMs y DVDs interactivos. En la figura 1.1 aparece una representación de un sistema multimedia aislado.

A diferencia del anterior, en un *sistema multimedia de comunicación entre iguales* pueden identificarse dos o más sistemas informáticos independientes. Desde el punto de vista del sistema multimedia todos los equipos son iguales, no existiendo ninguna diferencia funcional entre ellos. Como ejemplos paradigmáticos podemos mencionar los sistemas de videoconferencia o de voz sobre IP. Un esquema de esta clase de sistemas multimedia puede verse en la figura 1.2.

Por último, un *sistema multimedia cliente/servidor* es aquél en el que participan dos o más equipos informáticos, pero en el que uno se diferencia funcionalmente del resto. Sirva como ejemplo un servicio de vídeo bajo demanda utilizando técnicas de *streaming* para la distribución de los datos. En la figura 1.3 se esquematiza este tipo de sistemas.

Otra forma de catalogar los SMI es mediante el uso principal al que están destinados. De esta forma, tendremos sistemas de base de datos multimedia, sistemas multimedia de presentación y de conferencia.

Se denomina *sistema de base de datos multimedia* a un sistema multimedia cuya principal misión es la de almacenar información multimedia indexándola de alguna manera que permita consultarla en base a ciertos criterios. En este campo destacan los sistemas de gestión de bases de datos multimedia [Narasimhalu, 1996].

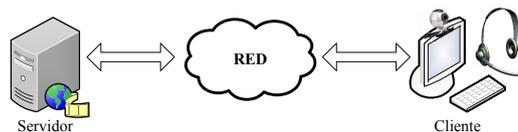


Figura 1.3: Sistema multimedia cliente/servidor

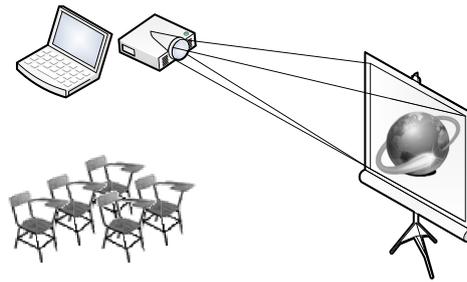


Figura 1.4: Sistema multimedia de presentación

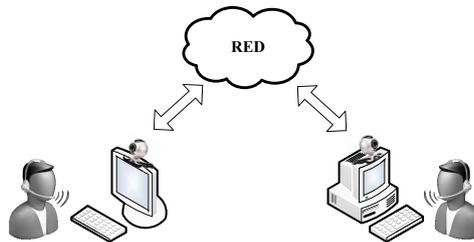


Figura 1.5: Sistema multimedia de conferencia

Un *sistema multimedia de presentación* es aquél destinado a la exposición y transmisión de información a través de diferentes medios, habitualmente, a varios usuarios. En la figura 1.4 aparece una representación de este tipo de sistemas.

Finalmente, se llama *sistema multimedia de conferencia* a un sistema multimedia que permite establecer una comunicación en tiempo real entre dos o más usuarios integrando diferentes tipos de información. La figura 1.5 muestra los elementos que intervienen en estos sistemas.

1.1.2. Diseño de sistemas multimedia interactivos

El diseño de un SMI no dista demasiado del diseño de cualquier aplicación convencional desde el punto de vista de la ingeniería del software. Al igual que con el resto de aplicaciones, es necesario seguir una serie de fases durante el *ciclo de vida* del software. Existen diferentes modelos con sus distintas fases para representar el ciclo de vida de un producto software, siendo el más clásico el modelo de cascada o *waterfall* [Royce, 1970], representado en la figura 1.6.

- La etapa de *Especificación de requisitos* involucra una comunicación entre el ingeniero del software y el cliente para acordar una descripción de la funcionalidad de la aplicación. Es, por tanto, el cliente quien proporciona unos requisitos que la aplicación final debe satisfacer.
- Durante la etapa de *Análisis* el ingeniero del software junto con el cliente debe delimitar, a través de esbozos del diseño final, entre otras cosas, cuál es el alcance de la aplicación así como la funcionalidad final que ésta debe proporcionar.
- En la fase de *Diseño* el ingeniero del software debe realizar una descripción más detallada de cómo la aplicación va a proporcionar la funcionalidad que se le presupone. Para ello, hará uso de diagramas y representaciones formales para especificar modelos de datos y arquitectónicos de la aplicación.

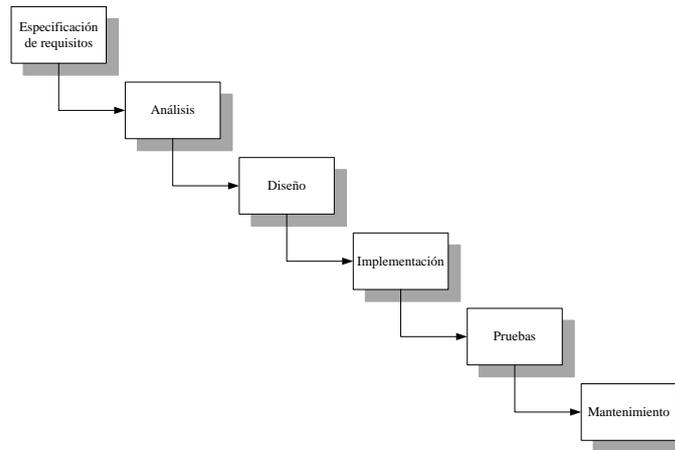


Figura 1.6: Modelo de ciclo de vida del software en cascada

- La etapa de *Implementación* implica la construcción de la aplicación utilizando un determinado lenguaje de programación.
- Durante la fase de *Pruebas* la aplicación construida es sometida a diferentes pruebas para verificar que ésta es correcta tanto funcional como operativamente.
- En la etapa de *Mantenimiento* la aplicación está en producción y es sometida regularmente a mejoras y cambios, tanto funcionales como correctivos y paliativos, con el fin de alargar su tiempo de vida.

Durante el desarrollo de cualquier producto software, especialmente en el caso de los SMI, es necesario prestar especial atención al diseño de la interfaz de usuario. La *Human Computer Interaction* (HCI) o interacción humano-computador estudia aquellas cuestiones relacionadas con las interfaces de usuario. Principalmente, se tienen en cuenta dos aspectos durante el diseño de las interfaces de usuario: la usabilidad y la accesibilidad.

La *usabilidad* es una medida de lo fácil y cómodo que es de usar una interfaz por un usuario. En [Shackel, 1990] se define en términos de efectividad, aprendibilidad, flexibilidad y actitud. Posteriormente, en [Nielsen, 1993] se amplió la definición y se enunciaron los siguientes atributos de la usabilidad:

- Tiempo de aprendizaje: es el tiempo que le lleva a un usuario aprender cómo realizar un conjunto de tareas.
- Eficiencia: el sistema debe ser productivo. Un usuario que haya aprendido cómo utilizar la interfaz debe alcanzar un cierto nivel de productividad.
- Errores: la interfaz debe favorecer que el usuario no cometa errores al realizar una tarea.
- Retención: los conocimientos de los usuarios acerca del manejo del sistema deben ser fácilmente memorizables, de modo que, tras un periodo de no uso, un usuario debe poder utilizarlo sin dificultad.
- Satisfacción: es la actitud subjetiva que muestran los usuarios hacia la interfaz. La interfaz debe resultar atractiva a los usuarios.

En el campo de la usabilidad se han realizado multitud de estudios con el fin de definir guías y pautas para el diseño de interfaces de usuario usables. Algunos ejemplos pueden encontrarse en [Galitz, 2007], [Lynch y Horton, 1999] y [Shneiderman y Plaisant, 2004].

El otro aspecto a cuidar durante el desarrollo de la interfaz de usuario es la accesibilidad. Se puede definir la *accesibilidad* como el grado en que la interfaz de usuario permite ser utilizada por cualquier usuario, independientemente de sus capacidades físicas o técnicas. Actualmente, existe un amplio consenso internacional en la necesidad de realizar interfaces accesibles. No en vano, por iniciativa del W3C se han definido una serie de pautas de accesibilidad en el diseño de interfaces de usuario [Chisholm *et al.*, 1999].

1.1.3. Ámbitos de aplicación

Muchos son los campos de aplicación de los SMI hoy en día. Por citar algunos ejemplos, los SMI se utilizan en la enseñanza, la difusión de noticias y publicidad, el comercio electrónico, el entretenimiento, etc. En los siguientes apartados se expondrán algunas de las posibles aplicaciones de los SMI, si bien no pretende proponer una lista exhaustiva, sino más bien un pequeño esbozo del abanico de posibilidades en que los SMI pueden llegar a utilizarse.

1.1.3.1. Enseñanza

Más allá de la enseñanza tradicional en la que el profesor impartía las clases de forma magistral y los alumnos eran receptores pasivos de conocimiento, en los últimos años se ha profundizado en el estudio de los beneficios que los SMI proporcionan al proceso de enseñanza-aprendizaje.

Los primeros productos para la enseñanza asistida por computador, *Computer Aided Learning* (CAL), venían lastrados por la posibilidad de utilizar únicamente texto para plasmar los contenidos didácticos, lo que hacía casi imposible representar ciertos tipos de información, de ahí que no tuvieran, inicialmente, el éxito que se les auguraba. Fue gracias a los avances técnicos cuando se pudieron incluir objetos multimedia dentro de estos paquetes CAL, favoreciendo el aprendizaje y los modos de interacción respecto a los paquetes tradicionales [Hutchings *et al.*, 1992].

En [Collins *et al.*, 1989] se expone que las habilidades de comprensión, composición, razonamiento y experimentación se adquieren no sólo por la mera transmisión de hechos, sino también por la interacción del aprendiz con los contenidos. Esto nos lleva a la definición de *aprendizaje interactivo* propuesta por [Taylor, 1980]:

«[...] el más valioso aspecto del computador en la educación [...], con los estudiantes actuando como verdaderos estudiantes más que como meros espectadores».

Diferentes autores han demostrado cómo los nuevos avances tecnológicos en el campo multimedia han venido posibilitando nuevas formas de comunicación, lo que redundará en una mayor riqueza de los materiales y, en última instancia, en una mejora de los mismos [Steinmetz y Nahrsstedt, 1995] y [Mishra y Sharma, 2005]. Por ello, la introducción de entornos de enseñanza virtuales multimedia en las aulas tiene un enorme potencial [Padmore *et al.*, 2006].

1.1.3.2. Entrenamiento

Cada vez les resulta más necesario a las empresas introducir innovaciones técnicas en sus procesos productivos para sobrevivir en un mercado competitivo. Como consecuencia de ello, el personal debe adaptarse a estos cambios de forma que pueda asumirlos y contribuya a la mejora de la productividad empresarial. En este sentido, se antoja imprescindible un programa de formación continua por parte de las empresas para mantener a sus empleados actualizados tecnológicamente.

Algunos estudios demuestran que el entrenamiento interactivo incrementa la retención de conocimientos al tiempo que acorta la curva de aprendizaje [Shulman, 1992], todo ello con una gran flexibilidad ya que es posible proporcionar entrenamiento *just-in-time* según las necesidades individuales de cada usuario [Israel, 1992].

En aquellas situaciones en las que durante el desarrollo de su labor un empleado deba manejar un material delicado o incluso peligroso para las personas son muy recomendables las técnicas de simulación. En [Mollet y Arnaldi, 2006] se desarrolla un trabajo para el entrenamiento de los empleados de una factoría de equipamiento militar utilizando técnicas de realidad virtual.

La dispersión geográfica de las factorías de una empresa es otro de los motivos que fundamentan el uso de los SMI para el entrenamiento de los empleados. Así, no es necesaria la existencia de un centro específico de formación, ahorrándose al mismo tiempo los costes que conllevarían los desplazamientos del personal. Además, habría que sumar el ahorro que supondría la utilización de aplicaciones multimedia para simular el manejo de maquinaria de alto coste, antes de la adquisición de la misma, para garantizar su correcto uso desde el principio.

En el capítulo 4 se expone un prototipo de plataforma de e-learning síncrono desarrollado en el ámbito de la presente tesis para la formación del personal de una multinacional en la que existe gran dispersión geográfica entre sus factorías.

1.1.3.3. Noticias y publicidad

Las agencias de publicidad consideran cada vez más necesario dotar de interactividad a los anuncios publicitarios para seducir a los clientes. Con la utilización de los SMI, es posible desarrollar campañas atractivas, bien sea mediante anuncios estáticos en páginas web, o *banners*, o mediante otros medios como el marketing directo a través del correo electrónico.

Los canales de noticias se ven también beneficiados por el uso de los SMI. En [Pañeda, 2004] se describe un servicio de noticias bajo demanda basado en audio y vídeo. Esto permite a los usuarios consultar las noticias de actualidad en cualquier momento y lugar, si bien existen estudios que aducen que la utilización de técnicas multimedia para la transmisión de noticias no es conveniente porque distrae al usuario de la verdadera historia de la noticia [Sundar, 2000].

Por otro lado, la difusión, previa suscripción, de señales de vídeo o televisión utilizando conexiones de banda ancha es un servicio que ofrecen cada vez más operadoras de acceso a Internet. Comúnmente recibe el nombre de IPTV y utiliza técnicas de *streaming* para hacer llegar los contenidos a los usuarios cuando éstos deseen desde servidores especializados.

1.1.3.4. Comercio electrónico

A pesar de los difíciles comienzos del comercio por Internet, o *e-commerce*, debido a la desconfianza que mostraban los usuarios [Hoffman *et al.*, 1999], el negocio que genera ha aumentado en los últimos años, y las previsiones indican que cada vez tendrá más peso en las cuentas de resultados de las empresas.

Los comerciantes están interesados en hacer atractivos al cliente los productos que ofertan, y es aquí donde entran en juego los SMI. La utilización de tecnologías multimedia permite captar la atención del cliente hacia los productos más fácilmente y posibilita asimismo una interactividad con los mismos. Por ejemplo, es posible visualizar un producto en tres dimensiones y girarlo para apreciar todas sus características como puede verse en la figura 1.7.

Otro campo de aplicación de los SMI es como soporte a los Terminales Punto de Venta (TPV), destinados a la gestión de negocios de venta al público. Algunos estudios detallan los beneficios derivados de la utilización de estos elementos en la venta al por menor [Weber y Kantamneni, 2002]. Sin embargo, no todos los TPV se utilizan para la gestión, otros pueden utilizarse, incluso, para realizar mercadotecnia visual [Schwartz, 2005]. Ejemplos de TPV pueden encontrarse en [Landa *et al.*, 1987] y [Yin *et al.*, 2005].



Figura 1.7: Detalle de la página web de Nokia



Figura 1.8: Second Life aplicado a la enseñanza

1.1.3.5. Entretenimiento

Uno de los primeros campos de aplicación de las tecnologías multimedia y, por extensión, de los SMI fue en el ámbito del entretenimiento. Muchos videojuegos se beneficiaron del desarrollo tecnológico posibilitando así la creación de mundos virtuales tridimensionales en los que el jugador podía sumergirse.

Un famoso mundo virtual inmersivo que ha venido ganando popularidad en los últimos años es el que representa *Second Life*. Dentro de él, un usuario puede ser representado por su propio *avatar* de forma que en dicho mundo se establecen relaciones sociales y comerciales entre los participantes. Muchos autores han puesto de manifiesto el potencial que este tipo de mundos virtuales ofrecen tanto en el ámbito educativo [Cheal, 2007], [Witmer *et al.*, 1996] y [Chenghung *et al.*, 2007]; como en el colaborativo [Snowdon y Munro, 2001]. La figura 1.8 representa una de las posibles aplicaciones de Second Life para la enseñanza, en donde cada alumno participaría en la clase virtual a través de su *avatar*, permitiendo una interacción cara a cara entre los participantes dentro de un mundo virtual 3D.

Otro fenómeno de ámbito mundial posibilitado por las tecnologías que ofrecen los SMI es *YouTube*. YouTube es un portal web que ofrece la posibilidad de compartir vídeos entre usuarios

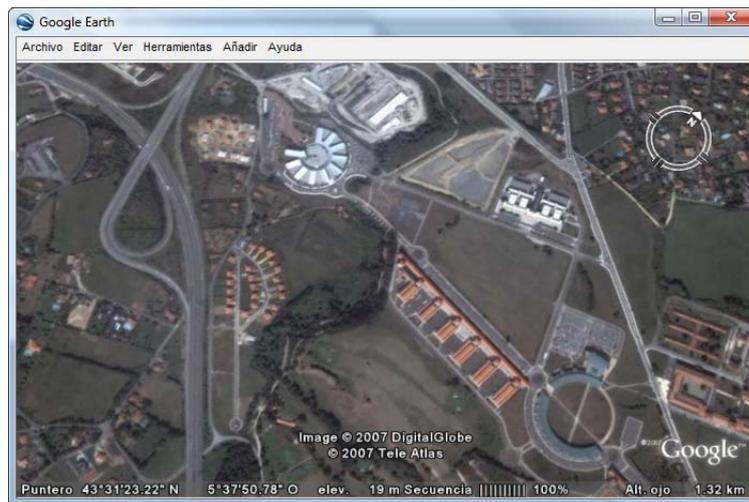


Figura 1.9: Ventana principal de Google Earth

utilizando técnicas de *pseudo-streaming*, que permiten a un usuario visualizar un vídeo mientras se descarga.

Por citar otro de los muchos ejemplos de aplicación de los SMI de forma exitosa en el campo del entretenimiento, podemos mencionar *Google Earth*. Se trata de una herramienta cartográfica que permite visualizar fotos por satélite de casi cualquier parte del globo terráqueo. En la figura 1.9 puede verse el aspecto que presenta la aplicación.

1.2. El e-learning

Con el auge mediático que ha tenido en los últimos años, existe un concepto extendido, más o menos meridiano, de lo que es el e-learning. Bajo este término confluyen dos aspectos que deben tenerse en cuenta. El primero es el aspecto pedagógico, ya que se trata de un proceso de enseñanza y aprendizaje al que le son aplicables las teorías y principios inherentes a cualquier proceso pedagógico clásico. La segunda cuestión está relacionada con la tecnología, pues es necesaria una base tecnológica para que el proceso pueda ser llevado a cabo.

De esta forma, dependiendo del punto de vista del autor, existen infinidad de definiciones del término, destacando el aspecto pedagógico o el tecnológico. Todas ellas tienen en común la identificación de un proceso de aprendizaje a distancia en el cual es necesaria la utilización de tecnologías multimedia para la transmisión de los contenidos educativos. Por tanto, se podría definir el *e-learning* como aquel proceso de enseñanza-aprendizaje a distancia que utiliza como soporte tecnologías multimedia. En este sentido, en [Sherry, 1996] se define el aprendizaje a distancia como:

«[...] aquel proceso de enseñanza y aprendizaje en el que existe una separación temporal y/o espacial entre profesor y alumno, dirigido por este último en lugar de por el profesor, y con una comunicación no continua entre ambas partes mediante material impreso u otras tecnologías».

Es común identificar diferentes generaciones a lo largo de la historia de la enseñanza a distancia atendiendo a las tecnologías utilizadas en cada una de ellas. Así, las primeras formas de enseñanza a distancia tuvieron lugar en Europa e implicaban la utilización del correo postal para el

intercambio de material formativo entre el profesor y el alumno, y la utilización del teléfono o entrevistas presenciales para tutorías puntuales. Esto suponía un lastre en el proceso de aprendizaje por la lentitud del canal de comunicación entre ambas partes.

A partir de mediados del siglo pasado, se comenzaron a utilizar tecnologías de difusión como la televisión o la radio para transmitir el conocimiento a través de emisiones en directo de clases magistrales [Cambre, 1991]. Todo ello sin realizar apenas cambios en el modelo instructivo tradicional. A esta modalidad de enseñanza en la que se utiliza la televisión como medio para hacer llegar los contenidos didácticos se le conoce por *t-learning* (*television learning*) o televisión educativa.

La gran problemática que este tipo de medios presentaban era la falta de interactividad entre profesor y alumno. Los alumnos no podían enviar sus comentarios al profesor por la naturaleza unidireccional del medio que se utilizaba. Además, no siempre el profesor experto en una materia determinada conseguía atraer el interés del televidente, y se prefería a un buen comunicador antes que a un buen formador. De todas formas, gracias a la televisión digital interactiva, existen todavía esfuerzos por utilizar este medio como difusor de contenido pedagógico [López Nores *et al.*, 2005] y [Jokipelto, 2005].

En una segunda generación, se empezó a utilizar Internet como medio para la transmisión de los contenidos didácticos. Así, era común el uso del correo electrónico y los tablones de anuncios (BBS), al igual que la audioconferencia y la videoconferencia para transmitir información educativa sonora y visual. En muchas ocasiones, la migración del material formativo para su uso a distancia se limitaba a su digitalización para ponerlo a disposición de los alumnos a través de un portal web. En esta etapa aún no se había producido ningún cambio en el aspecto instructivo del proceso de enseñanza-aprendizaje.

Ya por aquel entonces, algunos autores [Webster y Hackley, 1997] y [Zhang, 1998] destacaban los beneficios que suponían la gran riqueza de medios, gracias a la utilización de nuevas tecnologías, para la efectividad del proceso de enseñanza, siempre que se adoptase un estilo interactivo, al tiempo que ponían de manifiesto el esfuerzo extra necesario por parte del profesor para implicar a los alumnos.

En la actual generación, la tercera, se utilizan de forma intensiva las características de Internet para construir una comunidad virtual a través de la cual los alumnos pueden aprender de forma colaborativa [Weller, 2007]. Se busca en todo momento que los alumnos tengan la percepción de que estas comunidades virtuales son una extensión natural de los campus tradicionales y evitar que les embargue una sensación de aislamiento. De esta forma, entra en juego el concepto de *aprendizaje social*, que define el proceso de aprendizaje como una interacción entre las competencias que ha adquirido históricamente una comunidad social con la experiencia personal de cada uno de sus miembros [Wenger, 2000]. Específicamente, el concepto de *comunidad de práctica* define a un conjunto entrelazado de individuos que comparten un dominio de interés común y acerca del cual se comunican [Gannon-Leary y Fontainha, 2007], a diferencia de en una *comunidad de aprendizaje*, donde el objetivo que se persigue es el incremento del conocimiento de sus participantes a través de métodos educativos formales.

En una futura generación, la cuarta, se utilizarán dispositivos móviles tales como teléfonos móviles, ordenadores portátiles, PDAs, iPods, etc.; lo que supondrá un cambio hacia una enseñanza móvil y una revolución en el modelo instruccional con respecto al modelo tradicional de la enseñanza presencial. A colación de esto, se acuñó el término *m-learning* (*mobile learning*) para referirse a los eventos de e-learning que se desarrollan sobre este tipo de dispositivos [Motiwalla, 2007]. Debido a su éxito y su gran difusión, la tendencia actual es portar los sistemas de e-learning a una versión m-learning [Georgiev *et al.*, 2006], [Trifonova *et al.*, 2005]. Algunos autores van un paso más allá y utilizan el término *u-learning* (*ubiquitous learning*) para referirse al e-learning aplicado a cualquier dispositivo inalámbrico y destacar así su ubicuidad [Hwang, 2006].

A raíz de la popularidad del e-learning, se han venido acuñando en los últimos años diferentes términos para referenciar distintas modalidades de éste. Alguno de estos términos son los

siguientes:

- *b-learning* (*Blended learning*). Se utiliza para referirse a la combinación de enseñanza presencial con enseñanza a distancia, esto es, una enseñanza semipresencial. En algunos casos, puede encontrarse en la literatura un uso del b-learning para indicar combinaciones de distintas modalidades de enseñanza a distancia. En cualquiera de sus usos, el b-learning siempre identifica una combinación de diferentes modalidades de enseñanza.
- *c-learning* o *Computer-Supported Collaborative Learning* (CSCL). Define un aprendizaje colaborativo entre varios usuarios mediante la utilización de computadores. En este caso la figura de instructor es prescindible, puesto que los usuarios aprenden gracias a las interacciones que se producen entre ellos.
- *e-training*. Es el entrenamiento o capacitación en el ámbito corporativo, esto es, dirigido al personal de una empresa y conducido mediante e-learning.
- *d-learning* (*Distance Learning*). Se utiliza para referenciar la enseñanza a distancia en general, no necesariamente a la que utiliza medios electrónicos.

Aunque, a priori, el aprendizaje a distancia puede ser un instrumento válido en cualquier nivel de enseñanza, hay autores [Chassie, 2002] que ciñen su ámbito de aplicación exclusivamente a la educación superior. Esto se debe a que la motivación y la capacidad de organización necesarias para atender a un curso a distancia es mayor que en modalidades presenciales, características más habituales en personas que cursan educación superior respecto a personas más jóvenes aún inmaduras.

A pesar de este inconveniente, recientes estudios han puesto de manifiesto el tremendo potencial de crecimiento del e-learning a nivel global. Por ejemplo, un estudio de la *United States Distance Learning Association* (USDLA) [USDLA, 2008] de agosto de 2007 estima que más del 96 % de universidades de Estados Unidos ofrecen alternativas de educación a distancia, y que aproximadamente 4 millones de alumnos estaban recibiendo educación a distancia a finales de 2006 [Ebersole, 2007]. Otro informe por parte de *Learning Light* [LearningLight, 2008] de enero de 2007 sobre el estado del e-learning en el Reino Unido cifra en más de 240 millones de euros el mercado del e-learning con un incremento de hasta el 25 % al año [Helmer, 2007].

La enseñanza a distancia utilizando técnicas de e-learning presenta una serie de ventajas respecto a la enseñanza presencial tradicional:

- Mayor productividad: ya que el uso de nuevas tecnologías facilita la retención de los conocimientos por parte del alumno [Shulman, 1992].
- *Just-in-time*: los contenidos formativos pueden ser accedidos en cualquier momento y desde cualquier lugar. De esta forma se rompen las limitaciones geográficas y temporales que presenta la enseñanza tradicional.
- Flexibilidad: gracias a la utilización de sistemas multimedia interactivos el alumno puede diseñar su propio itinerario de aprendizaje a través de los contenidos formativos, según sus necesidades. Así, es posible adaptar la formación a cada alumno de forma específica.
- Coste: aunque inicialmente alto por el esfuerzo que supone la creación del material formativo y adaptarlo a las nuevas tecnologías, el coste de este tipo de soluciones es menor puesto que desaparecen los relacionados con desplazamientos.

No todo son ventajas al utilizar soluciones de e-learning para llevar a cabo acciones formativas. Está claro que la enseñanza a distancia en cualquiera de sus modalidades difícilmente podrá llegar a sustituir a la enseñanza presencial tradicional, principalmente por la ausencia de comunicación no verbal, inherente a la interacción cara a cara. Pero puede ser un complemento ideal a la enseñanza presencial que debe ser tenido en cuenta en el diseño de planes formativos.

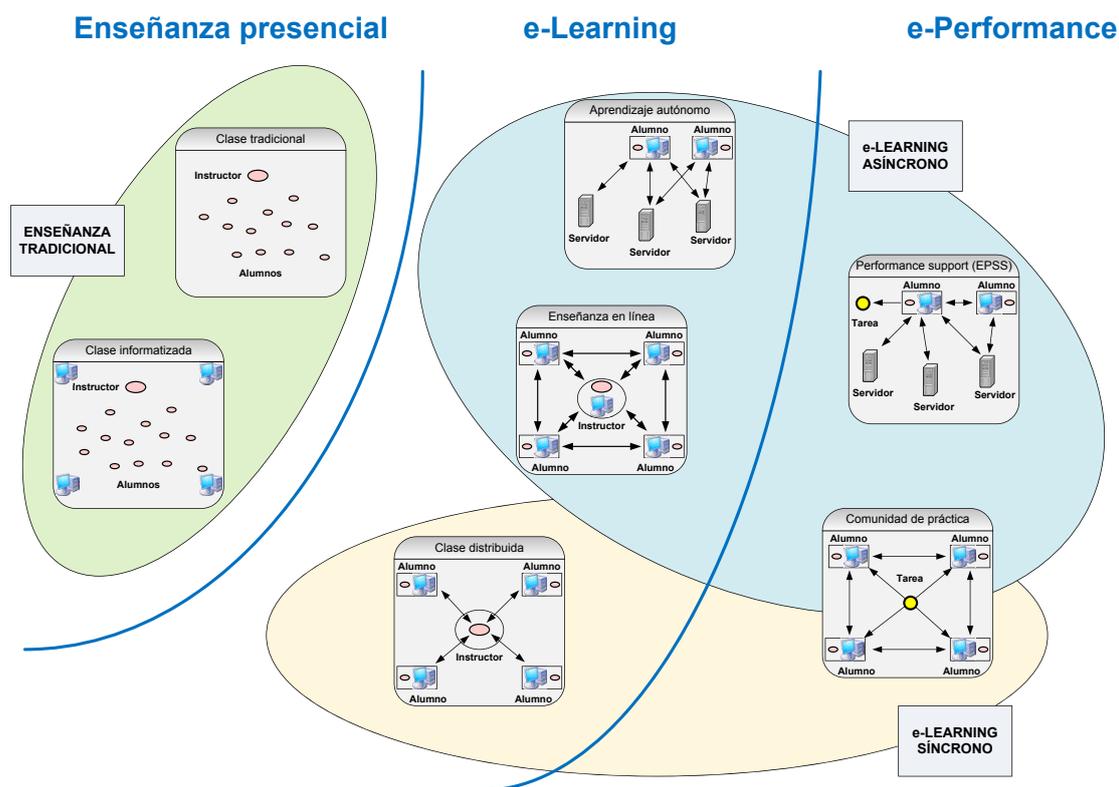


Figura 1.10: Modalidades de enseñanza

1.2.1. Modalidades de enseñanza

En la figura 1.10 aparecen representadas las diferentes modalidades de enseñanza [Paquette, 2001]. En la parte izquierda, se ubican las modalidades presenciales, donde instructor y alumnos deben coincidir en el espacio y en el tiempo para que el proceso de enseñanza y aprendizaje pueda llevarse a cabo. Habitualmente, este sitio de encuentro será un aula que podrá estar equipada con equipos informáticos.

En la parte central de la figura 1.10, se representan aquellas modalidades que hacen uso de las nuevas tecnologías pero apenas cambian los modelos de instrucción tradicionales. El *aprendizaje autónomo* consiste en utilizar material formativo a través de un portal web o CD-ROM de forma individual sin colaboración ni soporte por parte de otros alumnos o el instructor. Por contra, en la *enseñanza en línea* el aprendizaje del alumno está guiado por un instructor y se basa en interacciones asíncronas, principalmente mediante foros de discusión, entre instructor y alumnos o entre alumnos. Por último, las *clases distribuidas* utilizan técnicas como la videoconferencia para poner en contacto a instructor y alumnos, sin apenas cambios en el modelo de instruccional.

Finalmente, en la parte derecha de la figura 1.10, aparecen aquellas modalidades relacionadas con el concepto de e-performance. El término *e-performance*, habitualmente utilizado en el ámbito de la formación de recursos humanos, consiste en la mejora del desempeño individual utilizando las nuevas tecnologías como soporte. En este campo destacan los *Electronic Performance Support System (EPSS)*, que, según [Gery, 1991], consisten en:

«[...] un entorno electrónico integrado disponible y de fácil acceso para cada empleado y que está estructurado para proporcionar acceso en línea inmediato e individualizado

Tabla 1.1: Comparativa entre e-learning asíncrono y síncrono

e-learning asíncrono	e-learning síncrono
Acceso intermitente bajo demanda	Tiempo real
Grabado / producido con anterioridad	En directo
Disponible en cualquier momento (<i>just in time</i>)	A veces planificados y de duración concreta
Individual o colaborativo de forma intermitente	Colectivo y, a veces, colaborativo
Aprendizaje independiente	Presencia virtual simultánea entre aprendices e instructores
Autodidacta (<i>self-paced</i>)	Aprendizaje concurrente con otros

a todo el conjunto de información, software, orientación, asistencia, datos, imágenes, herramientas y sistemas de monitorización que le permiten desempeñar su trabajo con el mínimo soporte e intervención por parte de otras personas».

De esta forma, mediante los EPSS, los empleados aprenden según van realizando las tareas gracias a la ayuda *just-in-time* que le proporcionan estos sistemas. Por otro lado, las comunidades de práctica, sirven para que los expertos difundan los conocimientos entre todos los participantes de la comunidad y así, ayudarlos en el desempeño de las tarea de cada uno de ellos.

1.2.2. Clasificación de los sistemas de e-learning

Es habitual clasificar los sistemas de e-learning teniendo en cuenta las limitaciones temporales que imponen a los participantes del proceso de aprendizaje. Así, tendremos dos tipos de sistemas:

- Asíncronos: entre el profesor y los alumnos puede existir una separación tanto espacial como temporal.
- Síncronos: entre el profesor y los alumnos únicamente existirá una separación espacial.

Los sistemas de e-learning asíncronos no imponen ningún tipo de restricción temporal al desarrollo del proceso de aprendizaje. De esta manera, no es necesario que el profesor y los alumnos estén sincronizados temporalmente, estando a disposición de estos últimos todo el material formativo en cualquier momento.

En cuanto a los sistemas síncronos, éstos imponen una sincronización entre profesor y alumnos, de tal forma que, si bien geográficamente pueden estar muy distantes, deben coincidir en el tiempo para que el proceso de aprendizaje pueda llevarse a cabo.

En la tabla 1.1 aparece una comparativa de las características del e-learning asíncrono frente al síncrono. Por otro lado, en la tabla 1.2 se muestran diferentes ejemplos de sistemas que dan soporte al desarrollo de actividades de e-learning asíncrono y síncrono.

Además de para referirse a la combinación de enseñanza a distancia y presencial, en muchas ocasiones suele utilizarse el término b-learning para denominar a la utilización simultánea de las modalidades síncrona y asíncrona del e-learning en una misma acción formativa.

Tabla 1.2: Ejemplos de sistemas de e-learning asíncronos y síncronos

e-learning asíncrono	e-learning síncrono
E-mail	Mensajería instantánea
Foros de discusión	Chat en línea
Entrenamiento basado en web	<i>Live Webcasting</i>
<i>Podcasting</i>	Audioconferencia
DVD	Videoconferencia
Entrenamiento por computador	<i>Web conferencing</i>

1.2.2.1. Sistemas de e-learning asíncrono

Los sistemas asíncronos, a diferencia de los síncronos, ofrecen como ventaja que las discusiones y aportaciones de los participantes quedan registradas y el usuario puede estudiarlas con detenimiento antes de ofrecer su respuesta.

Algunas de las variantes de sistemas de e-learning asíncrono son:

- Cursos autodidactas (del inglés *self-paced courses*): son unidades didácticas en las que el ritmo de avance viene impuesto por el alumno. Su ventaja más obvia es la conveniencia, esto es, un alumno puede realizar el entrenamiento que desee cuando lo estime oportuno. Esto incluye el entrenamiento *just-in-time*, donde una persona realiza el entrenamiento que necesite para realizar una tarea concreta.

Para desarrollar estas unidades didácticas se suelen utilizar herramientas de autoría destinadas al e-learning, que siguen diferentes estándares y que pueden materializarse en varios entornos:

- Internet o una red local.
- CD-ROM o DVD.

Por otro lado, es habitual que estas unidades presenten una serie de características comunes:

- Multimedia: combinan texto, gráficos, animaciones y audio y vídeo para mejorar el proceso de aprendizaje.
 - Interactividad: siguen una estrategia instruccional que ayuda a los alumnos a practicar lo aprendido.
 - *Bookmarking*: permiten al alumno parar la unidad en cualquier momento para reanudarla en el mismo punto en otro momento.
 - Seguimiento: notifican el rendimiento del alumno y cuál es el avance en la unidad didáctica.
- Grupos de discusión: se trata de colecciones de conversaciones que ocurren a lo largo del tiempo, no siendo necesaria una continuidad temporal entre las mismas. Un alumno podría publicar un problema en un grupo de discusión para, algún tiempo después, ser respondido por otro alumno. Otros podrían unirse a la misma conversación proponiendo nuevas respuestas o problemas. Los grupos de discusión suelen utilizarse para dar soporte a un conjunto de personas que participan en el mismo proceso de aprendizaje o para comunicar personas realizando tareas similares. Resultan especialmente útiles para proporcionar respuestas de expertos a un grupo de personas.

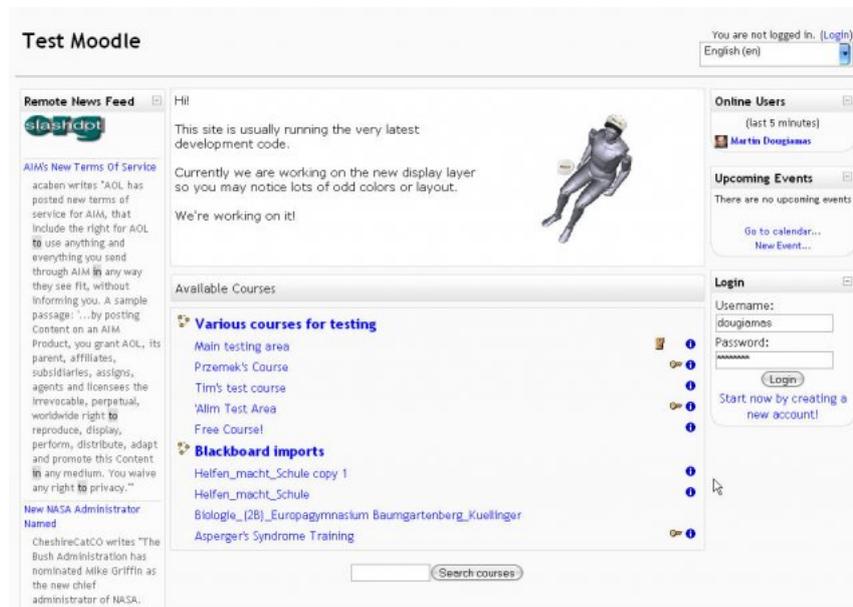


Figura 1.11: Aspecto de un curso desarrollado sobre Moodle

- Encuestas: pueden ser utilizadas para verificar la consolidación de conocimientos por parte de los alumnos.
- Wikis: consisten en páginas web que pueden ser editadas por varios usuarios de forma que es posible la construcción de documentos de forma colaborativa.
- Blogs: recopilan de forma cronológica textos o artículos de varios autores. Al igual que las wikis, permiten el desarrollo de una idea a través de la *escritura colaborativa*.
- Entornos de simulación virtuales: permiten sumergir al alumno en un entorno en el que puede practicar las tareas que desarrollaría en su puesto de trabajo habitualmente. Esto es muy útil para el desempeño de labores peligrosas o que conlleven la manipulación de material especialmente sensible. O incluso para representar escenarios realistas en los que se enfrente al alumno a situaciones problemáticas con las que pudieran encontrarse en su vida profesional.

Para dar soporte a este tipo de sistemas, al tiempo que administrar y proporcionar un seguimiento de la oferta formativa que se ofrece a los usuarios, es común utilizar un tipo concreto de herramientas informáticas. Se trata de los *Learning Management Systems* (LMS). Algunos ejemplos de este tipo de herramientas son Moodle, WebCT o .LRN. En las figuras 1.11 y 1.12 puede verse el aspecto de un sitio web basado en Moodle y WebCT respectivamente. Cuando, además, ofrecen servicios de gestión de contenidos se les conoce por *Learning Content Management Systems* (LCMS).

Otro tipo de herramientas son las destinadas a la creación de contenidos. Además, del tradicional PowerPoint, existen infinidad de herramientas específicas. Por poner un par de ejemplos representativos, podemos citar el propio .LRN o Hot Potatoes, este último destinado a la creación de preguntas de test, de respuesta corta y crucigramas.

El esfuerzo en los últimos años en el campo del e-learning asíncrono ha estado dirigido a desarrollar estándares y especificaciones que permitiesen la reutilización del contenido formativo

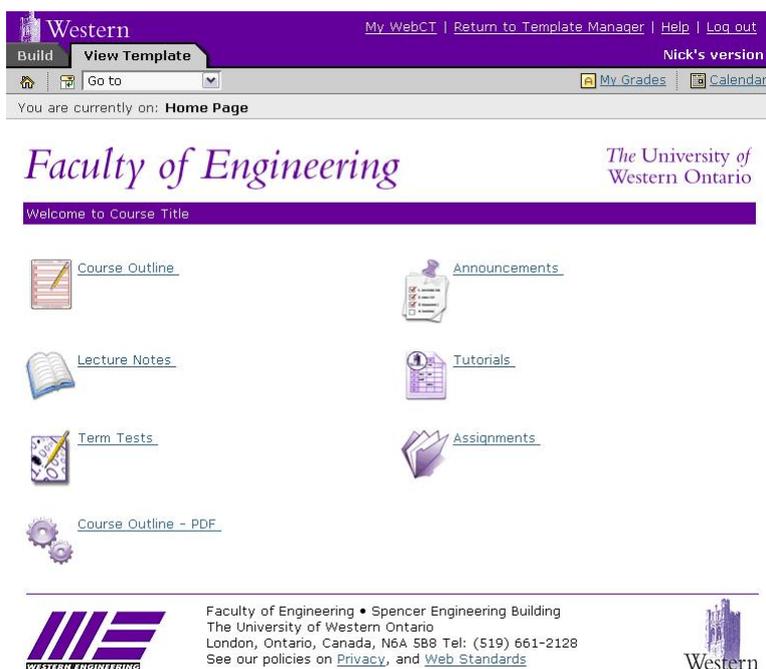


Figura 1.12: Aspecto de un curso desarrollado sobre WebCT

desarrollado. Estos contenidos suelen agruparse en pequeñas cápsulas de conocimientos conocidas como *objetos de aprendizaje reusables*.

1.2.2.2. Sistemas de e-learning síncrono

Aquel proceso de enseñanza-aprendizaje síncrono para cuyo desarrollo se utilizan medios electrónicos recibe el nombre de *e-learning síncrono*. Más concretamente, según [Hyder *et al.*, 2007] se define el e-learning síncrono como:

«[...] aprendizaje en directo, en tiempo real, interactivo y habilitado por la utilización de medios electrónicos».

Alguna de las características que presentan los sistemas de e-learning síncrono, y que los distinguen de las modalidades asíncronas del e-learning son:

- En directo. Esto significa que las actividades de e-learning síncrono se realizan en vivo, es decir, no son grabadas con anterioridad, aunque sí pueden hacer uso de material producido antes de la sesión.
- En tiempo real. Aunque las actividades de e-learning síncrono pueden llegar a grabarse para una posterior reproducción, éstas son de carácter eminentemente de tiempo real, sin que puedan pausarse ni retroceder.
- Facilitada. Por norma general, existirá la figura de facilitador cuya misión será la de velar por que las interacciones que se produzcan entre los participantes de la sesión de e-learning estén conducidas a la adquisición de conocimientos. Facilitará la consolidación de los mismos.

- Interacción orientada al aprendizaje. Esto distingue al e-learning síncrono de otras actividades desarrolladas en tiempo real como pueden ser las videoconferencias o la demostración de productos en línea.

Existen una serie de tecnologías que dan soporte al e-learning síncrono. A continuación se enumeran algunas de ellas:

- Teleconferencia. Dentro de esta categoría se encuadran los sistemas de audio y videoconferencia:

- Audioconferencia. Consiste en la comunicación a través de señales de audio utilizando, principalmente, el teléfono. Debido a su bajo coste y a la facilidad de implantación, muchas empresas optan por llevar a cabo planes de entrenamiento síncrono de esta forma. Además, aún a pesar de que algunos la consideran una tecnología anticuada, obtiene unos resultados aceptables y facilita la movilidad de los empleados con la utilización de teléfonos móviles. Una ventaja que ofrece es la posibilidad de grabar las conferencias para después ofrecerlas para su descarga y reproducción en dispositivos tipo iPod. Estos clips de audio reciben el nombre de *podcasts*.

Las audioconferencias rara vez se utilizan de forma aislada y sí en conjunción con otro tipo de herramientas como pueden ser la presentación de diapositivas.

- Videoconferencia. Implica la comunicación mediante señales de audio y vídeo. Aunque inicialmente fue utilizada como herramienta de encuentro empresarial, desde hace unos años se ha comprobado el gran potencial que ofrece el vídeo a pantalla completa en el campo del e-learning para emular en cierta forma la interacción cara a cara (*face-to-face*). Además, con la migración a soluciones IP en lugar de utilizar redes dedicadas, esta herramienta se ha convertido en una solución más que factible en cuanto a su coste.

La videoconferencia es adecuada para grupos dispersos en varias localizaciones geográficas. En este sentido, dos son los campos en que suelen utilizarse. La *telepresencia* se caracteriza por el uso de imágenes de alta definición y está destinada a entornos en los que es imprescindible una gran precisión visual, como la medicina a distancia. Por otro lado, están los sistemas de videoconferencia basados en webcam, o de escritorio. Sin embargo, la distinción entre uno y otro campo es cada vez más difícil por la convergencia que tiende a producirse entre ambos.

- *Webcasting*. Es un término que podría traducirse por difusión web. Consiste en la difusión de audio y vídeo utilizando técnicas de *streaming* a través de la red, ya sea en vivo o bajo demanda. Habitualmente existirá un único emisor y múltiples receptores, por lo que se podrán utilizar técnicas de difusión multicast para hacer fluir los datos en un solo sentido.

Su principal uso es la traslación de las clases magistrales desde un ámbito presencial a otro remoto. Es especialmente indicado para las situaciones en que existe un gran número de participantes, por lo que la interacción entre ellos no es deseable o, de producirse, es prácticamente nula, ya que de otra forma los recursos necesarios tanto de red como computacionales serían muy notables. Por norma general, el vídeo transportará el busto parlante del profesor, por lo que la calidad visual necesaria no es elevada.

Una clase magistral impartida con esta tecnología puede ser grabada para ser reproducida posteriormente. Incluso, puede ser redefinida como un objeto de aprendizaje reusable.

- Juegos y simulaciones. Son entornos virtuales que responden a las entradas de los usuarios. Las simulaciones sirven para que los participantes aprendan con la práctica y valoren las consecuencias de sus actos en un entorno seguro. Por otro lado, los juegos y simulaciones

pueden estimular las habilidades psicomotoras del individuo. Muchos son los autores que proclaman los beneficios de la verosimilitud, los entornos de aprendizaje inmersivos y los escenarios realistas basados en problemas.

- *Web conferencing*. Bajo este término se engloban aplicaciones para Internet altamente interactivas que incluyen un amplio conjunto de funcionalidades para el aprendizaje colaborativo. Estructuralmente estas aplicaciones están pensadas para escenarios multipunto a multipunto, de forma que podrían utilizarse con grupos muy amplios. Sin embargo, exhiben todo su potencial con grupos pequeños aprendiendo de forma colaborativa. En ese caso, es preferible utilizar una metodología facilitadora respecto a una didáctica tradicional. Más importante si cabe, la principal característica de esta tecnología es que posibilita la creación de redes y comunidades virtuales entre los participantes en las sesiones de e-learning.

Es habitual el uso del término e-learning síncrono para referirse al *Web conferencing*, sin precisar que también existen el resto de tecnologías mencionadas anteriormente y que también posibilitan el aprendizaje síncrono.

El e-learning en su modalidad síncrona ha sufrido una gran evolución a lo largo de muy poco tiempo. Fruto de esta evolución es la falta de buenas prácticas que recomienden cómo planificar, diseñar y llevar a cabo actividades de este tipo. Otro efecto colateral es la gran variabilidad en las herramientas que dan soporte a este tipo de enseñanza. Es frecuente que con el tiempo estas herramientas incorporen nuevas funcionalidades, lo que hace su catalogación en uno u otro campo más difícil.

En el capítulo 2 se realiza un amplio estudio de muchas de las herramientas de e-learning síncrono que se pueden encontrar en el ámbito comercial y de investigación, detallando sus características técnicas y funcionales.

1.2.3. Beneficios del e-learning síncrono

La utilización de herramientas de e-learning síncrono tiene una serie de ventajas:

- Permiten la conexión de alumnos dispersos, lo cual resulta especialmente apropiado para organizaciones distribuidas geográficamente. Específicamente, las empresas pueden obtener grandes beneficios formando a los empleados en sus propios lugares de trabajo.
- Permiten interactuar y colaborar en tiempo real, mimetizando las relaciones que se producen dentro de un aula tradicional. De este modo, se produce una relación natural entre los participantes de la sesión, que redundan en un *fluir* espontáneo de la misma. Las preguntas se plantean de forma inmediata al tiempo que las respuestas y aclaraciones se ofrecen directamente. Las relaciones sociales que permiten estas herramientas crean una sinergia de aprendizaje entre todos los participantes.
- Proporcionan una sensación de inmediatez, lo cual es muy útil para transmitir contenidos de última hora o información temporalmente sensible. Además, dado que la presencia del instructor es patente, sirve para aliviar las posibles ansiedades de los alumnos al enfrentarse a una experiencia de aprendizaje mecánica e impersonal diferente de la tradicional.
- Fomentan la creación de una comunidad de aprendizaje. Los alumnos comparten ideas entre sí y se benefician del intercambio de experiencias. Esto favorece el desarrollo de una sensación de conexión entre los alumnos. A largo plazo, mejora aspectos tan importantes como el trabajo en grupo y estimula la creación de una identidad grupal común. Además, fomenta el consenso y consolida valores como el respeto a las ideas del resto, lo que repercute en una moral más fuerte de cada uno de los miembros.

- Ayudan a balancear la dinámica de la clase al facilitar una experiencia de aprendizaje igualitaria entre los alumnos. Evita que alumnos más extrovertidos dominen el desarrollo de la clase frente a alumnos más tímidos como puede ocurrir en la enseñanza presencial tradicional. También, al evitarse la interacción cara a cara, se evita que los caracteres sociales o raciales afecten a las actividades de grupo.

Por otra parte, las herramientas que permiten a los alumnos responder de forma anónima fomentan la participación, puesto que se pierde el miedo a responder de forma errónea y quedar mal antes el instructor o los compañeros. Por el contrario, si se identifican las respuestas de los alumnos se introduce una competencia entre los alumnos en los ejercicios que el instructor plantea. Además, gracias a la utilización de distintos medios para impartir las sesiones, se produce una conexión más fuerte con los alumnos independientemente del modo de aprender que tengan.

- Incluyen funcionalidad adicional como pueden ser las pizarras compartidas que, gracias a la posibilidad de realizar anotaciones, fomentan el trabajo colaborativo. Adicionalmente, la compartición de aplicaciones aumenta las alternativas para un rápido y efectivo aprendizaje a través del trabajo en grupo.

1.3. Motivación

Con el paso de los años, según avanza el proceso de globalización, aumentan las necesidades formativas de las personas, especialmente para adquirir el conocimiento necesario para el desarrollo competente de su actividad profesional. Ya no basta con adquirir unos conocimientos mínimos válidos durante toda la vida laboral, sino que es preciso diseñar planes formativos a lo largo de la misma conducidos a actualizar conocimientos a la par de los avances tecnológicos. Es lo que se conoce como *formación continua*. Desde hace ya tiempo, las organizaciones empresariales se han concienciado de la importancia de ésta y, actualmente, les resulta imprescindible plantear políticas conducidas a la capacitación de personal como eje de su estrategia empresarial.

Por otro lado, es muy común que una gran multinacional esté constituida por diferentes factorías dispersas por un buen número de países. En ese caso, todas ellas deben regirse por las mismas directrices en sus procesos productivos y de control de calidad, impuestas a nivel corporativo buscando un fin común. De esta forma, existe un conjunto de procedimientos, normas y tecnologías que interviene en el día a día de la empresa, el *know-how*, que debe ser compartido entre todas sus sedes. Esto unido a que dicho conocimiento evoluciona según se introducen avances técnicos productivos y de gestión, necesarios para incrementar la competitividad de la empresa, otorga una mayor importancia si cabe al diseño de un buen plan formativo a nivel corporativo.

Por tanto, las empresas deben ofrecer a sus empleados la posibilidad de actualizar sus conocimientos de forma continua, para lo cual es habitual que dispongan de un departamento específico de formación de recursos humanos, que se encarga de la planificación y desarrollo de actividades orientadas a la formación de los mismos. Comúnmente, las grandes corporaciones destinan un centro específico para la impartición de estas actividades formativas. Ésta, aunque una solución muy flexible, es sólo factible para grandes empresas por el elevado coste que conlleva.

Varios son los factores que influyen en el coste final de un plan formativo. Por un lado, el mantenimiento de un centro de formación específico acarrea unos costes logísticos importantes, dependiendo del número de aulas y su dotación técnica. Desde otro punto de vista, deben tenerse en cuenta otros costes indirectos derivados de la utilización de una modalidad presencial para la capacitación de los empleados. Así, el coste final se ve influido por el número de horas de trabajo útil consumidas en la impartición de un curso. Esto es debido a que, habitualmente, este tipo de cursos se dividen en sesiones que se llevan a cabo durante la jornada laboral, en cuyo caso la actividad productiva de la factoría puede verse afectada por el menor tiempo que dedican los

empleados al desempeño de sus funciones. Por otro lado, el coste se incrementa cuando éstos deben desplazarse desde su puesto de trabajo habitual al centro de formación donde recibirán el curso. De esta forma, además del tiempo que se invierte en los desplazamientos, y en consecuencia, en horas hábiles de trabajo, también cuentan los gastos propios de locomoción, más acusados cuanto más lejos se encuentra el centro de formación.

Tabla 1.3: Gastos que deben afrontarse en función de la modalidad de enseñanza

Concepto	Presencial	A distancia
Mantenimiento centro de formación	✓	
Personal cualificado (instructores)	✓	✓
Desarrollo del material	✓	✓
Desplazamientos	✓	

Más allá de los gastos producto de un estrategia de enseñanza presencial, existen otros inevitables independientemente de la modalidad de enseñanza adoptada. Entre ellos podemos citar la disposición de personal cualificado para la impartición de cursos, el desarrollo del material formativo, etc. Si bien estos últimos resultan ineludibles, los primeros se pueden paliar en parte, y en algunos casos eliminarse, con la utilización de modalidades de enseñanza a distancia, puesto que ya no son necesarios desplazamientos y la planificación temporal de la actividad se flexibiliza; amén de las otras ventajas que aporta el e-learning vistas en el apartado 1.2. En la tabla 1.3 se detallan aquellos gastos que se eliminan con la utilización de una práctica educativa a distancia.

Por todas estas razones, la educación a distancia es una buena alternativa para la formación de personas que cobra cada día más importancia gracias, sobre todo, a la posibilidad de impartir contenidos *a la carta* y el ahorro de costes que supone. De ahí que el e-learning en sus modalidades síncrona y asíncrona esté recibiendo un gran impulso dentro de los entornos corporativos para la capacitación de los recursos humanos.

Especialmente en los sistemas de e-learning síncrono se ha venido produciendo una importante evolución en los últimos años, que ha tenido como consecuencia que cada vez más organizaciones lo consideren a la hora de diseñar sus planes formativos. Sin embargo, debido a la convergencia funcional que se está produciendo entre éstos y los sistemas de conferencia tradicionales, la distinción entre uno y otro tipo de herramientas es cada vez más difícil. Además, cada fabricante añade o excluye funcionalidades en base a criterios comerciales que dificultan aún más delimitar el campo de actuación de cada una de las herramientas. Resulta necesario, por tanto, determinar la funcionalidad mínima necesaria para una aplicación de e-learning síncrono dependiendo del entorno en el que vaya a operar. En la presente tesis, se caracterizan los requisitos funcionales de una herramienta de e-learning síncrono en el ámbito de la formación en grandes corporaciones. Para ello, debe tenerse en cuenta, principalmente, el perfil del usuario que utilizará la herramienta, ya sea ejerciendo el papel de instructor o de alumno, y las características de las redes de datos sobre las que opera la herramienta.

Una vez que los requisitos funcionales mínimos han sido establecidos, se implementa un prototipo de plataforma de e-learning síncrono para la formación de los recursos humanos de una gran corporación. En concreto, este trabajo se enmarca dentro de un proyecto de colaboración entre la Universidad de Oviedo y ArcelorMittal, denominado *eTipo*¹, para la implantación de una plataforma de e-training en su departamento de recursos humanos.

¹Referencia: CN-06-038.

1.4. Objetivos

Los dos principales objetivos que se plantea la tesis son:

1. La caracterización de los sistemas multimedia interactivos aplicados en entornos de e-learning síncrono, y en especial, en el ámbito de la capacitación de personal dentro de grandes corporaciones empresariales.
2. El diseño de un prototipo de herramienta siguiendo una metodología iterativa de prototipado, evaluación de alternativas y optimización, cuyo resultado final constituye un diseño óptimo para el caso concreto del e-training siguiendo los requisitos deducidos de la previa caracterización de las herramientas de e-learning síncrono.

1.4.1. Plan de trabajo

Para alcanzar los objetivos planteados, se desglosan a continuación los principales hitos en el plan de trabajo previsto:

- Caracterizar, en función de las herramientas existentes en la actualidad, las funcionalidades que presentan las aplicaciones de e-learning síncrono y extraer la funcionalidad básica común que presentan todos estos sistemas.
- Proponer un conjunto mínimo de requisitos para una herramienta de e-training dentro de grandes corporaciones a partir de la caracterización de las herramientas de e-learning síncrono y la características de los usuarios y la red con la que trabajará la herramienta.
- Evaluar la diferentes alternativas estándares para el diseño de cada una de las funcionalidades que una herramienta de e-training debe ofrecer.
- Plantear un diseño de prototipo de herramienta de e-learning síncrono optimizado para el caso de aplicación en la formación de los recursos humanos de una gran empresa.
- Implementar el diseño propuesto en una plataforma prototipo de e-learning síncrono que será utilizada en actividades formativas reales.
- Realizar diferentes pruebas para extraer la satisfacción de los usuarios en el uso del prototipo desarrollado durante sesiones de e-learning reales. Estas pruebas deben servir para refinar los sucesivos diseños para la implementación de la funcionalidad mínima que una herramienta de e-training debe satisfacer.

1.5. Organización de la documentación

A partir de este punto, la documentación de la presente tesis se organiza del siguiente modo:

- En el capítulo 2 se realiza un amplio estudio de las herramientas disponibles para realizar actividades de e-learning síncrono. Son analizadas tanto las herramientas comerciales como las destinadas a la investigación científica. Como conclusión del estudio se identifican cuáles son las características comunes que presentan todas estas herramientas. Así, se define una funcionalidad tipo que caracterizará a una herramienta de e-learning síncrono. Además, se define un mínimo de funcionalidad necesaria para la implantación de una herramienta de e-learning síncrono para la capacitación de personal dentro de una gran corporación.

- En el capítulo 3 se enumeran las diferentes tecnologías que posibilitan la implementación de una herramienta de e-learning síncrono. Para facilitar el análisis, se consideran una a una las diferentes capas que conforman una herramienta de este tipo y las opciones tecnológicas para su implementación.
- En el capítulo 4 se expone el prototipo desarrollado en base a los requisitos funcionales deducidos en el capítulo 2 y las alternativas tecnológicas expuestas en el capítulo 3, exponiendo las decisiones de diseño tomadas. Para la obtención del diseño final se sigue un proceso iterativo de evaluación de alternativas y optimización para el caso concreto de formación dentro de una gran corporación.
- En el capítulo 5 se describen las pruebas realizadas para valorar el funcionamiento del prototipo desarrollado en el capítulo 4. Se llevan a cabo dos tipos de pruebas: pruebas de laboratorio y pruebas sobre actividades de e-learning síncrono reales. Asimismo, se realiza un estudio de la satisfacción de los usuarios en el uso del prototipo mediante encuestas. Estos estudios sirven para refinar los requisitos funcionales del prototipo deducidos en el capítulo 2.
- En el capítulo 6 se presenta un resumen de los objetivos alcanzados y de las aportaciones de la presente tesis. A partir de estos datos se extraen unas conclusiones y se plantean posibles líneas de investigación que se pudieran derivar del presente trabajo.

Capítulo 2

Herramientas de e-learning síncrono

En este capítulo se realiza un amplio estudio de las herramientas que posibilitan el desarrollo de actividades de e-learning síncrono. Éstas se clasifican en dos grandes áreas, las herramientas comerciales y las concebidas en entornos académicos y de investigación. En el primero de los grupos se consideran también las soluciones basadas en hardware. El objetivo es detallar las características de cada solución para extraer una funcionalidad común a todas ellas que servirá para proponer una serie de requisitos mínimos que cualquier herramienta de e-learning síncrono debe cumplir. Además, a partir de las similitudes que presentan las interfaces de todas las herramientas, se propone una interfaz modelo que servirá de referencia en la implementación de este tipo de herramientas.

2.1. Introducción

Tanto en el mundo comercial como en el puramente investigador existen multitud de herramientas que pueden llegar a utilizarse para desarrollar sesiones de e-learning síncrono. Mediante un exhaustivo análisis de las características de estas herramientas es posible deducir un mínimo de funcionalidad común a todas ellas. Este mínimo de funcionalidad nos servirá para caracterizar este tipo de herramientas y encuadrar otras herramientas dentro del ámbito del e-learning síncrono para verificar si cumplen ese mínimo requerido.

Se ha realizado un estudio analizando hasta veinte herramientas comerciales y un buen número de herramientas desarrolladas en el campo de la investigación científica. Gracias a este análisis, se ha podido comprobar que, actualmente, las plataformas que permiten la comunicación en tiempo real mediante audio y vídeo entre interlocutores situados en diferentes localizaciones físicas han ido incrementando su potencial de tal forma que las diferencias entre unas y otras se han reducido ostensiblemente, a pesar de que originariamente fueron concebidas para distintos fines: e-learning, e-meeting, conferencias web, herramientas colaborativas, compartición de documentos, trabajo en paralelo sobre aplicaciones compartidas, etc. En mayor o menor medida, todas estas plataformas son válidas para su uso en e-learning síncrono. Dependiendo del enfoque original de cada una, ofrecerán mayores posibilidades para la realización de ciertas tareas y estarán más limitadas en otros aspectos.

Habitualmente, una herramienta de e-learning síncrono presenta las siguientes funcionalidades:

- Audio y videoconferencias: los participantes en la clase virtual se intercambian flujos de audio y vídeo con contenido didáctico; por ejemplo, las explicaciones del instructor o las preguntas de los alumnos. En este sentido, cobra especial importancia el audio del instructor en el que, generalmente, se transmitirá la mayor parte del contenido pedagógico del proceso de enseñanza.
- Mensajería instantánea (IM): una persona se comunica con otra, o muchas entre sí, mediante mensajes de texto tipografiados manualmente. Esta sería una solución apropiada cuando el intercambio de audio entre todos los participantes de la clase no sea posible por restricciones del ancho de banda disponible en la red, aunque, en otras ocasiones, es una

funcionalidad complementaria. Además, es posible mantener una lista con la disponibilidad de cada participante para establecer una conversación mediante lo que se conoce como *control de presencia*.

- Pizarra virtual: permite a un grupo de personas comunicarse mediante la exposición de diapositivas, la escritura de comentarios, dibujos o anotaciones, resaltando elementos o apuntándolos utilizando un puntero virtual.
- Compartición de aplicaciones: es muy útil cuando se desea demostrar el funcionamiento de una aplicación software a usuarios remotos. Un instructor podría dejar el control de una aplicación al alumno para que realice tareas y practique en el manejo de la misma.
- Escritorio compartido: permite que un usuario pueda ver el escritorio de otro usuario exactamente de la misma forma que lo ve el primero. Esto es una funcionalidad complementaria a la compartición de aplicaciones y, en algunas soluciones, es incluso posible ceder el control del equipo para trabajar sobre él de forma remota.
- Transferencia de ficheros: el instructor puede estar interesado en distribuir algún tipo de fichero entre el resto de participantes y no sólo los contenidos didácticos que se imparte. Es habitual ofrecer la posibilidad de intercambiar ficheros entre los participantes.

2.2. Herramientas de e-learning síncrono comerciales

En este apartado se analizan las herramientas que existen en el mercado para realizar actividades de e-learning síncrono, centrando la atención en aquellas herramientas con una mayor cuota de mercado. No obstante, el mercado del e-learning síncrono es un mercado tremendamente cambiante y el peso dentro del mismo de cada una de las herramientas dentro del mismo fluctúa con el tiempo.

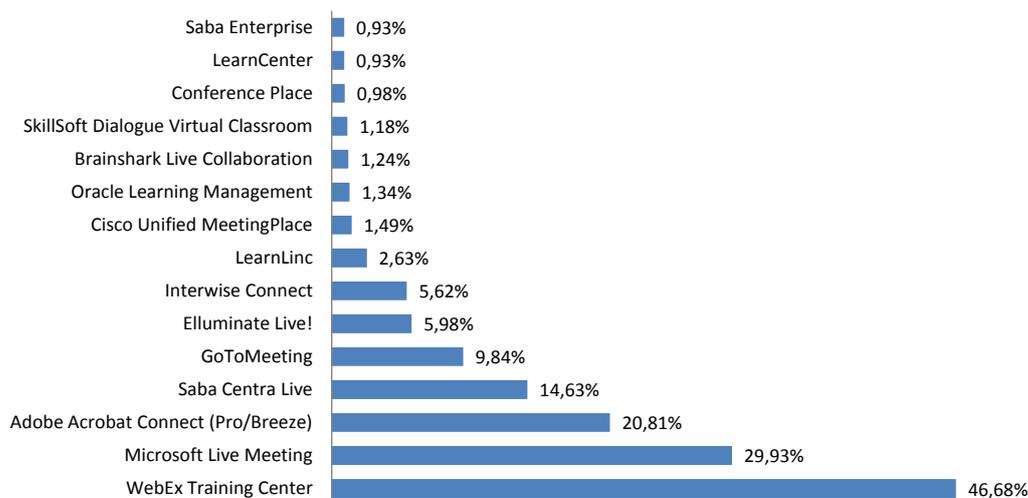
Un estudio de mercado realizado por *The eLearning Guild* [Guild, 2008] en junio de 2007 [Wexler *et al.*, 2007] arroja los resultados indicados en la figura 2.1. En dicha figura, pueden verse las cuotas de mercado que mantienen las herramientas de e-learning síncrono más representativas.

