

UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA



# “Aspectos relevantes en el diseño de aplicaciones multimedia para Internet”

---

Agustín J. González

Profesor Depto. Electrónica, UTFSM

[agustin.gonzalez@usm.cl](mailto:agustin.gonzalez@usm.cl)



# Contenido

---

- Conceptos
- Análisis de la transmisión de audio en Internet
- Comportamiento de Internet: retardo, sus variaciones y pérdida de paquetes
- Análisis de la transmisión de video en Internet
- Aspectos no cubiertos
- Aplicaciones: [MBone](#), [VRVS](#), [Realnetworks](#), WindowsMedia, [QuickTime streaming Server](#), [FFMPEG](#)



# ¿Qué es mutlimedia?

---

- Interacción Hombre-computador involucrando texto, gráficos, voz, y vídeo (<http://computing-dictionary.thefreedictionary.com/Multimedia>)
- adj. Que utiliza conjunta y simultáneamente diversos medios, como imágenes, sonidos y texto, en la transmisión de información. Dic. de la Real Academia Española, [www.rae.es](http://www.rae.es)
- “Uno debería hablar de multimedia sólo cuando medios continuos y discretos son utilizados.” R. Steinmetz & K.Nahrstedt.
- El término se ha relajado y algunos hablan de “multimedia” aún en presencia de sólo un medio continuo. Por ejemplo: la radio en Internet.



# Términos relacionados

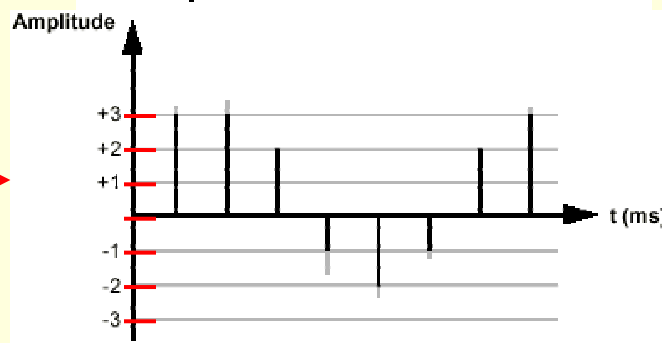
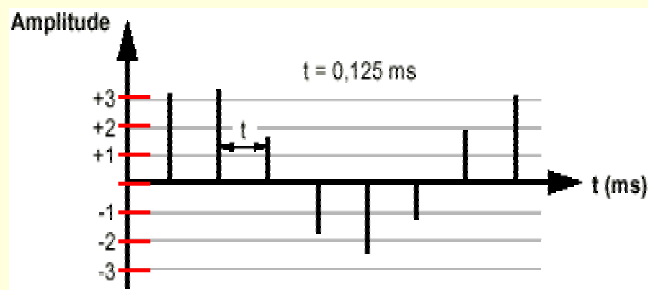
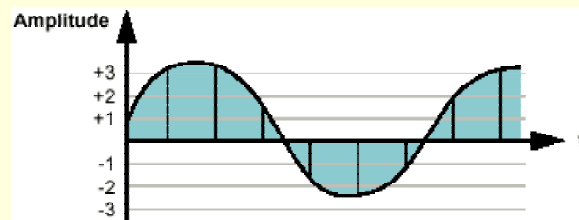
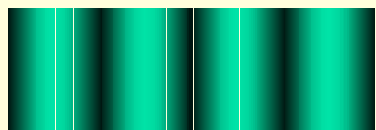
---

- **Streaming:** Reproducción a medida que el contenido es recibido. No se genera copia local. Ej. Televisión, radio por IP, videoconferencia.
- **En vivo:** el contenido es generado en ese momento. Se dice también sincrónico.
- **En demanda:** El contenido está almacenado.
- **Interactivo:** dos usos:
  - podemos hacer pausa, rebobinar, adelantar etc.
  - Cuando hay interacción entre los dos extremos: video conferencia.
- Si no hay simultaneidad, podemos optimizar la codificación al no tener restricciones temporales. Si no hay interacción, en vivo es similar a en demanda.

# Análisis de la transmisión de audio en Internet

- El sonido es una onda generada por la perturbación de la presión en el ambiente (onda de presión).
- Los humanos oímos sonidos entre 20 y 20 [KHz]
- El micrófono es el transductor sonido a señal eléctrica, cuya señal es discretizada en tiempo y en amplitud.

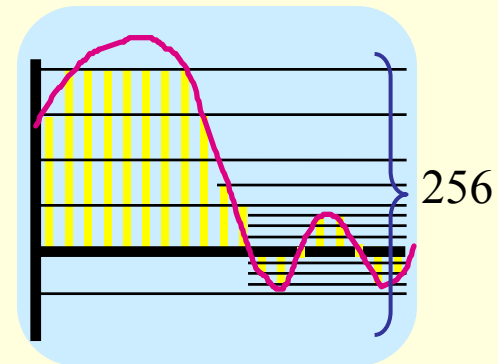
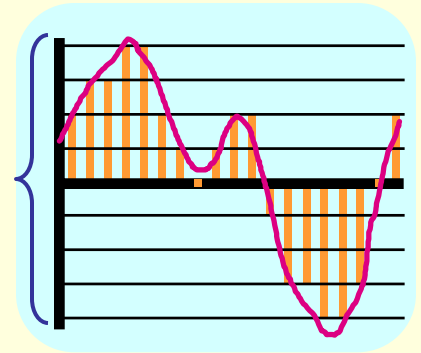
Onda de presión



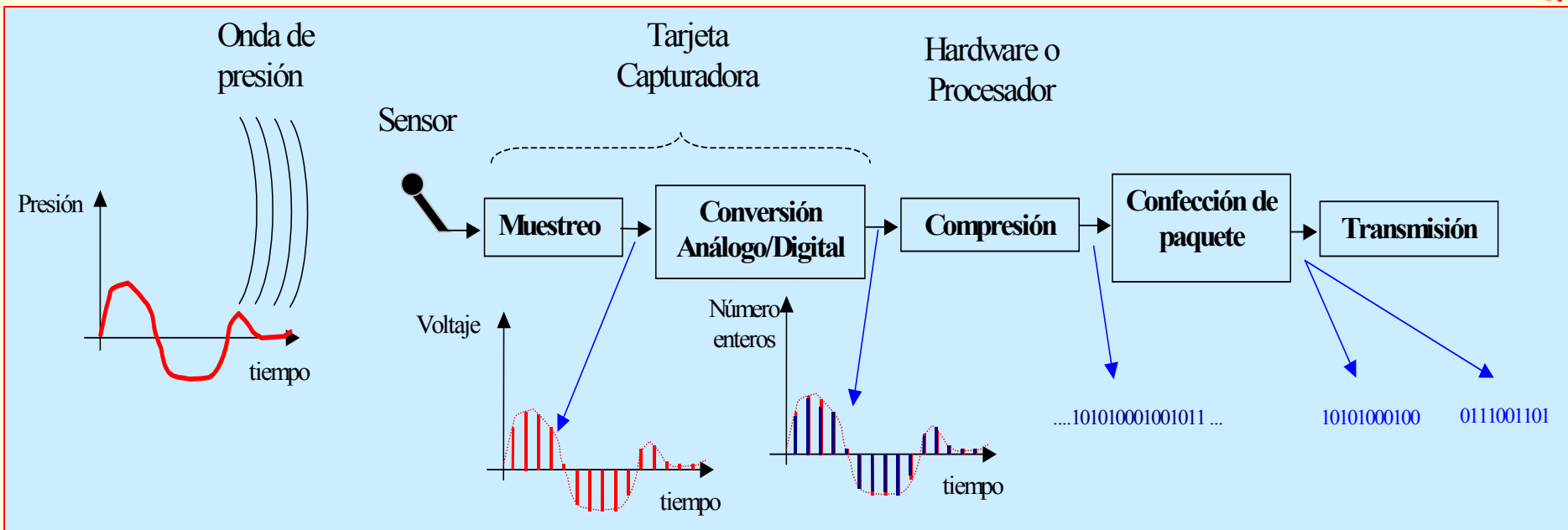
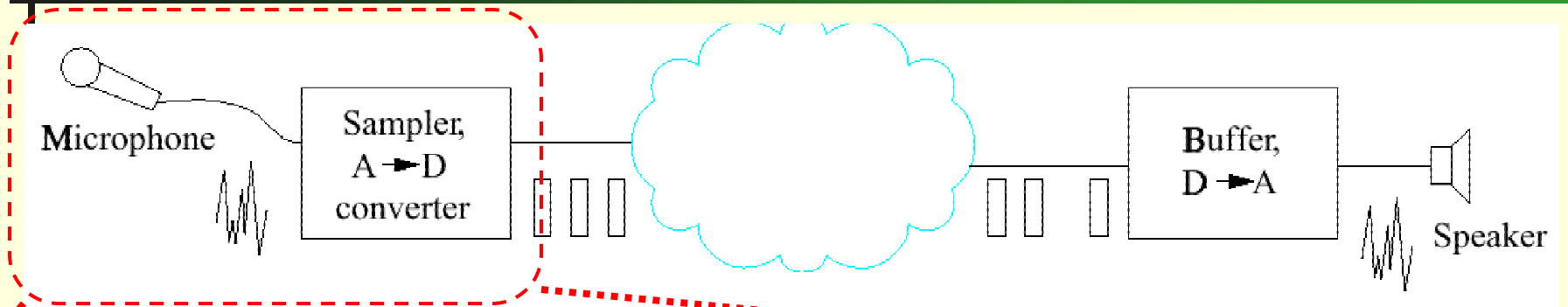
# Sonido/voz Digital

