



## 6 Capa de Aplicación

### 6.4 Multimedia

RdE 2014-2015

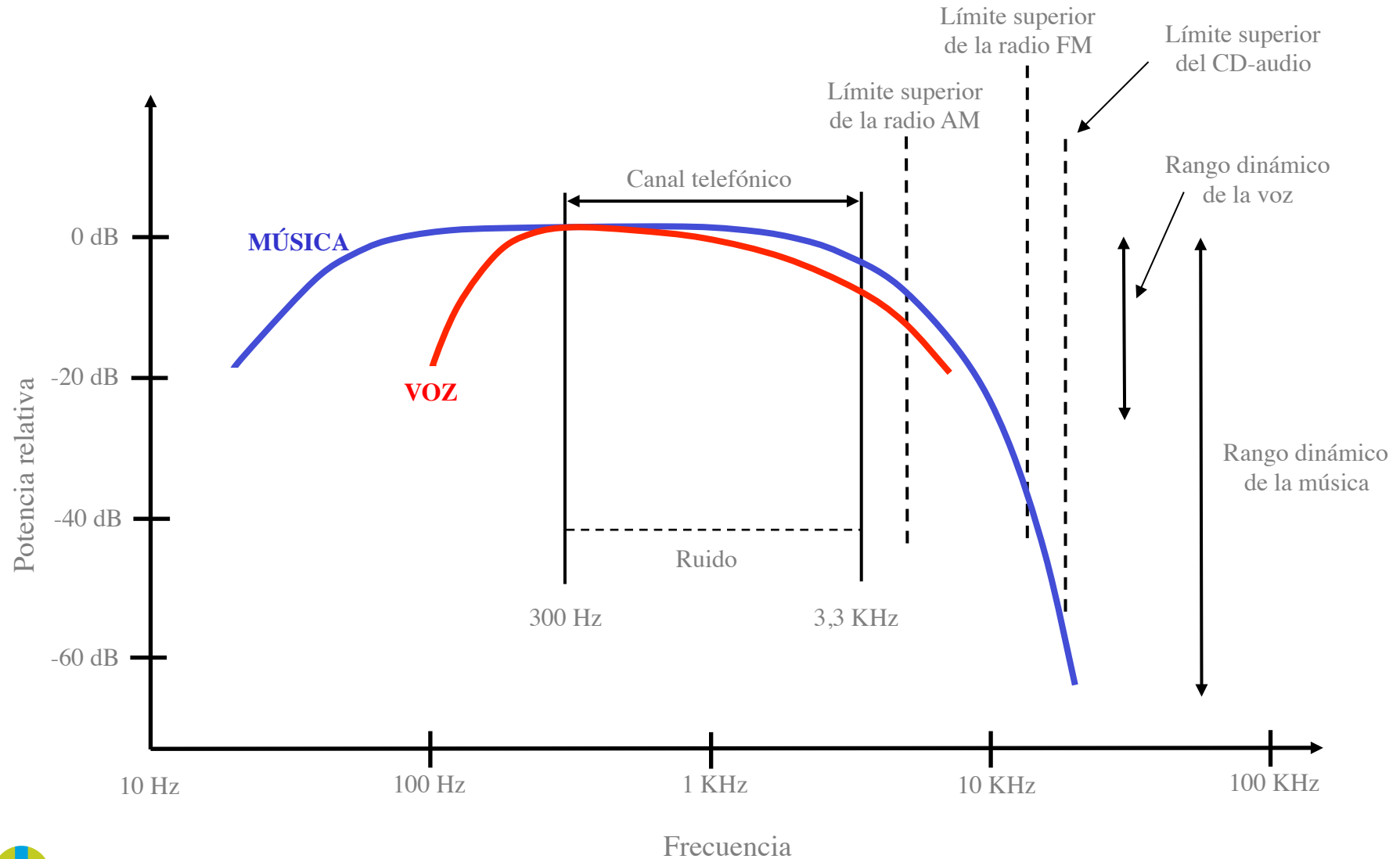
## 6 Guión del Tema 6

- 6. Capa de transporte:
  - 6.1 DNS.
  - 6.2 Correo electrónico.
  - 6.3 WWW.
  - **6.4 Multimedia.**

## 6.4 Índice

- 1. Fundamentos de telefonía. Conversión analógico-digital**
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

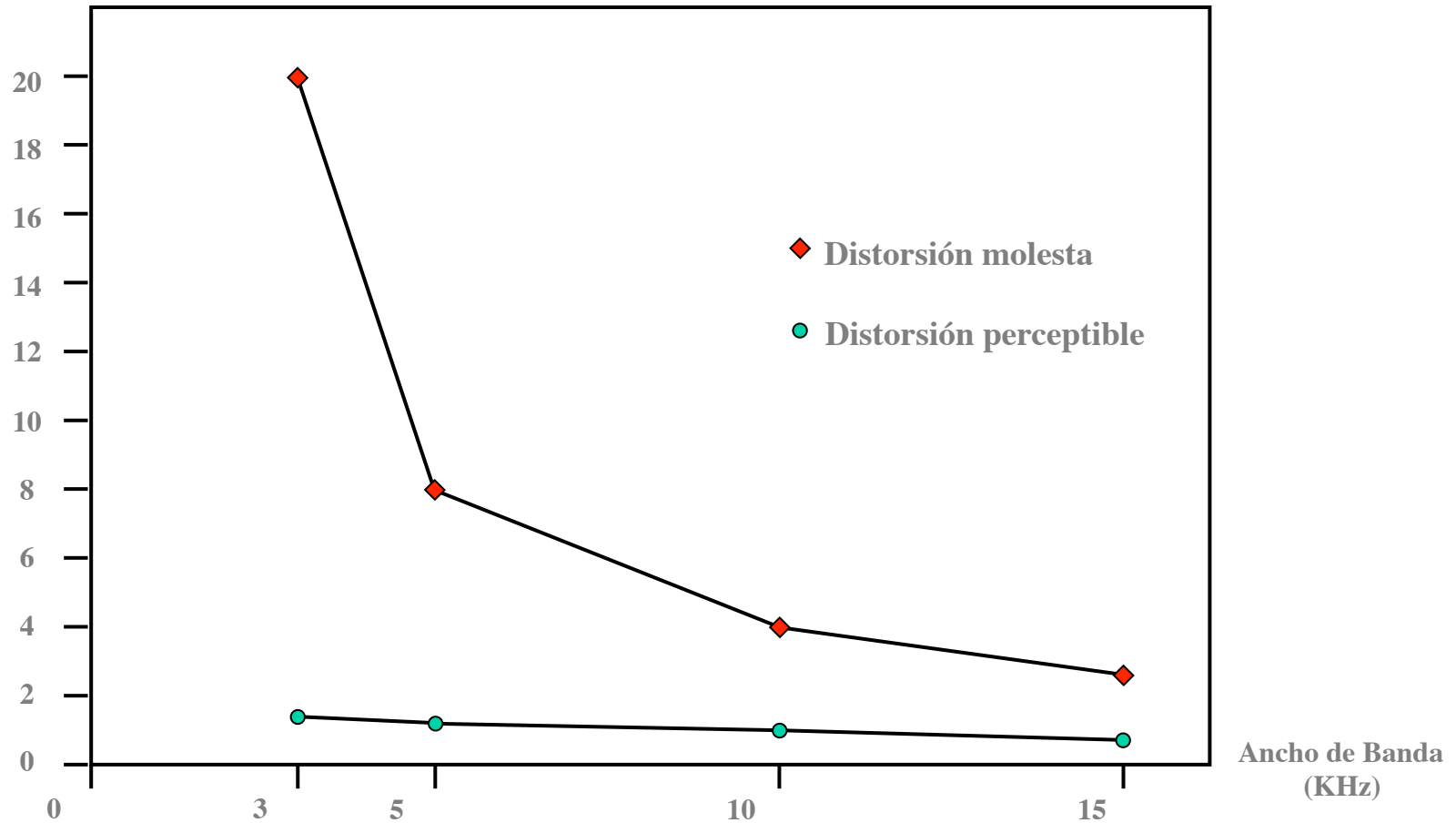
## 6.4.1 Espectro acústico de la voz y la música



## 6.4.1 Anchura de canal y distorsión

Al reducir la anchura del canal no solo se reduce el ancho de banda, también se reducen los requerimientos de calidad, ya que se tolera una mayor distorsión.

Distorsión (%)



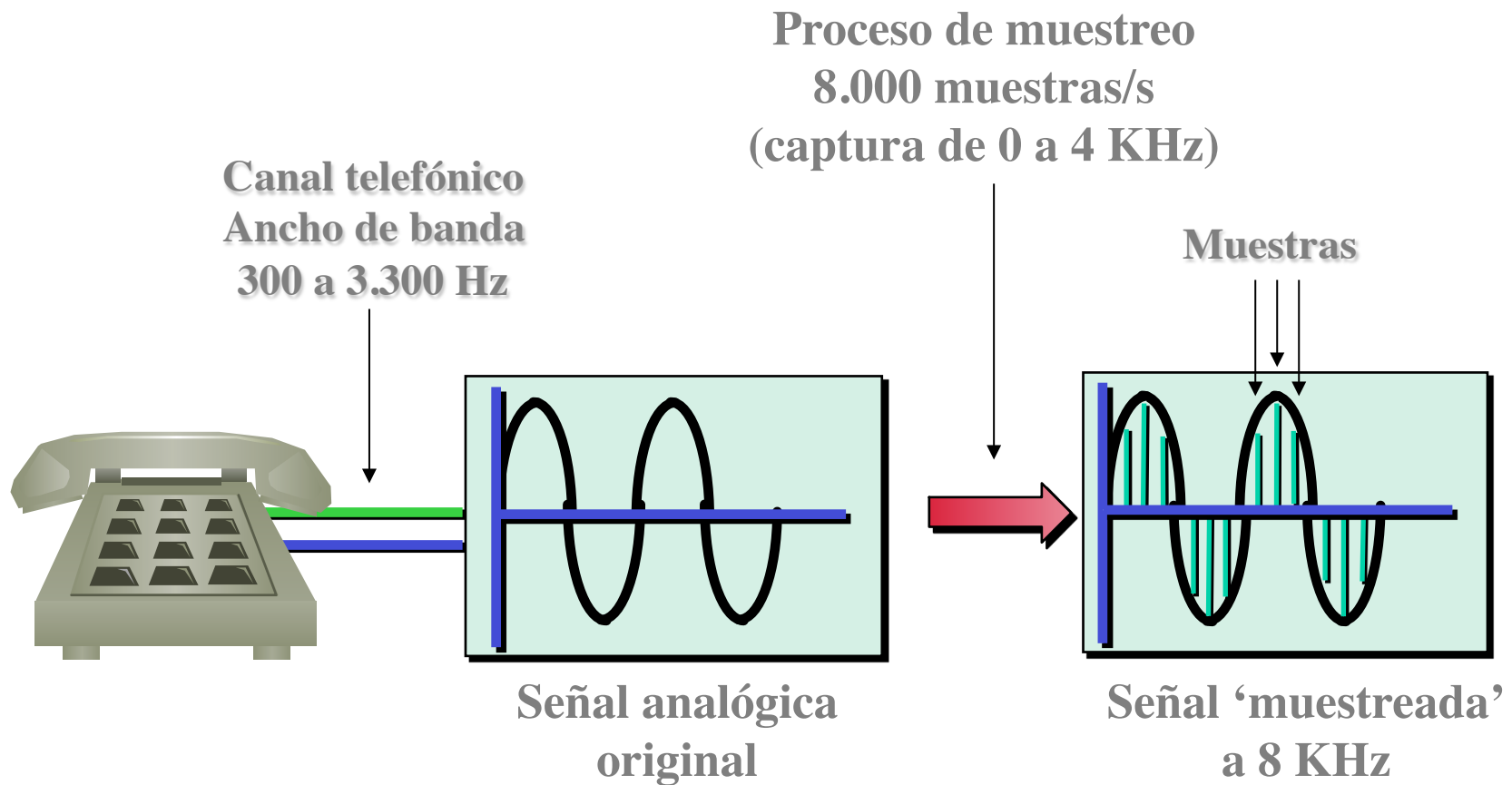
## 6.4.1 Diseño de la red telefónica

- Hacia 1930 la banda de frecuencias en una comunicación telefónica era de 300 a 3300 Hz, aproximadamente. Esta limitación se debía a la necesidad de abaratar los teléfonos y los equipos de transmisión. La relación señal/ruido (lo que se conoce como el 'rango dinámico') tampoco era demasiado elevada, en torno a 40dB.
- Se observó que estas condiciones tan 'pobres' eran suficientes para permitir una comunicación inteligible.
- Cuando, hacia 1960, se empezó a implantar la telefonía digital el objetivo fue igualar (no mejorar) la calidad que tenía la telefonía analógica.

### 6.4.1 Teorema de muestreo de Nyquist (1924)

- La digitalización de una señal analógica ha de hacerse muestreando al menos al doble de la frecuencia máxima que se pretende capturar.
- Ejemplos:
  - ❑ Canal telefónico: 3,3 KHz  $\rightarrow$  Muestreo 8 KHz
  - ❑ Audio HiFi: 20 KHz  $\rightarrow$  Muestreo 44,1 KHz

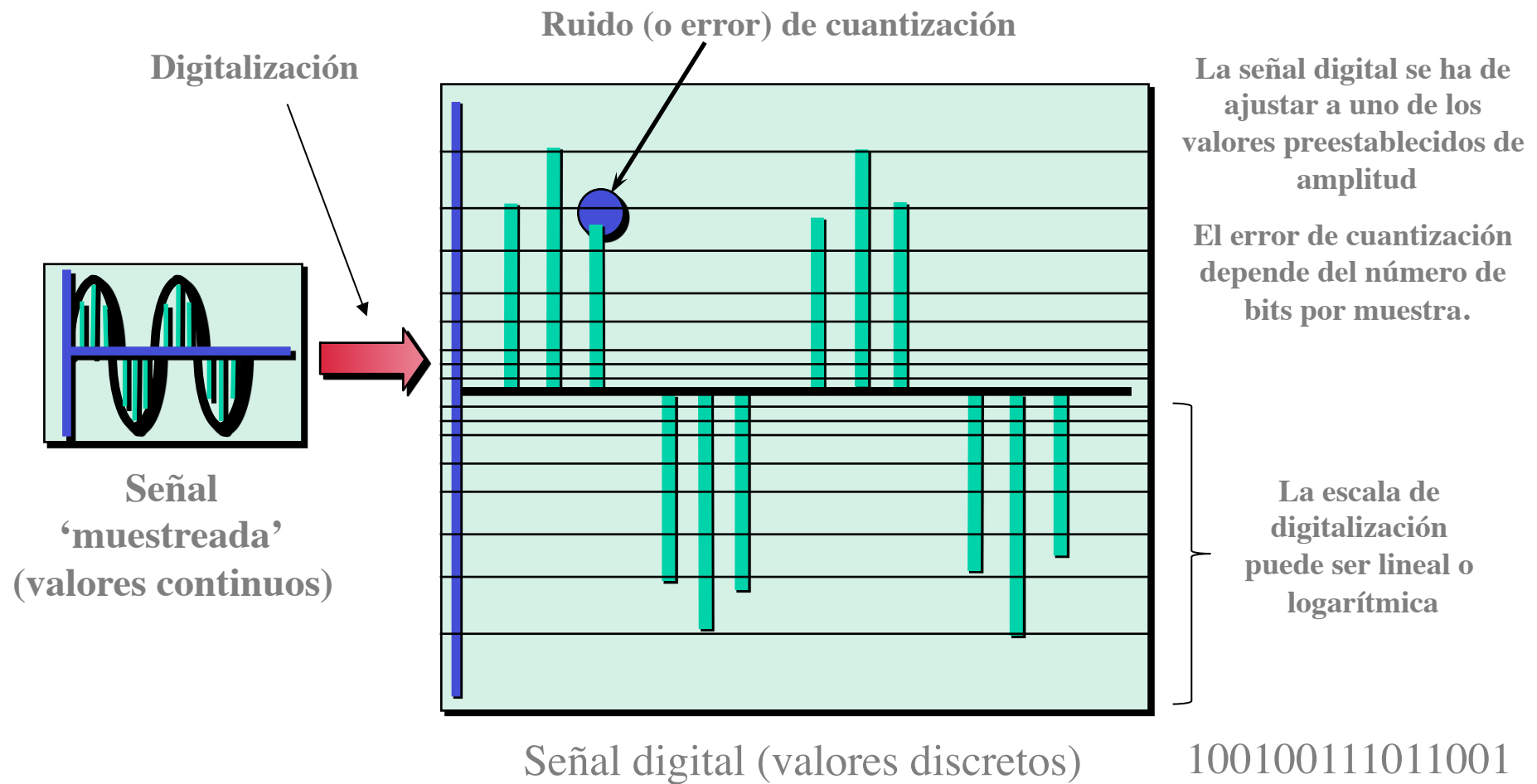
## 6.4.1 Conversión analógico-digital: muestreo de la señal







## 6.4.1 Conversión analógico-digital



## 6.4.1 Compromiso entre calidad y recursos (costo)

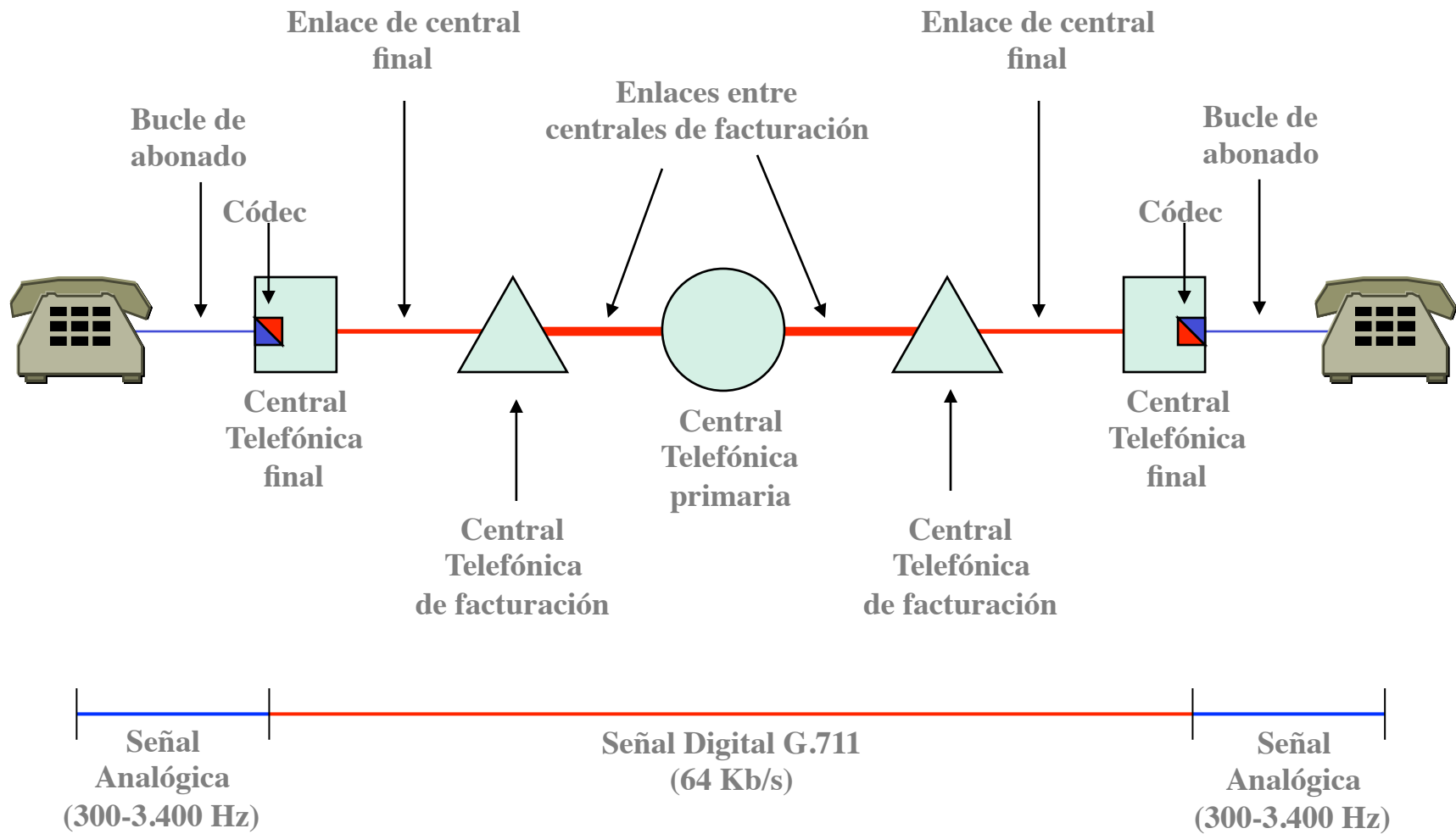


- Para digitalizar una señal analógica hay que fijar dos parámetros:
  - ❑ **Frecuencia de muestreo:** cuanto más alta mayor es el ancho de banda que se captura de la señal analógica.
  - ❑ **Número de bits por muestra:** cuanto mayor es menor es el error de cuantización y mayor la relación señal/ruido de la señal digital.
- En ambos casos un valor mayor supone una mayor fidelidad en la información digitalizada, pero también un mayor caudal en los datos generados.

## 6.4.1 Telefonía digital G.711 PCM

- La digitalización de audio para telefonía se viene haciendo en la red troncal desde los 1960s, por varias razones fundamentalmente:
  - ☐ Mayor calidad de sonido, especialmente en largas distancias cuando es necesario regenerar la señal.
  - ☐ Mayores facilidades para multiplexar varias conversaciones.
- En 1972 la ITU-T estableció el estándar G.711 para la telefonía digital, también llamado PCM (*Pulse Code Modulation*).
- El muestreo se hace con una frecuencia de 8 KHz, es decir una muestra cada 125  $\mu$ s.
- La amplitud se representa en una escala logarítmica utilizando 8 bits por muestra.
- El caudal utilizado es de 64.000 bits por segundo. Este caudal se utilizó después para el canal B de RDSI (Red Digital de Servicios Integrados).
- Existen dos variantes de G.711:
  - ☐ G.711  $\mu$ -law: usada en Norteamérica y Japón.
  - ☐ G.711 A-law: usada en el resto del mundo.

## 6.4.1 Comunicación entre teléfonos analógicos en una red moderna



## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
- 2. Audio digital. Estándares. Compresión**
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía sobre Internet
9. Protocolo SIP

## 6.4.2 Audio digital no comprimido

Tipo	Frec. de Muestreo (KHz)	Ancho de banda (KHz)	Bits por muestra	Relación S/R (dB)	Canales	Caudal
Sonido telefónico (G.711)	8	4	8 log	48	1	64 Kb/s (canal B de RDSI)
CD-DA (Compact Disc – Digital Audio) (Linear PCM)	44,1	22,05	16 lineal	96	2	1,411 Mb/s (CD-ROM 1x)
DVD-Audio	192 (máx.)	96	24 lineal (máx.)	144 (máx.)	2	9,216 Mb/s (capacidad de un DVD)

## 6.4.2 Diseño de la alta fidelidad (Hi-Fi)

- El diseño de la alta fidelidad se hizo con un criterio maximalista, es decir conseguir una calidad de sonido que abarque todo el rango de frecuencias que puede captar el oído humano (20-20.000 Hz), de forma que resulte indistinguible del original, para su almacenamiento o distribución.
- Como consecuencia de ello los parámetros elegidos fueron:
  - ☐ Anchura de canal: 20 KHz mínimo.
  - ☐ Relación señal/ruido: 90 dB mínimo.
- Cuando se estandarizó el CD-DA el objetivo fue mantener, e incluso mejorar, la calidad que tenía la Hi-Fi existente. La disminución en alguno de los parámetros del CD-DA podría haber dado lugar a una merma perceptible de calidad.

## 6.4.2 Algunos codecs de audio digital

	Formato	Ancho de banda (KHz)	Caudal (Kb/s)	Uso
<b>No comprimido</b>  <b>Bajo Retardo y, en general, baja calidad (telefonía)</b>	<b>PCM (G.711)</b>	3,3	64	Voz
	G.719	20	32 a 128	Voz y música
	SB-ADPCM (G.722)	7	48/56/64	Voz
	<b>SB-ADPCM (G.722.1)</b>	7	16/24/32	Voz y música
	<b>SB-ADPCM (G.722.1C)</b>	14	24/32/48	Voz y música
	<b>AMR-WB (G.722.2)</b>	7	12,65-23,85	Voz (GSM)
	<b>MP-MLQ (G.723.1)</b>	3,3	6,3/5,3	Voz
	ADPCM (G.726)	3,3	16/24/32/40	Voz
	LD-CELP (G.728)	3,3	16	Voz
	<b>CS-ACELP (G.729A)</b>	3,3	8	Voz
	LPC-10E (FS 1015)	3,3	2,4	Voz
	MELP (STANAG-4591)	3,3	0,6-2,4	Voz
	<b>ILBC (RFC 3951)</b>	3,3	13,33/15,2	Voz
<b>No comprimido</b>  <b>Elevado Retardo y alta calidad (música)</b>	CD-DA / DAT	22/24	705,6/768	Voz y música
	MPEG-1 Layer I	16/22/24	192-256	Voz y música
	MPEG-1 Layer II	16/22/24	96-128	Voz y música
	MPEG-1 Layer III (MP3)	16/22/24	64	Voz y música
	MPEG-2 AAC	16/22/24	32-44	Voz y música



## 6.4.2 Codecs de audio estandarizados

La mayoría de los *códecs* estandarizados de compresión de audio provienen de dos organismos:

- ❑ La ITU-T: son los estándares G.7xx pensados para telefonía (voz) aunque algunos son aptos para música y sonidos diversos. Suelen tener bajo retardo y poco consumo de CPU.
- ❑ La ISO: son los estándares de audio de MPEG pensados para películas (por tanto música y sonidos diversos). Suelen tener elevado retardo y gran consumo de CPU. Generalmente no aptos para telefonía.