Uno de los motivos para la rápida evolución del mercado en los últimos años es el descubrimiento por parte de las grandes empresas del gran potencial del e-learning síncrono. Esto ha propiciado una serie de movimientos empresariales que han supuesto la entrada en escena de empresas cuyo ámbito de negocio no estaba relacionado con el e-learning. Por citar algún ejemplo, WebEx y Centra, las anteriormente referencias en el mercado de las soluciones síncronas, fueron objeto de sendas adquisiciones por parte de Cisco y Saba Software respectivamente. Otras grandes compañías se han posicionado en el mercado, casi siempre a través de compras. Así, Adobe adquirió Macromedia y con ella su producto Macromedia Breeze, pasando a denominarse Acrobat Connect Professional; AT&T compró Interwise y cambió el nombre de su producto estrella a AT&T Connect; Microsoft adquirió GrooveNetworks para integrar el producto Groove Virtual Office en su paquete ofimático Office; y, por último, Google entró en el mercado con la adquisición de Marratech.

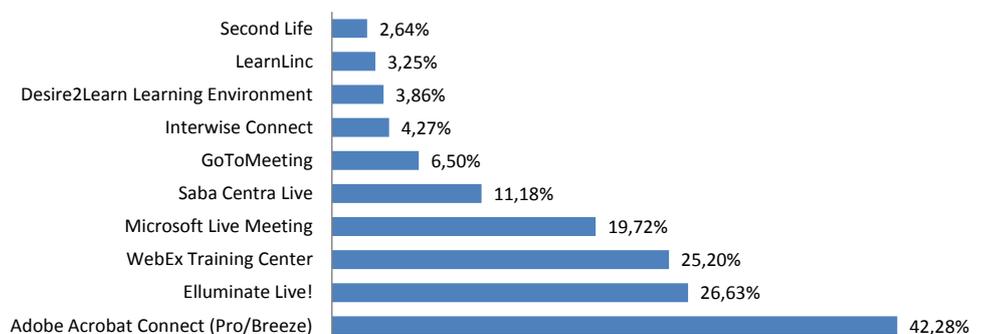
A continuación, se analizan individualmente las herramientas de e-learning síncrono más utilizadas.

2.2.1. WebEx Training Center

WebEx Training Center [WebEx, 2008] es el producto de WebEx específico para el e-learning síncrono. Otras variantes disponibles son: *Meeting Center* para el trabajo colaborativo; *Event Center* para realizar presentaciones a grandes audiencias; *Sales Center* para la promoción de ventas, y *Support Center* para implantar servicios de atención al cliente.



(a)



(b)

Figura 2.1: Cuotas de mercado en el ámbito (a) empresarial y (b) educativo y gubernamental

Su característica más importante es que es multiplataforma; soporta Windows, Linux, Mac OS y Solaris. Los clientes emplean únicamente un navegador web en el que sólo hay que instalar un plug-in y su conexión debe ser de al menos de 56 kbps. Se trata de un servicio ASP de pago por uso. Esto implica que los alumnos o conferenciantes, así como el instructor o ponente, deben acceder a la clase virtual por mediación de la red de WebEx, basada en una tecnología denominada *MediaTone*. Esta arquitectura es la que, según WebEx, más seguridad proporciona y garantiza una fiabilidad de conexión del 99,99 %. El otro posible despliegue de la red ofrecido por WebEx es instalar la arquitectura *MediaTone* en la red corporativa de la empresa que contrata el servicio. En cualquier caso, WebEx seguiría realizando las operaciones de mantenimiento y gestión de la misma.

Además de las características propias para e-learning síncrono, se incluyen otras adicionales de las que cabe mencionar las siguientes:

- Transmisión del sonido mediante VoIP (*Voice over IP*), lo que permite un mejor control de la calidad del audio y la compatibilidad con teléfonos VoIP. También permite gestionar el audio a través de conferencias telefónicas tradicionales empleando números proporcionados

por WebEx.

- Integración con Outlook, para gestionar los horarios de las conferencias e informar de las mismas.
- Hasta cuatro vídeos simultáneos en pantalla, de alumnos o profesores indistintamente.
- Posibilidad de combinar varios instructores, impartiendo clase desde ubicaciones diferentes.
- El instructor puede realizar encuestas o sondeos durante el desarrollo de la clase en cualquier momento.
- El instructor puede grabar la clase para ofrecerla en diferido a otros alumnos.
- La presentación del instructor puede estar escalada, ampliada o reducida, en un grado diferente en la pantalla de cada alumno.
- Permite realizar seguimientos de la evaluación de los alumnos.
- El instructor dispone de puntero virtual.
- El instructor puede controlar las operaciones que el alumno realiza en su equipo cliente.
- Posibilidad de dividir a los participantes en grupos más pequeños para la realización de actividades específicas.

Para programar una clase el instructor debe acceder a un portal de WebEx. Ahí configura opciones tales como los usuarios que podrán acceder a la clase y el horario de la misma. Cada persona accede a la clase mediante un applet Java, arrancando la aplicación WebEx Training Center, de la que se muestran tres capturas en las figuras 2.2, 2.3 y 2.4. En la imagen de la figura 2.2 puede verse la interfaz de la aplicación, que se compone de tres partes diferenciadas: la zona en la que se muestran las presentaciones o contenidos; el menú y la barra de herramientas para realizar modificaciones sobre la presentación, en la parte superior; y en la parte derecha, las ventanas deslizantes desde las que se puede realizar videoconferencia (en este caso se ve al instructor), chatear, tomar notas personales, ver los participantes, realizar encuestas entre los alumnos, y separar a los participantes en grupos de trabajo independientes (estas dos últimas operaciones sólo las puede realizar el instructor). Por ejemplo, desde estas ventanas el profesor puede otorgar a un alumno, o incluso a todos, el privilegio de realizar anotaciones sobre la presentación o de intervenir oralmente. Cualquiera de estas ventanas puede ser redimensionada o ampliada a pantalla completa. Mediante la pequeña barra de herramientas de la parte inferior derecha los alumnos pueden contestar a las encuestas o preguntas tipo test, así como *levantar la mano* para indicar al profesor que tienen alguna consulta que hacerle.

En las figuras 2.3 y 2.4 se presentan detalles particulares de la aplicación. En la primera se muestra la función de compartir o presentar por pantalla una aplicación que esté ejecutándose en el equipo del instructor. La segunda figura corresponde también al equipo del instructor, que puede ver una presentación tal y como los alumnos la están viendo y tener por encima varias ventanas flotantes, que los alumnos no ven.

2.2.2. Adobe Acrobat Connect Professional

Antes conocido como Macromedia Breeze, *Acrobat Connect Professional* [Adobe, 2008] es un producto con una filosofía similar a la de WebEx, ya que se presenta como un servicio ASP albergado en servidores propios, aunque es posible instalarlo, bajo licencia, en una máquina del cliente. El principal carácter diferenciador de este producto es el uso exhaustivo que hace

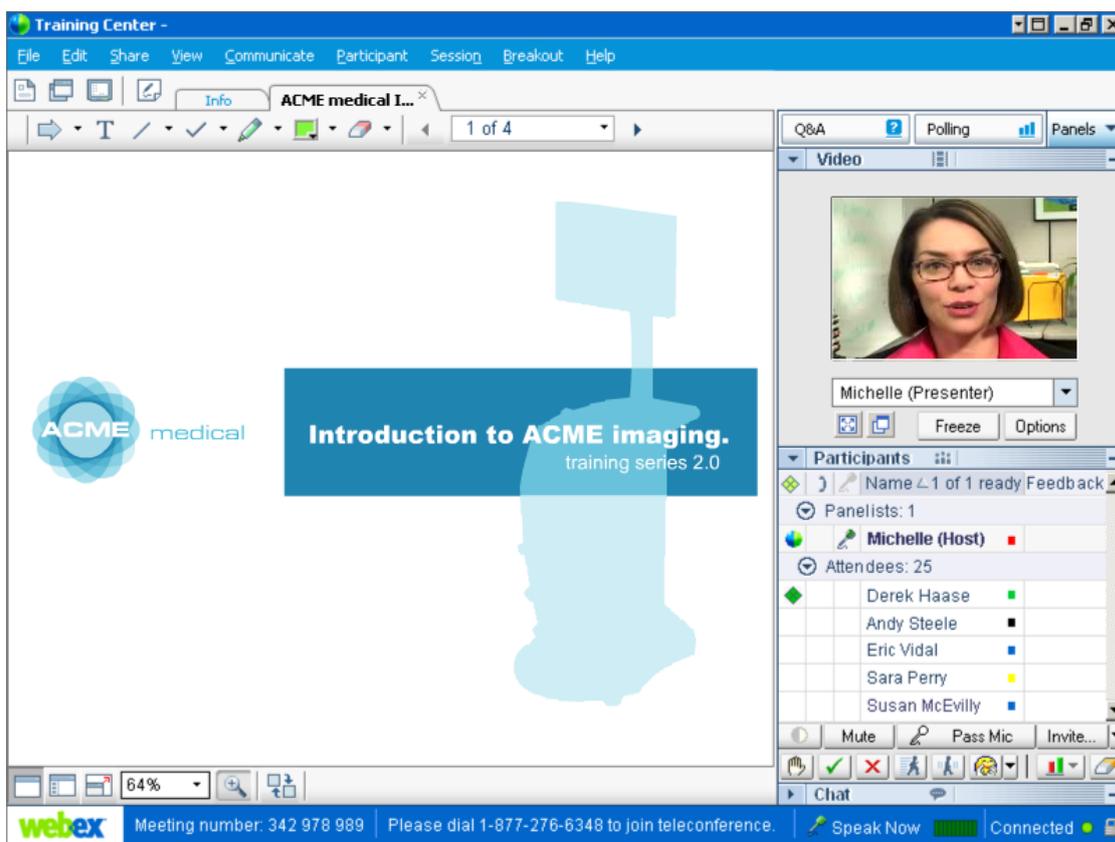


Figura 2.2: Pantalla principal de WebEx Training Center

de la tecnología Flash. Además, junto con WebEx Training Center, es la solución que presenta una interfaz más atractiva como puede observarse en la figura 2.5. Ésta se divide en segmentos denominados *pods*, cada uno de ellos conteniendo una funcionalidad determinada.

La solución Acrobat Connect Professional se presenta como tres productos relacionados: Acrobat Connect Professional, que permite realizar conferencias web en tiempo real; *Adobe Presenter*, para crear presentaciones bajo demanda en PowerPoint; y *Adobe Connect Enterprise*, que proporciona el servicio ASP.

Aunque inicialmente está orientado al e-meeting y al trabajo colaborativo, gracias a la inclusión de características adicionales lo hacen muy apropiado para el e-learning síncrono y, no en vano, es uno de los productos más utilizados en este campo. Entre las funcionalidades que ofrece se pueden citar:

- Utiliza tecnología VoIP para las conferencias de audio entre todos los participantes en la reunión. Es posible silenciar, dejar en espera e incluso expulsar participantes. También es posible atender a la reunión utilizando el teléfono tradicional a través de pasarelas.
- Las actividades que se desarrollan sobre la plataforma se organizan en *salas de reuniones* personales que están siempre disponibles y son configurables.
- Además de permitir compartir aplicaciones y el escritorio, es posible controlar de forma remota el escritorio o las aplicaciones de otros usuarios.

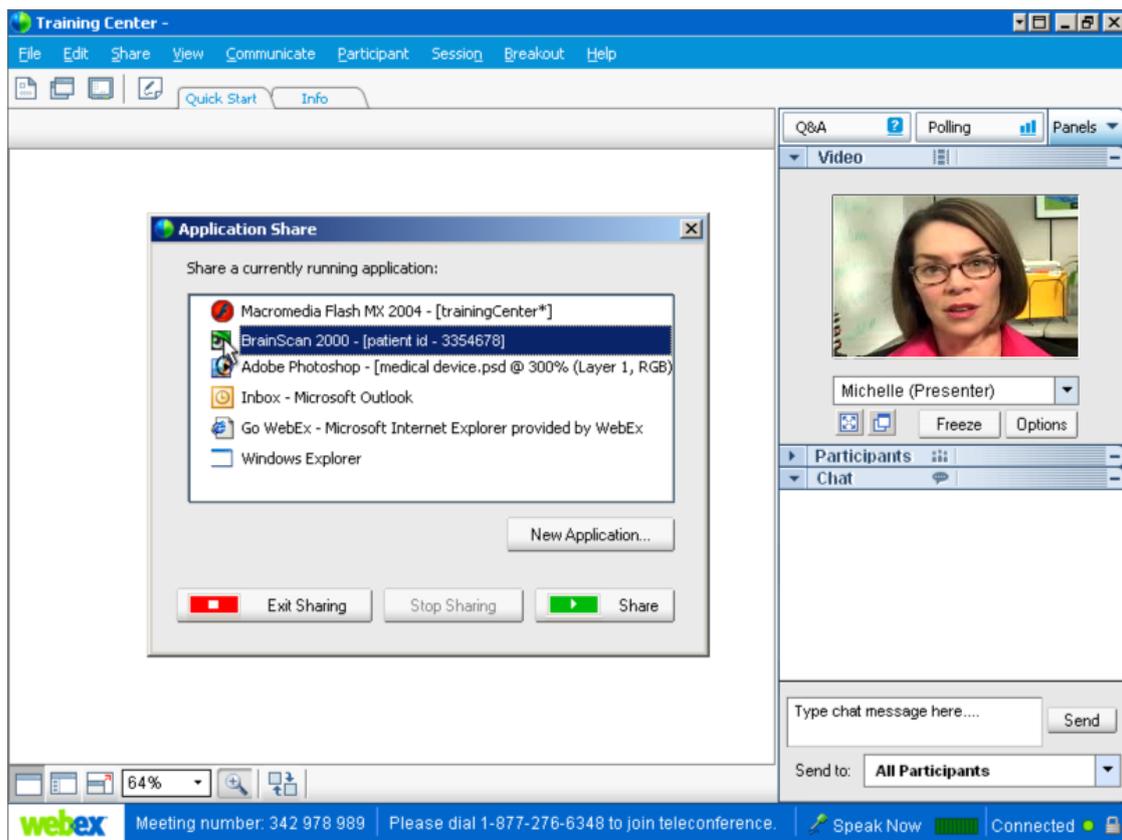


Figura 2.3: Funcionalidad para compartir un aplicación en WebEx Training Center

- Proporciona un SDK (*Software Development Kit*) para crear contenidos personalizados.
- Permite compartir documentos siempre que éstos sean imprimibles y previa conversión a formato Flash.
- Permite grabar las reuniones virtuales para ofrecerlas posteriormente bajo demanda.
- Ofrece la posibilidad de llevar a cabo estadísticas de la participación de los usuarios en las reuniones.
- Permite la integración con Outlook. Los usuarios pueden participar en las reuniones desde su calendario de Outlook.
- Es posible la asistencia de varios presentadores en la misma reunión virtual. En este sentido, se puede asignar este y otros privilegios a cada uno de los participantes de la reunión.
- El instructor puede plantear encuestas en cualquier momento de la clase.
- Ofrece un modo de preparación que brinda al instructor la posibilidad de modificar la sala de reunión o cargar los contenidos de presentación sin que el resto de usuarios puedan ver los cambios hasta que se desactive este modo.

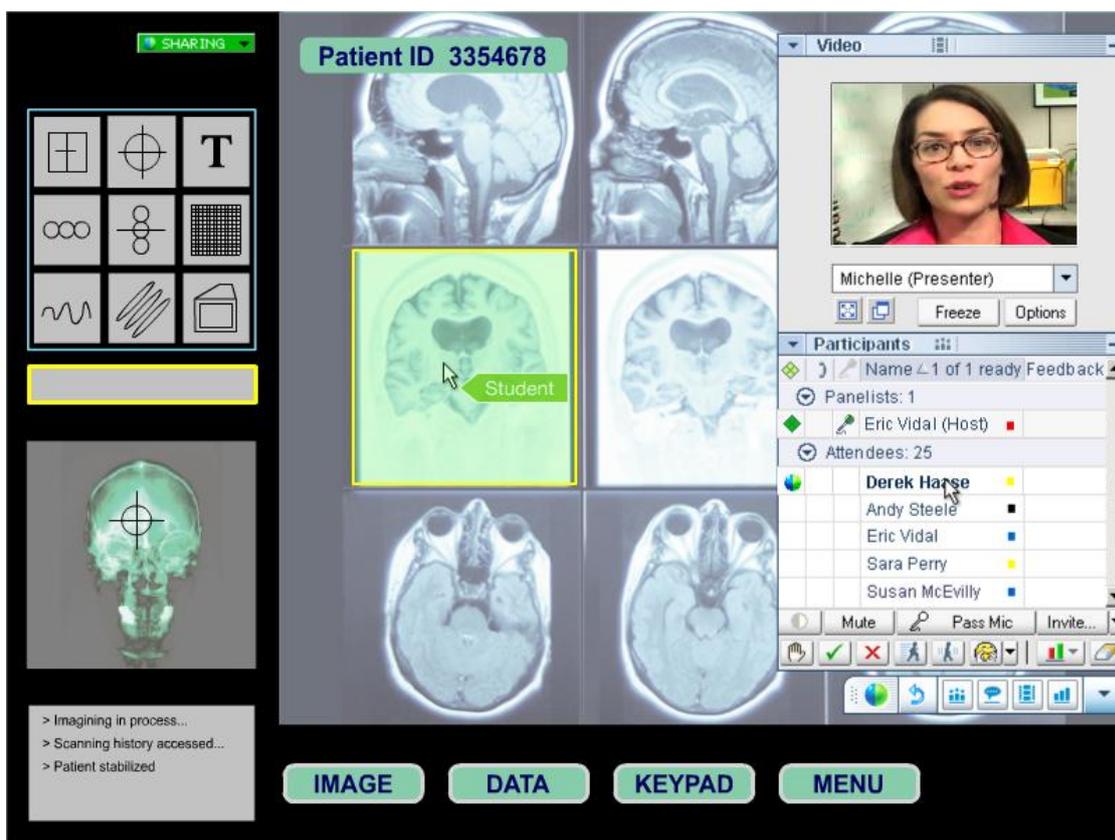


Figura 2.4: Interfaz de usuario del instructor en WebEx Training Center

Un participante puede acceder a la sala de reuniones virtual a través de un enlace que identifica dicha sala. Una vez que todos los participantes se encuentran conectados, el instructor o presentador puede comenzar la clase compartiendo su escritorio, una de las ventanas dentro del mismo o una aplicación concreta. La filosofía de funcionamiento de Acrobat Connect Professional es compartir los contenidos a utilizar durante la clase virtual a través de la compartición de la pantalla del presentador con el resto de participantes. Aunque también es posible compartir documentos directamente si éstos son convertidos previamente a formato Flash.

2.2.3. Microsoft Office Live Meeting

Este producto de Microsoft está basado en la aplicación Placeware, cuyo uso para e-meetings estaba bastante extendido, antes de que Microsoft adquiriera la compañía propietaria de la misma. *Live Meeting* [Microsoft, 2008b] está también basado en ASP. Se puede acceder a las reuniones mediante dos tipos de clientes: uno consiste en un programa instalado en el equipo del participante, del que puede verse una captura en la figura 2.6, mientras que el otro es un applet Java que se ejecuta accediendo a través de un navegador web. Esta última opción permite que la aplicación no esté únicamente limitada a usuarios con Windows.

Debido a su orientación hacia las reuniones de trabajo, Live Meeting ofrece menos funcionalidades adicionales que los productos anteriores:

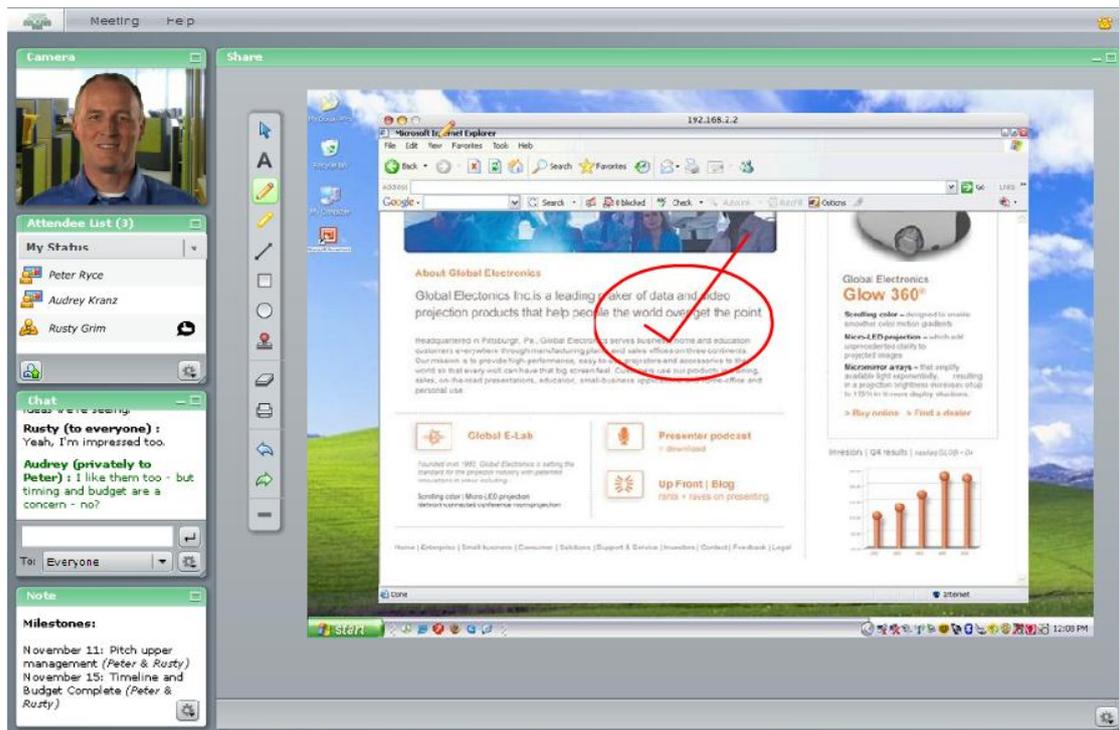


Figura 2.5: Pantalla principal de Acrobat Connect Professional

- La calidad del streaming de vídeo es muy limitada, con el objetivo de optimizar el rendimiento.
- Transmisión del sonido mediante VoIP, aunque sólo permite comunicación punto a punto entre asistentes a la reunión, excepto para el ponente, que es punto a multipunto.
- Posibilidad de compartir un área específica de la pantalla, ocultando el resto.
- Integración perfecta con Outlook y el resto de aplicaciones de Office, incluyendo botones de acceso directo para compartir inmediatamente documentos.
- El organizador de la conferencia puede grabar la reunión.
- El conferenciante dispone de puntero virtual.

La creación de la reunión es igual que con WebEx Training Center, accediendo a un portal web de Microsoft. La interfaz de la aplicación difiere sensiblemente de las de los dos productos anteriores. Existe una zona principal en la que se realiza la presentación y varias ventanas alrededor, cuyas funciones son mostrar un esquema de la presentación para facilitar la navegación por la misma, presentar la lista de participantes en la reunión así como su estado (presente, ausente, interviniendo...) en un gráfico a modo de mesa de reunión virtual y, por último, permitir el intercambio de mensajes de texto a través de la ventana de chat. También se puede activar la ventana de videoconferencia, pero no simultáneamente con la presentación en directo.

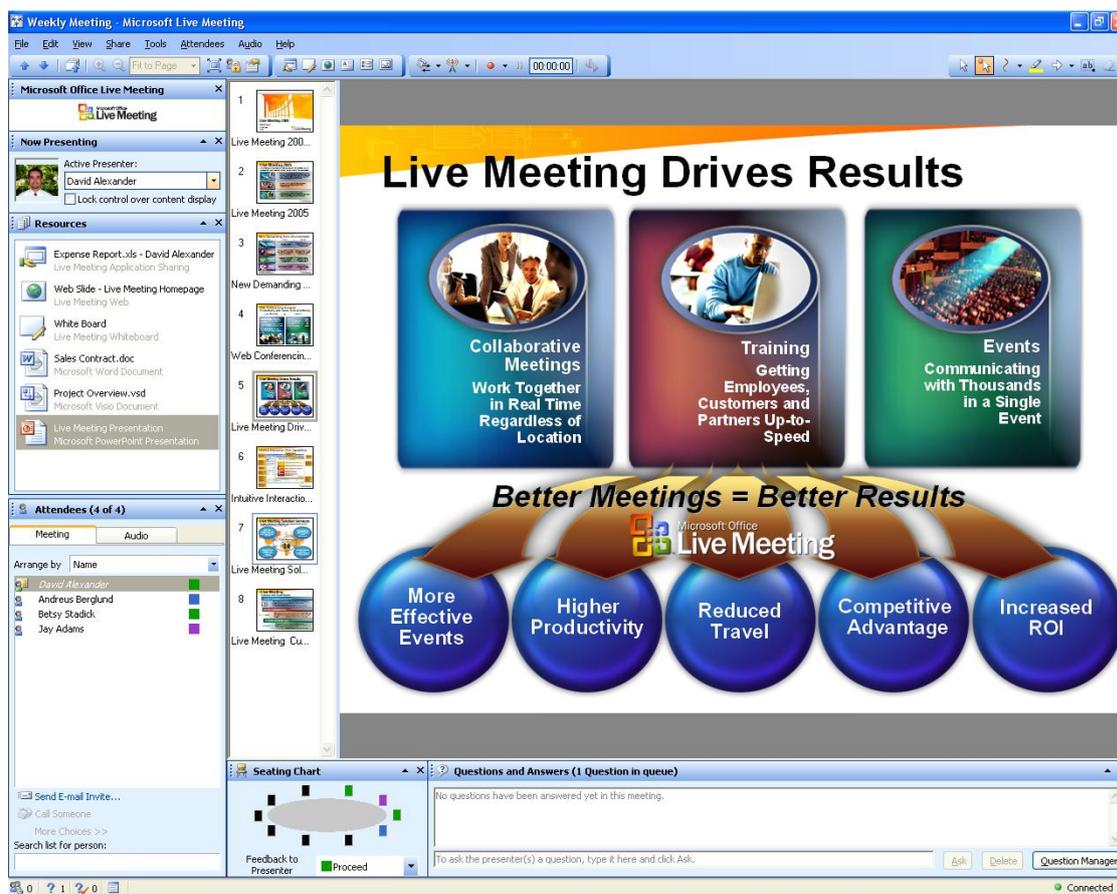


Figura 2.6: Pantalla principal de Live Meeting

2.2.4. Elluminate Live!

Elluminate Live! [Elluminate, 2008] es una de las herramientas que más ha crecido en los últimos años. Se trata de una plataforma disponible como producto completo cliente-servidor para su integración en la empresa y también como servicio ASP. Dispone de casi todas las virtudes de WebEx Training Center, Acrobat Connect Professional y Saba Centra Live en lo referente a las clases virtuales, destacando por encima de estos productos en el apartado de accesibilidad. Es multiplataforma, está traducido al español, consigue una transmisión VoIP de buena calidad y dispone de un SDK para añadirle nuevas funcionalidades. Se encuentra disponible en tres versiones: *Academic Edition*, *Enterprise Edition* y *Lite Edition*. Entre las funcionalidades que ofrece, además de las que se le suponen a una herramienta de e-learning síncrono, están:

- Audio de alta calidad utilizando VoIP.
- Importación de documentos de PowerPoint, StarOffice y OpenOffice.
- Posibilidad de realizar encuestas entre los alumnos en cualquier momento de la clase.
- Indicadores del estado de cada participante al igual que emoticonos que sirven para indicar su estado de ánimo.

- Creación dinámica de contenidos para ser compartidos durante la clase virtual.
- Posibilidad de copiar y pegar elementos desde el escritorio del instructor.

2.2.5. Saba Centra Live

Al igual que WebEx, Centra organiza su solución en diferentes productos, cada uno con su aplicación específica: *Saba Centra for eMeeting*, para reuniones colaborativas, *Saba Centra for Web Seminars*, para comunicaciones corporativas y grandes audiencias; y *Saba Centra Knowledge Center*, permite almacenar datos y recursos de forma centralizada. *Saba Centra Live* [Saba, 2008] es el heredero de Centra Symposium y sirve de núcleo para el desarrollo del resto de herramientas. Los nombres de estos módulos varían según el país, aunque se refieren siempre a las mismas aplicaciones, que funcionan bajo un núcleo central, denominado *Saba Centra 7*.

Saba Centra 7 sólo funciona en entornos Windows. Los requisitos para los clientes son similares a los de WebEx Training Center: un navegador web y una conexión mínima de 28,8 kbps. Los servidores deberán funcionar sobre máquinas con Windows e IIS (*Internet Information Server*). Aquí radica la principal diferencia entre la solución de Saba Centra y la de WebEx, mientras WebEx ofrece un servicio, Saba Centra vende el software completo, tanto cliente como servidor. Saba Centra Live se basa en una arquitectura de servidores distribuidos con administración centralizada. La adición de nuevos equipos que operen como servidores permitirá escalar el sistema para mayores audiencias (hasta 250 usuarios simultáneos en una sola sesión). En cualquier caso, Saba Centra ofrece la opción de instalar y mantener el despliegue del servicio (se seguiría vendiendo el software completo), así como la posibilidad de contratar el producto como un servicio ASP, de manera análoga a WebEx Training Center.

Hay que reseñar que Saba Centra Live, al igual que WebEx Training Center, es mucho más que una aplicación que permite dar clases virtuales en tiempo real. Además de los módulos que incluye para dar soporte a conferencias o reuniones colaborativas, el despliegue de Saba Centra Live contempla la posibilidad de instalar servidores de contenidos; es decir, repositorios en la red desde los que se puede descargar clases grabadas, contenidos bajo demanda o, simplemente, documentos o presentaciones de cursos. Del mismo modo, se puede disponer de portales web que ofrezcan servicios que van mucho más allá de las clases virtuales, tanto a alumnos (consultar horarios y participantes, acceder a material adicional del curso, foros de discusión...) como a instructores (gestionar los horarios, restringir la participación en los cursos, reservar recursos adicionales para clases que lo precisen, herramientas de notificación automática a alumnos...). Saba Centra también provee a los instructores de herramientas especializadas como un potente software de autoría para la creación de material para sus clases, y una aplicación de postproducción que permite al profesor sincronizar una presentación ya existente con voz, vídeo y otros contenidos multimedia, para emitirla posteriormente en falso directo.

Saba Centra Live posee todas las características adicionales de WebEx Training Center, excepto que no contempla la posibilidad de que el sonido se transmita mediante conferencia telefónica, y además incluye:

- Software en español (en total, doce idiomas diferentes).
- Integración completa con el entorno BackOffice de Microsoft.
- Soporte a bases de datos Oracle.
- Tecnología *Web Safari*, que permite navegar simultáneamente a todos los participantes por las webs que recorra el profesor. En WebEx Training Center esta funcionalidad se logra compartiendo la ventana del navegador.

- API (*Application Programming Interface*) disponible para integrar el producto con otras aplicaciones de la empresa y añadirle funcionalidades.

La forma de manejar la aplicación es análoga a la de WebEx Training Center, salvo que el instructor, en vez de acceder a un portal web, crea el aula virtual solicitándolo a un servidor perteneciente a la empresa. Todas las operaciones de gestión y mantenimiento se efectúan a través de un interfaz web que envía los mensajes pertinentes al servidor. Respecto a la interfaz de la aplicación, aún a pesar de ser menos llamativa que la de WebEx Training Center o Acrobat Connect Professional, ofrece las mismas funcionalidades y tiene la misma estructura, salvo que las opciones presentes en la parte derecha de la pantalla en WebEx Training Center se encuentran ahora a la izquierda de la misma.

Una pequeña diferencia es que Saba Centra Live permite tener en esta parte izquierda un esquema de las diapositivas de PowerPoint, Flash o PDF que se estén visualizando, para acceder rápidamente a una en concreto sin tener que salir del modo de pantalla completa.

2.2.6. GoToMeeting

GoToMeeting [GoToMeeting, 2008] es una de las últimas herramientas en adquirir una cuota de mercado significativa. Siendo originalmente una aplicación para el e-meeting, incorpora funcionalidades que la hacen apropiada para su uso en actividades de e-learning. GoToMeeting permite conferencias con hasta 15 usuarios simultáneos; 25 en el caso de la versión *Corporate*, pero siempre utilizando exclusivamente audio. Al igual que otras soluciones, ofrece una herramienta alternativa, *GoToWebinar*, cuando el objetivo es difundir seminarios web a un elevado número de usuarios.

La filosofía de funcionamiento es la misma en la que se basa Acrobat Connect Professional: la clase virtual comienza cuando el instructor comparte su escritorio o una aplicación específica con el resto de la clase. A partir de ese momento, el instructor puede incluso cederle el control de su teclado o ratón a un alumno para que opere de forma remota.

Algunas de las funcionalidades que presenta son:

- Integración con herramientas de mensajería instantánea para unirse a una reunión virtual desde las aplicaciones de mensajería como MSN Messenger.
- Posibilidad de que el instructor delegue su puesto en cualquier alumno.
- Se pueden ocultar ventanas e iconos cuando se comparte el escritorio con el resto de la clase.

2.2.7. Otras herramientas de interés

Existen otros muchos productos comerciales relacionados con facetas del e-learning síncrono, sobre todo en lo relativo a efectuar presentaciones para múltiples audiencias, realizar videoconferencias multipunto a multipunto y compartir documentos, pizarras virtuales o aplicaciones. Salvo que se indique lo contrario, las aplicaciones que se describen en esta sección cumplen las características básicas comunes de las plataformas de e-learning vistas anteriormente.

No se comentan aquí productos que se podrían catalogar como auxiliares, es decir, aplicaciones exclusivamente de vídeo o audioconferencia empleadas para complementar plataformas de e-learning o de e-meeting que, bien carecen de esas funcionalidades, o bien no las desarrollan con suficiente calidad. Dentro de esta categoría de productos se encontraría *Microsoft Portrait*, que es un prototipo nacido de un proyecto de investigación que Microsoft ofrece gratuitamente. Dispone de funcionalidades básicas de chat, audio y videoconferencia. Tiene por objeto lograr comunicaciones fluidas con muy poco ancho de banda (incluso por debajo de 10 kbps). Lógicamente la

calidad y el tamaño del vídeo son mínimos. Su otro punto fuerte es la integración con otras tecnologías de Microsoft (.NET Messenger Service, Windows y MSN Messenger, Live Communications Server 2003) y la compatibilidad con estándares públicos (SIP, ILS) y diferentes dispositivos (PC, Pocket PC, Smartphones); aunque, por supuesto, siempre bajo entornos Windows. En resumen, Portrait es en esencia un NetMeeting moderno, en cuanto a la optimización del tratamiento de los flujos de audio y vídeo se refiere, pero está desprovisto de las herramientas colaborativas que otras aplicaciones pueden proporcionar de forma más eficiente.

Otras opciones empleadas en la práctica se basan en asegurar o mejorar mediante otra aplicación la comunicación oral, esto es, intentar garantizar que en todo momento habrá una audioconferencia fluida, independientemente de las incidencias que puedan afectar al vídeo o a la presentación. *Net2Phone* y *Skype* son los dos productos más demandados para este fin. Ambos permiten establecer comunicaciones orales punto a punto a través de Internet con VoIP y también llamar a teléfonos convencionales; aunque para esta última tarea *Net2Phone* tiene mejores características y a menor coste. *Skype* permite efectuar también audioconferencias con un máximo de cuatro participantes.

Además de aplicaciones de vídeo y audioconferencia, otro grupo de herramientas auxiliares bastante empleadas son las de mensajería instantánea, es decir, chat. El objetivo de su uso es mantener un cauce de comunicación fiable y con una exigencia mínima de ancho de banda. Las aplicaciones de chat permiten, además de intercambiar mensajes de texto, efectuar un control básico de presencia, es decir, comprobar que el resto de interlocutores está conectado. Este tipo de aplicaciones se emplea asiduamente como apoyo en las clases y conferencias virtuales, dado que la mayoría de usuarios ya están familiarizados con su uso. *MSN Messenger*, de Microsoft, es el cliente de chat más usado. Cuenta con la posibilidad de realizar llamadas telefónicas y videoconferencias, aunque no son su principal cometido. Otras aplicaciones de chat muy extendidas son las que se basan en el protocolo XMPP (*Extensible Messaging and Presence Protocol*); el hecho de que el protocolo sea abierto permite la existencia de múltiples clientes diseñados para distintas plataformas. XMPP fue creado por la comunidad *Jabber*. La funcionalidad de chat de algunos de los productos comentados en este documento se basa en este protocolo, lo que permite que usuarios de la aplicación de e-learning puedan comunicarse con otras personas que no tengan acceso a la misma en un momento dado.

LearnLinc [iLinc, 2008] es un producto específico de e-learning que posee muchas de las funcionalidades de *WebEx Training Center* y *Saba Centra Live*, pero que se centra en la conferencia web, permitiendo la interacción en tiempo real, utilizando audio VoIP y vídeo multipunto. Además, incluye muchas de las funcionalidades que soportan los productos más extendidos: compartición de aplicaciones y escritorio, encuestas, división de los participantes en grupos, vigilar la pantalla de los alumnos, etc. Implementa un completo control de presencia que permite conocer la participación de los alumnos en la clase a través de un conjunto de iconos que representa el estado del alumno.

NetMeeting [Microsoft, 2008a] es una aplicación que lleva bastantes años sin renovarse. Se puede describir como un Live Meeting modesto, tan modesto que es gratuito, viene integrado en Windows. Aun así, es la plataforma escogida por muchas empresas para efectuar videoconferencias. Existen versiones para Linux y otros sistemas operativos. Por supuesto, la aplicación no requiere de un ASP para iniciar las conferencias. No posee ninguna característica adicional orientada al e-learning. Permite efectuar correctamente las operaciones básicas comentadas anteriormente y está traducido al español. Existe un SDK para personalizar la aplicación y combinarla con otros productos para añadirle más funcionalidades. Microsoft ya no proporciona soporte a este producto.

Se trata de una aplicación muy fácil de manejar. Cuando dos usuarios inician una conferencia, se puede ir aumentando el número de participantes. De las herramientas descritas es la menos indicada para el e-learning, ya que tareas como la compartición de pantalla están mucho menos optimizadas, tanto en rendimiento como en aspectos estéticos; no se puede seleccionar qué parte

de una aplicación o de una pantalla se desea que sea vista por el resto de participantes.

AT&T Connect [AT&T, 2008], antes conocido por Interwise, es un software que se vende íntegramente y su filosofía tiene la misma base que Saba Centra Live (servidores distribuidos). De alguna manera, se puede considerar a AT&T Connect como el hermano pequeño de este último. Es un producto muy polivalente, mucho más encaminado al e-learning que las soluciones de Microsoft. En cualquier caso, no hay motivo, ni siquiera económico, para otorgar mayor reconocimiento a esta aplicación frente a WebEx Training Center, Acrobat Connect Professional o Saba Centra Live.

Lotus Sametime [IBM, 2008] es una aplicación claramente orientada al trabajo cooperativo compartiendo escritorio y aplicaciones, que permite efectuar videoconferencia, exclusivamente punto a punto, pero con una calidad poco notable. Es posible realizar audioconferencias entre los participantes de la clase virtual utilizando técnicas VoIP. Lotus Sametime se comercializa en tres versiones, desde la versión más básica denominada *Entry*, hasta la más completa representada por la versión *Advanced*; a medio camino se sitúa la versión *Standard*.

Horizon Wimba Classroom [Wimba, 2008], hasta hace poco Horizon Live, es otro servicio ASP enfocado al e-learning, tanto síncrono como asíncrono, que se diferencia del resto por tener muchas más posibilidades en el terreno de la enseñanza asíncrona, incluyendo capacidad de integración con entornos como WebCT y Blackboard. Por ello, la mayoría de los clientes de este producto pertenecen a comunidades universitarias.

Intercall [Intercall, 2008], antes Raindance, es también un servicio ASP, muy enfocado a las conferencias y al e-meeting, principalmente entre miembros de diferentes empresas. Ofrece dos clientes, uno con todo el conjunto de funcionalidades disponibles y el otro con un mínima funcionalidad, ideal para usuarios que sólo desean recibir el vídeo del instructor.

Marratech [Marratech, 2008] es una aplicación multiplataforma y multilenguaje para e-meeting basada en la filosofía de Saba Centra Live de servidores distribuidos. Posee todas las funcionalidades esperadas para una aplicación de e-learning síncrono, salvo detalles como la imposibilidad de realizar encuestas o votaciones sobre la marcha. Es la plataforma que más información ofrece acerca del despliegue de la red y de su escalabilidad, así como de la seguridad de los servidores. Una de sus virtudes es la más que notable calidad de la voz en las transmisiones con VoIP. La interfaz de usuario es bastante similar a las de Saba Centra Live o WebEx Training Center. Dispone de una API para poder personalizar y comunicar el software con otras aplicaciones. Marratech se puede contratar también como servicio ASP.

e/pop Web Conferencing [WiredRed, 2008] es otro producto basado en servidores distribuidos que destaca por las opciones disponibles para presentar y compartir documentos, en particular, de PowerPoint. Tanto en características como en la cantidad de información disponible es una aplicación bastante similar a Marratech; la mayor diferencia entre ambas es que aquella exigía menos requerimientos técnicos que ésta. Sólo funciona bajo entornos Windows. Se proporciona un SDK para ampliar y personalizar la aplicación, y una API para que otras aplicaciones puedan interactuar con el servidor. El diseño de la interfaz de la aplicación sigue la tendencia de la mayoría.

Groove Virtual Office [Groove, 2008] no es realmente una aplicación de e-meeting. Está orientada a la compartición de archivos y a la colaboración entre grupos de trabajo sobre proyectos comunes. Por tanto, no está pensada en ningún momento para aplicaciones de e-learning, pero sus funciones de presentación de documentos cumplen con las funcionalidades básicas. La aplicación ha sido creada por el mismo equipo que desarrolló Lotus Notes y se basa en redes P2P (punto a punto). Esto quiere decir que se adquiere como un producto integral y que no necesita de ningún servidor para funcionar entre los clientes. En cualquier caso, para obtener mejores prestaciones en la videoconferencia la compañía recomienda instalar servidores intermedios entre los usuarios.

MegaMeeting [MegaMeeting, 2008] es un producto que se centra en la calidad de la videoconferencia y posee menos funcionalidades en otros aspectos. Aún así las versiones Professional (accesible como servicio ASP) y Enterprise (solución integral) permiten realizar exposiciones y el

resto de tareas básicas de manera aceptable para el e-learning.

VoxWire Web Conferencing [Voxwire, 2008], al igual que MegaMeeting, se trata de una aplicación que consigue audioconferencias de muy buena calidad. Pero en el resto de los aspectos se ciñe únicamente a una funcionalidad general: presentar contenidos al resto de conferenciantes. También permite intercambiar mensajes de texto (chat), pero no ofrece opciones para trabajar de forma compartida con aplicaciones.

HotConference [HotConference, 2008] es un producto basado en servicios ASP. Dispone de todas las funcionalidades básicas y permite cambiar el modo en el que los clientes visualizan la presentación con extrema facilidad; además de que el instructor puede controlar qué están viendo en todo momento.

PictureTalk [Pixion, 2008] es una solución multiplataforma con una pronunciada orientación hacia el e-learning; por ejemplo, contempla la posibilidad de realizar preguntas tipo test o encuestas entre los asistentes, característica incluida únicamente en los productos específicos de e-learning síncrono. Existen tres versiones: Personal, Professional y Enterprise. Las dos primeras se ofrecen como servicios ASP, mientras que la tercera se puede implantar íntegramente en una red propia.

Wave Three [Wave Three, 2008] es una plataforma de la que se venden dos módulos: *IP Communications* y *Conference Server*. El primero permite realizar videoconferencias punto a punto entre dos clientes que empleen este software. El segundo módulo es necesario para desplegar uno o más servidores intermedios que posibiliten la videoconferencia multipunto. La aplicación es multiplataforma y posee varias de las funcionalidades presentes en otros productos, además de la videoconferencia, pero no están integradas de la forma más óptima para su uso en e-learning, al igual que sucedía con las alternativas de Microsoft. En cualquier caso, hay que destacar que posee una gran interoperabilidad con otros programas, debido a que se ajusta fielmente a multitud de estándares, tanto de vídeo como de audio y otros protocolos de control.

Similar a HotConference, *WebConference* [WebConference, 2008] proporciona más funcionalidades añadidas enfocadas al e-learning y al e-meeting. Se oferta como servicio ASP sin limitación en el número de clientes simultáneos, aunque también puede alojarse en servidores propios.

2.2.7.1. Soluciones basadas en hardware

Aunque este trabajo está centrado en productos software hay que reseñar que existen compañías, como Polycom, Tandberg y Cisco, que ofrecen soluciones hardware específicas con su correspondiente software. Estas soluciones requieren una inversión mucho mayor, ya que cada cliente necesita un equipo específico, además de la licencia del software. La calidad de la videoconferencia es netamente superior a la del resto de alternativas, ya que el hardware está diseñado específicamente para este fin. Estos productos disponen de tantas funcionalidades como WebEx Training Center o Acrobat Connect Professional, y su interfaz suele ser bastante similar a la de éstos. En definitiva, las soluciones consistentes en hardware y software específicos son las más especializadas para e-learning y e-meeting, y las que mejores prestaciones ofrecen, aunque no tienen la flexibilidad de un producto íntegramente software.

2.2.8. Comparativa

En la tabla 2.1 se resumen las características de los productos comerciales revisados.

2.3. Herramientas de investigación

Dentro del ámbito de la investigación existen pocas herramientas que proporcionen toda la funcionalidad que requiere un sistema de e-learning síncrono completo. Habitualmente, se encuentran

Tabla 2.1: Comparativa de las plataformas de e-learning síncrono comerciales

	Servicio ASP	Despliegue					Funcionalidad										Conteni.	A/V	Sesión		Extra														
	Plataforma implantable en equipos propios	Multiplataforma	Acceso mediante <i>plug-in</i> en navegador web ^a	Integración en entornos empresariales ^b	Gestión de horarios (Outlook y Notes)	Múltiples idiomas (español incluido)	Presentación de contenidos específicos ^c	Audioconferencia	Videoconferencia	Puntero virtual	Pizarra virtual	Chat	Control de presencia	Transferencia de ficheros	Preguntas, encuestas, votaciones	Aplicaciones y escritorio compartidos	Compartir captura instantánea de pantalla	Control remoto	Navegación web simultánea ^d	Guardar capturas de la presentación	Diferentes modos de visualización ^e	Anotar sobre las presentaciones	Realizar y guardar anotaciones personales	Audioconferencia multipunto	Videoconferencia multipunto	Voz sobre IP	Figura de instructor	Varios instructores	División en grupos de trabajo	Vigilar la pantalla del alumno	Realimentación instantánea ^f	Grabación y almacén de clases en servidor	Grabación individual de clases en clientes	API / SDK	Comunicaciones seguras (SSL, HTTPS)
WebEx Training Center	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Acrobat Connect Pro.	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Office LiveMeeting	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Elluminate Live!	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Centra Symposium	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
GoToMeeting	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
LearnLink	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
NetMeeting	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
AT&T Connect	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
SameTime	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Horizon Wimba	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Intercall	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Marratech	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
e/pop	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Groove	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
MegaMeeting	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
VoxWire	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
HotConference	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
PictureTalk	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Wave Three	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
WebConference	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

^aLos productos que no disponen de esta característica para entrar en una clase o conferencia virtual, emplean programas cliente que hay que instalar previamente en los equipos. Algunos de los productos señalados en esta columna que incluyen esta característica, ofrecen las dos opciones (navegador web y cliente instalable).

^bPor ejemplo Office, BackOffice, SharePoint o LDAP.

^cPor ejemplo PowerPoint, PDF, Word, Excel, Autocad, imágenes o vídeos. No todas las aplicaciones señaladas en esta columna soportan los mismos tipos de documentos o contenidos. El único contenido aceptado por todas es el de presentación PowerPoint (PPT).

^dCasi todos los productos anuncian esta funcionalidad de manera explícita, aunque la mayoría de éstos se limitan a compartir la ventana del navegador web del instructor como una aplicación más, lo cual imposibilitaría la interacción de cada alumno con su navegador local.

^ePor ejemplo pantalla completa.

^fLevantar la mano, mostrar acuerdo o desacuerdo, emoticonos, etc.

soluciones parciales que vienen a cubrir parte de dicha funcionalidad. En ocasiones, incluso, se utilizan varias herramientas independientes, cada una con una funcionalidad específica, para llevar a cabo actividades de enseñanza a distancia síncrona de forma satisfactoria.

Una herramienta pionera fue la desarrollada por Latchman y otros [Latchman *et al.*, 1999]. Esta herramienta está diseñada para capturar el audio y vídeo del instructor durante el desarrollo de una clase tradicional y transmitirlos al equipo de los alumnos remotos. Dado que la resolución del vídeo del instructor es baja, este sistema utiliza dos cámaras adicionales, una para la captura de las notas de clase, y otra para las anotaciones manuscritas; ambas producidas por el instructor. El vídeo así capturado por estas dos cámaras se visualiza en dos ventanas independientes en la pantalla del equipo de los alumnos. Existe también una ventana de chat a través de la cual los participantes de la clase se pueden comunicar unos con otros de forma instantánea.

Actualmente, una aproximación como ésta, basada exclusivamente en flujos de vídeo, está obsoleta. Hoy en día, las notas de clase suelen venir en forma de documento electrónico que se transmite directamente a los alumnos. De esta forma, se evitan los problemas relacionados con la resolución del vídeo. Asimismo, las notas manuscritas se realizan sobre las notas de clase para mejorar el proceso de aprendizaje. En [Latchman *et al.*, 2001] los autores presentan varias mejoras e introducen el uso de la herramienta en un entorno mixto de e-learning síncrono y asíncrono.

Otra de las primeras herramientas desarrolladas es la descrita en [Deshpande y Hwang, 1999]. Los autores proponen un sistema de enseñanza a distancia basado en el concepto de clase virtual multimedia interactiva. Este sistema ya maneja diapositivas electrónicas siguiendo una filosofía cliente/servidor, para lo cual utiliza una herramienta denominada *SlideCast*. Esta herramienta transporta las anotaciones textuales que realiza el instructor sobre las diapositivas como un flujo de vídeo. Sin embargo, con bajo ancho de banda las técnicas de codificación clásicas no codifican este tipo de datos con la suficiente fidelidad. Para solucionarlo, los autores proponen una nueva codificación específica para este caso. Por otra parte, la interfaz de los alumnos consiste en un cliente específico para recibir el vídeo del instructor y sus anotaciones, y un navegador web con un plug-in instalado para recibir las diapositivas electrónicas. Sin embargo, el uso de ventanas independientes es una característica no deseable. La plataforma global resulta en un sistema complejo que requiere una MCU para centralizar todas las comunicaciones. Más tarde, en [Deshpande y Hwang, 2001], los mismos autores proponen diversas mejoras al sistema original.

En [Bouras *et al.*, 2002] los autores desarrollan una aplicación para la enseñanza a distancia síncrona que integra un módulo muy simple para el control del turno de palabra en la clase, basado en una máquina de estados, que no requiere soporte por parte de una MCU. El instructor actuaría como realizador de la clase otorgando el turno de palabra a alumnos concretos para que puedan dirigirse oralmente al resto de la clase. Sin embargo, no se presenta ningún estudio de la sobrecarga que impone en el instructor este modo de gestionar la clase. La aplicación desarrollada puede operar tanto en redes RDSI como en redes IP. Entre las funcionalidades que esta herramienta presenta se encuentran la videoconferencia, utilizando la recomendación H.323, la pizarra virtual, el chat y la compartición de aplicaciones.

WVOC (*Web-based Virtual Online Classroom*) es un entorno para el desarrollo de clases virtuales en línea [Yang y Liu, 2007]. Este entorno está a su vez integrado por otros dos: un entorno de comunicación instruccional (ICE) y un entorno de aprendizaje colaborativo (CLE). El objetivo del ICE es proporcionar soporte a las interacciones síncronas entre los instructores y los alumnos utilizando técnicas de difusión de vídeo, chat y pizarras electrónicas. Este entorno incluye servidores de difusión de medios, servidores de cursos bajo demanda, servidores de streaming y servidores de recursos. La difusión de los contenidos se realiza a través de técnicas de streaming de vídeo bajo demanda. La sincronía se consigue con el avance o retroceso dentro del vídeo a instancias del instructor. En definitiva, se trata de una arquitectura híbrida, inicialmente orientada al e-learning asíncrono, adaptada para su uso de forma síncrona, pero que adolece de falta de funcionalidad específica para el e-learning síncrono, además de ser extremadamente compleja.

Otra herramienta interesante es la que se propone en [Xinyou *et al.*, 2006]. Se trata de un sistema altamente interactivo destinado a la educación a distancia cuyo objetivo es emular un aula tradicional. Esta herramienta está totalmente orientada al aprendizaje colaborativo, en el que el instructor interactúa de forma continua con los alumnos en tiempo real, y está dividida en

varios subsistemas, todos ellos gestionables desde un panel de control. El subsistema de vídeo se utiliza para transmitir la imagen del instructor, las diapositivas y los vídeos educativos a los alumnos. El subsistema de vídeo transporta la voz del instructor hacia los alumnos y viceversa, esto es, una pregunta de cualquier alumno al resto de participantes de la clase. Por otra parte, el subsistema seminario se utiliza como área de discusión, en la cual los alumnos y el instructor pueden intercambiar vídeos, mensajes de audio o de texto, y cualquier otro contenido educativo. En esta herramienta, tanto el panel de control como el resto de subsistemas requiere una ventana independiente en la pantalla, a excepción del subsistema seminario que puede necesitar de más. Todas estas ventanas deben ser dispuestas en pantalla por el usuario, lo que aumenta la complejidad de la herramienta e implica la necesidad de que los usuarios de la misma estén muy familiarizados con el uso de entornos de ventanas avanzados. Por último, los mismos autores proponen un modelo de clase virtual que se encuentra integrado en su sistema de educación a distancia [Xinyou y Yan, 2006]

En [Snow *et al.*, 2005] se propone un sistema para habilitar la educación a distancia síncrona a través de Internet denominado NEW (*Network EducationWare*). Este sistema es muy sencillo de configurar y de operar, es de código abierto, totalmente gratuito para instituciones educativas y es capaz de operar de forma aceptable con un ancho de banda reducido. Sin embargo, no se trata de una solución monolítica, sino que se construye utilizando herramientas ya disponibles. Para la videoconferencia se utiliza la herramienta VIC, para el audio *SpeakFrealy* y para la pizarra WBD; añadiendo una ventana de interfaz para cada una de ellas. Estas tres herramientas de cada cliente se comunican con una capa que emula el comportamiento de una red multicast. Por último, se proporciona un módulo para la gestión del turno de palabra que es necesario para coordinar la operación de todo el sistema. De nuevo, estamos ante una herramienta con múltiples ventanas independientes que deben ser dispuestas en pantalla por el usuario. Una herramienta que realmente se considere de fácil uso debería integrar toda la funcionalidad en una sola ventana.

[Fung y Ledesma, 2005] detallan la utilización de la plataforma VITL (*Virtual Integrated Teaching and Learning Environment*) en la Universidad Baptista de Hong Kong durante el síndrome respiratorio que acaeció en abril de 2003 y que impidió el normal funcionamiento de las clases. Esta herramienta proporciona una interfaz muy simple y efectiva basada exclusivamente en tecnologías Flash, para lo que requiere la existencia de servidores Flash de comunicación intermedios. Uno de sus mayores problemas es que la funcionalidad que ofrece la pizarra es limitada.

En [Higuchi *et al.*, 2006] se desarrolla un sistema de instrucción multimedia interactivo denominado IMPRESSION. En este sistema, los materiales educativos multimedia se almacenan en servidores web públicos. Más tarde, el instructor puede recuperar los materiales desde su terminal para su manipulación. Todas las operaciones que realiza el instructor sobre los materiales (recuperación, anotaciones, resalte. . .) se reflejan en los terminales de los alumnos a través de un servidor que gestiona la clase. De igual manera, las operaciones que realizan los alumnos se sincronizan con el equipo del instructor automáticamente. Sin embargo, esta solución está únicamente pensada para la compartición de material multimedia a través de Internet, luego es necesaria una herramienta adicional de videoconferencia para transmitir vídeo entre los participantes de la clase.

Para tratar de proporcionar una mayor sensación de presencia y simular aún mejor el espacio de un aula tradicional, en [Massei *et al.*, 2006] se desarrolla un banco de pruebas llamado VIL (*Virtual Immersive Learning*). Este sistema resulta demasiado complejo y caro en comparación con los beneficios que reporta, una reconstrucción 3D de un aula tradicional que aparece en la interfaz de todos los participantes a la clase virtual.

Otra herramienta interesante es ConferenceXP [Pahud, 2008]. Se trata de una plataforma de investigación de código abierto promovida por Microsoft y orientada al trabajo colaborativo sobre Internet2. Para ello, utiliza técnicas multicast para hacer llegar los contenidos de la clase a sus participantes. Organiza la funcionalidad que ofrece al usuario en módulos que denomina *capabilities* y proporciona además una API a partir de la cual construir nueva funcionalidad. Entre la

funcionalidad que implementa se encuentra la videoconferencia, la compartición de el escritorio y aplicaciones, el chat y la pizarra virtual. Especialmente para esta última funcionalidad, ConferenceXP está centrado en actividades que precisen de la utilización de Tablet PC. Presenta como principal ventaja la modularidad de la aplicación y su ampliabilidad gracias a la posibilidad de desarrollar nuevas funciones para complementar las ya existentes. Su gran inconveniente son los recursos de ancho de banda que requiere debido a su orientación a Internet2, que la hacen poco apropiada para otro tipo de redes; y la interfaz de usuario, organizada en diferentes ventanas, que la convierte en una solución bastante compleja.

2.3.1. Conclusiones

Del análisis de las herramientas que se ha realizado en el apartado anterior, se puede comprobar cómo la mayor parte de soluciones propuestas desde el ámbito de la investigación no cubren satisfactoriamente toda la funcionalidad que se le presupone a una herramienta de e-learning síncrono. En lugar de ello, se centran en funcionalidades concretas, sin constituir en ningún caso una plataforma completa que posibilite el desarrollo satisfactorio de actividades formativas a distancia. Esto trae como consecuencia la necesidad de utilizar varias herramientas específicas para realizar acciones formativas de e-learning síncrono.

2.4. Caracterización de una herramienta de e-learning síncrono

A partir de las conclusiones extraídas del análisis de las diferentes herramientas, especialmente las comerciales, se deduce la funcionalidad que puede implementar una herramienta de e-learning síncrono. Sin embargo, no es necesario que una herramienta ofrezca todo este rango de funciones para que sobre ella puedan celebrarse actividades de e-learning síncrono de forma satisfactoria. Incluso, puede ser en ocasiones necesaria una deliberada reducción de funcionalidad dependiendo del entorno en el que la herramienta vaya a ser utilizada. En esta tesis, el principal ámbito en el que se consideran las herramientas de e-learning síncrono es en el de la formación dentro de grandes corporaciones. Por tanto, en primer lugar se describe todo el rango de funciones que una aplicación para la enseñanza síncrona puede ofrecer, para a continuación detallar cuál es el mínimo necesario, teniendo en cuenta las características especiales de la formación dentro de una gran empresa. Para finalizar, en el apartado 2.4.3 se describe una interfaz de usuario común, extraída a partir de las similitudes que presentan las herramientas de e-learning síncrono, a través de la cual ofrecer dichas funcionalidades.

El principal inconveniente al que debe hacer frente el departamento de formación de una gran empresa es el alto rechazo, en general, que presentan los empleados de mayor edad frente a las nuevas tecnologías. Este rechazo es más acusado cuanto mayor es la edad de los empleados. En consecuencia, ésta es una consideración que debe guiar el proceso de definición de requisitos de una herramienta de este tipo. Además, otro de los factores que tiene vital importancia es la red sobre la que opera la herramienta. Las grandes corporaciones suelen disponer de delegaciones en varios países, por lo que sus redes corporativas abarcan grandes distancias geográficas, al tiempo que dan servicio a gran cantidad de usuarios, cada uno con un determinado enlace a la red, no siempre con un ancho de banda mínimo aceptable. Esto implica que el efecto de la herramienta, especialmente en ancho de banda consumido, sobre la red debe ser mínimo puesto que, habitualmente, el ancho de banda disponible será escaso, especialmente en horario de oficina.

2.4.1. Funcionalidad de una herramienta de e-learning síncrono

A continuación se repasarán las funcionalidades típicas que un producto de e-learning síncrono incluye de cara al usuario, comentando su utilidad e importancia. En esta sección se recogen

las opciones que un usuario, alumno o instructor, de una aplicación de e-learning síncrono tiene a su disposición, una vez que el producto se encuentra instalado y configurado correctamente. No se entra a discutir sobre la base tecnológica de las distintas funcionalidades y tampoco se contemplan las tareas de administración del sistema, pues no son responsabilidad de los usuarios. Aunque el grueso de los servicios que proporcionan estas aplicaciones está orientado al desarrollo de la clase virtual en directo, no se debe descuidar el resto de funciones auxiliares disponibles antes y después de la celebración de la clase o conferencia.

2.4.1.1. Antes y después de la clase

Instructor y alumnos acceden a un portal web desde el que se gestiona la celebración de las clases y en el que deben estar registrados para poder acceder. Los instructores acceden a este portal para programar el horario de las clases, limitar el acceso a las mismas mediante una contraseña que los alumnos deberán introducir, designar específicamente los alumnos que pueden asistir a la clase seleccionándolos de una lista de alumnos registrados, dejar contenidos en el servidor de medios para su posterior empleo en las clases o decidir qué clases futuras grabar en el servidor. Otras posibilidades que tienen los instructores son las relacionadas con el envío de avisos a los alumnos: ya sea desde un cliente de correo o desde el portal web pueden enviar de forma automatizada correos electrónicos a los alumnos informando del horario de las clases o explicándoles algún aspecto sobre el desarrollo de éstas. En los sistemas en los que se puede acceder a la clase desde un navegador web, estos correos pueden incluir un enlace directo para que los alumnos entren en la clase virtual en el momento indicado. Del mismo modo estos enlaces directos pueden estar disponibles en el portal web una vez que el instructor haya programado la clase. Las posibilidades de los alumnos son lógicamente menores; normalmente pueden confirmar su asistencia a las clases, consultar horarios y asistir a clases en diferido, si tienen permiso para ello.

Tras finalizar la clase, el instructor puede decidir si dejar a disposición de los alumnos la grabación de la clase, en el caso de que hubiera sido grabada, y cualquier otro material generado durante la clase. Más allá de la enseñanza estrictamente síncrona, es bastante habitual complementar estos portales con elementos de aprendizaje asíncrono. Por ejemplo, el instructor suele proporcionar material complementario para los alumnos en un repositorio, que no tiene por qué residir en la misma ubicación que los contenidos grabados o generados en las clases, al cual se accede a través del mismo portal web. También puede abrirse un foro de debate tras cada clase en el que los alumnos discuten acerca de los contenidos de la misma, aclaran dudas y pueden recibir una respuesta colectiva del instructor. Además de la integración con clientes de correo y organizadores personales (Microsoft Outlook, Lotus Notes), estos portales web también suelen integrarse bien con servicios de directorio, como los basados en LDAP (Lightweight Directory Access Protocol). Esto permite, entre otras cosas, escoger a los instructores y alumnos de entre los registros de un servicio de directorio y, a partir de los datos obtenidos, poder gestionar sus perfiles de usuario de forma transparente; evitando, entre otras cosas, tener un nuevo registro independiente y exclusivo para los usuarios de la plataforma de e-learning.