- Podemos reducir datos si el muestreo se reduce: para voz basta 8 KHz, para captar nuestro rango audible se usa 44.1 KHz.
- Manejamos sólo valores discretos para la amplitud.
- Para una buena calidad de voz 8.192 niveles ( $2^{13}$ ) son suficientes.
- Si el salto discreto lo hacemos pequeño para valores pequeños y grande para los mayores, el número de niveles se puede reducir a 256 niveles (8 bits) por muestra. Se consigue así un error de cuantización parejo para todo el rango.

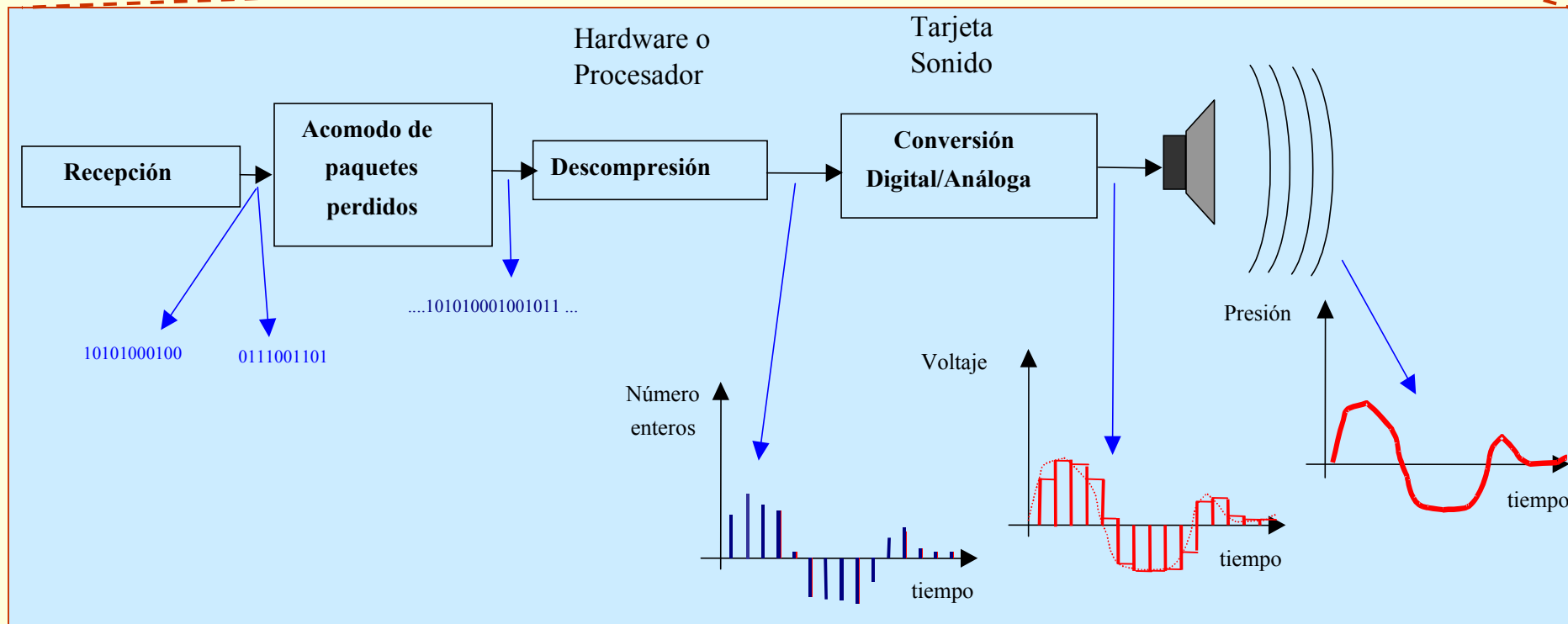
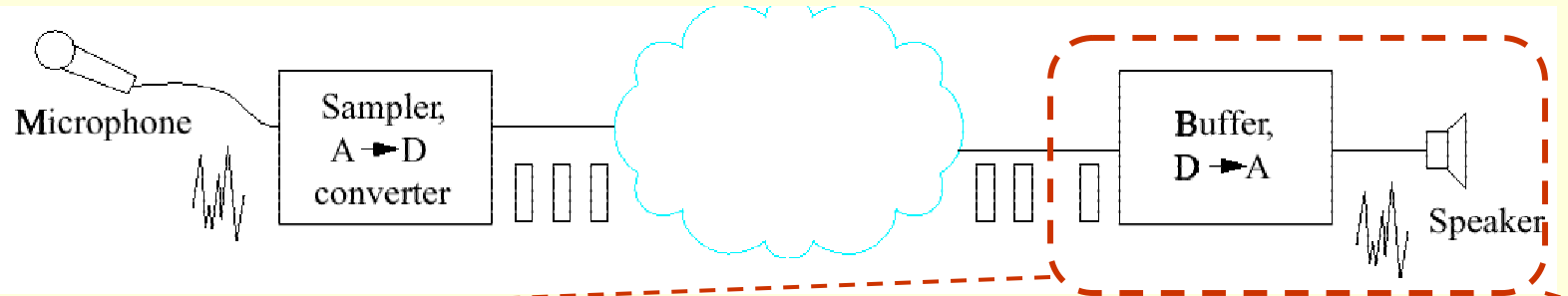
8.192



# De ondas de presión a paquetes



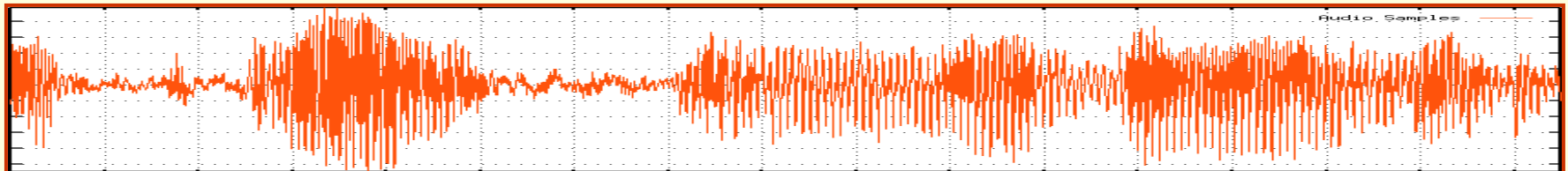
# Proceso inverso: paquetes a onda de presión





# Compresión/codificación

- Una vez que cada muestra de sonido es digitalizada, se acumula entre 20 a 40 ms de audio (en voz => 160 a 320 bytes), se comprimen y estructura un paquete para su transmisión.
- La compresión es la eliminación de redundancia y eventualmente información "poco" relevante. Objetivo es reducir ancho de banda o almacenamiento .  
$$1 \text{ min} \Rightarrow 44100[\text{muestra} / \text{s}] * 2[\text{canales}] * 2[\text{byte} / \text{muestra}] * 60[\text{s} / \text{min}] \approx 10 \text{ MB}$$
$$\Rightarrow \sim 10 [\text{min}] \text{ a } 128 [\text{Kbps}]$$
- Silencio es redundancia: no se transmite.

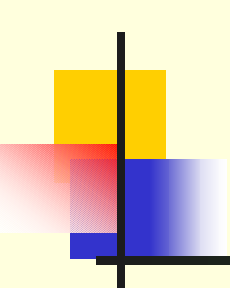




# Ideas usadas en compresión

---

- En lugar de enviar cada muestra codificada, enviar sólo las diferencias que al ser menores se pueden representar con menos bits.
- Usando las muestras previas predecir la siguiente y enviar la diferencia entre la predicción y la señal real.
- Descomponer la señal como la suma de un conjunto de frecuencias preestablecidas y enviar la amplitud de cada una de ellas.
- Modelar analíticamente el tracto vocal y enviar los parámetros que mejor se ajustan el modelo al trozo de señal.
- Similar al anterior enviando también la señal de error entre lo obtenido con el modelo y lo real.



