

# **ESCUELA POLITÉCNICA NACIONAL**

**FACULTAD DE INGENIERÍA ELÉCTRICA Y  
ELECTRÓNICA**

**DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA DE VIDEO-  
VIGILANCIA, EN BASE A UN SISTEMA DE VIDEO EMBEBIDO,  
PARA MONITOREO REMOTO A TRAVÉS DE INTERNET**

**PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN  
ELECTRÓNICA Y TELECOMUNICACIONES**

**GUALOTO PILLALAZA ROBERTO CARLOS**

**DIRECTOR: ING. FERNANDO FLORES**

**Quito, febrero 2008**

## **DECLARACIÓN**

Yo, Gualoto Pillalaza Roberto Carlos, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

**Gualoto Pillalaza Roberto Carlos**

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por Gualoto Pillalaza Roberto Carlos, bajo mi supervisión.

**Ing. Flores Fernando**  
**DIRECTOR DE PROYECTO**

## **AGRADECIMIENTO**

Agradezco a Dios por mantenerme siempre lleno de fuerza, esperanza y vida, ya que en todo paso que he dado he sentido su importante presencia en mi vida.

**Roberto Carlos**

## **DEDICATORIA**

El presente trabajo se lo dedico mis padres: Luis y María y a mis cinco hermanos: Paty, Marcia, Susana, Edison y Wilfrido por su apoyo y sus consejos en todo este caminar, así mismo se lo dedico a Mónica mi esposa y a mi hija Fátima que viene en camino.

**Roberto Carlos**

<b>PRESENTACIÓN.....</b>	<b>ix</b>
<b>RESUMEN .....</b>	<b>x</b>
<b>CAPÍTULO I.....</b>	<b>1</b>
<b>1 VIDEO DIGITAL: INICIOS, PROTOCOLOS Y NUEVOS AVANCES .....</b>	<b>1</b>
1.1 Introducción.....	1
1.2 Sistemas de Televisión.....	4
1.2.1 Generación de la imagen .....	9
1.2.1.1 Sensor de imagen .....	9
1.2.1.2 Tecnología CCD.....	10
1.2.1.3 Tecnología CMOS .....	12
1.2.1.4 Resolución.....	12
1.2.2 Medios ópticos.....	15
1.2.3 Barrido entrelazado frente al barrido progresivo.....	17
1.2.4 Dispositivos de video.....	18
1.2.5 Tipos de señales de video.....	22
1.3 Digitalización de Video.....	23
1.3.1 Recomendación ITU 601.....	24
1.3.2 Necesidad de comprimir.....	27
1.3.2.1 Principios de compresión.....	28
1.3.2.2 Redundancia .....	29
1.3.2.3 Técnicas de compresión.....	30
1.3.3 Formatos de video digital y reproductores.....	37
<b>CAPÍTULO II.....</b>	<b>42</b>
<b>2 VIDEO SOBRE REDES IP .....</b>	<b>42</b>
2.1 Introducción.....	42

2.1.1	Evolución necesaria de la Internet.....	42
2.1.2	Requerimientos para la transmisión de multimedia .....	43
2.1.3	Calidad de Servicio .....	45
2.2	Redes IP.....	48
2.2.1	Direccionamiento IPv4 .....	50
2.2.2	Direccionamiento IPV6 .....	52
2.3	Video sobre IP.....	53
2.3.1	Broadcast sobre IP.....	54
2.3.2	Streaming.....	56
2.3.3	Stack de protocolos TCP/IP para streaming.....	58
2.3.3.1	Protocolo RSVP.....	59
2.3.3.2	Protocolo RTP .....	62
2.3.3.3	Protocolo RTCP.....	65
2.3.3.4	Protocolo RTSP .....	66
2.3.4	SIP.....	67
2.3.5	H.323 .....	72
2.4	Aplicaciones de video digital.....	75
2.4.1	Televisión IP .....	75
2.4.2	Video-Vigilancia .....	78
<b>CAPÍTULO III.....</b>		<b>80</b>
<b>3</b>	<b>DISEÑO DEL SISTEMA DE VIDEO-VIGILANCIA .....</b>	<b>80</b>
3.1	Introducción.....	80
3.2	Descripción del sistema de video-vigilancia.....	84
3.3	Diseño del sistema .....	87
3.3.1	Internet como red de video-vigilancia.....	87
3.3.2	Sistema de video embebido .....	88