2.4.1.2. Privilegios del instructor

Aunque no se trata de una funcionalidad propiamente dicha, la figura del instructor, profesor o moderador es muy importante. El instructor no deja de ser un participante más en la clase virtual, pero con más privilegios o posibilidades que el resto. Normalmente, además de tener a su disposición más opciones que el resto de usuarios, tiene la potestad de dar o quitar privilegios a los participantes, tales como permitir realizar anotaciones o intervenir en la clase, ya sea mediante el chat o por videoconferencia. El instructor puede llegar a otorgar a un participante los mismos privilegios que él posee, pudiendo haber varios instructores que combinen sus expo-

siciones desde diferentes lugares. En cualquier caso, el instructor inicial nunca podrá perder el control sobre la clase virtual. El instructor también puede controlar opciones específicas de los participantes, de manera individual o colectiva; por ejemplo, puede dar libertad a los alumnos acerca del modo de presentación de la aplicación en sus pantallas o puede obligarles a un modo específico de visualización, como pudiera ser presentar los contenidos a pantalla completa. En algunas aplicaciones, puede incluso controlar el trabajo del alumno observando lo que éste está visualizando en su pantalla, sin que el alumno lo sepa. En definitiva, el instructor disfruta de una serie de privilegios con el fin de que pueda desarrollar la clase de la manera que crea más conveniente. A veces estos privilegios constituyen formas particulares de hacer frente a problemas en la conexión de los participantes, como puede ser una situación en la que varios alumnos no puedan seguir correctamente la clase debido a una congestión en la red. En este caso el instructor puede desactivar el vídeo y la voz de estos participantes para que puedan seguir la presentación de forma más fluida. Esta acción también podría ser efectuada por los alumnos implicados.

2.4.1.3. Control de presencia

Para el usuario, el control de presencia no es más que la facultad de poder ver quiénes están asistiendo a la clase o conectados de alguna forma a la misma¹. El control de presencia es una parte fundamental, ya que para casi cualquier tarea se necesita tener conocimiento de las personas que asisten a la clase, por ejemplo, para saber de quién proviene un mensaje de texto o quién está escribiendo en la pantalla. Todas las aplicaciones muestran en pantalla una lista de los asistentes a la clase, con símbolos explicativos acerca del estado de cada uno (ausente, hablando, instructor, privilegios concedidos...).

2.4.1.4. Mensajería instantánea

Se trata de una funcionalidad simple pero indispensable, porque permite la comunicación entre los participantes independientemente del estado de la red (excepto cuando no exista conexión), ya que el ancho de banda requerido es insignificante. Las posibilidades de esta funcionalidad son las típicas de los programas de intercambio de mensajes: chat común a todos los participantes, chat privado con uno o varios, cambiar la fuente del texto, etc. Uno de sus usos más frecuentes es para efectuar preguntas al instructor sin la interrupción de la clase que provocaría una intervención oral. Suele estar disponible en primer plano cerca de la lista de asistentes (control de presencia).

2.4.1.5. Audioconferencia

Normalmente, si así se desea, se puede tener una audioconferencia, prescindiendo del vídeo. La voz requiere bastantes recursos de red, aunque menos que el vídeo y su importancia es mucho mayor que la de éste. Es una tarea prioritaria que la voz del instructor llegue en las mejores condiciones posibles a todos los alumnos. Cada vez que un participante habla está generando un flujo de voz que debe ser enviado al resto de usuarios. Es decir, está consumiendo ancho de banda saliente de su conexión y ancho de banda entrante de las conexiones del resto de participantes. Esto quiere decir que si todos hablan a la vez no habrá recursos suficientes y muy probablemente no se entenderá nada. Por esta razón suele estar limitado el número máximo de interlocutores simultáneos. Para poder controlarlo, un participante que desee intervenir debe pulsar normalmente un botón antes de hablar; si en ese momento ya se encuentran hablando el número máximo de personas, entonces deberá esperar. Cuando la transmisión de voz se basa en VoIP multipunto (multicast) no se está enviando realmente un flujo diferente por cada receptor, lo que permite que este límite en el número de personas sea bastante alto. En esta situación el

¹Algunos productos contemplan el caso de que existan usuarios que no asisten a la clase pero pueden chatear con los que sí están en ella. Se trataría de usuarios fantasmas.

límite de interlocutores lo suele imponer la lógica más que la técnica: al igual que en cualquier conversación llega un momento en el que no se entiende nada al haber muchas personas hablando a la vez. Las audioconferencias pueden ser entre todos los participantes o privadas entre dos o más personas.

2.4.1.6. Videoconferencia

La transmisión de vídeo es la funcionalidad que más ancho de banda requiere. Por eso el primer efecto de una conexión con poco ancho de banda se percibe en el vídeo. Cuando esto ocurre lo mejor es desactivarlo, ya que una transmisión de vídeo deficiente no aporta nada a la clase y además consume un ancho de banda que puede ser aprovechado por otras funcionalidades para que la clase virtual se siga desarrollando sin vídeo pero de forma fluida. Las ventanas en las que aparece el vídeo de los interlocutores suelen aparecer en un primer plano de la pantalla y se puede escoger su tamaño y calidad. La calidad escogida para el vídeo determina también el ancho de banda que éste requiere. Normalmente se pueden cambiar parámetros del mismo, tales como su resolución, los cuadros por segundo, el códec utilizado para comprimirlo, etc. Lo más habitual es que los alumnos visualicen sólo el vídeo del instructor y éste, el de varios alumnos. Cuantos más vídeos se visualicen de forma simultánea, peores resultados se obtendrán. Al igual que sucede con la transmisión de sonido, si la videoconferencia es multipunto (multicast) se consigue reducir la cantidad de datos transmitidos, lo que mejora notablemente el rendimiento. En condiciones favorables, un escenario de videoconferencia aceptable y bastante habitual suele ser visualizar simultáneamente cuatro vídeos de 320x240 píxeles a 15 cuadros por segundo cada uno.

2.4.1.7. Presentación de contenidos

Es el aspecto más importante en una aplicación de e-learning síncrono, dado que casi siempre una clase virtual se desarrollará en torno a unos contenidos o documentos que serán presentados. En todas las herramientas existe un área principal destinada a presentar los contenidos que el instructor seleccione. Estas herramientas soportan una amplia variedad de tipos de archivo, generados por diferentes aplicaciones: PowerPoint, Project, Word, Excel, Acrobat (PDF), Autocad, imágenes (JPG, GIF, PNG...), etc. Los tipos de archivo más compatibles son los de PowerPoint, soportados por todas las aplicaciones. Los archivos PDF son también altamente compatibles y por eso se suele recomendar convertir cualquier tipo de documento no soportado a este formato. Los contenidos pueden estar indistintamente almacenados en el equipo del instructor o en un servidor de medios, y pueden ser seleccionados sobre la marcha o haber sido incluidos previamente al comienzo de la clase virtual. Con el consentimiento del instructor, cualquier participante puede presentar documentos. Cuando los contenidos se añaden sobre la marcha desde un equipo, se debe esperar unos segundos a que éstos sean adaptados adecuadamente para poder ser servidos a todos los participantes. Habitualmente, también durante ese intervalo de tiempo los contenidos son subidos a un servidor de medios.

Una vez que se está presentando un contenido, la ventana en la que éste es visualizado puede ser redimensionada. Éste es un factor bastante importante, ya que muchas veces será interesante maximizar la presentación a pantalla completa e ir siguiendo las explicaciones orales del instructor. Normalmente, cada alumno puede controlar de forma individual la visualización de la presentación en su equipo; esto es, que podrá aumentar o disminuir el tamaño del documento en pantalla o navegar por el mismo hacia delante o hacia atrás, siempre que el instructor lo permita. Otra característica disponible, sobre todo en las presentaciones con PowerPoint, es una ventana resumen de las diapositivas de las que consta la presentación, que permite ir rápidamente a un punto concreto de la misma. Este resumen suele ser en forma de esquema textual, aunque algunas aplicaciones también lo muestran mediante miniaturas de las diapositivas. Hay que destacar que ni los instructores ni los alumnos necesitan tener instalados en sus equipos los programas con los

que se han generado los documentos, siempre que los formatos de éstos sean soportados. Cuando se quiere mostrar un contenido cuyo formato no es soportado por la aplicación de e-learning, entonces sí se necesita disponer de la aplicación con el que fue generado y la solución consiste en compartir esta aplicación con el resto de participantes; lo cual constituye otra funcionalidad aparte.

Una característica íntimamente ligada con la presentación de contenidos es el puntero virtual. Así se denomina al cursor del dispositivo apuntador del instructor, cuando es visible por el resto de participantes. Su utilidad es análoga a la de un puntero en una presentación real: destacar o indicar elementos concretos de la presentación. Su uso es muy sencillo, basta pulsar un botón para que el resto de asistentes vea el puntero y volver a pulsarlo para que deje de estar visible.

2.4.1.8. Pizarra virtual

El nombre de esta característica es bastante descriptivo. No es más que un sencillo editor gráfico, como el incluido con Windows, que permite dibujar formas geométricas, insertar imágenes, escribir textos...; sobre un lienzo virtual, que es compartido por los asistentes a la clase.

2.4.1.9. Anotación sobre las presentaciones

Se trata de una posibilidad de la que habitualmente disponen todas las aplicaciones. El que cualquier participante pueda escribir o dibujar sobre un documento que está siendo visualizado permite destacar o corregir sobre la marcha algún elemento de la presentación. El funcionamiento es idéntico al de la pizarra virtual ya que básicamente se está haciendo uso de esta herramienta sobre un fondo constituido por la transparencia o página del documento presentado.

2.4.1.10. Compartición de escritorio y de aplicaciones

Es una funcionalidad presente en casi todos los productos. Posibilita que cualquier participante comparta una aplicación determinada o su escritorio completo, de forma que el resto de asistentes vean lo que él está realizando. El contenido compartido es mostrado en el área de presentación de los participantes, como si de un documento se tratase. Es la mejor solución a la hora de compartir contenidos no soportados específicamente por la aplicación de e-learning. El instructor puede usar esta función para explicar los pasos a seguir para realizar una operación en una aplicación determinada o para indicar cómo configurar un dispositivo. También puede existir la posibilidad de compartir una parte de pantalla de dimensión ajustable, por ejemplo una octava parte de la superficie de la pantalla desde la esquina inferior derecha; en este caso cualquier ventana, aplicación u objeto que se encuentre en esa zona será visible para todos los participantes, mientras que lo que esté fuera de la misma no lo será. Por último, otra variante destacable dentro de la compartición del escritorio es la posibilidad de capturar el contenido de la pantalla en una imagen mediante una tecla rápida y presentarla automáticamente a los demás participantes.

Esta funcionalidad se implementa mediante técnicas de streaming de vídeo. A pesar de que, normalmente, el tamaño de la ventana o del escritorio suele ser significativo, no se necesita una tasa de refresco tan alta como para el vídeo en que se muestra el rostro de un participante. Por tanto, una tasa de cuadros baja, del orden de unos pocos cuadros por segundo, es suficiente para transmitir el contenido de una ventana o del escritorio, de ahí que el ancho de banda que consume esta funcionalidad no es más elevado que el que consume la videoconferencia.

2.4.1.11. Control remoto

Permite que un usuario controle remotamente una aplicación o el escritorio de otro participante. Es el paso siguiente a la compartición del escritorio y aplicaciones. Además de ver, se puede actuar. Dado que es una operación que conlleva riesgos para el usuario que cede el control de su equipo,

en todo momento dispone de una tecla rápida o de un botón para cancelar el permiso de control remoto.

2.4.1.12. Navegación sincronizada por Internet

Muchas aplicaciones desempeñan esta función compartiendo la ventana del navegador web. Ésta es una de las opciones posibles. Otra, mucho más interesante, es que el instructor determine qué páginas web se van a visitar y que entonces el navegador integrado en la aplicación de cada usuario redirige a esas páginas. De esta forma, salvo en los accesos a páginas cuya carga sea muy lenta, se necesita mucho menos ancho de banda, ya que cada participante está accediendo individualmente a las páginas web en lugar de tener que recibir constantemente el contenido mostrado por el navegador del instructor cuando está siendo compartido. De una u otra forma, se trata de una función muy empleada por los instructores.

2.4.1.13. Transferencia de ficheros

Permite que los participantes intercambien ficheros entre sí. Basta con seleccionar uno o varios usuarios de la lista de asistentes y escoger los ficheros que se les desea enviar.

2.4.1.14. Anotaciones personales

Es una funcionalidad poco frecuente en los productos de e-learning síncrono. Permite a los usuarios escribir texto en una zona reservada para ello y que no es visible por el resto de participantes. El espacio reservado para escribir estas notas suele consistir en una pequeña ventana parecida a la del chat. Los usuarios pueden guardar estas notas en cualquier momento.

2.4.1.15. Almacenamiento de contenidos

Normalmente los instructores pueden grabar una clase en directo en el servidor para que pueda ser visualizada posteriormente. En la grabación se recoge la pantalla del instructor en todo momento. Otra posibilidad es que los alumnos puedan grabar su pantalla a lo largo de toda la clase en sus equipos. La activación de esta característica requiere únicamente la pulsación de un botón para comenzar la grabación. También es muy interesante que exista la opción de guardar contenidos puntuales presentados durante la clase, como pueden ser páginas de documentos, diagramas de la pizarra virtual o diapositivas modificadas con anotaciones realizadas durante la clase.

2.4.1.16. División en grupos

Esta funcionalidad permite al instructor separar a los alumnos en grupos de trabajo independientes. Se trata de una opción que puede requerir bastante ancho de banda, dependiendo de cómo se transporte los datos, pues cada grupo de trabajo pasa a ser, a casi todos los efectos, una conferencia distinta, excepto porque el instructor tiene acceso a todos los grupos. Dentro de cada grupo los participantes disponen de las mismas funcionalidades que en la clase virtual completa. Sólo el instructor puede volver a reunir los grupos.

2.4.1.17. Preguntas y sondeos

Abarca todas las posibles acciones que el instructor pueda efectuar para consultar la opinión o los conocimientos de los alumnos. Dentro del apartado de preguntas académicas se contemplan diferentes estructuras: verdadero o falso, múltiples opciones, rellenar con texto, etc. El instructor determina el tiempo máximo del que disponen los alumnos para responder cada pregunta. La zona desde la que el instructor confecciona las preguntas y donde los alumnos responden a éstas

suele situarse encima de la ventana de chat. Cuando termina el tiempo o todos los alumnos han respondido el instructor puede consultar y guardar las respuestas de cada alumno. Del mismo modo puede visualizar estadísticas globales acerca de las respuestas de los alumnos para comprobar sus conocimientos de forma global o su grado de seguimiento de la clase. El instructor o moderador también puede pedir la opinión de los asistentes efectuando un sondeo anónimo; es decir, se conocerá el resultado pero no el origen de cada respuesta. En este caso el resultado podrá ser visto por todos los participantes.

2.4.1.18. Realimentación instantánea de los alumnos hacia el instructor

Dentro de este apartado se engloban los diferentes mecanismos de comunicación rápida, casi instantánea, de los que disponen los alumnos. Más correctamente habría que referirse a comunicación intuitiva. Estos mecanismos suelen consistir en botones de acceso rápido que envían directamente un mensaje al instructor, como puede ser levantar la mano para efectuar una pregunta, dar una respuesta afirmativa o negativa, mostrar estados de ánimo mediante emoticonos, etc.

2.4.2. Requisitos mínimos de una herramienta de e-learning síncrono

En el apartado anterior se ha visto todo el rango de funcionalidades que una herramienta de e-learning síncrono puede ofrecer. Sin embargo, no siempre es necesaria toda esa funcionalidad, alguna de las cuales sólo se utiliza en casos concretos. Por ejemplo, la compartición de aplicaciones no es necesaria si la clase no trata sobre el manejo de una determinada aplicación. Lo mismo ocurre para otras funciones, todo depende del ámbito en que vaya a ser utilizada la herramienta; si se pretende abarcar un gran número de situaciones está claro que cuanto más funcionalidad ofrezca, más posibilidades se ofrece al instructor y a los participantes para interactuar. No obstante, si el ámbito donde va a utilizarse la herramienta se determina de antemano es posible eliminar aquella funcionalidad menos adecuada.

En el capítulo 4 se aborda el diseño de un prototipo de herramienta de e-learning síncrono para la formación de los recursos humanos dentro de una gran corporación. Estas corporaciones y los usuarios que utilizan estas herramientas dentro de la misma tienen unas características específicas, que determinan las funciones necesarias que la herramienta debe ofrecer. Específicamente, es necesario considerar los siguientes condicionantes:

- **Simplicidad.** La gran mayoría de usuarios a los que está destinada la herramienta final tiene escasos o nulos conocimientos de informática, incluso, presentan un gran rechazo frente a las nuevas tecnologías. Por esta razón, es imprescindible el diseño de la misma y de su interfaz con especial meticulosidad; ésta debe ser muy simple de utilizar por parte de sus usuarios y al mismo tiempo no generar un rechazo inicial al usuario novel. En consecuencia, se acepta una deliberada reducción de interactividad dentro de la plataforma en beneficio de una mayor simplicidad de manejo.

Por otra parte, las herramientas comerciales analizadas en el apartado 2.2 constituyen soluciones muy complejas, habitualmente basadas en plataformas compuestas por conjuntos de servidores con una infraestructura específica. La simplicidad en el diseño de la herramienta debe cubrir no sólo aspectos como la usabilidad y accesibilidad de la interfaz, sino la arquitectura global de la plataforma, de manera que el mantenimiento sea el mínimo necesario para evitar destinar personal especializado para su administración. Por tanto, se debe prescindir del uso de servidores que centralicen las comunicaciones o actúen de controladores de la clase virtual; el prototipo a desarrollar debe ser una herramienta *peer-to-peer*.

- **Herramienta integrada.** Al hilo del punto anterior, la herramienta que utilizan los usuarios para seguir las sesiones de e-learning, a diferencia de gran parte de las herramientas de

investigación disponibles, debe constituir una sola unidad, y no un conglomerado de herramientas con funcionalidades específicas (videoconferencia, mensajería instantánea, pizarra compartida...). De este modo, resulta en una herramienta integrada que será más fácil de utilizar por personas con escasos conocimientos informáticos; se evitará la existencia de varias ventanas que puedan confundir al usuario. Además, al implementar toda la funcionalidad dentro de la misma aplicación, es posible optimizar aquellas partes que se consideran críticas frente a otras que lo son menos.

- Bajo ancho de banda. En ocasiones, la capacidad del enlace de red de los usuarios de la plataforma es un recurso valioso. Este hecho es más notorio si cabe dentro de una red corporativa, especialmente en horario de oficina. Es habitual que las grandes corporaciones empresariales dispongan de dos redes de datos independientes: la red de industrial, destinada a interconectar equipos cuya función es controlar y monitorizar el proceso industrial; y la red de datos, destinada al transporte de datos de carácter ofimático y que habilita el acceso a Internet. Por esta razón, los datos que fluyen a través de la red procedentes del prototipo a desarrollar deben compartir el ancho de banda con otros tipos de tráfico. Además, existe la posibilidad de que se conecten a una sesión de e-learning alumnos desde fuera de la red corporativa a través de Internet, por lo que el rango de anchos de banda sobre los que debe operar la plataforma se diversifica en gran medida. Por tanto, el ancho de banda consumido por el tráfico de la plataforma debe ser el mínimo posible que permita ampliar el abanico de anchos de banda sobre los que la plataforma puede operar.
- Evitar sobrecarga cognitiva del instructor. Aunque en menor medida, a veces, el instructor que debe impartir una clase a un grupo de alumnos utilizando la herramienta carece de una habilidad aceptable en el manejo de aplicaciones informáticas. Si a esto se le une el, frecuentemente, elevado número de alumnos presentes en una sesión de e-learning, la sobrecarga que puede recaer en el instructor según evoluciona la clase puede llegar a ser elevada. Algunos de los factores que influyen en la carga del instructor son los siguientes:
 - Número de alumnos: cuanto mayor es el número, más trabajo tendrá que realizar el instructor para atender correctamente las dudas que pudieran plantear.
 - Perfil de los alumnos: si los alumnos sobre los que se lleva a cabo la acción formativa tienen un completo desconocimiento de la materia que se trata, el número de cuestiones que éstos plantearán será mayor, con las consiguientes interrupciones de la clase para que el instructor ofrezca las oportunas explicaciones. Además, el carácter y los prejuicios que los alumnos mantengan hacia la herramienta y el proceso formativo también influyen puesto que les hará participar en diferente grado durante el desarrollo de la clase.
 - Materia a tratar: dependiendo del tema sobre el que versa la acción formativa a los alumnos les resultará más o menos difícil adquirir los conocimientos planteados, lo que repercutirá en un mayor esfuerzo por parte del instructor para consolidar dichos conocimientos.
 - Medio utilizado para plantear las dudas: la utilización de uno u otro medio (audio, mensajes de texto, anotaciones...) para que un alumno plantee una duda al instructor le supondrá diferente esfuerzo para su comprensión. Por ejemplo, una cuestión planteada por un alumno a través del canal de audio supone la interrupción de la clase para atender dicha cuestión. Además, si varios alumnos plantean sus cuestiones al mismo tiempo es realmente difícil que el instructor pueda atenderlas a todas. Sin embargo, una pregunta realizada a través de un mensaje de texto permite al instructor abordarla en el instante que considere oportuno sin necesidad de interrumpir la clase. Además, la ocurrencia de varias cuestiones de forma simultánea no supone problema alguno

porque éstas quedan registradas en el panel de mensajes identificando al alumno que la planteó. En la figura 2.7 aparecen representadas ambas situaciones.

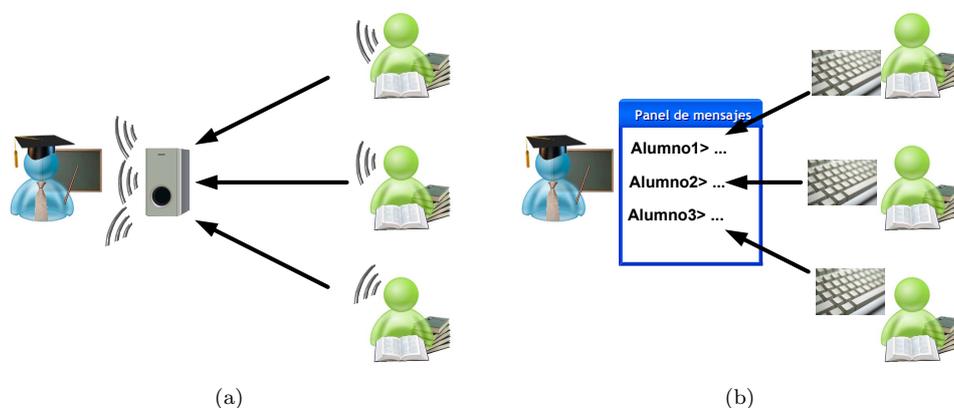


Figura 2.7: Planteamiento de preguntas de los alumnos al instructor utilizando (a) el canal de audio y (b) mensajería instantánea

Por todos estos condicionantes antes mencionados, se busca una herramienta esencialmente expositiva que permita al instructor transmitir una serie de conocimientos interrumpiendo la clase únicamente cuando éste estime oportuno. En base a esto, se hace necesario un proceso de selección para determinar la idoneidad de cada funcionalidad para este ámbito concreto. En la tabla 2.2 se indican las funciones que son necesarias implementar.

La utilización de un canal de audio a través del cual el instructor transmitirá la mayor parte de su discurso formativo se hace del todo imprescindible. Además, la calidad del audio debe ser la más alta posible, pues resulta vital que los alumnos reciban con nitidez las explicaciones del instructor. Es por esto que el audio debe tener prioridad sobre el resto de medios dentro de la herramienta.

Dependiendo del uso que se le dé al canal de vídeo puede constituir una funcionalidad más o menos interesante. Si el objetivo del vídeo es transmitir únicamente el busto parlante del instructor, la información que proporciona a los alumnos es mínima; únicamente sería útil para consolidar la sensación de presencia del instructor y evitar la sensación de desamparo en los alumnos [Chassie, 2002] y [Weller, 2007]. Por contra, si puede resultar más útil la videoconferencia si se utiliza para que el instructor muestre, a través de la cámara, algún otro tipo de información, por ejemplo, la secuencia de pasos de un proceso, una pieza industrial o la portada de un libro. En ese caso, será necesario utilizar una alta resolución espacial y temporal, lo cual puede ir en contra de las restricciones de ancho de banda a las que antes se hacía mención.

Tabla 2.2: Requisitos mínimos para una herramienta de e-training

Imprescindibles	Opcionales	Innecesarios
Audioconferencia	Videoconferencia	Escritorio compartido
Control de presencia	Puntero virtual	Navegación síncrona
Chat	Aplicaciones compartidas	Transferencia de ficheros
Presentación de diapositivas		División en grupos
Anotaciones		Emoticonos

Por otro lado, el control de presencia, además de reforzar la sensación de presencia de los participantes en la sesión de e-learning, es útil para que el instructor lleve registro del seguimiento de la clase por parte de los alumnos. Teniendo en cuenta que el ancho de banda que esta funcionalidad consume y los grandes beneficios que reporta es fácil comprender por qué todas las herramientas de e-learning síncrono la ofrecen.

Con la mensajería instantánea ocurre lo mismo que con el control de presencia: reporta más beneficios que inconvenientes. No en vano, en gran parte de las ocasiones ambas funcionalidades aparecen combinadas. La mensajería instantánea, o chat, permite la comunicación entre todos los participantes de la sesión de e-learning a través de mensajes de texto. Esto es útil para fomentar las relaciones entre los miembros, al tiempo que permite a los alumnos plantear dudas al instructor o a otros alumnos.

La presentación de diapositivas constituye una de las funcionalidades más importantes, pues aporta la mayor parte de los contenidos que se utilizan como soporte durante el desarrollo de una actividad formativa sobre una herramienta de e-learning síncrono.

Mediante el puntero virtual el instructor puede apuntar elementos dentro de una diapositiva para facilitar así la comprensión de su discurso. Únicamente es útil cuando la densidad de las diapositivas es lo suficientemente alta como para tener que señalar una parte concreta de la misma.

Las anotaciones se pueden utilizar para realizar concreciones sobre las diapositivas de la presentación, o también sobre una diapositiva en blanco a modo de pizarra compartida. Pueden además resultar útiles para que los alumnos, a instancias del profesor, resuelvan un ejercicio planteado.

El escritorio y las aplicaciones compartidas pueden llegar a ser una funcionalidad contraproducente por el ancho de banda que consumen en la red. Aunque pueden llegar a ser útiles para hacer uso de material no manejable directamente por la herramienta, esto limitaría el uso de la herramienta en alguno de los enlaces de red por el bajo ancho de banda disponible. Sería necesario descartar esta funcionalidad o la videoconferencia, que son dos de las que más recursos consumen. De cualquier forma, se puede prescindir tanto del escritorio como de las aplicaciones compartidas porque los beneficios que reportan son escasos, especialmente el primero. La compartición de aplicación podría resultar más útil si en una clase virtual se trata de mostrar el funcionamiento de una aplicación.

La transferencia de ficheros, al igual que la posibilidad de almacenar los contenidos de la clase, es del todo prescindible por la existencia de un portal corporativo desde el que los alumnos pueden descargar los ficheros necesarios, y el instructor almacenarlos.

Otra de las funcionalidades prescindible son los emoticonos de que disponen los alumnos para mostrar su estado de ánimo. Para usuarios con bajos conocimientos informáticos, o incluso con grandes reparos hacia las nuevas tecnologías, la comprensión del significado de cada uno de ellos y su utilidad es realmente baja.

2.4.3. Interfaz de usuario de una herramienta de e-learning síncrono

En esta sección se analizará el modelo de interfaz de usuario en el que se basan todas las aplicaciones de e-learning síncrono analizadas. En este tipo de aplicaciones existe una estrecha relación entre las tareas que se pueden desempeñar y las diferentes partes del interfaz; por lo que, habiendo descrito ya profusamente las funcionalidades, aquí se efectuará una descripción gráfica muy ilustrativa. Es muy significativo, y por eso se menciona nuevamente, que todas las soluciones software estudiadas poseen el mismo interfaz de usuario indistintamente para profesores y alumnos, tanto en la aplicación empleada durante el desarrollo de la clase como en los portales web que dan soporte auxiliar fuera del marco de la clase virtual, y que sirven para dar acceso a la misma.

2.4.3.1. Pantalla principal

En cualquier aplicación orientada al e-learning síncrono, la pantalla principal se encuentra dividida en varias zonas claramente diferenciadas. Siempre habrá una zona más grande que será el área de presentación. Excepto cuando se esté visualizando algún contenido a pantalla completa, también existirá una zona lateral en la que se encontrarán las funcionalidades más comúnmente utilizadas. Además de estas dos zonas, que son las más relevantes, se distinguen otras dos: en la parte superior de la pantalla, un menú y una barra de herramientas; y en la inferior, una barra de estado. En la figura 2.8 se muestran imágenes de cuatro aplicaciones: WebEx Training Center, Saba Centra Live, Elluminate Live! y WebConference.

En ellas se pueden diferenciar claramente las cuatro zonas comentadas (en la captura de Saba Centra Live la barra de estado está oculta). Éstos no son los únicos productos que tienen la pantalla principal estructurada de esta forma; todo lo contrario, la excepción son los que no la tienen así. Pero basta con estos ejemplos arbitrarios para corroborar la similitud, casi igualdad, existente entre las interfaces de usuario de las distintas aplicaciones de e-learning síncrono existentes. Junto a las capturas mencionadas se enumeran las cuatro zonas, así como las acciones más importantes que se llevan a cabo en cada una de ellas.

2.4.3.2. Menú principal, barra de herramientas y barra de estado

Mediante el menú situado en la parte superior se tiene acceso a opciones generales de la aplicación, como las relacionadas con la ayuda y la configuración del programa, y a la mayor parte de funcionalidades del programa, independientemente de que algunas se encuentren también disponibles de forma más directa en la zona lateral. Particularmente, en esta zona se suelen encontrar los accesos para insertar contenidos, compartir escritorio y aplicaciones, solicitar control remoto del escritorio de un alumno por parte del instructor, transferir ficheros, dividir a los alumnos en grupos y grabar la clase. Estas opciones suelen estar incluidas en el menú y también en la barra de herramientas, para agilizar su uso.

Algunos productos incluyen en la barra de herramientas los botones empleados para levantar la mano, contestar rápidamente o mostrar estados de ánimo. En otras aplicaciones estos botones están situados en la zona lateral, cerca de la lista de participantes. Un ejemplo de menú principal con su barra de herramientas puede verse en la figura 2.9.

La función de la barra de estado suele ser meramente informativa, mostrando información acerca de la transmisión de audio y del estado de la conexión. Ocasionalmente puede encontrarse en la barra de estado el botón que es preciso pulsar antes de comenzar a hablar o alguna otra opción de frecuente acceso. En la figura 2.10 se muestra la barra de estado de WebEx Training Center.

2.4.3.3. Área de presentación

El área de presentación es la zona hacia la que los participantes dirigen su atención la mayor parte del tiempo, ya que la información presentada a los participantes es mostrada aquí. Dentro de esta área se podrá observar, casi siempre en el borde superior o inferior, una o varias barras de herramientas que permitirán efectuar ciertas acciones relacionadas con el contenido mostrado: si se trata de una presentación o de un documento, habrá controles para avanzar y retroceder dentro del mismo, aumentar o disminuir su tamaño en pantalla y también para anotar sobre él; si se está navegando por Internet de forma sincronizada, aparecerá por ejemplo la barra de direcciones; y si se está compartiendo la pizarra virtual, figurarán las herramientas de dibujo correspondientes. En la barra de herramientas también suele ser habitual encontrar botones que cambian el modo de visualización de la aplicación.

En la parte inferior del área de presentación de WebEx Training Center, mostrada como ejemplo en la figura 2.11, se distinguen los controles para aumentar o disminuir el tamaño del contenido

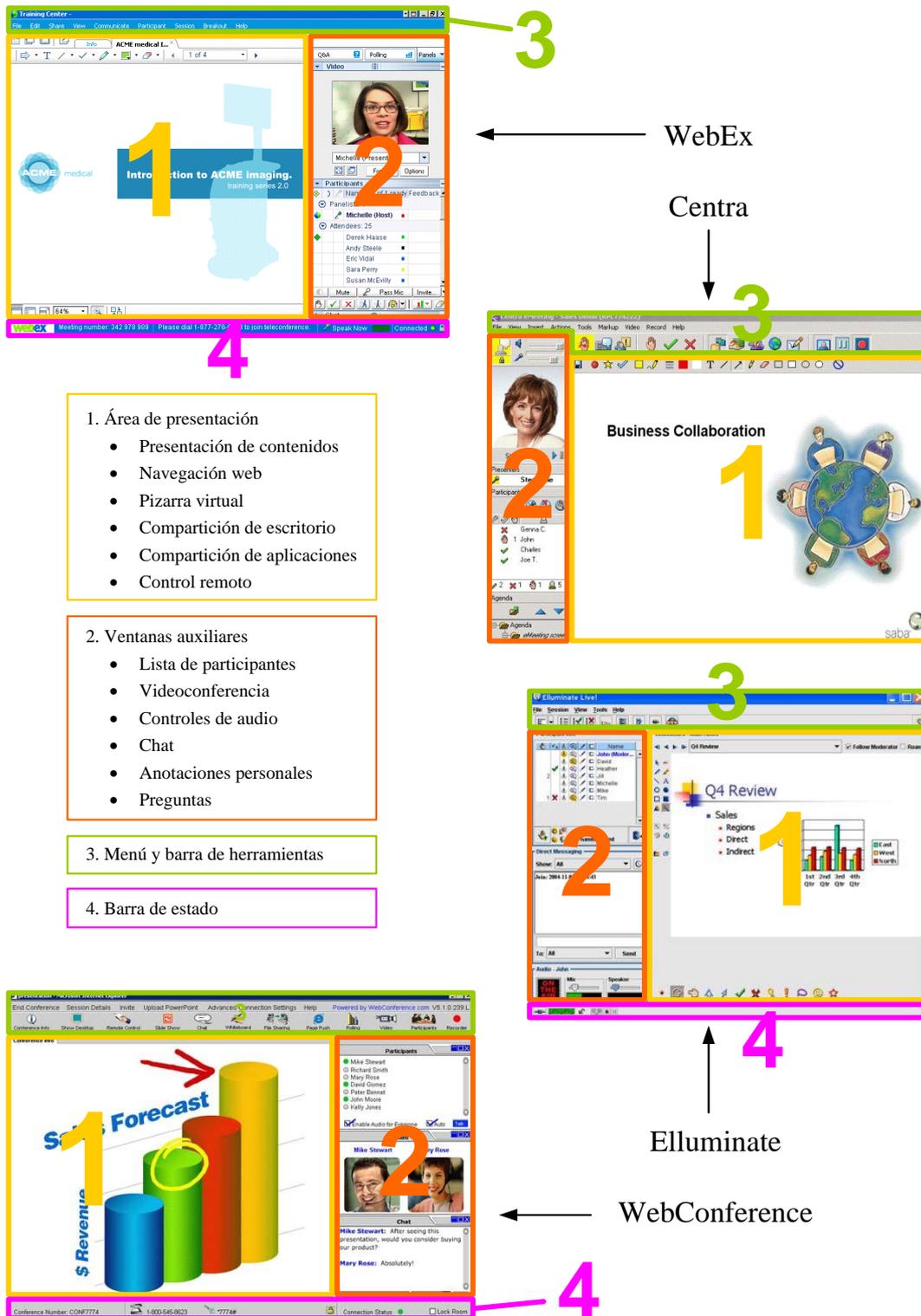


Figura 2.8: Ventana principal de diferentes aplicaciones de e-learning síncrono



Figura 2.9: Menu principal y barra de herramientas de WebConference

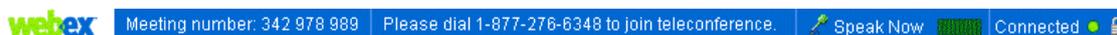


Figura 2.10: Barra de estado de WebEx Training Center

del área, un botón con el que el profesor fuerza que lo que él está viendo aparezca en las pantallas de los alumnos, y tres botones para mostrar a pantalla completa toda la aplicación, sólo el área de presentación o el área de presentación y un esquema del documento actual. En la parte superior de la imagen se distingue la barra de herramientas que contiene todas las opciones de navegación y edición. En este caso, sobre esta barra de herramientas se ven unas pestañas que sirven para acceder a los diferentes documentos, aplicaciones u otros objetos que estén siendo compartidos o presentados en la clase virtual; otras aplicaciones muestran esta información en la zona inferior del área de presentación. Los cuatro botones de la parte superior izquierda de la imagen son accesos rápidos para compartir documentos, aplicaciones, escritorio y pizarra virtual. Como las utilidades que arrancan estos botones se desarrollan en el área de presentación, no se puede considerar que su colocación sea inadecuada, aunque como ya se ha comentado, una ubicación más común sería en el menú principal o en una barra de herramientas fuera del área de presentación. A veces también se puede visualizar el esquema de una presentación o de un documento en esta área, en lugar de en una ventana auxiliar.

2.4.3.4. Ventanas auxiliares

Se suelen encontrar agrupadas en un lateral de la aplicación. En esta zona se concentran muchas de las funcionalidades más importantes, como la lista de usuarios, el chat y los vídeos de los participantes. Desde el menú principal o mediante pestañas presentes en esta misma zona, se puede elegir las ventanas que se quiere tener a la vista. Cuantas más ventanas se muestren a la vez, menos espacio tendrá cada una de ellas. En algunas aplicaciones estas ventanas están fijadas en un lateral, pudiendo modificar la altura de cada una y el ancho de la columna en la que se encuentran. En otras, además se podrán sacar del lateral y situarlas en cualquier lugar como ventanas o paneles flotantes. Lo ideal es que estas ventanas se comporten como lo que se podría denominar un sistema de ventanas acoplables y redimensionables; esto es, que al acercarlas a otra ventana auxiliar o al borde de la pantalla o de la aplicación, se posicionen automáticamente de la mejor forma posible. Del mismo modo es deseable tener pocas ventanas auxiliares, dos o tres, pero pudiendo condensar en cada una de ellas varias funcionalidades y poder rotar entre sus opciones mediante pestañas exclusivas de cada ventana.

En la figura 2.12 se muestran varios conjuntos de ventanas correspondientes a diferentes productos. Se puede constatar cómo todos los productos dedican la zona lateral a tareas similares. Se incluyen dos capturas de e/pop Web Conferencing, como muestra de disposiciones distintas en una misma aplicación. WebEx Training Center es extraordinariamente flexible a la hora de colocar las ventanas. Además de permitir redimensionar las ventanas y tenerlas flotantes, como se ejemplifica en la figura 2.13, muestra en todo momento los títulos de todas las ventanas existentes, de modo que unas puedan estar desplegadas y otras minimizadas u ocultas, pero estando siempre a la vista del usuario para que tenga en cuenta todas las opciones a las que puede acceder.

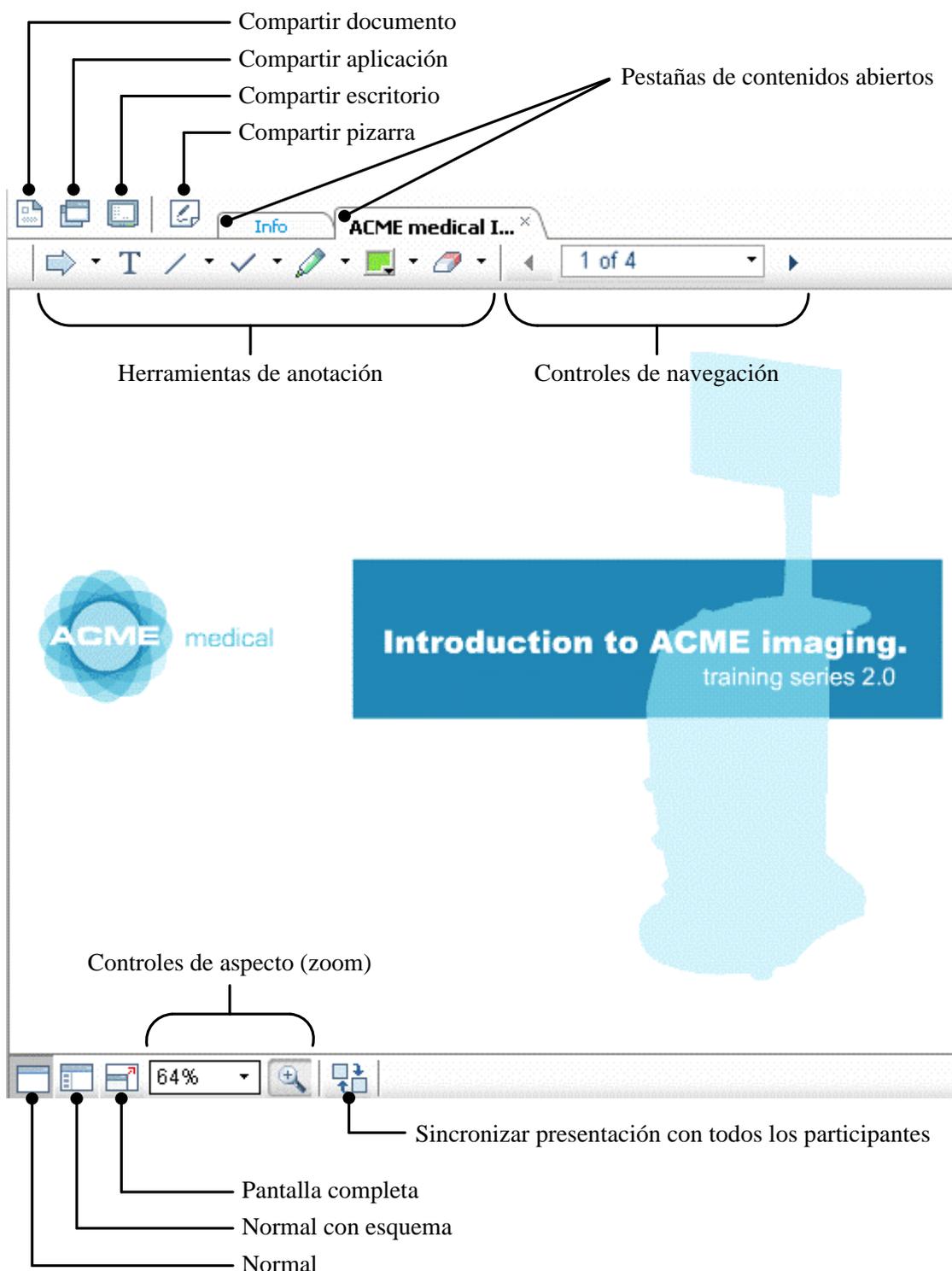


Figura 2.11: Área de presentación de contenidos de WebEx Training Center

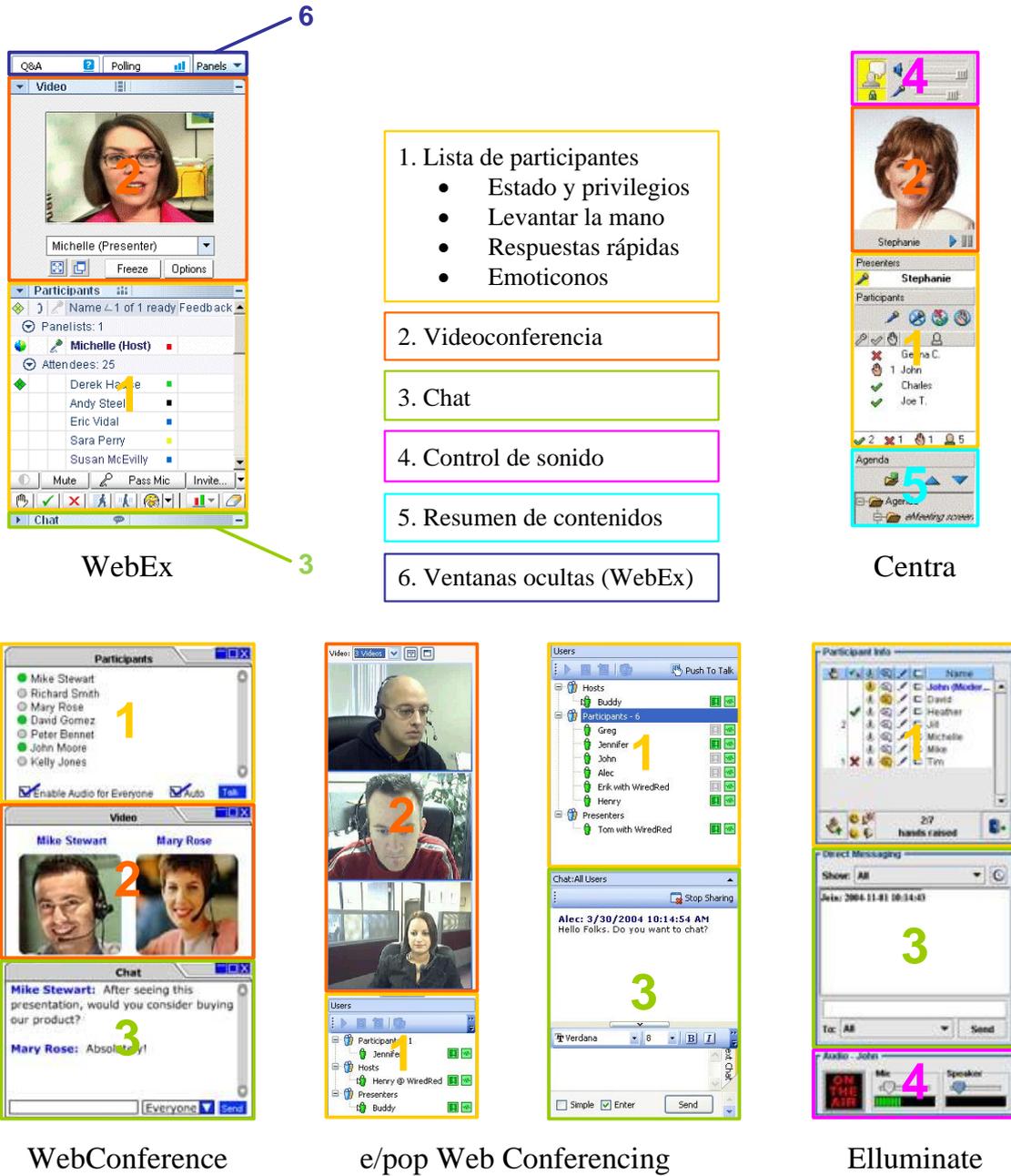


Figura 2.12: Ventanas auxiliares de varias aplicaciones de e-learning síncrono

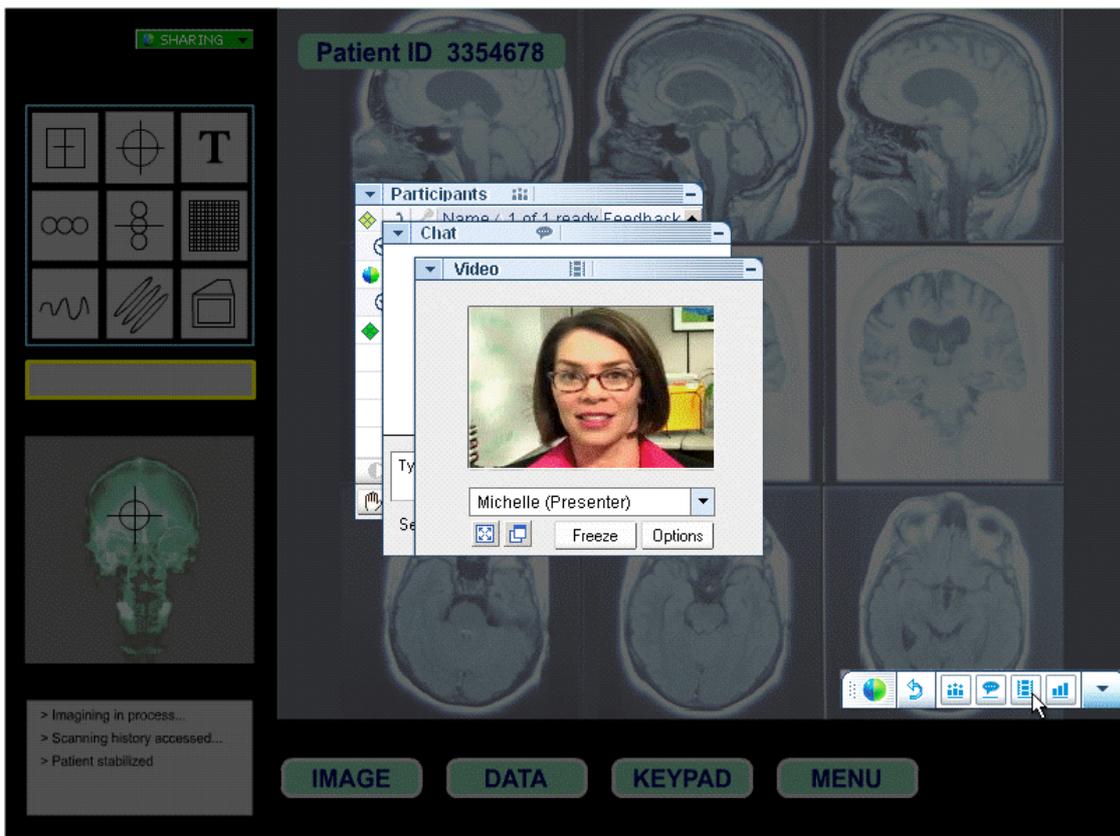


Figura 2.13: Ventanas auxiliares en WebEx Training Center

Capítulo 3

Análisis de tecnologías para el e-learning síncrono

En este capítulo se analizan las diferentes alternativas tecnológicas que posibilitan la implementación de un SMI aplicable al e-learning síncrono. Inicialmente, se abordan las cuestiones relativas al transporte de datos sobre las redes IP, prestando especial interés en las técnicas de multidifusión que habilita IP Multicast, y cómo se utiliza el protocolo RTP para el transporte de medios continuos sobre redes de paquetes. Posteriormente, se valoran los protocolos de señalización que nos permiten establecer sesiones multimedia entre varios participantes, así como las tecnologías que nos permiten procesar la información que se maneja dentro de una herramienta de este tipo. Ya por último, se introducen los protocolos que nos permiten gestionar el turno de intervención de cada uno de los participantes en las sesiones.

3.1. Introducción

Las tecnologías en las que se basan las soluciones de e-learning síncrono existentes se pueden agrupar en diferentes bloques según su ámbito de aplicación. La gestión de las sesiones de los usuarios conectados y la transmisión de flujos multimedia, independientemente de la naturaleza de éstos, son, sin duda, los apartados para los que existe mayor variedad de opciones tecnológicas y, al mismo tiempo, los que más importancia tienen. Dentro del campo de los flujos multimedia, la multitud de formatos de codificación y/o compresión de audio y vídeo empleados conforman otro importante grupo de estudio, aunque de menor trascendencia tecnológica que las técnicas de transporte de los flujos.

Otro aspecto a considerar es la transmisión y coordinación de flujos de datos distintos al audio y el vídeo; es decir, toda la información que se intercambia cuando se efectúa una presentación de contenidos, se comparten aplicaciones, se utiliza el chat o la pizarra compartida, se transfieren ficheros, etc.; dentro de una herramienta de e-learning síncrono. La mayor parte de las tecnologías consideradas se refieren a la transmisión de información multimedia por Internet y se pueden circunscribir al amplísimo conjunto de recomendaciones y estándares tecnológicos en los que se basan las comunicaciones a través de Internet. En este sentido, las dos organizaciones de estandarización más relevantes en este campo son la IETF (*Internet Engineering Task Force*) y la ITU-T (Sector de Normalización de las Telecomunicaciones de la Unión Internacional de Telecomunicaciones). Ambas contribuyen de forma constante al desarrollo tecnológico de Internet elaborando recomendaciones, que si no alcanzan el grado de estándar, muchas veces lo son de facto. Tradicionalmente, contribuyen de forma conjunta a la difusión y asentamiento de sus desarrollos, pero en ocasiones, las presiones parciales de la industria contribuyen a retrasar los avances, llegando a enfrentar recomendaciones de ambos organismos entre sí.

En el resto del capítulo se analizan, inicialmente, las diferentes tecnologías disponibles para el transporte de datos a través de las redes de paquetes y la influencia de éstas en las transmisiones. Este análisis comprende desde las técnicas de multidifusión, que permiten distribuir los datos de la forma más eficiente posible, hasta los protocolos de señalización, que gestionan el establecimiento

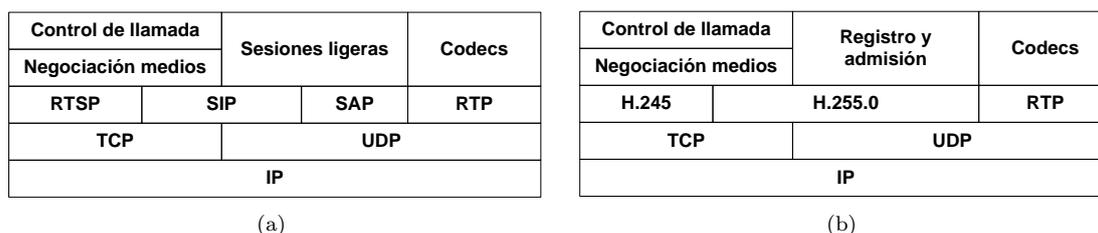


Figura 3.1: Pilas de protocolos de la (a) IETF y de la (b) ITU-T

y la negociación de las sesiones multimedia. A continuación, se describen las alternativas para trabajar con diferentes tipos de medios, especialmente orientados a su uso en herramientas de e-learning síncrono. Para terminar, se introduce el concepto de gestión del turno de palabra en el ámbito de conferencias web.

3.2. Transporte de la información multimedia

La principal diferencia que existen entre los estándares de la IETF y la ITU-T resultan de su orientación inicial. Mientras que los estándares de la IETF fueron concebidos siguiendo un modelo cliente/servidor, según la filosofía de la mayor parte de protocolos de Internet, los estándares de la ITU-T giran en torno a las redes RDSI. Estas diferencias se concentran, principalmente, en los protocolos de la capa de presentación y de sesión; los estándares de la IETF se organizan en torno al protocolo SIP, mientras que los estándares de la ITU-T se organizan en torno a la recomendación H.323. En la figura 3.1 se muestran las pilas de protocolos que se utilizan para las aplicaciones multimedia.

Ambas pilas de protocolos comparten el protocolo de red (IP) y los de transporte (UDP y TCP). Además, otro punto en común es el protocolo RTP, perteneciente a la IETF pero asumido por la ITU-T dentro de su pila de protocolos. Dentro de la norma H.323 existen dos protocolos que cumplen la misión equivalente a gslinksipSIP: H.245 y H.255.0.

Independientemente de la alternativa tecnológica elegida, bien sea la que representa la pila de protocolos ITU-T o IETF, éstas se organizan en capas, cada una con una funcionalidad determinada. En la figura 3.2 se muestra un resumen de los protocolos que pueden utilizarse en una aplicación multimedia y la misión de cada uno de ellos. En los siguientes apartados se analizan en detalle cada una de estas capas describiendo su misión y los problemas a los que debe hacer frente.

3.2.1. Comportamiento de las redes IP

Idealmente, para el correcto funcionamiento de las aplicaciones multimedia que deben transmitir datos a través de una red IP resultaría conveniente conocer cuál es el tipo y la congestión actual en la red subyacente. De esta forma, estas aplicaciones pueden adecuar su modo de funcionamiento al estado de la red. Este tipo de aplicaciones que son capaces de adaptarse a las situaciones cambiantes de la red se denominan *net-aware applications*.

El problema que presentan las redes IP es que, precisamente, el protocolo IP proporciona una capa de abstracción que oculta los detalles de los enlaces de red y su topología a las aplicaciones. Es común representar esta arquitectura que siguen las redes IP mediante un modelo de reloj de arena como el que se puede ver en la figura 3.3.

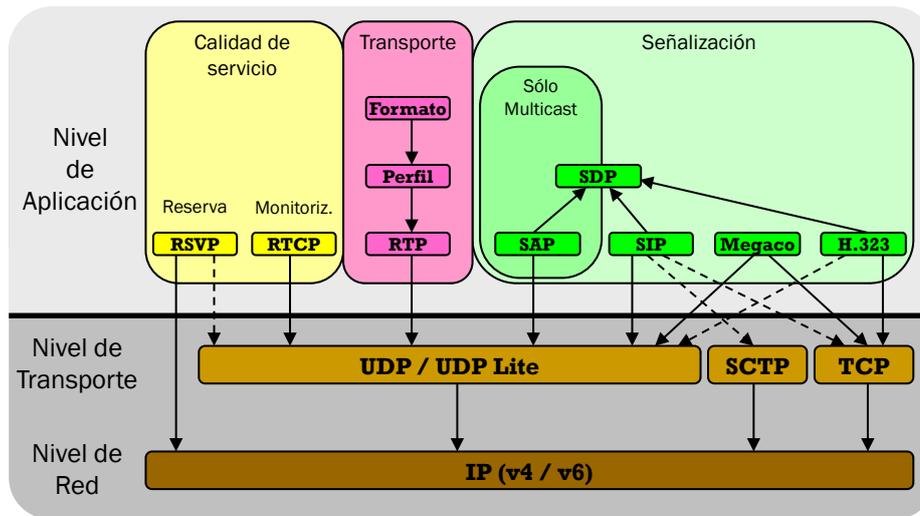


Figura 3.2: Funciones de cada protocolo de la pila multimedia

Ya que las aplicaciones no tienen una forma directa de conocer los detalles de la red sobre la que operan, una forma de recabar datos sobre la misma es mediante la observación de su comportamiento, lo que nos permitirá caracterizarla en base a unos parámetros. Por ejemplo, en [Paxson, 1999] se analizan las trazas de paquetes de 20 800 conexiones TCP entre 35 sitios de Internet para estimar diferentes dinámicas de la red como las pérdidas de paquetes, los paquetes fuera de orden, etc. En [Jain y Dovrolis, 2003] se describe una metodología para estimar el ancho de banda entre dos equipos conectados a Internet. De forma adicional, se pueden encontrar informes periódicos con resultados de mediciones del estado global de Internet en [ITR, 2008] e [ITA, 2008].

Habitualmente, en la caracterización de una red IP suelen considerarse los siguientes parámetros:

- Pérdida de paquetes, que representa la fracción de todos los paquetes enviados que no llegan a su destino.
- Replicación de paquetes, que mide el porcentaje de paquetes de los que llegan varias copias al receptor.
- Corrupción de paquetes, que representa la fracción del total de paquetes que llega al receptor pero con uno o varios bits de información modificados respecto al paquete original emitido por el emisor.
- Tiempo de tránsito en la red, que mide el tiempo que tarda en recibirse un paquete enviado desde el emisor en el receptor.
- Tamaño máximo del paquete, indica el tamaño máximo de un paquete que puede ser procesado por la red.

Ahora bien, a la hora de medir debemos tener presente que existirán diferencias entre unas y otras mediciones dependiendo de una serie de factores, algunos de los cuales son:

- Localización de las estaciones de medición: habrá que tener en cuenta tanto la ubicación geográfica como el número de enlaces y proveedores que debe atravesar el tráfico analizado.

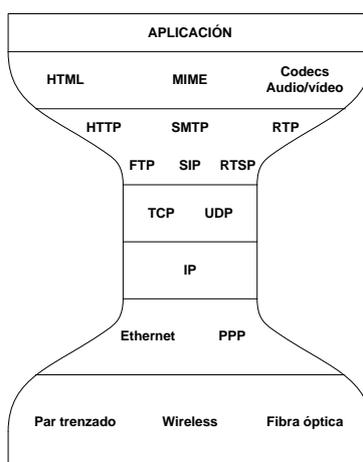


Figura 3.3: Modelo del reloj de arena de una red IP

- La hora del día en que se toman: suelen existir unas franjas horarias en las que el tráfico en la red es mayor que en otras, influenciando a otros tráficos.
- Tipo de red que debe atravesar el tráfico: las redes cableadas tendrán, en general, un mejor comportamiento que las redes inalámbricas.
- Otros tráficos en la red: habrá que considerar también el efecto de otros tráficos sobre las mediciones.

En los siguientes apartados se explican con más detalle en qué consisten cada uno de los parámetros que caracterizan una red IP.

3.2.1.1. Pérdida de paquetes

Existen diferentes métricas que se pueden utilizar para medir la pérdida de paquetes que se produce en una red. Así, podemos medir la pérdida media de paquetes para estimar cuál es la congestión actual de la red, mientras que midiendo los patrones que siguen esas pérdidas o la correlación entre ellos podemos analizar su comportamiento dinámico.

Gracias a los experimentos que se han ido realizando [Paxson, 1999], [Bolot y Vega-Garcia, 1996] y [Yajnik *et al.*, 1999], se puede decir que el ratio medio de pérdidas de paquetes en Internet no es constante y que cambia suavemente, aunque en ocasiones se pueden producir cambios bruscos.

En las redes actuales, aunque es posible diseñar el núcleo (*backbone*) de la misma libre de pérdidas, esto no implica que la red en su conjunto esté libre de pérdidas [Perkins, 2003], puesto que muchos enlaces de la red aún presentan pérdidas. En la figura 3.4 aparece un gráfico obtenido de [ITR, 2008] representando la pérdida media y máxima de paquetes a escala global durante el periodo comprendido entre el 7 de abril y el 7 de mayo de 2008. En un tono más claro se muestran los valores máximos, mientras que en uno más oscuro los valores medios. Puede observarse cómo el ratio medio de pérdidas oscila entre el 7 y el 10 % habitualmente.

Desde el punto de vista de una aplicación, resultaría deseable que las pérdidas de paquetes que se pudieran producir sucedieran como eventos aislados, de tal forma que estén distribuidos aleatoriamente en el tiempo. De este modo, una aplicación podría recuperarse más fácilmente de este tipo de problemas frente a un escenario en el que las pérdidas ocurrieran en ráfagas, esto es, que se perdieran varios paquetes consecutivos.

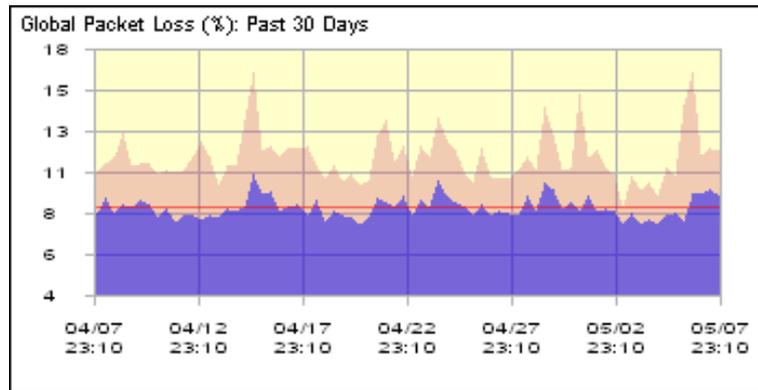


Figura 3.4: Ratios medio y máximo de pérdidas de paquetes en Internet durante los meses de abril y mayo de 2008

Por desgracia, los resultados que se obtienen de los diferentes estudios realizados nos dicen que las pérdidas no se distribuyen aleatoriamente, sino que la probabilidad de que un paquete se pierda aumenta si el anterior se ha perdido. Se puede concluir que en el 90 % de las ocasiones las pérdidas son de paquetes individuales frente al 10 % en ráfagas. De cualquier forma, la probabilidad de que las pérdidas ocurran en grandes ráfagas disminuye con el tamaño de éstas según una relación como la que puede verse en la figura 3.5.

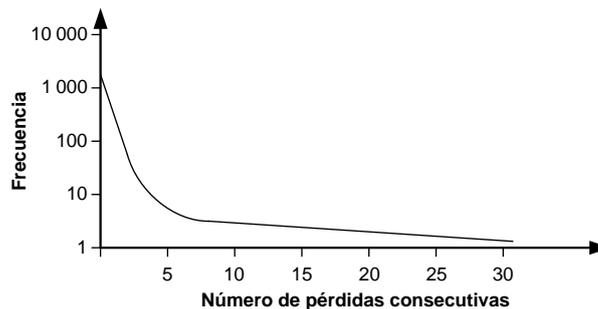


Figura 3.5: Distribución de frecuencia de las pérdidas de paquetes en ráfagas

3.2.1.2. Replicación de paquetes

Más infrecuente que las pérdidas, aunque posible, es la replicación de paquetes. Esto suele ser síntoma de un mal funcionamiento o una mala configuración de los dispositivos de red, esto es, conmutadores y enrutadores.

En realidad, esto no tendría por qué ser un problema para las aplicaciones. Deberían ser capaces de detectar que se trata de un paquete duplicado y, simplemente, descartarlo. Puede llegar a ser un problema si la frecuencia de paquetes duplicados es tal que reduce el ancho de banda disponible en la red. En ese caso, conviene detectar esa situación para reemplazar o reparar el dispositivo de red averiado o mal configurado.

De todas formas, la replicación de paquetes es un fenómeno inusual [Paxson, 1999], y como las aplicaciones pueden recuperarse fácilmente de una situación así, no se le suele prestar demasiada importancia.

3.2.1.3. Corrupción de paquetes

Otro de los fenómenos que ocurren en las redes IP es la corrupción de paquetes. Éste es más notorio en redes inalámbricas, más propensas a las interferencias que las redes cableadas. En todo caso, la utilización de IP junto con un protocolo de transporte como UDP o TCP, gracias al servicio de *checksum* que proporcionan, debería provocar la detección de la mayor parte de los paquetes corruptos.

Para las aplicaciones, la llegada de un paquete corrupto es percibida como un paquete perdido, debido a que es descartado en las capas de red (IP) o transporte (TCP o UDP). Esto, que a priori parece el comportamiento más adecuado, puede no ser lo más conveniente en algunas ocasiones. A veces, es conveniente para las aplicaciones recibir esos paquetes corruptos para intentar recuperar la información que contienen, especialmente en aplicaciones multimedia, en las que el coste de retransmitir el paquete habitualmente no es asumible.