# Contenido sensible al tiempo: Introducción de marcas de tiempo

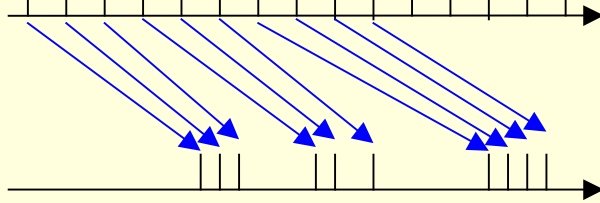
---

- El sonido es una señal continua, si cada muestra es recibida podremos mantener la relación temporal del contenido.
- Sin embargo,
  - la eliminación del silencio,
  - la pérdida de paquetes,
  - la necesidad de sincronización con otros mediosobligan incorporar marcas de tiempo en cada paquete para su posterior reproducción en forma sincronizada.
- Las marcas de tiempo permiten también estimar las variaciones de retardo en la red y ajustar el retardo de reproducción.

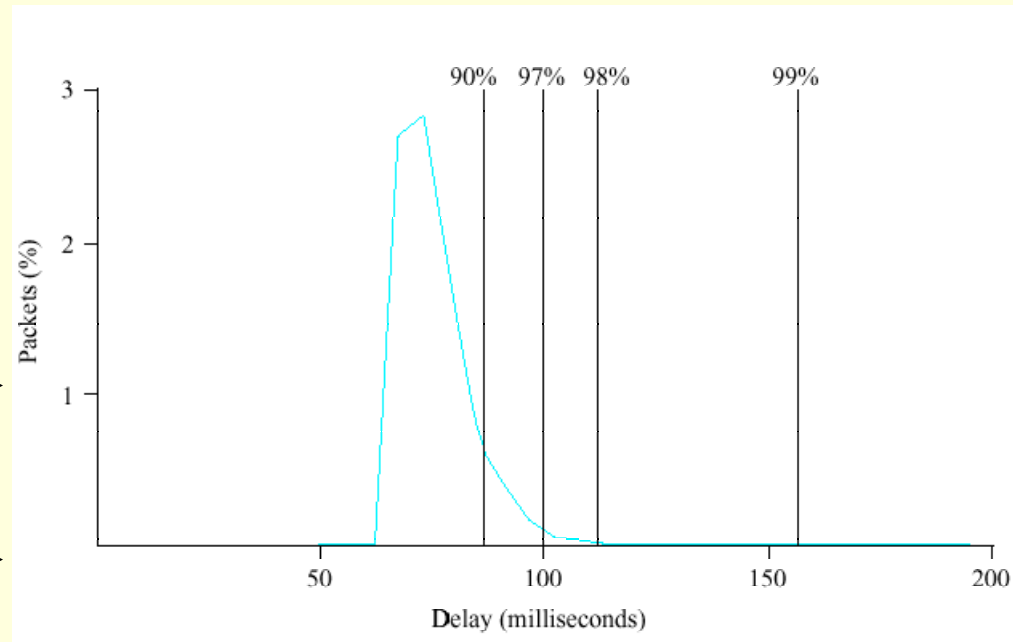
# Servicio ofrecido por Internet

- Best-effort: Mejor esfuerzo, básicamente no garantiza nada.
- Principales problemas: retardo variable y pérdidas de paquetes.

Transmisión de  
paquetes



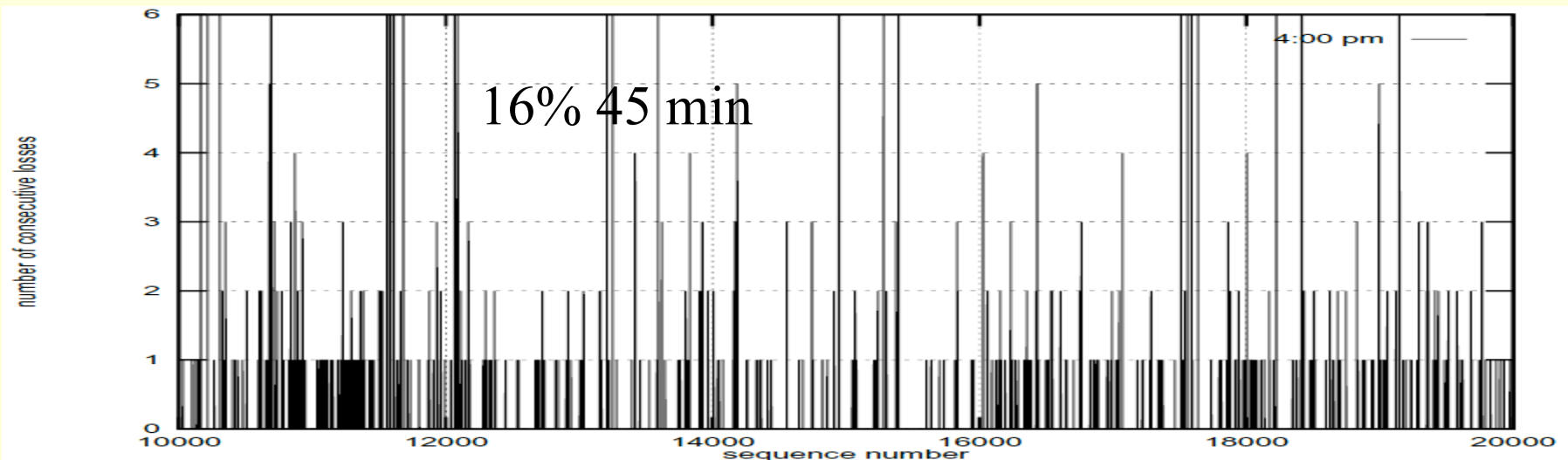
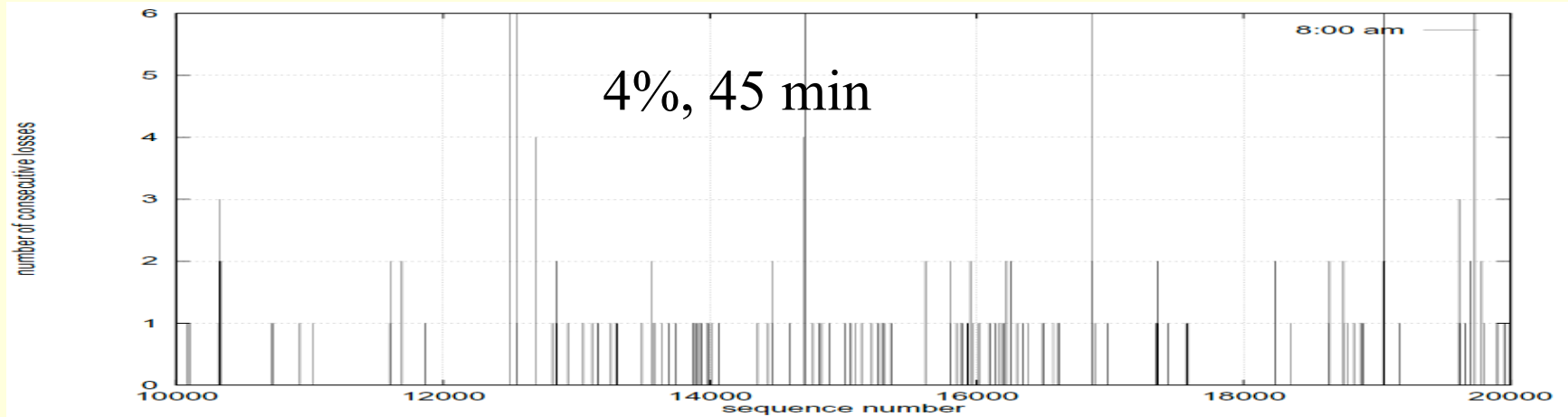
Recepción de  
paquetes