## 6.4.2 Audio digital comprimido

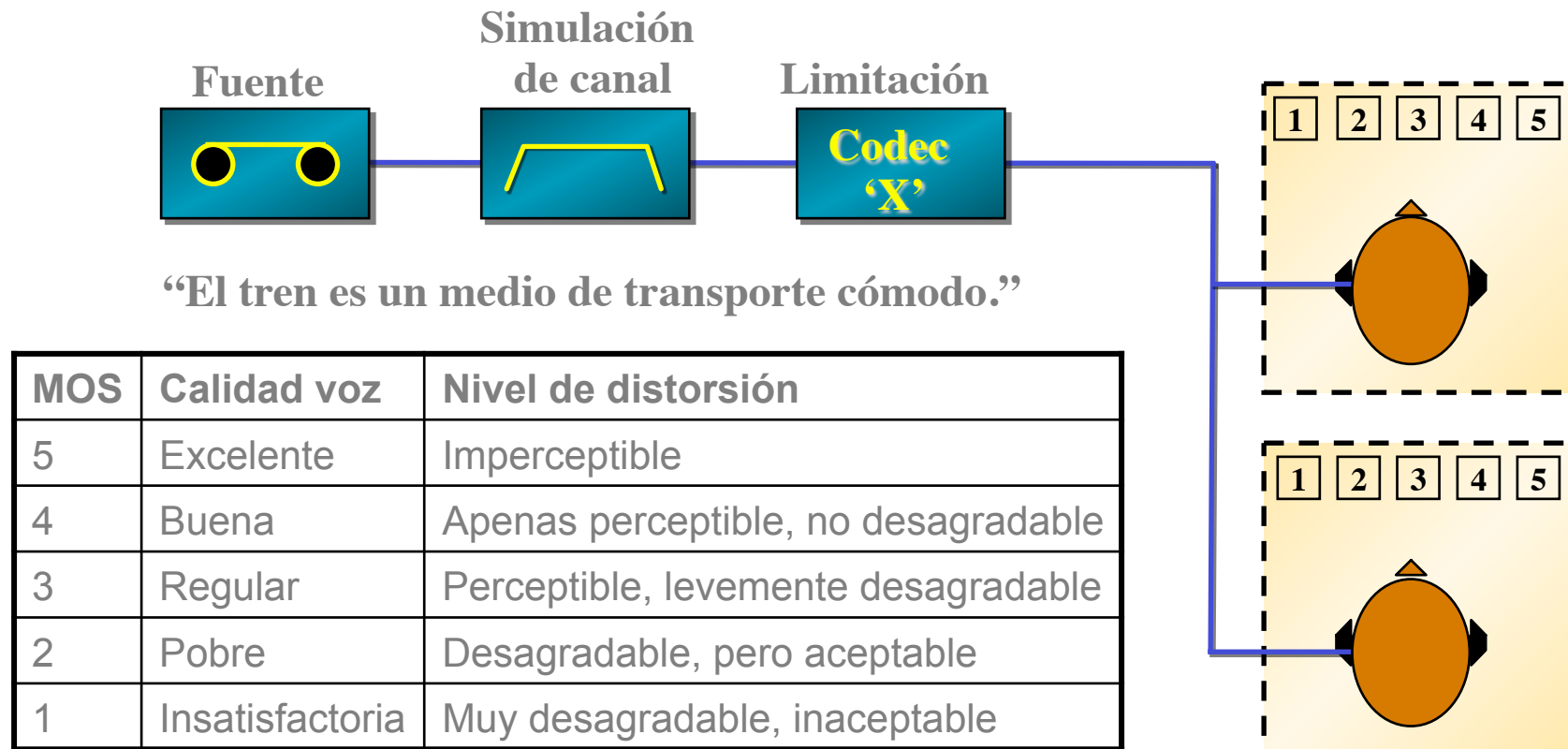
- En telefonía y videoconferencia se suele utilizar *codecs* G.xxx (estándares ITU-T) que dan un bajo retardo y una calidad adecuada para la voz, pero no para la música (excepto algunos como G.719 y G.722).
- La parte de compresión audio de MPEG (estándares ISO) es más eficiente (mejor ratio de compresión) y da mayor calidad, pero consume mucha CPU e introduce mucho retardo por lo que no suele emplearse en aplicaciones interactivas.
- Generalmente a más compresión menor calidad y mayor consumo de CPU.

## 6.4.2 Retardo y complejidad de *codecs* ITU vs ISO

Codec	Ancho de banda	Caudal	Retardo	Complejidad (CPU)
G.719	20 kHz	64-256 Kbps	40 ms	17,7 MIPS
MP3	20 kHz	64-384 Kbps	$\geq 54$ ms	$> 100$ MIPS
MPEG-4 AAC LD	20 kHz	24-192 Kbps	$\geq 20$ ms	$> 130$ MIPS

## 6.4.2 Comparación de codecs: MOS

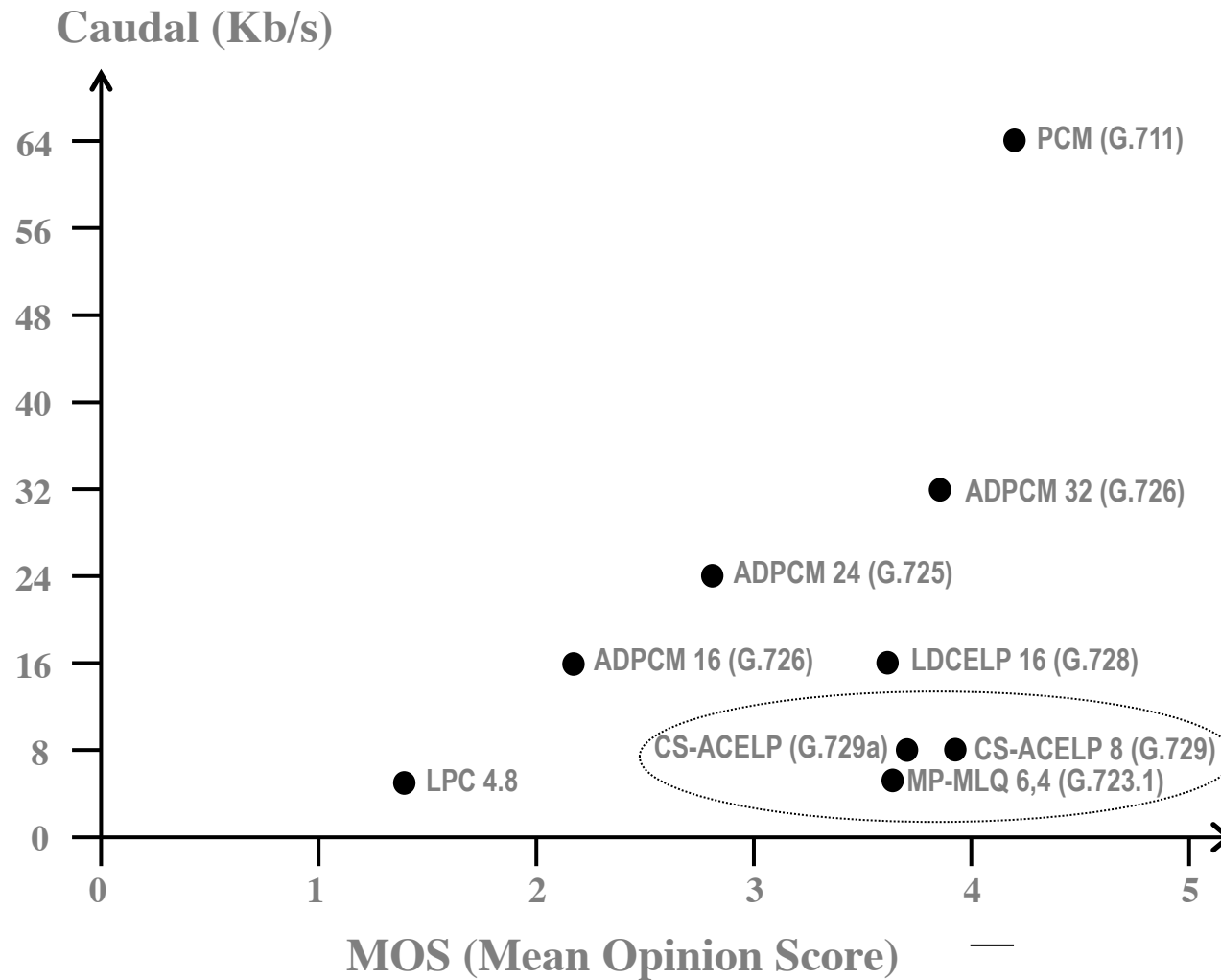
La comparación de calidad de *codecs* de audio no puede hacerse por técnicas objetivas. Se suele recurrir a encuestas de opinión puntuando entre 1 y 5, obteniendo lo que se denomina un **MOS** (*Mean Opinion Score*)



G.711 tiene un MOS de 4,2



## 6.4.2 Compresión vs calidad



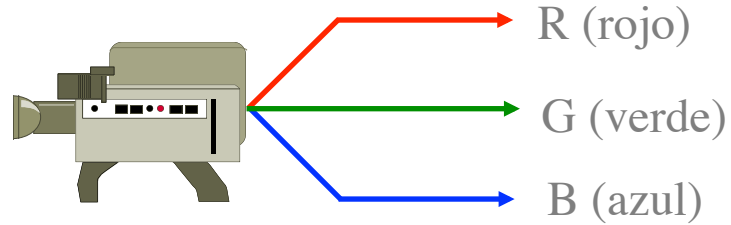
Normalmente para tener alta compresión con calidad y bajo retardo hay que ejecutar el algoritmo de compresión en hardware, en unos chips llamados **DSP** (*Digital Signal Processor*)



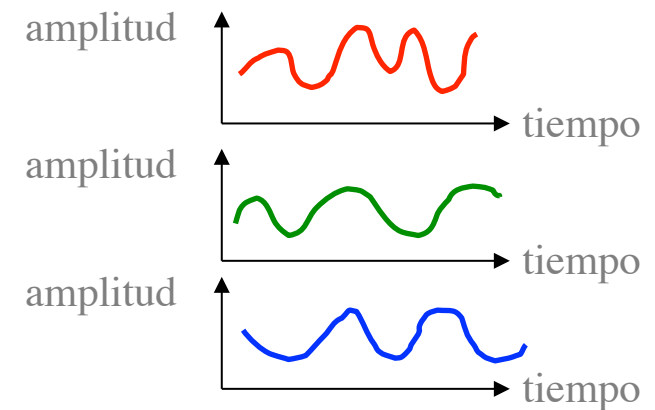
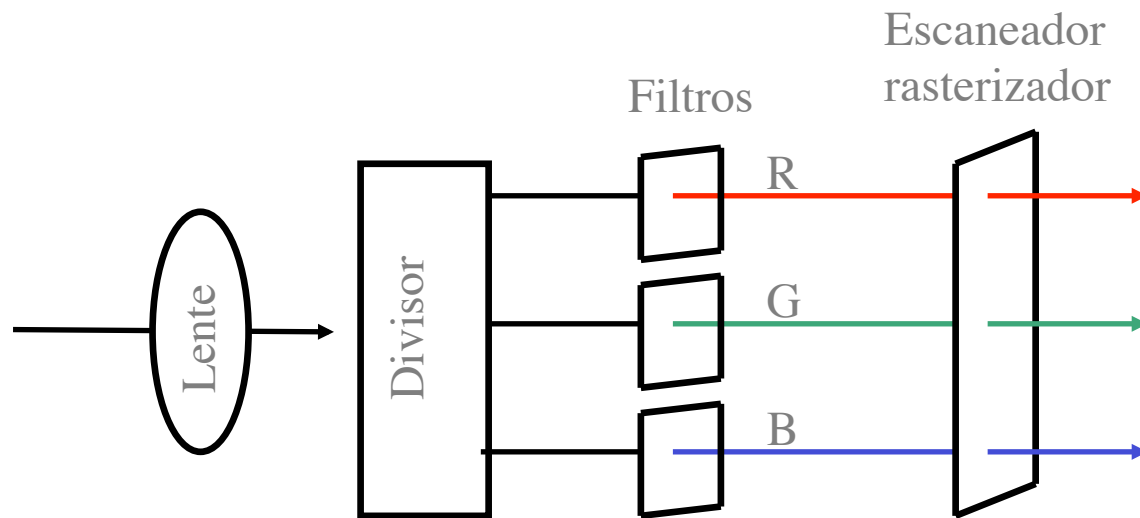
## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
- 3. Vídeo digital. Estándares. Compresión**
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

### 6.4.3 Señal de vídeo analógica



La imagen capturada se descompone en tres señales que corresponden a los colores primarios. A partir de ellos se puede reconstruir cualquier color

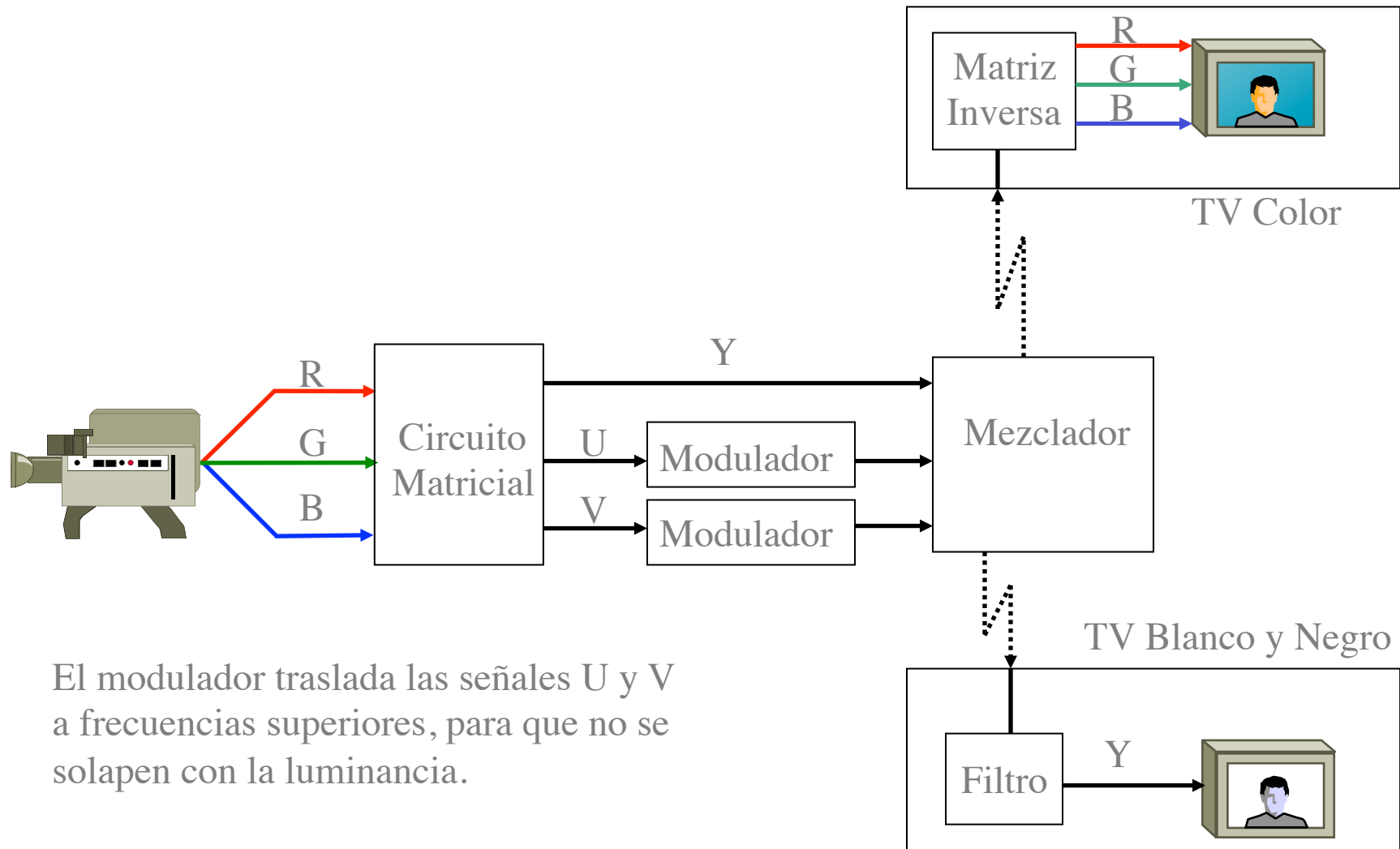


### 6.4.3 Fundamentos de la TV en color

- Las señales R-G-B obtenidas se transforman en otras tres, una de luminancia (Y) y dos de crominancia (U y V). Esta conversión se hace:
  - ❑ Para mantener la compatibilidad con televisión en blanco y negro (en ese caso se ignora la crominancia).
  - ❑ Para dar más ancho de banda a la luminancia, ya que el ojo es menos sensible a la crominancia.
- En el sistema PAL (europeo) la transformación se realiza aplicando las siguientes fórmulas:
  - ❑  $Y \text{ (Luminancia)} = 0,30 R + 0,59 G + 0,11 B$
  - ❑  $U \text{ (Crominancia)} = 0,493 (B - Y) = -0,15 R - 0,29 G + 0,44 B$
  - ❑  $V \text{ (Crominancia)} = 0,877 (R - Y) = 0,62 R - 0,52 G - 0,10 B$



### 6.4.3 Funcionamiento de la TV en color



El modulador traslada las señales U y V a frecuencias superiores, para que no se solapen con la luminancia.

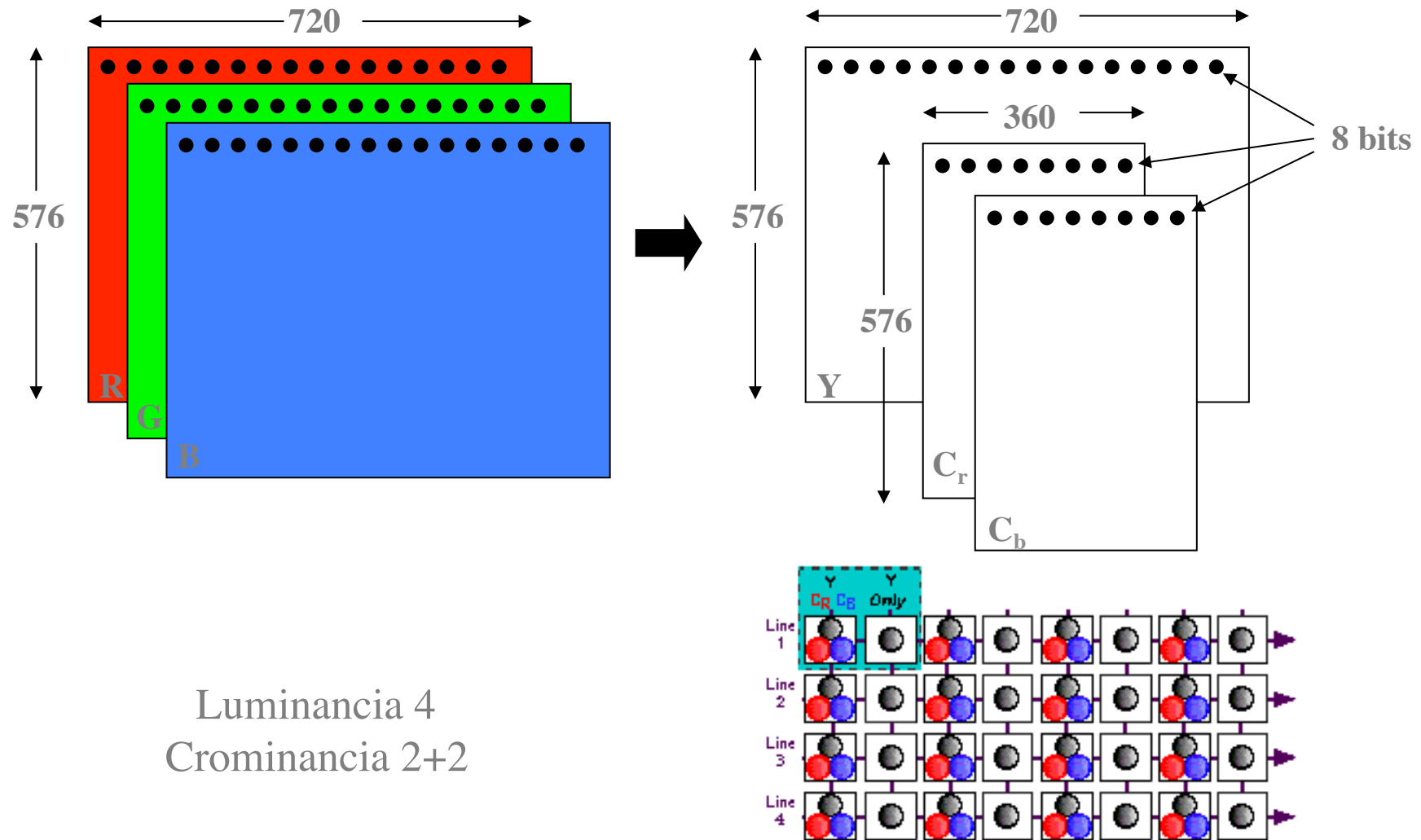
### 6.4.3 Video digital “no comprimido”

- El formato de vídeo digital utilizado como referencia en estudios de TV es el estándar ITU-R CCIR-601, también llamado D1.
- En CCIR-601 (y en otros formatos digitales) las componentes de crominancia se denominan  $C_r$  y  $C_b$  (en vez de U y V).
- En CCIR-601 PAL la parte de luminancia de cada fotograma se representa como una imagen de 720x576 píxels. Las de crominancia con 360x576:
  - ❑ Luminancia (Y):  $720(h) \times 576(v) \times 8 \text{ bits} \times 25 \text{ fps} = 82,944 \text{ Mb/s}$
  - ❑ Crominancia ( $C_r$ ):  $360(h) \times 576(v) \times 8 \text{ bits} \times 25 \text{ fps} = 41,472 \text{ Mb/s}$
  - ❑ Crominancia ( $C_b$ ):  $360(h) \times 576(v) \times 8 \text{ bits} \times 25 \text{ fps} = 41,472 \text{ Mb/s}$
- Caudal total:  $82,944 + 41,472 + 41,472 = 165,888 \text{ Mb/s}$

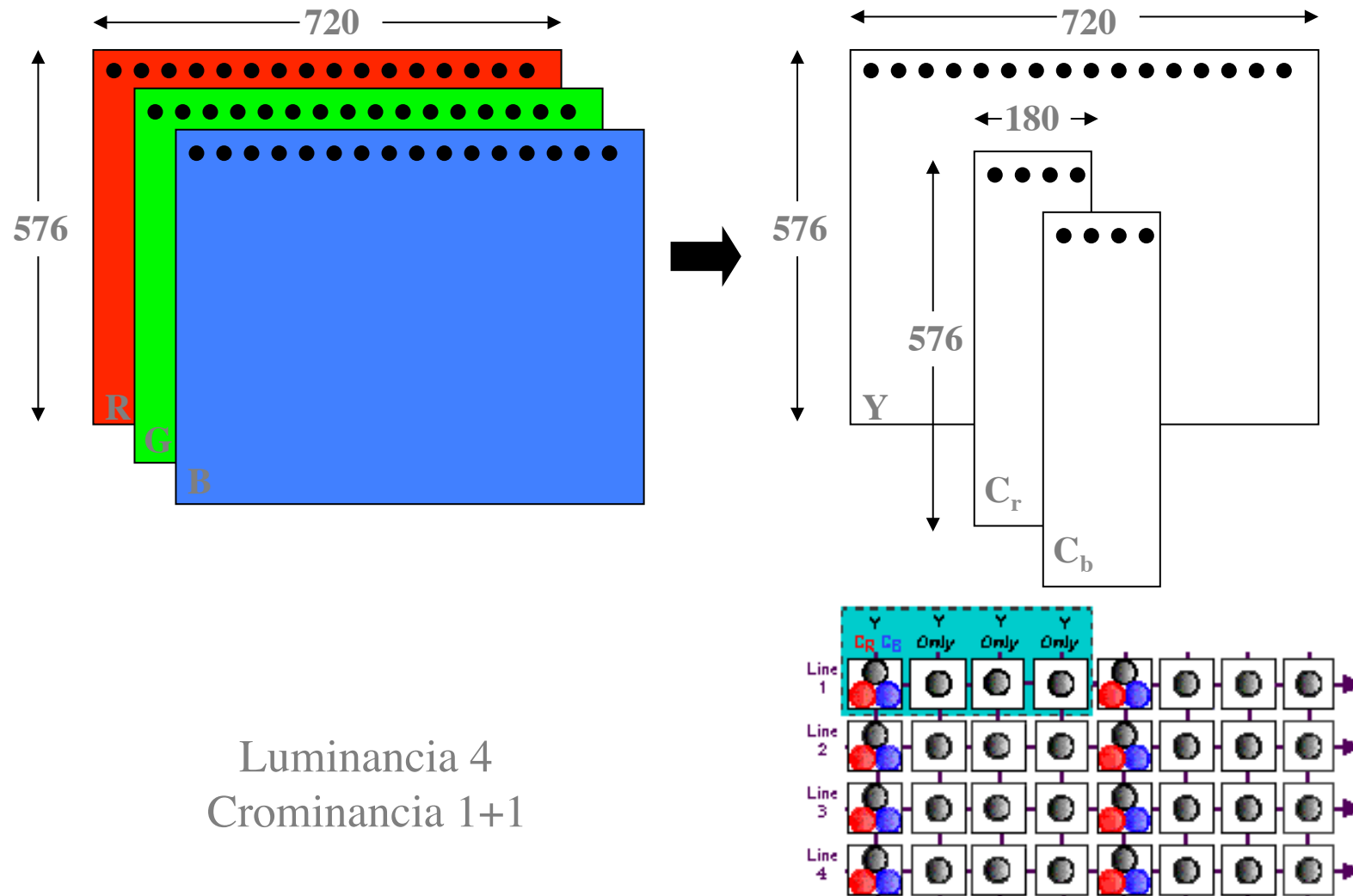
### 6.4.3 Submuestreo

- La reducción de la resolución en las componentes de crominancia se denomina **submuestreo** (equivalente a reducir la frecuencia de muestreo al digitalizar una onda).
- El submuestreo se basa en la menor sensibilidad del ojo humano a la crominancia.
- El submuestreo de CCIR-601, llamado 4:2:2, reduce la información de crominancia a la mitad. Sin submuestreo el caudal total sería 248,832 Mb/s.
- La información de crominancia puede reducirse aún más aplicando submuestreo 4:1:1 ó 4:2:0. Esto degrada un poco la calidad de color, pero la diferencia respecto a 4:2:2 es pequeña y sólo suele ser percibida por profesionales o en situaciones extremas.