3.2.1.4. Tiempo de tránsito en la red

El tiempo de tránsito es el tiempo que permanece el paquete en la red hasta que llega al destino. Esto incluye el tiempo desde que la aplicación emisora entrega el paquete a la pila de protocolos de red hasta que la aplicación receptora recibe el paquete.

Hay una serie de factores que influyen en el tiempo de tránsito:

- La longitud del camino a recorrer: en este sentido tiene más importancia el número de routers o de proveedores que se atraviesan que la distancia física real.
- El tamaño de las colas en los routers: es el factor que más influye y será tanto más influyente cuanto más grande sea la congestión de la red. Con una congestión elevada las colas de reenvío en los routers serán más largas, con lo que el tiempo de salto de un paquete de un enlace de red a otro será mayor.

Habitualmente, en lugar de trabajar con el tiempo de tránsito, muy difícil de medir directamente, se suele considerar el tiempo de ida y vuelta, es decir, el tiempo que transcurre desde que se envía una petición hasta que se obtiene una respuesta. A este tiempo se le conoce como *Round-Trip Time* (RTT) y se tiende a aproximar por el doble del tiempo de tránsito.

Otro término que se utiliza, y que es especialmente importante para las aplicaciones con restricciones temporales como pueden ser las aplicaciones multimedia, es el jitter. El *jitter* es la variación en el tiempo de tránsito, que no es siempre constante por la naturaleza cambiante de la red.

Desde el punto de vista de las aplicaciones, resulta más sencillo trabajar con una red que introduzca un retardo constante, siempre que éste se sitúe dentro de unos márgenes adecuados. En esta dirección, existen diferentes estudios para estimar la latencia máxima admisible por los usuarios en sistemas de voz sobre IP (VoIP). En [Mehta y Udani, 2001] se propone una división como la que aparece en la figura 3.6. De cualquier forma, la latencia final que perciben los usuarios es el resultado de una combinación de tiempos, además de la latencia de red, aunque sea ésta la que mayor peso tenga. Alguno de estos tiempos son:

- Tiempo de adquisición de los datos: es el tiempo que transcurre desde que se captura el primer dato hasta que se captura el último, todos ellos a enviar en el mismo paquete.
- Tiempo de codificación: es el tiempo que se invierte en codificar los datos ya capturados.
- Tiempo de empaquetado: es el tiempo que se necesita para formatear los datos y enviarlos a la red.

- Latencia de red: es el tiempo en el que un paquete está circulando por la red hasta que llega al receptor. Este incluye también el tiempo que se invierte en la pila de protocolos de emisor y receptor.
- Tiempo de desempaquetado: el tiempo que se invierte en extraer los datos útiles del paquete que llega por la red.
- Tiempo de decodificación: es el tiempo que se necesita para decodificar los datos antes de su reproducción.
- Latencia en el buffer de reproducción: es el tiempo que se añade para que la reproducción de los datos sea correcta y esté sincronizada.

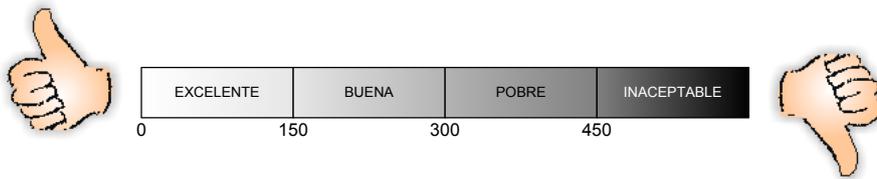


Figura 3.6: Calidad percibida en función de la latencia de red en audioconferencias (milisegundos)

La latencia máxima que se puede permitir en una conferencia de audio entre dos personas para que se considere óptima se estima en 150 milisegundos. En términos de la voz, la latencia representa el tiempo en que tarda en llegar al receptor una palabra pronunciada por el emisor. De ahí que deba fijarse un límite empírico para esta latencia para que ambos interlocutores no solapen sus conversaciones. A modo de comparación, el sistema telefónico tradicional siempre cumple esa premisa [Chong y Matthews, 2004]. Algunos autores cifran su latencia habitual entre 45 y 75 milisegundos.

En la figura 3.7 aparece un gráfico que muestra el RTT medio medido a escala global entre los meses de abril y mayo de 2008 según las mediciones publicadas en [ITR, 2008]. En azul se representa el RTT medio mientras que en rojo aparece el máximo. Conviene recordar que el tiempo de tránsito en la red se aproxima por la mitad del RTT.

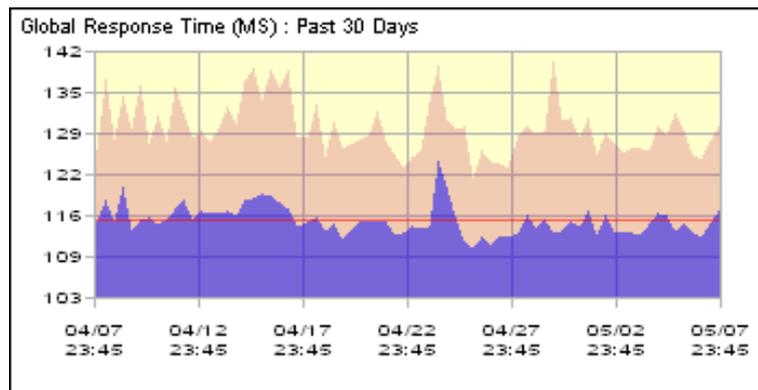


Figura 3.7: RTT medio y máximo en Internet durante los meses de abril y mayo de 2008

El problema al trabajar con la red aparece cuando este tiempo de tránsito no es constante. En ese caso, las aplicaciones multimedia deben estimar cuál es esa variación para adecuar su buffer

de reproducción para que éste no se vacíe y se produzca un corte en la reproducción [Perkins, 2003], un ejemplo de lo cual se muestra en la figura 3.8.

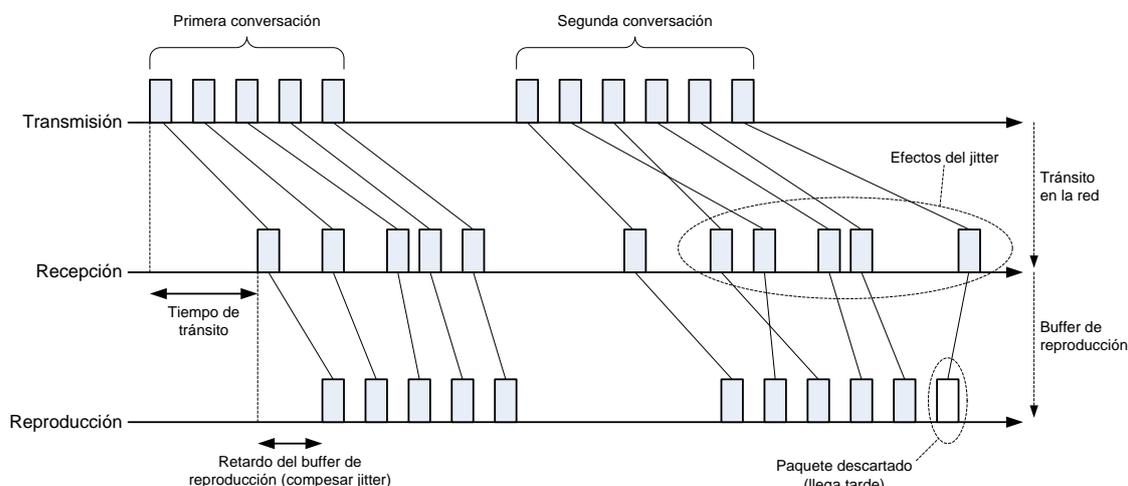


Figura 3.8: Efectos del jitter de la red sobre las aplicaciones multimedia

Existen diferentes técnicas para estimar el jitter:

- Mediante el ratio de llegada de paquetes: si el ratio de envío de paquetes es constante, el jitter se aproxima mediante la variación en el ratio de llegada de paquetes.
- Como la diferencia entre el tiempo de envío y el de recepción para cada paquete: esta forma, al contrario que la anterior, no requiere que el ratio de envío sea constante. Sin embargo, el problema que presenta reside en que es muy difícil medir el tiempo de tránsito absoluto de un paquete, puesto que se necesita una gran sincronización entre los relojes de emisor y receptor.

Para las conferencias de audio se suele estimar como admisible un jitter de hasta 75 milisegundos [Chong y Matthews, 2004].

3.2.1.5. Tamaño máximo del paquete

La MTU (*Maximum Transfer Unit*) representa el tamaño máximo de un paquete en bytes que la red puede transportar. Alguno de los valores típicos de MTU son:

- PPP: 576 bytes.
- IEEE 802.3 Ethernet: 1 500 bytes.

El resto de protocolos de enlace presentan una MTU superior a las anteriores. Es por esto que en las aplicaciones multimedia suele tomarse como tamaño de paquete 576 ó 1 500 bytes, aunque este último es más habitual.

En el caso de que el tamaño de un paquete exceda la MTU de la red, éste debe ser fragmentando en diferentes trozos para que el tamaño de éstos sea menor o igual que la MTU. En ningún caso las aplicaciones multimedia deben confiar en este mecanismo para enviar sus datos. La pérdida de un fragmento de paquete supondrá la pérdida del paquete entero, con lo que la probabilidad de que un paquete se pierda se vería multiplicada, acentuando los efectos sobre la aplicación.

3.2.1.6. Conclusiones

La principal conclusión que se puede extraer es que las aplicaciones en ningún caso deben suponer un comportamiento ideal de la red. Como hemos visto en los anteriores apartados, existen diferentes situaciones que afectan al funcionamiento de las aplicaciones: pérdidas de paquetes, corrupción de paquetes, retrasos, reordenamientos, etc.

De igual manera, estos efectos son percibidos de distinta forma en diferentes puntos de la red, es decir, existe una heterogeneidad inherente a la red que debe tenerse en cuenta.

Por un lado, las aplicaciones deben diseñarse de tal forma que estén preparadas para soportar una pequeña fracción de paquetes perdidos sin afectar excesivamente a su comportamiento. Además, estas pérdidas se producirán, en la mayor parte de los casos, como eventos aislados, rara vez como ráfagas muy cortas y, excepcionalmente, como ráfagas del orden de fracciones de segundo.

Por otro lado, existe la posibilidad de que los paquetes lleguen duplicados o corruptos. Una aplicación multimedia tendrá un mejor comportamiento si es capaz de recuperar los datos de un paquete corrupto.

Por último, es necesario tener en cuenta que durante el tránsito de los paquetes por la red se pueden producir reordenamientos, con lo que los paquetes pueden llegar al receptor en un orden distinto al que fueron enviados desde el emisor. Igualmente, el tiempo de tránsito no es constante, existe un jitter que es necesario estimar para que la reproducción de los datos sea continua.

3.2.2. IP Multicast

IP Multicast es una tecnología que nos permite ahorrar ancho de banda en la red reduciendo el número de conexiones individuales necesarias para transportar la misma información a múltiples receptores. Entre las aplicaciones que pueden beneficiarse de sus características están las aplicaciones de videoconferencia, enseñanza a distancia, herramientas colaborativas, etc.

La figura 3.9 muestra cómo los datos desde un emisor llegan a varios receptores a través de IP Multicast. Los routers que se sitúan en el camino de datos entre la fuente y los receptores deben encargarse de hacer llegar la información que envía ésta replicando los paquetes de datos cuando sea necesario. En un color más intenso se muestra el camino de los datos a través de la red multicast.

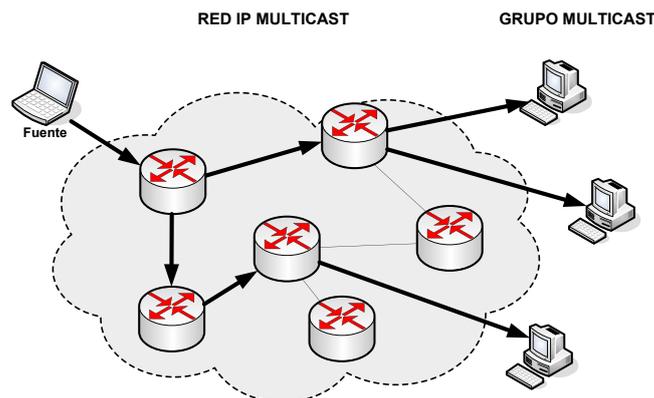


Figura 3.9: Transmisión de datos en una red IP Multicast

3.2.2.1. Grupo multicast

La tecnología IP Multicast se basa en el concepto de *grupo multicast*, que representa un grupo de receptores que expresan su interés en recibir un flujo de datos particular. Este grupo no tiene ninguna limitación física ni geográfica, los receptores podrían ubicarse, teóricamente, en cualquier lugar de Internet. Cualquier receptor que desee recibir los datos emitidos deberá suscribirse al grupo multicast utilizando IGMP [Cain *et al.*, 2002]. Por el contrario, para enviar datos a un grupo multicast no es necesaria la suscripción al mismo.

Cada grupo multicast es identificado mediante una dirección IP dentro de un determinado rango destinado a las direcciones multicast. De esta forma, una dirección IP Multicast representa a un conjunto de receptores que reciben los datos a través del grupo identificado mediante dicha dirección IP.

La ICANN es la sucesora de la IANA en el control de la asignación de las direcciones IP, identificadores de protocolos y diferentes labores de gestión de los nombres de dominio de Internet. Específicamente, reserva para las direcciones IP Multicast una clase de tipo D que abarca el rango de direcciones desde 224.0.0.0 hasta 239.255.255.255. Este rango de direcciones se divide a su vez en distintos subrangos:

- Direcciones reservadas para el enlace local: su ámbito se restringe a un segmento de red local y cubre el rango 224.0.0.0 a 224.0.0.255.
- Direcciones de ámbito global: son válidas en todo Internet y su rango abarca las direcciones desde 224.0.1.0 hasta 238.255.255.255.
- Direcciones de ámbito limitado: este tipo de direcciones se pueden utilizar en el ámbito de la red corporativa de una única organización. Su rango es desde la dirección 239.0.0.0 hasta 239.255.255.255.

3.2.2.2. Internet Group Management Protocol (IGMP)

IGMP [Cain *et al.*, 2002] se utiliza para suscribir dinámicamente hosts individuales a grupos multicast en una LAN particular. Los hosts utilizarán mensajes IGMP enviados a su router local para realizar la suscripción a grupos multicast. De esta forma, le informan de su deseo de recibir los datos enviados a dicho grupo, y el router descubre a qué subredes debe enviar los datos que recibe dirigidos a ese grupo multicast.

Adicionalmente, bajo IGMP, los routers esperan por mensajes IGMP y, periódicamente, enviarán peticiones para descubrir qué grupos están activos o inactivos en una subred particular.

3.2.2.3. Árboles de distribución multicast

Los routers con capacidad multicast crean árboles de distribución de los datos que determinan el camino que sigue el tráfico multicast a través de la red para hacerlo llegar a todos los receptores que previamente se suscribieron al grupo multicast. Hay dos tipos básicos de árboles de distribución: los árboles de fuente y los árboles compartidos.

La forma más simple de árbol de distribución es un *árbol de fuente*. Es un árbol cuya raíz es el origen del tráfico multicast y cuyas ramas se expanden a través de la red hasta los receptores. Debido a que este tipo de árbol utiliza el camino más corto a través de la red, también se le conoce como *árbol del camino más corto* (SPT).

A diferencia de los árboles SPT, que tienen la raíz en la fuente, los *árboles compartidos* utilizan una única raíz común en un punto concreto de la red que recibe el nombre de *rendezvous point*. El tráfico de todas las fuentes fluye hasta el punto de encuentro, desde el cual se distribuye a todos los receptores del grupo multicast.

Tanto los árboles SPT como los compartidos están libres de bucles. Los paquetes únicamente se replican cuando el árbol se bifurca. Además, ambos tipos de árboles deben actualizarse dinámicamente puesto que los miembros de un grupo multicast pueden unirse y abandonar el grupo en cualquier instante. Cuando todos los receptores en una rama particular abandonan el grupo, los routers podan esa rama y dejan de enviar datos por ese camino.

Los árboles SPT tienen la ventaja de que crean el camino óptimo entre la fuente y los receptores, lo cual garantiza una latencia de red mínima. Pero tienen la desventaja de que los routers deben mantener la información del camino para cada fuente. En una red con miles de fuentes y grupos aparecerán pronto problemas relacionados con los recursos de los routers.

Los árboles compartidos tienen la ventaja de requerir la mínima información de estado en cada router. Por contra, su principal desventaja es que, en ocasiones, el camino que siguen los paquetes entre la fuente y los receptores no es el óptimo, lo que implica la introducción de latencias adicionales en el nivel de red.

3.2.3. Efectos del protocolo de transporte

Como se ha visto en el apartado anterior, la utilización de IP para el transporte de medios continuos en tiempo real supone una serie de inconvenientes. Es por esto que, en la práctica, nunca se utiliza IP únicamente, sino que es necesario disponer un protocolo de transporte encima de la capa IP que trata de corregir o atenuar estas deficiencias. Habitualmente, en las herramientas de e-learning síncrono se utilizan dos protocolos de transporte: TCP y UDP. TCP es un protocolo fiable y orientado a conexión, mientras que UDP no. La transmisión de paquetes UDP se efectúa de una manera más ágil que la de paquetes TCP, ya que se evitan etapas de negociación y los paquetes UDP contienen menos información de control que los TCP. Debido a estas diferencias, cada protocolo es empleado con diferentes fines pero ambos son igualmente relevantes.

Actualmente, todas las aplicaciones multimedia emplean UDP como protocolo de transporte a bajo nivel. Esto se debe a que las principales virtudes de TCP (control de flujo y de congestión, detección de paquetes perdidos y duplicados, recepción de los mismos en el orden correcto, mecanismos de reenvío...) se convierten en desventajas en el terreno de la transmisión de flujos multimedia, donde es preferible tener bajos retardos en la comunicación a asegurar que toda la información llega correctamente. Pero UDP no es ni mucho menos un protocolo diseñado con esta finalidad, por lo que deberá ser complementado con algunos mecanismos de control específicos similares a los de TCP, que actuarán en niveles superiores. Por contra, el protocolo TCP es el elegido para la transmisión de otro tipo de información, principalmente de señalización, cuya integridad debe asegurarse. Otra razón por la que se ha venido empleando UDP es porque el diseño original de TCP no contempla transporte multicast, una característica muy deseable en cualquier aplicación multimedia de tiempo real en Internet. Este argumento ha ido perdiendo peso con el paso del tiempo, ya que en los últimos años se han desarrollado protocolos multicast fiables basados en la filosofía de TCP, pero no compatibles con él, lo que constituye su mayor handicap. Por el momento ninguna plataforma de e-learning síncrono ha abandonado el uso de UDP en favor de estos nuevos protocolos.

En definitiva, para el transporte de flujos de datos con restricciones de tiempo real se utiliza UDP por su mejor rendimiento temporal, ya que no espera recibir confirmación de la llegada de cada uno de los paquetes. Sin embargo, UDP no aumenta la fiabilidad de las conexiones, los paquetes se pueden perder, retrasar e incluso sufrir un reordenamiento, los mismos problemas que presentaba la red IP, sólo añade puertos para realizar la multiplexación sobre la misma dirección de red y un servicio de checksum que, a diferencia de IP, afecta a todo el contenido del paquete en lugar de a la cabecera exclusivamente.

Además de TCP y UDP existen otros protocolos de transporte diseñados para dar una base más sólida y adecuada a los servicios de transporte en tiempo real, tanto para transporte de datos como para transporte de señalización. El objetivo de estos protocolos es descargar a las capas

superiores de las tareas adicionales que UDP no puede gestionar, como el control de la congestión y de los errores. En especial SCTP (*Stream Control Transmission Protocol*) y UDP Lite pueden coexistir e incluso sustituir en un futuro próximo a UDP como protocolo de transporte base para los sistemas de tiempo real en Internet.

El protocolo SCTP [Stewart *et al.*, 2000], aunque diseñado para el transporte de los mensajes de señalización SS7 de la red telefónica tradicional sobre redes IP, es válido para un rango mucho mayor de aplicaciones. SCTP proporciona, entre otras, las siguientes funciones: transferencia de datos fiable, definición de múltiples flujos para una misma conexión, posibilidad de seleccionar entrega ordenada o desordenada por separado para cada flujo, monitorización, fragmentación y reunificación de datos, y control de congestión y flujo compatibles con TCP. Asimismo, contempla funciones adicionales de seguridad, como protección contra ataques por denegación de servicio o por enmascaramiento. Por su diseño, es muy adecuado para protocolos de señalización de sistemas de tiempo real, como *Session Initiation Protocol* (SIP). El estudio de su uso como protocolo de transporte para SIP [Rosenberg *et al.*, 2005] ha puesto de manifiesto algunas de sus posibles ventajas respecto a UDP en el contexto de los sistemas en tiempo real, derivadas fundamentalmente del hecho de que SCTP es un protocolo orientado a la conexión: proporciona mecanismos para retransmitir con rapidez segmentos perdidos y control de congestión, y permite fragmentación en el nivel de transporte, evitando recurrir a la fragmentación de datagramas IP. Estas ventajas frente a UDP también las presenta TCP, pero por otro lado SCTP permite obtener prestaciones ligeramente mejores que TCP, en especial cuando se producen pérdidas.

UDP Lite [Larzon *et al.*, 2004] se ha diseñado pensando en aplicaciones para dispositivos móviles inalámbricos en las que, respecto a la mayoría de las aplicaciones de Internet, existe una mayor frecuencia de errores. Es un protocolo de transporte que difiere del UDP tradicional únicamente por la sustitución del campo longitud de la cabecera UDP por un nuevo campo que indica el número de octetos sobre los que se debe aplicar el algoritmo de comprobación de validez. La comprobación es pues parcial, de modo que los segmentos quedan divididos en una parte sometida a verificaciones (generalmente las cabeceras de los distintos niveles) y una parte en la que se aceptan errores (datos de la aplicación). Si el valor del rango de comprobación es igual al de la longitud del segmento, UDP Lite no difiere de UDP; si el valor es menor, se admiten errores en la parte del segmento no cubierta por el algoritmo de comprobación. Esta modificación permite que las aplicaciones multimedia reciban el mayor número posible de los segmentos que componen los flujos de audio o vídeo. De este modo, el impacto negativo en la calidad de reproducción en los receptores se reduce drásticamente para muchos formatos de codificación.

3.2.4. Real-time Transport Protocol

Como se comenta en el apartado 3.2.3, UDP presenta muchas carencias para el transporte de datos multimedia en la red, por no mencionar TCP, que resulta poco idóneo para este fin debido a la imposibilidad de utilizar técnicas de transporte multicast. Por esta razón, las deficiencias de UDP han de ser compensadas añadiendo mecanismos de control de errores y de congestión adaptables a las necesidades concretas de cada aplicación. La solución que propone la IETF, y que es empleada por la totalidad de herramientas de e-learning síncrono, es el protocolo de transporte en tiempo real RTP (*Real-time Transport Protocol*) [Schulzrinne *et al.*, 2003a]. Aunque es independiente del protocolo subyacente, las implementaciones actuales de RTP se apoyan en la práctica en el protocolo UDP. Si las capas inferiores lo permitiesen, será posible el transporte de los datos tanto en entornos unicast como multicast. Esto requiere que todos los mecanismos contemplados por el protocolo sean escalables.

El protocolo RTP tiene como función el transporte de extremo a extremo para aplicaciones que requieren transmisión de datos sometidos a restricciones temporales, como pueden ser audio, vídeo o datos de simulación. Sin embargo, no garantiza la entrega de los paquetes en un tiempo máximo ni otro tipo de calidad servicio. En lugar de ello, confía en los servicios de capas inferiores

para llevarlos a cabo. De este modo, las funciones de reserva de recursos y garantía de calidad de servicio, necesarias en gran parte de las aplicaciones de conferencia multimedia, quedan fuera del alcance del protocolo. Por tanto, es posible que los paquetes RTP enviados lleguen retrasados, fuera de orden o incluso se pierdan.

RTP fue diseñado para ser adaptable a las necesidades de cualquier aplicación, por lo que está deliberadamente incompleto, proporcionando una estructura y un conjunto de mecanismos comunes, en lugar de algoritmos concretos. Por norma general, RTP vendrá integrado como parte de las aplicaciones en lugar de implementarse como una capa separada. Esto permite seguir, si se desea, un enfoque de procesamiento integrado de los diferentes niveles que componen las aplicaciones; lo cual posibilita implementar un mecanismo, por ejemplo el control de congestión, en la capa que se considere más conveniente, a nivel de transporte o a nivel de aplicación, o incluso en más de una. A esto se le conoce como *application-level framing*. Por esta razón, cuando se utilice RTP para una aplicación en particular requerirá, además de la especificación RTP [Schulzrinne *et al.*, 2003a], los siguientes documentos:

- un documento de especificación de un perfil, que define un conjunto de códigos para el campo `payload`¹ y su traducción a formatos de datos, (*i.e.* formatos de codificación). Un perfil puede, además, definir extensiones a la norma RTP para un determinado tipo de aplicaciones.
- documentos de especificación del formato de la carga útil, que definen cómo un determinado tipo de datos, por ejemplo audio o vídeo codificado, debe ser empaquetado y transportado sobre RTP.

Junto con su perfil de audio y vídeo [Schulzrinne *et al.*, 2003b], RTP se ha diseñado para dar soporte a conferencias multimedia con múltiples participantes, aunque es asimismo extensible a otras aplicaciones. Huelga decir, que RTP es el estándar *de facto* para todas las soluciones VoIP. En definitiva, RTP es el protocolo idóneo para la transmisión de casi toda la información que se intercambia en una sesión de e-learning síncrono: audio, vídeo, diapositivas, instrucciones para la presentación de documentos, órdenes al compartir aplicaciones, trazos de la pizarra, etc; unos por tratarse de medios continuos y el resto por las ventajas que supone la utilización de RTP para su transporte y que se analizan en el apartado 4.2.

Entre las funciones que RTP proporciona están, entre otras, funciones de sincronización, detección de pérdidas, seguridad, multiplexación de flujos provenientes de varias fuentes e identificación de participantes y contenidos. Esto se consigue mediante la utilización de marcas de tiempo y números de secuencia, entre otros mecanismos. Las funciones señaladas están repartidas entre dos componentes: el protocolo de transporte de datos RTP y su protocolo de control *Real-time Transport Control Protocol* (RTCP). Como consecuencia, una *sesión* RTP está compuesta por dos direcciones de transporte, habitualmente una dirección de red y dos puertos UDP, para el tráfico RTP y RTCP respectivamente. En ese caso, se recomienda que el puerto para los datos RTP sea un número par, y para los datos de control RTCP se utilice el puerto impar inmediatamente superior.

RTCP es un protocolo de control de la transmisión basado en el intercambio periódico de paquetes de control entre los participantes. Su uso requiere que el protocolo de transporte subyacente realice la multiplexación de los paquetes de control y datos, que se transmiten por separado. La definición del protocolo de control como componente adicional permite que aplicaciones que requieran un mayor nivel de control que el proporcionado por RTCP lo reemplacen con otros protocolos de control del transporte específicos. Algunas implementaciones, incluso, no hacen uso de protocolo de control alguno, incluido RTCP, por lo que pierden toda la funcionalidad que este último proporciona:

¹Este campo, como indica su nombre en inglés, determina la carga útil que transporta un paquete RTP.

- Identificación de participantes y sincronización entre distintas sesiones RTP: los paquetes RTCP contienen un identificador textual, el nombre canónico (CNAME), único para cada participante en una sesión multimedia. Este identificador, a diferencia del identificador de fuente (SSRC) presente en la cabecera de los paquetes RTP, es único para todas las sesiones RTP en las que participe el usuario. Por tanto, el nombre canónico permite determinar qué flujos de datos proceden de un mismo participante, de modo que sea posible sincronizar distintas sesiones RTP dentro de una sesión multimedia. Además, el nombre canónico permite identificar a cada participante de forma persistente, independientemente de si en algún momento se produce un conflicto de identificadores de fuente. Esta característica es fundamental y se explota al máximo en cualquier aplicación de e-learning síncrono, donde los participantes suelen emitir o recibir varios flujos al mismo tiempo, por ejemplo audio, vídeo, anotaciones, etc.
- Realimentación acerca de la calidad de recepción: esta información, proporcionada en distintos tipos de paquetes de informe, sirve de base para monitorizar y diagnosticar problemas de transmisión, en especial en entornos multicast. Esta información también es aprovechada por mecanismos de control de congestión adicionales.
- Estimación del número de participantes: el envío de paquetes de control RTCP por parte de todos los participantes en una sesión permite que cada uno de ellos pueda estimar de forma dinámica el número total de participantes. Así, pueden ajustar la frecuencia de transmisión de paquetes RTCP para mantener el límite de ancho de banda de control establecido en la sesión. Este intervalo de transmisión será mayor cuanto mayor sea el número de participantes en la sesión RTP. De esta forma, RTCP es escalable según el número de usuarios.
- Proporcionar información mínima de control de la sesión: como nombre, dirección o teléfono de cada participante. Esta función opcional se basa en la definición de unos pocos tipos de paquetes RTCP que proporcionan la información necesaria para realizar un control mínimo de la sesión. Las aplicaciones con requisitos especiales de control pueden sustituir o complementar estos mecanismos con protocolos de control más complejos.

En esta capa se suelen implementar mecanismos de control de congestión, que emplean la información proporcionada por los paquetes RTCP acerca del estado de la red. Algunas variantes de las técnicas de control de congestión se han aplicado satisfactoriamente en sistemas basados en RTP [Sisalem y Schulzrinne, 1998a] y [Sisalem y Schulzrinne, 1998b], comprobándose que se consigue un reparto equitativo del ancho de banda disponible entre varios flujos RTP, y entre éstos y otros flujos TCP. A estos algoritmos que comparten adecuadamente el ancho de banda disponible con tráfico TCP se les conoce como *TCP-friendly* [Widmer *et al.*, 2001].

Tanto RTP como RTCP admiten diferentes extensiones. Es posible definir paquetes de datos y de control específicos de una aplicación con fines experimentales. También es posible definir cabeceras adicionales en los paquetes RTP de datos así como dentro de la carga útil del paquete. Incluso a través del perfil se puede definir un significado diferente a campos concretos de la cabecera RTP.

En ocasiones, especialmente en entornos VoIP con bajo ancho de banda, el tamaño de la cabecera es significativo respecto al tamaño de la carga útil de los paquetes de datos. Una forma de optimizar las prestaciones de RTP cuando se transmite a través de enlaces de ancho de banda pequeño es comprimir el conjunto de la cabecera RTP/UDP/IP de cada paquete [Casner y Jacobson, 1999] de manera análoga a como se realiza con las cabeceras TCP/IP [Jacobson, 1990].

Por último, hay que apuntar que la ITU-T recomendó el uso de RTP incluso antes de que se convirtiera en un estándar². De modo que no hay dudas en cuanto a la aceptación de este

²El protocolo RTP es el estándar número 64 de la IETF.

protocolo por parte de los organismos reguladores y de las empresas desarrolladoras; téngase en cuenta que la categoría de estándar de la IETF sólo se otorga, tras muchos años de estándar no oficial, cuando una tecnología es verdaderamente aceptada por todos los entes reguladores y ampliamente empleada por la industria.

La IETF ha publicado más de cuarenta *Request For Comments* (RFC) dedicados a tipos de carga concretos; en ellos se especifica la forma de optimizar su transmisión con RTP. Además, el perfil de audio y vídeo posee el estatus de estándar, pero la IETF no ha publicado más RFC documentando perfiles, exceptuando el que explica el protocolo RTP seguro (SRTP) [Baugher *et al.*, 2004]. Sí se han publicado sendos RFC [Ott *et al.*, 2006] y [Ott y Carrara, 2008] que amplían los dos perfiles anteriores, cuya filosofía puede ser muy interesante de cara a mejorar el rendimiento de las aplicaciones de e-learning síncrono. En ellos se contempla el uso de mensajes RTCP adicionales a modo de realimentación para conseguir latencias de transmisión menores.

3.2.4.1. Multiplexación de sesiones

Es muy común que una sesión multimedia lleve consigo asociadas una o más sesiones RTP, cada una de las cuales se corresponde con un tipo de datos distinto. Una sesión RTP queda identificada por un par de direcciones de transporte (dirección de red y número de puerto). Por ejemplo, en una videoconferencia en la que se transmite audio y vídeo existirán dos sesiones RTP, a cada una de las cuales se asocia un par de direcciones de transporte distintas³.

Los flujos de audio y vídeo no deberían transportarse en una única sesión RTP y demultiplexarse en base al tipo de carga o el campo SSRC. Los paquetes entrelazados con diferentes tipos de datos (audio, vídeo...) pero utilizando el mismo SSRC podrían provocar varios problemas:

1. Si dos flujos de audio comparten la misma sesión RTP y el mismo valor SSRC, y uno de esos flujos cambia la codificación y, en consecuencia, adquiere un tipo de formato RTP diferente, no habría forma de identificar qué flujo ha cambiado su codificación.
2. Un SSRC se define para identificar un único espacio de marcas de tiempo y de números de secuencia. Entrelazar múltiples tipos de datos podría necesitar diferentes espacios de tiempo si la frecuencia de muestreo de los datos difiere, y podría necesitar diferente espacio de números de secuencia para decidir qué tipo de datos ha sufrido la pérdida de un paquete.
3. Los informes periódicos RTCP solamente describen un espacio de marcas de tiempo y de números de secuencia por SSRC y no llevan un campo de tipo de carga útil, con lo que sería imposible componer los informes de recepción y emisión.
4. Un mezclador RTP no sería capaz de combinar flujos entrelazados de diferentes tipos de medios en un único flujo.
5. Transportar múltiples tipos multimedia en una misma sesión RTP imposibilita:
 - el uso de diferentes caminos de red o asignación de recursos de red para cada uno.
 - la recepción de un subconjunto de los datos multimedia, por ejemplo sólo el audio si el vídeo excede la capacidad del ancho de banda.
 - implementaciones de los receptores que utilizan procesos separados para diferentes tipos multimedia, mientras que utilizando sesiones RTP separadas para cada tipo multimedia permite desarrollos mono o multi-proceso.

³Habitualmente, tendrán la misma dirección de red y diferentes números de puerto. No es así en el caso de distribución multicast, donde se suele utilizar diferente dirección de grupo multicast para cada sesión RTP.

Si se utiliza un identificador de fuente distinto para cada tipo multimedia pero enviándolos en la misma sesión RTP se evitan los tres primeros problemas pero no los dos últimos.

Por otra parte, multiplexar varias fuentes relacionadas del mismo tipo multimedia en una sesión RTP utilizando diferentes SSRC es la norma para sesiones *multicast*. Los problemas anteriormente citados no aparecen: por ejemplo, un mezclador RTP puede combinar varias fuentes de audio; y el mismo tratamiento es aplicable para cada una de ellas.

3.2.4.2. Sistemas intermedios

El diseño de RTP prevé el uso de sistemas intermedios de nivel RTP, además de los sistemas terminales o aplicaciones de usuario. El fin de estos sistemas es conectar dos dominios del nivel de transporte para realizar una transformación en los datos RTP que fluyen entre ambos dominios. En ocasiones, hay casos en los que todos los participantes en una sesión RTP pueden no querer recibir la información multimedia en el mismo formato, especialmente en sesiones que se extienden a varias redes, unas con ancho de banda significativamente mayor que otras. En lugar de que todos se vean obligados a utilizar la configuración más restrictiva, es decir, la que menor ancho de banda consume y por tanto menos calidad ofrece, se puede introducir un sistema intermedio que realice una adaptación de las señales. Asimismo, algunos participantes pueden disponer del ancho de banda necesario para seguir una sesión multicast, pero no tener acceso multicast por encontrarse tras un cortafuegos. Para estos casos, un repetidor de nivel RTP, puede hacer de intermediario. En estos casos, se requiere instalar dos repetidores (o uno bidireccional), uno a cada lado del cortafuegos; el primero introduce los flujos procedentes de la red multicast en la red unicast, y el segundo realiza la operación inversa.

En función de la operación que realicen sobre los flujos de datos RTP existen dos tipos de sistemas intermedios:

- Traductores: las operaciones que realizan se limitan a flujos individuales, luego les salen tantos flujos como les entran, y éstas pueden abarcar desde no alterar los datos, reenviándolos tal cual le llegan; modificar algún parámetro de transporte, lo que permite conectar dominios con diferentes protocolos de transporte o grupos multicast con participantes que no tiene la posibilidad de utilizar IP Multicast; eliminar la encriptación de los datos; e, incluso, recodificar por completo el flujo utilizando otro formato.
- Mezcladores: recibe flujos de datos RTP de varias fuentes y los combina para formar un nuevo flujo. De esta forma, un mezclador RTP es visto como una nueva fuente de datos por el resto de participantes de la sesión. La combinación puede consistir en una simple selección de un flujo de entre todos los flujos a combinar, o bien puede ser una operación más compleja que implique la fusión de todos los flujos en un nuevo flujo recodificado.

Las cualidades de los sistemas intermedios es aprovechada por la mayoría de las aplicaciones de e-learning síncrono, con el fin de ofrecer a todos los usuarios la mejor calidad posible de los medios transmitidos para cada ancho de banda de las redes de acceso.

3.2.5. Calidad de servicio

La calidad de servicio en las transmisiones de información multimedia, y en particular en las plataformas de e-learning síncrono, suele ser un aspecto clave sobre el que se desea tener control, aunque rara vez se consiga. Hay que tener en cuenta que por la red circulan multitud de flujos de comunicaciones transportando todo tipo de datos, algunos con requisitos temporales, como la mayoría de los flujos que maneja una aplicación de e-learning síncrono, y otros no. La mayoría de las veces el hardware de la red ignora que existen estos flujos con dependencias temporales, por lo que no les concede prioridad alguna. Incluso cuando la naturaleza de los datos transmitidos

es conocida por la red, los dispositivos que la conforman pueden no considerar necesario dar prioridad o asignar más recursos a estos flujos.

En la actualidad los mecanismos que permiten implementar alguna clase de calidad de servicio (QoS) se dividen en dos grupos principales, según su filosofía de servicio:

- Servicios integrados (IntServ): se utiliza este término para un servicio de Internet que contemple los modelos de tráfico *best-effort*, tráfico compuesto por flujos de información en tiempo real, y la compartición controlada del enlace asignando tráfico de diferentes tasas binarias. En este modelo, la tradicional entrega *best-effort* de los paquetes coexiste con unas opciones mejoradas de entrega de los mismos basadas en las especificaciones de ciertas clases de QoS [Braden *et al.*, 1994]. RSVP [Braden *et al.*, 1997] es el protocolo que se utiliza en este modelo para solicitar a la red una calidad de servicio específica para un flujo de datos.
- Servicios diferenciados (DiffServ): este modelo permite asignar diferentes niveles de servicio a diferentes usuarios de Internet, entendiendo como usuarios a tipos de aplicaciones. En general, en el uso en Internet se entiende este término como cualquier mecanismo que no dependa completamente de la reserva de recursos hecha para cada flujo de datos, y que trate a los distintos usuarios de forma diferente, definiendo varios niveles de calidad de servicio.

Los mecanismos de control de QoS mediante señalización RSVP, y otras técnicas compatibles de servicios integrados, actúan de forma menos gruesa que los basados en servicios diferenciados, ya que son los más precisos a la hora de distinguir los diferentes niveles de calidad de servicio, pues cada flujo es catalogado de forma individual. Aunque la complejidad de los primeros es bastante más elevada, las aplicaciones multimedia que implementan técnicas de QoS suelen emplear RSVP. El protocolo RSVP puede implementarse en los nodos terminales y en los routers. El receptor de un flujo es el que inicia la señalización con paquetes RSVP para solicitar la reserva de recursos a lo largo del camino entre cada par de nodos que soporten el protocolo; esto obliga a almacenar los estados de la reserva en cada uno de los puntos (routers y otros nodos) donde ésta se realiza, lo que supone un problema de escalabilidad, pues se necesita almacenar un estado por cada flujo y nodo, y estos estados han de retransmitirse desde cada nodo al resto de forma periódica para que todos posean la información actualizada. Todos los nodos que tengan este tipo de mecanismo QoS habilitado emiten y reciben continuamente paquetes de señalización, por ello no se trata de un protocolo fácilmente escalable pues los paquetes de señalización pueden literalmente apropiarse de los recursos de red. La consecuencia es que RSVP no puede adoptarse en todos los escenarios; y uno de los más comprometidos son las sesiones de e-learning síncrono a través de Internet.

El modelo de servicios integrados tiene la ventaja de poder desarrollar toda la política de QoS y la clasificación de tráfico en conjuntos específicos de máquinas que limitan con redes que no están sometidas al control de QoS; aunque también se pierde el control sobre las decisiones de los routers intermedios, lo que hace que los resultados en la práctica sean bastante impredecibles. A este hecho se le une que su uso a menudo es cuestionado porque en las situaciones en las que realmente se espera su eficacia, esto es, cuando los recursos de red están casi al máximo de utilización, se observa que el tráfico de señalización puede llegar a superar el alivio de ancho de banda conseguido mediante la reserva y/o el desalojo de recursos. Por ello, en la práctica estas técnicas basadas en servicios integrados suelen implementarse en mecanismos de racionamiento, como los que emplean proveedores de Internet y empresas para limitar el ancho de banda consumido por el tráfico punto a punto.

Para remediar la impredecibilidad de DiffServ en las transmisiones entre varios usuarios pertenecientes a organizaciones o dominios distintos la IETF ha introducido el concepto de *bandwidth broker* [Nichols *et al.*, 1999], que consiste en un agente intermedio que conoce las políticas a seguir en materia de QoS para la transmisión de los flujos de estos usuarios. Este agente utilizará la información de la que dispone para enviar señales de control a un número limitado de nodos intermedios. A pesar de esta medida, seguirá habiendo un gran número de routers intermedios

que no conocerán la política de QoS que sería conveniente aplicar. En cualquier caso, el empleo de estos agentes mejora considerablemente las comunicaciones extremo a extremo sin necesidad de tener un gran control sobre la red que separa ambos extremos. Aunque RSVP es el protocolo que típicamente se ha asociado a la gestión de QoS en las transmisiones multimedia por Internet, los mecanismos basados en DiffServ han ido mejorando sus prestaciones reales. Esta circunstancia, unida a la mayor libertad de diseño y despliegue que proporcionan los protocolos bajo este modelo, ha llevado a que se contemple su uso en escenarios tan hostiles como el de VoIP a través de Internet [Escribano *et al.*, 2002].

3.2.6. Protocolos de señalización

Varias son las misiones que debe cumplir un protocolo de señalización:

- Establecimiento y finalización de las sesiones multimedia.
- Negociación y selección de características de cada sesión, en especial los formatos y la información multimedia que se desea y es posible intercambiar. Esta función se debe realizar tanto durante la etapa de establecimiento de la sesión como durante el transcurso de ésta.
- Gestión de participantes, para incorporar y dar de baja usuarios.
- Localización de usuarios y traducción de direcciones, lo que requiere un procedimiento de adaptación de diversos tipos de direcciones, como números de teléfono tradicionales, direcciones de correo electrónico, correo de voz o páginas web.

El conjunto de protocolos H.323 [Union, 2003], propuesto por la ITU-T, y el protocolo SIP (*Session Initiation Protocol*) [Rosenberg *et al.*, 2002b], propuesto por la IETF, constituyen los protocolos de señalización que han tenido más éxito y están actualmente más difundidos. SIP fue diseñado con posterioridad con la pretensión de que presentase dos ventajas frente a H.323, como son una mayor flexibilidad para incorporar nuevas funciones y una implementación más sencilla de realizar y depurar. Mientras que SIP es un protocolo ligero similar a otros desarrollados previamente por la IETF y de amplia extensión, como HTTP, la recomendación H.323 sigue un esquema similar al sistema tradicional de conmutación de circuitos, basándose en la señalización de RDSI, y otras recomendaciones de la serie H.

Existen infinidad de trabajos en los que, más o menos visceralmente, se defienden diferentes posturas a favor de SIP o de H.323. Ejemplos de estos trabajos pueden verse en [Schulzrinne y Rosenberg, 1998], [Dalgic y Fang, 1999], [Microtronix, 2003] y [Glasmann *et al.*, 2003]. Sin embargo, aunque las arquitecturas en las que se basan H.323 y SIP son sustancialmente distintas, a medida que se definen nuevas ampliaciones a SIP y aparecen nuevas versiones de H.323, ambos protocolos se diferencian menos en cuanto a las funciones y posibilidades que ofrecen.

Un factor común a las arquitecturas de conferencia basadas en ambos estándares es el protocolo de transporte utilizado para el intercambio de datos multimedia durante las sesiones que, aunque no está prefijado, en la práctica es RTP, sin que haya ninguna alternativa hasta la fecha. Como precondiciones generales para una arquitectura que debe dar soporte, al menos, a un sistema global de telefonía que coexista con y sustituya finalmente a la red telefónica tradicional se deben tener en cuenta las siguientes:

- Escalabilidad hasta un gran número de usuarios de cualquier parte del mundo, así como un gran número de llamadas activas de forma simultánea, del orden de millones.
- Permitir gestionar la red, de modo que sea posible aplicar políticas de control y tarificación.
- Proporcionar métodos de selección de calidad de servicio.

- Interoperabilidad entre diferentes fabricantes, protocolos y versiones de protocolos.
- Facilidad de ampliación.

A continuación se comparará, a grandes rasgos, las arquitecturas de control de sesiones multimedia H.323 y SIP, atendiendo a criterios de complejidad, ampliabilidad, escalabilidad, servicios ofrecidos y funcionamiento.

En primer lugar, la complejidad de la recomendación H.323 es muy superior a la del protocolo SIP. Mientras que la especificación base de la primera supera las 700 páginas, el protocolo SIP, junto con sus primeras ampliaciones y los protocolos de descripción de sesiones en los que se apoya, queda especificado en alrededor de 130 páginas. Las 37 cabeceras de SIP, todas con un reducido número de valores y parámetros, contrastan asimismo con los cientos de elementos definidos en la recomendación H.323. Dado que SIP es un protocolo de codificación textual, es posible realizar analizadores y generadores de mensajes mediante scripts en lenguajes como *Perl* o *Tcl*, facilitando su integración con los entornos web. Sin embargo, H.323 requiere generalmente el uso de herramientas específicas mucho más complejas. Asimismo, el establecimiento de llamadas H.323 supone la intervención de varios de los protocolos asociados, en un proceso en el que los elementos que intervienen deben mantener el estado de la sesión. Por el contrario, todas las operaciones del protocolo SIP se basan en el intercambio de peticiones y respuestas individuales. Este factor tiene una gran influencia en la escalabilidad de la arquitectura.

La ampliabilidad es un factor clave para el éxito de los protocolos de sesiones multimedia, dado que al regular servicios de uso muy extendido, están sujetos a futuras necesidades de ampliación procedentes de muy diversas fuentes y campos, al tiempo que es crucial que los dispositivos diseñados para distintas versiones y por distintos fabricantes sean compatibles entre sí en la ejecución de un conjunto mínimo de funciones. Los autores de SIP han tomado como precedente la evolución de dos protocolos muy extendidos: HTTP y SMTP. Este hecho les ha llevado a prever mecanismos para garantizar la compatibilidad entre diferentes versiones del protocolo, así como métodos de ampliación basados en el registro de nuevas características a través de la IANA. En la especificación de H.323 también se reservan algunos campos para futuras ampliaciones específicas del fabricante; sin embargo, éstas sólo se pueden realizar en aquellas partes donde se ha previsto de forma explícita. Además, no existen mecanismos para determinar qué ampliaciones admite un dispositivo, por lo que el uso de estas ampliaciones conlleva riesgos de incompatibilidad entre distintos fabricantes.

En cuanto al uso de formatos multimedia estandarizados, las sesiones controladas mediante SIP pueden emplear de forma estandarizada cualquier formato multimedia registrado a través de la ICANN [Freed y Borenstein, 1996a] y [Freed y Borenstein, 1996b], utilizándose para ello una descripción en formato SDP. En contraste con este procedimiento plenamente abierto, las sesiones H.323 sólo pueden emplear formatos registrados a través de una entidad central, la ITU-T, que hasta el presente sólo ha estandarizado formatos promovidos por la propia entidad, la mayoría de los cuales son de uso limitado por problemas de patentes. Con respecto a la modularidad del sistema, el protocolo SIP comprende funciones de señalización básica, localización y registro de usuarios. Otras funciones, como la garantía de calidad de servicio, la descripción de contenidos, los servicios de directorio y el control de la conferencia se delegan en otros protocolos. De este modo cada protocolo empleado en una arquitectura SIP es plenamente intercambiable, mientras que la arquitectura establecida por la recomendación H.323 define un conjunto de componentes acoplados que cubren un mayor rango de funciones.

Dado el gran número de sesiones que pueden llegar a procesar los componentes de la infraestructura telefónica de Internet, la simplicidad de la gestión de sesiones es clave. En este aspecto, SIP no requiere mantener estados en elementos intermedios y utiliza fundamentalmente conexiones UDP o de otros protocolos que requieren menos recursos que TCP. Por el contrario, H.323 requiere que los elementos intermedios mantengan el estado a lo largo de toda la sesión, y emplea UDP a partir de la versión 3, debiendo mantener la compatibilidad con versiones previas

basadas en conexiones TCP. Tanto H.323 como SIP admiten sesiones con múltiples participantes y transporte multicast. No obstante, a diferencia de la especificación de H.323, la especificación de SIP no define ninguna entidad central para coordinar las sesiones, estableciendo un modelo de control plenamente distribuido. Además, al basarse fundamentalmente en UDP, SIP puede distribuir mensajes en modo multicast. Por ello, a igualdad de recursos disponibles, el número de participantes en una sesión controlada mediante SIP puede llegar a ser mayor que en una sesión H.323. Por otra parte, la versión inicial de H.323 se diseñó para utilizarse en redes de área local, por lo que, aunque las últimas versiones han incorporado métodos de localización de usuarios, carece de algunos mecanismos, como prevención de bucles. Sin embargo, SIP se ha diseñado para redes de área extensa, por lo que cuenta con mecanismos flexibles de localización y registro de usuarios definidos con suficiente generalidad, como búsqueda a través de un número indefinido de servidores proxys, búsqueda en paralelo y ramificación.

A medida que SIP y H.323 han ido evolucionando de forma competitiva, los servicios que ambos ofrecen se van pareciendo cada vez más unos a otros. H.323 proporciona servicios de negociación de contenidos mucho más completos que los que pretende proporcionar SIP, permitiendo especificar complejas combinaciones de formatos multimedia admitidos por los dispositivos, junto con diversos parámetros de estos formatos. SIP prevé únicamente el intercambio de conjuntos de formatos admitidos. Sin embargo, los mecanismos adicionales proporcionados por H.323 rara vez son necesarios y en la práctica se usa sólo un subconjunto de ellos equivalente a los mecanismos de SIP.

Establecer una llamada con H.323 versión 1 ó 2 requiere establecer, previamente a la conexión H.323, conexiones TCP para H.225.0 y H.245, cuya misión se menciona en el apartado 3.2.6.1. Por el contrario, el establecimiento de una llamada SIP sólo requiere un intervalo de ida y vuelta de paquetes en el caso más común, puesto que se realiza mediante un mensaje transportado generalmente sobre UDP. A partir de la versión 3, H.323 permite iniciar llamadas mediante UDP. El establecimiento de llamadas es normalmente más rápido con SIP; sin embargo, H.323 puede solventar con mayor rapidez situaciones de error. No obstante, con la progresiva incorporación de SCTP como protocolo de transporte para señalización SIP, se espera mejorar las prestaciones en situaciones con un elevado porcentaje de errores en la transmisión. En ambos protocolos el proceso de finalización de llamadas se lleva a cabo mediante un único mensaje. En cuanto al control durante la llamada, tanto SIP como H.323 proporcionan servicios suplementarios que permiten modificar y comprobar el estado de una sesión durante su desarrollo, como desconexión provisional, identificación de llamadas, llamadas en espera o redirección y transferencia de llamadas.

Tras este repaso a las similitudes y diferencias más destacables de ambos protocolos, se procederá a describir someramente en los siguientes apartados la arquitectura en la que se basa cada uno de ellos, sin entrar en detalles como la semántica de los mensajes. A continuación se describe el protocolo de control H.248 (también conocido como Megaco), que complementa a H.323 y a SIP, y otros dos protocolos desarrollados por la IETF para anunciar y describir sesiones.

3.2.6.1. Recomendación H.323

Aunque no se ha mencionado hasta ahora, H.323 forma parte del conjunto o “paraguas” de recomendaciones de la ITU-T para videoconferencia (H.3xx). Aunque el protocolo H.323 es el único de los H.3xx aplicable a la transmisión multimedia por Internet cuando se carece de mecanismos de garantía de QoS entre los enlaces y no se puede desarrollar íntegramente la transmisión sobre una infraestructura de red específica, como ATM o RDSI. Los sistemas basados en señalización H.323 se articulan en torno a cuatro tipos de componentes: terminales, pasarelas, gatekeepers y unidades de control multipunto (MCU), cuya interacción se esquematiza en la figura 3.10.

- *Terminales*: son los clientes situados en redes IP. Se comunican bidireccionalmente y en

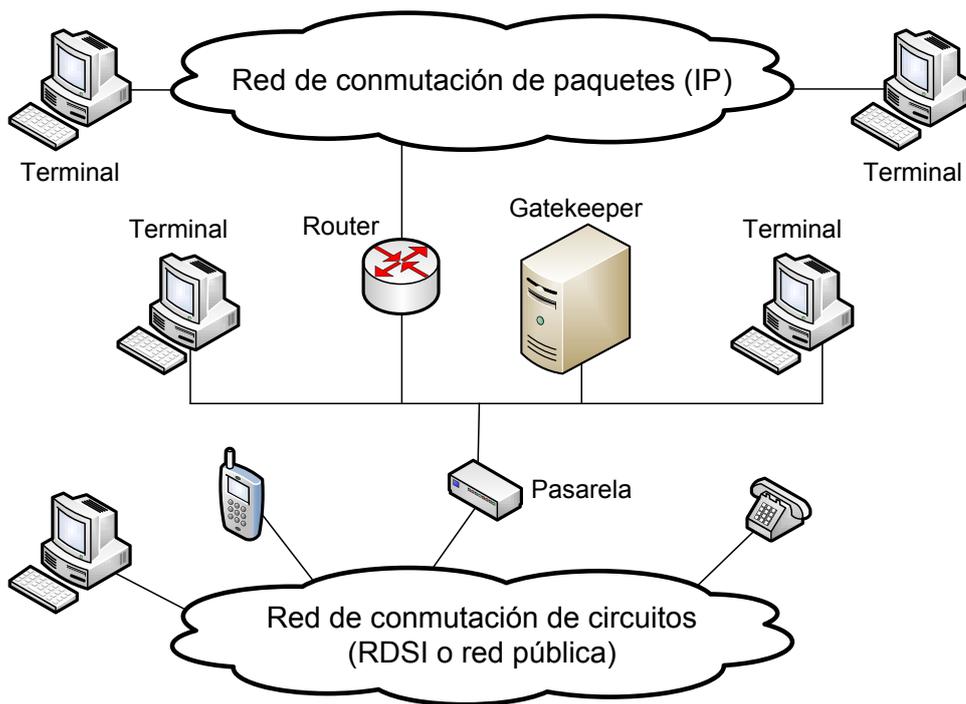


Figura 3.10: Elementos en una red H.323

tiempo real con el resto de elementos de los sistemas H.323. Deben ser capaces de comunicarse mediante los protocolos de control H.225 y H.245, así como de transmitir datos en tiempo real sobre RTP. Asimismo, pueden contener varios codificadores y decodificadores de formatos de sonido y vídeo, de entre los cuales, el formato de sonido G.711 (PCM) es obligatorio, con objeto de garantizar una mínima compatibilidad entre todos los terminales.

- *Pasarelas*: encargadas de conectar redes de conmutación de paquetes con redes de conmutación de circuitos, ya sean públicas o privadas. Deben ser capaces de establecer y controlar llamadas en ambos tipos de redes, traduciendo los distintos formatos y procedimientos de transmisión. Su presencia, del mismo modo que la conexión a redes de conmutación de circuitos, es opcional.
- *Gatekeepers*: su presencia en un sistema es también opcional; sin embargo, cuando están presentes deben desempeñar cuatro funciones: control de admisión, control de ancho de banda, traducción de direcciones alias o números de teléfono a direcciones de transporte y viceversa, y gestión de zonas. Cuando existe un gatekeeper en el sistema, todos los dispositivos deben registrarse y obtener permiso de él antes de realizar llamadas.
- *Unidades de control multipunto (MCU)*: permiten que se establezcan conferencias entre tres o más elementos. Generalmente constan de un controlador multipunto y un número variable de procesadores multipunto.

En terminología H.323, un canal es equivalente a una conexión de nivel de transporte. La recomendación H.323 especifica cuatro tipos de canales:

- Canal de señalización de llamada: por él circula la información relativa al control de llamadas

y servicios adicionales. El protocolo de control utilizado se especifica en las recomendaciones H.450 y H.225.0.

- Canal de control H.245: dedicado al control de la transferencia de información multimedia y la negociación de capacidades.
- Canal de control RAS: destinado a la comunicación entre los terminales y sus gatekeepers asociados. El protocolo RAS (registro, admisión y estado) está especificado en la norma H.225.0 y permite que los terminales se registren ante un gatekeeper y soliciten permiso para establecer llamadas.
- Canales lógicos para información multimedia: por donde circulan los datos de audio, vídeo y cualquier otro medio. Cada uno de ellos se transporta mediante una sesión RTP.

En las últimas versiones de H.323 todos los canales emplean protocolos de transporte no fiables (UDP), excepto el canal de control H.245, que requiere el uso de un protocolo fiable, como TCP. En la tabla 3.1 se muestran algunas de las recomendaciones más relevantes incluidas en la especificación H.323.

Tabla 3.1: Recomendaciones incluidas en la especificación H.323

Recomendación	Función
H.245	Funciones de control
H.225.0	Establecimiento y control de conexión
H.332	Gestión de conferencias de gran tamaño
H.235	Seguridad (privacidad, autenticación, etc.)
H.246	Interoperación con servicios basados en conmutación de circuitos
H.450.x	Servicios adicionales (transferencia y redirección de llamadas, etc.)
H.26x	Formatos de vídeo
G.7xx	Formatos de audio
T.120	Intercambio de datos (herramientas colaborativas)
X.680	Sintaxis ASN.1 empleada por los protocolos
X.691	Reglas de codificación del empaquetado (PER)

3.2.6.2. Session Initiation Protocol

La arquitectura de sesiones multimedia de la IETF, a diferencia de la arquitectura H.323, está constituida por un amplio conjunto de protocolos independientes, cada uno de los cuales cumple funciones complementarias. Esta arquitectura engloba dos protocolos de control específicos de Internet: SIP y SAP. SIP cumple las funciones de establecimiento de sesiones mediante invitación, así como de modificación y finalización de sesiones. SAP es un protocolo de anuncio de sesiones desarrollado para entornos multicast. Adicionalmente, se ha definido el protocolo Megaco para el control de las pasarelas o sistemas de interconexión entre redes de diferentes tipos, en especial entre redes de conmutación de paquetes, como Internet, y redes de conmutación de circuitos, como la red telefónica. Se analizarán más en detalle estos protocolos en posteriores apartados.

H.323 puede estar presente entre los protocolos de señalización en la arquitectura IETF, puesto que, al igual que cualquier otro protocolo que cumpla los requisitos de control de sistemas multimedia, puede sustituir a SIP en las funciones de señalización. Asimismo, es posible utilizar de forma combinada ambos protocolos, por ejemplo, estableciendo sesiones mediante el protocolo SIP y llevando a cabo el control de la sesión mediante los protocolos de la recomendación H.323.

La arquitectura de control de la IETF, basada en SIP, establece un modelo de sesiones descentralizado; es decir, no contempla, en principio, la existencia de un registro central para los participantes en cada sesión, sino que éstos comparten el concepto de sesión simplemente como una abstracción en común. SIP forma parte de una arquitectura integrada y estandarizada vinculada a otros protocolos de Internet, como el correo electrónico y HTTP. Cabe recordar que las funciones de garantía de calidad de servicio quedan delegadas en otros protocolos. Además, el protocolo de transporte en tiempo real preferente es RTP, junto con RTCP, que desempeña algunas de las funciones de control de sesiones incluidas en la recomendación H.323. Tanto SIP como RTP son independientes del protocolo de transporte subyacente. Todas estas funciones quedan completadas con los siguientes servicios:

- Servicios de directorio, delegados en el protocolo LDAP (*Lightweight Directory Access Protocol*) [Hodges y Morgan, 2002].
- Búsqueda de pasarelas, servicios proporcionados por el protocolo TRIP [Rosenberg *et al.*, 2002a], que no depende de ningún protocolo de señalización en concreto, por lo que se puede utilizar tanto con SIP como con H.323.
- Autenticación, autorización y tarificación, que pueden desempeñar los protocolos RADIUS [Rigney *et al.*, 2000] y Diameter [Calhoun *et al.*, 2003].

El protocolo SIP, cuya operación se basa en mensajes de petición y respuesta, reutiliza muchas de las reglas de codificación, códigos de error y campos de cabecera de HTTP. Las funciones de control de llamadas (redirección, transferencia, cambio de formatos y codificación, etc.) que proporciona están integradas, por tanto, con la infraestructura web del mismo modo que los sistemas de programación que utilizan interfaces CGI. El uso de tipos MIME para describir los contenidos tratados por los mensajes SIP hace posible, por ejemplo, devolver cualquier tipo de contenido web ante un mensaje de inicio de llamada. Las sesiones multimedia controladas por SIP pueden constar de varias sesiones RTP sin que todos los participantes tengan por qué participar en todas las sesiones RTP. SIP es un protocolo cliente-servidor de señalización al que se pueden añadir nuevos métodos y capacidades, cuyo registro se debe realizar a través de la IANA. Igualmente, define respuestas de rechazo ante elementos desconocidos, de modo que aplicaciones con diferentes grados de implementación y versiones de SIP puedan interoperar en la medida de lo posible.

Los dos tipos de elementos presentes en la arquitectura SIP son los agentes de usuario (UA) y los servidores. Los agentes de usuario constan, a su vez, de dos componentes: los agentes de usuario clientes (UAC) y los agentes de usuario servidores (UAS). Ambos se encuentran en todos los agentes de usuario, permitiendo la comunicación entre diferentes agentes de usuario mediante peticiones y respuestas de tipo cliente-servidor. Los UAC tienen como misión el envío de peticiones SIP, mientras que los UAS están encargados de atender tales peticiones y remitir las correspondientes respuestas.

Los servidores SIP pueden ser de tres tipos no excluyentes: servidores de redirección, de registro y proxys. Aunque una llamada básica se puede realizar con SIP sin que intervengan servidores, las funciones avanzadas del protocolo no se pueden llevar a cabo sin su participación. Las funciones básicas de los servidores SIP son la localización de usuarios y la resolución de nombres. Puesto que generalmente los agentes de usuario clientes no conocen la dirección IP del destinatario de una llamada, sino su nombre (en forma de dirección de correo electrónico) o un número de

teléfono, necesitan enviar en primer lugar al servidor correspondiente un mensaje de invitación al nombre o número para que localice al destinatario. El servidor puede conocer la dirección del destinatario o recurrir a otros servidores para continuar la búsqueda. Cuando las llamadas se redirigen, la ruta seguida se registra en los mensajes SIP, de modo que a la hora de generar respuestas se pueda conocer el camino de retorno hasta el origen del mensaje inicial. Con objeto de proporcionar la mayor movilidad posible a los usuarios a través de diferentes terminales, los servidores SIP pueden emplear cualquier medio para localizar agentes de usuario, ya sea el sistema DNS, ejecutando programas auxiliares o accediendo a bases de datos. En algunos casos, un servidor puede determinar que es posible acceder a un destinatario a través de varios servidores, pudiendo bifurcar la búsqueda, con la posibilidad de unificar las múltiples respuestas recibidas antes de devolverlas al agente que inició la llamada. Los servidores pueden mantener el estado de las llamadas a dos niveles de detalle o no mantener estado alguno. Al igual que en la red telefónica convencional, pueden mantener el estado completo de cada llamada; sin embargo, este comportamiento es opcional, puesto que limitaría la escalabilidad de los sistemas SIP. El modo de funcionamiento recomendado y más frecuente es que los servidores mantengan únicamente el estado de cada transacción por separado. Una *transacción* SIP se puede definir como el conjunto de peticiones y respuestas intercambiadas desde que un agente de usuario cliente envía una petición hasta que recibe una respuesta definitiva. Por este motivo, si los servidores sólo mantienen el estado de cada transacción por separado, no tienen conocimiento de las llamadas existentes en un cierto instante, sino sólo de cada una de las peticiones y respuestas asociadas que la componen. De este modo, los servidores no tienen que mantener el estado de cada una de las llamadas, constituyendo así un sistema altamente escalable. Asimismo, el comportamiento de los servidores SIP en cuanto al mantenimiento de estados se puede modificar de forma dinámica en función de las circunstancias. En la figura 3.11 se esquematiza la relación entre los componentes descritos.