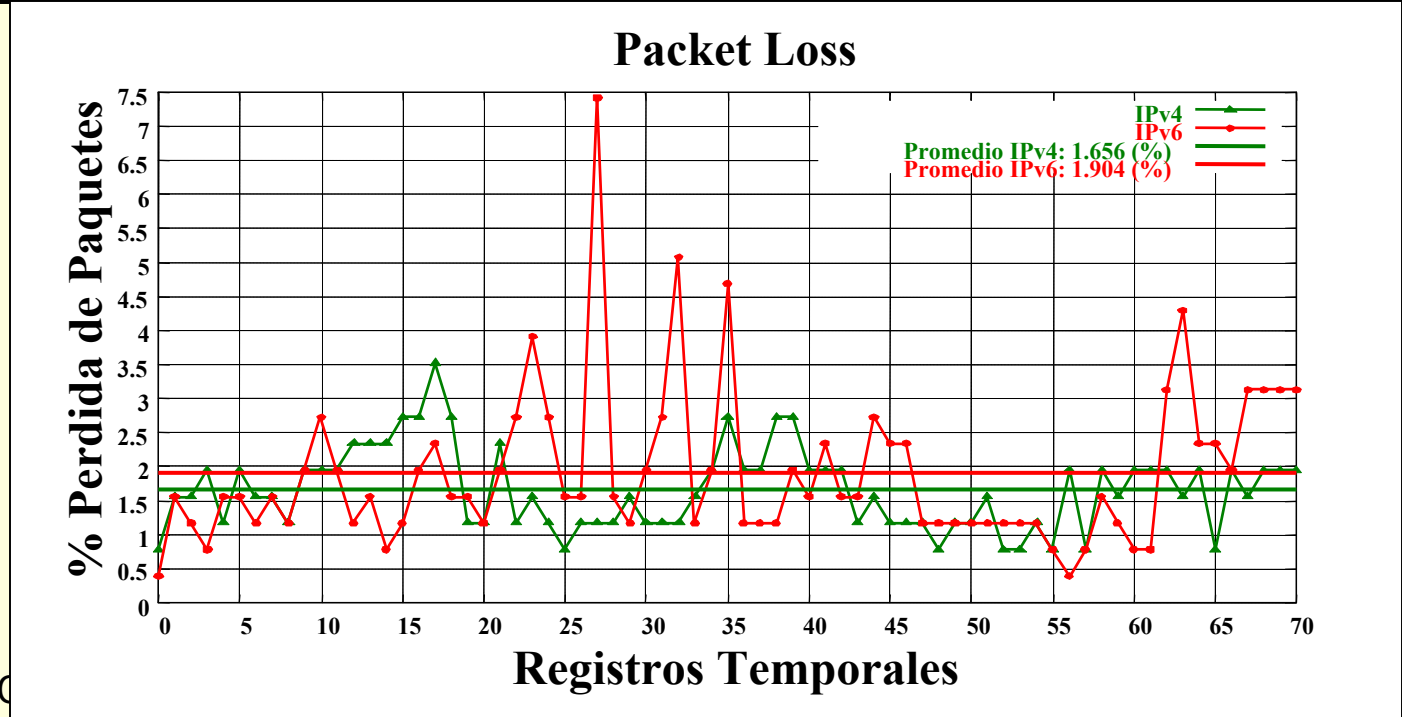
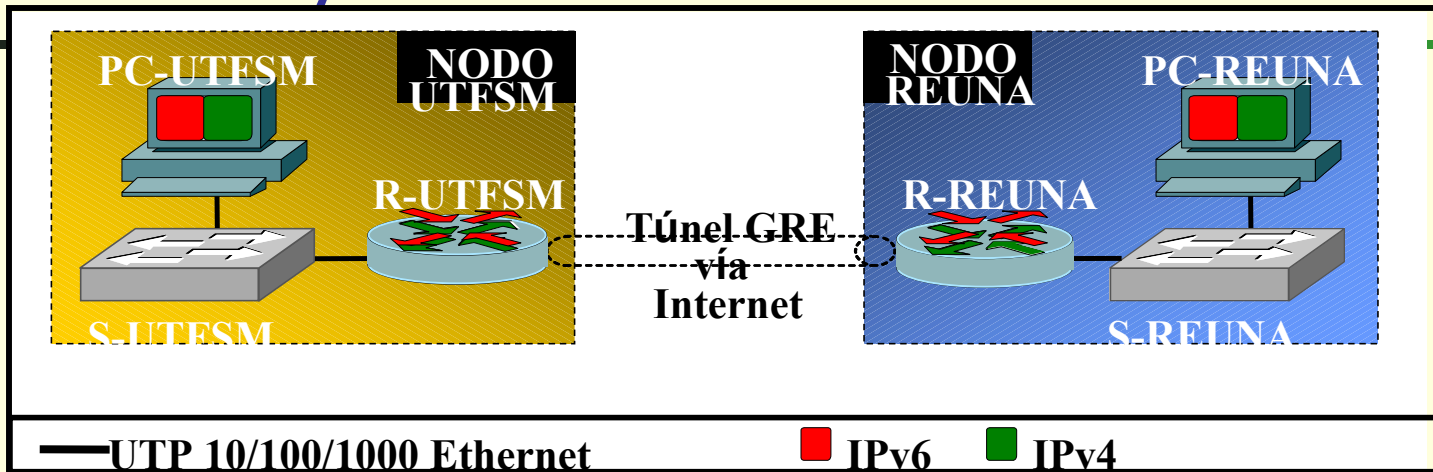
Ejemplo: One-way delay para una ruta en Internet

# Pérdida de paquetes en Internet

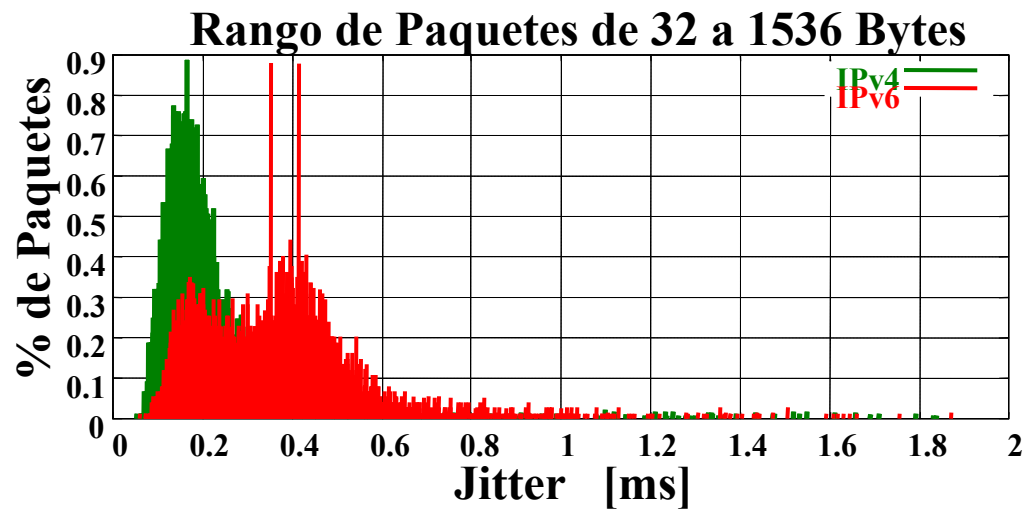
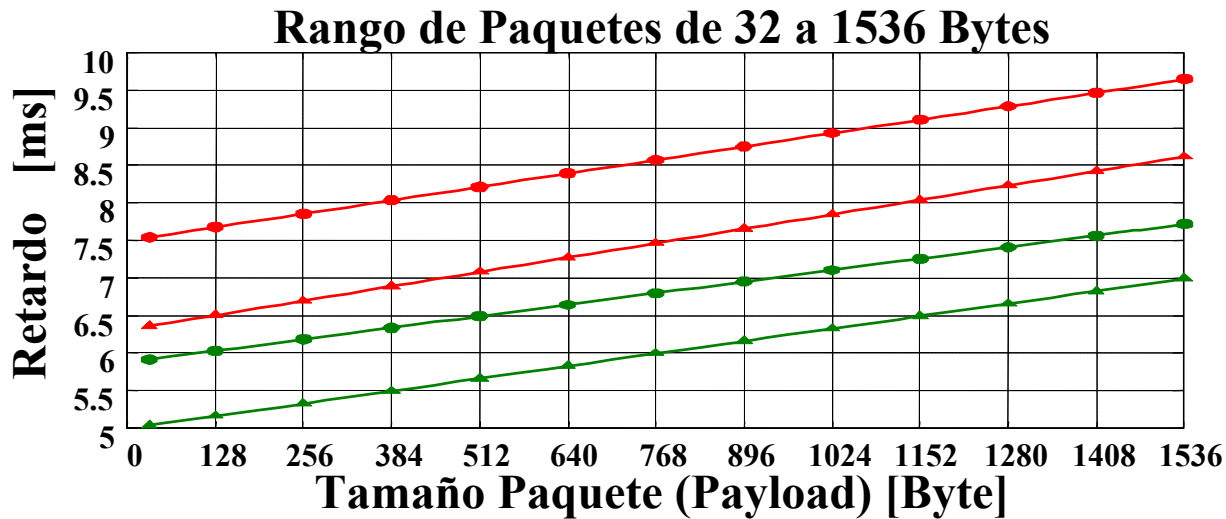
- Mediciones entre Inglaterra y Francia a las 8:00 am y 4 pm