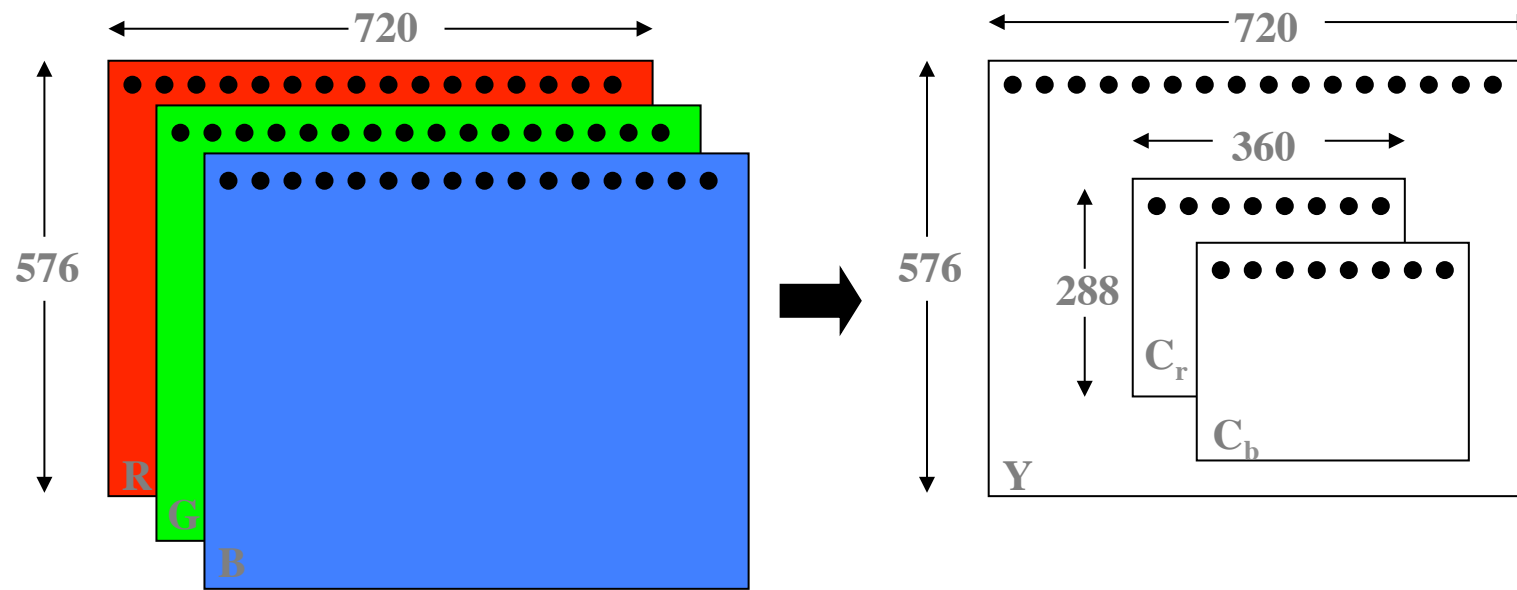
### 6.4.3 Submuestreo 4:2:2



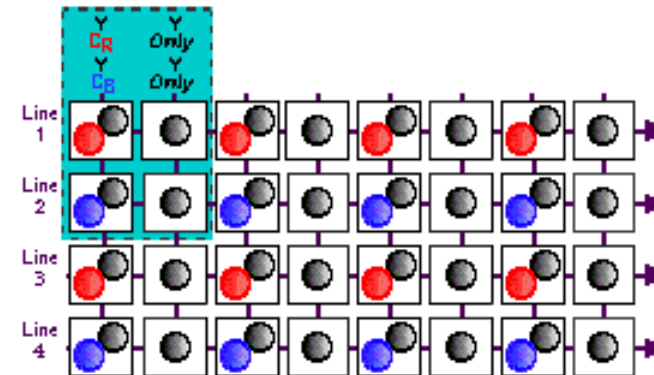
### 6.4.3 Submuestreo 4:1:1



## 6.4.3 Submuestreo 4:2:0



Luminancia 4  
Crominancia 2+0



### 6.4.3 Caudales de video digital en diferentes calidades (a 25 fps)

Calidad	Sub-muestreo	Bits/pixel	Resolución	Caudal no compr.	Caudal compr.
CIF (VHS)	4.1:1	12	352 x 288	30 Mb/s	56-256 Kb/s
SD (CCIR-601)	4:2:2	16	720 x 576	166 Mb/s	256-512 Kb/s
HD 720p	4:2:2	16	1280 x 720	369 Mb/s	1-2 Mb/s
HD 1080p	4:2:2	16	1920 x 1080	829 Mb/s	2-4 Mb/s
Dig. cinema 2K	4:4:4	36	2048 x 1080	3822 Mb/s	100 Mb/s
Dig. cinema 4K	4:4:4	36	4096 x 2160	7644 Mb/s	200 Mb/s

Conclusión: La compresión en vídeo digital es una necesidad

### 6.4.3 Compresión de vídeo

- Para la compresión de vídeo se aplican dos técnicas:
  - ❑ **Compresión espacial o intraframe:** se aprovecha la redundancia de información que hay en la imagen de cada fotograma, como en la imágenes JPEG.
  - ❑ **Compresión temporal o interframe:** se aprovecha la redundancia de información que hay entre fotogramas consecutivos.
- La compresión temporal siempre lleva implícita la espacial.
- La compresión temporal aumenta el consumo de CPU, pero mejora la eficiencia.



### 6.4.3 Compresión espacial de vídeo

- Consiste en la compresión de cada fotograma de forma independiente. Permiten la edición no lineal (edición independiente de cada fotograma):
  - ❑ M-JPEG: los fotogramas se comprimen con JPEG. Usado en las cámaras DV (25 Mb/s). No es un estándar.
  - ❑ M-JPEG2000: los fotogramas se comprimen con JPEG2000. Usado en vídeo de muy alta resolución (digital cinema). Forma parte del estándar ISO JPEG2000 (parte 3).
- Generalmente en estos casos el audio se registra como LPCM (*Linear Pulse Code Modulation*, calidad CD sin compresión).

## 6.4.3 Compresión temporal

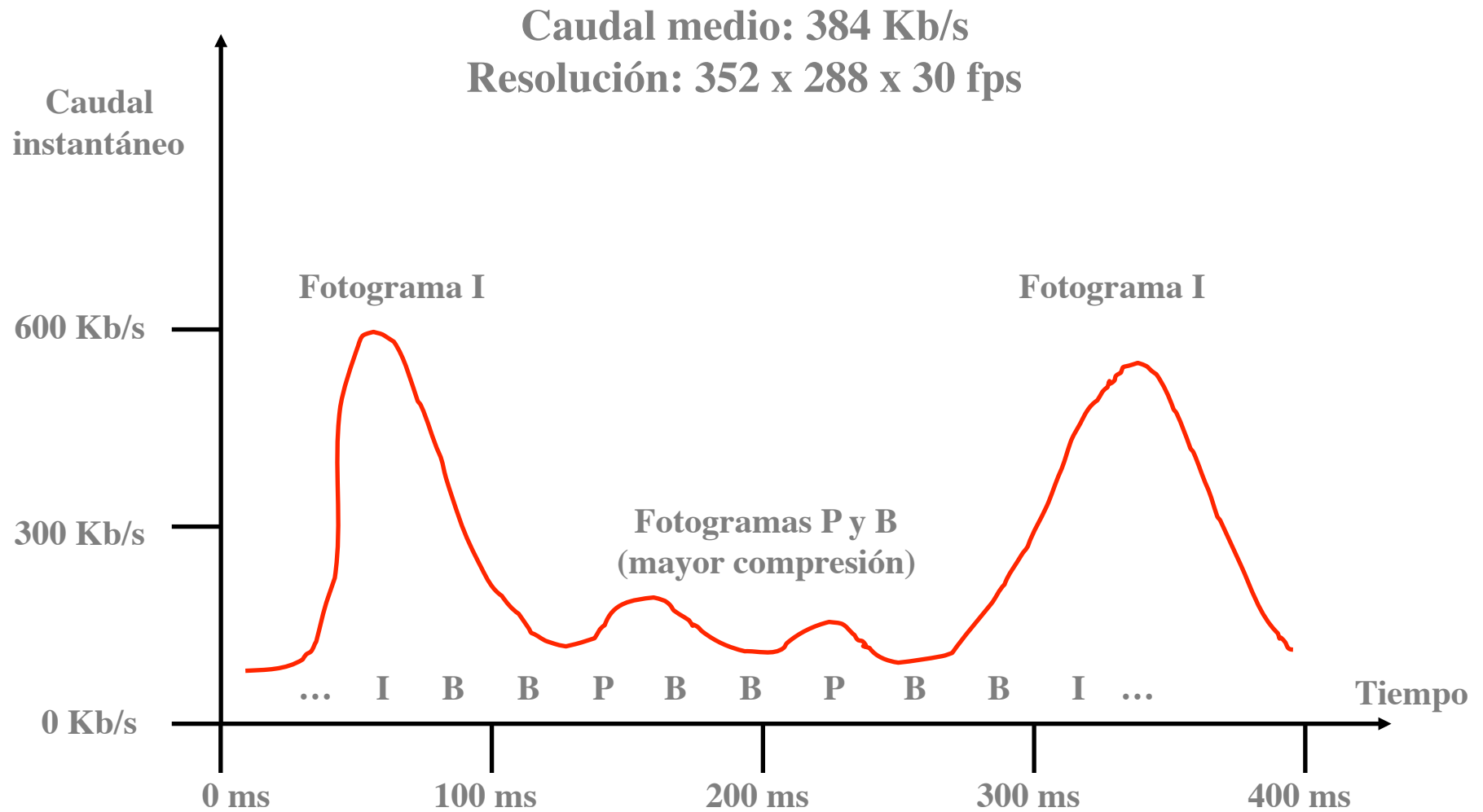
Se definen tres tipos de fotogramas:

- **I (Intra):** autocontenidos, solo compresión espacial (como JPEG).
- **P (Predictive):** referido al P/I anterior. Compresión temporal por extrapolación mediante macrobloques. Un macrobloque pueden ser:
  - ☐ Inalterado: no modificado respecto al fotograma de referencia.
  - ☐ Desplazado: (p. ej. un balón en movimiento) se describe por un vector de movimiento y eventualmente una corrección (diferencia respecto al original).
  - ☐ Nuevo: (p. ej. lo que aparece detrás de una puerta que se abre) se describe partiendo de cero por compresión espacial (como un fotograma I).
- **B (Bidireccional):** compresión temporal con interpolación; referido al P/I anterior y al P/I posterior. Máxima compresión, máxima complejidad de cálculo. Suaviza la imagen, reduce el ruido.

### 6.4.3 Compresión temporal

- Para detectar la redundancia de información los algoritmos de compresión temporal dividen la imagen de cada fotograma en una cuadrícula de **macrobloques**.
- Cada macrobloque está formado por un cuadrado de 8x8 ó 16x16 píxels. Por eso a veces cuando se pierden paquetes la imagen se ve ‘cuadriculada’.
- El tamaño de un fotograma (especialmente P ó B) puede variar mucho en función de la cantidad de información que deba incluir respecto al fotograma anterior.
- Los fotogramas B al ser una interpolación del fotograma anterior y posterior aumentan la complejidad de cálculo y la latencia.

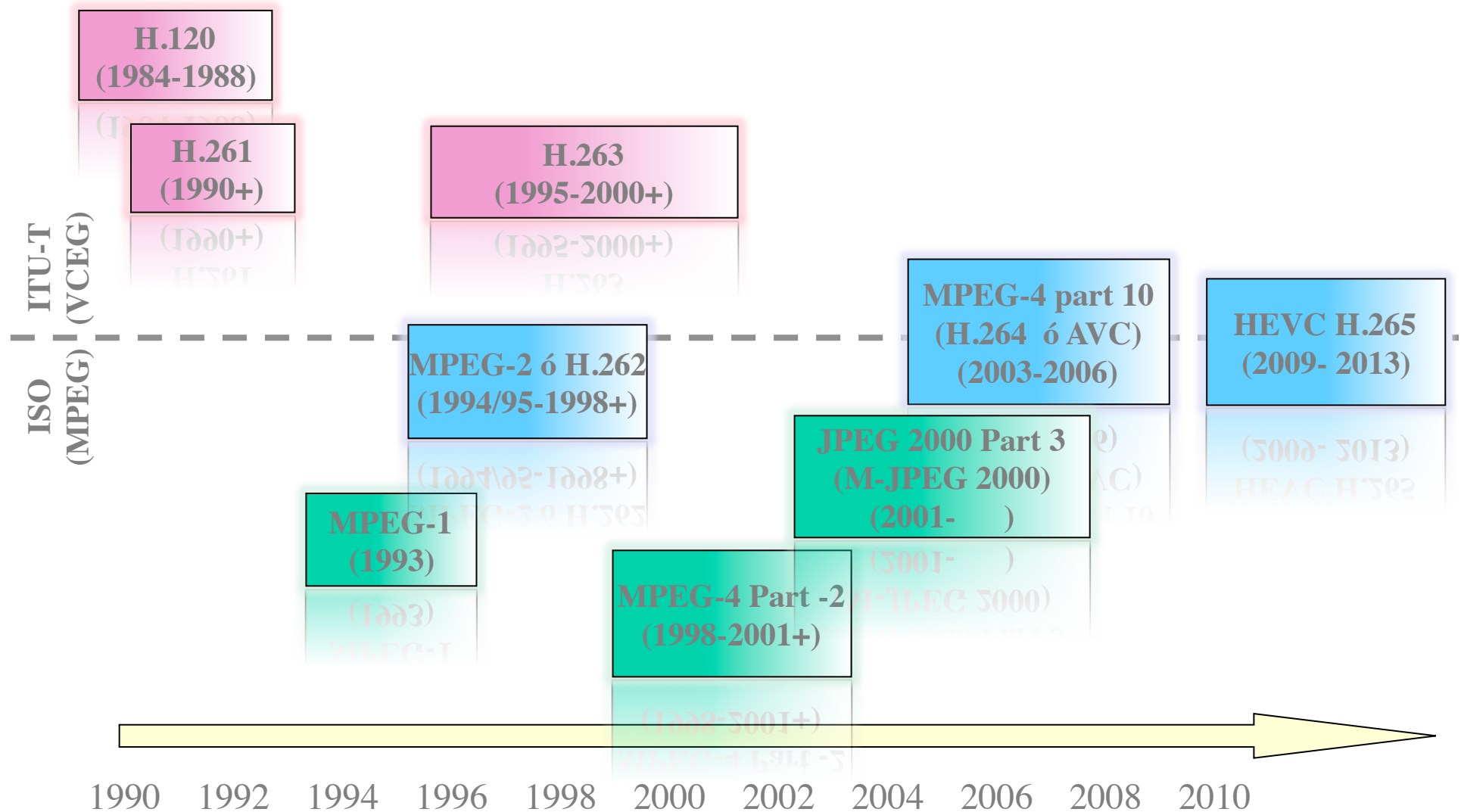
## 6.4.3 Caudal de una videoconferencia



### 6.4.3 Estándares de compresión de vídeo

- Los estándares de compresión de vídeo se llevan a cabo en el seno de dos grupos de trabajo:
  - ❑ El **MPEG** (*Moving Pictures Expert Group*) de la **ISO**: sus estándares se denominan MPEG-n ( $n = 1,2,4,7\dots$ ). Abarcan la compresión de audio y video.
  - ❑ El **VCEG** (*Video Coding Experts Group*) de la **ITU-T**: sus estándares se denominan H.26x ( $x = 1,3,4,5\dots$ ). Solo abarcan la compresión de video, la de audio se hace según los estándares de telefonía (G.xxx).
- Los algoritmos básicos son los mismos y en muchos casos pueden utilizar los mismos los ASICs (mismos chips).

## 6.4.3 Estándares de compresión de vídeo



## 6.4.3 MPEG 1 y 2

- **MPEG-1**

- ☐ Orientado a vídeo en CD-ROM (vídeo progresivo).
- ☐ Objetivo: Calidad VHS. Caudal típico 1,5 Mb/s.
- ☐ Útil para tele-enseñanza, aplicaciones de empresa, negocios, etc.

- **MPEG-2**

- ☐ Extensión compatible de MPEG-1 ‘hacia arriba’ (mayor calidad).
- ☐ Orientado a TV digital (vídeo entrelazado) calidad SD (*broadcast*) también HD. Caudales de 4 a 100 Mb/s.
- ☐ Útil para un rango amplio de aplicaciones mediante diferentes **niveles** (calidades) y **perfiles** (algoritmos).
- ☐ Es el más utilizado hoy en día (TV digital, DVDs, cámaras digitales).

### 6.4.3 Niveles y perfiles en MPEG-2

Nivel	Resolución max.	Calidad
Bajo	352 x 288	MPEG-1 (CIF, VHS)
Principal	720 x 576	SD ( <i>'Broadcast'</i> o CCIR-601)
Alto-1440	1440 x 1152	HD 4:3
Alto	1920 x 1152	HD 16:9

Perfil	Finalidad
Simple	Solo fotogramas I y P (para codecs de bajo costo)
Principal	El más utilizado
SNR	Alta calidad, escalable
Espacial	Para video tridimensional
Alto	Permite submuestreo 4:2:2 (además del 4:2:0)



### 6.4.3 Combinaciones permitidas de niveles/ perfiles en MPEG-2 y caudales máximos

Perfiles		Simple	Principal	SNR	Espacial	Alto
Submuestreo		4:2:0	4:2:0	4:2:0	4:2:0	4:2:0/2
Niveles	Alto 1920 x 1152 (HDTV 16:9)	No	80 Mb/s	No	No	100 Mb/s
	Alto-1440 1440 x 1152 (HDTV 4:3)	No	60 Mb/s	No	60 Mb/s	80 Mb/s
	Principal 720 x 576 (CCIR 601)	15 Mb/s	<b>15 Mb/s</b>	15 Mb/s	No	50 Mb/s
	Bajo 352 x 288 (MPEG1)	No	4 Mb/s	4 Mb/s	No	No

## 6.4.3 MPEG 3, 4 y HEVC

- **MPEG-3:** Grupo inicialmente pensado para HD, finalmente disuelto. Objetivo resuelto por reparametrización de MPEG-2.
- **MPEG-4:** ha aprobado multitud de estándares diferentes denominados ‘partes’ de los que dos tienen que ver con algoritmos de compresión de video:
  - ❑ **Parte 2:** basado en algoritmos DCT, como MPEG-1 y 2 pero un poco más eficiente. Resoluciones desde 176 x 144 hasta 1920 x 1080 y caudales desde 64 Kb/s hasta 38 Mb/s. Utilizado en *codecs* DivX, Xvid y Nero Digital.
  - ❑ **Parte 10:** también llamado H.264/AVC (*Advanced Video Coding*). Calidad comparable a MPEG-2 o MPEG-4 parte 2 con la mitad de caudal. Rango amplio de resoluciones (16 niveles) y algoritmos (14 perfiles). Se usa en algunas emisiones de HDTV y en videoconferencia (H.264).
- El siguiente estándar recién aprobado es el del grupo denominado HEVC (*High Efficiency Video Codec*) H265 que es una colaboración de la ISO y la ITU-T.

### 6.4.3 Estándares de vídeo ITU-T (VCEG)

- Desarrollados para videoconferencia:
  - ☐ H.120: Histórico, no se usa.
  - ☐ H.261(1980): desarrollado para RDSI (caudal constante). Poco utilizado actualmente.
  - ☐ H.262: es el MPEG-2. No se usa en videoconferencia.
  - ☐ H.263 (1995): Desarrollado para IP (caudal variable).
  - ☐ H.264 (2003): Coincide con MPEG-4 Parte 10.
  - ☐ [H.265](#) HEVC (2013): Ratificado en 2013.
- La videoconferencia generalmente usa un caudal mas bajo que el cine y tiene menos acción (los vectores de movimiento se restringen a +/- 15 pixels).
- Estos estándares no especifican la compresión de audio. Para ello se emplean los estándares G.xxx.

### 6.4.3 Compresión M-JPEG2000: Digital Cinema

- Resoluciones: 2K (2048x1080) y 4K (4096x2160).
- Caudal de vídeo hasta 250 Mb/s.
- 4K ofrece una calidad superior al cine de 35mm, comparable a una actuación en directo (límite del ojo humano).
- Codificación intraframe: Motion-JPEG2000 (MJ2, MJP2).
- Audio: muestreo de 24 bits, PCM a 48 ó 96 kHz sin comprimir. Hasta 16 canales.
- Se está estableciendo un estándar a través del DCI (Digital Cinema Initiatives) <http://www.dcinovies.com/>

## 6.4.3 Formatos compresión de vídeo

Codec	Compresión temporal	Complejidad Algoritmo	Eficiencia	Usos
M-JPEG	No	*	*	Cámaras DV
M-JPEG 2000	No	**	**	Digital Cinema
H.261	Fot. I y P	**	****	Videoconf. (antigua)
MPEG-1	Fot. I, P y B	***	***	VCD
MPEG-2, H.262	Fot. I, P y B	***	***	DVD, SDTV, HDTV
MPEG-4 parte 2	Fot. I, P y B	****	****	DivX, Xvid
H.263	Fot. I, P y B	****	****	Videoconf.
MPEG-4 Parte 10 H.264	Fot. I, P, B y otros	*****	*****	HDTV, videoconf.

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
- 4. Protocolos RTP y RTCP**
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

## 6.4.4 Protocolo de transporte para tráfico en tiempo real

- TCP presenta dos inconvenientes para tráfico en tiempo real:
  - ☐ En caso de reenvío de un paquete (por error o pérdida) la copia no llegará normalmente a tiempo de ser reproducida en su sitio.
  - ☐ El mecanismo de *slow-start* puede provocar que en caso de reinicio el caudal de la comunicación sea menor que el mínimo necesario.
- Por ese motivo normalmente se utiliza UDP.
- El único caso en que se utiliza TCP para tráfico en tiempo real es como 'truco' para cruzar los cortafuegos (que a veces no dejan pasar UDP). En esos casos se suele crear un buffer grande que retrasa la reproducción 10-30 segundos, de forma que los reenvíos sí pueden llegar a tiempo y el efecto del *slow-start* puede compensarse con el contenido del *buffer*.

## 6.4.4 Características de UDP

- UDP da un servicio '*best effort*' (del mejor esfuerzo) a nivel de transporte, igual que lo hace IP a nivel de red. En particular:
  - ☐ UDP no reenvía paquetes perdidos
  - ☐ UDP no descarta duplicados
  - ☐ UDP no garantiza que se respete el orden
  - ☐ UDP no realiza control de congestión



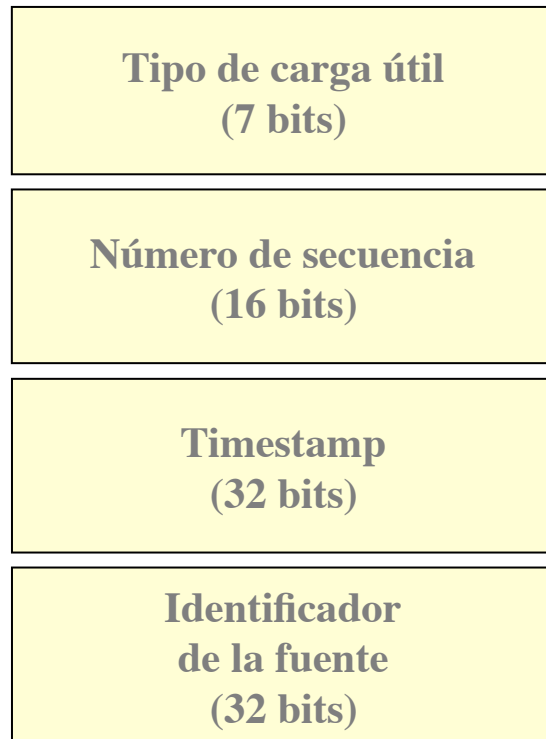
## 6.4.4 Protocolos RTP y RTCP

- Los protocolos RTP (*Real Time Protocol*) y RTCP (*Real Time Control Protocol*) complementan UDP cuando se envía tráfico en tiempo real.
- Sus funciones son:
  - ☐ Detectar los paquetes perdidos e informar de ello al emisor (el cual normalmente no reenviará esos paquetes).
  - ☐ Detectar y descartar duplicados.
  - ☐ Detectar paquetes que llegan fuera de orden.
- RTP y RTCP pueden funcionar en *unicast* y en *multicast*.