En la figura 3.12 se describe el establecimiento de una sesión, en este caso una audioconferencia, entre dos interlocutores utilizando SIP. Inicialmente, **Equipo1** realiza una invitación a **Equipo2** a través de su proxy local, el cual la redirige al proxy que corresponde al destinatario, y éste a su vez a este último. El terminal **Equipo2** notificará a **Equipo1** que está avisando al usuario de la llamada entrante. En el momento que éste conteste se establecerá la sesión de audio entre ambos terminales de forma directa. Más ejemplos de llamada pueden encontrarse en [Johnston *et al.*, 2003].

En los sistemas SIP, tanto los mensajes de señalización como la información multimedia intercambiada en una sesión pueden transmitirse en unicast o multicast. El caso correspondiente a señalización e información multimedia transmitida en multicast es equivalente a las sesiones iniciadas por medio del protocolo SAP, que serán descritas más adelante. En la práctica, SIP se ha usado típicamente para llamadas entre dos participantes. El uso generalizado de SIP para conferencias con cualquier número de participantes presenta dificultades adicionales que requieren un esquema general para llamadas con múltiples participantes [Rosenberg, 2006].

Generalmente un servidor SIP realiza de forma simultánea más de una función de las tres indicadas anteriormente. Gracias a una infraestructura de servidores SIP, es posible gestionar las llamadas de forma distribuida entre equipos personales, equipos de proveedores de servicios y pasarelas corporativas, con la consiguiente flexibilidad y control por parte del usuario, que puede mantener la privacidad de sus datos personales en todo momento. Asimismo, un mensaje SIP puede pasar por un número indeterminado de servidores desde que un agente de usuario cliente lo envía hasta que llega al agente de usuario servidor destinatario.

Los servidores de registro registran los diferentes dispositivos en el momento en el se conectan, de modo que puedan proporcionar las direcciones de éstos cuando se reciban invitaciones dirigidas a ellos. Para ello, aceptan y procesan peticiones de registro. Se contempla la existencia de un tiempo máximo de validez de cada registro, definible por el servidor, tras el cual se debe renovar el registro. Asimismo, existen mecanismos para cancelar todos los registros contenidos en un servidor de registro.

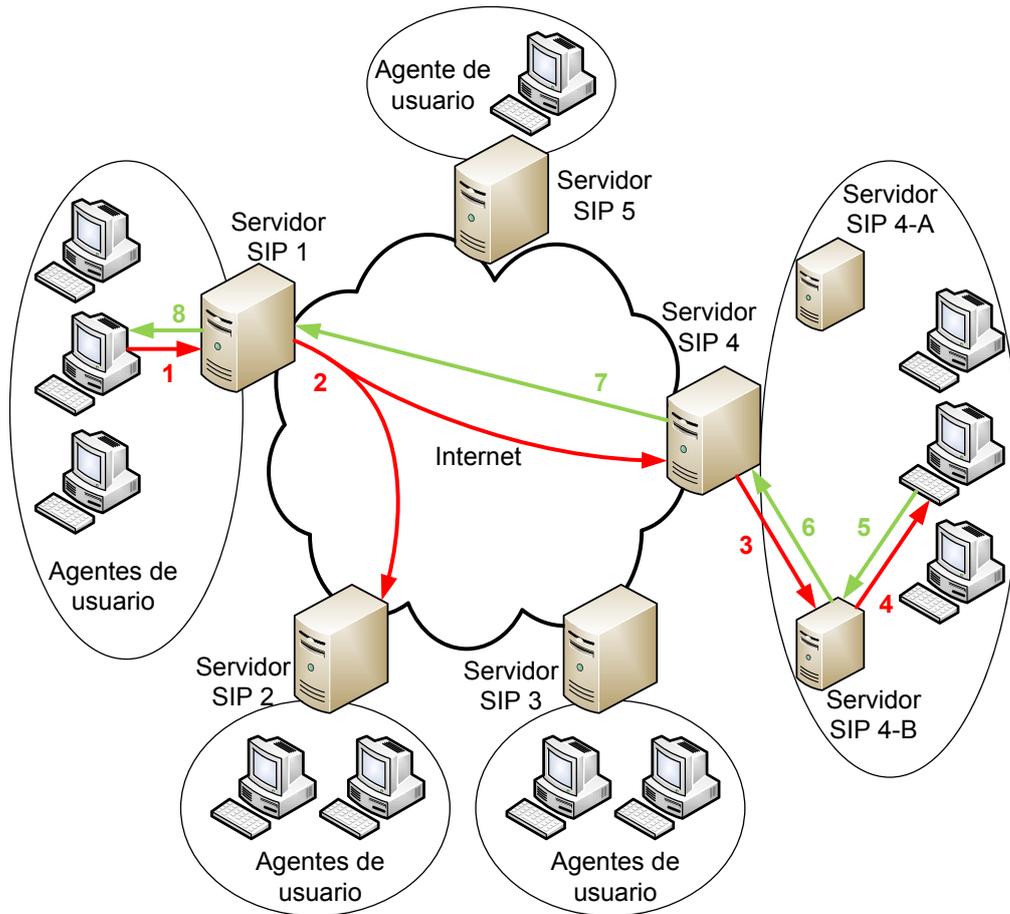


Figura 3.11: Elementos en una red SIP

La función de los proxys es similar a la de los proxys HTTP o los agentes de transferencia de mensajes SMTP: reciben solicitudes y deciden a qué otro servidor las deben remitir, alterando además algunos campos de la solicitud. Por tanto, actúan como intermediarios en las transacciones que procesan. En la figura 3.13 se esquematiza el funcionamiento como proxy de los servidores SIP. El elemento al que el proxy reenvía la petición puede ser tanto otro servidor proxy, como un servidor de redirección o un agente de usuario servidor, que es el caso ilustrado. Los proxys suelen actuar también como servidores de registro para todos los dispositivos que representan.

El servidor proxy puede recurrir a un servidor de localización para determinar la dirección en la que actualmente está disponible el destinatario, `dest@adhoc.org`, correspondiente a la dirección original, `dest@habitual.org`, indicada por el UAC que inició la llamada. Los proxys son, al igual que en el caso del protocolo HTTP, particularmente útiles como representantes de salida/entrada de/a redes corporativas, proporcionando servicios de búsqueda de direcciones, control de cortafuegos y gestión de políticas corporativas. Asimismo, pueden cumplir funciones de control de salida a pasarelas para redes telefónicas tradicionales.

Los servidores de redirección, a diferencia de los proxys, no inician transacciones, sino que, cuando reciben solicitudes desde un agente de usuario cliente, remiten al mismo agente un mensaje indicando el o los servidores con los que debe ponerse en contacto, en un procedimiento similar al de búsqueda iterativa del sistema DNS. A diferencia de los agentes de usuario servidores, no

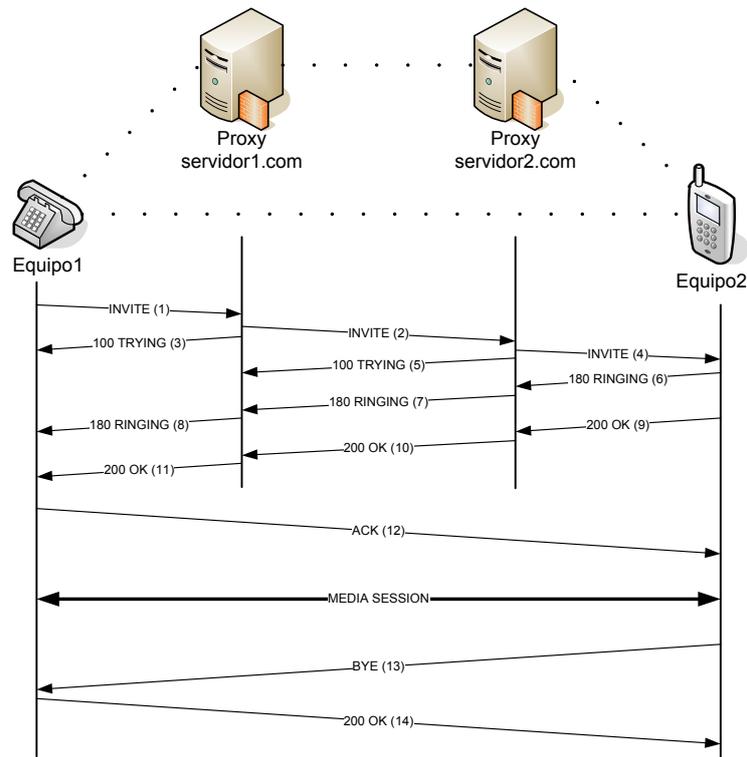


Figura 3.12: Establecimiento de una sesión con SIP

aceptan llamadas. Normalmente, los servidores de redirección gestionan un mayor número de mensajes que los proxies, pero con menores necesidades de procesamiento.

El protocolo SIP, gracias a sus mecanismos de redirección y su esquema de nombres, puede proporcionar servicios de movilidad en cuatro aspectos:

- Movilidad personal: una misma dirección personal puede asociarse a diferentes dispositivos, como teléfonos móviles, ordenadores portátiles y de bolsillo, teléfonos Ethernet (VoIP), de modo que la búsqueda del usuario proporcione la dirección de un dispositivo u otro en función de las circunstancias.
- Movilidad de servicios: posibilidad de tener acceso a servicios como libros de direcciones, funciones especiales o preferencias desde distintas localizaciones.
- Sesiones: traslado de sesiones activas entre distintos terminales a medida que el usuario se traslada, por ejemplo, por un edificio.
- Movilidad de los terminales: poder utilizar un terminal mientras que el usuario se conecta a través de diferentes subredes. Este tipo de movilidad debe ser admitido por protocolos de capas inferiores, como IEEE 802.11 o IP móvil, y requiere en principio el uso de protocolos adicionales como DHCP y DNS dinámico.

Cabe destacar que, del mismo modo que SIP ha tomado muchos de sus elementos de la infraestructura web existente, gran parte de los nuevos servicios web en desarrollo en la actualidad son asimilables a los servicios de movilidad contemplados en la arquitectura de los sistemas SIP.

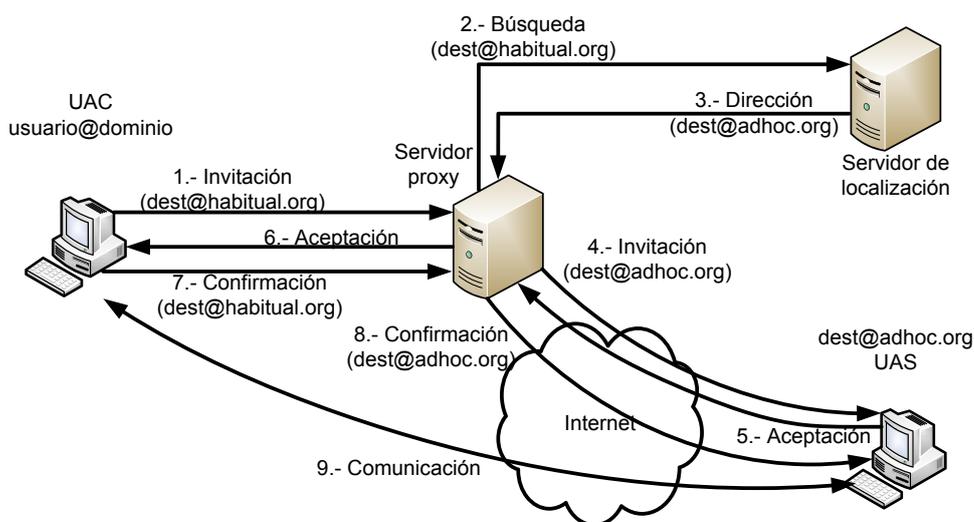


Figura 3.13: Funcionamiento de un proxy SIP

Para garantizar la privacidad de los participantes en sesiones, SIP admite varios modos de cifrado, con mensajes de señalización basados en criptografía de clave pública. Existe un lenguaje de programación específico para servicios de telefonía basados en SIP denominado CPL (*Call Processing Language*) [Lennox *et al.*, 2004]. CPL es un lenguaje basado en XML diseñado para permitir que los usuarios creen de forma sencilla, mediante interfaces gráficas, servicios de telefonía, generando programas que se ejecutan en servidores SIP. Aunque su desarrollo está ligado al de SIP, es un lenguaje independiente del protocolo de señalización utilizado.

Cada llamada SIP tiene asociado un código único que identifica a todos los participantes en una sesión multimedia invitados por un mismo agente. Una llamada puede dar lugar a una rama de llamadas, identificada por la combinación del código de la llamada, el origen y el destinatario. Cada llamada dentro de una rama se distingue mediante un número de secuencia. Con objeto de que los proxys SIP sólo tengan que mantener el estado de cada respuesta individual, en lugar de la conferencia por completo, cada mensaje de señalización contiene el identificador de llamada. Durante la fase de iniciación de llamadas, se realizan varias funciones que se pueden agrupar en dos bloques:

- Localización y determinación de disponibilidad del destinatario, y selección de terminal [Rosenberg y Schulzrinne, 2002b]: las respuestas de los servidores SIP pueden contener una lista con la descripción de los distintos terminales a los que tiene acceso un destinatario, especificando asimismo las posibilidades y direcciones de cada uno de ellos.
- Negociación de tipos de datos y de su codificación: el mensaje de invitación de SIP contiene una lista de los tipos de datos y codificaciones que se pretende utilizar en una sesión. En el caso de sesiones de dos participantes, el destinatario de la invitación proporciona en su respuesta un subconjunto de la lista de tipos original, indicando los tipos que está dispuesto a utilizar, por lo que la negociación de contenidos se realiza con la mayor rapidez posible. Este modelo de negociación, conocido como oferta/respuesta se ha definido por el momento para SDP, aunque es generalizable para otros formatos de descripción.

3.2.6.3. Megaco / H.248

El protocolo Megaco/H.248 [Cuervo *et al.*, 2000], resultado de la colaboración entre la IETF y la ITU-T, se emplea para controlar las pasarelas, indistintamente si éstas conectan una red de conmutación de circuitos, como la red de telefonía básica, con una red de conmutación de paquetes basada en IP, o dos redes de conmutación de paquetes, por ejemplo una basada en SIP y otra en H.323. El protocolo es comúnmente conocido en cualquiera de sus dos denominaciones: H.248, según la ITU-T, y Megaco, nombre dado por la IETF.

Las pasarelas están formadas por un controlador de pasarela de medios *Media Gateway Controller* (MGC) y una o varias pasarelas de medios *Media Gateway* (MG). Las MG actúan únicamente como pasarelas de los medios transmitidos, realizando sobre los mismos las conversiones o adaptaciones que sean necesarias. Las MGC realizan el resto de funciones de una pasarela, como las de conversión de la señalización, y además son las encargadas de controlar el funcionamiento de las MG. Todos estos elementos pueden estar presentes en un mismo dispositivo o repartidos físicamente entre varios, aunque también existen pasarelas software. Megaco/H.248 es un complemento a los protocolos H.323 y SIP. El mismo protocolo de señalización al que complementa es el empleado por los MGC para comunicarse entre sí, de manera análoga a cualquier otro par de dispositivos de la red.

3.2.6.4. Session Announcement Protocol

La especificación del protocolo SAP [Handley *et al.*, 2000] establece las reglas que los directorios de sesiones multicast deben seguir en la transmisión periódica de anuncios de sesiones, tales como el límite de ancho de banda, el algoritmo a seguir para el cálculo del intervalo de transmisión, las direcciones y el puerto de destino de los anuncios, las condiciones bajo las cuales se considera que las sesiones anunciadas dejan de estar activas, etc. Puesto que SAP es un protocolo diseñado para el anuncio de sesiones predefinidas, a diferencia de SIP, no contempla ningún proceso de negociación.

Los anuncios SAP contienen la descripción de sesiones junto con una cabecera de autenticación, que debe estar siempre presente salvo en circunstancias excepcionales. El campo indicador del tipo de datos de la descripción puede contener el tipo MIME de la descripción transportada en el anuncio. Si la descripción tiene formato SDP, el uso de este último campo es opcional. La cabecera de autenticación tiene dos funciones: autenticar al creador de la sesión descrita en un anuncio y verificar la misma identidad cuando se solicite la modificación o suspensión de una sesión. SAP admite el uso de formatos distintos a SDP para describir sesiones, aunque se desaconseja con el fin de asegurar la mayor interoperabilidad posible; por otra parte es obligatorio que todo agente SAP manipule correctamente descripciones en formato SDP.

3.2.6.5. Session Description Protocol

El protocolo SDP, especificado en [Handley y Jacobson, 1998], es un lenguaje de descripción diseñado exclusivamente para describir sesiones multimedia en tiempo real, siendo útil para invitaciones, anuncios, y cualquier otra forma de inicio de sesiones. En la actualidad, es el protocolo de descripción más extendido para anunciar, y también negociar, las características de las sesiones multimedia en Internet. En este ámbito, por descripción se entiende proporcionar en formato estándar la información necesaria para que posibles participantes (usuarios y aplicaciones) puedan tener conocimiento de una sesión y unirse a ella. SDP describe las capacidades (de recepción, captura, reproducción. . .) de los sistemas y proporciona varias alternativas de configuración a elegir. Puesto que SDP no es más que un formato de descripción, las descripciones SDP se pueden transportar mediante distintos protocolos, como SAP, SIP, RTSP, HTTP y correo electrónico. Aunque SDP ha sido desarrollado por la IETF, la ITU-T lo ha adoptado como protocolo de

descripción de sesiones para muchas de sus recomendaciones. Asimismo, el lenguaje SMIL del W3C utiliza descripciones SDP.

Como ya se ha mencionado, SDP se emplea actualmente para desempeñar dos funciones: descripción de sesiones y negociación de capacidades, aunque no fue diseñado con esta última tarea en mente. Las especificaciones de los protocolos SIP, RTSP, SAP y Megaco designan a SDP como formato de descripción preferente para sus funciones de especificación y negociación de contenidos y formatos multimedia. Cada descripción SDP proporciona la siguiente información:

- Nombre y propósito de la sesión.
- Tipo y formato de los medios que la componen.
- Intervalo o intervalos de tiempo en los que se desarrolla la sesión.
- Parámetros necesarios para poder recibir e interpretar los datos: direcciones, puertos, formatos, etc.
- Información complementaria relativa a los recursos de red necesarios para participar en la sesión: ancho de banda e información de contacto del responsable de la sesión.

La codificación de las descripciones, al igual que en el protocolo SIP, es textual. Es posible establecer sesiones de ámbito privado cifrando la descripción SDP, proceso que depende del mecanismo utilizado para transportar la descripción. También es posible anunciar una sesión en privado, incluyendo entonces las claves y la información acerca del esquema de cifrado en el propio anuncio.

Como la especificación original de SDP no contempla la función de negociación de capacidades, no se trata del mecanismo más adecuado para conferencias que se puedan iniciar de forma espontánea, como una llamada telefónica. La capacidad de negociar y seleccionar las características de cada sesión multimedia es una función básica de cualquier arquitectura de control por lo que se requiere una solución específica para este problema. Para cubrir esta carencia se han definido ampliaciones a SDP [Rosenberg y Schulzrinne, 2002a] y [Andreasen, 2002], que presentan aún alguna limitación pues deben mantener la compatibilidad con la especificación original.

El empleo de tipos MIME en sistemas de conferencia multimedia está relacionado con SDP de dos formas:

- Indicación del formato de las descripciones de sesiones: las descripciones SDP pueden distribuirse mediante correo electrónico o transacciones HTTP; el tipo MIME asignado a estas descripciones es `application/sdp`, al que conviene asociar aplicaciones que interpreten la descripción y permitan lanzar, con los parámetros adecuados, las aplicaciones necesarias para participar en la sesión. Este mismo tipo MIME es el utilizado por SIP, Megaco y SAP para indicar que SDP es el formato de descripción de sesiones utilizado.
- Especificación de formatos multimedia en descripciones de sesiones: las descripciones SDP utilizan tipos MIME para identificar los formatos empleados en una sesión multimedia. Existe una especificación [Casner y Hoschka, 2003] del procedimiento a seguir para registrar formatos RTP como tipos MIME, donde además se registran los formatos definidos en el perfil para conferencias de audio y vídeo sobre RTP.

3.3. Procesamiento de la información multimedia

Dada la cantidad de medios que maneja una herramienta de e-learning síncrono es necesario considerar cuáles son las alternativas tecnológicas de que se dispone para trabajar con cada uno de ellos. En los siguientes apartados se valoran las diferentes alternativas en función del medio a procesar.

3.3.1. Audio y vídeo

Habitualmente, las señales multimedia que se intercambian durante una sesión de e-learning síncrono requieren un ancho de banda tal que hace imposible su transmisión por la red sin comprimir. Esto es especialmente notable cuando se trabaja con audio y vídeo. Por ejemplo, para el caso de la voz humana muestreada con una frecuencia de 8 kHz, con 16 bits por muestra y un único canal, la tasa de bits por segundo que implica es la indicada en 3.1.

$$16 \text{ bits} \times 1 \text{ canal} \times 8000 \text{ hercios} = 128000 \text{ bps} \quad (3.1)$$

Para el caso de una señal de vídeo RGB con un tamaño de imagen de 768x576 y 25 cuadros por segundo necesita el un ancho de banda indicado en 3.2.

$$768 \text{ columnas} \times 576 \text{ filas} \times 24 \text{ bits} \times 25 \text{ cuadros/s} = 256420800 \text{ bps} \quad (3.2)$$

Debido a esto, se hace necesario una etapa de codificación que reduzca el ancho de banda necesario. Los codecs que se utilizan en las soluciones síncronas, debido a su orientación al tiempo real, interesa que presenten las siguientes características:

- Baja latencia: de esta forma se añade la menor latencia posible al procesamiento y transmisión de los datos, tanto para codificarlos como para decodificarlos, para así cumplir los requerimientos de tiempo real.
- Alta compresión: así son necesarios menos recursos en la red y es posible que un mayor número de usuarios participen en una conferencia multipunto de forma simultánea.
- Robusto frente a pérdidas: debido a las características de la red sobre la que se transportan los datos, vistas en el apartado 3.2.1, los codecs utilizados deben estar preparados para decodificar una señal en la que falten paquetes, es decir, no sea continua la señal.

Dentro de las arquitecturas Windows, el estándar de facto para la captura y posterior codificación de audio y vídeo es DirectShow, aunque existen otras opciones como la API Win32 *Video for Windows* o librerías de terceros que se construyen a partir de éstas. DirectShow organiza los flujos multimedia como secuencias de filtros interconectados formando un grafo, que realizan sucesivas operaciones sobre el flujo de datos. Dependiendo del tipo de operación que realicen, existen tres tipos de filtros:

- Filtros de fuente: se conectan al principio de la cadena de filtros y son el origen de los datos que forman el flujo y sobre los que el resto de filtros actúan.
- Filtros de transformación: realizan alguna operación sobre los datos que les llegan y transforman en flujo, por ejemplo, lo codifican en un determinado formato. Pueden encadenarse varios filtros de transformación dentro del mismo grafo DirectShow.
- Filtros sumidero: constituyen el final del flujo de datos y se conectan al final de la cadena de filtros. Filtros de este tipo sirven para renderizar el flujo, volcarlo a disco, a la red, etc.

Recientemente, Microsoft ha liberado el SDK *Media Foundation*, cuyo objetivo es reemplazar a DirectShow y las otras API y librerías para el procesamiento de medios digitales a partir de Windows Vista.

3.3.1.1. Audioconferencia

La audioconferencia es una de las funcionalidades más importantes que ofrecen las herramientas de e-learning síncrono. Consiste en la transmisión del discurso oral de un usuario a uno o varios receptores. Por ello, es necesario capturar en tiempo real la señal acústica para, a continuación, realizar su codificación.

Varios son los parámetros que deben configurarse antes de iniciar el proceso de captura de audio, independientemente de la técnica utilizada:

- Frecuencia de muestreo: es el número de muestras que se obtienen a partir de la señal analógica. El número mínimo de muestras necesario viene determinado por el *Criterio de Nyquist*. Este criterio implica que una señal continua puede ser reconstruida totalmente a partir de muestras si la frecuencia de muestreo es, al menos, el doble de su ancho de banda.

Por tanto, para una señal de audio que transporte la voz humana, es decir, comprendida en un rango entre 0 y 4000 Hz, necesita una frecuencia de muestreo de 8000 Hz para que pueda ser recuperada. Otras frecuencias de muestreo habituales son 11025 Hz, 22050 Hz, 44100 Hz y 48000 Hz; aunque para la audioconferencia se utilizan 8000 Hz prácticamente en todas las ocasiones.

- Número de canales: un flujo de audio puede contener una o más señales independientes. Así, tendremos audio *mono* (un canal), *estéreo* (dos canales) o multicanal. En VoIP prácticamente se utiliza un sólo canal en todas las soluciones.
- Tamaño de la muestra: representa la resolución de cada una de las muestras. Valores típicos son 8 y 16 bits, aunque casi siempre se utilizan 16 bits por la baja resolución que se obtiene con 8.
- Periodo de captura: los dispositivos de captura de audio suelen implementar un buffer de captura donde almacenan las muestras capturadas. El periodo de captura indica el tiempo durante el cual el dispositivo debe capturar muestras de audio, que debe almacenar en el buffer, antes de pasárselas a la aplicación. De esta forma, este tiempo también determina la latencia mínima que tiene la primera muestra capturada, ya que, aunque ya esté almacenada en el buffer de captura, no llega a la aplicación hasta que se captura la última muestra del periodo. Por tanto, interesa que este periodo de captura sea lo más bajo posible. En VoIP, son comunes valores de 20 ó 30 milisegundos.

Dependiendo del tipo de audio a codificar se utilizan unas u otras técnicas. Es muy común distinguir entre audio de banda ancha, por ejemplo música, o banda baja, representado por la voz humana. En la audioconferencia únicamente se transmite voz, luego sólo se tienen en cuenta las técnicas que se utilizan en la codificación de la voz humana:

- PCM (*Pulse Code Modulation*): es un proceso de modulación para convertir una señal de audio analógica en una secuencia de bits. Un ejemplo de implementación es G.711.
- ADPCM (*Adaptive Differential Pulse Code Modulation*): en lugar de cuantificar la señal de audio directamente como hacen los codecs PCM, se cuantifica la diferencia entre la señal y una predicción de la misma. Un estándar que sigue esta técnica es G.726.
- CELP (*Code-Excited Linear Prediction*): es una técnica de codificación de la voz humana con numerosas variantes que consiste en buscar la codificación más adecuada optimizando la señal decodificada como realimentación. Ejemplos de codecs que utilizan esta técnicas son G.723.1, G.728, G.729 y Speex.

- LPC (*Linear predictive coding*): se utiliza para procesar señales de audio y codificar la voz humana. Consiste en utilizar un modelo predictivo para aproximar la señal. Un codec que implementa esta técnica es iLBC.

Estas técnicas constituyen algoritmos de referencia que sirven de orientación para que implementaciones concretas de codecs realicen la codificación y la decodificación siguiendo estos algoritmos. En particular, los codecs que se utilizan de forma habitual en VoIP son los que aparecen en la tabla 3.2. En dicha tabla se enumeran los codecs junto con características como la frecuencia de muestreo que debe tener el audio a codificar, la tasa de bits de los datos ya codificados, el tamaño del paquete tanto en tiempo como en bytes, y la ancho de banda nominal. Esta última es el resultado de añadir a cada paquete la sobrecarga que introducen las capas RTP (12 bytes), UDP (8 bytes), IP (20 bytes) y Ethernet (18 bytes).

Tabla 3.2: Codecs disponibles para la audioconferencia

Codec	Frec. muestreo (kHz)	Bitrate (kbps)	Duración paquete (ms)	Tam. paquete (bytes)	A.B. nominal (kbps)
G.711	8	64	20	160	87,2
G.723.1	8	5,3	30	20	20,8
	8	6,4	30	24	21,9
G.726	8	16	–	–	–
	8	24	20	60	47,2
	8	32	20	80	55,2
	8	40	–	–	–
G.728	8	16	30	60	31,5
G.729	8	8	20	20	31,2
GSM	8	12,2	20	–	23,2
	8	13	20	33	36,4
iLBC	8	13,33	30	50	28,8
	8	15,20	20	30	38,4
Speex	8	–	20	–	–
	16	–	20	–	–
	32	–	20	–	–

G.711, también conocido por audio PCM, es un estándar para la compresión de audio telefónico propuesto por la ITU-T. Se utiliza para representar señales de audio con frecuencias de la voz humana, es decir, con una frecuencia de muestreo de 8 000 Hz. Existen dos versiones del algoritmo: μ -law, utilizado en América del Norte y Japón, y A-law, utilizado en el resto del mundo.

G.723.1 es un algoritmo de codificación de voz estandarizado por la ITU-T. Tiene dos modos de trabajo, a 5,3 y 6,4 kbps, utilizando como entrada de datos audio PCM con 16 bits por muestra. Consigue uno de los más altos ratios de compresión de todos los estándares de la ITU-T.

Por su parte, G.726 trabaja con tasas de bits de 16, 24, 32 y 40 kbps. Es un codec que puede ser utilizado para la compresión de la voz humana en aplicaciones de telefonía o para su almacenamiento.

G.728 es un codec de codificación de voz, también de la ITU-T, que destaca por su bajo retardo de codificación. Ha ganado popularidad en aplicaciones de vídeo, satélite y móviles dado que obtiene calidad de voz parecida a G.726 a 32 kbps utilizando la mitad de ancho de banda.

G.729 es otro de los algoritmos de compresión propuestos por la ITU-T. Ofrece alta calidad y compresión a costa de una mayor complejidad que el resto. Se basa en la técnica CELP y recibe como entrada muestras de 16 bits de audio PCM.

GSM es un codec que se utiliza en telefonía celular y presenta gran número de variantes. Proporciona buena calidad de voz, pero no tan buena como G.728. Su mayor ventaja es la simplicidad del algoritmo de codificación que utiliza, que hace que los recursos que necesite sean mínimos.

iLBC [Andersen *et al.*, 2004] es un codec de audio especialmente indicado para VoIP por la robustez que exhibe, degradando de una forma controlada la calidad de la voz ante la pérdida de paquetes.

Speex es un codec de código abierto que utiliza la técnica CELP para codificar la señal de audio. Tiene tres modos de funcionamiento: banda estrecha (8 kHz), banda ancha (16 kHz) y banda ultra-ancha (32 kHz), permitiendo cambiar el modo de funcionamiento dentro del mismo flujo de bits. Por esta razón, la tasa de bits que genera no es constante.

3.3.1.2. Videoconferencia

La videoconferencia consiste en el transporte de la señal de vídeo a uno o varios destinatarios. Dependiendo del número de participantes a los que se les permite el envío de su flujo de vídeo, la videoconferencia puede ser punto a punto o multipunto. Es habitual que las herramientas de e-learning síncrono ofrezcan esta funcionalidad, al menos, para transmitir el vídeo del instructor al resto de participantes.

Para la captura del vídeo se deben tener en cuenta los siguientes parámetros:

- Resolución: que representa el tamaño (ancho por alto) de los cuadros a capturar. Cuanto mayor sea la resolución, mayor será el número de bits necesario para representar cada cuadro. Una resolución típica cuando se trabaja en videoconferencia es 320x240.
- Formato de imagen: determina el formato, e implícitamente el tamaño en bytes, de cada cuadro a capturar. Es posible utilizar diferentes formatos para los cuadros, entre los que se encuentran los formatos RGB (de componentes de color) o YUV, que son representaciones comprimidas aprovechando la mayor sensibilidad del ojo humano frente a la luz (luminancia) que frente al color (crominancia).
- Cuadros por segundo: es la tasa de cuadros que se captura por unidad de tiempo. Cuantos más cuadros se capturen mayor resolución temporal tendrá el vídeo al tiempo que mayor ancho de banda consumirá.

Dejando a un lado los codecs de vídeo sin pérdidas, principalmente orientados a la edición de vídeo, todos los codecs que se utilizan en videoconferencia degradan la señal original en mayor o menor medida. Esto es debido a que así, se consiguen unos ratios de compresión más altos. Algunos de estos codecs son:

- MJPEG (Motion JPEG) es una forma de codificar el vídeo en la que cada cuadro se comprime por separado como una imagen JPEG. Se utiliza comúnmente en videocámaras IP, aunque no es el más adecuado para las soluciones de e-learning síncrono por el alto ancho de banda que consume.
- MPEG-4 es un grupo de estándares de codificación de audio y vídeo. Al igual que otras versiones anteriores, MPEG-2 y MPEG-1, organiza las funcionalidades que ofrecen las implementaciones del estándar alrededor del concepto de perfil, que permite definir conjuntos específicos de capacidades que pueden ser implementados para cumplir con objetivos particulares. En concreto, la parte 2 del estándar propone un codec de compresión de elementos

visuales, que define el perfil ASP (*Advanced Simple Profile*), que en ocasiones se utiliza en videoconferencia.

- H.261 es un estándar de codificación de vídeo de la ITU-T destinado a la transmisión de este tipo de señales a través de redes RDSI en las cuales el ancho de banda disponible es múltiplo de 64 kbps. Tiene varios modos de operación con unas tasas de bits entre 40 kbps y 2 Mbps. De igual manera, admite dos tamaños de cuadro: 352x288 y 176x144.
- El estándar H.263 apareció como una evolución del anterior para entornos de bajo ancho de banda, además de incorporar algunas mejoras respecto al anterior. Soporta cinco resoluciones, tres más que H.261 y, en general, mejora el ratio de compresión significativamente.
- H.264, también conocido por MPEG-4 Parte 10 AVC, es el resultado del trabajo conjunto de el grupo MPEG y la ITU-T. Supone una notable mejora respecto a su antecesor H.263. Está preparado para trabajar en entornos de muy bajo ancho de banda con una calidad aceptable. También es válido para situaciones en las que se necesita una mayor resolución, de ahí que también se utilice en discos Blue-ray para almacenar películas de alta definición. La mayor parte de los sistemas de videoconferencia incorporan H.264 junto a H.263 y H.261.
- VC-1 es un codec de vídeo que pretende ser la alternativa a H.264. Su principal objetivo es dar soporte a la compresión de contenido entrelazado sin tener que convertirlo primero a progresivo. En algunas situaciones, compite directamente con H.264 en calidad de compresión y fidelidad de la señal de vídeo.

3.3.2. Pizarra virtual

En los SMI destinados al e-learning síncrono el instructor envía de forma continua flujos de información hacia los alumnos, y éstos también envían algún tipo de información de realimentación de vuelta al instructor. Específicamente, además de audio y vídeo, otra de las funcionalidades más importantes de estos sistemas es la pizarra virtual, que permite a los participantes de una sesión de e-learning compartir materiales educativos como presentaciones de diapositivas u otros documentos. Adicionalmente, se pueden realizar anotaciones sobre dichos documentos de forma que el instructor concrete aspectos de su contenido que no puedan estar del todo claros o los complementa; también los alumnos podrían anotar sobre el documento, de forma que se permitiría un cierto nivel de trabajo colaborativo.

Existen gran cantidad de trabajos sobre pizarras virtuales, especialmente orientados al trabajo sobre Tablet-PC. Uno de los primeros trabajos fue el realizado en [McCanne, 1992], en él se describe una aplicación de pizarra virtual (*wb*) basada en transporte IP Multicast que utiliza gráficos PostScript para cada una de las diapositivas y las anotaciones. Más tarde, en [Floyd *et al.*, 1997], los autores añaden una capa de transporte multicast fiable a la misma permitiendo un alto número de participantes. Un trabajo posterior [Tung, 1997] amplió la funcionalidad de *wb* para añadir la posibilidad de navegar por el historial de operaciones de dibujo realizadas, además de incorporar otras nuevas funcionalidades.

Un aspecto muy importante de las herramientas de pizarra virtual dentro de los sistemas de e-learning síncrono es que permitan realizar anotaciones libres, es decir, a mano alzada, sobre los contenidos, típicamente las diapositivas. El sistema *Xlibris*, desarrollado en [Schilit *et al.*, 1998], permite la anotación de documentos utilizando formas libres. Más tarde, en [Golovchinsky y Denoue, 2002], se propone un nuevo algoritmo, probado sobre *Xlibris*, para preservar la correcta posición de las anotaciones insertadas en los documentos digitales cuando son visualizados en diferentes dispositivos con distintas relaciones de aspecto.

Por otro lado, la recomendación T.126, propuesta por la ITU-T, define una pizarra virtual como un conjunto de áreas de trabajo divididas a su vez en capas, cada una con sus propios contenidos.

Una primera capa es la destinada a la representación de los contenidos, por ejemplo diapositivas, mientras que por encima de ella se sitúan una o varias capas de anotaciones, sobre las que los usuarios pueden realizar trazos de pizarra. Por último, por encima de las capas de anotaciones se sitúan las capas de punteros virtuales, en las que se renderizan los punteros de los usuarios cuando éstos desean apuntar algún elemento del área de trabajo.

En los siguientes apartados se describen las tecnologías disponibles para la implementación de una pizarra virtual, organizadas en función de las tres partes en que se organiza típicamente una pizarra virtual: presentación de contenidos, anotaciones sobre los contenidos y puntero virtual.

3.3.2.1. Presentación de contenidos

La mayor parte de los materiales educativos de los que se hace uso durante una sesión de e-learning síncrono provendrán de documentos habitualmente utilizados en clases expositivas. Sería conveniente que los materiales que se utilizan durante una clase magistral tradicional fueran directamente utilizables en una sesión de e-learning, como permiten la mayor parte de las herramientas. Es este sentido, los materiales formativos suelen desarrollarse en una gran variedad de formatos, lo que conduce a que algunos deban adaptarse de alguna forma para que sean válidos para su uso en determinadas herramientas. Éstas, habitualmente utilizan una serie de formatos:

- **PowerPoint:** es una aplicación de Microsoft que permite crear presentaciones de diapositivas en un formato propietario. Permite la creación de animaciones y su uso está muy extendido en el campo de la enseñanza. El gran inconveniente de esta tecnología es que se trata de un formato cerrado, con lo que fundir el documento con las anotaciones realizadas sobre el mismo resulta muy difícil.
- **PDF:** es un formato de documento universal, destinado a principalmente a documentos imprimibles, independiente de la plataforma que almacena los gráficos internamente en formato vectorial, por lo que el tamaño que suelen ocupar estos documentos es relativamente bajo en comparación con otras alternativas. Contiene toda la información necesaria para la representación impresa de la información, por lo que no es necesario realizar ningún tipo de ajuste para su impresión. Una de sus principales ventajas es que, a través de impresoras PDF, es posible convertir cualquier documento imprimible en PDF. Gracias a herramientas de generación de fichero PDF, es posible fundir las anotaciones que se realicen con el documento original.
- **Flash:** es un conjunto de tecnologías de Adobe, es decir propietarias, que permiten añadir información multimedia dentro de las páginas web. De esta forma, es necesario un plug-in que se instala en el navegador para reproducir los contenidos Flash, del cual existen versiones para los navegadores web más extendidos, hasta el punto de que, según el fabricante, se encuentra instalado en 820 millones de equipos conectados a Internet. Flash se utiliza muy a menudo para crear animaciones e incrustar vídeo dentro de las páginas web. Recientemente, incluso permite el streaming de audio y vídeo utilizando el mismo plug-in. Adicionalmente, flash puede trabajar con gráficos vectoriales o mapas de bits, por lo que el tamaño de sus ficheros suele ser pequeño.
- **Imágenes vectoriales:** en lugar de utilizar un formato de documento original, se convierte previamente cada una de sus páginas en una imagen vectorial que será la que se envía hacia el resto de participantes en la sesión. Un ejemplo de este tipo de soluciones es SVG, un estándar del W3C para la representación de gráficos vectoriales basados en el formato XML. Con SVG, es posible incluso implementar animaciones utilizando el lenguaje SMIL. Representa el equivalente a la tecnología Flash, con la salvedad de que SVG es un estándar abierto.

Existen otros formatos de documentos también muy frecuentes en los entornos de enseñanza como pueden ser documentos Word, Excel, Autocad, etc. Está claro que, dada la gran variedad de formatos, es muy difícil, por no decir imposible, que una sola herramienta soporte de forma nativa todos estos tipos de documentos. Algunas herramientas de e-learning síncrono, como se vio en el capítulo 2, utilizan la compartición de aplicaciones para utilizar formatos que la propia herramienta no soporta de forma nativa. Así, bastaría con compartir la aplicación que sirve para renderizar documentos en ese formato para que los participantes en la sesión puedan compartir sus contenidos. Esto tiene el problema de que anotar sobre dichos contenidos para después fundirlos dentro del mismo fichero es extremadamente complejo.

Otra solución consiste en adoptar uno o varios formatos nativos al que deben convertirse todos los documentos en otros formatos para que puedan ser utilizados en las sesiones de e-learning. En ese caso, el formato o los formatos elegidos deberían presentar una serie de características:

- Compacto: el formato de representación no debe ser voluminoso para acelerar su transmisión por la red. El ancho de banda necesario para transmitir un documento con este formato debe ser pequeño.
- Particionable: debe ser posible comenzar a renderizar páginas de un documento sin haberlo recibido completamente. De esta forma, es posible servir las páginas que componen el documento bajo demanda, no siendo necesaria la espera por la recepción completa del documento para comenzar su exposición.
- Universal: el formato debe estar ampliamente extendido de forma que gran parte de los contenidos educativos se puedan utilizar directamente sin ningún tipo de modificación. En cualquier caso, la universalidad también implica que existan conversores a este formato para otros tipos de documento.
- Editable: interesa también que dicho formato sea editable de tal forma que se puedan combinar los documentos originales con las anotaciones y trazos de pizarra que se realicen sobre los mismos. Así, es posible almacenar un documento junto con las anotaciones que el instructor y los alumnos hicieron sobre el mismo.

3.3.2.2. Anotaciones

Una anotación es un trazo con una determinada forma, o a mano alzada, que realiza un usuario sobre los contenidos que se exponen dentro de la sesión de e-learning síncrono. Es una funcionalidad muy útil para que el instructor concrete aspectos que hayan podido quedar dudosos o complemente la información que las contienen diapositivas. De la misma forma, también es útil para que los alumnos tomen sus propias notas, para que anoten la diapositiva a instancias de profesor para resolver algún problema que les proponga, o para que planteen una cuestión al resto de participantes.

La posibilidad de anotar los contenidos digitales que se emplean durante una clase no es sólo una característica del e-learning síncrono. En los últimos tiempos han aparecido multitud de trabajos destinados al uso de sistemas de anotación, o tinta digital, en ámbitos educativos presenciales y no presenciales, favorecidos por la gran facilidad y versatilidad de los equipos Tablet PC para realizar anotaciones digitales. En [Mock, 2004] se describe un sistema que integra anotaciones de formas libres con diapositivas durante la presentación de una clase. En este sistema, únicamente el instructor tiene un Tablet PC con el que proyecta la presentación sobre una pantalla de pared.

El sistema desarrollado en [Huang, 2003] soporta la anotación de documentos de forma colaborativa utilizando Tablet PC. Está basado en una cache compartida para almacenar los documentos y tecnología IP Multicast para comunicar todos los equipos.

Otro de los sistemas que trabaja sobre Tablet PC es el descrito en [Anderson *et al.*, 2004]. Consiste en una aplicación de presentación de clases magistrales que permite las anotaciones

manuscritas sobre las diapositivas. Dicha aplicación, denominada *Classroom Presenter*, fue mejorada en [Anderson *et al.*, 2007] para permitir una mejor compartición de las anotaciones digitales entre instructores y alumnos. Este sistema está principalmente enfocado al trabajo en entornos inalámbricos locales.

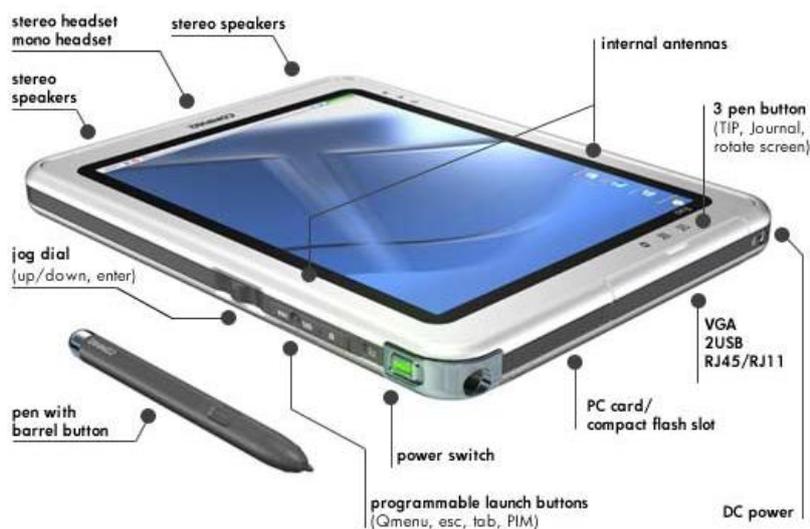


Figura 3.14: Tablet PC

En la figura 3.14 se muestra un ejemplo de Tablet PC. Estos equipos constituyen un paso intermedio entre una PDA y un computador portátil, ya que ofrecen la potencia de estos últimos y permiten escribir a través de su pantalla táctil. Para ofrecer toda esta funcionalidad necesitan una versión específica de Windows XP. En concreto, utilizan Windows XP Tablet PC Edition. Este sistema operativo ofrece un SDK para el desarrollo de aplicaciones que utilizan tinta electrónica.

3.3.2.3. Puntero virtual

El puntero virtual representa un cursor del que puede hacer uso un participante durante el desarrollo de una sesión de e-learning síncrono para apuntar algún elemento de los contenidos que actualmente estén siendo visualizados. Inmediatamente aparecería un cursor en el equipo del resto de participantes que seguiría los movimientos del ratón o elemento apuntador, por ejemplo un *stylus* para un Tablet PC, que utilice el usuario para indicar una posición en pantalla. También podría utilizarse una tableta digitalizadora cuando los contenidos a apuntar requiere una gran precisión.

Se han realizado varios trabajos sobre punteros virtuales o *telepointers*, especialmente en el campo del trabajo colaborativo para la edición gráfica. Así, en [Adams *et al.*, 2005] se realiza un estudio del impacto que tiene el uso de un puntero virtual para la retención del conocimiento y la satisfacción de los alumnos recibiendo educación a distancia. El estudio concluye que gracias al puntero virtual se mejoró entre un 30 y un 40% la retención de conocimientos a largo plazo, al tiempo que los alumnos percibían menos sensación de lejanía por parte del instructor.

En [Bu *et al.*, 2006] se describe un algoritmo para estimar los movimientos del cursor en función del los movimientos pasados y así evitar los efectos del jitter en la suavidad del movimiento.

En [Dyck *et al.*, 2004] se propone, tras un análisis de las implementaciones existentes, una nueva implementación de un puntero virtual que utiliza técnicas como el transporte UDP, predicción del movimiento, compresión de los mensajes, corrección de errores y buffering de los datos en el

receptor.

Ya por último, la IETF ha desarrollado un formato para el transporte de punteros virtuales en tiempo real sobre RTP [Civanlar y Cash, 2000]. Bajo este estándar, es posible transportar la posición del puntero, así como el estado de los botones asociados al puntero (si se utilizan) y el cursor a utilizar.

3.3.3. Mensajería y control de presencia

La mensajería instantánea o chat consiste en el intercambio de mensajes de texto entre varios usuarios casi en tiempo real. Esto permite establecer comunicaciones entre participantes cuando otros medios, por ejemplo el audio, no están disponibles, o bien para servir de complemento al resto de medios que permiten la interacción entre usuarios. Habitualmente, la mensajería instantánea suele venir asociada con el control de presencia, que permite conocer cuándo un usuario está disponible para mantener una conversación.

Gracias a la extensibilidad de SIP, se han definido un gran número de extensiones, entre las cuales se encuentra SIMPLE (*SIP for Instant Messaging and Presence Leveraging Extensions*), que constituye un estándar abierto. Estas extensiones se definen como paquetes de eventos que puede ofrecer una entidad SIP [Rosenberg *et al.*, 2002c] y [Rosenberg, 2004]

XMPP [Saint-Andre, 2004] es un protocolo, abierto y basado en XML, inicialmente pensado para la mensajería instantánea y el control de presencia. Sin embargo, gracias a que su funcionamiento consiste en el streaming de elementos XML en tiempo real, las aplicaciones del protocolo son numerosas. Aunque la infraestructura de red es independiente del protocolo, habitualmente se suele implementar siguiendo una filosofía cliente/servidor utilizando conexiones TCP.

Las siglas IRC (*Internet Relay Chat*) [Oikarinen, 1993] se corresponden con un protocolo para el intercambio de mensajes de texto entre dos o más personas en tiempo real. En este protocolo los clientes se corresponden con cada uno de los participantes en la conversación, y los servidores proporcionan el servicio trasladando los mensajes entre clientes. Este servicio se organiza en salas de conversación o canales, donde los usuarios pueden conversar acerca de un tema concreto. Es un protocolo textual, por lo que es fácilmente depurable.

La ITU-T, a través de su recomendación T.140, define una codificación para los mensajes de texto que se intercambian en una sesión de mensajería instantánea. Para complementar esta recomendación, la IETF propone un estándar para el transporte de este tipo de mensajes en tiempo real a través de RTP [Hellstrom y Jones, 2005].

3.4. Control del uso de la palabra

Típicamente, en las sesiones de e-learning síncrono participan un elevado número de participantes que desean interactuar entre sí. En una conferencia entre dos participantes es muy factible suponer que ambos se pondrán de acuerdo para determinar cuál es el turno de palabra en un momento dado. Así, es posible mantener una conversación interactiva sin que uno interrumpa el discurso del otro. Sin embargo, cuando el número de participantes en la conferencia crece, es más difícil que se pongan de acuerdo de forma espontánea quien tiene el turno de palabra, más si cabe teniendo en cuenta que no existe una interacción cara a cara.

Esta situación, que ocurre para las conversaciones de audio, puede generalizarse a otras situaciones. Por ejemplo, para determinar cuál es el usuario que tiene permiso para acceder a un recurso compartido, como puede ser el acceso a un canal de vídeo, realizar anotaciones sobre una pizarra virtual, u obtener el foco de una aplicación compartida. En estos casos, es interesante también controlar quién tiene permiso para conceder estos privilegios. Todas éstas, son funciones que debe ofrecer un protocolo de control del turno de palabra.

Muchos son los estudios que se han realizado sobre el control del turno de palabra dentro de grandes conferencias. Un primer trabajo es el descrito en [Dommel y Garcia-Luna-Aceves, 1997], en el que se introducen los requisitos para el desarrollo de un protocolo de gestión del turno de palabra en herramientas multimedia colaborativas, cumplimentado más tarde por [Koskelainen *et al.*, 2006]. En [Malpani y Rowe, 1997] se introduce una herramienta para la gestión del turno de palabra en seminarios web con un gran número de usuarios desarrollados sobre la red multicast MBONE. En la tesis [Boyd, 1996] se aborda el problema del control de turnos en las aplicaciones colaborativas como un problema más amplio partiendo de una aproximación socio-psicológica, emulando el modelo humano de interacción conversacional. Otros autores [Fantar *et al.*, 2004] proponen el uso de SIP para gestionar los turnos, si bien esta aproximación es desaconsejada por la propia especificación de SIP [Rosenberg *et al.*, 2002b]. En [Qiu *et al.*, 2002] se propone un algoritmo para el control de turnos cuyo funcionamiento ha sido evaluado sobre Internet2 con una herramienta de videoconferencia interactiva.

Hasta hace poco tiempo no han aparecido soluciones estándar para el control de turnos dentro de una conferencia. La serie de protocolos T.120 implementa, dentro del control de la conferencia, la asignación de turnos de palabra. No obstante, esta norma está muy relacionada con el marco de conferencias de H.323, habitualmente desarrollado sobre redes RDSI. La IETF ofrece un protocolo denominado BFCP cuya utilización no está vinculada al uso de otros protocolos de la IETF.

3.4.1. Protocolo BFCP

El protocolo BFCP (*Binary Floor Control Protocol*) [Camarillo *et al.*, 2006] se utiliza dentro de una conferencia para gobernar el acceso a recursos compartidos. Estos recursos pueden representar múltiples situaciones, por ejemplo el permiso para intervenir oralmente en la conferencia, realizar anotaciones sobre la pizarra compartida, etc.; BFCP no interpreta el significado de cada recurso, sólo se encarga de gestionar los turnos de uso de dicho recurso.

En esta especificación se consideran una serie de entidades:

- **Participantes:** estas entidades representan participantes en la conferencia que pueden solicitar la concesión de turnos al servidor de control. Pueden existir otros participantes en la conferencia que no tengan permiso para solicitar turnos.
- **Moderador:** notifica al servidor de control la decisión de conceder o no un turno a un participante que lo solicita.
- **Servidor de control de turnos:** admiten o rechazan solicitudes de concesión de turnos por parte de los participantes. Además, consulta al moderador la concesión de los turnos cuando son solicitados por los participantes.

Dentro de la conferencia, cada uno de los turnos (*floor*) está relacionado con un determinado recurso. La concesión de un turno a un participante le habilita para utilizar el recurso asociado de alguna forma. En un caso típico, la concesión de un turno, permite a un participante enviar su audio a la conferencia para ser oído por el resto de participantes. No obstante, en una conferencia pueden existir multitud de turnos, cada uno representando un recurso distinto.

BFCP está diseñado para trabajar con un reducido ancho de banda por lo que utiliza una codificación binaria que redundante en un tamaño de mensaje pequeño. Para el transporte de estos mensajes se realiza utilizando TCP, por lo que se evitan problemas relacionados con la pérdida de un mensaje que de otra forma podría ocurrir si no se utilizara un protocolo fiable como TCP. El principal inconveniente del uso de TCP para el transporte es que la fragmentación de los mensajes debe hacerse a nivel de aplicación, lo que añade cierta complejidad adicional a las aplicaciones. Asimismo, el servidor de control de turnos debe mantener una conexión TCP con cada uno de

los participantes que puedan solicitar turno, con lo que la escalabilidad del protocolo puede verse limitada.

Los participantes en una conferencia pueden conocer la existencia de los turnos dentro de la misma, siguiendo el mismo mecanismo de negociación de medios que se utiliza en SIP para establecer las sesiones multimedia mediante SDP [Camarillo, 2006]. De esta forma, BFCP se integra directamente en un marco de conferencias gobernados por SIP.

Capítulo 4

Diseño de un prototipo de herramienta: evaluación y optimización

En este capítulo se presenta el proceso de diseño de un prototipo de herramienta de e-learning síncrono que se utiliza para la capacitación del personal dentro de una gran corporación. Este proceso está guiado por los requisitos deducidos en el capítulo 2 y las alternativas tecnológicas analizadas en el capítulo 3, y consistirá en un proceso de prototipado iterativo que supondrá la evaluación de las diferentes alternativas de diseño y la optimización progresiva hasta el diseño final.

El diseño que se plantea se divide en capas. Una primera capa, la capa de transporte, es la encargada de realizar la transmisión de los datos entre los diferentes participantes. En este sentido, se definen y evalúan los distintos escenarios posibles a partir de la tecnología de distribución seleccionada. La capa de procesamiento de medios especifica los medios que el prototipo puede utilizar. En su descripción, se abordarán las alternativas de diseño dentro del prototipo, determinando qué alternativa es la óptima para el caso de la capacitación en grandes corporaciones. La capa de control tiene como misión el establecimiento y finalización de las sesiones. Para su diseño, se evalúan las diferentes alternativas disponibles y la solución óptima para el entorno concreto en que operará el prototipo. Por último, para el diseño de la interfaz de usuario se han tenido en cuenta las interfaces de las herramientas analizadas en el capítulo 2 con el fin de homogeneizar el prototipo con las mismas, ya que, en gran medida, todas presentan similitudes en sus interfaces de usuario.

4.1. Introducción

En el capítulo 2 se realiza un amplio análisis de las herramientas que existen en el mercado y en el ámbito investigador para llevar a cabo actividades de e-learning síncrono. Como conclusión del capítulo, en el apartado 2.4, se plantean las características comunes que exhiben las herramientas de e-learning síncrono, y se plasman los requisitos mínimos que debe satisfacer cualquier herramienta de este tipo cuando se aplica a la formación de personal dentro de grandes corporaciones.

Un ejemplo de gran corporación es ArcelorMittal, con quien la Universidad de Oviedo desarrolla un proyecto de colaboración denominado eTipo¹ para la implantación de una plataforma de e-training en su departamento de recursos humanos. Este proyecto ha servido de marco para el desarrollo de la presente tesis, para lo cual se ha implementado un prototipo de plataforma de e-learning síncrono siguiendo los requisitos mínimos, enumerados en el apartado 2.4.2. Dichos requisitos fueron deducidos de la caracterización de la red de ArcelorMittal y del perfil de sus empleados como ejemplo paradigmático de una gran corporación. Asimismo, determinan que deben ser utilizados los siguientes medios:

¹Referencia: CN-06-038.

- **Audio.** Transporta la mayor parte del mensaje que el instructor hace llegar a los alumnos y, en consecuencia, tendrá prioridad sobre el resto de medios dentro de la plataforma.
- **Vídeo.** Habitualmente se utiliza para transportar únicamente el busto parlante del instructor. Esto sirve para reforzar la sensación de presencia de la figura del instructor en los alumnos, de forma que no se sientan aislados y se evitan posibles ansiedades que pudieran generarse en el alumno o en el instructor en relación al proceso de enseñanza [Chassie, 2002] y [Weller, 2007]. En algunas ocasiones puede resultar conveniente su uso para que el instructor muestre a los alumnos objetos o piezas industriales, o secuencie los pasos de una determinada acción. En ese caso, será necesario utilizar una alta resolución espacial y temporal respectivamente.
- **Mensajería instantánea.** Permite la comunicación entre todos los participantes de la sesión de e-learning a través de mensajes de texto. Esto es útil para fomentar las relaciones entre los miembros, al tiempo que permite a los alumnos plantear dudas al instructor o a otros alumnos.
- **Control de presencia.** Además de reforzar la sensación de presencia de los participantes en la sesión de e-learning, el control de presencia es útil para que el instructor lleve registro de la asistencia a la clase por parte de los alumnos.
- **Presentación de diapositivas.** Constituyen la mayor parte de los contenidos que se utilizan como soporte durante el desarrollo de una actividad formativa sobre la plataforma prototipo. Deben poder utilizarse los formatos más comunes como PowerPoint y PDF.
- **Anotaciones.** Se utilizan para realizar concreciones sobre las diapositivas o sobre una diapositiva en blanco a modo de pizarra compartida. Pueden resultar útiles para la resolución de ejercicios planteados.
- **Puntero virtual.** Su principal cometido es apuntar elementos dentro de una diapositiva.

En lo que resta de capítulo se describen en profundidad las diferentes alternativas de diseño para cada una de las partes del prototipo, para, a continuación, realizar una evaluación de cada una de ellas y concluir una solución óptima para el caso concreto de aplicación en ámbitos corporativos.

4.1.1. Modelo de interacción

La principal misión de una herramienta destinada al e-learning síncrono consiste en hacer llegar las explicaciones y comentarios del instructor a los alumnos de forma que éstos puedan asimilar e interiorizar las competencias que están siendo entrenadas durante una sesión. Además, debe permitir una cierta reciprocidad en las comunicaciones, esto es, la información no debe fluir únicamente desde el instructor hacia los alumnos. Si fuera el caso, no habría forma de que un alumno planteara dudas o sugerencias al instructor, ni éste podría adaptar el ritmo de la clase en función de esta información de realimentación que obtiene de los alumnos. Por tanto, tiene que existir tráfico de información en ambos sentidos, aunque no necesariamente todo tipo de tráfico debe ser simétrico. Adicionalmente, se puede permitir que el tráfico fluya también entre alumnos, con lo que se posibilita así implantar estrategias conducidas a fomentar el aprendizaje y el trabajo colaborativo entre alumnos e instructor.

Desde un punto de vista práctico, el objetivo inicial que se busca cumplir con la plataforma a desarrollar es portar las clases magistrales que se imparten de forma presencial en el centro de formación de una gran corporación industrial mediante una metodología esencialmente expositiva a un entorno en línea que permita la participación de los empleados en la actividad formativa, independientemente del lugar en que se encuentren (véase figura 4.1). Se hace necesario, pues,

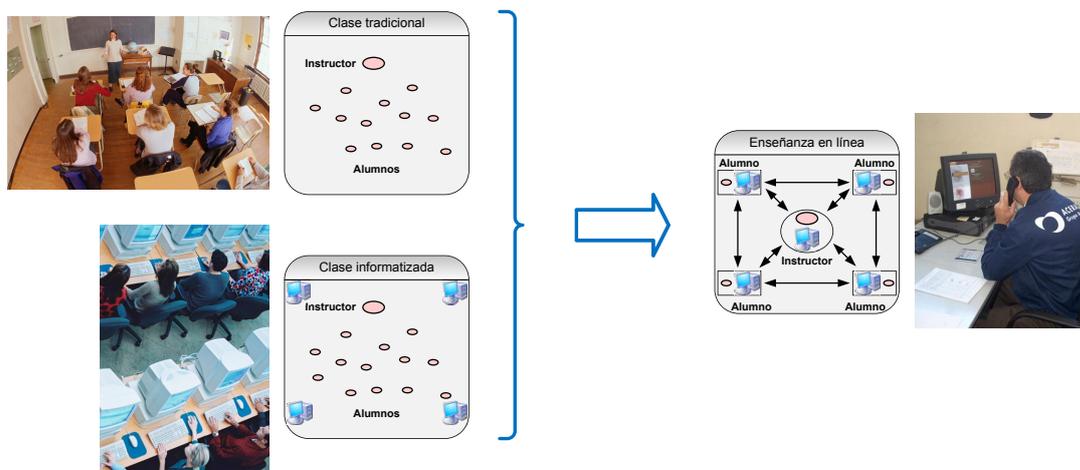


Figura 4.1: Traslación de las clases presenciales a un entorno en línea

reproducir las condiciones de un aula en un entorno virtual. En este sentido, a lo largo del presente trabajo se denominará *clase virtual* a las interacciones que se producen entre un conjunto de usuarios utilizando una agregación de medios: audio, vídeo, pizarra compartida, mensajería instantánea, etc.; que permiten reproducir más o menos fielmente las relaciones que se producen en un aula tradicional.

En una situación ideal todo tipo de interacciones estarían permitidas entre los miembros de una clase virtual. Esto es, tanto el instructor como los alumnos se podrían comunicar con cualquier otro miembro de la sesión utilizando todos los medios de los que la clase virtual está compuesta (audio, vídeo, mensajería instantánea, pizarra compartida...). Sin embargo, existen una serie de condicionantes que dificultan, y en algunos casos imposibilitan, esta aproximación para todos los medios. Por un lado, la disparidad en los anchos de banda de red disponibles en cada una de las ubicaciones donde los alumnos pueden localizarse, lo cual resulta incluso más determinante si dicha ubicación se encuentra fuera de la red corporativa de la empresa, posiblemente Internet. Interesa que una plataforma de este tipo sea capaz de amoldarse a situaciones de bajo ancho de banda para que todos los alumnos puedan participar en las actividades independientemente de la capacidad de su enlace de red y su localización geográfica. Esta es una cuestión importante que tiene gran implicación a la hora de plantear un modelo de interacción instructor-alumnos y alumnos-alumnos dentro de la plataforma.

Tampoco conviene olvidar que los alumnos pueden plantear dudas al instructor al tiempo que éste puede instarles a resolver un determinado ejercicio. Así, la plataforma debe proporcionar un mecanismo de realimentación para que los alumnos puedan comunicarse con el instructor de forma simple y eficiente. Además, esta comunicación debe llevarse a cabo de tal forma que no suponga una sobrecarga cognitiva excesiva para el instructor, puesto que, en ocasiones, el número de alumnos que participan en una determinada actividad puede ser elevado.

Partiendo de esta base, se debe definir cuál debe ser el modelo de interacción permitido en la plataforma, es decir, las diferentes interacciones entre los participantes a una sesión de e-learning que estarán permitidas. Básicamente, existen tres modelos de interacción:

- Modelo 1-N. Este es un modelo en el que existe una única fuente de datos y N receptores de información pasivos. El *webcasting* es la técnica más representativa que cumple este modelo; la información fluye en un solo sentido desde el emisor a los receptores. El instructor

representaría el emisor, mientras que los alumnos serían meros receptores sin posibilidad de enviar ningún tipo de información de realimentación al instructor. En la figura 4.2 puede verse una representación de este modelo.

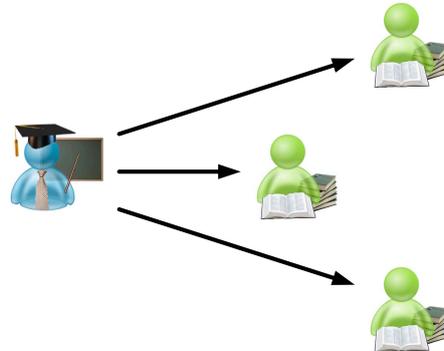


Figura 4.2: Modelo de interacción 1-N entre instructor y alumnos

- Modelo N-1. En este modelo la información fluye en sentido contrario al anterior. Existen N fuentes, representadas por cada uno de los alumnos, y un único receptor, el instructor. Es una solución que permite a los alumnos proporcionar realimentación al instructor acerca del seguimiento de la sesión de e-learning. En este caso, el instructor se convertiría en un mero receptor pasivo, con lo que se anula la figura del instructor como facilitador del proceso de aprendizaje. Además, tiene el gran inconveniente de que toda la información fluye hacia el instructor, con la posible sobrecarga que se puede generar en él, dependiendo del número de alumnos. Este modelo aparece representado en la figura 4.3.