# Pérdidas de paquetes y retardo entre UTFSM y REUNA

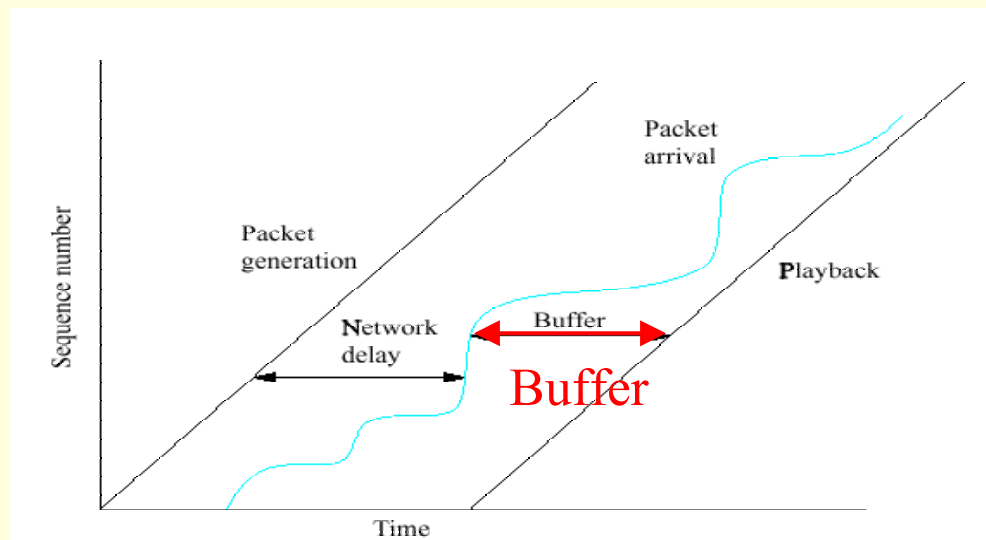
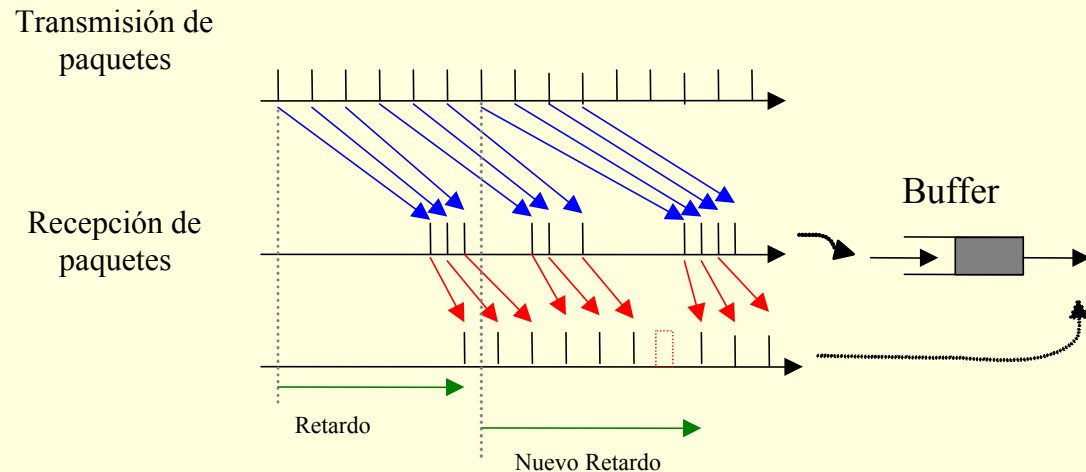


# Retardo entre UTFSM y REUNA



# ¿Cómo hacemos frente al retardo y pérdidas?

- Retardo: Se uniforma usando un buffer como “colchón”.



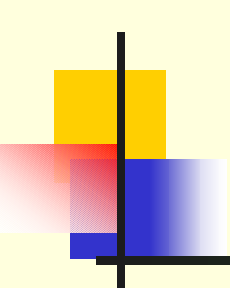




# ¿Cómo hacemos frente a las pérdidas?

---

- Básicamente: aminorar la carencia o agregar redundancia inteligentemente.
- Retransmisiones (TCP) sólo son posibles en casos no interactivos. En éstos usamos UDP.
- Opciones sin agregar información extra:
  - Introducir ruido en lugar del paquete perdido
  - Repetir el paquete previo
- Opciones agregando redundancia (**F**orward **E**rror **C**orrection):
  - Usar códigos correctores por ejemplo cada dos paquetes enviar uno de paridad de los dos primeros.
  - En cada paquete enviar una versión de menor calidad del anterior.

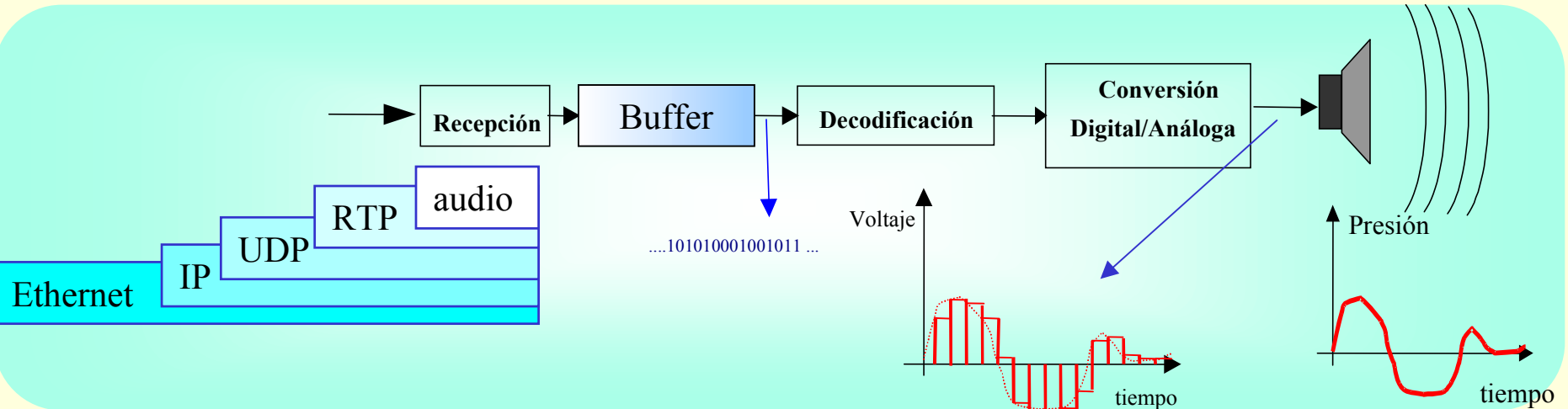
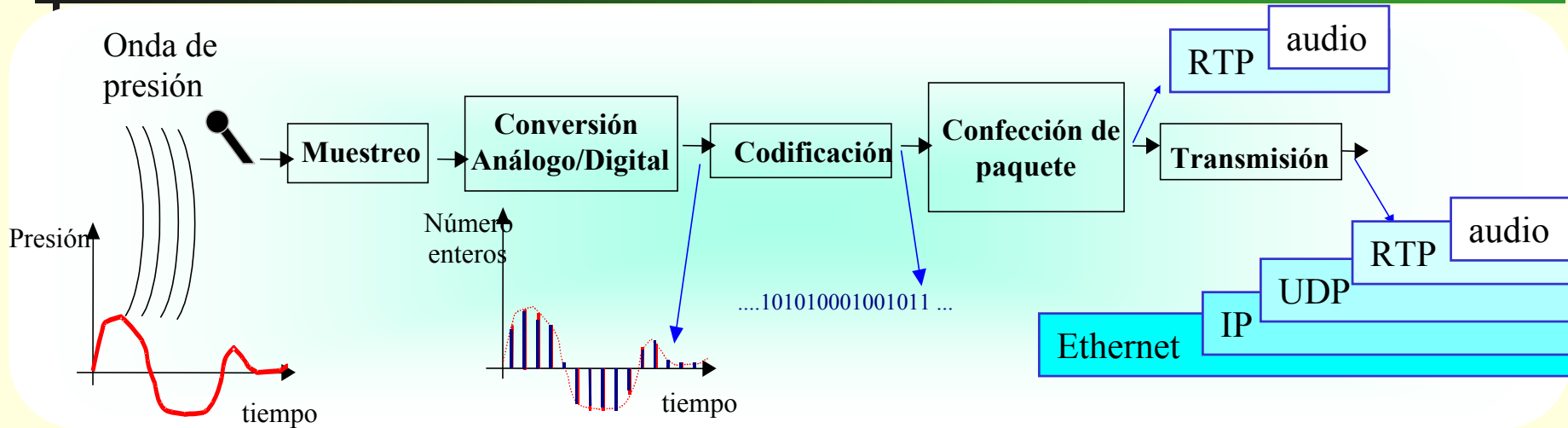