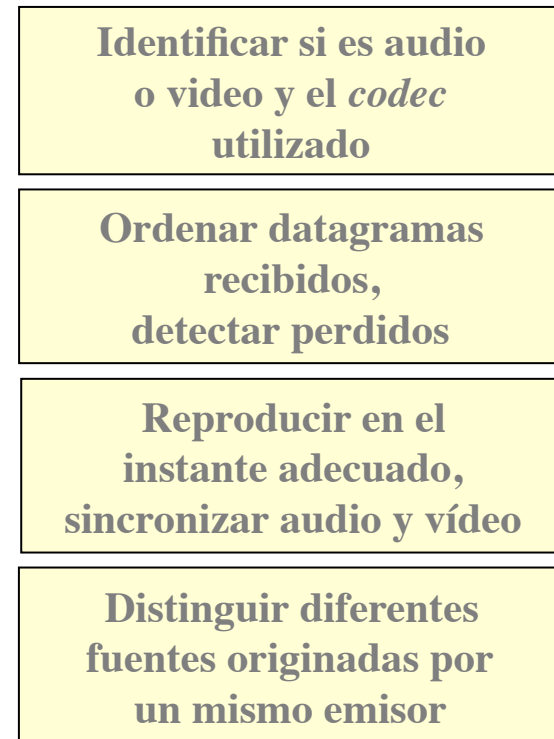
## 6.4 Estructura de un paquete RTP



La cabecera RTP incluye:



Con esto el receptor puede:



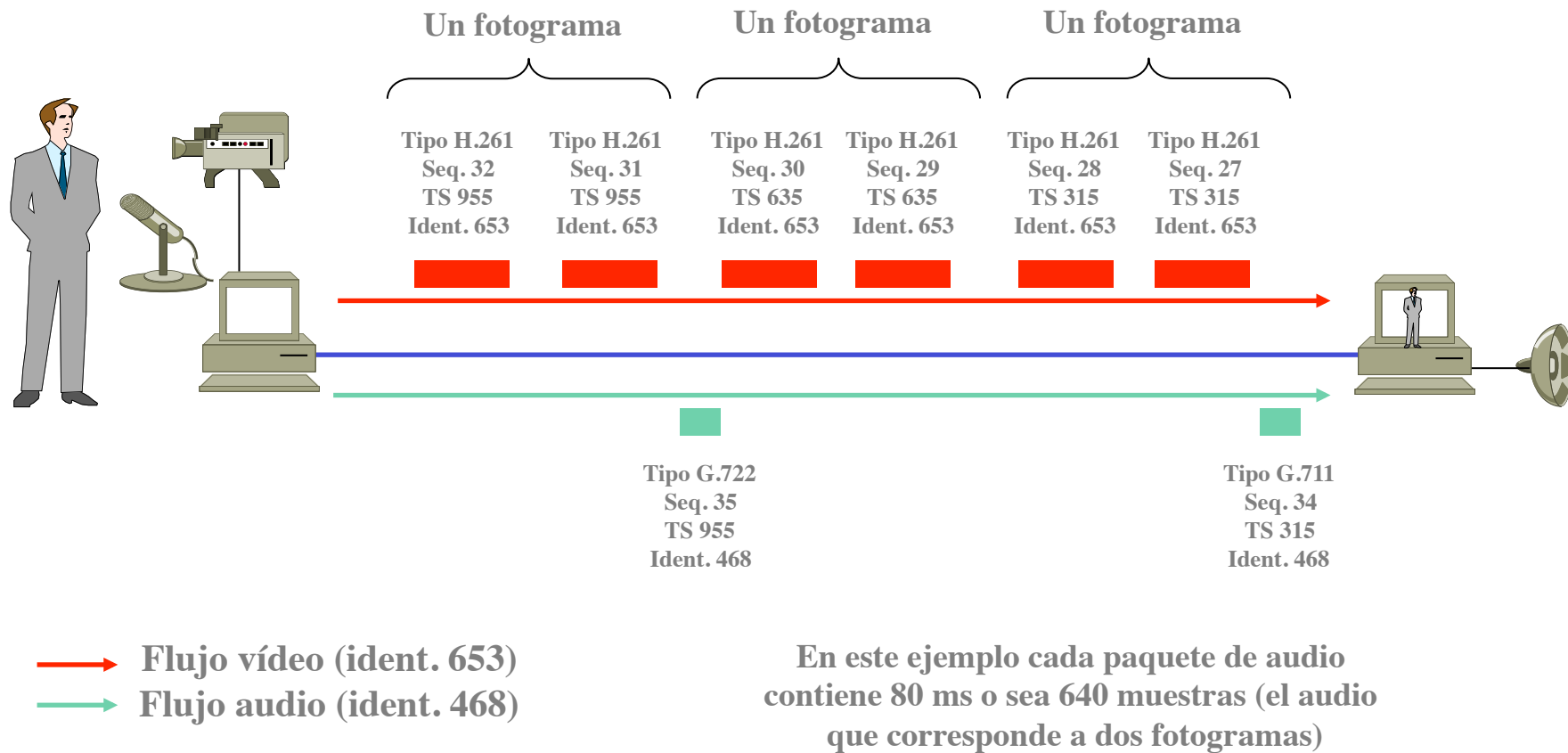
## 6.4.4 Valores Campo “tipo carga útil” en RTP

Valor	Significado
0	G.711 $\mu$ -law
9	G.722
14	MPEG Audio
15	G.728
26	M-JPEG
31	H.261
32	MPEG 1 video
33	MPEG 2 video

El valor se puede cambiar ‘al vuelo’ en una transmisión, por ejemplo un audio de G.711 a G.728

## 6.4.4 Flujos RTP en una videoconferencia

A 25 fps se emite un fotograma cada 40 ms



## 6.4.4 Mensajes RTCP

Tipo de mensaje	Enviado por	Contenido
SR ( <i>Sender Report</i> )	Emisores	Información sobre paquetes enviados y timestamp de sincronización
RR ( <i>Receiver Report</i> )	Receptores	Información sobre paquetes recibidos, perdidos, <i>jitter</i> y retardo medidos
SDES ( <i>Source Description</i> )	Emisores	Descripción de un emisor, p. ej.: nombre, e-mail, teléfono, identificación, etc.
BYE	Emisores y receptores	Cuando un participante se despide del grupo

## 6.4.4 RTCP

- Los mensajes RTCP no llevan información de audio/vídeo, sólo información de control.
- Los mandan tanto emisores como receptores. En una emisión *multicast* el RTCP convierte a todos los receptores del tráfico RTP en emisores del grupo *multicast* correspondiente (debido al envío de los *Receiver Report*).
- En principio una emisión RTP *multicast* con muchos receptores podría llegar a tener una parte importante de tráfico RTCP.
- Para evitarlo el tráfico RTCP se limita a un 5% del tráfico RTP. De ese 5% el 25% se asigna a los emisores (mensajes SR y SDES) y el 75% restante a los receptores (mensajes RR).

## 6.4.4 Control de congestión con RTP

- El receptor de un flujo RTP informa periódicamente al emisor del número de paquetes perdidos.
- El emisor normalmente supondrá que la pérdida se debe a congestión en la red (como hace TCP).
- En ese caso la aplicación puede intentar paliar el problema, por ejemplo pasando a un *codec* con mas compresión, bajando la resolución del vídeo o la frecuencia de muestreo del audio.
- RTP y RTCP permiten detectar las situaciones de congestión, pero en ningún caso intentan aplicar medidas correctoras o paliativas. Eso siempre queda al albedrío de la aplicación.

## 6.4 Resumen

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
- 5. Calidad de servicio**
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía sobre Internet
9. Protocolo SIP



## 6.4.5 Redes con calidad de servicio

- Desde principios de los 90 se ha experimentado en Internet con diferentes protocolos que permiten priorizar determinado tipo de tráfico considerado más importante o 'urgente'.
- Decimos que las redes que incorporan dichos protocolos ofrecen 'Calidad de Servicio' (QoS, *Quality of Service*) mientras que las que no lo hacen dan un servicio '*best effort*' (del mejor esfuerzo).
- La realidad es que la mayor parte de la Internet sigue funcionando en modo '*best effort*'. Ninguno de los protocolos de Calidad de Servicio se ha extendido de forma significativa.
- Algunos consideran que cuando aparecen problemas de congestión es hoy en día más sencillo y más barato ampliar capacidad que implementar QoS.

## 6.4.5 Requisito de las redes multimedia

- Las aplicaciones de audio-video en tiempo real han sido tradicionalmente los principales usuarios de redes con QoS, ya que el servicio se degrada rápidamente cuando la red va lenta.
- Los principales factores que influyen en esa degradación son los siguientes:
  - ☐ Pérdida de paquetes.
  - ☐ Retardo excesivo en la entrega de los paquetes.
  - ☐ Fluctuación excesiva en el retardo (*Jitter*).
- Los problemas se pueden deber a congestión en la red o a que simplemente se esté inyectando un caudal superior a la capacidad disponible

## 6.4.5 Multimedia sin QoS

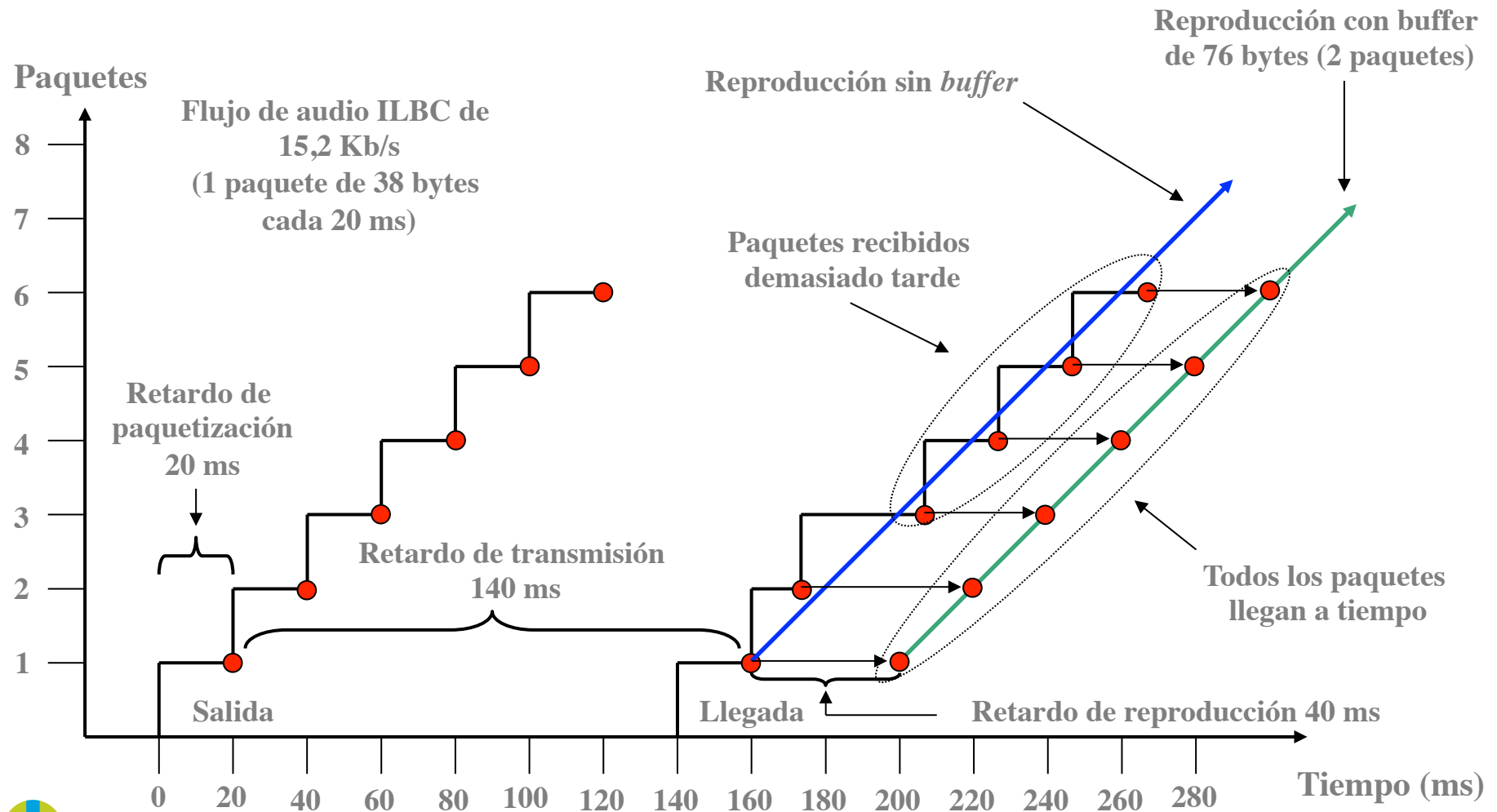
La necesidad de utilizar aplicaciones multimedia en redes sin QoS ha provocado el desarrollo de técnicas ‘paliativas’ tales como:

- Usar grandes búferes en el receptor. Esto solo es posible en aplicaciones no interactivas.
- Ajustar dinámicamente el ancho de banda utilizado.
- Reconstruir la información de los paquetes perdidos a partir de información colindante.
- Introducir un canal de comunicación en paralelo con información redundante.

## 6.4.5 Uso de grandes búferes

- El *jitter* puede compensarse introduciendo un buffer en el receptor que retrase la reproducción del flujo de audio o vídeo.
- Con un buffer suficientemente grande puede compensarse cualquier *jitter*, por grande que sea.
- Sin embargo esto puede introducir un retardo inaceptable cuando se trata de una comunicación interactiva (videoconferencia o telefonía).
- Las mayores exigencias en cuanto a retardo se dan en las aplicaciones de telefonía.

## 6.4.5 Compensación del *jitter* con un buffer de reproducción



## 6.4.5 Retardo máximo admisible

- La recomendación G.114 de la ITU-T establece retardos máximos en telefonía:
  - ☐ Se considera que un retardo por debajo de 150 ms es imperceptible y permite una buena calidad de comunicación.
  - ☐ Entre 150 y 400 ms el retardo se considera perceptible pero tolerable.
  - ☐ Por encima de 400 ms se considera excesivo e inaceptable.
- En comunicación unidireccional (p. ej. conferencias, clases magistrales) se puede tolerar un retardo bastante mayor.
- En aplicaciones donde el receptor no tiene ninguna posibilidad de interacción (p. ej. emisora de radio por Internet) se genera un retardo de reproducción de hasta 30 segundos, con lo que se pueden tolerar congestiones momentáneas (y se puede funcionar con TCP sin problemas).

## 6.4.5 Recomendación ITU-T G.114

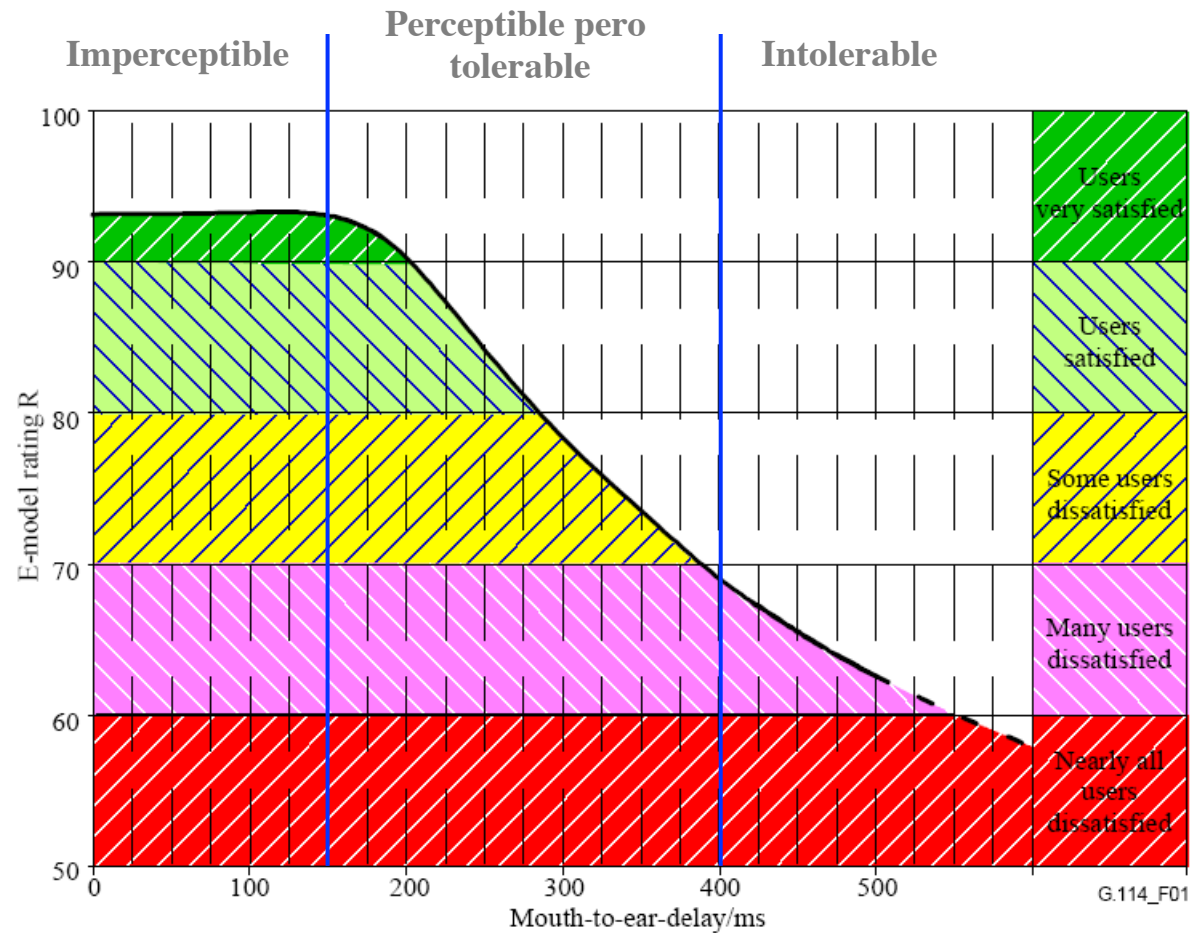


Figure 1/G.114 – Determination of the effects of absolute delay by the E-model

### 6.4.5 Ajuste dinámico del ancho de banda

- El receptor puede (mediante RTP) detectar los paquetes perdidos, e informar al emisor de la tasa de pérdidas mediante RTCP (*Receiver Report*).
- El emisor puede entonces reducir el caudal bajando la calidad (ej. video de 720x576 a 352x288) o pasando a un códec más eficiente (de H.263 a H.264), o ambos.
- Periódicamente el emisor debería probar a volver a la situación inicial por si la saturación era momentánea.



## 6.4.5 Reconstruir los paquetes perdidos

El receptor puede intentar minimizar el efecto producido por la pérdida de un paquete de varias maneras, por ejemplo:

- En vídeo se puede reconstruir un macrobloque ausente a partir de los macrobloques vecinos, o de los fotogramas anteriores o posteriores.
- En audio se puede rellenar un hueco extrapolando o interpolando a partir de las muestras de audio anteriores o posteriores.

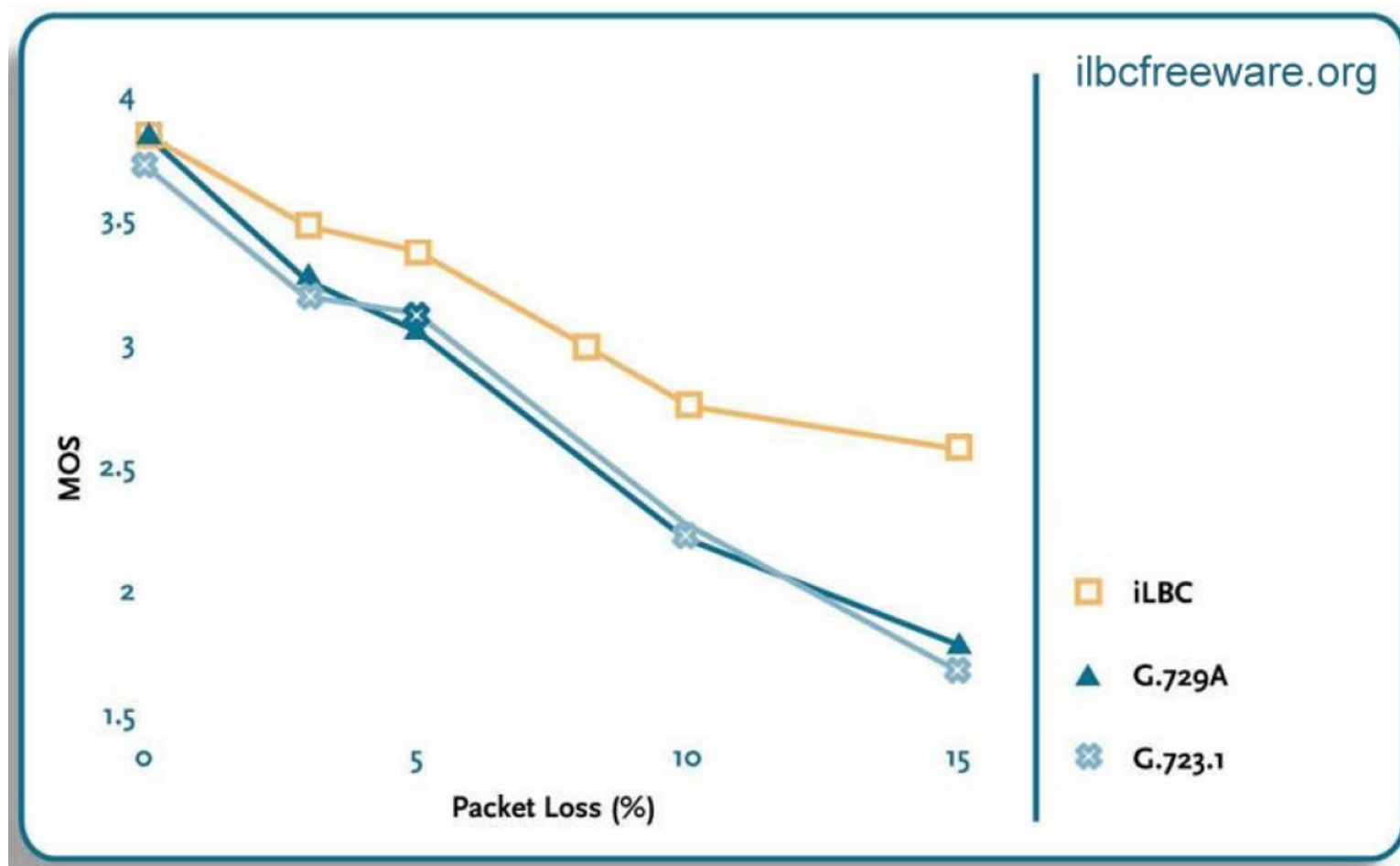
### 6.4.5 Información redundante

- Algunos sistemas envían un flujo paralelo con información redundante utilizando códigos correctores FEC (*Forward Error Correction*).
- Esto permite al receptor reconstruir los paquetes perdidos sin que el emisor los tenga que reenviar y sin tener que enviar el tráfico duplicado.
- Normalmente estos mecanismos se ajustan dinámicamente, de forma que cuando no se pierden paquetes no se envía información redundante.

### 6.4.5 Códec ILBC

- Para compensar la poca implantación de técnicas QoS en Internet se han desarrollado códecs resistentes a la pérdida de paquetes.
- El códec ILBC (*Internet Low Bitrate Codec*) ha sido desarrollado por el IETF (RFC 3951) con el objetivo de permitir una calidad aceptable aun en el caso de que produzca la pérdida de una cantidad significativa de paquetes.
- ILBC es utilizado por Skype, Google Talk, Yahoo Messenger, Gizmo5, QuteCom y Ekiga.

## 6.4.5 Comparación de los codecs ILBC, G.729A y G.723.1



## 6.4 Resumen

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
- 6. Videoconferencia. Estándares H.32x**
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

## 6.4.6 Aplicaciones de audio-vídeo en tiempo real

Aplicación	Sentido	Retardo tolerable	Espectadores	Multicast
Audio/Video conferencia (telefonía)	Bidirecc.	150-400 ms	Uno o varios	Apropiado
Audio-Vídeo bajo demanda*	Unidirec.	5-10 s	Uno	No
Emisión en directo (radio-TV por Internet)	Unidirec.	10-30 s	Muchos	Muy Apropiado

(\*) En el audio-vídeo bajo demanda el usuario puede controlar la emisión, de ahí el requerimiento de un retardo no excesivo

## 6.4.6 Videoconferencia

- Comunicación interactiva por medio de audio y video. Opcionalmente puede haber compartición de datos.
- Puede ser:
  - ☐ Punto a punto
  - ☐ Punto a multipunto
  - ☐ Multipunto a multipunto

## 6.4.6 Requisitos/Características de la videoconferencia

- Compresión/descompresión en tiempo real.
- Retardo máximo 400 ms.
- Movilidad reducida.
- Normalmente aceptable audio de calidad telefónica.
- Necesidad de sincronizar audio y vídeo.
- Necesidad de protocolo de señalización (servicio orientado a conexión).