3.4	Implementación .....	97
3.4.1	Configuración del sistema .....	101
3.4.2	Pruebas.....	113
<b>CAPÍTULO IV.....</b>		<b>119</b>
<b>4</b>	<b>CONCLUSIONES Y RECOMENDACIONES .....</b>	<b>119</b>
	<b>BIBLIOGRAFÍA .....</b>	<b>121</b>
	<b>ANEXOS .....</b>	<b>123</b>



## **PRESENTACIÓN**

Los sistemas de video-vigilancia en los últimos años han experimentado un gran desarrollo; la era digital de la información la ha enrumado hacia la convergencia de tecnologías, es así como las redes y sistemas de comunicación dan origen a la video-vigilancia IP(Internet Protocol), la convergencia hacia IP le ha permitido integrarse a estándares y protocolos internacionalmente adoptados.

Los sistemas de video-vigilancia IP dejan atrás a los tradicionales CCTV (Closed Circuit Television Systems) analógicos, brindan mejores prestaciones y junto al desarrollo y expansión de la Internet hacen posible la video-vigilancia remota. La Internet por su parte a pesar que no fue diseñada para aplicaciones como el video en tiempo real, se ha adaptado mediante protocolos complementarios para hacerlas posible.

En el presente proyecto se plantea el uso de un equipo que posee embebido los componentes necesarios para la video-vigilancia a través de Internet, empleando tecnología streaming.

## **RESUMEN**

El objetivo del proyecto es diseñar e implementar un sistema de video-vigilancia a través de Internet, empleando un sistema de video embebido.

En el primer capítulo se describen los principios del video digital, los principales estándares de compresión y la televisión y la video-vigilancia como aplicaciones del video IP.

En el segundo capítulo se revisa el video IP, calidad de servicio y requerimientos de un sistema video sobre Internet.

En el tercer capítulo, previo una revisión de los sistemas de video-vigilancia se realiza el diseño y configuración del sistema de video-vigilancia, además de pruebas y evaluación del sistema de video embebido.

En el cuarto capítulo se establece las principales características, ventajas y desventajas del sistema de vide-vigilancia.

# CAPÍTULO I

## 1 VIDEO DIGITAL: INICIOS, PROTOCOLOS Y NUEVOS AVANCES

### 1.1 Introducción

Etimológicamente la palabra Video proviene del verbo latín videre que significa "Ver", estimula el sentido de la visión que aporta con la mayor cantidad de información en situaciones de la vida diaria, permitiéndonos apreciar formas, colores, texturas y movimientos.