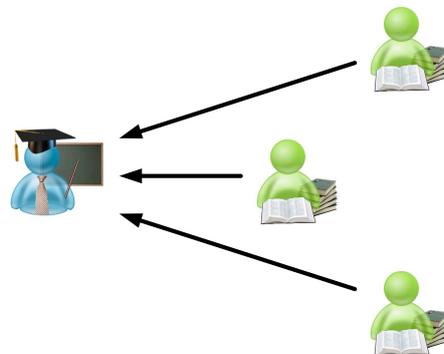


Figura 4.3: Modelo de interacción N-1 entre instructor y alumnos

- Modelo N-N (Colaborativo). En este modelo, por contra, cualquiera de los participantes en la sesión puede actuar como emisor de datos, permitiendo así una participación colaborativa. No sólo un alumno puede enviar y recibir información del instructor, sino que también puede comunicarse con otro alumno, posibilitando así las interacciones entre todos los miembros de la comunidad. De esta forma, no existe un individuo distinto al resto en el sentido de que la información fluya o parta de él como ocurría en los otros dos casos. En la figura 4.4 se representan las interacciones que pueden producirse bajo este modelo.

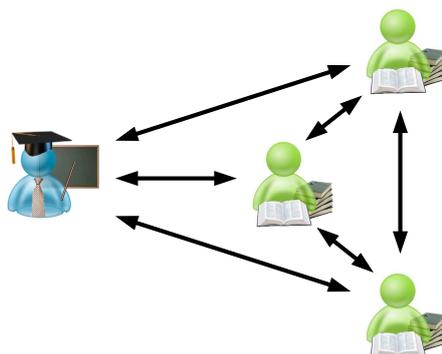


Figura 4.4: Modelo de interacción N-N entre instructor y alumnos

Dentro de la clase virtual no existe un único modelo de interacción. Por ejemplo, la elección de un modelo 1-N para todos los medios que componen la clase virtual dificultaría enormemente el proceso de aprendizaje por la unidireccionalidad de la comunicación. Tampoco es factible un modelo N-1 puesto que desaparecería la figura del instructor ya que no podría comunicar ningún tipo de información a los alumnos, sólo éstos enviarían información al instructor. Por último, un modelo de interacción N-N para todos los medios de la clase virtual implicaría un elevado ancho de banda consumido en la red, lo que supondría una gran limitación por los requisitos que impondría sobre la red subyacente.

En lugar de eso, cada uno de los medios que componen la clase virtual viene caracterizado por un modelo de interacción determinado, dependiendo de si resulta beneficioso obtener realimentación de los alumnos a través de dicho medio. Es importante también destacar que el modelo de interacción tiene grandes implicaciones sobre la forma en que se transporta un determinado medio en la red y los requisitos que impone sobre ésta. En general, es más fácil transportar un medio cuyo modelo de interacción es 1-N que otro con un modelo de interacción N-N. En el último caso, se transportarían a través de la red, al menos, N flujos de datos en el mejor de los casos (si está habilitado el transporte multicast), y en el peor de los casos (si sólo son posibles conexiones unicast) un número de flujos determinado por la fórmula 4.1:

$$\text{Flujos} = N \times (N - 1) \quad (4.1)$$

Por otro lado, un modelo de interacción 1-N implica, al menos, un flujo de datos desde el instructor cuando existe la posibilidad de utilizar técnicas multicast, y N flujos si es necesario establecer conexiones unicast con cada uno de los alumnos de la sesión.

En principio, el modelo de interacción para todos los medios será 1-N como mínimo. En el caso de que resulte conveniente proporcionar realimentación desde los alumnos al instructor utilizando un determinado medio, se utilizará entonces este medio bajo un modelo N-N.

En la tabla 4.1 se muestran resumidas las ventajas e inconvenientes que presenta utilizar cada uno de los medios como información de realimentación desde los alumnos hacia el instructor.

La utilización del audio de los alumnos como fuente de realimentación hacia el instructor tiene la gran ventaja de que permite una comunicación más natural y espontánea, puesto que los alumnos pueden plantear sus dudas y comentarios al tiempo que el instructor imparte la clase. Además, permite la interacción de los alumnos entre sí, de forma que esta interacción contribuye a la consolidación de una comunidad virtual gracias a las relaciones sociales que se producen. Sin embargo, tiene el inconveniente del ancho de banda en la red que consume, puesto que es necesario transportar el audio de cada uno de los participantes al resto, con lo que la

escalabilidad de la plataforma se vería comprometida. Otro problema que plantea es debido al carácter espontáneo de las cuestiones que los alumnos realizan. Éstas no son registradas por lo que no pueden ser pospuestas para un tiempo después si en el momento no pueden ser respondidas. En el caso de recibir varias preguntas al mismo tiempo, si no existe un mecanismo adecuado de gestión del turno de palabra, se puede crear una gran sobrecarga cognitiva en el instructor que se ve obligado a atender a todas las cuestiones planteadas en un periodo corto de tiempo. En consecuencia, el audio no se utilizará como medio de realimentación para evitar estas situaciones y al mismo tiempo habilitar el prototipo para su utilización en entornos de muy bajo ancho de banda.

Tabla 4.1: Utilización de los distintos medios como fuente de realimentación

Medio	Ventajas	Inconvenientes
Audio	Comunicación más natural Espontaneidad Fomentar comunidad virtual Permite el trabajo colaborativo	Ancho de banda moderadamente alto Posible sobrecarga del instructor
Vídeo	Asegurar presencia de los alumnos Permite el trabajo colaborativo	Gran ancho de banda Aporta poca información
Mensajes de texto	Dudas quedan registradas Bajo ancho de banda	Necesaria redacción Poca espontaneidad
Control de presencia	Identificación participantes Fomentar comunidad virtual	No asegura presencia
Diapositivas	Permite el trabajo colaborativo	Aporta poca información
Anotaciones	Permite resaltar partes importantes Realizar esquemas/gráficos	Posible sobrecarga
Puntero virtual	Apuntar detalles importantes Realizar ejercicios	Posible sobrecarga

Por otro lado, la utilización del vídeo de los alumnos no proporciona prácticamente ningún tipo de información de realimentación. Únicamente puede servir para asegurar la presencia de cada uno de los alumnos frente al equipo, permitiendo al resto de participantes conocer sus rasgos físicos, lo cual, dependiendo de la audiencia, puede ser beneficioso o contraproducente: beneficioso porque ayuda a consolidar la comunidad virtual mediante la identificación de cada uno de los miembros; y contraproducente porque puede acentuar en un momento dado las diferencias culturales, étnicas y raciales creando prejuicios que de otra manera no existirían. También posibilita en cierta manera el trabajo colaborativo, por ejemplo, si éste supone la demostración de la secuencia de pasos de un proceso determinado y la resolución del vídeo utilizada es apropiada para mostrar estas fases. El gran inconveniente de utilizar el vídeo con un modelo N-N es el gran ancho de banda necesario para transmitir el vídeo de todos los participantes, sensiblemente mayor que el necesario para transmitir el audio. Si a esto se une que los beneficios de su utilización son menores que la utilización del audio, se descarta el vídeo como medio de realimentación.

El intercambio de mensajes de texto entre los participantes de la sesión, además de servir para fomentar las relaciones sociales entre ellos, puede utilizarse como información de realimentación de los alumnos hacia el instructor. De esta forma, plantearían sus dudas escribiendo un mensaje que enviarían al instructor. En ese caso, es más apropiado un entorno de mensajería instantánea

global, en el que todos los participantes pueden ver los mensajes que escriben el resto, frente a un entorno en el que la comunicación alumno-instructor se realice de forma privada, pues así, varios alumnos pueden plantear la misma duda y, de esta forma, se evitan cuestiones repetidas. La gran ventaja que supone la utilización de mensajería instantánea es que permite al instructor abordar las cuestiones de los alumnos en el momento que estime oportuno sin que un alumno pueda interrumpir la clase de forma discrecional. Esto es debido a que, una vez que el alumno plantea su cuestión, ésta queda reflejada en el panel de mensajes junto con la persona que la planteó. Otra ventaja que el intercambio de mensajes de texto ofrece es el menor ancho de banda con respecto al audio, lo que lo hace especialmente recomendable en situaciones de bajo ancho de banda. Por contra, el principal inconveniente de esta técnica es que un alumno debe redactar su cuestión para plantearla al instructor, lo que suele resultar difícil cuando el tema que se está tratando es nuevo para el alumno. Además, pueden introducirse ambigüedades en la redacción que obliguen a un nuevo intercambio de mensajes para matizar la pregunta que redunde en una mayor lentitud a la hora de atender las dudas de los alumnos. Este problema será mayor o menor en función del perfil de alumnos que estén recibiendo instrucción, puesto que influye en gran medida los conocimientos lingüísticos de los mismos.

Aunque no puede considerarse un medio por sí mismo puesto que no sirve para transferir conocimientos, el control de presencia sirve, en cierta medida, para proporcionar realimentación a todos los participantes de la sesión de e-learning. En este caso, sirve para informar de la identidad de los participantes conectados, lo cual repercute en que cada uno de los alumnos pierda la sensación de aislamiento y se sienta parte de una comunidad virtual. Está claro, por tanto, la gran ventaja que supone que todos los participantes sean conscientes del número y la identidad de todos los miembros de la sesión, por lo que es del todo imprescindible adoptar un modelo de interacción N-N para el control de presencia. Sin embargo, que un alumno aparezca conectado no implica directamente que asiste y realiza un seguimiento de la sesión de e-learning. Esta funcionalidad sólo garantiza que está conectado a la sesión.

La posibilidad de exponer una presentación de diapositivas es, a priori, una cualidad que se le ofrece al instructor. Se puede abrir esta posibilidad también a los alumnos para fomentar el trabajo colaborativo. Así, se intercambiarían los papeles entre instructor y alumnos, lo cual puede ser útil cuando se realiza la labor tutorial y un alumno expone el trabajo realizado. De cualquier forma, este medio no permite plantear dudas o sugerencias, por lo que no es adecuado para proporcionar información de realimentación por parte de los alumnos.

Las anotaciones sobre las diapositivas constituyen uno de los principales mecanismos para ofrecer al instructor realimentación de los alumnos. Las anotaciones pueden ser utilizadas por los alumnos para resaltar partes concretas dentro de una diapositiva, o bien el instructor podría instar a uno o varios alumnos a anotar la diapositiva con esquemas o texto como respuesta a una pregunta. De esta forma, se puede facilitar una nueva práctica educativa que la meramente expositiva. También pueden ser utilizadas por el instructor para complementar la información que contienen las diapositivas que se exponen o concretar algunos aspectos que no pudieran quedar claros. Asimismo, dado que todos los participantes de la sesión pueden realizar anotaciones, se puede utilizar una pizarra en blanco para desarrollar una estrategia colaborativa. Como punto débil de utilizar las anotaciones como fuente de realimentación se puede mencionar la posible sobrecarga que se generaría si muchos participantes anotaran la misma diapositiva, convirtiendo la identificación del autor de cada anotación en una tarea ardua. Este problema se solventaría si el instructor dispusiera de un mecanismo de gestión de turnos que pudiera determinar quien tiene permiso para realizar anotaciones sobre una determinada diapositiva.

Ya por último, el puntero virtual, además de servir al instructor para apuntar detalles concretos dentro de una diapositiva, puede utilizarse para que los alumnos indiquen partes de la diapositiva que le resultan dudosas, o incluso para apuntar elementos como respuesta a una pregunta realizada por el instructor. De esta forma, al igual que ocurría con las anotaciones, se proporcionan mecanismos para innovar en el modelo de instrucción para que el instructor no se vea limitado

al modelo de clase magistral tradicional. El inconveniente que presenta es la necesidad de limitar el número de punteros visibles en cada momento para no producir una excesiva sobrecarga, debiendo identificar cada uno de ellos con su participante correspondiente.

A continuación, se resumen los modelos de interacción permitidos dentro de la plataforma para cada uno de los medios que se utilizan:

- Audio: 1-N. Sólo el instructor está habilitado para enviar audio a la sesión de e-learning.
- Vídeo: 1-N. Ocurre lo mismo que con el audio.
- Mensajería instantánea: N-N. Permite la realimentación hacia el instructor por parte de los alumnos.
- Control de presencia: N-N. Posibilita que cada participante en la sesión de e-learning sea consciente de quiénes participan en la misma.
- Presentación de diapositivas: 1-N. Únicamente el instructor puede iniciar la exposición de presentaciones de diapositivas.
- Anotaciones: N-N. Es otro de los medios que permite la realimentación desde los alumnos hacia el instructor. Éste puede instar a los alumnos a escribir o anotar la diapositiva actual para consolidar los conocimientos adquiridos. También puede ser utilizado por el instructor para resaltar aspectos en la diapositiva.
- Puntero virtual: N-N. Se puede utilizar para apuntar elementos dentro de la diapositiva actual. Será de ámbito colaborativo, permitiendo a los alumnos activar un puntero virtual para apuntar zonas de la diapositiva a instancia del instructor.

4.2. Arquitectura de red

Está claro que uno de los aspectos más importantes a tener en cuenta en el diseño de una plataforma de e-learning síncrono es la red subyacente y cómo los datos son transportados por la misma². Esto es así porque, habitualmente, los recursos de red son limitados y el comportamiento de la misma no es constante, tal y como se expuso en el apartado 3.2.1. Por consiguiente, debe prestarse especial atención al diseño arquitectónico de la plataforma prototipo y su despliegue en la red para, por un lado satisfacer los requisitos planteados consumiendo los menos recursos posibles y por otro, hacer frente a estados cambiantes de la red, especialmente ante situaciones dramáticas de pérdida de paquetes y retardos.

Uno de los requisitos más importantes que se plantean en el desarrollo del prototipo es la simplicidad de gestión de la plataforma. Para ello, se debe evitar la utilización de servidores en la medida de lo posible; el prototipo debe ser una herramienta eminentemente *peer-to-peer*. Esto tendrá gran impacto sobre el diseño de la plataforma prototipo.

Como se ha visto en el apartado anterior, sobre la plataforma fluyen varios tipos de datos, entre los que se encuentran el audio y el vídeo del instructor, los mensajes de texto que intercambian los participantes de una sesión de e-learning, etc. Cada uno de estos medios viene caracterizado por un modelo de interacción ya definido y, en consecuencia, puede ser transportado de forma independiente al resto. Toda la información multimedia que se intercambia en el ámbito de una sesión de e-learning dentro del prototipo estará asociada a una sesión RTP, independientemente del tipo de medio al que pertenezca. En concreto, se utiliza una o varias sesiones RTP específicas para cada medio ya que, como se justificó en el apartado 3.2.4, es conveniente distribuir los medios en sesiones independientes. Por esta razón, es posible utilizar un mecanismo distinto

²En el apéndice A se proporciona una descripción de la red de *Arcelor* en Europa.

para la distribución de cada uno. Por simplicidad, en lo que resta de apartado se analizan las alternativas para el transporte de cada medio individual.

A pesar de que el protocolo RTP está orientado al transporte de flujos de datos continuos como audio y vídeo, y, de hecho, es el estándar de facto para el transporte de esta clase de medios sobre redes de paquetes, el resto de medios no continuos que se utilizan dentro de la clase virtual (pizarra compartida, mensajes de texto...) también serán transmitidos utilizando RTP. Esta aproximación supone una serie de ventajas:

- Gracias a que RTP puede trabajar encima de UDP e IP Multicast, se pueden utilizar técnicas de multidifusión para el transporte de todos los medios.
- Se aprovechan las características de RTP para todos los medios, incluidos los no continuos. De esta forma, se pueden utilizar los números de secuencia de la cabecera RTP para detectar paquetes duplicados, perdidos o que llegan en distinto orden en el que fueron enviados.
- Gracias a las marcas de tiempo es posible sincronizar de forma global todos los medios que componen la clase virtual, y no sólo el audio y vídeo. Esto posibilita grabar las clases que transcurren en tiempo real para después reproducirlas bajo demanda, con una ordenación temporal idéntica a la original.
- Dado que todos los datos que se transportan por RTP están asociados, a través de un SSRC, a un determinado flujo, se les puede relacionar de forma automática con el participante que los generó, sin necesidad de etiquetar apropiadamente cada uno de los datos.
- Algunos routers intermedios por los que fluye el tráfico de la plataforma son capaces de identificar el tráfico RTP. De este modo, es posible diferenciar este tráfico respecto al resto y proporcionarle una QoS distinta [Blake *et al.*, 1998].

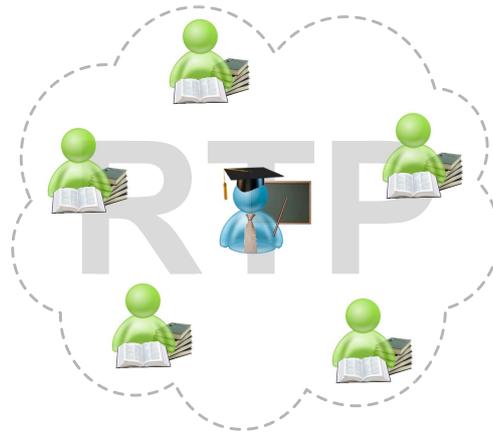
Las soluciones que se proponen para el transporte de medios varían en función del tipo de transporte a nivel de red que se utilice. En concreto se han valorado tres situaciones diferentes:

- Multicast: se utilizan técnicas IP multicast para el transporte de los datos.
- Multi-unicast: se crean múltiples enlaces punto a punto, uno por cada uno de los destinos hacia los que deban enviarse los datos.
- Híbrida: es una combinación de las dos anteriores.

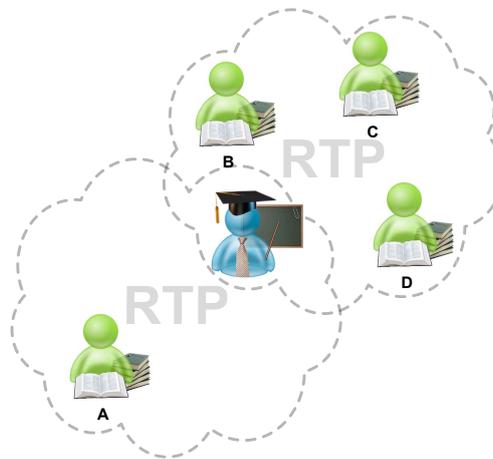
La primera decisión de diseño a tomar es si para distribuir un determinado medio utilizaremos una única sesión RTP o varias. La utilización de una sola sesión RTP, como se refleja en la figura 4.5(a), tiene como principal ventaja que todos los participantes tienen una visión global de la sesión y, gracias a las funcionalidades de RTCP, pueden detectar cuándo el enlace de un determinado participante está congestionado. Además, permite que cada participante mantenga la lista de participantes en la sesión sin necesidad de utilizar un protocolo de más alto nivel, simplemente utilizando de nuevo la funcionalidad de RTCP.

Por contra, mediante la utilización de varias sesiones, como ocurre en la figura 4.5(b), es posible modificar el flujo de datos entre una y otra sesión e, incluso, enviar unos flujos específicos a unos determinados participantes. Su principal desventaja es que los participantes no pueden conocer el estado global de la sesión y es necesario la utilización de un protocolo como SIP para mantener una lista de participantes. En esta situación, el alumno A tendría la percepción de que la clase virtual está compuesta por él mismo y el instructor; el resto de alumnos no serían visibles directamente a través de RTP.

Una vez se ha decidido la utilización de una o varias sesiones RTP para el transporte de un medio, es necesario determinar las posibilidades de distribución del medio dentro de una sesión



(a) Sesión única



(b) Múltiples sesiones

Figura 4.5: Transporte de un medio mediante (a) una única sesión RTP y (b) varias sesiones RTP

RTP. Desde este punto, por simplicidad, se supone que existe una única sesión RTP que engloba a todos los participantes.

En el esquema de la figura 4.6 aparecen representadas las diferentes posibilidades para la distribución de cada medio dentro de la plataforma.

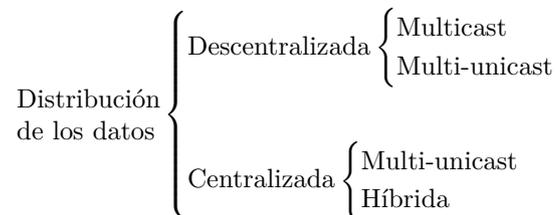


Figura 4.6: Posibles alternativas de distribución de los datos

En el siguiente apartado se describen las ventajas e inconvenientes de la utilización de cada una de estas alternativas y se plantea la solución óptima para dos escenarios concretos: la capacitación de personal dentro de una red corporativa y la formación a distancia dentro de unos estudios universitarios.

4.2.1. Evaluación de alternativas y optimización

Inicialmente, se puede pensar que la elección de una u otra arquitectura para el transporte de los medios dentro de la plataforma únicamente afecta al total del ancho de banda de red consumido. No en vano, cuantas más réplicas de los datos circulen, tanto mayores serán los recursos de red necesarios. Por tanto, habrá que minimizar el número de réplicas que se envíen a la red. No obstante, existen también otros factores condicionados por dicha elección y que deben ser estimados, puesto que influyen en el comportamiento final de la plataforma. En concreto, éstos son los parámetros que se analizan:

- Configuración de la red. Algunas de las soluciones planteadas requieren que la red esté configurada de una forma determinada o, en su defecto, que su configuración pueda ser modificada de forma adecuada. Tal es el caso cuando se utilizan técnicas multicast para el transporte de los datos. Como se vio en el apartado 3.2.2, estas técnicas requieren que los routers de la red sean conscientes del tráfico multicast y lo redirijan correctamente. Desgraciadamente, esto no es la norma general, y muchos de los routers distribuidos por Internet filtran este tráfico. Por tanto, la posibilidad de utilizar transporte multicast es, en la actualidad, limitada.
- Ancho de banda. El aspecto de diseño que tiene mayor impacto sobre el ancho de banda que consume la plataforma es su modelo de interacción. En un modelo 1-N los datos fluyen exclusivamente desde el instructor hacia los alumnos, de forma que existe una única fuente de datos en toda la plataforma. Por el contrario, en un modelo N-N, los alumnos pueden enviar datos hacia el instructor y a otros alumnos. Así, el número de fuentes de datos en este último caso crece linealmente con el número de participantes.
Por otro lado, debe considerarse el número de réplicas de cada uno de los flujos que deben transmitirse por la red. No es lo mismo transportar un único flujo de datos desde una fuente que llegue a todos los participantes, que enviar un flujo individual replicado para cada uno.
- Retardo. En aplicaciones de e-learning síncrono en que se busca posibilitar la interactividad entre instructor y alumnos, es de crucial importancia el retardo en las comunicaciones. En el apartado 3.2.1.4 se indicaron las latencias máximas admisibles para sistemas de conferencias de voz, que constituyen los datos con mayores requerimientos de bajo retardo. La utilización de una arquitectura determinada para la plataforma puede afectar al retardo que perciben los usuarios.
- Escalabilidad. Interesa que la plataforma pueda ser utilizada por muchos usuarios al mismo tiempo, sin que ello implique una degradación excesiva de la calidad de servicio proporcionada a medida que más participantes se conectan. La elección de una u otra arquitectura limita el número máximo de usuarios simultáneos que puede soportar la plataforma.
- Recursos de la máquina. Deben tenerse en cuenta también las necesidades computacionales que impondrá la plataforma a los clientes dependiendo del rol que jueguen en las sesiones de e-learning y el modelo de interacción. Así, si el instructor es el único participante que transmite información, la carga que soportará su equipo será mayor que la que soportan los usuarios que sólo reciben datos, puesto que es necesario realizar un proceso de captura y codificación de la información en el equipo del instructor. En este sentido, debe estimarse

la carga que supone tanto la generación como la visualización de los datos, así como la transmisión de éstos por la red. Requiere más carga computacional enviar N réplicas del mismo flujo a otros tantos destinos unicast que una única réplica a un grupo multicast.

Es posible disponer de uno o varios servidores intermedios para centralizar el transporte de los datos y así liberar al instructor y los alumnos. De esta forma, es razonable pensar que la potencia computacional de estos elementos es más elevada que la de los equipos de los participantes en la sesión. Igualmente, en un modelo de interacción 1- N la potencia necesaria en el equipo del instructor es mayor que en los equipos de los alumnos, en contraste con un modelo N - N donde todos los equipos soportarían un carga computacional parecida. En todo caso, siempre resulta más factible sustituir el equipo del instructor por uno de mayor potencia que sustituir N equipos de todos los participantes en la sesión.

De todo esto se deduce que la mayor parte de la carga computacional debe situarse en servidores intermedios o, en el peor de los casos, en los equipos de los instructores, nunca en el de los alumnos, ya que no es posible conocer a priori el tipo de máquinas de estos últimos. Habitualmente nos encontraremos con alumnos con equipos con una potencia de cálculo heterogénea.

Todos estos parámetros no se pueden tratar de forma ortogonal evaluando cada uno de forma independiente, ya que de una u otra forma se encuentran relacionados entre sí. Por ejemplo, un sistema que implique transportar todo el tráfico por un único enlace de red, además de tener un posible cuello de botella, redundará en una plataforma que requerirá unos anchos de banda altos en dicho enlace.

En los siguientes apartados se evalúan cada una de las posibilidades de distribución de los datos dentro de la plataforma prototipo y se estima el efecto sobre cada uno de los parámetros anteriormente citados.

4.2.1.1. Transporte utilizando multicast

El caso ideal es aquél en el que el transporte de los flujos que fluyen por la plataforma, en su mayoría desde el instructor hasta los alumnos, se realice mediante técnicas de *multicasting*. De esta manera, el flujo enviado por un participante de la clase virtual llegaría al resto directamente. Recaería sobre la red la misión de hacer llegar los datos desde la fuente a todos los receptores. Puede apreciarse en la figura 4.7 cómo se llevaría a cabo la transmisión de los datos.

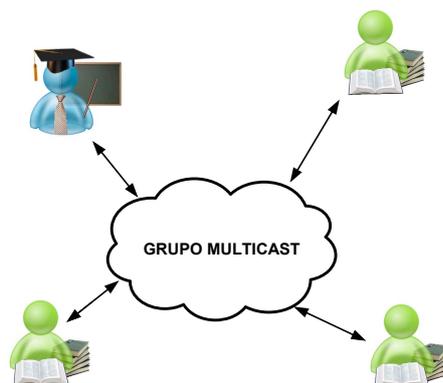


Figura 4.7: Transporte distribuido utilizando técnicas multicast

Para que esta solución fuese viable, sería necesario utilizar una dirección multicast³ que serviría para implementar un grupo de difusión al que deberían unirse el instructor y los alumnos que deseen participar en la clase.

Las principales ventajas de una solución basada en la difusión descentralizada de los datos utilizando técnicas multicast son:

- ✓ Ancho de banda: Es de todas las soluciones analizadas la que menos ancho de banda necesita, puesto que un único flujo de datos enviado desde la fuente llegará a todos los participantes en la sesión utilizando el menor ancho de banda posible.
- ✓ Retardo: Debido a que los datos fluyen desde la fuente directamente hasta el resto de participantes, el retardo sufrido por los datos es mínimo puesto que no es necesario ningún procesamiento intermedio de las señales.
- ✓ Escalabilidad: En esta solución no existe un punto intermedio por el que deban pasar las comunicaciones. En este sentido, el número de usuarios conectados a la clase virtual podrá crecer sin que ello resulte en un rendimiento menor como consecuencia del aumento lineal del ancho de banda necesario, directamente relacionado con el número de fuentes de datos.
- ✓ Recursos de la máquina: La fuente de datos, que a priori es quien más carga computacional debe soportar con esta solución, deberá transmitir un único flujo que la red se encargará de hacer llegar a todos los participantes de la clase. Por tanto, los recursos computacionales son, exclusivamente, los necesarios para generar la información.

Por contra, éstos son los inconvenientes que la distribución descentralizada multicast plantea:

- ✗ Configuración de red: Para que esta solución sea factible es necesario que los elementos de la red subyacente permitan el tráfico multicast. En la actualidad el tráfico multicast suele restringirse a los ámbitos corporativos. En consecuencia, si se pretende escalar esta solución a un nivel global es muy probable que no sea posible puesto que las redes por las que deban fluir los datos con toda seguridad impedirán el tráfico multicast.

4.2.1.2. Transporte utilizando multi-unicast

En aquellas situaciones en las que no es posible realizar el transporte multicast habrá que utilizar unicast. En este caso, cada fuente de datos es responsable de hacer llegar su flujo de datos al resto de participantes de la clase y, para ello, debe transmitir tantas réplicas como participantes en la clase menos uno. Sin embargo, también es posible la utilización de un servidor que se encargue de transportar los datos entre los participantes de la clase. Una fuente enviaría sus datos al servidor, que los reenviaría a todos los participantes.

Las dos posibles soluciones, que implican diferentes infraestructura de red, se describen a continuación.

4.2.1.2.1. Transporte descentralizado utilizando multi-unicast El equipo que actúa como fuente de datos deberá establecer conexiones punto a punto con el resto de participantes en la sesión. En la figura 4.8 pueden verse las distintas conexiones a establecer suponiendo un modelo de interacción N-N.

Esta opción tiene sus ventajas e inconvenientes:

- ✓ Configuración la red: Como todas las conexiones que se establecen son unicast, la red que subyace bajo la plataforma no debe cumplir ninguna condición específica.

³Dentro del rango de direcciones 234.0.0.0 - 239.255.255.255.

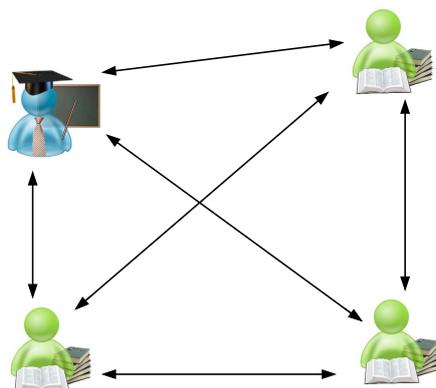


Figura 4.8: Transporte distribuido utilizando técnicas multi-unicast

- ✓ Retardo: Debido a que las conexiones son directas entre la fuente y los receptores, el retardo que introduce la etapa de transporte en los datos es mínimo.
- ✗ Recursos de la máquina: Todas las conexiones entre la fuente y los receptores deben partir desde el mismo punto, por tanto, el equipo fuente debe tener la suficiente capacidad de procesamiento para poder establecer tantas conexiones como participantes menos uno en la clase.
- ✗ Ancho de banda: Debido a que las conexiones son punto a punto, en el peor de los casos, desde cada fuente partirá un flujo independiente hacia cada uno del resto de participantes conectados a la clase virtual. Por consiguiente, el ancho de banda de subida necesario en la fuente será elevado. Tampoco se tiene en cuenta que se podrían estar enviando dos flujos iguales a la misma subred destino (donde podría estar habilitado el transporte multicast) por el mero hecho de que haya dos participantes conectados desde esa red a la clase virtual.
- ✗ Escalabilidad: Dado que el número de flujos que circularán por la red crecerá de forma exponencial con el número de participantes en la clase, la escalabilidad de esta solución es bastante baja.

4.2.1.2.2. Transporte centralizado utilizando multi-unicast Para paliar las carencias en cuanto a recursos computacionales y ancho de banda que presentaba la anterior alternativa, se puede situar un servidor en medio de las transacciones entre los participantes de la sesión como se ve en la figura 4.9. Este servidor ejercerá de *relay*, redirigiendo el tráfico que llega de un participante al resto, para lo cual, deberá crear las réplicas del flujo original necesarias.

Como la anterior, esta solución tiene sus ventajas:

- ✓ Configuración la red: Todas las conexiones a establecer para llevar a cabo el transporte de los datos son unicast, por lo que las restricciones que se aplican a la red son nulas.

Existen unos factores de los que a priori no se puede decir si constituyen una ventaja o un inconveniente para esta solución:

- Recursos de la máquina: En cierta manera, el problema de los recursos necesarios para reportar todas las conexiones se atenúa, dado que esta responsabilidad ya no recae en el equipo de la fuente sino en un servidor central. Es de suponer que la capacidad de proceso

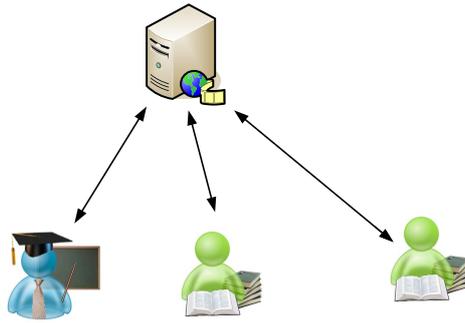


Figura 4.9: Transporte centralizado utilizando técnicas multi-unicast

del servidor será mayor que la de cualquier PC que utilice un participante, por lo que, en principio, deja de ser un problema tan acuciante.

- **Ancho de banda:** Al igual que ocurre con los recursos de la máquina necesarios el problema del ancho de banda desaparece desde el punto de vista de la fuente. En este caso, sólo necesitará enviar un único flujo hacia el servidor, mientras que éste necesitará un ancho de banda proporcional al número de usuarios conectados a la clase virtual. Se puede suponer que el ancho de subida disponible en el servidor es mayor que el de cualquier equipo de los participantes, por lo que el problema del ancho de banda se limita a un posible cuello de botella, dependiendo del número de participantes y el modelo de interacción, constituido por el servidor.

En cuanto al ancho de banda consumido en la red seguirá siendo elevado, puesto que se enviarán flujos replicados a cada uno de los participantes independientemente de que, en determinadas situaciones, se pueda utilizar la multidifusión.

Esta alternativa también tiene sus inconvenientes:

- ✗ **Retardo:** Por el hecho de introducir un elemento en medio del transporte de los datos, debido al procesamiento que debe realizar el servidor para redirigir los datos enviados por un participante al resto, se introduce un retardo extra.
- ✗ **Escalabilidad:** El número de flujos que circularán por la red seguirá dependiendo del número de participantes en la clase. Por tanto, tanto la capacidad de proceso como el ancho de banda necesarios en el servidor aumentarán exponencialmente con el incremento del número de participantes en la clase.

4.2.1.3. Transporte utilizando técnicas híbridas

Aunque dentro de esta categoría caben muchas posibilidades, tienen en común la utilización de técnicas de distribución multicast con enlaces punto a punto. La clave reside en aplicar el transporte multicast siempre que se pueda, en el peor de los casos, en las redes locales desde las que se conectan los alumnos, pues es la técnica que más beneficios reporta. Dos son las aproximaciones que se pueden seguir para llevar a cabo esta solución:

- **Transporte descentralizado.** No existe un punto central por el que deban fluir todos los datos.
- **Transporte centralizado.** Existe un servidor central que concentra y dirige todo el tráfico de la sesión.

Analizaremos cada una de las opciones en los siguientes apartados.

4.2.1.3.1. Transporte descentralizado utilizando técnicas híbridas Esta técnica consiste en la utilización de varios servidores (relays) que se encargarán de hacer de puente entre dominios en los que es posible la utilización del transporte multicast interconectados mediante redes que no permiten este tipo de tráfico. De esta forma, la fuente debe enviar el flujo de datos hacia su relay local que lo redirige al resto de relays, ubicados en dominios multicast distintos, y éstos a su vez lo redirigen a cada uno de los participantes a través de un grupo multicast. En la figura 4.10 aparece representado un esquema de las conexiones que se producen con esta solución.

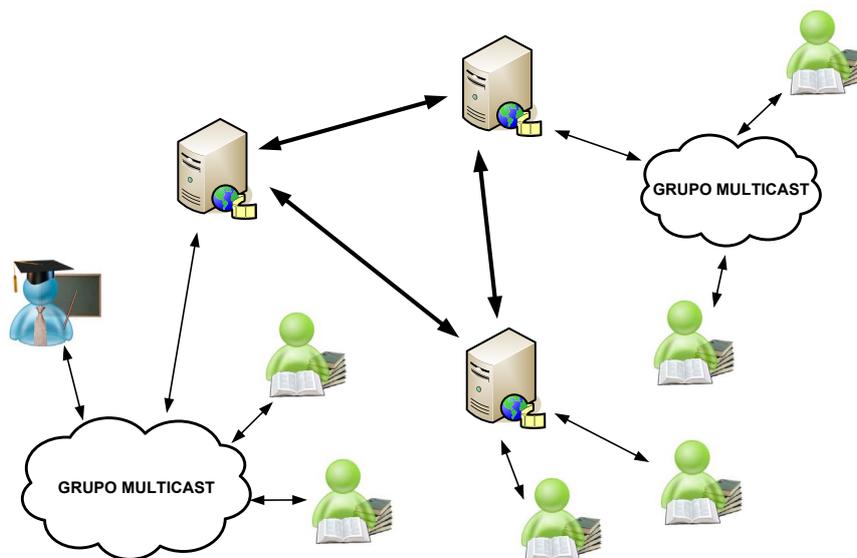


Figura 4.10: Transporte distribuido utilizando técnicas híbridas

En dicha figura, aparecen dos dominios multicast, es decir, dos ubicaciones geográficas en las que es posible utilizar transporte multicast. Cada una de estas regiones está gobernada por un relay que permite redirigir el tráfico de la sesión hacia el grupo multicast que controla y viceversa; reenvía el tráfico desde el grupo multicast al resto de relays de la sesión. Por ejemplo, si el instructor envía un flujo de datos lo hará a través del grupo multicast al que pertenece. Así, este flujo llega directamente a los alumnos del mismo grupo y al relay correspondiente. Éste último debe redirigir este flujo a los otros dos relays que, a continuación, lo reenvían a su grupo multicast.

Podrá destinarse un relay específico para aquellos alumnos cuyo proveedor de servicios filtre el tráfico multicast y, por consiguiente, no puedan suscribirse a un grupo multicast. De esta forma, este relay transportará los datos a cada uno de estos alumnos deslocalizados utilizando técnicas multi-unicast.

Esta opción presenta sus ventajas e inconvenientes:

- ✓ Configuración de red: Para la implementación de esta solución no es necesaria ninguna configuración específica de la red. La plataforma utilizará conexiones unicast entre los relays (y para los alumnos deslocalizados), lo que permite su uso a través de redes públicas, mientras que en destino, habitualmente redes locales, se utilizarán técnicas multicast. Con esto salvamos el problema del cierre al tráfico multicast de las redes intermedias al tiempo que aprovechamos la capacidad de redifusión en las redes locales.

- **Ancho de banda:** Si bien conseguimos ahorrar ancho de banda debido a la capacidad de utilizar multicast en destino, este ahorro será tanto más notable cuanto menos dispersos se encuentren los participantes de la clase. De esta manera, el caso ideal, sería aquél en el que todos los alumnos se encuentran en una misma subred en la que es posible realizar el multicasting. En ese caso, el ancho de banda necesario en toda la red sería el necesario para transportar un único flujo desde cada fuente; el mismo que para el caso visto en el apartado 4.2.1.1. Por contra, en el peor de los escenarios posibles todos los alumnos se localizarán en ubicaciones diferentes entre las que no pueden establecerse comunicaciones multicast. En este otro caso, estaríamos ante una solución similar a la vista en el apartado 4.2.1.2. En consecuencia, la aplicabilidad de esta solución depende en gran medida de la dispersión de los participantes.
- **Escalabilidad:** En principio, esta opción es más escalable que el uso de conexiones unicast. Por otra parte, lo habitual es encontrarnos que la mayoría de los participantes que asisten a una clase virtual se encontrarán localizados en las mismas zonas geográficas, en las que, muy posiblemente, se podrá de utilizar tráfico multicast. De cualquier forma, debido a que la comunicación entre los relays sigue realizándose a través de enlaces punto a punto, la escalabilidad de la plataforma dependerá del número de dominios multicast, siempre que el número de alumnos deslocalizados sea bajo.
- ✗ **Recursos:** La mayor parte del peso del transporte de los datos entre los participantes de la clase recae en los relays. Estos servidores deben redirigir el tráfico que circula por el grupo multicast que controlan al resto de relays y a los alumnos deslocalizados. Por tanto, para un flujo de datos desde una fuente, deben enviar una réplica a cada uno de los relays de la sesión y a cada uno de estos alumnos. Resulta razonable pensar que la carga computacional que esto supone sobre un equipo servidor es mínima.
 El problema más importante a resolver es qué hacer si el relay local deja de funcionar. Llegado el caso, todos los participantes que se encontraran bajo el dominio de dicho relay quedarían excluidos de la clase. En última instancia, cuantos más elementos de este tipo existan, menos tolerante a fallos será la plataforma en general.
- ✗ **Retardo:** Al existir uno o varios elementos intermedios en las comunicaciones entre los participantes de la clase, el retardo con el que los datos llegan desde una fuente a los receptores se verá incrementado.

4.2.1.3.2. Transporte centralizado utilizando técnicas híbridas En este otro caso, también se utilizan servidores relays para reducir el tráfico que fluiría por la red si únicamente se utilizaran conexiones punto a punto. La diferencia fundamental radica en la distinta función de un servidor respecto al resto. Existirá un servidor central por el que deberán pasar todos los flujos de datos. Esto, que a priori parece contraproducente, puede resultar útil si fuera necesario realizar un procesamiento intermedio de todos los flujos de datos que se intercambian los participantes de la sesión. Por ejemplo, en el caso de la sesión de vídeo, esta solución nos permitiría añadir una mosca⁴ a todos los flujos de vídeo de los participantes. En un caso extremo, este procesamiento de los datos podría suponer una recodificación de los mismos para enviar a cada relay un flujo específico⁵.

En la figura 4.11 puede apreciarse una representación de una solución para el transporte centralizado de datos utilizando técnicas híbridas. En dicha figura aparecen un servidor que centraliza todo el tráfico de la sesión y con el cual deben comunicarse todos los relays de cada uno de los dominios multicast. Este servidor central será el encargado de redirigir el tráfico a cada uno de

⁴Es una imagen que se utiliza como logotipo en las emisiones de televisión.

⁵Por ejemplo, para hacer frente a diferentes capacidades de los enlaces de red.

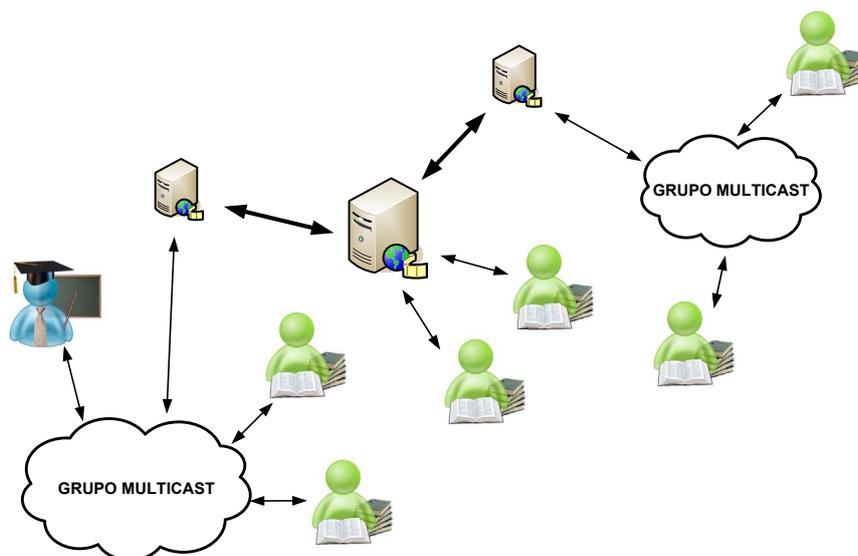


Figura 4.11: Transporte centralizado utilizando técnicas híbridas

estos relays para que éstos, a su vez, lo redirijan a su grupo multicast correspondiente. De igual forma, también se encargará de hacer llegar a los alumnos deslocalizados los flujos de la sesión, al tiempo que redirigirá los flujos generados por estos alumnos al resto de la sesión.

Esta solución presenta sus ventajas e inconvenientes:

- ✓ Configuración de red: Al igual que la anterior, esta solución no requiere ninguna configuración especial de la red subyacente, por lo que será aplicable a cualquier tipo de redes.
- Ancho de banda: Todo el tráfico de la sesión debe pasar por el servidor central, luego esto provocará un cuello de botella, que puede ser atenuado si en los relays se implanta algún tipo de política para filtrar los flujos que llegan a este servidor central, por ejemplo, enviando al servidor central un único flujo, ya sea mezclando previamente los flujos de todos los participantes que dependan del relay, o seleccionando uno de ellos en base a qué usuario esté en poder del turno de palabra.
- Escalabilidad: Si bien esta solución no es tan escalable como sería aquella que utiliza multicasting puro, el único posible cuello de botella en cuanto a rendimiento global de la plataforma sería el servidor central. Pese a todo, un dimensionamiento adecuado de este servidor posibilitará que se puedan establecer clases virtuales con alumnos distribuidos en varias localizaciones geográficas diferentes. Por tanto, la escalabilidad sigue dependiendo del número de localizaciones geográficas. De nuevo, al igual que ocurría con la anterior alternativa, los alumnos deslocalizados juegan un papel importante en la escalabilidad de la plataforma. En todo caso, resulta conveniente que su número sea bajo.
- Recursos: Con esta alternativa, descargamos a los relays del reenvío de todo el tráfico entre sí y depositamos esa responsabilidad en el servidor central. Por este motivo, la potencia de cálculo necesaria en este servidor puede llegar a ser elevada, dependiendo del número de alumnos de la clase, pero siempre será menor que la necesaria con un único servidor central y conexiones punto a punto con todos los participantes, sin la ocurrencia de relays intermedios.

- ✗ Retardo: El mayor problema que presenta esta alternativa es el retardo que sufren los datos en su viaje desde una fuente hasta los alumnos. Al retardo que ya de por sí introducen los relays locales hay que sumar el que ahora introduce el servidor central. Entonces, el problema se agrava doblemente.

4.2.1.4. Solución óptima

El prototipo de plataforma desarrollado tiene dos ámbitos de aplicación. El primero consiste en la utilización de la plataforma dentro de la red de una gran corporación. Para este caso, se toma como ejemplo la red de la empresa *ArcelorMittal*. En este caso, el uso principal que se le da a la plataforma es la capacitación de los recursos humanos para la realización de su actividad profesional. Por tanto, en ningún caso un alumno o instructor conectado a la plataforma se encontrará fuera de la red corporativa. Además, como se detalla en el apéndice A, en dicha red se puede habilitar el transporte multicast, en consecuencia, la arquitectura para la plataforma que más beneficios reporta implica la utilización de transporte descentralizado utilizando multicast.

El segundo ámbito de aplicación de la plataforma es la Universidad de Oviedo. Dentro de la red universitaria el transporte multicast se encuentra habilitado a nivel global, es decir, cualquier alumno o instructor de una clase virtual que se estuviera llevando a cabo dentro de la red de la Universidad de Oviedo podría participar en la clase independientemente del lugar desde el que se conecte, siempre que éste se encuentre dentro de la red de la Universidad. Sin embargo, ya que el uso que se le da a la plataforma en este caso es el de impartir docencia, bien sea correspondiente a cursos de doctorado, o para facilitar la acción tutorial de un profesor sobre sus alumnos, puede darse el caso de que un alumno quiera unirse a la clase virtual desde fuera de la red de la Universidad de Oviedo. En ese caso, se utiliza una solución híbrida como la que aparece en la figura 4.12.

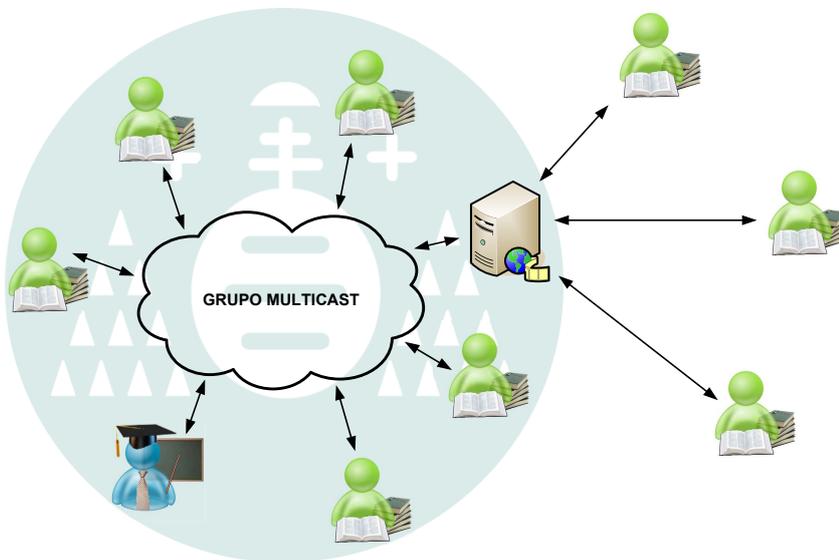


Figura 4.12: Arquitectura de la plataforma dentro de la Universidad de Oviedo

En dicha figura puede observarse cómo tanto el profesor como cinco alumnos se conectan a la plataforma desde dentro de la Universidad de Oviedo. En ese caso, utilizan un grupo multicast como mecanismo de transporte de los flujos de datos de cada uno de ellos. Por el contrario, existen tres alumnos deslocalizados, esto es, su ubicación no es conocida y se encontrará, en todo caso,

fuera de la red de la Universidad. De este modo, el transporte multicast no será posible para estos tres alumnos, por lo que se dispone de un servidor relay que se encarga de establecer conexiones punto a punto con cada alumno deslocalizado para reenviarles el tráfico que circula por el grupo multicast y hacer llegar a este último el tráfico que generan dichos alumnos.

4.3. Servidor relay RTP

Como se vio en el apartado 4.2, la utilización de un servidor relay va a depender de la arquitectura de la plataforma. Este relay tiene la misión de reenviar tráfico desde un grupo multicast a uno o varios destinos utilizando enlaces unicast. Dado que todo el tráfico que circula por la plataforma se agrupa en sesiones RTP, este servidor se comportará como un traductor RTP, cuyo funcionamiento fue descrito en el apartado 3.2.4.2. Específicamente, la misión que realiza es redirigir tráfico dentro de un grupo multicast a destinos fuera del grupo y viceversa, por esta razón recibe el nombre genérico de relay o repetidor.

El objetivo que se persigue con la utilización de un relay es emular el comportamiento de un grupo multicast independientemente de que la red física permita el transporte multicast o no para extender una sesión RTP más allá del alcance del grupo multicast. Así, un relay tendrá bajo su control un único grupo multicast. Su misión será, por tanto, redirigir el tráfico de este grupo a aquellos destinos que se encuentren fuera del grupo y al contrario. De esta forma, aquéllos participantes que se encuentren fuera de la red habilitada para multicast, gracias al relay, pueden enviar y recibir el tráfico del grupo multicast como si realmente pertenecieran a él. Se consigue que el nivel del transporte sea transparente para la herramienta independientemente de la ubicación del usuario. Así, la herramienta cliente enviará y recibirá flujos multimedia de la misma forma que si utilizara para ello un grupo de distribución multicast.

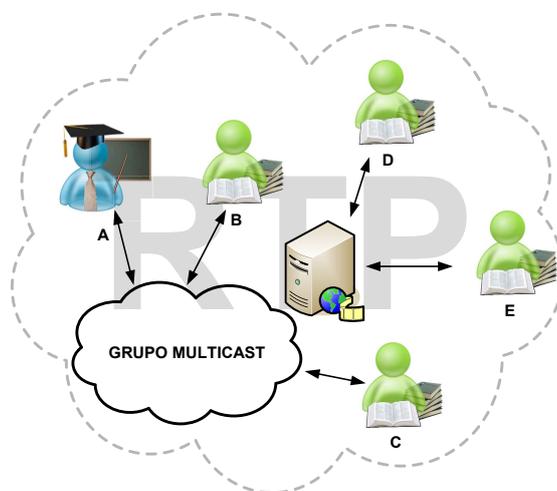


Figura 4.13: Relay RTP para la traducción de flujos de un ámbito multicast a otro unicast

En la figura 4.13 aparece representada una sesión RTP compuesta de cinco participantes: tres dentro de un grupo multicast y otros dos participantes en ubicaciones fuera del citado grupo multicast. La labor del relay consiste en reenviar los flujos de datos que emiten A, B y C al grupo multicast hacia D y E. De igual modo, el flujo de datos que emite D debe ser reenviado al grupo multicast y a E. Lo mismo ocurre para el participante E.

Para llevar a cabo su tarea, el relay debe mantener una tabla de enrutado RTP que determina

a qué destinos debe ser reenviado un flujo de datos a partir de la fuente a través de la que llega al relay. Existe dos modos de gestión de esta tabla:

- Rutas estáticas. Las rutas a las que deben reenviarse los flujos de datos están predefinidas y no pueden ser modificadas.
- Rutas dinámicas. La tabla de enrutado se construye de forma dinámica según se añaden participantes a la sesión RTP.

En los siguientes apartados se comentan las ventajas e inconvenientes de cada una de las alternativas.

4.3.1. Rutas estáticas

Cuando las rutas que se utilizan para redirigir el tráfico RTP le son proporcionadas al relay, se evita que éste tenga que analizar el tráfico para mantener estas tablas de enrutamiento. Habitualmente, estas rutas se especifican a través de ficheros de configuración o a través de un protocolo de nivel superior que se encarga de mantener la tabla de rutas y proporcionársela al relay. De esta forma se simplifica el diseño del servidor.

Esto presenta una serie de ventajas. Por un lado, para que un relay reciba tráfico de un grupo multicast es necesario que anteriormente se suscriba al grupo. Por consiguiente, es necesario conocer de antemano la dirección IP multicast que representa al grupo que el relay gestiona, lo cual es únicamente posible si se proporciona a través de una ruta estática. Además, esta opción permite utilizar varios servidores relay dentro de la misma sesión como se ilustra en la figura 4.14, puesto que de otra forma ambos relays no serían conscientes de la existencia del otro. Por otra parte, si un equipo que no pertenece a la sesión RTP intenta emitir un flujo de datos a la sesión a través del relay, inmediatamente será filtrado si no aparece en la tabla de enrutado. De esta forma se evitan ataques de tipo denegación de servicio por parte de usuarios que no aparecen en la tabla de enrutado.

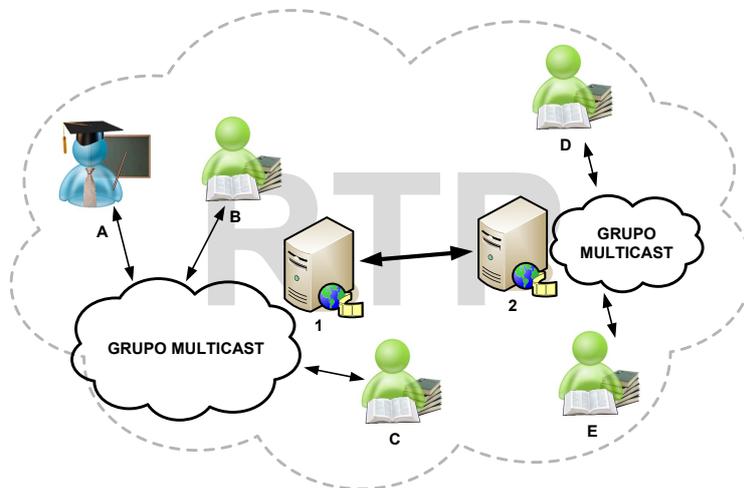


Figura 4.14: Utilización de varios relay RTP para la traducción de flujos

Esta alternativa, por contra, presenta un inconveniente. Si no se utiliza un protocolo de más alto nivel para mantener actualizada la tabla de enrutamiento, un usuario que quisiera unirse a la sesión RTP desde fuera del grupo multicast no podría si no aparece una ruta en dicha tabla que defina desde qué participantes puede recibir flujos el relay y hacia dónde deben ser redirigidos.

4.3.2. Rutas dinámicas

En esta otra alternativa, en cambio, el relay debe analizar el tráfico que le llega para mantener actualizada la tabla de rutas. Así, además de redigir el tráfico el relay debe encargarse de la gestión de la tabla de enrutado. Para ello descubrirá los participantes de la sesión en el momento que le envíen flujos RTP. Es decir, cada vez que reciba tráfico RTP, concretamente un paquete RTCP compuesto, de un nuevo destino, si este se encuentra fuera del grupo multicast que el relay gestiona, añadirá una entrada en la tabla de rutas para este nuevo participante. El tráfico que éste genere será redirigido al grupo multicast y a todos aquellos participantes que se encuentren fuera del grupo y de los cuales también se hayan recibido datos RTP.

Esta solución se asemeja más a un grupo multicast que la anterior. Cualquier equipo puede emitir a un grupo multicast pero sólo recibir datos de éste si se encuentra suscrito. De esta forma, la inclusión del destino en la tabla de rutas implica su suscripción automática al grupo multicast que emula el relay.

De nuevo, esta solución tiene sus ventajas e inconvenientes. La principal ventaja es la flexibilidad que permite en la composición de la sesión RTP a través de la tabla de rutas, ya que no es necesario especificar de antemano los participantes de la misma ni utilizar un protocolo diferente para mantener esta tabla. Por el contrario, el mayor problema que presenta es la debilidad que ofrece ante ataques del tipo denegación de servicio, ya que cualquier equipo podría emitir datos a la sesión y suscribirse a ella. Esto se puede paliar en parte si la inclusión de un destino en la tabla de rutas se realiza tras un número determinado de paquetes RTP válidos o bien tras la recepción de un paquete SDES de descripción de la fuente.

La solución implementada es una solución mixta, pues incluye rutas estáticas y dinámicas. Esto es, se le proporcionan al relay una serie de destinos (como por ejemplo el grupo multicast que gestiona y otros relays) a los cuales debe reenviar tráfico RTP, pero esto no es óbice para que el relay añada entradas en la tabla para reflejar la inclusión de nuevos participantes en la sesión RTP.

4.3.3. Network Address Translation

Uno de los mayores desafíos a los que tienen que hacer frente los SMI trabajando en red son los firewalls y los NAT. Estos últimos se utilizan para compartir una dirección IP pública, única en todo Internet, entre diferentes equipos. De este modo, una organización no se ve obligada a disponer de tantas direcciones IP públicas como equipos para que éstos puedan conectarse a Internet. En su lugar, dispone de una dirección IP pública, que es la que utilizan los equipos dentro de la red privada para acceder a Internet, tal y como se refleja en la situación 1 de la figura 4.15.

Este modo de operación se adapta bien al tradicional esquema cliente/servidor de los servicios que ofrece Internet, por ejemplo la navegación web, la consulta del correo electrónico, etc.; cuando el cliente se encuentra en la red privada. El router debería mantener una tabla dinámica relacionando conexiones entrantes y salientes de manera que pueda redirigir la respuesta del servidor, dirigida a la IP pública del router, hasta el usuario dentro de la red privada que realizó la petición. Concretamente, mantiene una tabla que asocia una dirección de transporte privada (IP y puerto) y una dirección de transporte pública remota. Las entradas de esta tabla tienen una validez temporal, más efímera para comunicaciones UDP que TCP. De esta forma, si no se produce ninguna comunicación que tenga como origen la dirección de transporte privada y como destino la dirección de transporte pública o viceversa, el router eliminará la entrada de la tabla.

Como también se aprecia en la situación 2 de la figura 4.15, el problema de la utilización de NAT aparece en las conexiones entrantes. Específicamente, en uno de los siguientes casos:

- Un equipo desde fuera de la red privada inicia la conexión con un equipo dentro de la red.

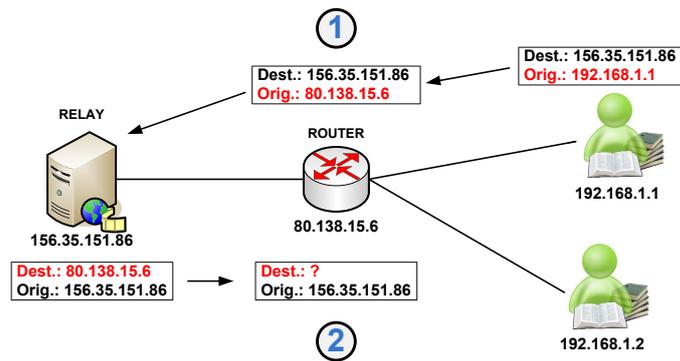


Figura 4.15: Comunicación entre dos equipos a través de un NAT

- Un equipo desde fuera de la red privada envía datos a un equipo dentro de la red una vez que el temporizador ha vencido y se ha eliminado la entrada correspondiente de la tabla NAT.

En ambos casos el router se encontraría en una situación en la que no podría determinar a qué equipo de la red privada redirigir el tráfico, con lo que filtraría los paquetes y los datos nunca llegarían a su destino, con el consiguiente corte en las comunicaciones.

Con el uso del relay para conectar alumnos desde fuera del grupo multicast a la clase virtual puede ocurrir alguna de las situaciones antes mencionadas. Cada una de las sesiones RTP que el relay maneja tiene asociadas dos direcciones de transporte: una para los datos y otra para la información de control RTCP. El relay tendrá siempre asignada una dirección IP pública, o en su defecto, se modificará la configuración de puertos en su router local para redirigir las conexiones de los alumnos directamente al relay. El problema aparece cuando el este último debe enviar los datos a un alumno que se encuentra detrás de un NAT. En ese caso estaremos en la situación 2 de la figura 4.15.

Para el tráfico RTCP no existirá ningún problema, pues el funcionamiento del protocolo de control se basa en el intercambio de paquetes entre todos los participantes, incluyendo el participante que se encuentra detrás del NAT. Así, cuando este participante envíe su primer paquete RTCP compuesto hacia el relay, el router creará una entrada en la tabla NAT que permitirá al relay comunicarse con el participante. Gracias a que el intercambio de paquetes RTCP es periódico, y que este periodo es siempre menor que el tiempo de vencimiento del temporizador del router, se garantiza así que las comunicaciones no se interrumpen.

Por contra, en el caso del tráfico de datos pueden existir problemas debido a que se utiliza una dirección de transporte diferente a la de control. Si el participante detrás del NAT no emite flujos a la sesión (sólo recibe), o el relay intenta enviarle flujos de otros participantes antes de que éste haya emitido nada, el tráfico será filtrado en el router. Esta situación se prolongará hasta que el participante emita datos RTP hacia el relay, momento en el que el router creará la entrada en la tabla NAT que permitirá la comunicación desde el relay hacia el participante.

La solución adoptada en la herramienta consiste en el envío periódico de paquetes UDP vacíos a través de la dirección de transporte de datos hacia el relay para que el router mantenga la asociación entre el relay y el participante. De esta forma, cuando el relay envíe flujos de datos hacia el participante no serán filtrados por el router.

4.4. Diseño de alto nivel de la herramienta

La plataforma prototipo desarrollada está compuesta de una herramienta cliente y, opcionalmente, de un servidor relay que se encarga de redirigir tráfico entre dominios multicast. La herramienta será idéntica para instructor y alumnos, de forma que el rol que cada participante juega en la sesión de e-learning no dependa de la herramienta en sí, sino del modo en que se unen a dicha sesión. Esto presenta una serie de ventajas. Por un lado, el cambio de papel dentro de una clase virtual se realiza de forma directa, no siendo necesario el entrenamiento en el manejo de otra herramienta. Un alumno dentro de la plataforma podría convertirse en instructor sin necesidad de utilizar una herramienta distinta. De igual modo, al tratarse de una única herramienta su mantenimiento es más simple puesto que toda la funcionalidad permitida a instructor y alumnos se engloba en el mismo producto software. Además, el manejo de la herramienta por parte de los usuarios es también más simple, puesto que no es necesario el uso de varias herramientas, cada una con una funcionalidad específica. De este modo, por tratarse de una herramienta integrada, es posible optimizar aquellas partes de la misma que se consideran cruciales para el correcto funcionamiento de la plataforma.

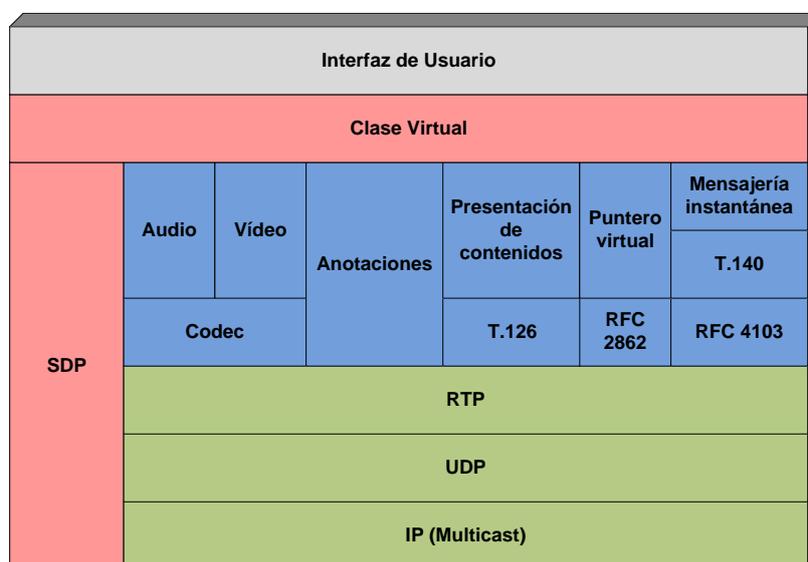


Figura 4.16: Modelo arquitectónico de la herramienta prototipo

Por otra parte, el diseño de este tipo de herramientas debe organizarse en capas, cada una con su propio cometido. En la figura 4.16 se refleja el modelo de capas seguido para el desarrollo de la herramienta cliente. En ella, aparecen representadas una serie de capas:

1. Transporte de medios: aparece resaltada en color verde y se encarga del transporte de los datos entre los participantes de la sesión de e-learning.
2. Gestión de medios: esta capa aparece en la figura 4.16 en color azul y su cometido es el de gestionar todos los medios que componen la clase virtual.
3. Control de la sesión: representada en rosa, sirve para administrar la clase virtual como un todo, especificando qué medios utilizar.
4. Interfaz de usuario: es la capa visible al usuario y a través de la cual interactúa con la plataforma.