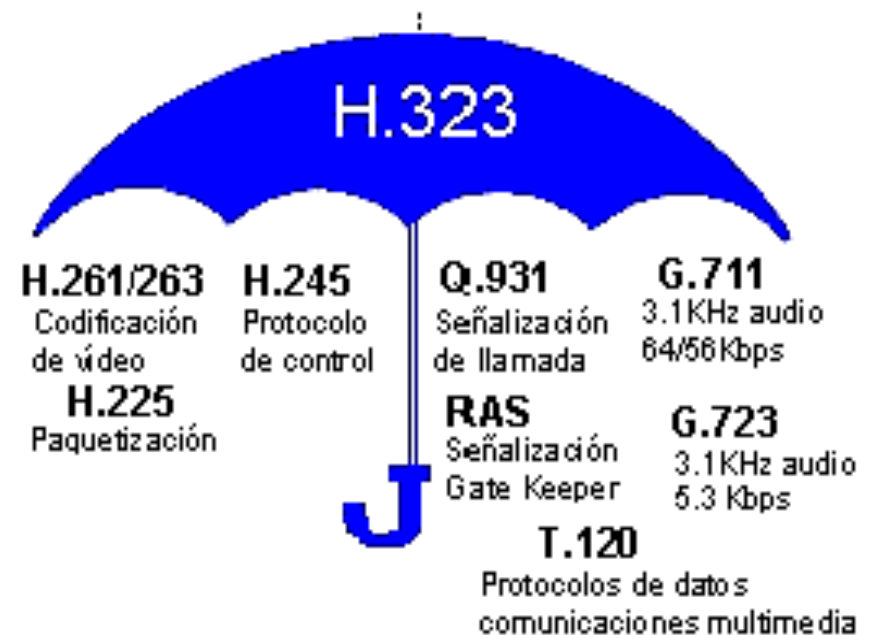
## 6.4.6 Estándares de Videoconferencia

- Los estándares H.32x (x = 0, 1, 2, 3 ó 4) de la ITU-T establecen todo lo relativo a videoconferencia. La 'x' depende del tipo de red utilizada.
- Son estándares 'paraguas' ya que en muchos casos se basan en otros estándares. Por ejemplo G.7xx para el audio, H-26x para el vídeo.
- La serie H de la ITU-T se refiere a sistemas multimedia y audiovisuales.

## 6.4.6 Estándares H.320 y H.323

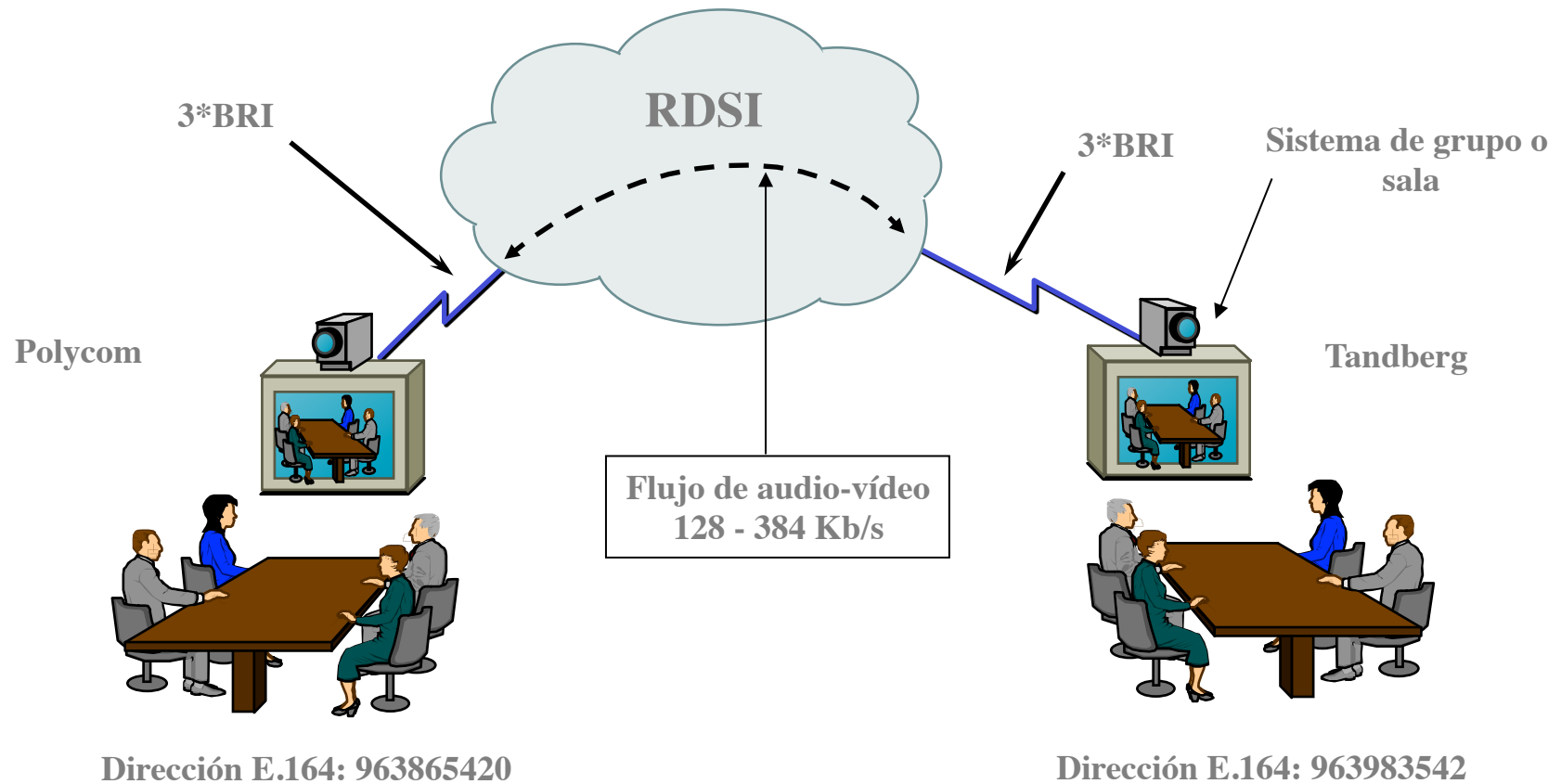


RDSI



IP

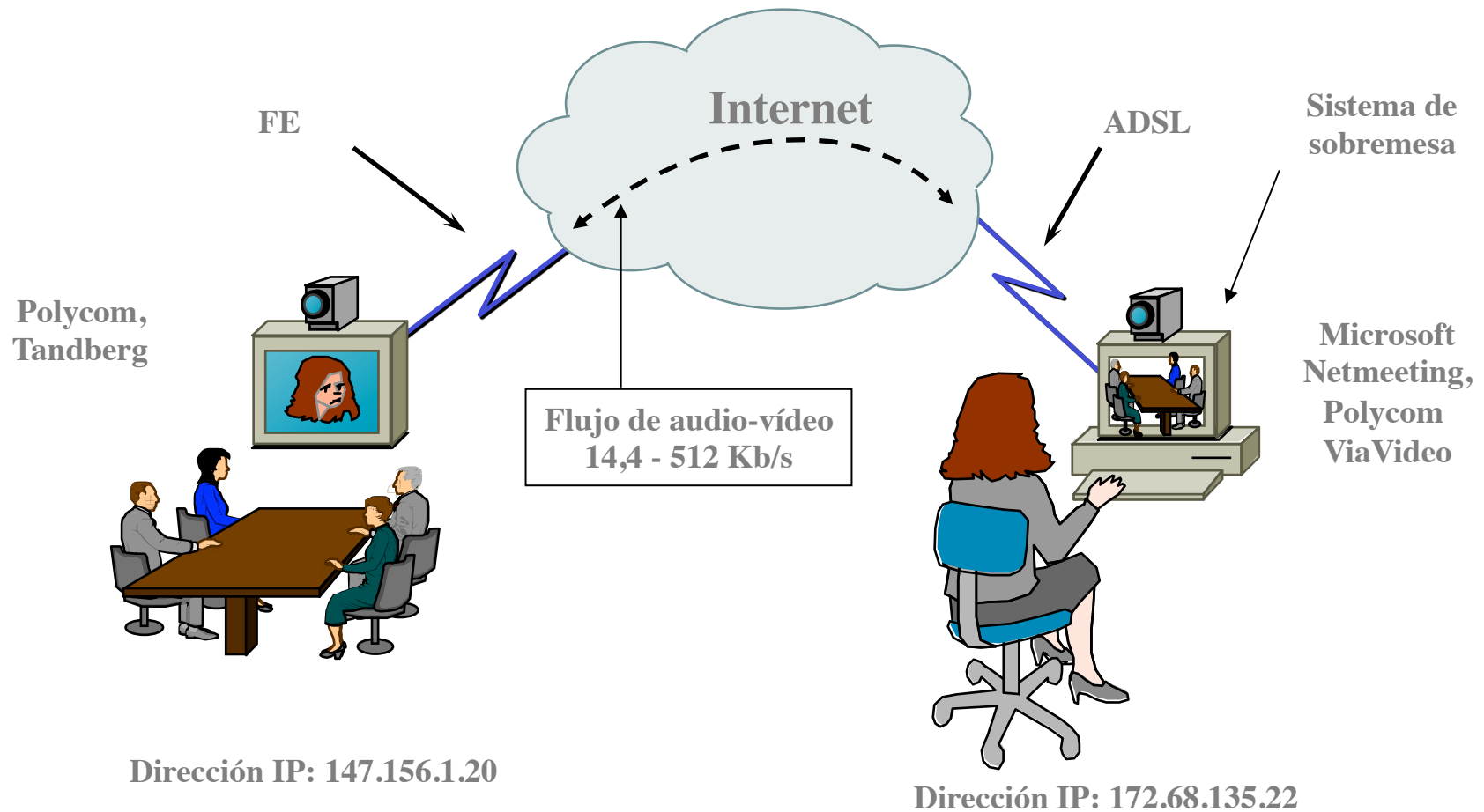
## 6.4.6 Videoconferencia H.320



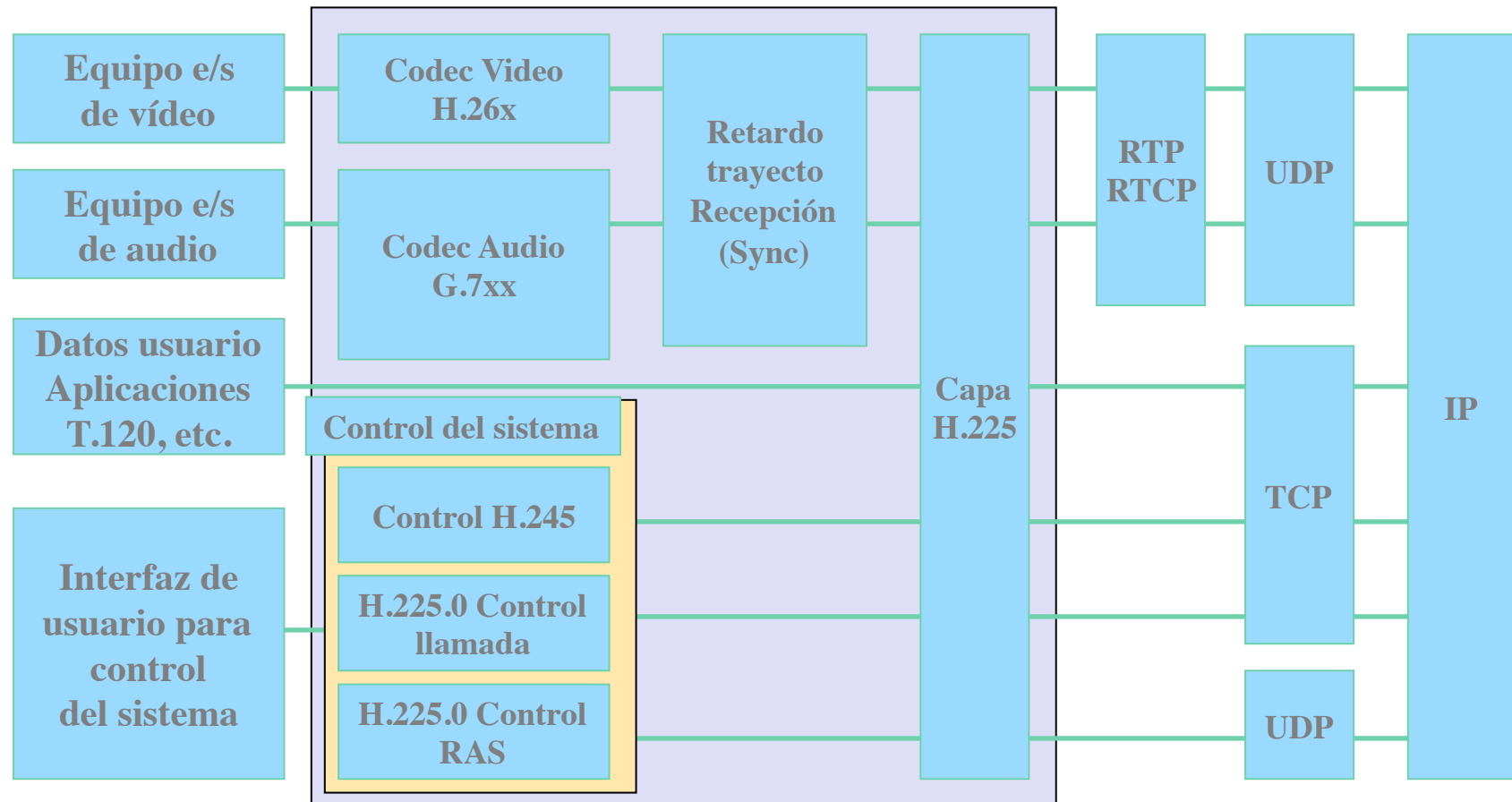
## 6.4.6 Direcciones E.164

- El formato de los números de teléfono se establece en el estándar E.164 de la ITU-T.
- Los números pueden tener un máximo de 15 dígitos decimales.
- Los primeros 1, 2 ó 3 dígitos representan el país, ej:
  - ❑ 1: Norteamérica (Estados Unidos y Canadá).
  - ❑ 34: España
  - ❑ 216: Túnez.
- La estructura de las direcciones dentro de cada país es decidida por el país. Normalmente es de tipo jerárquico con criterio geográfico.

## 6.4.6 Videoconferencia H.323



## 6.4.6 Arquitectura terminal H.323



## 6.4.6 Terminales de videoconferencia



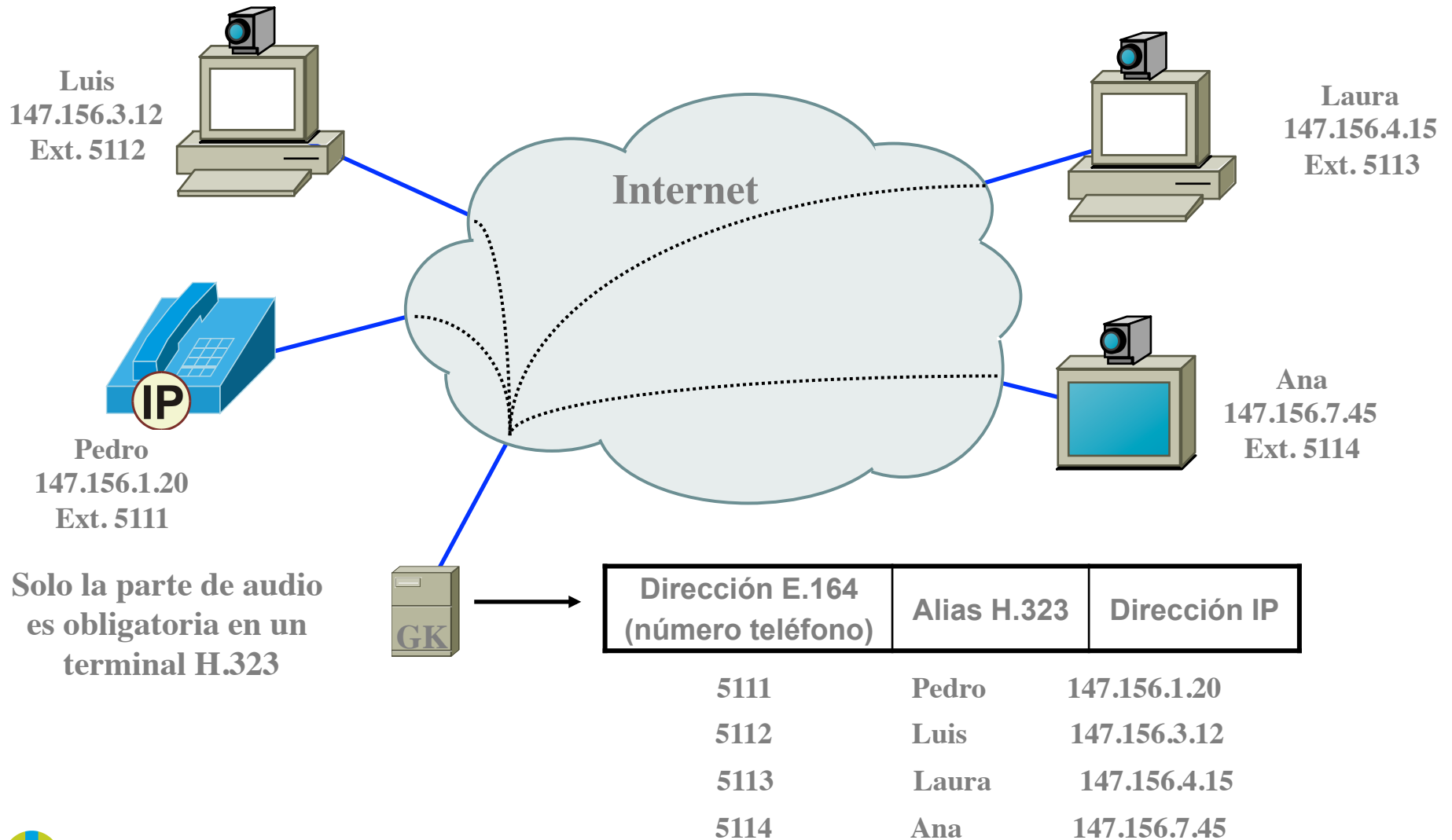
**Polycom ViaVideo**  
Video: H.261, H.263, H.263+  
Audio: G.711, G.722, G.728, G.723.1  
Caudal: 32-384 Kb/s (H.323)  
Formatos: CIF, QCIF  
Peso: 250 g  
Conexiones ent./sal.: USB, audio  
Precio: 500 euros



**Polycom ViewStation SP128**  
Video: H.261, H.263+  
Audio: G.711, G.722, G.728  
Caudal: 56-128 Kb/s (H.320), 56-768 Kb/s (H.323)  
Formatos: CIF, QCIF  
Peso: 2,7 Kg  
Conexiones ent./sal.: video v audio  
Precio: 5.000 euros

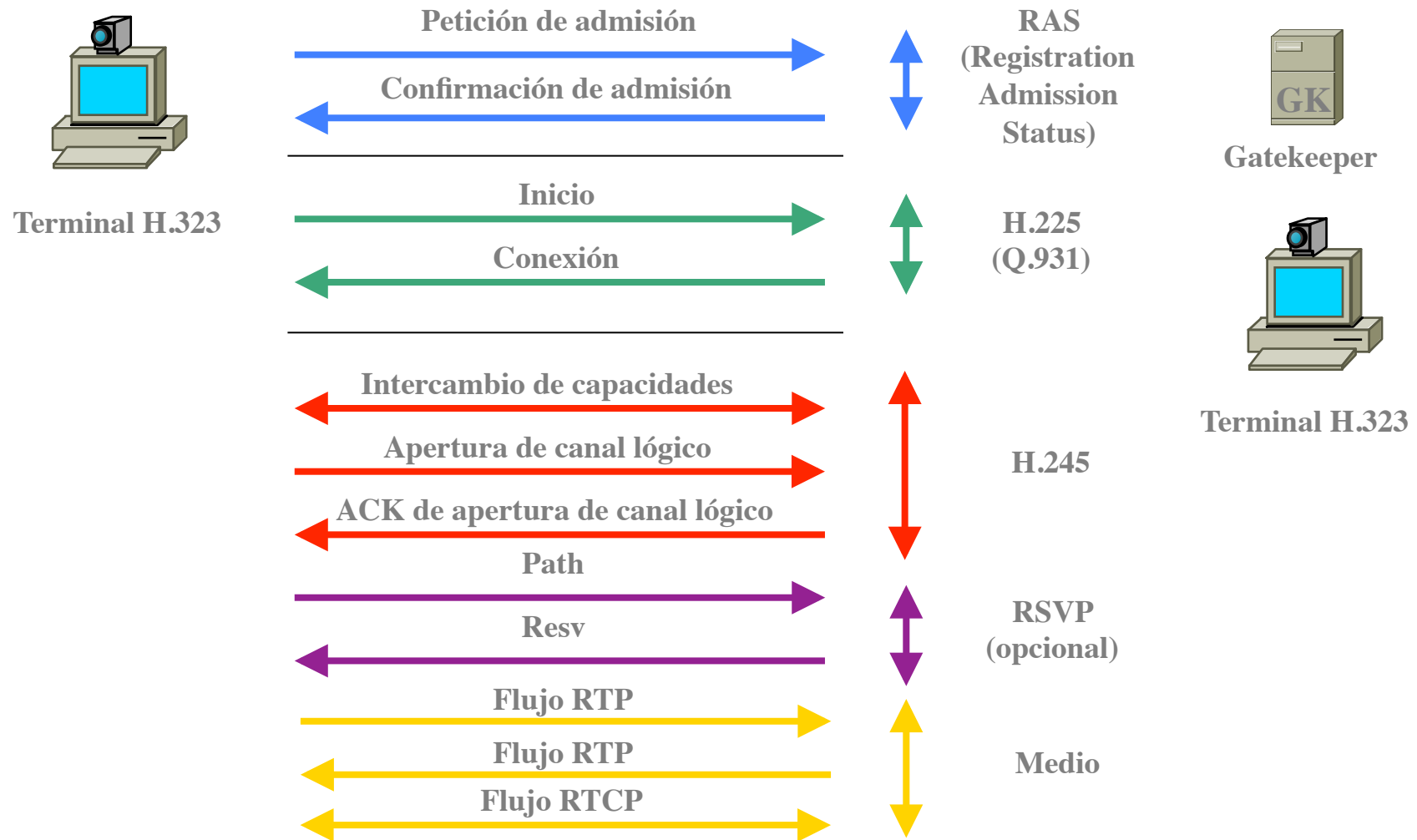


## 6.4.6 Videoconferencia H.323: Gatekeeper





## 6.4.6 Señalización H.323



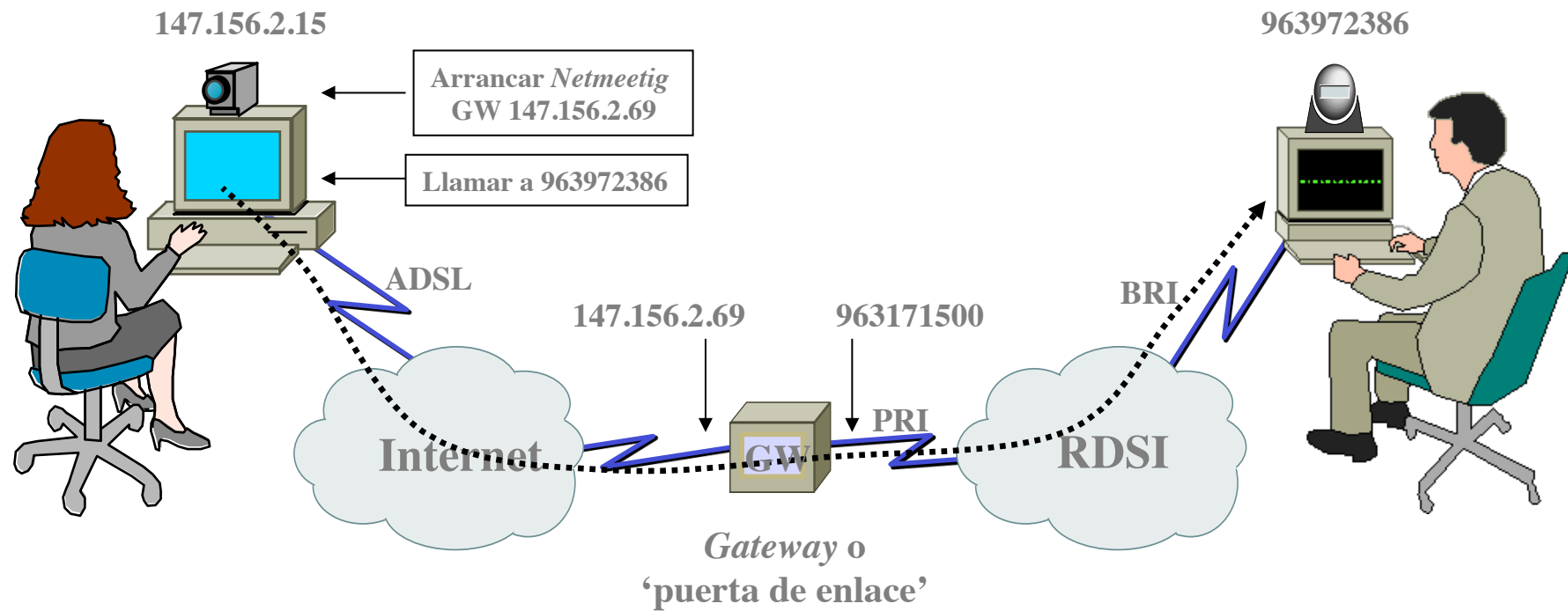
## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
- 7. Pasarelas e Interoperabilidad**
8. Telefonía Internet
9. Protocolo SIP

## 6.4.7 Elementos de videoconferencia

- **Terminal:** es el equipo que utiliza el usuario para comunicarse.
- **Gateway, pasarela o puerta de enlace:** interconecta redes diferentes: H.320 (RDSI) e Internet (H.323).
- **Gatekeeper o equipo selector:** permite el control de acceso. Realiza la equivalencia de direcciones E.164 o usuarios a direcciones IP.
- **MCU, Multipoint Control Unit o Unidad de control multipunto:** replica un flujo de audio/video para permitir multiconferencia.