# Estructura de la trama: **Real Time Protocol**

---

- RTP es un protocolo que estandariza el formato de los paquetes de contenido multimedia en Internet (audio, video y otros).
- Ofrece servicios
  - Identificación del tipo de contenido del paquete
  - Números de secuencia
  - Marcas de tiempo
- No ofrece:
  - Garantía de llegada a tiempo
  - Calidad de servicio
  - Entrega de paquetes en orden

# Resumen de audio en Internet





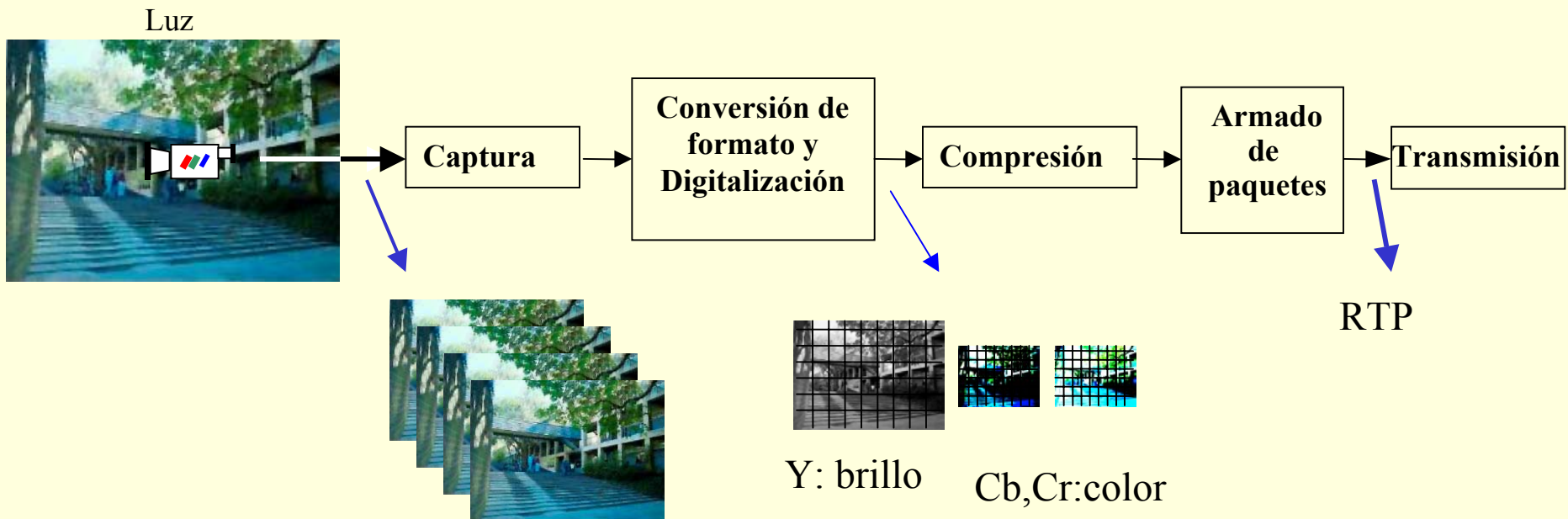
# Contenido

---

- ✓ Conceptos
- ✓ Análisis de la transmisión de audio en Internet
- ✓ Comportamiento de Internet: retardo, sus variaciones y pérdida de paquetes
- ■ Análisis de la transmisión de video en Internet
  - Aspectos no cubiertos
  - Aplicaciones: MBone, VRVS, Realnetworks, WindowsMedia, QuickTime streaming Server, FFMPEG

# Análisis de la transmisión de video en Internet

- Vemos en forma continua la luz reflejada por los objetos.
- El video son secuencias discretas de imágenes.





# Captura de video

---

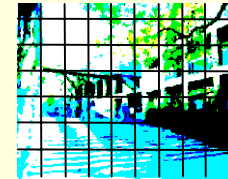
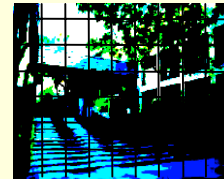
- Si proviene de una cámara grabadora (No webcam), ésta genera 30 cuadros por segundo en formato NTSC.
- La tarjeta muestrea y obtiene las componentes de luminancia y crominancia para cada pixel.
- Las resoluciones son típicamente múltiplos o sub-múltiplos de 320x240 para NTSC.

# Conversión de formato

- El ojo humano distingue mejor la intensidad de la luz o brillo que el color, por ello cada cuadro es almacenado con  $1/4$  de la resolución para cada componente de color.



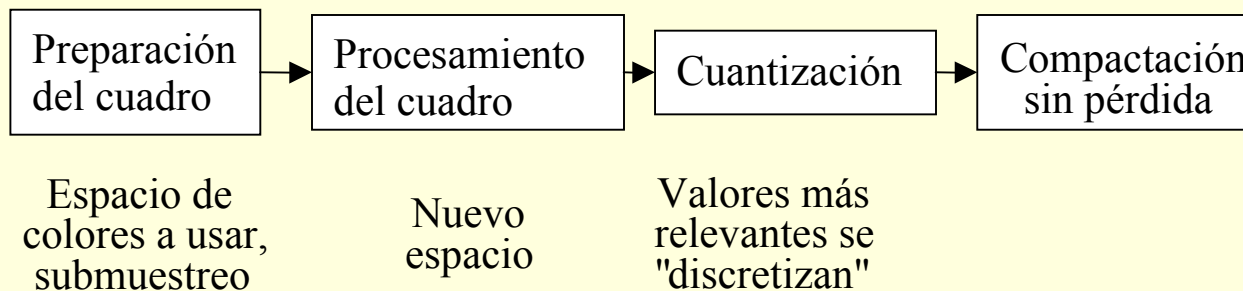
Intensidad



Crominancia del azul y rojo

# Compresión

- Se trata de eliminar la redundancia y aquello imperceptible para el ojo.
- Redundancia
  - Espacial: como con imágenes hay zonas regulares
  - Temporal: cuadros seguidos son parecidos
  - Psicovisual: no vemos los detalles
- Pasos en la compresión:







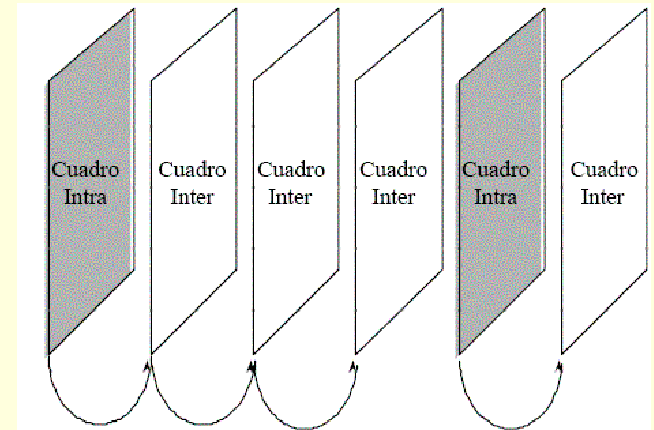
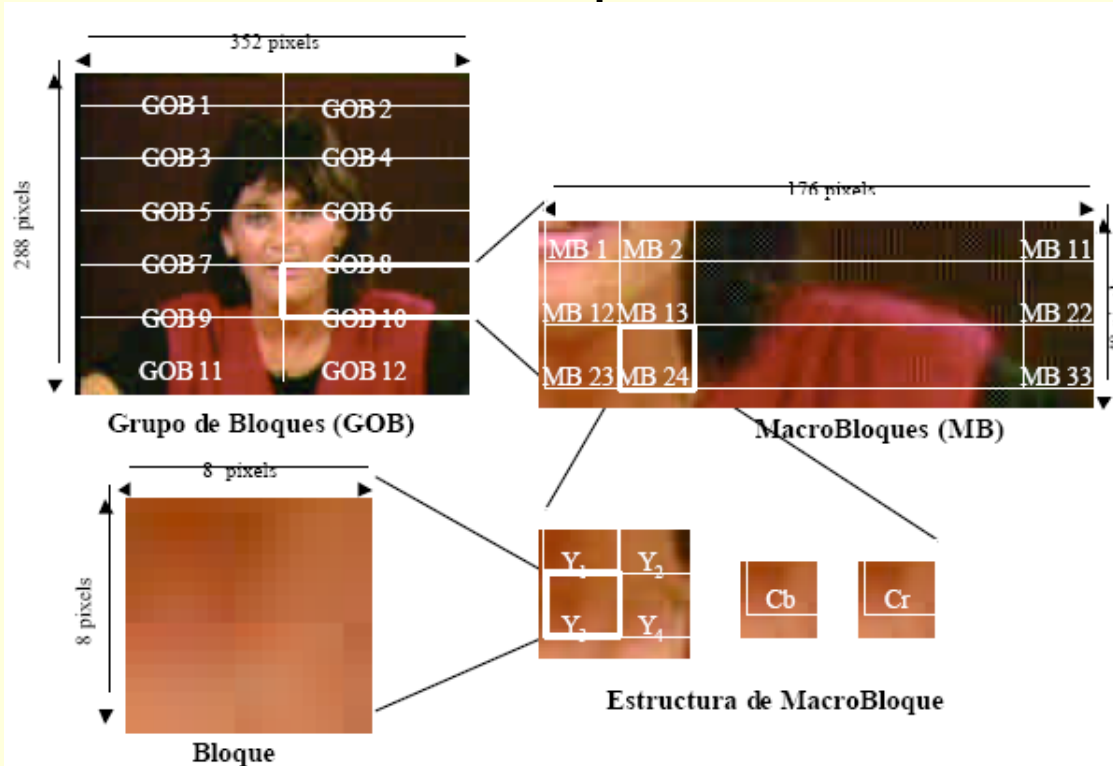
# Ideas usadas en la compresión

---

- No necesitamos 30 fps, 10 puede bastar.
- En muchos casos resoluciones de 320\*240 basta.
- Cada cuadro o frame se subdivide en cuadrados de 8x8 pixeles.
- Como notamos más los valores promedios que los detalles, éstos son codificados con mayor precisión, los detalles con menor o son eliminados.
- Se toman diferencias del cuadrado con respecto al cuadro anterior.
- Como la imagen se pudo correr, se busca el cuadro en un entorno (predicción del movimiento).

# Consideraciones de la red

- Cada cuadro no cabe en un paquete, entonces se agrupan varios rectángulos autónomos.
- Como hay pérdida de paquetes debemos enviar cada cierto rato cuadros comprimidos como imágenes (Intra)





# El Receptor

---

- Básicamente hace la operación inversa.
- El despliegue se hace en el monitor (RGB) por ello se ocupa mucha CPU en la descompresión y el cambio de formato de cada pixel.
- Hay alto movimiento de datos en los buses internos del computador que pueden frenar el sistema.
- El audio generalmente debe esperar al video para reproducción sincrónica.



# Aspectos no cubiertos sobre audio y video

---

- Sincronización intra-stream: proceso de reproducir el contenido respetando su dependencia temporal.
- Sincronización inter-stream: proceso de reproducir un contenido respetando la dependencia temporal con otro.
- Mezcla de múltiples contenidos de audio. Conferencias entre más de dos.
- Herramientas para programar aplicaciones multimedia.



# Aplicaciones

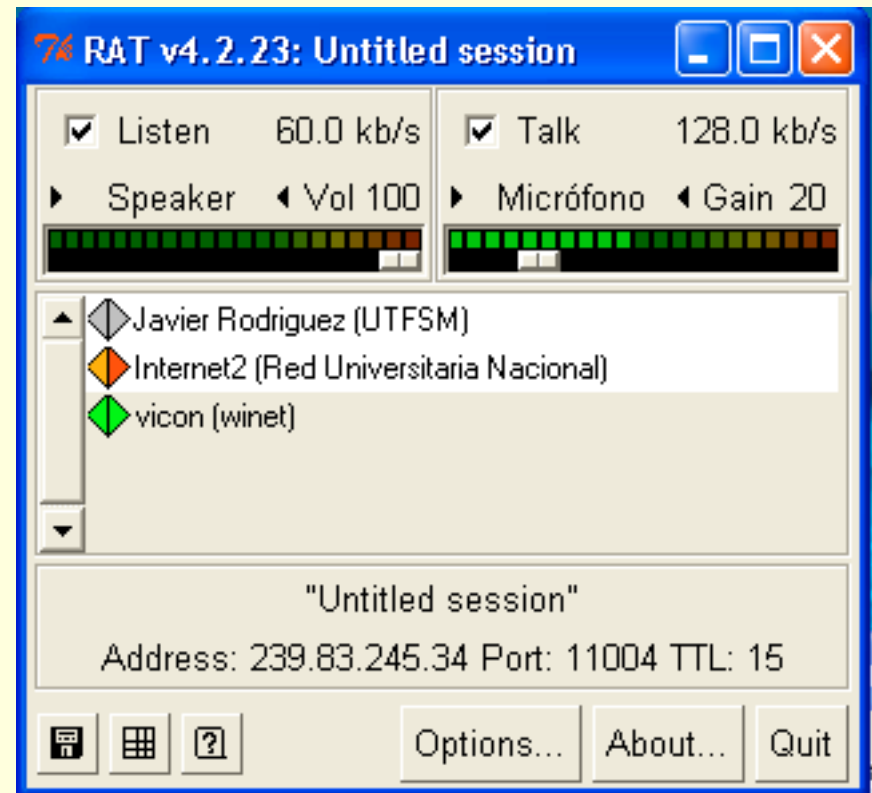
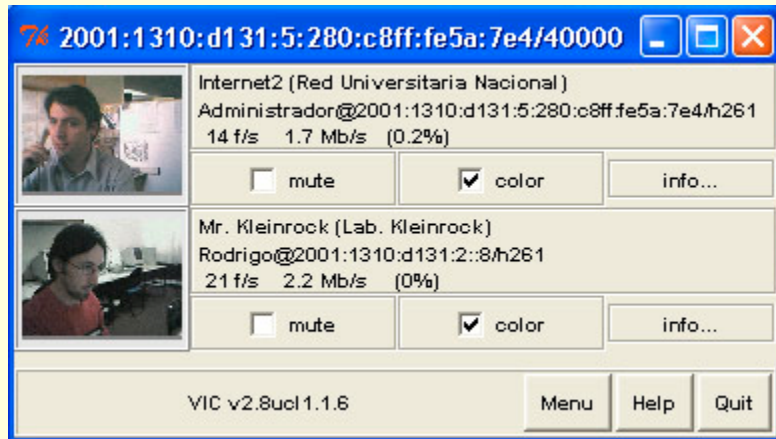
---

- Para video conferencia:
  - MBone,
  - VRVS,
  - NetMeeting,
  - Messenger,
  - .....
- Para video streaming:
  - RealNetworks,
  - Windows Media,
  - QuickTime streaming Server,
  - FFMPEG
  - .....

# Aplicaciones MBone

rat: robust audio tool

Vic: video conferencie tool



Soporte IPv6 !

# VIC

The screenshot displays the VIC interface with several components:

- Plot Windows:**
  - Kilobits:** range 0 to 500, 50/div
  - Packets:** range 0 to 100, 10/div
  - Frames:** range 0 to 50, 5/div
  - Missing:** range 0 to 5, 0.5/div
  - Misordered:** No data
- Video Window:** Cristian E. Henry
- RTP Statistics Table:**

	EWA	Delta	Total
Kilobits	247.9	169.0	95680
Frames	18.1	20.0	13143
Packets	40.2	31.0	19301
Missing	0.0	0.0	19
Misordered	0.0	0.0	0
Runts	0.0	0.0	0
Dups	0.0	0.0	0
Bad-S-Len	0.0	0.0	0
Bad-S-Ver	0.0	0.0	0
Bad-S-Opt	0.0	0.0	0
- Metadata Window:**

```

Cristian E. Henry
format: h261 [q=10] (352x288)
cname: Administrador@146.83.245.34
tool: vic-5.2/Windows NT-5.0-intel
email: chenny@reuna.cl
note:
srcid: 1653613589/146.83.245.34
last data Fri Oct 24 16:41:31 2003
last control Fri Oct 24 16:41:30 2003
          
```
- Taskbar:** Shows three video feeds: Rodrigo Guerra, Cristian E. Henry, and Javier Rodriguez.