En los siguientes apartados se describe en detalle la gestión que hace la herramienta de cada uno de los medios que pueden utilizarse en la clase virtual y cómo la interfaz de usuario ofrece dicha funcionalidad. De la misma forma, se detalla de forma pormenorizada la funcionalidad que debe proporcionar cada una de las capas y las decisiones de diseño tomadas para su implementación.

4.5. Procesamiento de medios

Esta capa de la herramienta prototipo es la que se encarga de toda la gestión de medios, que comprende desde la creación de las sesiones RTP a través de las que se envían y reciben los flujos de datos hasta la generación de la propia información que se transporta en dichos flujos. Cada uno de estos medios se gestionará de forma independiente al resto, ya que, como se vio en el apartado 4.2, en el prototipo a cada medio se le asocia una única sesión RTP independiente. Sin embargo, dentro de cada sesión circularán varios flujos, habitualmente uno por participante, aunque esto no limita que un participante de la sesión RTP pueda generar más de un flujo, como será el caso del instructor si abre varias presentaciones de diapositivas.

Además de la creación de las sesiones RTP de transporte y la generación de la información multimedia a enviar, otra de las misiones de esta capa es reproducir los flujos de datos multimedia que llegan al cliente. Esto, que a priori parece una cuestión menor, no lo es tanto cuando es posible recibir varios flujos del mismo medio, por ejemplo, si se recibieran dos flujos de audio pertenecientes a dos participantes distintos. Esta capa debe encargarse de la reproducción de los flujos de forma que la información sea inteligible. En este caso, se podría optar por fundir los flujos en uno solo o seleccionar aquél que debe ser reproducido. Para otros medios, como la mensajería instantánea, el problema es menos significativo puesto que es posible la reproducción de varios flujos de forma simultánea sin que ello afecte a la comprensión de los mismos.

La capa de procesamiento de medios maneja todos los medios de forma uniforme. Para ello identifica un determinado medio con una descripción SDP que permite definir, además del medio, parámetros como la codificación y otros aspectos de configuración. Como se verá en el apartado 4.9, esta aproximación es útil para configurar la clase virtual y especificar qué tipos de datos multimedia se utilizan dentro de la misma. Además, es posible pausar momentáneamente la reproducción de un flujo concreto o de un medio completo, es decir, todos los flujos de datos del mismo medio. Por ejemplo, pausar el flujo de audio de un participante, o pausar el audio de todos los participantes. Esta pausa sólo supone parar la reproducción del medio o flujo seleccionado, pero no implica que la herramienta deje de recibir datos por red. En todo caso, si se desea dejar de recibir datos de un medio determinado debe abandonarse la sesión RTP asociada a ese medio.

Otra característica que presenta la capa de procesamiento de medios, que se deriva del transporte multicast elegido para el intercambio de datos, es la posibilidad de utilizar todos los medios (audio, vídeo, pizarra. . .) bajo un modelo de interacción N-N. Esto quiere decir que el prototipo está preparado para trabajar con todos los medios de forma colaborativa, si bien por los requisitos mínimos deducidos en el apartado 2.4 y el perfil de los usuarios del prototipo, la interfaz de usuario limita las posibilidades de interacción de los usuarios ofreciéndoles para algunos medios un modelo 1-N.

4.5.1. Evaluación de alternativas y optimización

Para cada uno de los medios con los que trabaja el prototipo se realiza una evaluación de las diferentes alternativas para su representación y procesamiento. Se realizan evaluaciones iniciales para desarrollar sucesivas versiones del prototipo mediante un proceso iterativo que concluyen en el prototipo final descrito en el presente capítulo. A partir de cada una de las versiones desarrolladas se proponen diferentes optimizaciones orientadas a mejorar el prototipo en una siguiente iteración.

Se ha hecho especial hincapié durante el proceso de optimización en mejorar aquellas partes del prototipo más sensibles, especialmente, todas aquellas que están relacionadas con la presentación de contenidos.

A pesar de que el prototipo final es una herramienta plenamente operativa, aún sigue en marcha el proceso de mejora continua, orientado principalmente a mejorar la experiencia de los usuarios en el trabajo con la herramienta, y, en menor medida, en mejorar el rendimiento de partes concretas del prototipo.

Una consideración importante a tener en cuenta es que la evaluación y posterior optimización se realiza en base a las decisiones de diseño tomadas. Inicialmente, se analizan las alternativas para el desarrollo de cada una de las partes del prototipo que gestionan el procesamiento de medios, enumerando sus ventajas e inconvenientes. A partir de estos análisis, y mediante pruebas de laboratorio, se concluye cuál es el diseño óptimo para una plataforma prototipo de e-learning síncrono aplicada a la capacitación de recursos humanos dentro de grandes corporaciones.

La principal decisión que debe tomarse para el diseño de la capa de procesamiento de medios es cómo se gestionan a nivel global dentro de la plataforma, esto es, el modelo de interacción. Si bien, idealmente, el modelo ideal es aquél que permite comunicarse de igual a igual entre todos los participantes de la clase virtual, existen restricciones, derivadas de los requisitos que la herramienta debe cumplir, el perfil de los usuarios con los que trabaja, y la red sobre la que opera; que obligan a reducir intencionadamente la funcionalidad que el prototipo ofrece. Por tanto, habrá medios que permitirán una interacción colaborativa entre los participantes de la clase, y habrá otros que servirán para difundir el mensaje del instructor a los alumnos. Todo esto ya fue comentado en el apartado 4.1.1.

4.6. Procesamiento de medios: audio y vídeo

El audio y el vídeo pueden utilizarse de forma conjunta o independiente para transportar información educacional desde el instructor a los alumnos. Habitualmente, el vídeo transportará el busto parlante del instructor, mientras que el canal de audio se utiliza para transmitir su discurso.

Como se comentó con anterioridad, el audio y el vídeo, al igual que todos los medios en el prototipo, admiten la utilización de un modelo de interacción N-N. Esto quiere decir, que cualquier participante dentro de la clase virtual podría comunicarse con el resto a través de videoconferencia. Sin embargo, en el apartado 4.1.1, se argumenta el hecho de que los alumnos planteen sus dudas al instructor utilizando el canal de audio presenta más inconvenientes que ventajas, especialmente en situaciones de muy bajo ancho de banda disponible, por lo que se desecha su utilización de forma colaborativa y se opta por un modelo de interacción 1-N desde el instructor hacia los alumnos. De esta forma, la sesión RTP que se utiliza para transportar el audio sólo tendría un único flujo: el proveniente del instructor. En el caso del vídeo la situación es similar, pero con el agravante de que el ancho de banda que consume cada flujo de vídeo es significativamente mayor que el que consume el audio. Por este motivo, para ser estrictos, no se puede decir que la funcionalidad que ofrece el prototipo sea la videoconferencia, pues no se produce una comunicación bidireccional utilizando audio y vídeo.

La gran problemática que presenta el utilizar un modelo 1-N para la videoconferencia es la sensación de aislamiento que invade al instructor durante el desarrollo de la clase, pues no tiene constancia de la presencia física de los alumnos excepto a través del control de presencia y la mensajería instantánea. Además, tras realizar una indicación a los alumnos, debe esperar un tiempo por los mensajes de texto que éstos escriban que, dependiendo de las habilidades mecanográficas de los alumnos, puede ser mayor o menor. En cualquier caso, si no se implementa un mecanismo para informar al instructor de que un alumno está escribiendo un mensaje, esto tiende a incrementar la sensación de aislamiento del instructor.

Tanto en el caso del audio como del vídeo la información que se obtiene para generar dichos flujos desde el instructor procede de un proceso de captura en tiempo real. Es necesario capturar la imagen del instructor y su alocución, que componen su flujo de vídeo y audio respectivamente. En ambos casos se utilizará DirectShow como tecnología de soporte para la captura y la reproducción de los flujos de audio y vídeo que circulen por el prototipo. De esta forma, se construye una cadena de filtros enlazados (grafo DirectShow) para cada flujo. En las figuras 4.17 y 4.18 aparece un grafo DirectShow para la captura y transmisión de datos, y otro para la reproducción de datos provenientes de la red respectivamente. No obstante, aunque DirectShow es una tecnología muy extendida, hay situaciones para las que no existe un determinado filtro que se adecúe correctamente. En concreto, es necesario implementar un filtro DirectShow en los siguientes casos:

1. Ya que los flujos de datos deben ser enviados a la red utilizando RTP para el transporte, debe existir una conexión entre DirectShow y el nivel RTP. Esto se consigue con la implementación de un filtro sumidero cuya misión es enviar los datos del flujo a la red utilizando RTP.
2. Análogamente al anterior, en el equipo receptor es necesario obtener los datos de un flujo de la red para generar a partir de ellos el flujo que se reproduce bajo DirectShow. En este caso, se desarrolla un filtro fuente para obtener los datos del nivel RTP e introducirlos en el flujo DirectShow para su reproducción.
3. No todos los codecs útiles para la videoconferencia tienen un filtro asociado que les permita ser utilizado dentro de grafos DirectShow. Llegado el caso, habría que desarrollar un filtro a medida para el codec que se desee utilizar para codificar la señal de audio o vídeo.

En el apartado 3.3.1 se mencionan los codecs más adecuados para implementar soluciones de videoconferencia. Estos codecs se diferencian del resto, esencialmente, en la eficiencia del proceso de codificación en términos de ancho de banda utilizado. Esto debe ser así, porque habitualmente son utilizados en videoconferencias multipunto con un número de participantes significativo, por lo que el número de flujos de audio y vídeo puede ser elevado. De ahí que deba comprimirse la información al máximo para que cada flujo individual ocupe el menor ancho de banda posible.

Sin embargo, en el prototipo desarrollado únicamente el instructor está autorizado a utilizar audio y vídeo para transmitir información. Por tanto, sólo fluye un flujo de audio y otro de vídeo a través de la plataforma prototipo. Esto implica que, aún siendo un factor determinante el ancho de banda consumido por el prototipo, por las características de la red corporativa vistas en el apartado 2.4.2, ya no supone una decisión crítica la elección de un determinado codec con muy bajo ancho de banda, y otros factores ganan peso en la elección de uno u otro.

Una decisión a tomar es si el prototipo podrá utilizar varios codecs para codificar la señal de audio y vídeo. Por un lado, la posibilidad de soportar varios codecs favorece la compatibilidad del prototipo con otras herramientas de videoconferencia al tiempo que aumenta la flexibilidad en las comunicaciones, a costa de que el instructor configure de forma adecuada el codec a utilizar. Por el contrario, la utilización de un único codec redundante en una mayor optimización de la implementación y sencillez de configuración, pues el codec vendrá configurado correctamente por defecto.

En comparación con el vídeo, la información didáctica que se transmite a través del audio es mucho más importante. En condiciones de bajo ancho de banda o congestión de la red, es interesante priorizar un medio frente a otro, especialmente, teniendo en cuenta el elevado ancho de banda que consume el vídeo. Por tanto, el prototipo ofrece la posibilidad al instructor de dejar de transmitir uno u otro medio en cualquier momento durante la clase. Si las condiciones de la red mejoran, éste puede reanudar la transmisión del medio parado cuando lo considere oportuno.

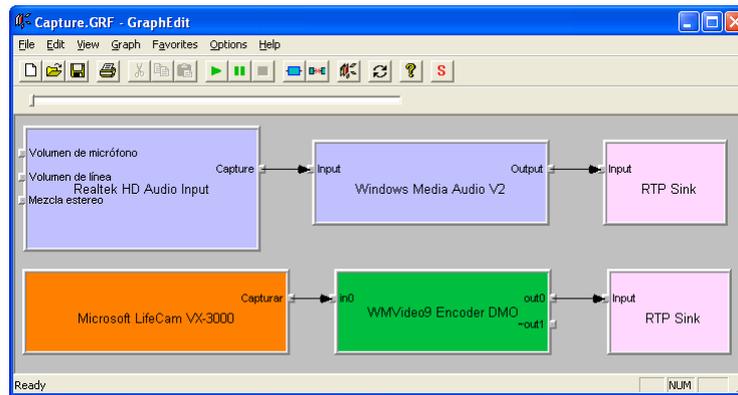


Figura 4.17: Grafo DirectShow para la transmisión mediante RTP de audio y vídeo

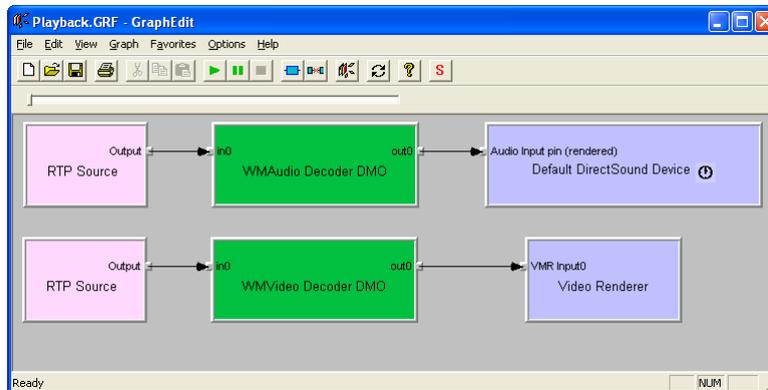


Figura 4.18: Grafo DirectShow para la reproducción de flujos RTP de audio y vídeo

4.6.1. Audio

El canal de audio se utiliza para transmitir de forma unidireccional, principalmente, el discurso pedagógico del instructor. Para ello, se utilizará un micrófono para capturar la voz del instructor. De esto puede deducirse que la gran mayor parte del audio que se transporta se corresponde con voz humana, luego esto tiene una implicación directa en el modo en que los datos pueden ser capturados y codificados. Específicamente, se nos plantean dos posibilidades:

- Codec de banda ancha: un codec de banda ancha está preparado para codificar cualquier tipo de señal de audio ofreciendo una alta fidelidad a costa de un mayor consumo de ancho de banda. Puede codificar señales de voz o música indistintamente.
- Codec de banda estrecha: este tipo de codec está destinado a codificar sólo la señal acústica de la voz humana, por lo que aprovecha características de este tipo de sonidos para optimizar el proceso de compresión.

4.6.2. Vídeo

El vídeo que se transmite desde el instructor hacia los alumnos apenas proporciona conocimiento pues, habitualmente, se trata del busto parlante del instructor mientras lleva a cabo su

explicación. Su principal utilidad es reforzar la sensación de presencia del instructor en los alumnos, para que éstos perciban la existencia del instructor con más fuerza que la que proporciona el control de presencia y la voz por sí solas.

De nuevo, la principal decisión a tomar durante el diseño del prototipo es el codec a utilizar para codificar la señal de vídeo. En este caso, la elección de uno u otro codec puede suponer una diferencia de anchos de banda consumidos en la red bastante significativa a diferencia de lo que ocurría con el audio.

4.6.3. Evaluación de alternativas y optimización

En la implantación de la funcionalidad de videoconferencia dentro del prototipo se pensó desde un primer momento en DirectShow frente a otras alternativas, pues constituye un entorno integrado en el que llevar a cabo la captura del audio y vídeo y su posterior codificación de forma uniforme. Además, la organización de los datos en flujos casa directamente con el concepto de flujo con el que trabaja RTP. Su modularidad a través de filtros que realizan operaciones sobre el flujo permiten ampliar las capacidades de la videoconferencia sin más que utilizar otros filtros diferentes a los que se usan actualmente.

Una vez fijado el marco de trabajo deben elegirse los codecs que se utilizan para codificar las señales de audio y vídeo para ser enviadas desde el instructor a los alumnos. Es especialmente interesante que los codecs presenten las siguientes características:

- **Baja latencia:** la latencia de codificación y decodificación debe ser lo más baja posible por los requerimientos de tiempo real de las comunicaciones.
- **Alta compresión:** cuanto mayor sea la capacidad de compresión de las señales de audio y vídeo, menor será el ancho de banda consumido en la red.
- **Robusto frente a pérdidas:** la red presenta pérdidas, por lo que el codec debe estar preparado para asumirlas sin provocar una merma excesiva de calidad.
- **Recursos mínimos:** el codec debe utilizar la menor cantidad de recursos posible para codificar y decodificar la señal.

Sin embargo, a partir de los requisitos que debe cumplir la herramienta junto con las características de la red y de los usuarios con los que trabaja, puede deducirse que estas características no tienen la misma importancia, con lo cual, puede plantearse una clasificación en función de su importancia, tal como se indica en la figura 4.19.

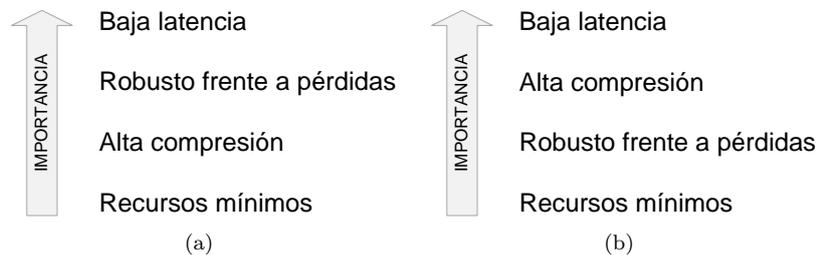


Figura 4.19: Características deseables de los codecs de (a) audio y (b) vídeo

Habitualmente, decodificar una señal de audio o vídeo requiere muchos menos recursos que codificarla. Cualquier equipo informático actual puede decodificar las señales de audio y vídeo en tiempo real sin demasiado esfuerzo. Esto es si cabe aún más cierto en el caso de la videoconferencia

por el tipo de señales que se manejan: vídeo de pequeñas dimensiones y voz humana. Luego el mayor problema que podría aparecer sería en la codificación de las señales en el equipo del instructor. Ahora bien, resulta factible sustituir el equipo del instructor por uno de mayor potencia si los requisitos computacionales necesarios no fueran correctamente satisfechos por el primero.

Motivado por la utilización de un modelo de interacción 1-N para la videoconferencia, únicamente circulan por la red un flujo de vídeo y otro de audio: los que provienen del instructor. De esta forma, el efecto sobre la red que impone la elección de uno u otro codec es poco significativo, aunque entre un codec y otro de vídeo la diferencia en ancho de banda puede llegar a ser notable.

La robustez del codec elegido frente a pérdidas de paquetes es importante, ya que el flujo que genera va a ser enviado por red y, especialmente en el caso de audio, interesa que la señal tenga la mejor calidad posible. Sin embargo, dadas las características de las redes corporativas, administradas por la misma organización, puede configurarse la red para que las pérdidas de paquetes sean mínimas, aunque pueden existir enlaces de red que todavía las presenten.

En cambio, las latencias para la codificación y la decodificación de las señales juegan un papel importante a la hora de sincronizar el audio y el vídeo con el resto de medios dentro del prototipo. Altas latencias supondrían que el audio y el vídeo del instructor no se reproducirían de forma síncrona con el resto de medios según imparte sus explicaciones.

Otra decisión a tomar es si el codec de audio elegido debe permitir la codificación de señales de banda ancha o sólo la voz humana (banda estrecha). Por norma general, un codec de banda ancha genera un flujo con una tasa de bits por segundo más alta que un codec de banda estrecha. Sin embargo, se ha visto que el poder de compresión de la señal no es de los factores más importantes de un codec para este prototipo. Por tanto, se puede considerar la utilización de un codec de banda ancha que permitiría codificar tanto señales de voz como música y otro tipo de sonidos.

En definitiva, la mejor solución para el prototipo a desarrollar es la utilización de un codec de vídeo de baja latencia y gran poder de compresión para que el ancho de banda consumido sea bajo y no afecte a otro tipo de flujos más importantes. Específicamente, se utiliza el codec VC-1 que ofrece una alta compresión del vídeo junto con una calidad aceptable de la imagen.

Para el caso del audio interesa, inicialmente, un codec de banda ancha que permita codificar cualquier tipo de señal de audio, no sólo la voz humana. De esta forma, el prototipo ofrece nuevas posibilidades como la transmisión de información sonora diferente al discurso del instructor. En concreto, para la implementación de la audioconferencia se utiliza el codec *Windows Media Audio V2*, similar al MP3, que ofrece en uno de sus modos un bajo retardo de codificación, clave para que la interacción entre los participantes de la clase se produzca en tiempo real.

Actualmente, se encuentra en desarrollo la funcionalidad necesaria para la utilización de dos codecs de audio de banda estrecha, como son iLBC y Speex, y un codec de vídeo de muy alta compresión, H.264, para permitir su transporte a través de la red utilizando RTP [Duric y Andersen, 2004], [Herlein *et al.*, 2008] y [Wenger *et al.*, 2005]. El uso de este tipo de codecs para las comunicaciones permitirá la utilización de la videoconferencia bajo un modelo de interacción N-N, pues las restricciones de la red podrán ser salvadas gracias al mínimo consumo de ancho de banda de estos codecs.

4.7. Procesamiento de medios: mensajería instantánea y control de presencia

La mensajería instantánea permite el intercambio de mensajes de texto entre los participantes de una clase virtual en tiempo real. Constituye el mecanismo básico a través del cual los alumnos proporcionan realimentación al instructor. Esta realimentación puede venir en forma de preguntas acerca de los contenidos de la clase, o como respuestas que hace el instructor a través del canal de audio. Ante una pregunta realizada por un alumno a través de un mensaje de texto, el proceder

habitual del instructor es hacer referencia a dicha pregunta para a continuación responderla. Gracias a la persistencia de las consultas realizadas mediante mensajes de texto, frente a lo efímero de una pregunta planteada de viva voz, el instructor puede optar por abordar la cuestión en el momento que lo estime más oportuno, porque ésta queda registrada en el panel de mensajes de la interfaz de usuario de todos los participantes de la clase virtual. Por su parte, el control de presencia permite conocer el estado de un participante dentro de la clase virtual (conectado, desconectado, ausente...).

La posibilidad de que un alumno se comunique con otro de forma privada a través de mensajes de texto debería evitarse. Esto podría conducir a que un cierto grupo de alumnos se distraiga del desarrollo de la clase. Además, la principal motivación de la mensajería instantánea es que los alumnos planteen preguntas al instructor; utilizando mensajes privados dos alumnos podrían plantear la misma cuestión al instructor de forma independiente. En consecuencia, es preferible utilizar un modelo de interacción N-N frente a uno N-1 (desde los alumnos al instructor), y que los mensajes de texto sean visibles a todos los participantes, de esta forma se pueden autolimitar las consultas al instructor; un alumno no planteará una pregunta que previamente haya realizado un compañero. La utilización de mensajes privados puede ser útil también para que el instructor le comunique a un alumno de forma individual cualquier incidencia, si bien lo mismo se consigue a través de una comunicación oral, aunque en este caso sin privacidad alguna.

Para llevar a cabo una solución que permita desarrollar conversaciones de texto entre los participantes de la clase virtual se nos plantean una serie de posibilidades. Una primera posibilidad es utilizar SIP, a través de sus extensiones SIMPLE, para permitir el intercambio de mensajes de texto entre los participantes. Sin embargo, esta solución está más pensada para el intercambio de mensajes entre dos participantes, con lo que las conversaciones multipunto son difíciles de llevar a cabo sin un elemento central que distribuya los mensajes entre todos los participantes de la clase virtual. De cualquier forma, en el caso de utilizar SIP para el control de la clase ese servidor intermedio sería la solución inmediata para la implementación de la mensajería instantánea, aunque es preferible la no utilización de servidores por la orientación *peer-to-peer* del diseño de la herramienta.

Otra de las alternativas posibles es utilizar un protocolo estándar específico para la mensajería instantánea como XMPP. De forma genérica, permite el transporte de documentos XML en tiempo real, con lo que la información de presencia, representada mediante XML, se puede incorporar fácilmente dentro del prototipo. Una de las principales desventajas que presenta esta solución para el ámbito concreto de aplicación del prototipo es la necesidad de un servidor específico que publique la información de presencia de todos los participantes. Esto chocaría con la aproximación *peer-to-peer* que se ha seguido en el desarrollo del prototipo completo, para que no sea imprescindible la figura de servidor central.

Tampoco IRC se adapta a los requisitos que se persiguen en el desarrollo del prototipo y descritos en el apartado 2.4.2, puesto que requiere la utilización de un servidor específico para centralizar las comunicaciones de texto entre los participantes.

4.7.1. Evaluación de alternativas y optimización

Uno de los principales requisitos que se busca satisfacer con el prototipo es la simplicidad de la herramienta. Esta simplicidad abarca desde el fácil manejo de la herramienta a través de la interfaz de usuario hasta la facilidad de gestión de la plataforma. Esto último implica una aproximación *peer-to-peer* donde no sean imprescindibles servidores. En todo caso, estos servidores se utilizarían para habilitar el transporte de los datos fuera de la red corporativa y de forma transparente a la herramienta. Por tanto, se desechan aquellas alternativas que implican la utilización de un servidor para la centralización de las conversaciones de texto.

La solución óptima propuesta pasa por la implementación utilizando la recomendación T.140 de la ITU-T que define un formato para los mensajes de texto y su transporte a través de RTP,

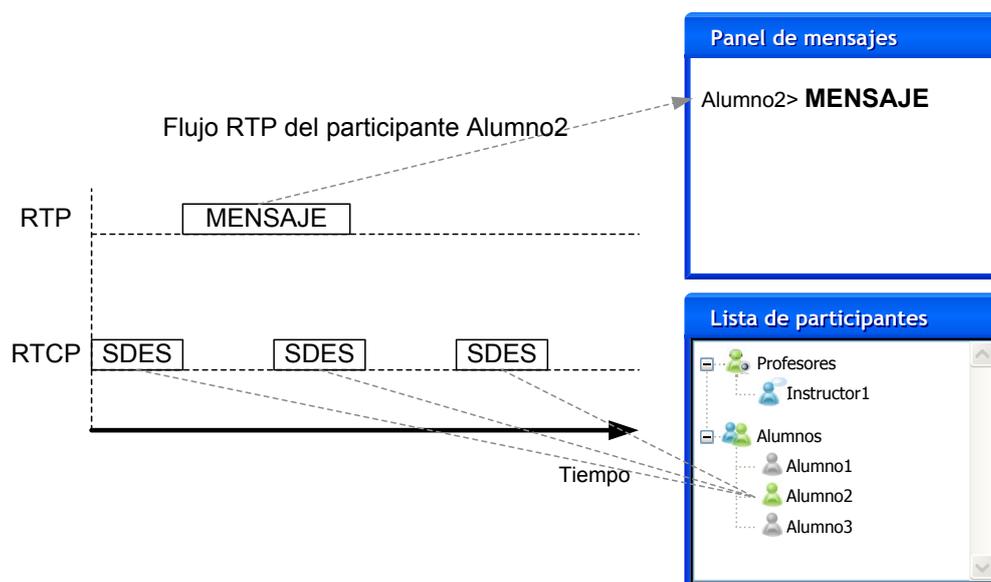


Figura 4.20: Implementación de la mensajería instantánea y control de presencia con RTP/RTCP

al igual que el resto de medios, utilizando el estándar definido por la IETF [Hellstrom y Jones, 2005]. Los mensajes de texto se transportarían a través del protocolo de transporte RTP, mientras que a través del protocolo de control RTCP se puede construir un control de presencia bastante simple, que nos permite determinar qué usuarios están conectados. Si bien esta solución no ofrece todas las características de control de presencia que ofrecen otras alternativas como SIP o XMPP (ausente, ocupado, no disponible, etc.), únicamente es necesario conocer el estado de conexión de cada usuario, puesto que la disposición o no a participar en la clase se presupone porque las clases virtuales se imparten a empleados durante su jornada laboral. La figura 4.20 muestra la utilización de RTP para permitir la mensajería instantánea dentro del prototipo.

Cada uno de los participantes que desee comunicarse con el resto de la clase utilizando mensajes de texto debe unirse a la sesión RTP correspondiente. En ese momento, comenzará a recibir periódicamente paquetes RTCP que identifican a cada uno de los participantes que está conectado a dicha sesión. Suponiendo que cualquier usuario que participe en la clase virtual debe hacerlo, al menos, a través de la mensajería instantánea, resulta en que los participantes de esta sesión RTP se corresponden con todos los participantes de la clase virtual. Se pueden utilizar las extensiones de RTCP para transmitir información más específica de cada uno de los usuarios como su nombre, teléfono, ubicación, una nota (que puede utilizarse como una notificación de que el usuario no está en su equipo), etc. Además, se pueden añadir extensiones propias de la aplicación, como por ejemplo la identificación del rol de cada participante; así, es posible distinguir a los instructores del resto de participantes de la clase virtual. Por otra parte, los mensajes de texto se organizan en paquetes de datos que constituyen un flujo RTP que se envían utilizando técnicas de difusión multicast, por lo que llegarán a todos los participantes de la clase virtual.

El gran inconveniente de esta solución es que, además de no ofrecer el gran abanico de posibilidades que permiten soluciones específicas (imágenes personales, estados de presencia complejos, emoticonos...), no es posible establecer comunicaciones privadas entre los participantes.

4.8. Procesamiento de medios: pizarra virtual

La pizarra virtual es una funcionalidad que ofrece el prototipo para que el instructor pueda exponer contenidos didácticos al resto de la clase, realizar anotaciones sobre dichos contenidos y apuntar elementos mediante un puntero virtual o telepuntero; estas dos últimas opciones, como se justifica en el apartado 4.1.1, también es conveniente que las puedan llevar a cabo los alumnos para ofrecer cierta información de realimentación al instructor.

De nuevo, al igual que ocurre con el audio y el vídeo, aún siendo posible aplicar un modelo de interacción N-N (colaborativo) entre todos los participantes de la clase para la pizarra virtual, se limita esta posibilidad modificando la interfaz de usuario para el caso de los alumnos, para que éstos no puedan compartir contenidos; ésta es una característica que se le ofrece sólo al instructor. En cambio, a los alumnos se les permite realizar anotaciones y utilizar el puntero virtual de la misma forma que lo utiliza el instructor, es decir, estos dos últimos medios se utilizan bajo un modelo de interacción N-N. Cuando el instructor decida cerrar el área de trabajo, bien sea porque considera finalizada la exposición de la información que contiene, o bien porque da por terminada la clase, se cerrará automáticamente al mismo tiempo en el resto de alumnos.

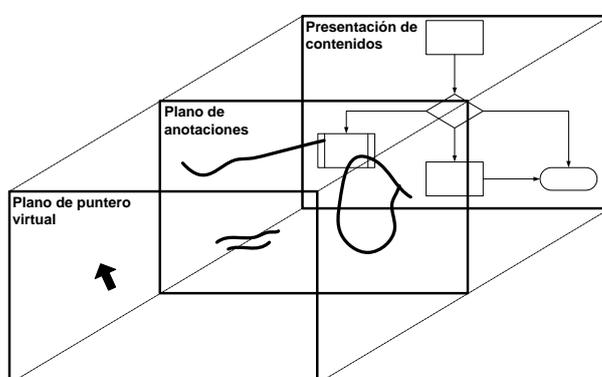


Figura 4.21: Organización en capas de un área de trabajo

En la implementación de la pizarra virtual se sigue una aproximación similar a la que se propone en la recomendación T.126 de la ITU-T. Esta solución consiste en organizar la pizarra en diferentes áreas de trabajo compuestas de capas, también denominadas planos, cada una de ellas con una determinada funcionalidad. En la figura 4.21 se muestra una representación de un área de trabajo en la recomendación T.126. La capa más inferior (fondo) está destinada a la presentación de los contenidos, y por tanto, se utiliza para la exposición de las diapositivas de fondo. Sobre esta primera capa, se disponen una o varias capas destinadas a renderizar las anotaciones que realizan el instructor y los alumnos sobre los contenidos que se muestran de fondo. Un tercer tipo de capas son las que se sitúan por encima de estas últimas, cuyo fin es el de contener los punteros virtuales de los usuarios para que éstos puedan apuntar elementos concretos dentro del área de trabajo. En un momento dado, la pizarra puede tener abiertas varias áreas de trabajo, si bien sólo una puede ser la activa.

Cuando existe más de una capa por nivel, esto es, dos o más capas de anotaciones o de puntero virtual, éstas se organizan por niveles. Las capas de anotaciones se sitúan siempre encima de la capa de presentación de contenidos, ordenándose en base a un índice Z, que determina qué capas estarán en primer plano. Lo mismo ocurre para las capas de los punteros virtuales; éstas se disponen delante de las capas de anotaciones y entre ellas se ordenan utilizando un índice Z.

Sobre la pizarra virtual el usuario puede realizar diferentes acciones. Además de anotar o utilizar el puntero virtual, también es posible avanzar o retroceder páginas, ir a una página

determinada, incrementar el factor de zoom con el que se dibujan los contenidos de la pizarra (incluyendo anotaciones) y desplazarse a través de los contenidos cuando el zoom es mayor que el 100%. En este último caso, la vista que obtendrá el participante será sólo una parte del área de trabajo completa, quedando el resto oculta. Para facilitar el proceso de impartición de la clase por parte del instructor, y para que en todo momento pueda controlar la forma en que los contenidos son expuestos a los alumnos, la vista actual del área de trabajo no es específica de cada participante, sino que es compartida por todos los asistentes a la clase virtual y está determinada por el instructor. De esta forma, si el instructor aplica un factor de zoom para resaltar una zona concreta de la diapositiva actual y sólo visualiza una parte del área de trabajo, se garantiza así que el resto de participantes de la clase están visualizando el área de trabajo de la misma forma que el instructor.

Todos los participantes dentro de un área de trabajo comparten el mismo sistema de coordenadas para expresar la posición de todos los elementos que se intercambian entre sí. Esto es aplicable tanto para las anotaciones que se realizan sobre el área de trabajo como para los desplazamientos del puntero virtual sobre los contenidos. Debido a las diferentes resoluciones de pantalla en los equipos de los participantes de la clase, las dimensiones de la zona en la que debe renderizarse el área de trabajo puede variar entre unos participantes otros, así que el área de trabajo debe ser escalada para ajustarse a esta zona de representación, pero siempre manteniendo su relación de aspecto (*aspect ratio*), es decir, la relación entre el ancho y el alto original. Esta relación viene determinada por las dimensiones de las diapositivas que se muestran en la capa de contenidos. En la figura 4.22 aparece un resumen del proceso de escalado para la renderización de un área de trabajo.

Tanto el área de trabajo como la zona donde ésta debe ser renderizada tienen unas determinadas dimensiones (x_W e y_W , para el área de trabajo, y x_R e y_R , para el área de renderizado) y relaciones de aspecto (AR_W , el área de trabajo, y AR_R , el área de renderizado). Si la relación de aspecto del área de trabajo es mayor que la correspondiente al área de renderizado, el ancho de la primera (x'_W) se establece para abarcar todo el ancho del área de renderizado, mientras que el alto (y'_W) se ajusta de acuerdo a la relación de aspecto del área de trabajo para evitar deformar su representación. En este caso, aparecen dos franjas horizontales de igual tamaño en la parte inferior y superior del área de renderizado, como se aprecia en la parte izquierda de la figura. Si, por contra, la relación de aspecto del área de trabajo es menor que la relación de aspecto del área de renderizado, entonces el área de trabajo se ajusta a la altura de esta última y la anchura se calcula para mantener la relación de aspecto, con lo que aparecerán dos franjas verticales a ambos lados en el área de renderizado, como se ve en la parte derecha de la misma figura. A continuación, si en cualquiera de los dos casos se aplica un factor de zoom (z), es posible que el área de renderizado no abarque el área de trabajo completa, cuyas nuevas dimensiones (x''_W e y''_W) se habrían modificado proporcionalmente con el factor de zoom. Por último, si el factor de zoom es mayor que 1 (100%), la vista del área de trabajo, compartida por todos los participantes, se calcula a partir de la posición de la esquina superior izquierda del área de renderizado (x_O , y_O) en relación con la posición de la misma esquina dentro del área de trabajo, que coincide con el origen de coordenadas, resaltado en la figura 4.22 con una estrella.

Para el transporte de los datos con los que trabaja la pizarra, igual que para el resto de medios, se utiliza RTP a pesar de que no está originalmente pensado para el transporte de este tipo de medios, ya que ofrece unas características muy interesantes, descritas en el apartado 4.2. Un problema que aparece por la utilización RTP sobre UDP para el transporte de los datos de la pizarra es la posibilidad de que los paquetes se pierdan. Esta situación, que para un medio continuo como el audio o el vídeo es asumible, para medios discretos es inaceptable. La pérdida de un paquete correspondiente a una diapositiva o a un trazo supone irremediablemente la pérdida de la diapositiva o el trazo concretos, incluso, dependiendo del formato de los contenidos, es posible que implique la pérdida de toda la presentación. Para mitigar estos problemas, se pueden utilizar técnicas de corrección de errores tipo FEC (*Forward Error Correction*) [Watson *et al.*,

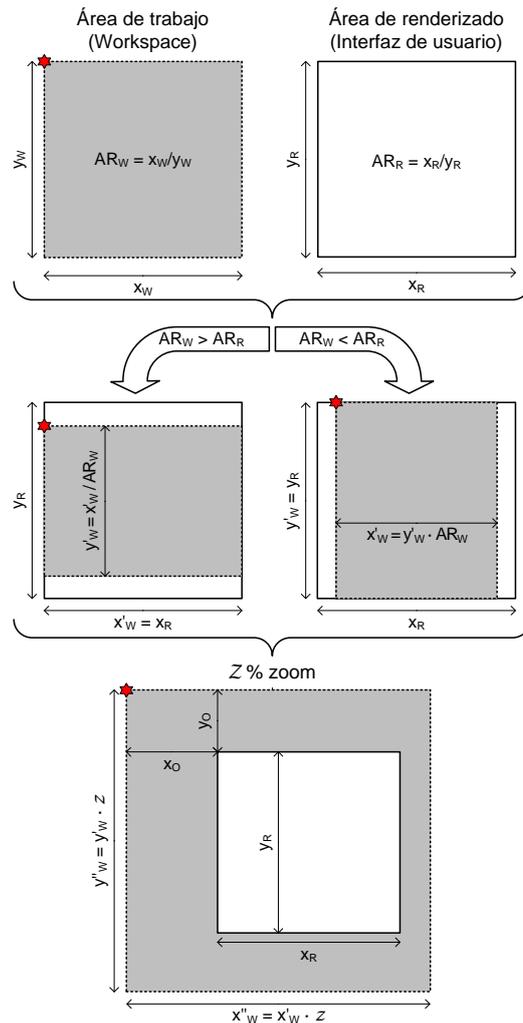


Figura 4.22: Renderización de un área de trabajo de la pizarra virtual

2007], muy útiles en entornos de tiempo real o situaciones en que no se admite señal de retorno (interacción 1-N). Estas técnicas consisten en añadir redundancia a los datos que se envían para detectar y corregir errores sin la necesidad de solicitar un reenvío de los mismos. Si bien cuanto más redundancia se incluya supone mayor capacidad de recuperación frente a errores, implica un mayor consumo de ancho de banda en la red así como más retardo en la recepción de los paquetes. En [Li, 2007] se propone un formato de empaquetamiento de datos que utilicen FEC sobre RTP.

Otro de los problemas más complicados a los que tiene que hacer frente una herramienta de pizarra virtual es el asociado al *late-join*, esto es, participantes que se unen con la clase iniciada. Para que estos usuarios puedan participar en la clase igual que el resto, deben recibir lo más pronto posible información que les sitúe en el mismo estado que los demás participantes. Este problema se palia en parte gracias a la utilización de RTP para transportar los datos de la pizarra. RTCP, el protocolo de control de RTP, envía periódicamente paquetes de control. Junto a estos paquetes de control, organizados en un mismo paquete RTCP compuesto, es posible

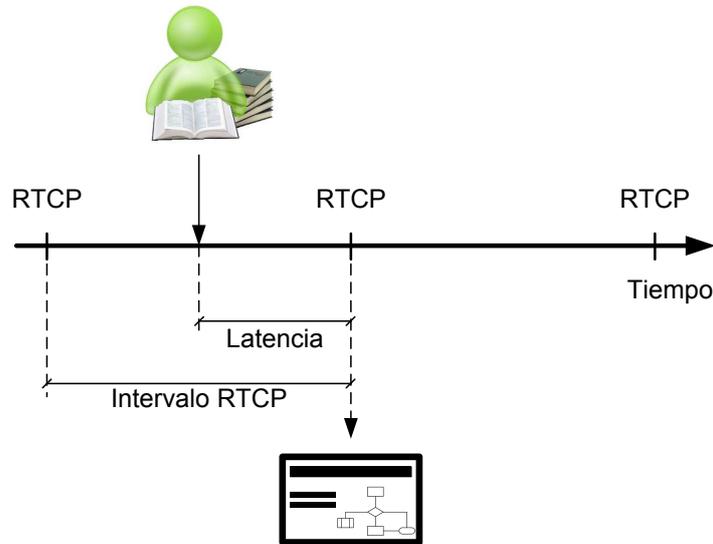


Figura 4.23: Latencia durante la unión a la clase de un nuevo participante

adjuntar información específica de la aplicación. De esta forma, se pueden utilizar estos paquetes periódicos para adjuntar información que permita recuperar el estado actual de la pizarra; así, tan pronto como un alumno se conecte y reciba el primer paquete RTCP compuesto conocerá el estado actual de la pizarra virtual como se representa en la figura 4.23.

4.8.1. Evaluación de alternativas y optimización

Los datos de la pizarra virtual se transportan utilizando RTP. No obstante, debido a las características de la pizarra virtual y su clara división en tres tipos de elementos (contenidos, anotaciones y punteros) independientes, se puede optar por dos alternativas:

- Un sólo medio: bajo esta opción, se considera la pizarra como un único medio. De esta manera, los datos que genere un usuario dentro de la misma área de trabajo (contenidos, anotaciones y puntero virtual) pertenecen al mismo flujo de datos. Sin embargo, esta aproximación choca con la separación de los distintos tipos de datos que recomienda el estándar RTP [Schulzrinne *et al.*, 2003a], pues esto impediría, por ejemplo, realizar un tratamiento específico a cada uno de los datos que componen la pizarra virtual como se comentó en el apartado 3.2.4.1.
- Múltiples medios: en este caso, los datos que maneja la pizarra se organizan en medios independientes, del mismo modo que la videoconferencia está compuesta de dos medios independientes: audio y vídeo. Así, existen tres medios relacionados: presentación de contenidos, anotaciones y puntero virtual. Cada uno de estos medios representa un tipo de datos distinto y, al transportarse mediante sesiones RTP independientes, permite procesar cada uno de ellos de forma separada, luego no es necesario recibir los datos de los tres medios para poder participar en la clase. Esta solución permite que un participante intervenga en la clase virtual a través de la presentación de contenidos, sin utilizar anotaciones ni puntero virtual; se consigue así que la pizarra virtual sea configurable. En todo caso, aunque se puede evitar la utilización de anotaciones y puntero virtual, la presentación de contenidos es

un requisito imprescindible para poder trabajar con la pizarra como se verá en el siguiente apartado.

En este caso, se deduce que la solución óptima para la implementación dentro del prototipo es dividir las tres partes funcionales de la pizarra (contenidos, anotaciones y puntero virtual) en diferentes medios, cada uno transportado en su sesión RTP independiente.

Por otra parte, a pesar del gran detalle de la recomendación T.126, no especifica ni limita el número de capas de anotaciones o de puntero virtual que pueden situarse por encima de la capa de presentación de contenidos que es única dentro del área de trabajo. Ante esta situación, se nos plantean tres posibilidades:

- Una única capa: de esta forma todas las anotaciones o todos los punteros virtuales se representarán en la misma capa. Esto implicaría que cada uno de los trazos debe estar etiquetado con el participante que lo realizó para identificar los trazos que pertenecen a un determinado usuario.
- Tantas capas como usuarios: en este caso se dispone de una capa independiente para cada uno de los usuarios que participa en la sesión. Así, las anotaciones no deben ser etiquetadas con el participante, pues la capa en la que se renderizan ya determina el participante al que pertenecen todos los trazos que contiene.
- Agrupar en capas: esta solución consistiría en agrupar las anotaciones/puntero virtual de varios participantes dentro de la misma capa. De algún modo sería una solución intermedia a las dos anteriores, que combinaría lo mejor de cada una de ellas.

Inicialmente, si se utilizan múltiples capas de anotaciones, una capa para cada participante de la clase virtual, supone las siguientes ventajas respecto a una solución de una única capa:

- Transporte uniforme con el resto de medios: gracias al uso de RTP para el transporte de las anotaciones, éste se realiza al igual que el resto de medios pero de forma independiente, por lo que es posible un tratamiento común, por ejemplo priorizarlos respecto a otros tráficos en la red, o distinto, dado que cada uno de ellos utiliza su propia sesión RTP. Además, sigue la filosofía propia de RTP de asignar un flujo por participante; en este caso, se va un paso más allá y se hace un mapeo directo entre flujo y capa de anotaciones.
- Ocultamiento de los trazos: en un determinado momento, un participante puede desear ocultar los trazos realizados por un participante para tener mejor visibilidad de los contenidos de la pizarra. Con una situación en la que las anotaciones se agrupan en capas por participante, la operación de ocultar los trazos de un participante tiene una complejidad $O(1)$ independientemente del número de trazos que haya realizado, puesto que la operación consiste en marcar la capa como no visible. En cambio, una solución de una capa compartida en la que cada trazo está etiquetado con el participante al que pertenece, supondría una complejidad $O(n)$, pues es necesario comprobar todos los trazos para marcarlos individualmente como no visibles. Esto puede llegar a ser costoso en clase virtuales con alto número de participantes y gran nivel de interacción mediante anotaciones.
- Menor ancho de banda: dado que los trazos no deben ser etiquetados de forma individual con el participante al que pertenecen, ya que éste se deriva de la capa a la que pertenecen los trazos. Esto supone un menor consumo de ancho de banda puesto que no es necesario enviar un identificador de usuario para cada uno de los trazos.

Por contra, una capa de anotaciones por participante tiene varias desventajas respecto a una capa de anotaciones global única para todos los participantes:

- Borrado de anotaciones ajenas: mediante esta solución tal cual, sin ningún tipo de mecanismo adicional, es imposible que un participante elimine un trazo que pertenezca a otro participante. Cada capa o plano de anotaciones pertenece a un único usuario y sólo este puede añadir o eliminar trazos a dicha capa. Incluso en el caso del instructor, éste no puede eliminar anotaciones que hayan realizado los alumnos. Para mitigar este inconveniente se ofrece la posibilidad de ocultar las anotaciones de dicho participante, no obstante, esto implica el ocultamiento de todos sus trazos.
- Más recursos: como los trazos se reparten entre varias estructuras de datos que representan las capas, los recursos necesarios para implementar esta solución son mayores que una solución basada en una única capa.

De esto se deduce que la solución óptima para el prototipo es la utilización de varias capas en el caso de las anotaciones, una por participante. Habitualmente el número de participantes en una clase virtual será bajo, por lo que no habrá problemas de recursos para gestionar las capas. Además, en raras ocasiones es necesario borrar trazos ajenos y, en última instancia, se puede instar al propietario a eliminar sus anotaciones.

Para el caso del puntero virtual se adopta una solución similar, más apropiada si cabe por tratarse de un medio continuo, puesto que el flujo que genera cada usuario se reproduce de forma independiente. En este tipo de medios, la relación entre flujo de datos y flujo RTP es directa al igual que ocurre con el audio y el vídeo.

Para el transporte de los medios discretos de la pizarra (contenidos y anotaciones), debido a su especial sensibilidad frente a las pérdidas de paquetes, se pueden utilizar técnicas FEC de corrección de errores para evitar las pérdidas de paquetes que pueden resultar fatales. No ocurre lo mismo para el puntero virtual, ya que al tratarse de un medio continuo la pérdida de un paquete implica una degradación de la señal, representada por saltos en los movimientos del puntero, pero no representaría un error crítico. Sin embargo, las pruebas realizadas sobre la red corporativa pusieron de manifiesto la ausencia de pérdidas de paquetes dentro de la red corporativa, luego esta funcionalidad puede obviarse en el principal escenario de aplicación del prototipo.

En conclusión, el transporte de los medios que componen la pizarra virtual y su traducción a elementos constitutivos de las áreas de trabajo se realiza de acuerdo a la figura 4.24. En los siguientes apartados se detalla más en concreto el diseño de cada uno de los medios que componen la pizarra virtual.

4.8.2. Presentación de contenidos

La presentación de contenidos se utiliza para que el instructor comparta materiales didácticos con el resto de la clase del mismo modo en el que un cañón proyector se utiliza en un aula tradicional para exponer los apuntes de clase. De esta forma, la presentación de contenidos en una herramienta de e-learning síncrono utiliza la metáfora de una pantalla de proyección para presentar los contenidos didácticos. Por tanto es una de las funcionalidades más importantes de una herramienta de educación a distancia, pues permite presentar la mayor parte del material sobre el que los participantes de la clase trabajan.

Como se ha visto anteriormente, la pizarra virtual se organiza en áreas de trabajo, cada una de las cuales está compuesta de una serie de capas. La presentación de contenidos se realiza dentro de estas áreas de trabajo de la pizarra virtual. Concretamente, el documento a compartir se visualiza en la capa de presentación de contenidos que se utiliza de fondo. Por tanto, si un instructor desea compartir más de un documento con el resto de la clase, deberá crear tantas áreas de trabajo como documentos quiera compartir.

Cada una de las áreas de trabajo de la pizarra puede contener una única capa o plano de presentación de contenidos, por lo que existe una relación uno a uno entre un plano de presentación

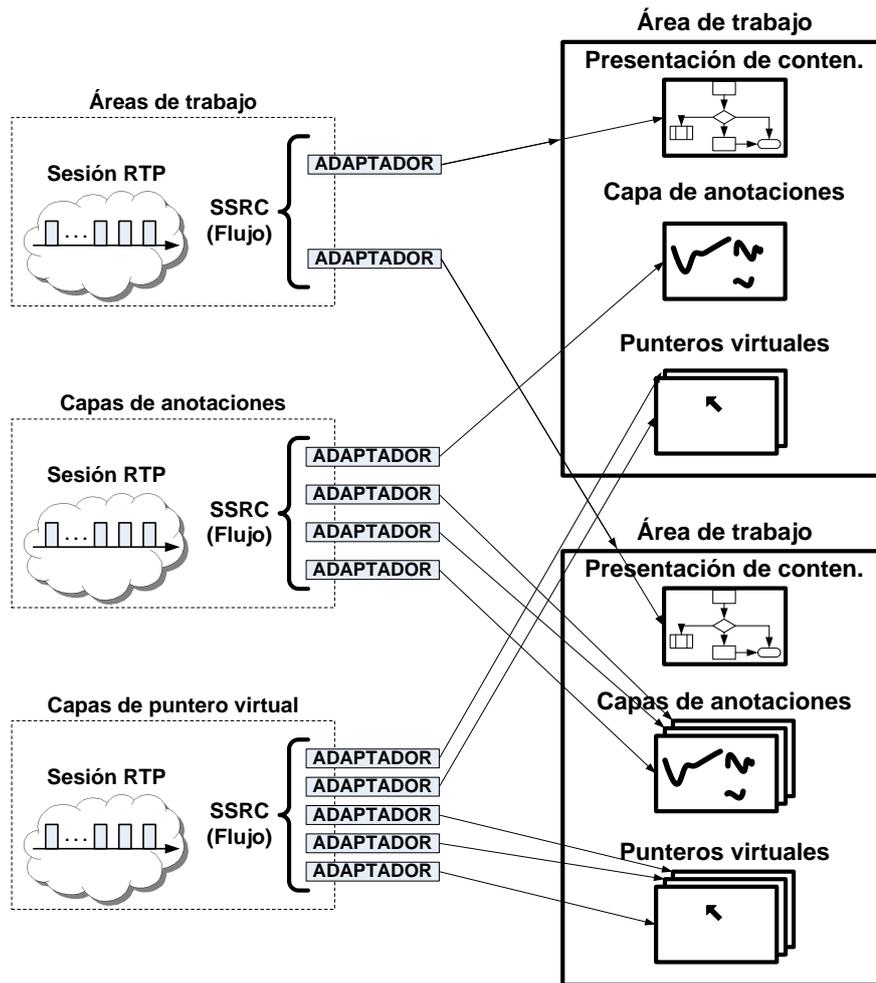


Figura 4.24: Mapeo de las estructuras de un área de trabajo en la red

de contenidos y el área de trabajo al que pertenece. En consecuencia, se puede identificar una presentación de contenidos con un área de trabajo, luego el tiempo de vida del área de trabajo es el tiempo de vida de este plano; la finalización del plano de presentación de contenidos de un área de trabajo implica la inmediata eliminación de esta última y de todas las estructuras de datos que contiene, incluyendo todas las capas que lo componen. El mayor inconveniente de este esquema es la imposibilidad para que un instructor ceda el control de un área de trabajo a otro instructor o un alumno. Un área de trabajo se relaciona con un flujo RTP y, por extensión, con un SSRC, el cual hará las veces de identificador del área de trabajo, ya que este identificador es único a lo largo de la sesión. Esto implica que un área de trabajo estará vinculada indefinidamente a un usuario, ya que un área de trabajo creado por otro participante siempre tendrá un SSRC distinto, pues los identificadores de fuente (SSRC) deben ser únicos dentro de la misma sesión RTP.

Aunque esta funcionalidad está pensada para la compartición de contenidos didácticos, también se puede emular una pizarra compartida sobre la que todos los participantes puedan trabajar de forma colaborativa sin más que crear una capa de presentación de contenidos en blanco, esto es, sin compartir documento alguno.

4.8.2.1. Evaluación de alternativas y optimización

Para que el prototipo lleve a cabo la presentación de contenidos es necesario tomar unas decisiones durante su fase de diseño. Las decisiones más importantes son las relacionadas con:

- Representación interna de los contenidos: cómo representar internamente los documentos que se comparten entre los usuarios.
- Difusión de los contenidos: cómo se transportan los contenidos desde el instructor a los alumnos.
- Navegación por los contenidos: cómo pueden los usuarios navegar a través de los contenidos de la pizarra.

En los siguientes apartados se analizan en detalle estas tres cuestiones planteando las posibles alternativas de solución.

4.8.2.1.1. Representación interna de los contenidos Una las primeras decisiones que debe tomarse es la forma con la que el prototipo maneja los documentos que los participantes en la clase virtual comparten a través de la pizarra. Existen varias alternativas para representar este tipo de información:

- Streaming de vídeo: algunas herramientas analizadas, especialmente dentro del ámbito de investigación, utilizan técnicas de streaming de vídeo para hacer llegar contenidos a los alumnos. De esta forma, representan los contenidos didácticos de forma interna como flujos de vídeo y, por tanto, su tratamiento no difiere de cómo se procesa el vídeo que transporta el busto del instructor. Los contenidos se proyectarían de la forma habitual, por ejemplo utilizando un cañón proyector, y se capturarían de la misma forma que se captura el vídeo que se utiliza en la videoconferencia.

A pesar de todo, esta solución es poco flexible puesto que no permite a los alumnos realizar anotaciones sobre los contenidos, con lo cual es muy difícil poder utilizar las anotaciones bajo un modelo de interacción colaborativo. Además, para que los detalles de los contenidos puedan ser visualizados con una mínima calidad, la resolución del vídeo debería ser más alta que la que habitualmente se utiliza en sistemas de videoconferencia, lo que repercutiría en un mayor consumo de ancho de banda del prototipo.

- Compartición de aplicaciones: esta alternativa sigue una filosofía *si no puedes con ellos, únete a ellos*, esto es, busca la colaboración con las herramientas que manejan los formatos de documentos utilizados durante la clase para renderizar esta información. Por ejemplo, en lugar de soportar de forma nativa documentos PowerPoint, es posible que el instructor comparta su aplicación Microsoft PowerPoint para que los alumnos visualicen la presentación de diapositivas.

Esto tiene como ventaja la posibilidad de ampliar en el futuro el número y tipos de formatos que permite el prototipo; sólo es necesario que exista una aplicación que los renderice. Sin embargo, esto presenta problemas a la hora de combinar los contenidos con las anotaciones realizadas sobre ellos, pues sólo las aplicaciones pueden modificar sus propios formatos.

- Soporte nativo: en este caso, el prototipo debe trabajar directamente con los formatos de documento en los que se proporcionan los contenidos didácticos. Dada la gran cantidad de formatos en los que están disponibles dichos contenidos, es muy improbable que una herramienta de este tipo pueda manejarlos todos, por lo que bastaría manejar un subconjunto de ellos, los más comunes, para que la mayor parte del material pudiera ser utilizado directamente, o bien existiera algún tipo de conversión para adaptar los contenidos a estos formatos. Por consiguiente, si se elige esta opción habría, además, que decidir si:

- Utilizar uno o varios formatos de forma nativa: aunque es una solución más flexible, dificulta el desarrollo del prototipo.
- Qué formato o formatos utilizar: deberían tenerse en cuenta las características ideales de un formato para el e-learning sincrónico vistas en el apartado 3.3.2.1.

Esta solución es la más flexible de todas pues nos permite acceder a la representación interna de los contenidos para representarlos y modificarlos, posibilitando así la combinación de trazos y contenidos. Asimismo, los contenidos serán enviados una única vez, a diferencia de lo que ocurren en las dos alternativas anteriores, en que el instructor está enviando continuamente información del estado de los contenidos.

Otra posibilidad viable para representar los contenidos es mediante su previa conversión a un formato intermedio no específico de ningún tipo de documentos. En [Pahud, 2008] se utilizan contenidos basados en representaciones gráficas vectoriales de cada una de las páginas de los documentos imprimibles, junto con una serie de metadatos que permiten acceder a informaciones del documento tales como el título, el creador, el idioma, etc. Este formato es utilizado para renderizar presentaciones PowerPoint convertidas utilizando la propia aplicación Microsoft PowerPoint y almacenadas como una sucesión de imágenes vectoriales.

En una primera versión del prototipo se implementó la presentación de contenidos a través del estándar SVG exclusivamente, que por tratarse de un estándar abierto, ofrece posibilidades de compatibilidad con soluciones de creación de contenidos, editores gráficos, etc. Además, tratándose de un formato de imagen vectorial, al reducido número de bits para representar una imagen hay que unirle la posibilidad de convertir cualquier formato imprimible a este formato, lo que lo hacía bastante indicado. En esta versión, el envío de las páginas que componían el documento a compartir debían ser convertidas previamente a formato SVG, típicamente utilizando una impresora virtual. A partir de la experiencia, se pusieron de manifiesto una serie de cuestiones:

- Poca fidelidad de los conversores SVG: la mayor parte de los conversores a SVG, la gran mayoría distribuidos como impresoras virtuales, realizan una conversión deficiente de los ficheros PowerPoint, formato mayoritario para la representación de contenidos. Esto suponía en ocasiones la imposibilidad de entender los contenidos de las presentaciones.
- Bitmaps en SVG: todos los conversores SVG tienen problemas a la hora de trabajar con formatos de datos que almacenan imágenes de mapas de bits. El proceso de conversión genera ficheros de texto, del orden de cientos de kilobytes, cuyo tamaño es incluso mayor que el necesario para almacenar el mapa de bits original. Esto repercutía en el ancho de banda necesario para transmitir estos documentos por la red.

Por tanto, se declinó la utilización de SVG como formato único de representación. En su lugar, se utiliza como formato principal una representación de las páginas de los contenidos basadas en imágenes vectoriales *Compressed Windows Enhanced Metafile* (EMZ) las cuales permiten, además de trabajar con cualquier formato imprimible, fundir junto con los contenidos las anotaciones que realizan los usuarios. Además, el tamaño de los ficheros resultantes de la conversión a imágenes de este tipo es relativamente pequeño en comparación con los resultados obtenidos con SVG.

4.8.2.1.2. Difusión de los contenidos Una vez descritas las alternativas para trabajar internamente con los contenidos, es necesario considerar cómo deben ser distribuidos entre todos los miembros de la clase virtual, para lo cual existen dos posibilidades:

- Bajo demanda: los contenidos se distribuirán desde el instructor a los alumnos cuando éste decida compartirlos con el resto de la clase. Esto supone un tiempo inicial desde que el instructor comparte los contenidos hasta que realmente puede trabajar sobre ellos. Este

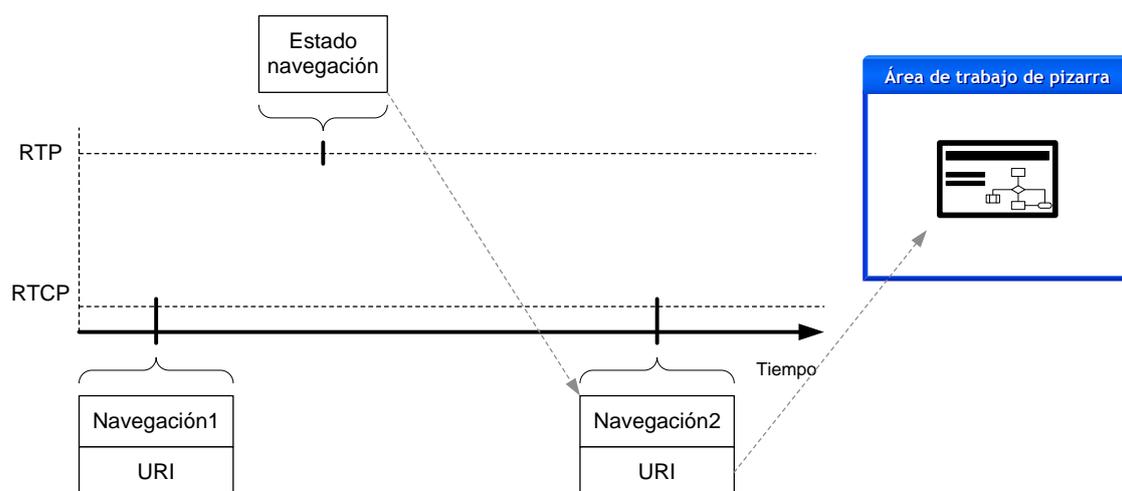


Figura 4.25: Implementación de la presentación de contenidos con RTP

tiempo puede reducirse en parte si el formato con el que se representan estos contenidos es particionable, es decir, si se pueden renderizar páginas sin que se haya recibido completamente el documento (constituiría un pseudo-streaming de documentos). Sin embargo, si no es factible esta posibilidad, esta latencia de inicio puede ser bastante elevada, especialmente si el tamaño del documento es grande. Además, el transporte de este tipo de datos a través de UDP, sin un mecanismo adecuado para repartir el ancho de banda disponible con otros tráficos, puede afectar negativamente a estos últimos. Asimismo, es necesario utilizar alguna técnica de corrección de errores para paliar el efecto de las pérdidas de paquetes, que para este tipo de información resultan fatales.

- En destino: supone que los alumnos deben descargarse los contenidos previamente al comienzo de la clase, típicamente, desde el portal web corporativo. Esto puede evitar latencias al comienzo de la clase puesto que los alumnos, ante una clase planificada, pueden descargarse los contenidos con la suficiente antelación para que no retrase el comienzo de la misma. Entonces, la presentación de contenidos se limitaría a un control remoto de cómo éstos se exponen a los usuarios, sin que el instructor deba hacer llegar los contenidos a los alumnos. Un alumno que no se hubiera descargado los contenidos previamente, no podría participar en la clase utilizando la pizarra virtual. También se evitan los problemas derivados del transporte de estos contenidos durante la impartición de la clase virtual, se ahorra ancho de banda, muy valioso en las redes corporativas como se explicó en el apartado 2.4.2, y el efecto sobre otros flujos, no siendo ya necesario implementar un mecanismo TCP-friendly para repartir de forma justa el ancho de banda disponible entre todos los tráficos.

Cuando se utiliza la opción de medios en destino, los contenidos educativos residen en ficheros que los alumnos deben descargarse en un directorio accesible por el prototipo. Para identificar el recurso que se utiliza como documento a visualizar de fondo en la pizarra virtual, se utilizan las características de RTCP para adjuntar al paquete compuesto que se envía periódicamente la URI que apunta al documento en cuestión. Podría utilizarse un código hash para identificar los ficheros independientemente del nombre que tengan. En la figura 4.25 se representa este esquema de funcionamiento.

El mayor inconveniente de esta técnica es que es necesario tener descargado todo el documento a visualizar independientemente del número de páginas del mismo que vayan a

utilizarse durante la clase virtual. Por ejemplo, el instructor puede estar interesado en mostrar a los alumnos una página concreta de un documento porque contenga una figura relacionada con el tema que se esté tratando. En ese caso, los alumnos deberían descargarse todo el documento que contiene esa figura previamente. Esto puede paliarse si el instructor planifica con antelación los materiales que utiliza durante la clase virtual.