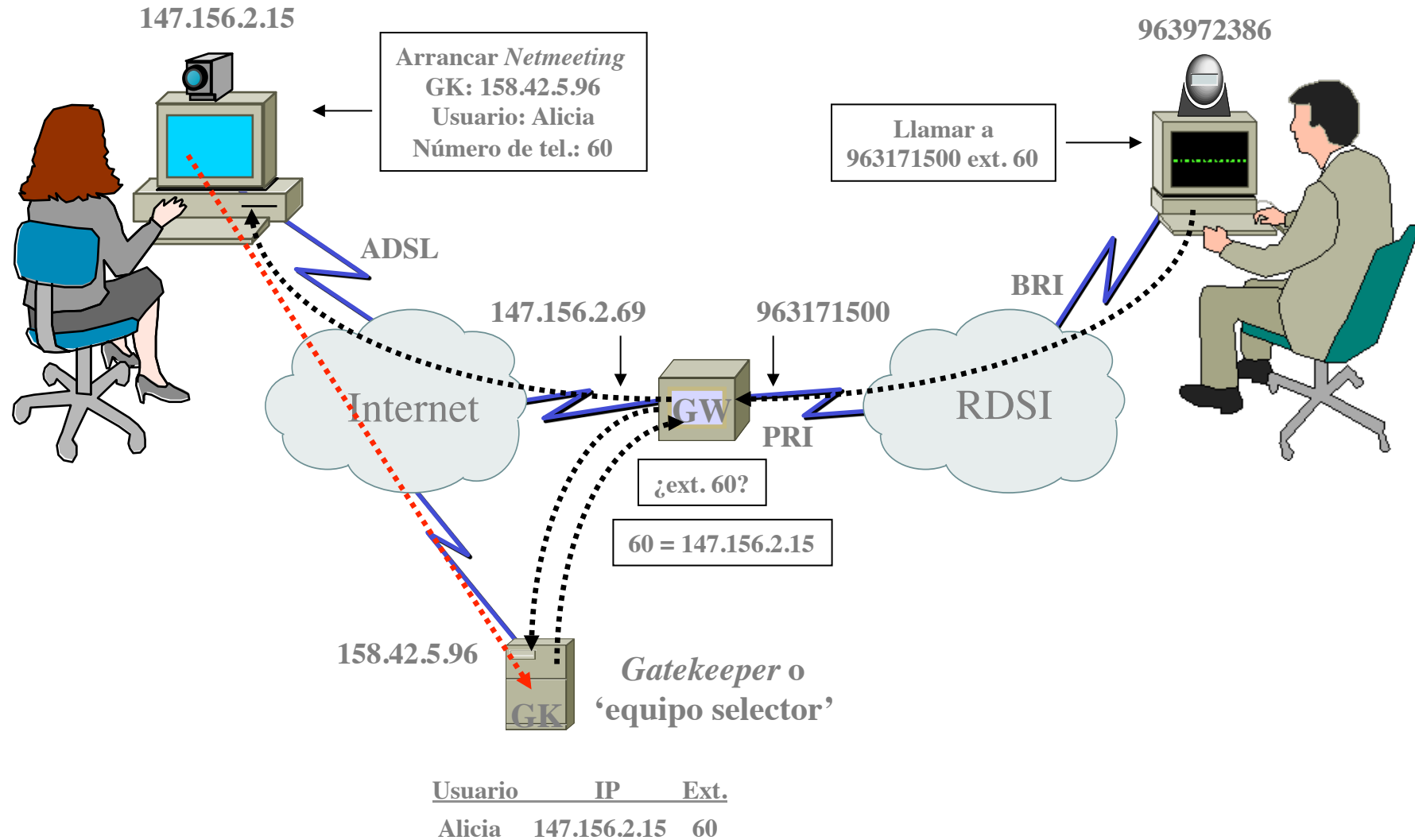
## 6.4.7 Pasarela (Gateway) H.320-H.323



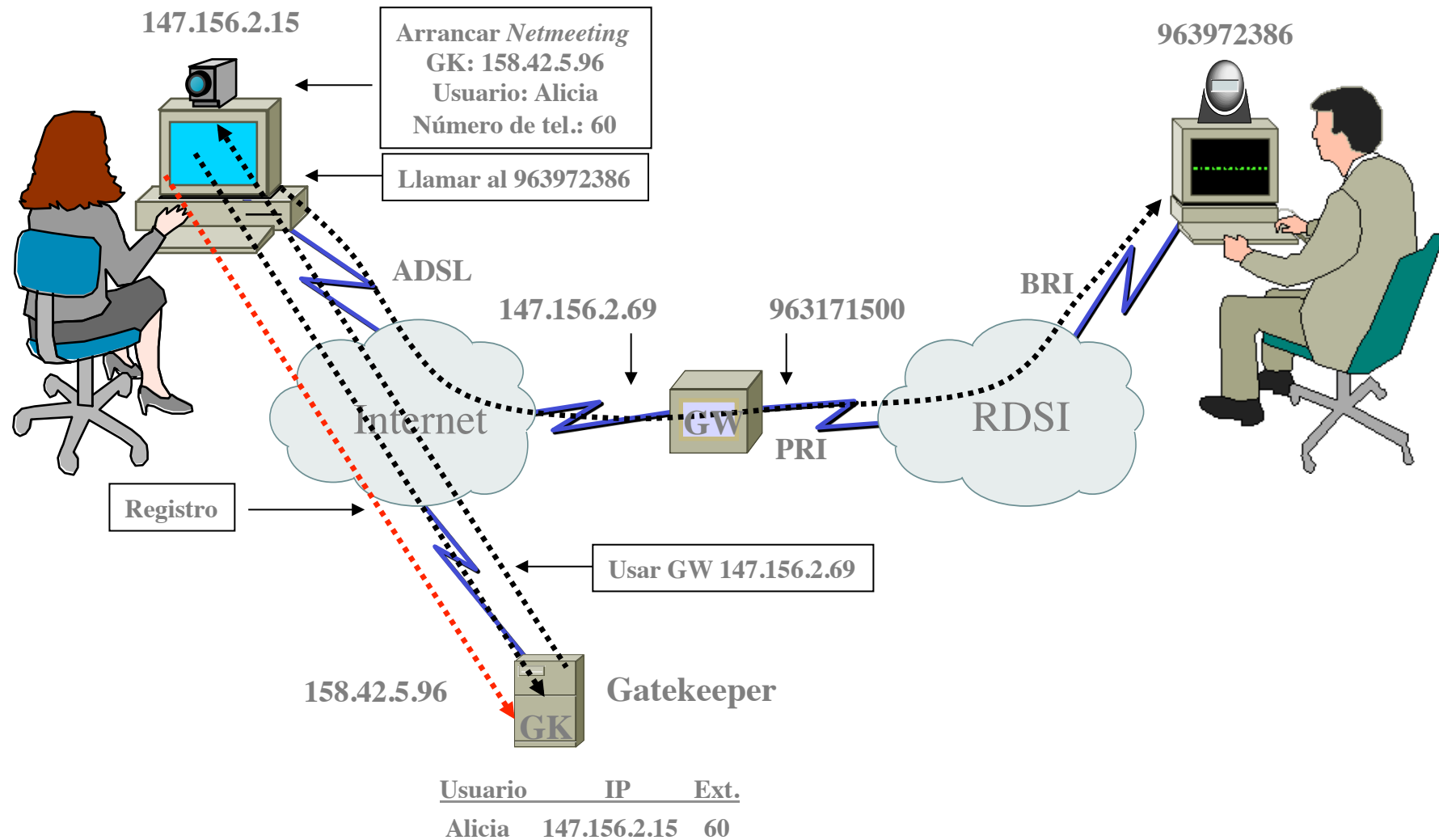
## 6.4.7 Funciones *Gateway* H.323

- Interoperabilidad entre audio/vídeo y estándares de red.
- Conversión de protocolo
  - ☐ Procedimientos de comunicación.
  - ☐ Formatos de transmisión.
- Opcionalmente: Transcodificación (conversión de formatos audio/video).

## 6.4.7 Gateway/Gatekeeper, llamada entrante



## 6.4.7 Gateway/Gatekeeper, llamada saliente



## 6.4.7 Funcionamiento del gatekeeper

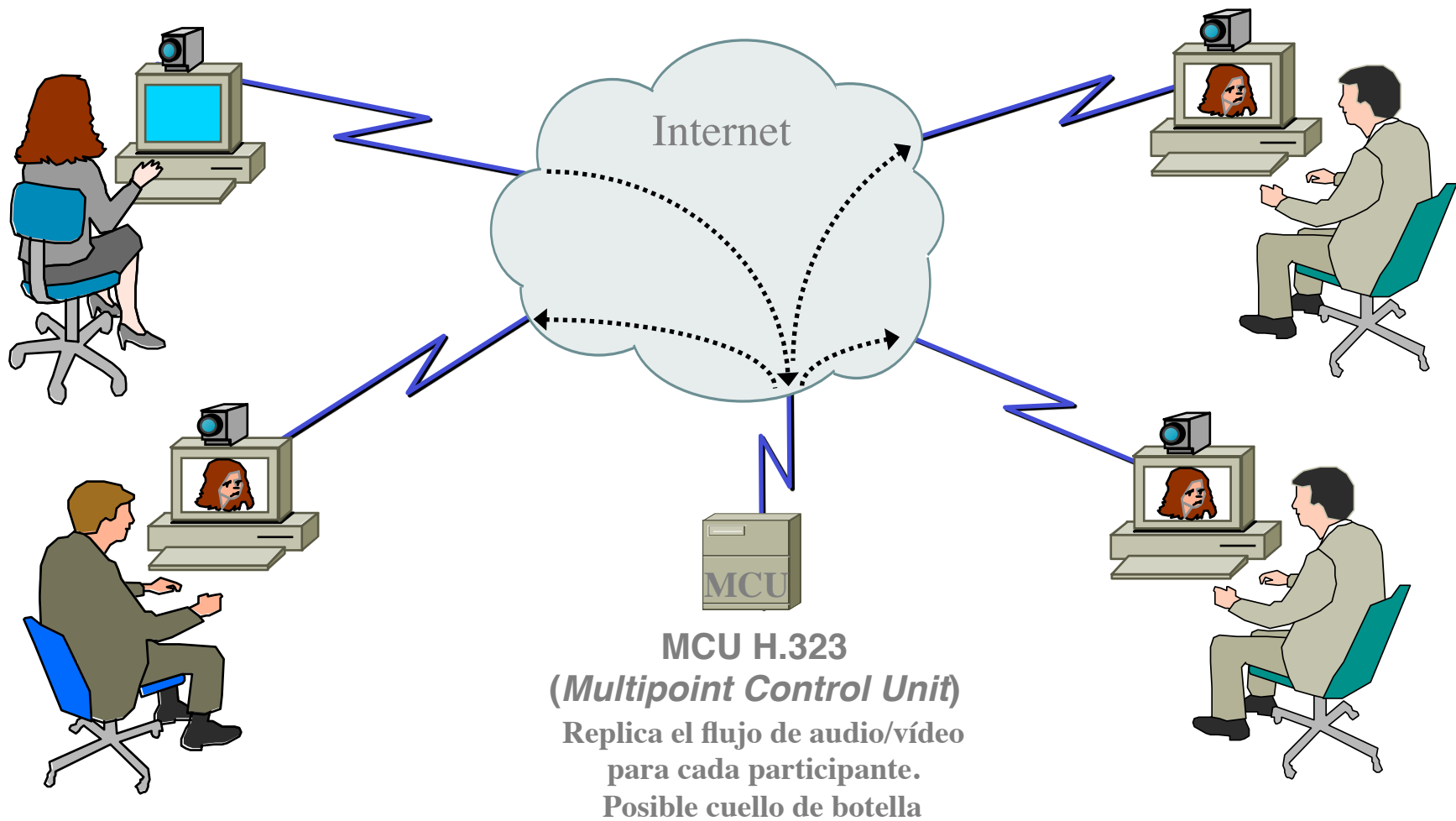
- El *Gatekeeper* puede validar el usuario/password en el momento del registro accediendo a un servidor RADIUS.
- El alias y la dirección E.164 permiten localizar a usuarios que utilicen diversos terminales o en redes que utilicen direcciones IP dinámicas.
- El *Gatekeeper* facilita el control de uso del servicio, permisos y autorizaciones, paso por cortafuegos, NATs, etc.



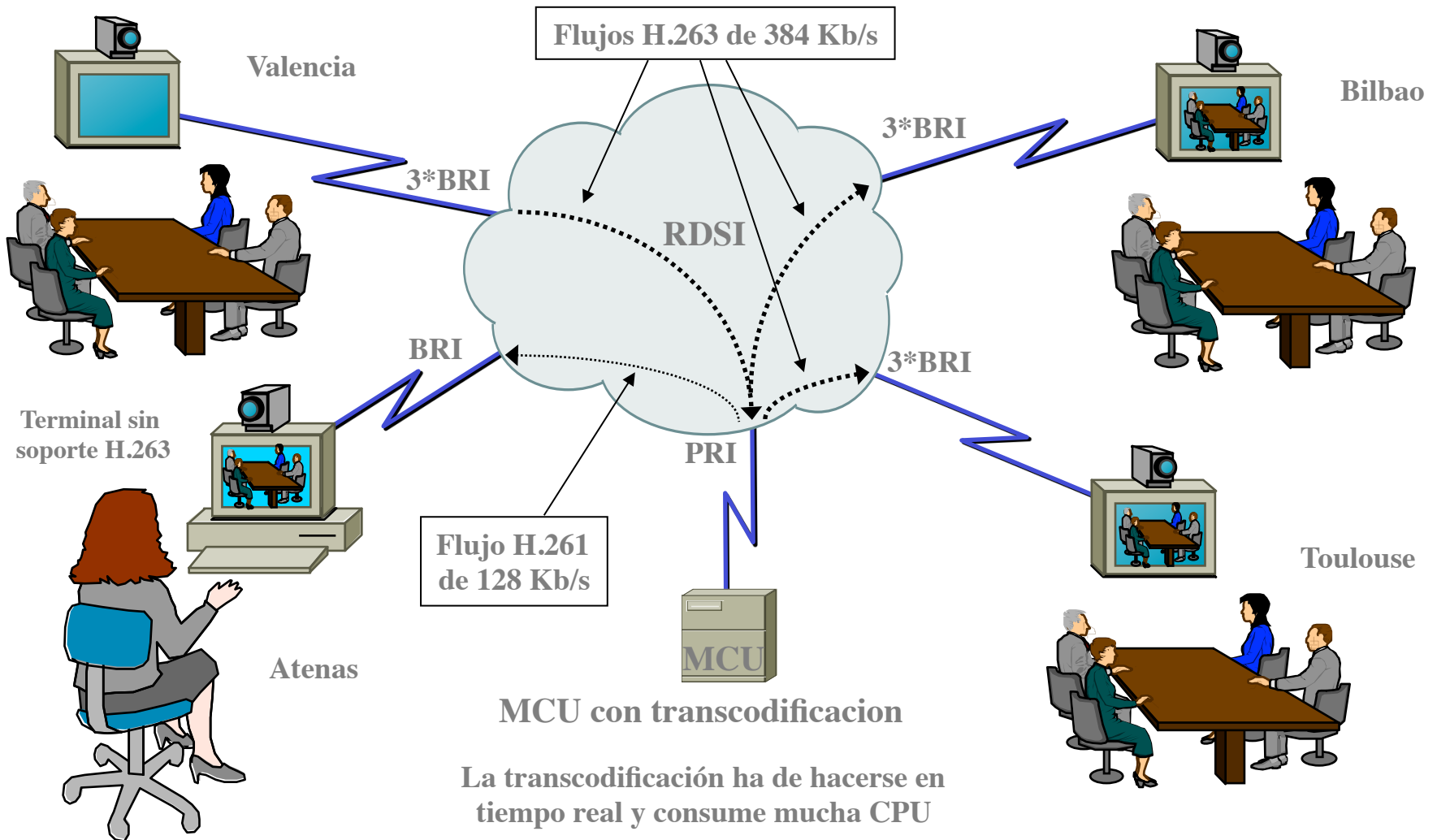
## 6.4.7 Procedimientos de llamada vía Gateway/ Gatekeeper

- Respuesta de voz interactiva (**IVR**, *Interactive Voice Response*):
  - ☐ Al marcar se escucha una locución que dice: ‘si sabe la extensión tecléela con un cero delante, si no espere y le atenderá la operadora’.
- Extensión por defecto:
  - ☐ Todas las llamadas se encaminan a una extensión determinada.
- Llamada directa del exterior (**DID**, *Direct Inward Dialing*):
  - ☐ Cada extensión recibe un número directo del exterior. Ej.: 96-386-3563 llama a la extensión 3563. Requiere obtener números extra del operador, lo cual tiene un costo.
- Enrutamiento **TCS4**:
  - ☐ La extensión se marca detrás del número, ej.: 96-386-3500#3563 llama a la extensión 3563. No disponible en España (los números extra se ignoran).

## 6.4.7 Videoconferencia multipunto H.323



## 6.4.7 Transcodificación



## 6.4 Índice

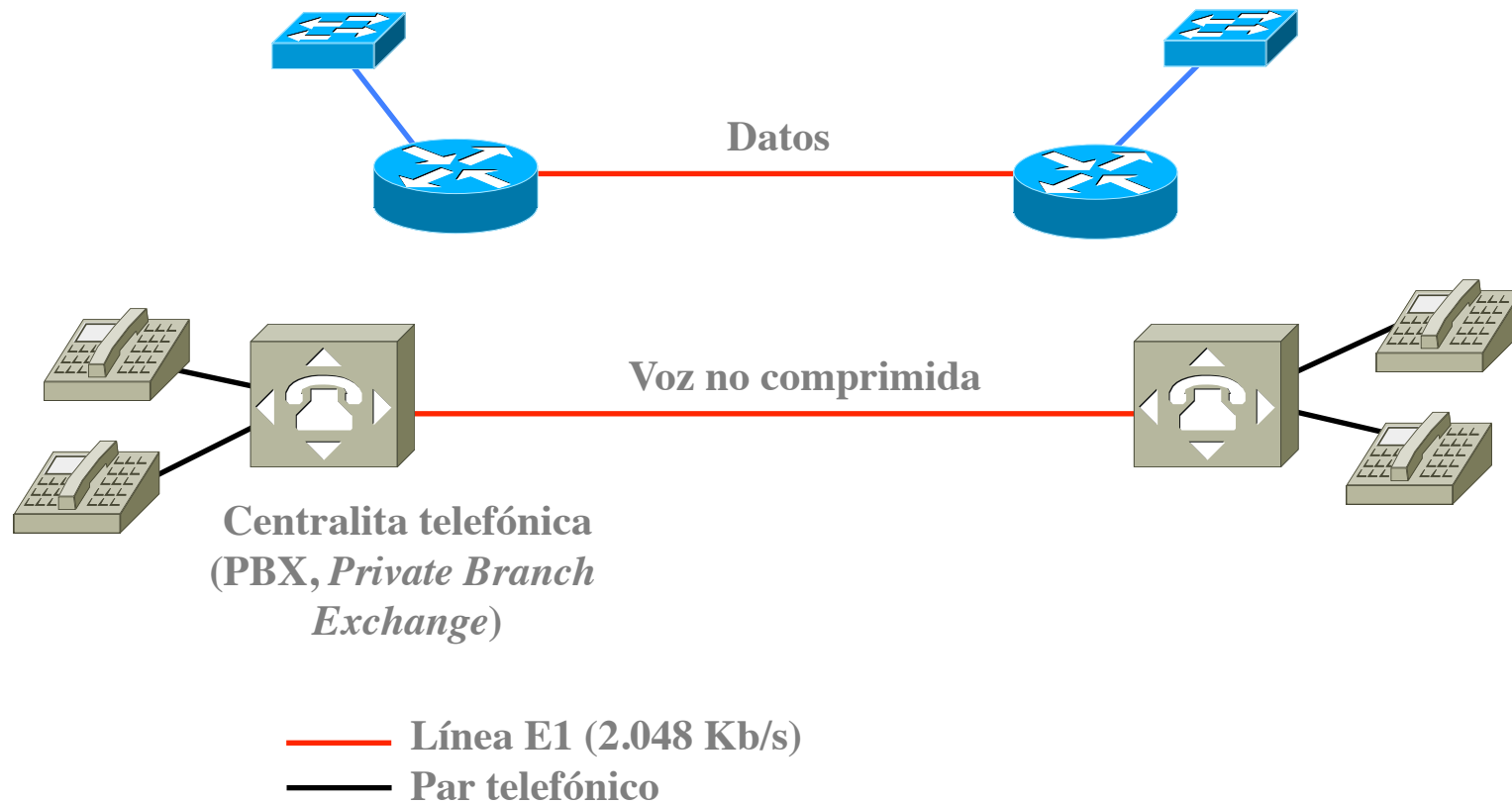
1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
- 8. Telefonía sobre Internet**
9. Protocolo SIP

## 6.4.8 Telefonía sobre Internet

- Pretende aprovechar la red IP para la comunicación telefónica.
- Requiere una red con bajo retardo y QoS, o bien una red sobredimensionada.
- Además de digitalizar la voz es necesario ofrecer todas las funciones propias de una red telefónica:
  - ☐ Señalización (llamada).
  - ☐ Funciones avanzadas: reenvío de llamadas, mensajería, etc.

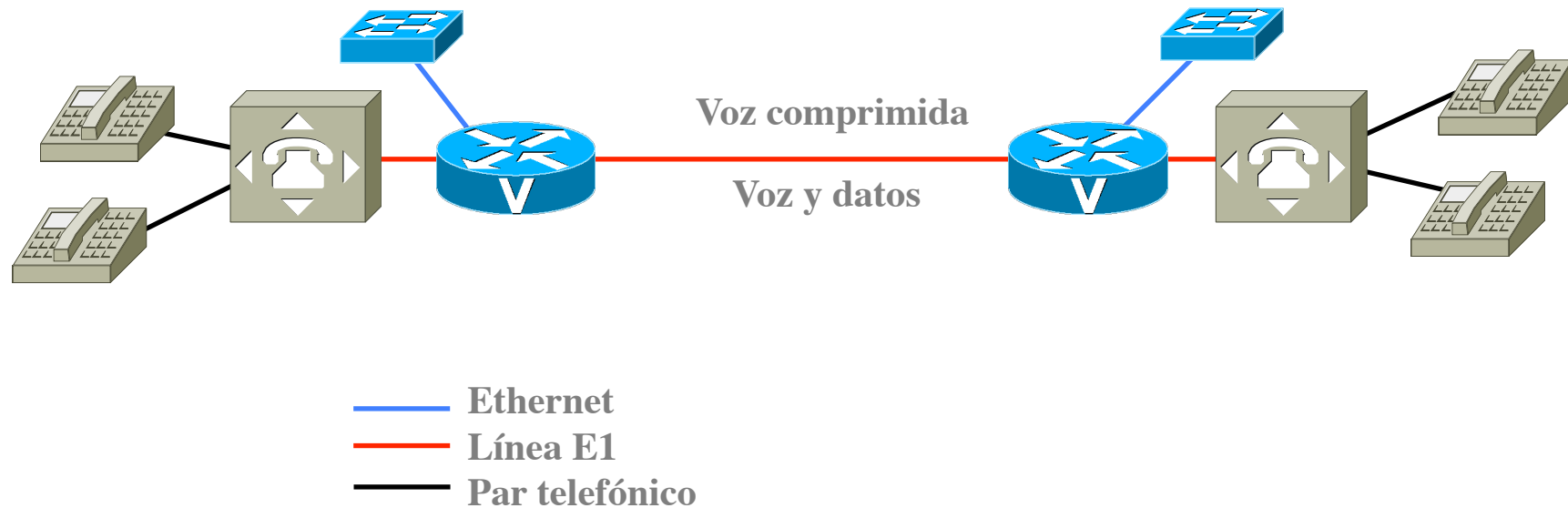
## 6.4.8 Evolución de la telefonía (I)

### Telefonía Tradicional



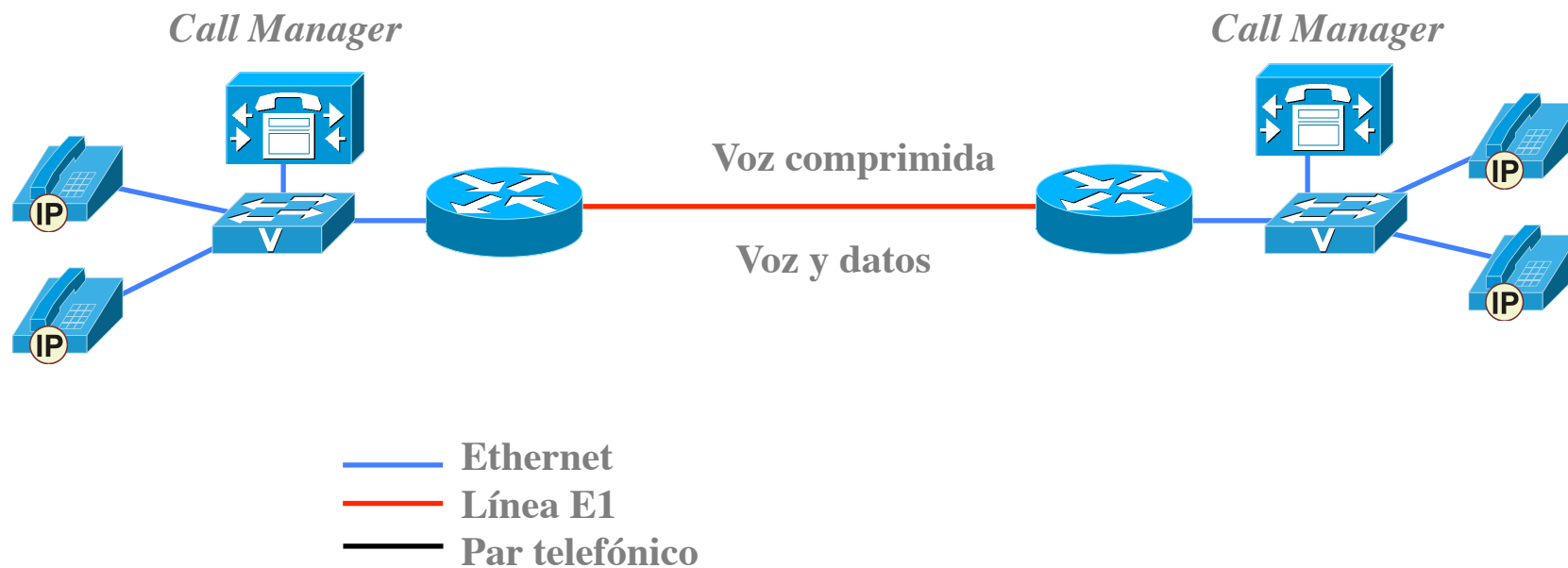
## 6.4.8 Evolución de la telefonía (II)