# VRVS

---

- Virtual Rooms VideoConferencing System
- Es un sistema distribuido que establece su propia red entre nodos (reflectores) distribuidos a lo largo del planeta.
- Es una red sobre la actual Internet (e Internet2).
- Su principal ventaja es la facilidad de uso. Los usuarios ven salas virtuales en las cuales se pueden reunir con audio, video, y datos a través de VNC (Virtual Network Computing)





# Real Networks

---

- Existen versiones para varios sistemas operativos (Solaris, Linux, Windows, etc.)
- Usado por canales de televisión difusión de videos etc..
- Radio UTFSM
- Clases en línea desde Berkeley: Herencia
- Canal 13
- TVN



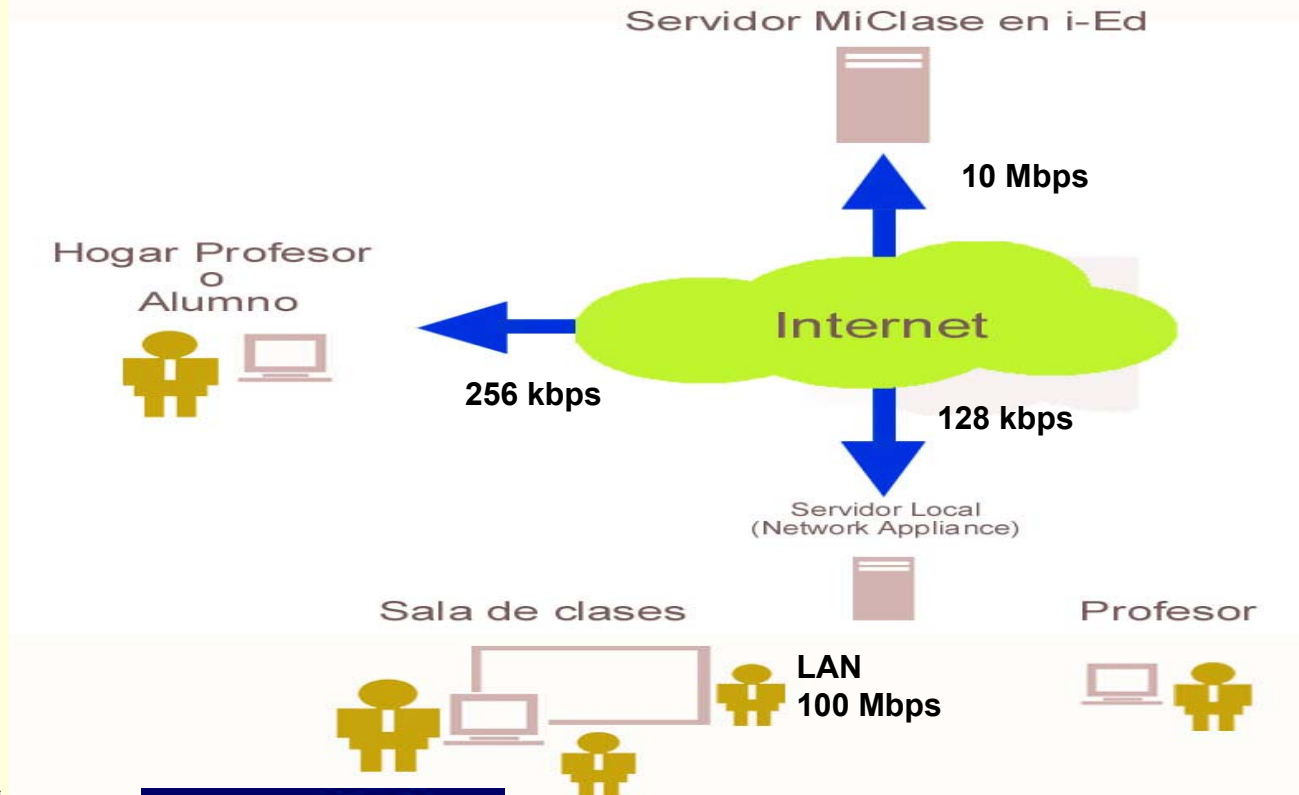
# Windows Media Player

---

- Sólo para usuarios Windows.
- Varias radios y canales de televisión ya lo usan:
- [Radio Cooperativa](#)
- Canal 13
- [TVN](#)

# Proyecto Red de Distribución de Contenidos para MiClase

- Idea: Usar videos en las clases de una escuela



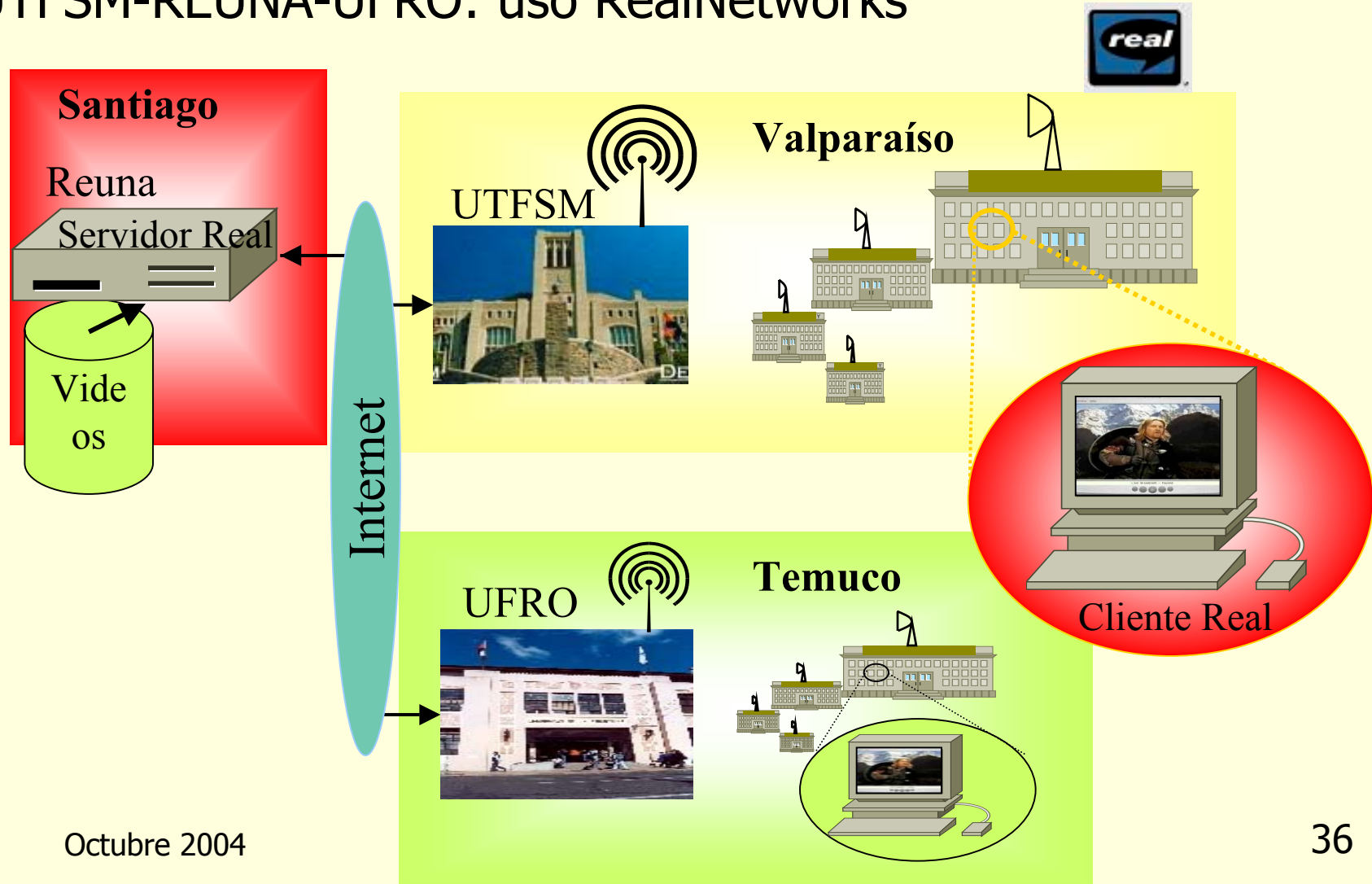
UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA





# Proyecto difusión de video en Valpo.

- UTFSM-REUNA-UFRO: usó RealNetworks



# Solución a nivel de red

## Red Inalámbrica en la V Región