En una de las versiones del prototipo se utilizó la técnica de difusión de los contenidos bajo demanda dentro de la red corporativa de ArcelorMittal. En esta versión, los contenidos SVG se enviaban desde el instructor hasta los alumnos página a página a través de RTP cuando se producía un cambio en la página actual. Se utilizaban técnicas FEC para hacer frente al problema de pérdida de paquetes. De esta experiencia se pudieron extraer las siguientes conclusiones:

- Picos de ancho de banda: cuando se navegaba de una página a otra el prototipo enviaba a través de RTP e IP Multicast la página destino a todos los participantes de la sesión, incluso aunque ya hubiera sido visitada. La mayor parte de las veces, estas páginas contenían mapas de bits que requerían un gran ancho de banda para su transmisión. Esto, ante la ausencia de un mecanismo TCP-friendly que repartiera equitativamente el ancho de banda disponible entre todos los tráficos, afectaba de forma notable al audio del instructor, que se cortaba por momentos o incluso dejaba de reproducirse, además de implicar un mayor tiempo de transmisión.
- Procesamiento elevado: para paliar las posibles pérdidas y corrupciones de paquetes cuando se enviaban las páginas de los documentos por red, se utilizaba un mecanismo de control de errores de tipo FEC. Dado el gran tamaño de cada página individual en formato SVG, los requisitos computacionales para el cálculo de la clave FEC resultaban muy pesados y requería, en ocasiones, del orden de varios segundos para llevarse a cabo, lo que introducía demasiada latencia en la navegación a través de las páginas. Además, esto comprometía el proceso de decodificación de audio y vídeo, especialmente al segundo pues requiere mayor recursos computacionales.

A partir de esta experiencia se desechó la posibilidad de ofrecer las páginas de los documentos bajo demanda. En lugar de ello, se habilita un espacio web específico para estos recursos en el sitio web corporativo, desde el que los alumnos pueden descargarlos. Esta alternativa es la más adecuada para entornos en los que no se puede garantizar un mínimo de ancho de banda, pues se evita que el transporte de las páginas influya negativamente otros tráficos.

4.8.2.1.3. Navegación por contenidos Para navegar a través de los contenidos que se están compartiendo dentro de un área de trabajo de la pizarra virtual es necesario definir un mecanismo que permita desplazarnos por las páginas que componen dicho área de trabajo. En concreto, éstas son las acciones que son necesarias:

- Navegar por las páginas que componen el área de trabajo.
- Cambiar el factor de zoom con el que se renderizan las diapositivas para ver con más detalle algún contenido de la página actual.
- Desplazarse de arriba a abajo y/o de izquierda a derecha dentro de una página que tiene un factor de zoom elevado y que no permite, por tanto, visualizar la página completa dentro del área de renderizado.

El protocolo en sí es bastante simple, sólo es necesario tomar algunas decisiones respecto a cómo llevar a cabo la implementación. La primera es si el protocolo debe ser con estado o sin estado. Un protocolo sin estado transportaría en cada mensaje de navegación la referencia completa del

estado de renderización actual, es decir, número de página, factor de zoom y desplazamiento respecto al origen de coordenadas actuales. Sin embargo, un protocolo con estado, únicamente transmite los cambios en el estado de la navegación, con lo que se consigue un ahorro de ancho de banda por el menor tamaño de los mensajes intercambiados. Por ejemplo, avanzar página sería un comando que obligaría a los participantes que lo reciben a conocer cuál es el estado previo, incluyendo el número de página anterior. Esto puede plantear problemas para alumnos que se unan con la clase ya iniciada. El mismo problema aparece si un alumno no recibe un paquete con información de navegación. En ese caso, perdería el estado de navegación de forma permanente sin poder recuperarlo, pues únicamente se transmite un estado parcial en cada uno de los paquetes de datos.

Durante una clase de e-learning síncrono interesa que el instructor pueda controlar en todo momento qué contenidos visualizan los alumnos. Esto se consigue si la navegación que realiza el instructor sobre los contenidos se realiza de forma síncrona respecto a los alumnos. En general, el creador de un área de trabajo debería tener potestad para dirigir la navegación dentro de la misma para el resto de participantes de la clase. Sin embargo, como el instructor es el único autorizado para crear espacios de trabajo de pizarra, en la práctica, sólo éste puede gobernar la navegación de los participantes en la clase. No obstante, esto no debería impedir que en un momento determinado un participante pueda navegar por los contenidos de forma local y asíncrona respecto al resto de la clase. El prototipo ofrece la posibilidad de que un alumno navegue asíncronamente de forma local, pero ante cualquier evento de navegación del instructor (cambio en la página actual, en el factor de zoom. . .) el alumno recuperaría el estado de navegación de todo el resto de la clase.

Esta solución que se plantea es válida para la navegación dentro de la misma área de trabajo, puesto que nos permite desplazarnos por los contenidos que esta área presenta. Sin embargo, no define un mecanismo para la navegación entre diferentes áreas de trabajo de pizarra. En una situación en la que existan diferentes áreas de trabajo abiertas en la pizarra, no ofrece mecanismo alguno para conocer cuál de estas áreas es la activa de forma global y, por tanto, la que todos los alumnos deberían estar visualizando. Como consecuencia de esto, si un alumno cambia el área activa de forma local, no podrá conocer en qué área de trabajo se encuentra el resto de la clase a menos que se le informe de ello por otros medios, por ejemplo, mediante una alocución del instructor.

Para la transmisión de las ordenes de navegación desde el instructor hasta los alumnos existen asimismo varias alternativas:

- Conexiones independientes: supone la transmisión de las ordenes de navegación de forma distinta a como se transmiten los medios que componen la pizarra virtual.
- Transporte dentro de los medios de la pizarra: aprovechando la capa de presentación de contenidos que siempre existe dentro de un área de trabajo de pizarra, se pueden transmitir las órdenes de navegación a través de la misma sesión RTP en la que éstos se transportan.

Como solución óptima, se utiliza la propia sesión RTP originalmente destinada a transportar los contenidos del área de trabajo para implementar un protocolo de navegación sin estado. De esta forma, cada uno de los flujos de dicha sesión transporta, además de la URI del recurso a utilizar como contenidos del área de trabajo, los mensajes del protocolo de navegación, tal como se refleja en la figura 4.25.

4.8.3. Anotaciones

Las anotaciones, también denominadas tinta virtual, son trazos que puede realizar el instructor sobre los contenidos de un área de pizarra para completar la información que éstos proporcionan o para aclarar algún concepto que pueda haber quedado poco claro durante su explicación. En ocasiones puede resultar de utilidad permitir a los alumnos que anoten sobre los contenidos

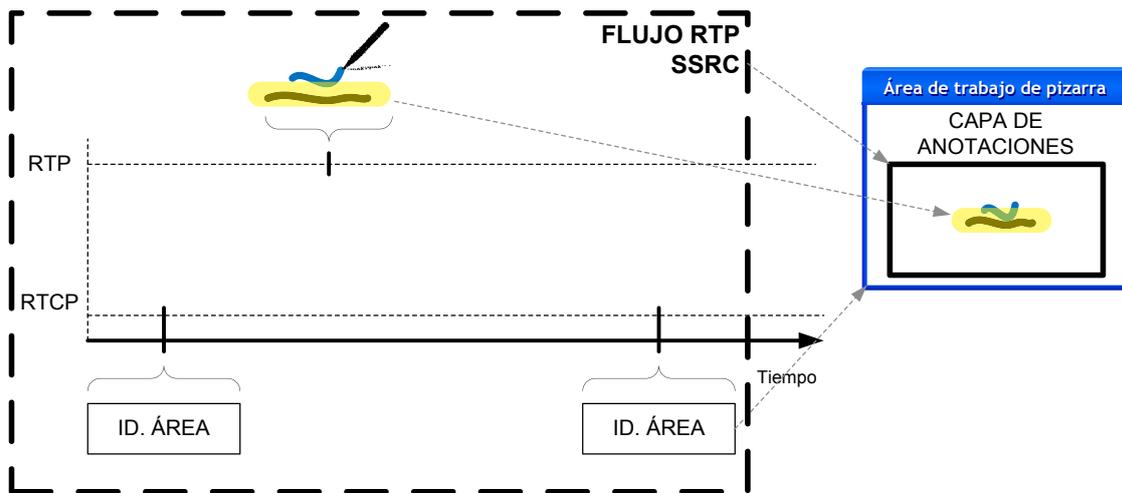


Figura 4.26: Implementación de la capa de anotaciones con RTP

compartidos para concretar alguna duda que le planteen al instructor, o para que, a instancias de una pregunta de éste, realicen una determinada anotación a modo de respuesta. Por este motivo, el modelo de interacción que se propone para este tipo de medios es un modelo colaborativo N-N, donde cualquier participante en la clase virtual puede realizar anotaciones sobre los contenidos.

Siguiendo la recomendación T.126 de la ITU-T, las áreas de trabajo de la pizarra virtual están organizadas en capas, de las cuales la más inferior consiste en la capa de presentación de contenidos. Sobre esta capa se sitúan una o varias capas de anotaciones y, sobre éstas a su vez, una o varias capas de puntero virtual. En la solución que se propone, cada capa tiene un propietario y éste es el único que puede interactuar con esa capa. Sólo un usuario puede realizar anotaciones sobre dicha capa. Lo habitual es disponer de una capa por cada uno de los participantes que puede realizar anotaciones sobre la pizarra virtual.

El SDK de Tablet PC de Microsoft es una buena alternativa para permitir la incorporación de tinta electrónica dentro del prototipo. Esta tecnología permite realizar anotaciones y proporciona estructuras de datos para manipularlas de forma individual. Esto nos permite su transporte por la red al resto de participantes. Aunque inicialmente está pensado para el uso sobre equipos Tablet PC, se puede portar a cualquier plataforma Windows.

Para el transporte de los trazos y la representación de una capa dentro de la clase virtual se aprovechan las características de RTP y RTCP respectivamente de cada flujo RTP dentro de la sesión RTP destinada al transporte de anotaciones. Sobre el protocolo de transporte se transmiten los datos de cada trazo de forma individual, mientras que se utilizan los paquetes RTCP periódicos para relacionar una capa con el área de trabajo al que pertenece. En este sentido, se produce una relación directa entre capa de anotaciones y flujo de datos dentro de la sesión RTP. En la figura 4.26 se ilustra esta situación. Esta solución presenta dos inconvenientes:

- Late-join: los participantes que se unen a una clase virtual ya comenzada deberían recibir lo más pronto posible todos los datos necesarios para que conozcan el estado de la clase virtual y puedan seguirla de forma efectiva. Sin embargo, con esta solución sólo se puede recuperar la existencia de las diferentes capas de anotaciones a través de los paquetes periódicos RTCP, pero no las anotaciones que fueron realizadas hasta el momento. Una posible solución a este problema implicaría que cada usuario propietario de cada capa de anotaciones reenviara la información de los trazos ya realizados, sin embargo esto podría provocar una explosión de paquetes cuando muchos participantes se conectan al mismo tiempo que podría afectar

a otros flujos dentro de la plataforma prototipo. De cualquier forma, esta circunstancia se asemeja a la situación en que un alumno llega tarde a clase y el profesor ya ha borrado la pizarra que contenía las explicaciones hasta ese momento.

- **Persistencia:** cuando un participante abandona la clase virtual automáticamente se elimina su capa de anotaciones y, con ella, todas las anotaciones que contiene. Esto puede resultar inconveniente en el caso del instructor; si éste se desconecta momentáneamente de la clase virtual, supondría que las anotaciones que ha realizado hasta el momento se perdieran, si bien este problema también sucede para el área de trabajo completa.

4.8.3.1. Evaluación de alternativas y optimización

Para la captura de los trazos y su representación en el equipo del resto de participantes de la clase existen, básicamente, dos alternativas:

- **Capturar trazos completos:** esto supone que no se envía ningún tipo de información a través de la sesión RTP hasta que el trazo está completo. De esta forma, se ahorra ancho de banda pues se reduce el número de paquetes de datos que es necesario transmitir. Por contra, esto puede resultar contraproducente cuando se anota sobre contenidos muy cargados, puesto que la anotación aparecerá en los equipos del resto de participantes de forma instantánea, por lo que puede ser difícil de localizar en diapositivas con muchos contenidos. Esto puede mitigarse en parte con la utilización del puntero virtual, así el resto de participantes pueden seguir los movimientos del cursor donde aparecerá la anotación a continuación.
- **Muestrear la generación de trazos:** esta alternativa implica el muestreo de los movimientos del cursor durante la generación de un trazo para enviar paquetes de datos periódicos a través de la sesión RTP, para que los equipos del resto de participantes puedan recomponer el trazo de forma progresiva. Esto facilita el reconocimiento de los trazos en contenidos muy cargados puesto que dicha generación progresiva atraerá la atención del usuario dentro de unos contenidos estáticos. Sin embargo, el gran inconveniente que presenta es el mayor ancho de banda que consume respecto a la anterior solución, que será tanto mayor cuanto mayor sea el periodo de muestreo de los movimientos del cursor.

Un problema que se deriva de la utilización de dos sesiones independientes para el transporte de las anotaciones y de los contenidos, que en última instancia representan el área de trabajo completo, es la necesaria identificación de un plano de anotaciones con el área de trabajo al que pertenece. Esto se hace mediante el identificador de fuente (SSRC) del flujo de datos que representa el área de trabajo dentro de la sesión RTP. Este SSRC es transportado en cada uno de los paquetes RTCP periódicos que se envían para el flujo que representa la capa de anotaciones. Sin embargo, pueden existir condiciones de carrera en que un plano de anotaciones se cree antes que el área de trabajo que lo contiene debido a los diferentes intervalos de transmisión RTCP dentro de cada una de las sesiones RTP. Esto es especialmente notorio cuando un usuario se conecta con la clase ya iniciada. Por tanto, es necesario tener en cuenta estos planos *huérfanos* para asociarlos más adelante cuando ya se cree el área de trabajo que los contiene.

La solución óptima para el prototipo a desarrollar es la captura de trazos completos por el menor ancho de banda que consume. Además, debido al menor número de paquetes y la utilización de técnicas FEC para la corrección de errores, resultan en una solución que consume pocos recursos de red y que está preparada ante la aparición de pérdidas de paquetes. Asimismo, como el número de paquetes a procesar es bajo, los recursos computacionales para el cálculo de las claves FEC no será excesivo.

4.8.4. Puntero virtual

El puntero virtual, denominado también telepuntero, es una funcionalidad que puede utilizar el instructor de una clase virtual para señalar o apuntar algún elemento concreto de la pizarra sin necesidad de realizar una anotación. De igual manera que sucedía con las anotaciones, resulta interesante ofrecer al resto de participantes de la clase virtual la posibilidad de utilizar su propio puntero para señalar algún elemento dentro de la pizarra. Esto puede ser muy útil para que un alumno plantee una cuestión sobre los contenidos de la pizarra, especialmente cuando éstos son muy densos, para apuntar algún aspecto de los contenidos. También es interesante para que el profesor plantee cuestiones a los alumnos que deban responder apuntando un elemento dentro de la diapositiva actual.

Al igual que ocurría con las capas de anotaciones, dentro de la misma área de trabajo de la pizarra virtual pueden coexistir varias capas de punteros virtuales. Cada una de estas capas se utiliza para representar el puntero de un determinado participante, por lo que existirán tantas capas dentro del área de pizarra como punteros virtuales activos. De nuevo, es necesario transportar en los mensajes RTCP el identificador del área de trabajo al que pertenece. La única diferencia respecto a cómo se gestionan las anotaciones es que, mientras que los planos de anotaciones se crean de forma automática una vez que un participante entra en la clase virtual, un plano que contiene el puntero virtual de un participante se crea exclusivamente cuando éste activa su puntero. Por tanto, inicialmente, dentro de un área de trabajo de la pizarra no existirán planos de puntero, puesto que, por defecto, todos los usuarios tienen deshabilitado su puntero virtual.

Para el transporte del puntero virtual se utiliza, como ya se mencionó anteriormente, una sesión RTP independiente de la de cualquier otro medio. El IETF define un estándar para el transporte en tiempo real de los datos de un puntero virtual a través de RTP [Civanlar y Cash, 2000]. Este estándar permite transportar por la red la posición del puntero virtual, el estado de los botones del ratón o puntero que se utilice para desplazar el cursor, y el icono a utilizar para representar visualmente el puntero virtual. Esto viene representado en la figura 4.27.

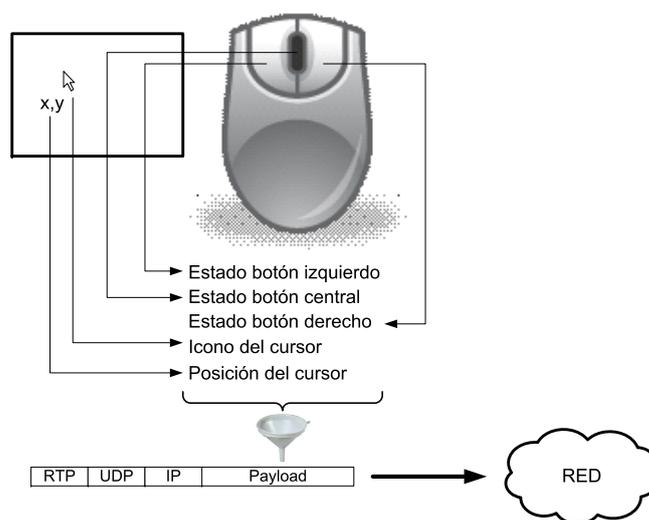


Figura 4.27: Datos que se transportan del puntero virtual

Uno de los problemas que trata de resolver este estándar del IETF, mediante la definición de un espacio de coordenadas independiente, es la posibilidad de que el espacio de trabajo común que comparten los participantes tenga diferentes medidas en uno u otro, dependiendo de las capacidades de los dispositivos de visualización utilizados por los usuarios. Sin embargo, esta

problemática no se presenta en la solución propuesta, ya que todos los participantes en la clase virtual conocen y están de acuerdo en las dimensiones del área de trabajo; a continuación, éste se renderiza en los equipos de los usuarios tal como se describió en el apartado 4.8. No obstante, para la compatibilidad con otras implementaciones, deben convertirse las coordenadas con las que trabaja el puntero virtual (coordenadas del área de trabajo) a las coordenadas correspondientes dentro del espacio que define el estándar, cuyas coordenadas se definen en el rango $[0, 2^{12} - 1]$. En la figura 4.28 se ilustra el proceso de traducción de coordenadas del área de trabajo (x,y) a coordenadas de transporte (x',y') , siendo (X,Y) las dimensiones del área de trabajo de pizarra.

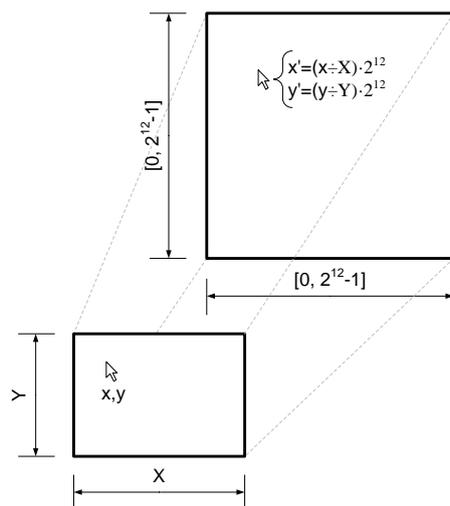


Figura 4.28: Conversión de unidades de espacio de trabajo a unidades de transporte

4.8.4.1. Evaluación de alternativas y optimización

Como sucede con el audio y el vídeo, el puntero puede considerarse también un medio continuo representado por los movimientos que realiza sobre el espacio de trabajo. De igual forma, para convertir esos movimientos en información digital que pueda ser transportada por la red al resto de participantes en la clase, debe realizarse un proceso de captura, para lo cual es necesario definir un cierto periodo para muestrear la posición del puntero. Este periodo debe ser un compromiso entre recursos utilizados, computacionales y de red principalmente, y fidelidad en los movimientos del puntero. En la figura 4.29 se representa este proceso de muestreo, de donde se deduce que la elección de un periodo de captura menor repercute en una mayor precisión en la reconstrucción de los movimientos del puntero a costa de un mayor número de mensajes que se envían a la red.

Pruebas realizadas en el laboratorio determinan que es necesario muestrear los movimientos del cursor con una frecuencia mayor o igual que 10 veces por segundo para que el movimiento sea fluido.

4.9. Control de la sesión

La capa de control se encarga de gestionar el acceso de un participante a la clase virtual. Entre sus tareas se encuentran la negociación y el establecimiento de las sesiones RTP necesarias para intercambiar información multimedia entre los participantes de la clase. Este proceso de unión de un usuario a una clase virtual consta de los siguientes pasos:

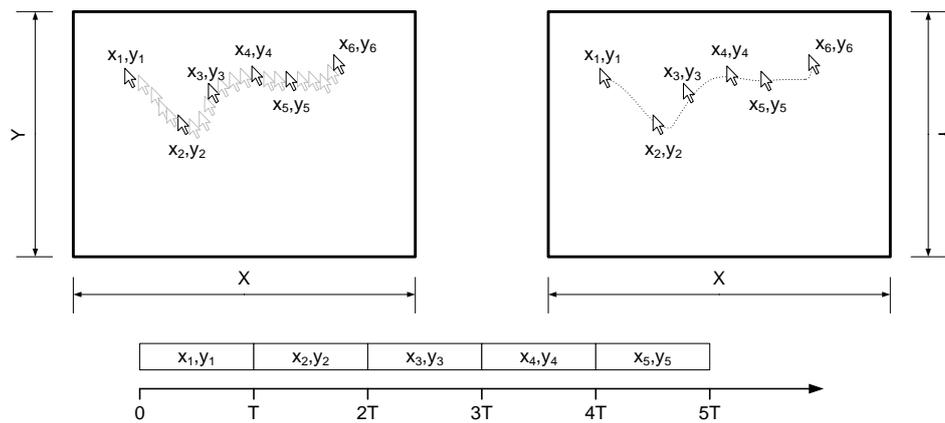


Figura 4.29: Muestreo del puntero local y su representación en un equipo remoto

1. Descubrir la clase: es necesario que los usuarios conozcan cómo unirse a la clase.
2. Negociación de medios: permite seleccionar la configuración de medios que se va a utilizar durante la clase virtual.
3. Unirse a la clase: supone la creación de las sesiones RTP de cada uno de los medios que componen la clase virtual.

Inicialmente, un usuario que desee participar en una clase virtual debe conocer cuál es el mecanismo para unirse a la clase. Interesa que este mecanismo sea lo más transparente al usuario, para que éste, a lo sumo, tenga que pinchar un enlace en una página web o proporcionar un identificador que previamente se le haya suministrado.

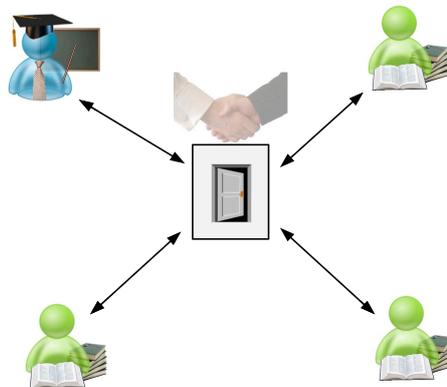


Figura 4.30: Negociación para la unión a la clase virtual

En el caso de que sea necesario un proceso de negociación de los medios que se intercambian en la clase virtual y/o sus parámetros, se hace necesaria la existencia de un elemento central desde el que los usuarios acceden a ésta, previa negociación de medios. La comunicación entre cada usuario y el elemento central se realiza mediante un protocolo de señalización. Una representación de esta comunicación puede verse en la figura 4.30. El grafo que conforman estas conexiones de señalización es totalmente independiente del grafo que constituyen las conexiones para el

transporte de datos analizadas en el apartado 4.2. Dependiendo de cómo se realice el proceso de negociación, este elemento central posibilita dos alternativas no excluyentes:

- Negociación homogénea: ante cualquier usuario que solicite unirse a la clase virtual, a través del protocolo de señalización correspondiente, el elemento central negociará con el la utilización de los mismos medios y parámetros de los mismos que para el resto de participantes en la conferencia. De esta forma, no existe diferencia en el tratamiento de cada uno de los participantes.
- Negociación heterogénea: en este caso, el elemento central puede ofrecer a cada participante un subconjunto de los medios que se están utilizando en la clase virtual. Por ejemplo, en una clase virtual en que se utilicen audio y vídeo, ante la llegada de un nuevo participante, en función de su identificador (específico del protocolo de señalización) el elemento central puede negociar con él la utilización de audio únicamente, así el participante desconocería la existencia de la sesión RTP de vídeo, por lo que sólo participaría en la clase a través del audio.

A continuación, se describen las posibles alternativas para el desarrollo de la capa de control de la sesión.

4.9.1. SIP

Una de las posibles realizaciones de la capa de control de la clase virtual sería a través de SIP. En este caso, debería implementarse una solución como la que se propone en [Rosenberg, 2006]. En ella, existe la figura de foco de la conferencia y todos los participantes deben comunicarse con éste para llevar a cabo la negociación de medios previa a la participación en la clase virtual.

La clase virtual se identificaría por una URI SIP del tipo `clase001@sip.arcelormittal.com`. La gestión de las URI SIP y su relación con las clases virtuales depende de la política de nombres de la empresa y de un administrador que se encargue de crear uno para cada clase individual⁶. En este último caso, además de especificar el URI correspondiente a la clase habría que indicar de alguna manera los medios que se utilizan durante la clase, para que el foco los negocie correctamente con cada participante.

A partir de este momento, los participantes pueden unirse a la clase virtual solicitando una petición de conexión a la URI `sip:clase001@arcelormittal.com`. Para ello deben conocer esta URI de alguna forma, por ejemplo, mediante su publicación en el portal web corporativo o su difusión a través del correo electrónico. En la figura 4.31 se ilustra el proceso de unión a la clase virtual.

- En primer lugar el participante envía un mensaje INVITE para establecer comunicación con el usuario `sip:clase001@arcelormittal.com`, que en realidad se corresponde con una URI de clase virtual (conferencia), pero esto es transparente al usuario. El contenido del mensaje SIP podría ser parecido al que aparece en el listado 4.1. En dicho listado se ha omitido el cuerpo del mensaje SDP que se utilizaría para negociar los medios.
- A continuación, el foco respondería con un mensaje del tipo 180 para indicar que la petición está en progreso. A partir de entonces, el foco puede decidir si darle al usuario la posibilidad de participar en la clase virtual y proseguir con la negociación de medios o enviarle una notificación de error para impedirle el acceso. Esta decisión puede tomarse en base a muchos criterios, como por ejemplo los medios que especifica el usuario en su mensaje SDP, la URI SIP del usuario⁷, el número de participantes en la clase, etc. Si se autoriza al usuario, se

⁶También es posible crear la URI de la clase dinámicamente cuando se une un instructor a la clase.

⁷Es posible mantener una base de datos con usuarios autorizados para acceder a la clase virtual.

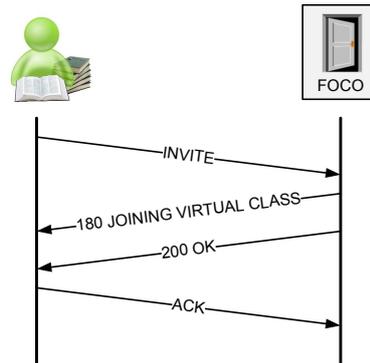


Figura 4.31: Unión a la clase virtual utilizando SIP

```

INVITE sip:clase001@arcelormittal.com SIP/2.0
Via: SIP/2.0/UDP equipo01.arcelormittal.com:5060;branch=z9hG4bKfw19b
Max-Forwards: 70
To: Clase virtual 001 <sip:clase001@arcelormittal.com>
From: Usuario 1 <sip:usuario1@arcelor.com>;tag=76341
Call-ID: 12345@equipo01.arcelormittal.com
CSeq: 1 INVITE
Contact: <sip:usuario1@equipo01.arcelormittal.com>
Content-Type: application/sdp
Content-Length: 142

—— Cuerpo del mensaje no mostrado ——
  
```

Listado 4.1: Petición INVITE de un usuario al foco de la clase virtual

continúa el proceso de negociación y el foco le envía un mensaje de tipo 200 junto con la configuración de medios que puede utilizar. Un ejemplo aparece en el listado 4.2. En este también se ha omitido el contenido del mensaje SDP.

- La negociación terminaría con un mensaje de tipo ACK enviado por el usuario al foco para indicarle que acepta la configuración de medios proporcionada. El usuario ya podría unirse a las sesiones RTP de los diferentes medios.

4.9.2. H.323

El proceso para el establecimiento y negociación de las sesiones sería muy similar si en lugar de utilizar SIP se utiliza H.323. Las únicas diferencias estriban en el nombre que se le da al punto central de negociado de las sesiones, MCU en H.323 y foco en SIP, y la diferente complejidad de implementación dependiendo del protocolo elegido. Por lo demás, la comunicación de mensajes sería parecida, con la salvedad de que en H.323 requeriría el intercambio de un mayor número de mensajes. De cualquier forma, la utilización de H.323 en lugar de SIP no aporta muchos más beneficios y, por contra, añade una complejidad extra derivada del mayor número de mensajes a procesar y la difícil depuración de los mismos al tratarse de mensajes binarios.

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP equipo01.arcelormittal.com:5060;branch=z9hG4bKfw19b
;received=10.90.90.90
To: Clase virtual 001 <sip:clase001@arcelormittal.com>;tag=a53a42
From: Usuario 1 <sip:usuario1@arcelor.com>;tag=76341
Call-ID: 12345@equipo01.arcelormittal.com
CSeq: 1 INVITE
Contact: <sip:clase001@sipfocus.arcelormittal.com>
Content-Type: application/sdp
Content-Length: 130
—— Cuerpo del mensaje no mostrado ——
```

Listado 4.2: Respuesta afirmativa del foco de la clase virtual

```
v=0
o=granda 2890844526 2890844526 IN IP4 156.35.151.86
s=Ejemplo de mensaje SDP
i=Este mensaje describe el uso del protocolo SDP
u=http://www.ietf.org/rfc/rfc4566.txt
e=Juan Carlos Granda <jcgranda@uniovi.es>
p=+34 985 18 26 38
c=IN IP4 225.45.3.56/236
b=CT:144
t=2877631875 2877633673
m=audio 49172 RTP/AVP 0
a=rtpmap:0 PCMU/8000
m=video 23422 RTP/AVP 31
a=rtpmap:31 H261/90000
```

Listado 4.3: Ejemplo de mensaje SDP

4.9.3. SAP

Como se discutió en el apartado 3.2.6.4, el protocolo SAP sirve para anunciar periódicamente, utilizando mensajes multicast a la dirección IP 224.2.127.254 y puerto 9875, sesiones multimedia, para lo cual es común adjuntar un mensaje SDP en el que se describe la configuración de medios que se utiliza durante la sesión. No ofrece ningún mecanismo que permita la negociación para unirse a la sesión, por tanto, los usuarios que pretenden unirse a la sesión deben hacerlo utilizando la configuración especificada en el mensaje SDP.

4.9.4. SDP

SDP es un protocolo que nos permite describir sesiones multimedia y especificar qué medios y con qué parámetros se intercambian en dicha sesión. Es utilizado en SAP y SIP para proporcionar la descripción de medios y la negociación de cada uno de ellos respectivamente. En este sentido SAP y SIP se convierten en meros vehículos de transporte del mensaje SDP, que es realmente el protocolo que se utiliza para describir la sesión y para realizar la negociación previa a la unión a la clase virtual por parte de un participante. Un ejemplo de mensaje SDP describiendo una clase virtual se muestra en el listado 4.3.

Gracias a la utilización de SDP, independientemente de que se se utilice conjuntamente con un protocolo de señalización, se consigue:

- **Flexibilidad:** la posibilidad de especificar los medios a utilizar junto con sus parámetros a través de un mensaje SDP permite que los usuarios participen en las clases virtuales con diferente configuración multimedia. De esta forma, puede adaptarse la configuración de la clase en función de la situación de la red, el equipo del participante, etc. Por ejemplo, ante una situación de congestión en la red o de bajo ancho de banda del enlace de un participante, se le ofrecerá una descripción SDP que implique la no utilización de vídeo durante la clase, puesto que es el medio que más ancho de banda consume.
- **Nuevos medios:** es posible añadir nuevos medios al conjunto de los que componen la clase virtual. Aquellos clientes que no reconozcan un medio dentro de una descripción SDP simplemente lo ignoran. Esto no impide que los usuarios se unan a la clase virtual, ya que pueden seguir participando a través del resto de medios que componen la descripción SDP.
- **Compatibilidad:** ésta es una solución estándar, ya que se basa en la utilización de un protocolo estandarizado, lo cual hace que la herramienta sea compatible con otras soluciones que utilicen SDP para la negociación de medios, por ejemplo, VoIP, IPTV, audio y vídeo bajo demanda, etc.
- **Ampliabilidad:** SDP es el protocolo que se utiliza para la descripción de sesiones y la negociación de medios en dos de los protocolos estándar de señalización: SIP y SAP. El hecho de que la configuración multimedia de la clase pueda ser descrita mediante SDP supone que la utilización de un protocolo de señalización como SIP es directa.

4.9.5. Evaluación de alternativas y optimización

Cada una de las alternativas propuestas en el apartado anterior tiene sus ventajas e inconvenientes. Algunas incluso, ofrecen funcionalidad innecesaria por el ámbito específico de utilización del prototipo. Por ejemplo, la utilización de SIP o H.323 es la solución que más funcionalidad nos reporta, en concreto, ofrece las siguientes funciones:

- **Descubrimiento de sesiones:** es un mecanismo universal que viene incorporado de forma transparente en el protocolo de señalización. Sería posible establecer las clases virtuales con participantes totalmente desconocidos sin más que éstos conociesen el identificador de la clase apropiado. Sin embargo, en un entorno corporativo es factible suponer que los empleados pueden utilizar otras técnicas para descubrir las sesiones, como puede ser el correo electrónico, el portal corporativo, etc.
- **Negociación de medios:** los protocolos de señalización permiten una completa negociación de la configuración multimedia de la clase virtual. De hecho, están pensados para poner de acuerdo a dos o más participantes cada uno dispuesto a utilizar un conjunto de medios determinados. Sobre estos conjuntos de medios no puede presuponerse un mínimo común, de ahí que sea necesario ese proceso de negociación para alcanzar ese mínimo común. No obstante, dentro de una misma empresa, sí resulta coherente pensar que los participantes en una clase virtual pueden utilizar una configuración multimedia más o menos común, ya que la configuración software y hardware de los equipos desde los que se conectan los participantes puede ser modificada apropiadamente.
- **Expulsar a usuarios:** con la utilización de un protocolo de señalización y un punto central de administración de la clase virtual, es posible expulsar a un participante sin más que el controlador de la clase le inste a finalizar las conexiones multimedia que tenga abiertas.

Para ello, utilizará el diálogo de señalización que mantienen todos los participantes con el controlador de la conferencia.

Para el caso que concierne al prototipo desarrollado, teniendo en cuenta el que el entorno sobre el que opera es dentro de una red corporativa para la capacitación de los recursos humanos, pueden hacerse una serie de suposiciones:

- Homogeneidad: la mayor parte de los equipos informáticos tendrán una configuración software y hardware parecida, por lo que podrán utilizar para la clase virtual configuraciones multimedia parecidas. Además, interesa que todos los empleados participantes de la clase reciban el mismo tipo de información y no se vean limitados en la recepción de un tipo de datos, por ejemplo vídeo. De esta forma, la configuración multimedia que utilizan todos los participantes está determinada de antemano, luego no es necesario un proceso de negociación individual con cada uno de los participantes a la clase virtual. Esto implica que únicamente es necesaria la descripción de los medios que se utilizan en la clase, esto es, un mensaje SDP.
- Estado de la red: ya que la red sobre la que opera el prototipo es una red corporativa se tiene acceso a la configuración y al estado de la misma. Es posible determinar el estado de los enlaces y realizar un cálculo inicial del ancho de banda disponible. Incluso, es posible asignar un nivel de calidad de servicio a los datos que transmite por la red el prototipo para asegurar un ancho de banda mínimo. Así, es posible deducir una configuración multimedia mínima que todos los usuarios podrían satisfacer a priori.
- Portal corporativo: existe un portal corporativo al que los empleados tienen acceso y desde el que pueden descargar los materiales así como las configuraciones multimedia para cada una de las clases virtuales.

La solución que se implementa en el prototipo es la utilización de SDP para describir todos los medios que se intercambian durante la clase virtual. Esta descripción, de la que se muestra un ejemplo en el listado 4.4, es utilizada por la herramienta para establecer las sesiones RTP necesarias para cada uno de los medios. Posteriormente, los medios son procesados de acuerdo al tipo de cada uno como se detalla en el apartado 4.5.

Para obtener la descripción SDP, los usuarios acceden al portal corporativo desde el cual pueden obtener el enlace a un fichero que contiene una descripción SDP de la configuración multimedia de la clase. Este fichero puede ser generado de forma dinámica para configurar cada una de las clases de forma adecuada, especificando los medios que se utilizarán en cada una de ellas.

No se utiliza un protocolo de señalización como SIP puesto que su principal cometido es el anuncio de las clases virtuales y esto ya se consigue gracias a la publicación de las configuraciones de la clase en el portal corporativo. Tampoco es necesario un protocolo del tipo SIP porque no es necesario el proceso de negociación de la sesión, la configuración multimedia de ésta se fija de antemano para que todos los empleados obtengan la misma experiencia del desarrollo de la clase.

La principal desventaja de esta solución es que, debido a la utilización de grupos multicast para el transporte de los datos y a la no utilización de un protocolo de señalización, es imposible expulsar de la clase virtual a un participante problemático, ya que éste seguiría recibiendo los flujos multimedia mientras permanezca suscrito al grupo multicast. Sin embargo, la aparición de participantes problemáticos en una clase virtual es improbable que suceda teniendo en cuenta que los participantes en la clase son empleados de la empresa y, por tanto, están bajo su tutela. En este caso, el miedo a represalias debería ser suficiente para evitar este tipo de situaciones.

```

v=0
o=ArcelorMittal 2890844526 2890844526 IN IP4 10.90.90.100
s=Virtual Class
t=0 0
m=text 11000 RTP/AVP 127
c=IN IP4 239.255.0.1/127
a=rtpmap:127 t140/1000
m=application 13000 RTP/AVP 100
c=IN IP4 239.255.0.2/127
a=rtpmap:100 etipoWB/1000
m=application 14000 RTP/AVP 101
c=IN IP4 239.255.0.3/127
a=rtpmap:101 etipoAnnot/1000
m=video/pointer 15000 RTP/AVP 102
c=IN IP4 239.255.0.4/127
a=rtpmap:102 pointer/90000

```

Listado 4.4: Ejemplo de descripción de una clase virtual con SDP

4.10. Interfaz de usuario

En este apartado se describe la interfaz de usuario definitiva del prototipo que permite a instructores y alumnos participar en una clase virtual. Las decisiones de diseño de la interfaz se han tomado a partir de la previa caracterización de las herramientas de e-learning síncrono y los requisitos deducidos en el capítulo 2. En todo momento, debe mostrarse la gran cantidad de información con la que trabaja el prototipo de la forma más simple posible sin olvidar el perfil de usuario al que está dirigido el prototipo. El usuario típico del prototipo es un empleado de una gran corporación empresarial, muchas veces con escasos conocimientos informáticos y muchos prejuicios ante las nuevas tecnologías, que participa en la clase para recibir capacitación en el desarrollo de su actividad laboral.

Al tratarse de una única herramienta que opera en modo *peer-to-peer* entre los participantes de la clase virtual, se proporciona la misma interfaz de usuario básica para instructor y alumnos, de tal forma que el cambio de rol de un participante se pueda realizar de forma directa. De esta forma, una única aplicación con la misma interfaz sirve para acceder a la clase virtual tanto a instructores como alumnos.

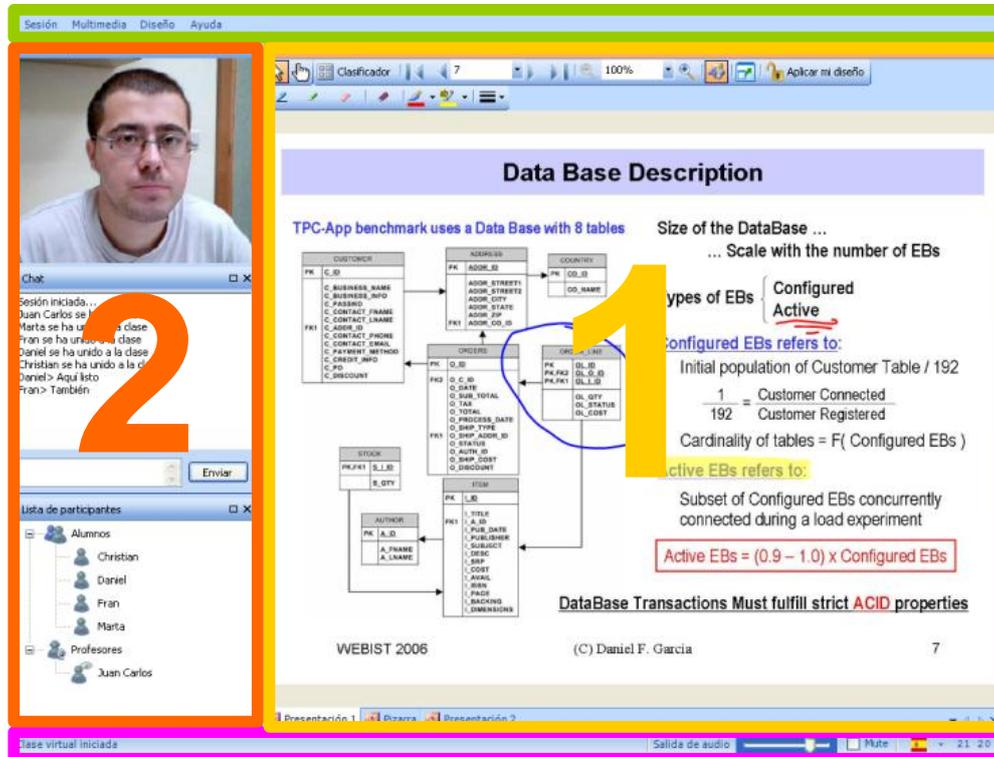
En los siguientes puntos se tratan los aspectos más importantes de la interfaz de usuario reflejada en la figura 4.32: el menú principal, la barra de estado, el área de presentación de contenidos y las ventanas auxiliares, que dan servicio a la mensajería instantánea, la lista de participantes y la videoconferencia.

4.10.1. Menú principal

El menú principal contiene accesos a todas las funcionalidades del prototipo, salvo a aquéllas que se refieren a aspectos muy específicos, cuya inclusión en el menú confundiría al usuario en lugar de facilitarle el manejo del prototipo. Este es el caso de las herramientas de anotación sobre los contenidos de la pizarra, que sólo están accesibles desde una barra de herramientas situada en el área de presentación. En la figura 4.33 se muestra el menú principal con sus diferentes submenús abiertos.

Se observa que el menú principal tiene cuatro elementos: **Sesión**, **Multimedia**, **Diseño** y **Ayuda**. **Sesión** agrupa las opciones relacionadas con el inicio y la finalización de las clases virtuales, y

3



4

Figura 4.32: Interfaz del prototipo donde se identifican cada una de las partes funcionales: (1) área de presentación, (2) ventanas auxiliares, (3) menú principal y (4) barra de estado

la configuración de los dispositivos de audio y vídeo que utiliza el prototipo. **Multimedia** es un submenú que, dependiendo del rol que juegue el participante en la clase virtual aparecerán más o menos opciones. Si el participante ejerce el papel de alumno dentro de la clase virtual tan sólo contiene la opción para mostrar el mezclador del sistema para tener acceso a la configuración de audio del sistema. En el caso del instructor o presentador, además de esta opción, aparecen todas las necesarias para dirigir y controlar el desarrollo de las clases virtuales. Mediante el submenú **Diseño** los usuarios pueden variar la disposición en pantalla de los diferentes componentes del prototipo. El submenú **Ayuda** no precisa mayores aclaraciones. Desde él se puede acceder al manual de la aplicación y a una pantalla con información sobre la misma.

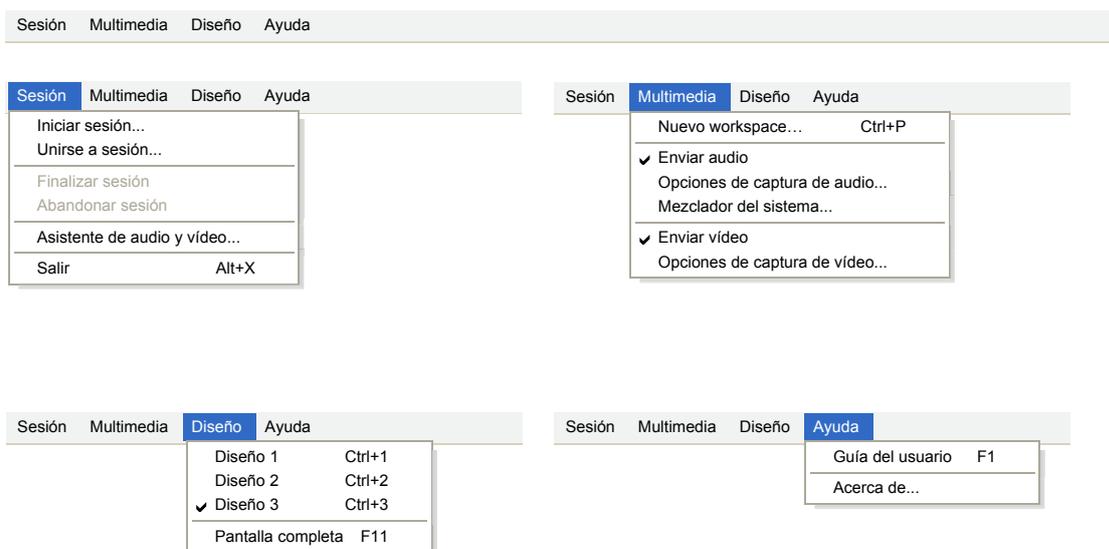


Figura 4.33: Menú principal del prototipo

4.10.1.1. Submenú Sesión

El primer elemento del menú principal de la gran mayoría de aplicaciones suele contener opciones relativas al comienzo y la finalización de una tarea desempeñada por la aplicación. Por ejemplo, las aplicaciones ofimáticas suelen agrupar en el primer submenú (**Archivo**) todas las opciones relativas al manejo de archivos (abrir, cerrar, guardar...). Del mismo modo, el submenú **Sesión** constituye el primer paso para trabajar con el prototipo, ya que a través de él se comienza y se termina la participación en una clase virtual.

El instructor que desee comenzar una sesión de e-learning deberá seleccionar **Iniciar sesión** para crearla y que esté disponible para el resto de participantes. Los alumnos, por su parte, entrarán en la clase virtual seleccionando **Unirse a sesión**. No obstante, no es necesario que el instructor inicie sesión para que los alumnos puedan acceder a la misma. En ambos casos aparecerá un cuadro de diálogo solicitando el nombre del usuario, que será el texto que aparecerá en la interfaz de usuario del resto de participantes, y una contraseña, cuya función es evitar que personas no autorizadas inicien una clase virtual o se unan a una ya existente.

En cualquier momento, un alumno podrá finalizar su participación en la sesión eligiendo la opción **Abandonar sesión**. Por otra parte, el abandono del instructor conlleva inexorablemente la terminación de la clase virtual para todos los participantes. Previendo que en el futuro el prototipo soporte más de un instructor simultáneo, se ha incluido el elemento de menú **Finalizar sesión**, que deberá ser el empleado por el instructor para abandonar la sesión. De este modo, en un escenario en el que pueda haber varios instructores, si uno de ellos quiere abandonar la clase virtual pero no finalizarla para el resto de alumnos e instructores, podrá seleccionar **Abandonar sesión**; mientras que si quiere dar por concluida la clase, seleccionará **Finalizar sesión**. Actualmente el instructor tendrá deshabilitada la opción **Abandonar sesión** y sólo podrá usar **Finalizar sesión**. Del mismo modo los alumnos tendrán deshabilitado el elemento de menú **Finalizar sesión**. Tras seleccionar cualquiera de las dos opciones se pedirá confirmación al usuario.

El elemento de menú **Asistente de audio y vídeo** mostrará un asistente que permitirá configurar los dispositivos de captura y reproducción de audio y vídeo que se utilizará durante la clase virtual.

Salir sirve para salir de la aplicación y de la sesión, en caso de estar conectado a alguna. Las

opciones **Finalizar sesión** y **Abandonar sesión** no terminan la aplicación tras desconectarse de la sesión.

4.10.1.2. Submenú Multimedia

Como se comentó anteriormente, este submenú cambia dependiendo del rol que ejerza el usuario dentro de la clase virtual. El alumno dispondrá de un subconjunto de las opciones disponibles dentro del submenú. En concreto, sólo tiene disponible la opción **Mezclador del sistema**. Desde este submenú el instructor puede decidir qué contenidos presentar en el área de presentación y controla los aspectos relativos al vídeo y al sonido que transmite a los alumnos. La primera opción del submenú **Nuevo área de trabajo** sirve para crear un nuevo área de trabajo de pizarra dentro del área de presentación. Al seleccionarla un cuadro de diálogo solicitará al instructor que elija los contenidos a mostrar como diapositivas de fondo de la pizarra, o bien omitirlos para utilizar el área de trabajo como si fuera una pizarra compartida. Una vez confirmadas las selecciones, los contenidos serán visualizados por todos los participantes de la clase virtual.

En este submenú se encuentran dos pares de opciones análogas para el audio y el vídeo. Las opciones **Enviar audio** y **Enviar vídeo** están activadas por defecto, ya que la desactivación de cualquiera de ellas provoca que se deje de transmitir el flujo correspondiente. Se utilizan para reducir el volumen de datos enviados a los alumnos por parte del instructor, con lo que disminuirán notablemente los requerimientos de ancho de banda de la clase virtual. Estas acciones permitirán hacer frente a congestiones de red y otras situaciones en las que no sea posible mantener activados todos los flujos multimedia sin acumular retardos o pérdidas que perjudiquen el correcto desarrollo de la clase. Los elementos del submenú **Opciones de captura de audio** y **Opciones de captura de vídeo** permiten consultar y modificar algunos de los parámetros con los que se capturan estos datos.

4.10.1.3. Submenú Diseño

Mediante este submenú, cualquier usuario puede modificar la disposición en pantalla de las ventanas auxiliares y el área de presentación según una serie de diseños predefinidos. Se ofrecen únicamente diseños predefinidos para evitar que el usuario se vea obligado a disponer las ventanas auxiliares en la pantalla, ya que eso requeriría unas habilidades en el manejo de entornos de ventanas complejos que no se le pueden suponer al usuario tipo de este prototipo.

Se ofrecen cuatro combinaciones: columna de ventanas auxiliares a la izquierda y el área de presentación a la derecha (**Diseño 1**, figura 4.34(a)), columna de ventanas auxiliares a la derecha y el área de presentación a la izquierda (**Diseño 2**, figura 4.34(b)), columna de ventanas auxiliares oculta y el área de presentación ocupando todo el espacio disponible (**Diseño 3**, figura 4.34(c)), y modo de **Pantalla completa** (figura 4.34(d)), en el que sólo se visualiza el área de presentación, mientras que el resto de elementos permanecen ocultos.

La activación de la última opción de este submenú, **Aplicar mi diseño**, obligaría a los alumnos a tener el mismo diseño de pantalla que el instructor, salvo a aquéllos a los que el instructor exima concediéndoles el privilegio *Diseño libre*.

4.10.2. Barra de estado

Esta zona de la interfaz se divide en tres bloques: dos de carácter informativo, y otro que permite controlar aspectos básicos del audio. Los elementos de la barra de estado destinados a informar al usuario están situados en los extremos izquierdo y derecho de la misma. A la izquierda se encuentra el área de notificación, donde aparecerán mensajes informativos para el usuario. Por ejemplo, si el cursor del ratón se encuentra sobre un elemento de un submenú del menú principal, en esta zona se explicará brevemente la utilidad del elemento resaltado.

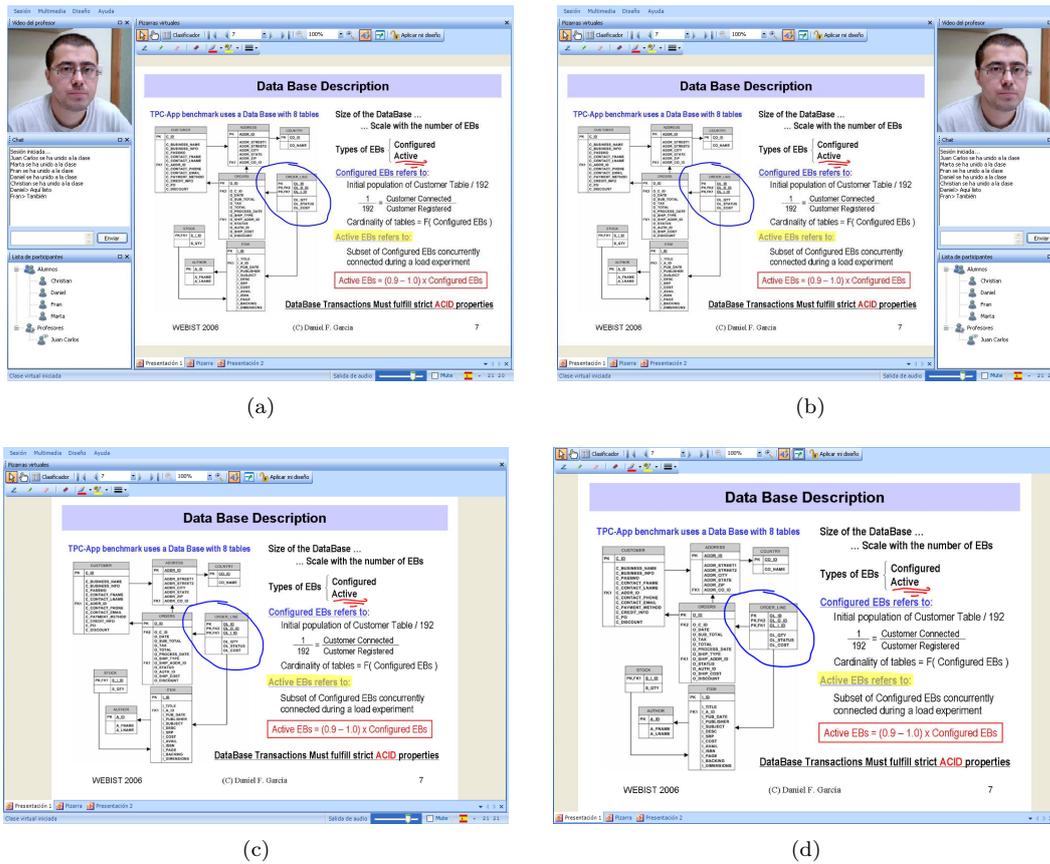


Figura 4.34: Posibles disposiciones de la interfaz del prototipo: (a) Diseño 1, (b) Diseño 2, (c) Diseño 3 y (d) Pantalla completa

En la parte derecha, además de mostrar la hora actual, se ha colocado un botón desplegable que permite seleccionar un icono con la bandera de un determinado país. Cada uno de estos iconos sirve de acceso directo para cambiar el idioma de los textos de la interfaz de usuario, que pasará a ser el correspondiente al país cuya bandera ha sido seleccionada. Por defecto, la aplicación muestra los textos en pantalla en español. Se ha decidido ofrecer estos accesos directos para lograr que cualquier usuario, independientemente de su pericia en el manejo de aplicaciones informáticas, sea capaz de cambiar de forma rápida e intuitiva el idioma de la interfaz de usuario. De esta forma, se pretende que el primer contacto del usuario con el prototipo sea lo menos traumático posible.

En la zona central, existe un control deslizante que permite al usuario controlar el volumen de los altavoces o auriculares (Salida de audio), junto con una casilla de verificación Mute que sirve para suprimir el sonido.

4.10.3. Ventanas auxiliares

En la interfaz propuesta se contemplan tres ventanas auxiliares ubicadas en una columna lateral, cada una de las cuales está dedicada a una funcionalidad relevante de la aplicación.

Las funcionalidades que residirán en estas ventanas auxiliares son la reproducción del vídeo

transmitido por el instructor, la mensajería instantánea y la lista de participantes en la clase. Nuevamente este planteamiento coincide con la filosofía de la mayoría de las aplicaciones de e-learning síncrono. Independientemente de esta circunstancia, optar por ubicar estas funcionalidades en una columna lateral está más que justificado. Las tres funcionalidades tienen en común que resulta muy interesante que puedan estar visibles en todo momento sin alterar o reducir ostensiblemente el área de presentación. Cuando el instructor transmite vídeo, es evidente que la videoconferencia debe ser un elemento visible. Además, tampoco se precisa un área más grande, ya que el vídeo transmitido habitualmente en una clase virtual será de un tamaño reducido. El chat es la herramienta de comunicación básica para que los alumnos le planteen dudas al instructor, y por ello es fundamental el poder acceder rápidamente al mismo. La lista de participantes no es tan importante para los alumnos, pero sí para el instructor, ya que le permite controlar la presencia de los usuarios.

4.10.3.1. Videoconferencia

En base a los requisitos formulados en el apartado 2.4.2, sólo el instructor podrá transmitir audio y vídeo. Todos los usuarios recibirán el vídeo del instructor o presentador, incluido él mismo. Este flujo de vídeo se muestra en la ventana auxiliar **Vídeo del profesor**.

4.10.3.2. Mensajería instantánea

La ventana auxiliar **Chat** se emplea para mantener conversaciones textuales. Posee dos zonas principales: el cuadro de escritura y el cuadro de lectura. El primero se emplea para enviar mensajes a la conversación general y el segundo para leer los mensajes enviados y recibidos hasta el momento.

La razón por la que se ha decidido permitir en cualquier momento la comunicación textual entre todos los usuarios es porque ante cualquier problema o duda que surja durante la sesión, el chat permitirá notificar la incidencia o solicitar ayuda, ya que no requiere ninguna preparación previa para su funcionamiento tras haber entrado en la sesión y será menos sensible a congestiones severas de la red.

Para esta ventana auxiliar se ha considerado incluir un menú contextual, desplegable al hacer clic con el botón derecho sobre esta ventana auxiliar, que permitiría guardar la conversación en un fichero en el equipo local.

4.10.3.3. Lista de participantes

La función de la ventana auxiliar **Lista de participantes** es mostrar la lista de participantes presentes en la clase virtual. Cuando un usuario entra en la clase, automáticamente su nombre aparece en la ventana auxiliar **Lista de participantes** de todos los participantes, bajo la categoría correspondiente a su rol (profesores o alumnos). Como ya se mencionó anteriormente, instructor y alumnos emplean la misma herramienta cliente en las clases virtuales, por lo que la única diferencia substancial entre ambos son sus privilegios, los cuales limitan las acciones que los usuarios pueden realizar durante la clase virtual.

4.10.4. Área de presentación

En el área de presentación se pueden diferenciar cuatro zonas: la barra de herramientas principal, la barra de herramientas de anotación, las pestañas de los contenidos abiertos y el área donde específicamente se visualizan los espacios de trabajo de la pizarra virtual.

Cuando se comparte un contenido, ya sea una presentación de PowerPoint, un documento PDF, una imagen SVG o una pizarra en blanco, éste se mostrará en el área de presentación de todos los participantes. En la parte inferior existirán tantas pestañas como contenidos haya

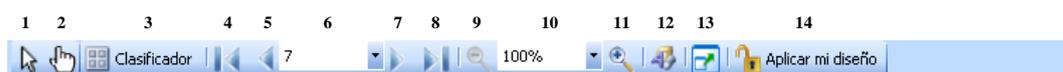


Figura 4.35: Barra de herramientas principal

abiertos. Las pestañas permiten cambiar ágilmente entre unos contenidos y otros. Al abrir por primera vez un documento, se mostrará en el área de presentación de todos los usuarios la primera diapositiva, ajustada al tamaño del área de presentación, manteniendo su relación de aspecto tal como se detalló en la figura 4.22. Para ello, todos los participantes de la clase virtual deben haber descargado previamente el documento a compartir. Si no fuera el caso, se le notifica al usuario que no puede abrir dicho documento y se le impide participar en el espacio de trabajo.

Si lo que se inicia es una pizarra compartida, aparecerá un lienzo en blanco ajustado igualmente al tamaño del área de presentación. Cuando se hace uso de las pestañas para cambiar a otro contenido abierto anteriormente (y que no haya sido cerrado), éste se mostrará en el área de presentación tal y como se dejó. Es decir, en el caso de una presentación, se visualizará la última diapositiva mostrada y todas las anotaciones efectuadas con anterioridad en la presentación se mantendrán; si se trata de una pizarra compartida, se mostrará con los dibujos que hubieran sido realizados previamente.

Cuando se cierra una presentación, todas las anotaciones realizadas se desechan y no aparecerán la siguiente vez que se abra el archivo. Lo mismo sucede con los dibujos y anotaciones de una pizarra compartida. Para cerrar el contenido que está siendo actualmente visualizado se puede pulsar en el aspa situada en la esquina inferior derecha, a la altura de las pestañas de selección de espacios de trabajo.

4.10.4.1. Barra de herramientas principal

La barra de herramientas principal del área de presentación es una parte fundamental de esta zona de la interfaz y por eso se ha fijado en la zona superior de ésta. Además, desplazándola a esta ubicación se consigue ganar un poco más de espacio para la columna lateral de ventanas auxiliares, respecto a la opción de situarla justo debajo del menú principal de la aplicación. Como cabría esperar, los botones de esta barra están íntimamente relacionados con el área de presentación y el espacio de trabajo de la pizarra mostrado actualmente. En la figura 4.35 se desglosan los diferentes botones y a continuación se menciona la función que desempeñan:

1. Puntero normal (deshabilitar puntero virtual).
2. Puntero virtual o telepuntero.
3. Mostrar el clasificador de páginas.
4. Retroceder a la primera página.
5. Retroceder a la página anterior.
6. Número de página actual.
7. Avanzar a la página siguiente.
8. Avanzar a la última página.
9. Disminuir el factor de zoom.
10. Factor de zoom actual.



Figura 4.36: Barra de anotaciones

11. Aumentar el factor de zoom.
12. Anotar sobre el espacio de trabajo de pizarra.
13. Presentación a pantalla completa.
14. Aplicar el diseño del instructor.

Como se puede observar, la denominación de la mayoría de los botones es autoexplicativa. El puntero normal no tiene mayor función que suprimir el puntero virtual cuando se está utilizando este último. Cuando el instructor activa el puntero virtual todos los usuarios pueden ver el puntero en sus pantallas. La herramienta **Clasificador** dispone las páginas del área de trabajo como miniaturas en pantalla para que el usuario pueda acceder rápidamente a una página determinada.

Los controles de navegación permiten avanzar y retroceder a través de las páginas del área de trabajo, mientras que los controles de zoom permiten aumentar o disminuir el tamaño con el que se visualizan los contenidos en pantalla.

El acceso a la utilidad **Anotar** se representa con un botón con texto para facilitar su localización, ya que es una de las funcionalidades más empleadas. Cuando un participante pulsa este botón aparece una nueva barra de herramientas específica para dibujar y escribir, situada por defecto justo debajo de la barra de herramientas principal. Esta barra de herramientas desaparece si se vuelve a pulsar sobre el botón **Anotar**.

Por último, los dos botones del extremo derecho de la barra de herramientas afectan a cómo se visualizan los contenidos del espacio de trabajo en pantalla. Mientras que **Pantalla completa** establece el modo de pantalla completa, **Aplicar mi diseño** obliga a que todos los participantes visualicen el mismo diseño de pantalla de la interfaz que el instructor, excepto si tuvieran concedido el privilegio para elegir el diseño libremente.

4.10.4.2. Barra de herramientas de dibujo

Esta barra aparece justo debajo de la barra de herramientas principal cuando cuando se pulsa el botón **Anotar** de la primera. Mientras no se vuelva a pulsar este botón la barra de herramientas de dibujo será visible.

En la figura 4.36 se muestran los botones de la barra de herramientas de anotación del área de presentación, cuyas herramientas se enumeran a continuación.

1. Pluma.
2. Resaltador.
3. Borrador.
4. Borrar todos.
5. Color de la pluma.
6. Color del resaltador.
7. Grosor de la pluma.

La funcionalidad y forma de empleo de todas las herramientas de dibujo son análogas a las de cualquier programa de edición gráfica. **Pluma** permite realizar trazos a mano alzada, **Resaltador** permite remarcar elementos de la pizarra, **Borrador** se utiliza para borrar trazos de forma individual, mientras que **Borrar todos** elimina todos los trazos que ha realizado el usuario sobre la página actual. Por otra parte, los controles **Color de la pluma** y **Color del resaltador** permiten modificar los colores con los que se realizan los trazos, mientras que **Grosor de la pluma** permite cambiar el ancho de los trazos realizados con la pluma.