### Telefonía tradicional sobre backbone IP (voz sobre IP)



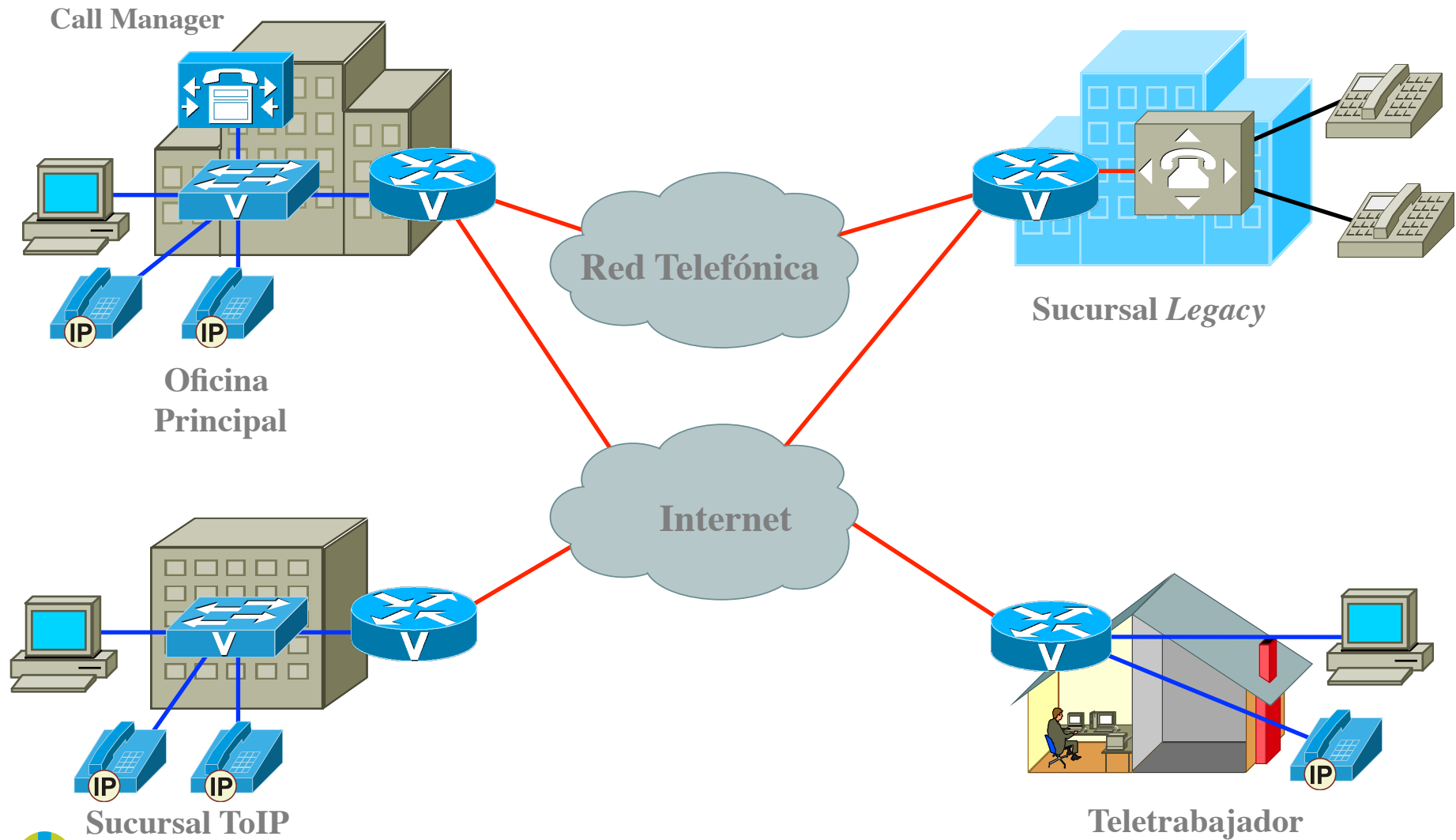
## 6.4.8 Evolución de la telefonía (III)

### Telefonía IP





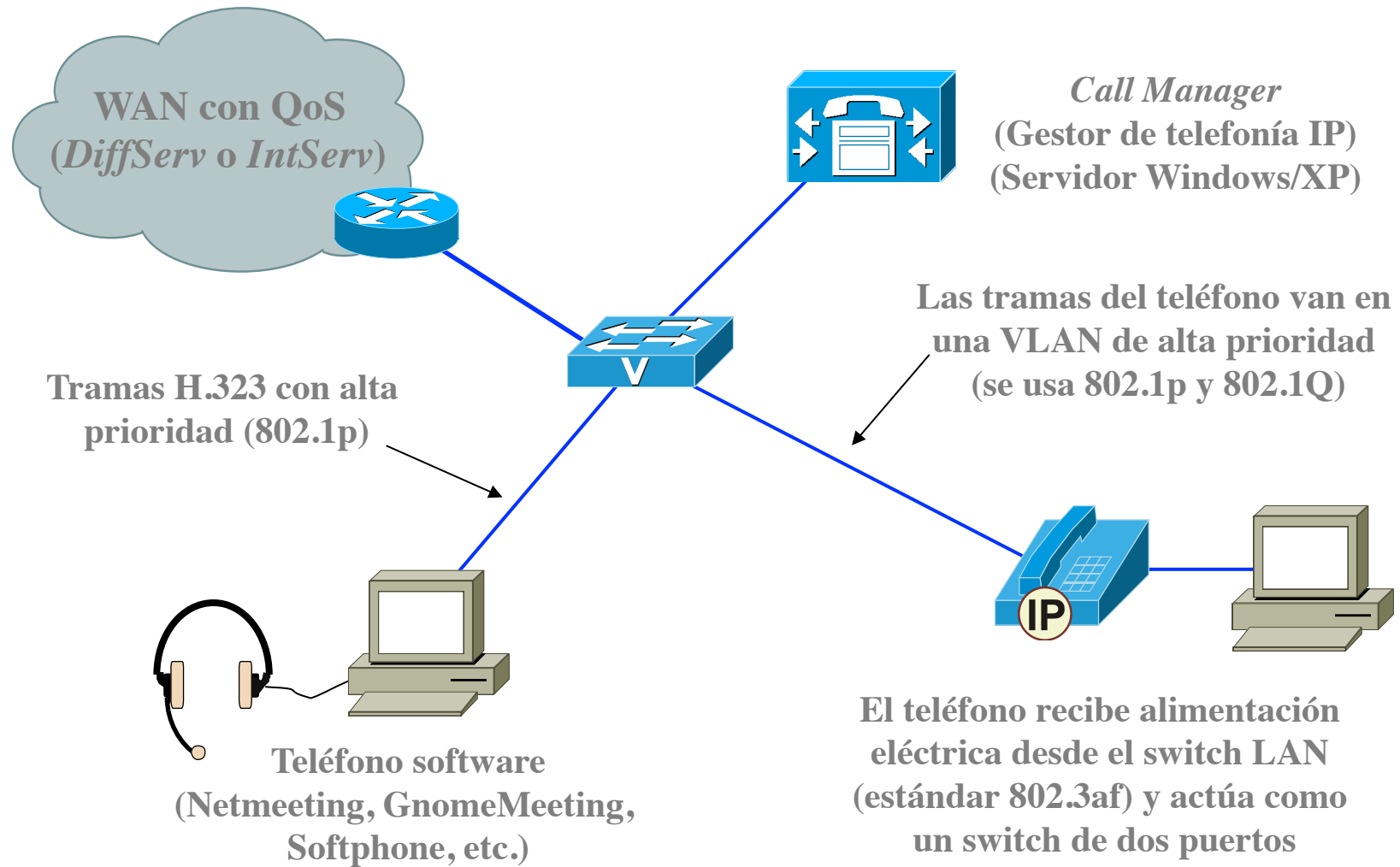
## 6.4.8 Ejemplo de red de telefonía IP



## 6.4.8 Telefonía IP

- Ventajas:
  - ☐ Integración de la red de datos y la red telefónica
  - ☐ Reducción de distancias (y costes) en la red telefónica
  - ☐ Fácil enrutamiento alternativo en caso de averías en la red (servicio no orientado a conexión)
  - ☐ Posibilidad de compresión de la voz (G.729, G.723.1)
  - ☐ Supresión de silencios y generación de ruido de confort
  - ☐ Servicios de alta calidad (G.722, G.722.1, G.719)
  - ☐ Posibilidad de integrar servicios: directorio telefónico con LDAP, envío de ficheros de audio por e-mail, lectura automática de e-mails por teléfono, navegación web asistida, etc.
- Inconvenientes
  - ☐ Degradación de la calidad cuando hay congestión
  - ☐ Mayores retardos, posibles problemas de ecos
  - ☐ Mayor costo de los teléfonos

## 6.4.8 LAN con telefonía IP



## 6.4.8 Teléfonos IP

Hard



**Cisco 7905**  
Audio G.711 y G.729a  
Precio: 200 euros



**Cisco 7960G**  
Audio G.711 y G.729a  
Incorpora conmutador de dos puertos 10/100  
Precio: 500 euros

Soft

**OpenPhone**  
<http://www.openh323.org/code.html>  
Precio: 0 euros



**Cisco SoftPhone**  
Audio G.711, G.723.1 y G.729a  
Precio: 150 euros



CEU

## 6.4.8 Tratamiento del sonido en videoconferencia y telefonía

- **Control automático de ganancia:** si la fuente sonora es más débil la ganancia se aumenta.
- **Supresión automática de ruidos:** un sonido constante se suprime (por ejemplo el ventilador de un proyector).
- **Supresión de silencios y ruido de confort:** el emisor deja de enviar cuando el sonido está por debajo de un umbral. El receptor genera entonces un ruido de fondo artificial llamado 'ruido de confort'.
- **Cancelación de eco:** en los sistemas manos libres es fundamental evitar realimentaciones altavoz-micrófono; muchas veces esto se consigue forzando una comunicación half duplex. Requieren cierto 'adiestramiento' (por ejemplo si se cambia la ubicación de los micrófonos).

## 6.4.8 Telefonía IP con H.323

- Un terminal H.323 solo está obligado a soportar audio, el vídeo es opcional.
- Por tanto con H.323, *gateways* y *gatekeepers* podemos ofrecer telefonía Internet sin necesidad de nuevos estándares.
- Sin embargo la mayoría de los fabricantes utilizan protocolos propietarios para dar soporte a las funciones adicionales (desvío de llamadas, mensajería, etc.).
- En la práctica es muy difícil desarrollar una red de telefonía IP multifabricante basada solo en estándares H.323.

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
- 9. Protocolo SIP**

## 6.4.9 SIP (Session Initiation Protocol)

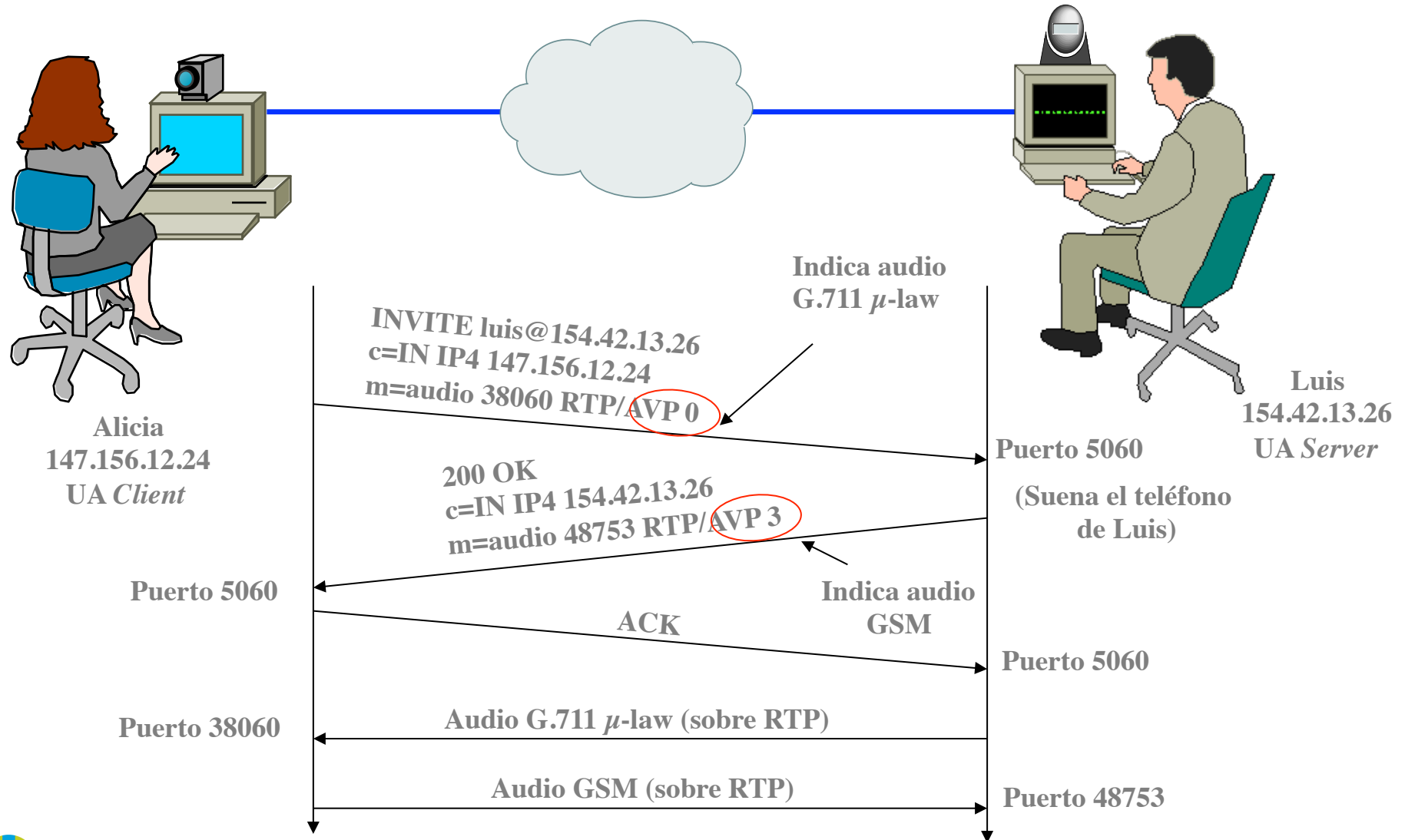
- En el mundo del IETF los estándares de la ITU-T siempre se han visto con recelo. En particular H.323 se considera un protocolo demasiado complejo.
- Esto dio lugar al desarrollo de un protocolo alternativo llamado SIP (*Session Initiation Protocol*, RFC2543, 3/99, 153 pág.) cuyo diseño está inspirado en HTTP y SMTP.
- En SIP las direcciones son URIs. Por ej.: sip:rector@ceu.es.
- Página principal del SIP: Universidad de Columbia: <http://www.cs.columbia.edu/sip>
- Con el tiempo SIP se ha complicado más de lo previsto. Pero aun así parece que terminará ganando la batalla a H.323.



## 6.4.9 Componentes de SIP

- UA (*User Agent*): Terminal SIP (*hard* o *soft*). Un UA puede actuar como:
  - ❑ UAC (*User Agent Client*): el que llama.
  - ❑ UAS (*User Agent Server*): el que es llamado.
- Servidores:
  - ❑ **Proxy**: actúa como intermediario, en representación de un UA para efectuar o atender llamadas. Ayuda a gestionar el paso por cortafuegos, establecer restricciones, etc.
  - ❑ **Registrar**: atiende peticiones de registro de los agentes activos en su dominio (bien suyos o visitantes)
  - ❑ **Redirect**: informa a los clientes de los agentes que se encuentran fuera de su ubicación natural

## 6.4.9 Llamada SIP directa entre dos UAs



## 6.4.9 Algunas características de SIP

- Los mensajes de control se envían '**fuera de banda**', es decir forman un flujo independiente (sockets distintos) de los datos intercambiados (en este caso la voz digitalizada).
- Los mensajes de control se envían a un puerto 'bien conocido' (el 5060). El puerto de los datos en cambio lo elige cada agente independientemente en el momento de iniciar la sesión.
- Los mensajes están en caracteres ASCII y tienen una sintaxis parecida a los del protocolo HTTP.
- Las direcciones IP de los agentes implicados y los puertos elegidos figuran en los mensajes SIP. Cuando hay un NAT en el camino éste ha de realizar las conversiones pertinentes.
- Todos los mensajes requieren **confirmación**, por lo que pueden enviarse sobre TCP o sobre UDP.
- En el proceso de conexión se realiza una negociación de capacidades. En el ejemplo si Luis no soportara el audio G.711  $\mu$ -law le enviaría a Alicia un mensaje '*600 Not Acceptable*' indicándole además los codecs que él soporta.

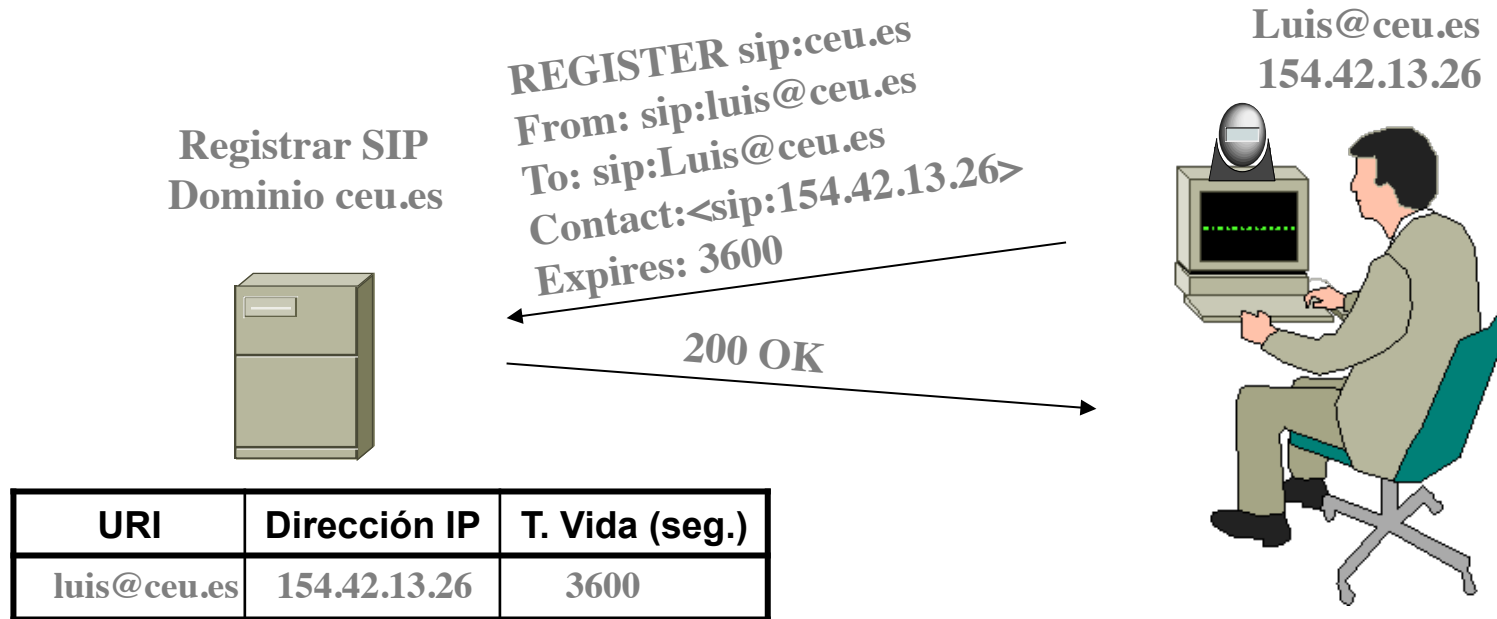
## 6.4.9 Direcciones SIP

- En el ejemplo anterior Luis tenía una IP fija.
- Esto no siempre es posible, bien porque el usuario obtenga su dirección de forma dinámica (por DHCP) o porque utilice diversos dispositivos, cada uno con una IP diferente (portátil, PDA, etc.).
- En SIP se prevé que las direcciones puedan ser:
  - ☐ IPs fijas: sip:[luis@152.44.13.26](mailto:luis@152.44.13.26)
  - ☐ URIs: sip:luis@ceu.es
  - ☐ Direcciones E.164: [43865@ceu.es](mailto:43865@ceu.es)
- Para que los URIs y las direcciones E.164 puedan funcionar hace falta que la red SIP tenga un *Registrar* y un *Proxy*.

## 6.4.9 Registrar SIP (Registrador)

- Cada vez que un usuario arranca una aplicación SIP envía un mensaje REGISTER al *Registrar* competente en ese dominio, que toma nota de su presencia.
- Los mensajes REGISTER tiene un tiempo de vida (por ejemplo una hora).
- Cada dominio tiene un *Registrar* que desarrolla para las direcciones SIP una función equivalente a la del DNS con los nombres.
- Los *Proxys* SIP lanzan consultas a los *Registrar*.

## 6.4.9 Proceso de registro SIP



El registro debe renovarse periódicamente (en este ejemplo cada hora)

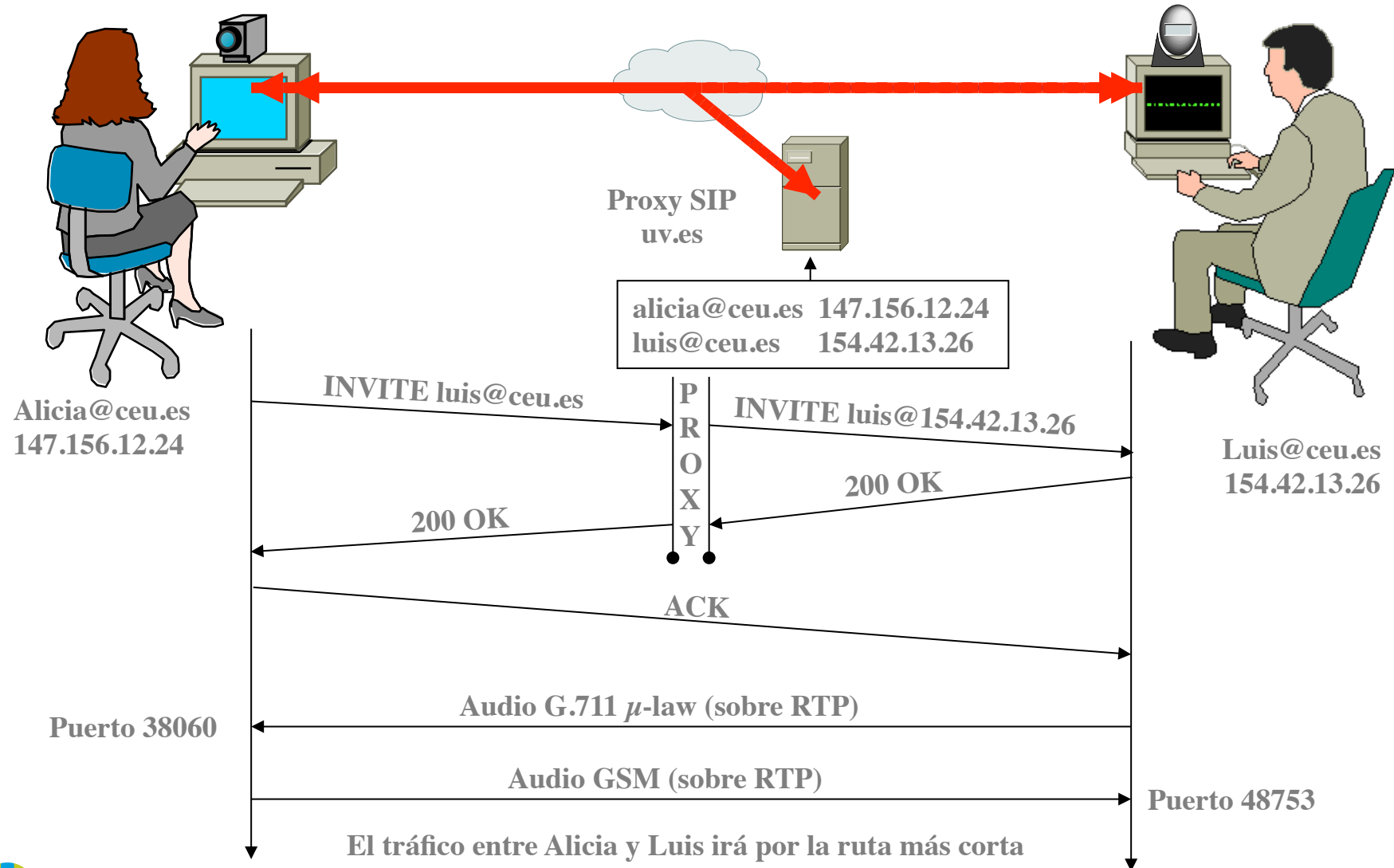
Normalmente el registro irá precedido de una validación por usuario/ password ante un servidor RADIUS (no mostrado en la figura)

El usuario puede registrarse desde cualquier dirección IP, aunque ésta no pertenezca al dominio ceu.es

## 6.4.9 Funciones del Proxy SIP

- El *Proxy* SIP es un *router* que actúa como intermediario en la comunicación entre agentes SIP.
- Su actuación se limita al proceso de conexión (señalización). Una vez establecida esta los agentes intercambian directamente los paquetes de datos.
- El *Proxy* consulta al registrar para localizar a un usuario dado. Generalmente el *Proxy* y el *Registrar* son el mismo equipo (normalmente un *router*)
- El *Proxy* permite forzar la autenticación de los usuarios y restringir el uso de los servicios de acuerdo con lo que tiene autorizado cada uno.

## 6.4.9 Llamada SIP mediante Proxy/Registrar





## 6.4.9 Descubrimiento del Proxy

- En el ejemplo anterior Alicia tenía que saber quién era su Proxy.
- El proxy se puede indicar en la configuración del agente, pero eso es poco flexible.
- En su lugar se utilizan registros tipo SRV (*Service*) en el DNS. Así se puede cambiar el Proxy sin tener que tocar la configuración de los agentes.

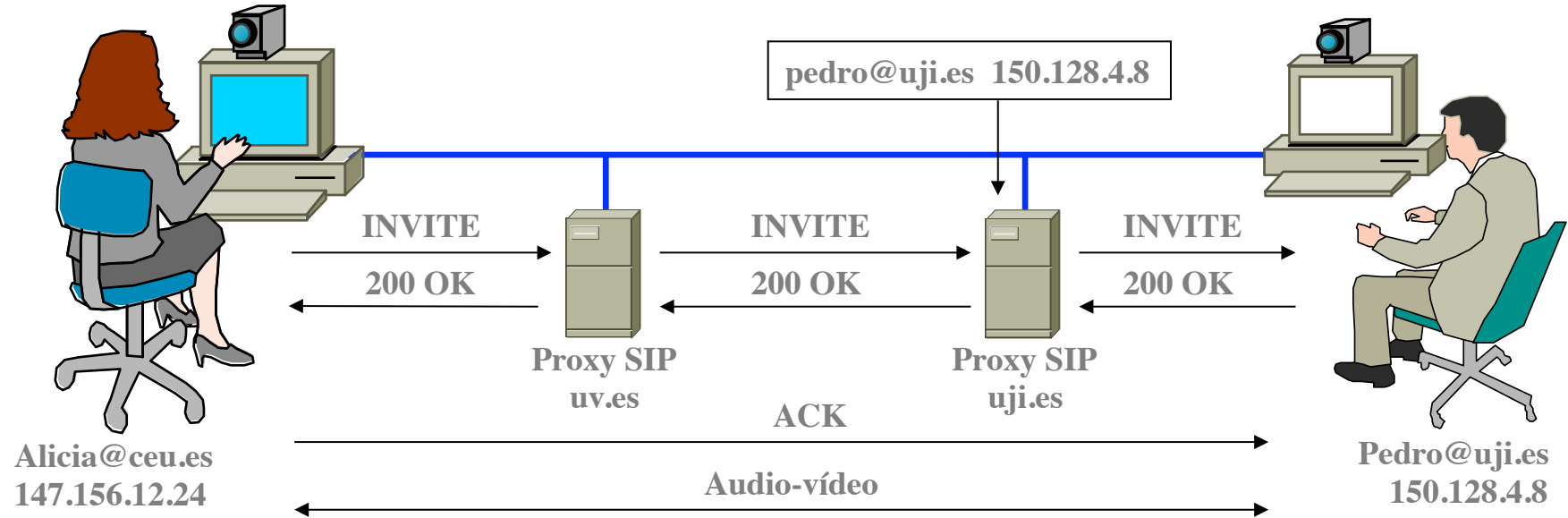
Definición en el DNS del proxy SIP para uv.es:

```
uv.es.      IN NAPTR 2 0 "s" "SIP+D2U" "" _sip._udp.uv.es.  
_sip._udp  IN SRV 0 0 5060 admundsen.red.uv.es.  
_sip._tcp  IN SRV 0 4 5060 admundsen.red.uv.es.
```

↑  
protocolo

↑  
Puerto (SIP)

## 6.4.9 Llamada SIP a un usuario remoto



- Cuando Alicia llama a Pedro sigue el mismo procedimiento que antes, salvo que ahora usa dos proxys, el de **uv.es** y el de **uji.es**.
- El proxy de uv.es averigua la dirección del Proxy de uji.es haciendo una consulta al DNS:

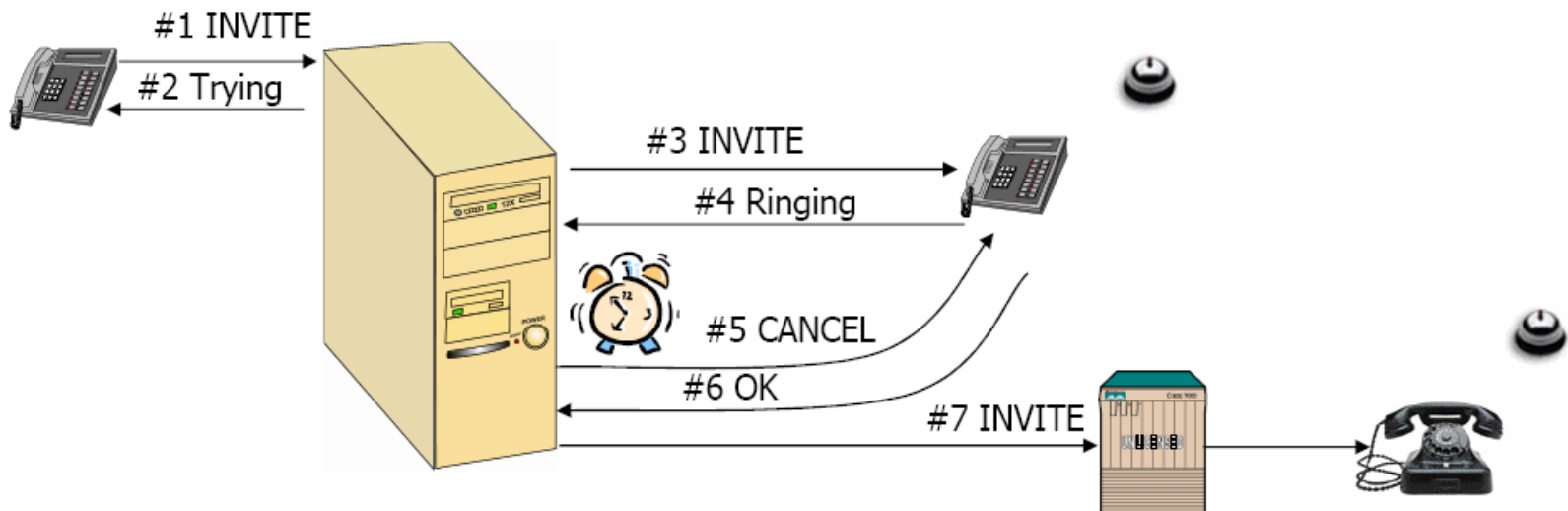
**DNS Query:** SRV uji.es UDP 5060?

**DNS Response:** 150.128.4.8

- Una vez establecida la sesión el tráfico entre Alicia y Pedro discurre por la ruta más corta entre ambos

## 6.4.9 Llamadas a múltiples agentes

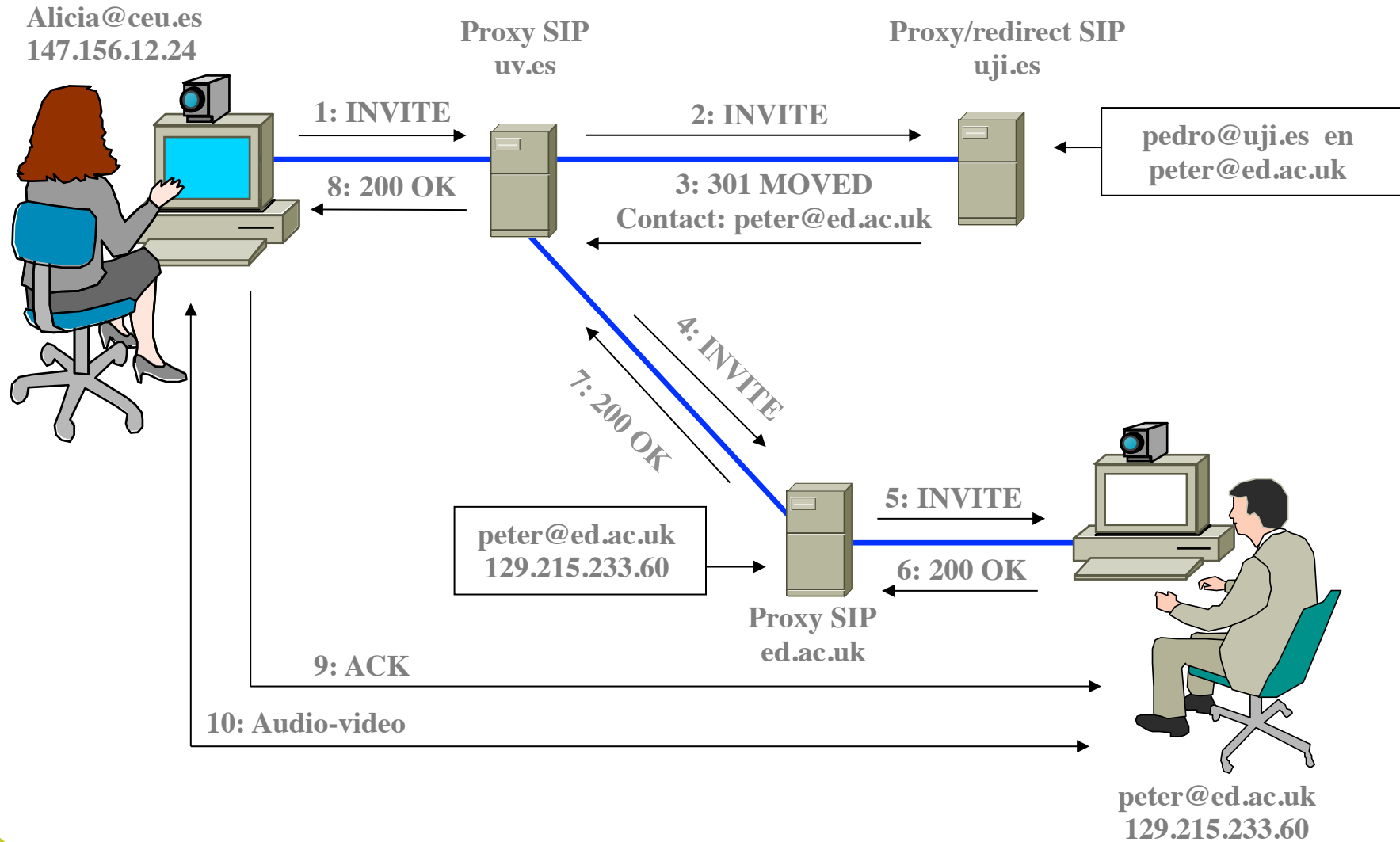
- Un mismo URI (p. ej. luis@ceu.es) se puede registrar desde diferentes agentes simultáneamente.
- En ese caso el proxy puede redirigir la llamada a todos ellos en paralelo o en serie, siguiendo una secuencia determinada previamente.



## 6.4.9 Redirección de llamadas SIP

- Un URI puede redirigirse (desviarse) temporalmente hacia otro.
- Supongamos que Pedro se va de Erasmus a la Universidad de Edimburgo; allí le asignan un nuevo identificador, `peter@ed.ac.uk`.
- Pedro recibirá las llamadas dirigidas a su nuevo URI, pero además quiere seguir recibiendo en su teléfono de Edimburgo las que le hagan a su URI permanente, `pedro@uji.es`.
- Para ello Pedro debe indicar al '*redirect server*' de `uji.es` que se ha trasladado temporalmente a [peter@ed.ac.uk](mailto:peter@ed.ac.uk).
- Normalmente el *proxy*, el *registrar* y el *redirect server* de una organización son el mismo equipo (un *router*).

## 6.4.9 Redirección de llamadas SIP



## 6.4.9 Comunicación SIP ↔ Red telefónica

- Comunicación física: se requiere un *gateway* o pasarela (normalmente un *router*) entre la red IP y la red telefónica pública.
- Comunicación lógica (direccionamiento):
  - ❑ SIP → Red telefónica: La posibilidad de utilizar direcciones E.164 desde SIP está prevista por defecto.
  - ❑ Red telefónica → SIP: hay que asignar direcciones E.164 a los agentes SIP. Esto se consigue con ENUM (E.164 NUmber Mapping), conjunto de protocolos que integran el espacio de direcciones E.164 en el DNS. ENUM permite además asignar URIs a las direcciones E.164.

## 6.4.9 Como funciona ENUM

- A cada dirección E.164 le corresponde un nombre en el DNS, de la siguiente forma (RFC 3761):

Dirección E.164: 3 4 9 6 3 5 4 3 8 6 5

Nombre DNS: 5.6.8.3.4.5.3.6.9.4.3.e164.arpa

- Los registros introducidos en el DNS son de tipo NAPTR (Name Authority Pointer) y le asignan uno o varios URIs equivalentes con diferentes prioridades. Ejemplo:

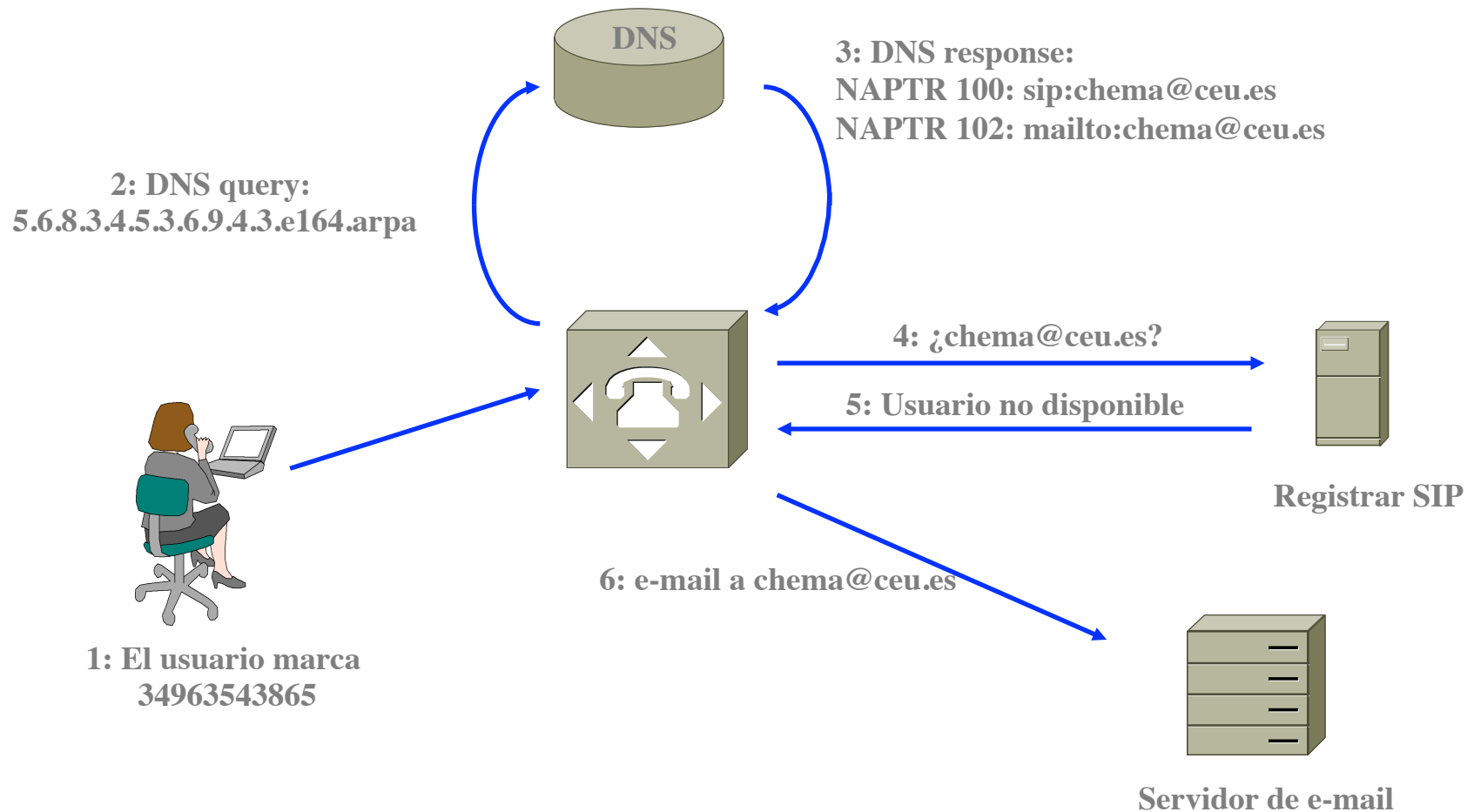
```
$ORIGIN 5.6.8.3.4.5.3.6.9.4.3.e164.arpa.
```

```
IN NAPTR 100 10 "u" "E2U+sip" "!^.*$!sip:chema@ceu.es!i"
```

```
IN NAPTR 102 10 "u" "E2U+email" "!^.*$!mailto:chema@ceu.es!i"
```

- Una vez configurado el 'mapping' las direcciones E.164 pueden ser utilizadas no solo desde la red telefónica convencional sino también desde terminales SIP

## 6.4.9 Ejemplo de uso de ENUM

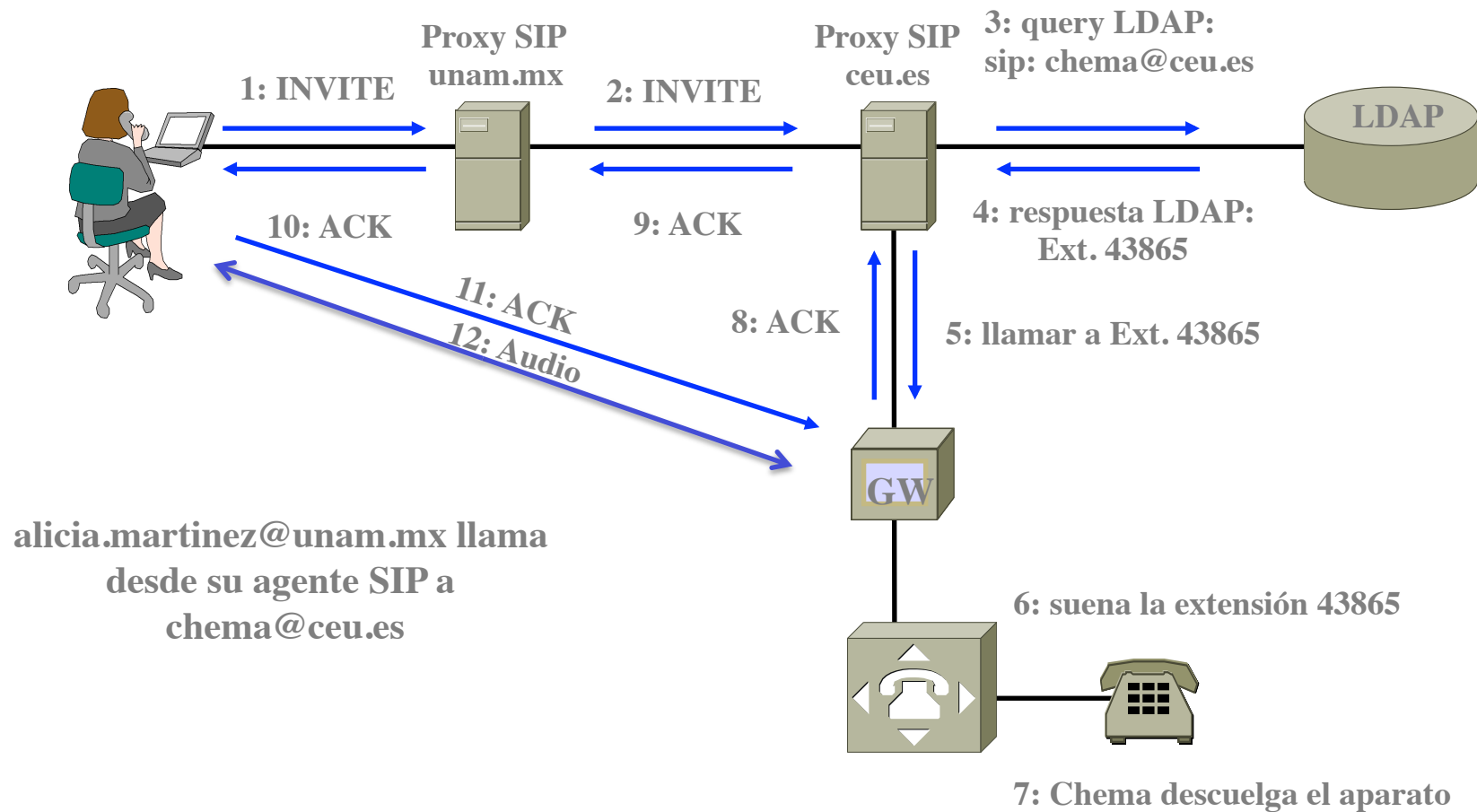




## 6.4.9 Aplicación de SIP: sip.edu

- El proyecto SIP.edu de Internet2 consiste en permitir el acceso por Internet a las redes telefónicas privadas de las universidades.
- La llamada se enruta por Internet y una vez en la universidad de destino se pasa a través de una pasarela a la red telefónica (privada) de la universidad, con lo que la llamada no tiene costo.
- El llamante utiliza un agente SIP y llama al URI del destinatario. El *Proxy* del destinatario traduce dicho URI en un número de extensión (dirección E.164) consultando una base de datos (directorio LDAP por ejemplo). Entonces la llamada es enrutada a través de la pasarela de la red privada al teléfono de la persona de destino. No se accede al DNS, no se usa ENUM.

## 6.4.9 Funcionamiento de SIP.edu



## 6.4 Bibliografía

[1] Tanenbaum, A. S., Computer Networks, 4ª Ed  
Pearson 2003, apartado 7.4.

[2] J.F Kurose y K.W. Ross, Redes de Computadores:  
un enfoque descendente basado en Internet, 2002.

## 6.4 Bibliografía

[3] Sobre RTP:

<http://www.ietf.org/html.charters/avt-charter.html>

[4] Sobre H.323: <http://www.h323plus.org/>

[5] Sobre SIP: <http://www.ipstel.org/SIPResources>

[6] Tutorial de SIP:

[http://www.ipstel.org/files/sip\\_tutorial.pdf](http://www.ipstel.org/files/sip_tutorial.pdf)

[7] Sobre SIP.edu: <http://www.internet2.edu/sip.edu/>

[8] Sobre H.265: <http://en.wikipedia.org/>



## 6 Capa de Aplicación

### 6.4 Multimedia (adicional)

RdE 2013-2014

## 6.4 Índice

- 1. Fundamentos de telefonía. Conversión analógico-digital**
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
- 2. Audio digital. Estándares. Compresión**
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía sobre Internet
9. Protocolo SIP

## 6.4.2 Audio MPEG-1

- Compresión psicoacústica con pérdidas (aprovecha las características del oído humano).
- Frecuencias de muestreo entre 8 y 48 KHz. Si se va a utilizar un caudal reducido es conveniente hacer el muestreo a baja frecuencia.
- Tres capas en orden ascendente de complejidad/calidad. Cada capa incorpora nuevos algoritmos, y engloba los de las anteriores :
  - ☐ Capa I: buena calidad con 192-256 Kbps por canal; no se utiliza.
  - ☐ Capa II: calidad CD con 96-128 Kbps por canal.
  - ☐ Capa III: calidad CD con 64 Kbps por canal.
- El más utilizado actualmente es Capa III por su mayor eficiencia. También se le conoce como MP3. El caudal puede estar entre 8 y 320 Kb/s.



## 6.4.2 Audio MPEG-2 AAC *Advanced Audio Coding*

- Algoritmo de compresión de audio de alta eficiencia y alta calidad incorporado en los estándares MPEG-2 parte 7 y MPEG-4 parte 3.
- Frecuencias de muestreo de 8 a 192 KHz. Caudales entre 8 y 529 Kb/s.
- Calidad comparable a MP3 con el 50-70% de caudal. Calidad CD estéreo con 96 Kbps.
- Soporte multicanal (idiomas) y 5.1 (5 canales más *surround*). Permite hasta 48 canales simultáneos.
- Hay una versión adaptada para bajo retardo, especialmente pensada para telefonía (AAC-LD, *Low Delay*).

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
- 3. Vídeo digital. Estándares. Compresión**
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
- 4. Protocolos RTP y RTCP**
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

## 6.4 Resumen

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
- 5. Calidad de servicio**
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía sobre Internet
9. Protocolo SIP

## 6.4 Resumen

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
- 6. Videoconferencia. Estándares H.32x**
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
9. Protocolo SIP

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
- 7. Pasarelas e Interoperabilidad**
8. Telefonía Internet
9. Protocolo SIP

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
- 8. Telefonía Internet**
9. Protocolo SIP

## 6.4 Índice

1. Fundamentos de telefonía. Conversión analógico-digital.
2. Audio digital. Estándares. Compresión
3. Vídeo digital. Estándares. Compresión
4. Protocolos RTP y RTCP
5. Calidad de Servicio
6. Videoconferencia. Estándares H.32x
7. Pasarelas e Interoperabilidad
8. Telefonía Internet
- 9. Protocolo SIP**