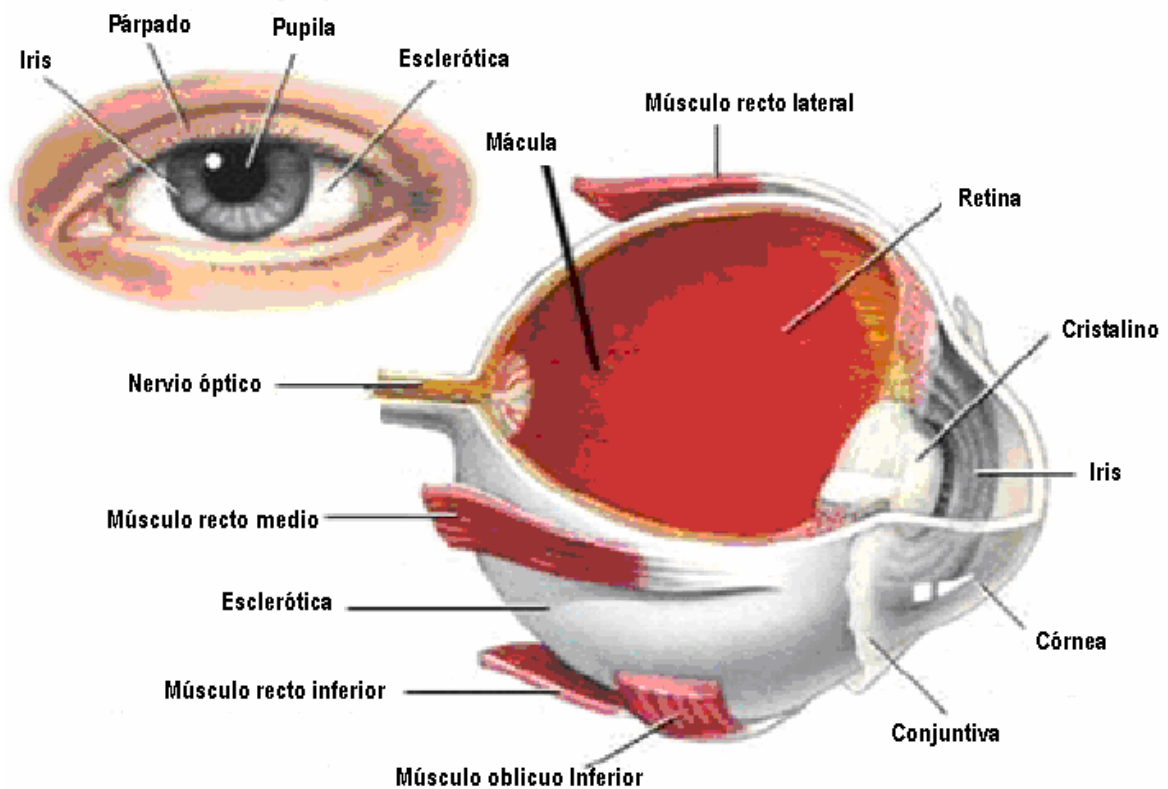


Figura 1.1 Sistema visual<sup>1</sup>

<sup>1</sup> SAGAMAR Deb, VIDEO DATA MANGEMENT AND INFORMATION RETRIEVAL, University of Southern Queensland, Australia.

El desarrollo de los sistemas de video ha sido posible gracias al estudio del comportamiento y características del sistema visual humano<sup>2</sup> cuyo principal órgano visual es el globo ocular, formado por un complejo entramado de nervios y tejidos; encargado de captar y transformar la energía luminosa en señales electroquímicas que son enviadas hacia áreas cerebrales encargadas de la proyección e interpretación visual de la información Figura 1.1. Los procesos que se llevan a cabo en el órgano visual, son análogos a los que se llevan a cabo dentro de una cámara de video.

El video se logra gracias a tres factores característicos: la persistencia retiniana, el fenómeno PHI y la frecuencia crítica de fluctuación. La persistencia retiniana retiene la imagen 1/15 de segundo después de que la imagen se retira de la vista; el fenómeno PHI provoca la sensación de movimiento y la frecuencia crítica de fluctuación provoca la sensación de continuidad de imágenes.



Figura 1.2 Espectro visible al ojo humano<sup>3</sup>

La visión percibe radiaciones en forma de energía luminosa en un rango de 380 nm (nanómetros<sup>4</sup>) para el color violeta, hasta 780 nm para el color rojo, en la Figura 1.2 se ilustran estos límites.

### Señal de video<sup>5</sup>

La señal de video contiene información de brillo, color y sincronismo que genera imágenes en movimiento en una pantalla del televisor, un monitor, etc.

<sup>2</sup> [www.whyfiles.org/163amd\\_eye](http://www.whyfiles.org/163amd_eye)

<sup>3</sup> <http://www.ugr.es/~jnieves/Textos/Tema2Psicofisica%20VisionColor%20I%20CONt%20Gelstalt.pdf>

<sup>4</sup> 1 nanómetro =  $1 \cdot 10^{-9}$  metros

<sup>5</sup> [http://www.rediris.es/media/gt/gt2004\\_2/CapturaVGA.ppt](http://www.rediris.es/media/gt/gt2004_2/CapturaVGA.ppt)

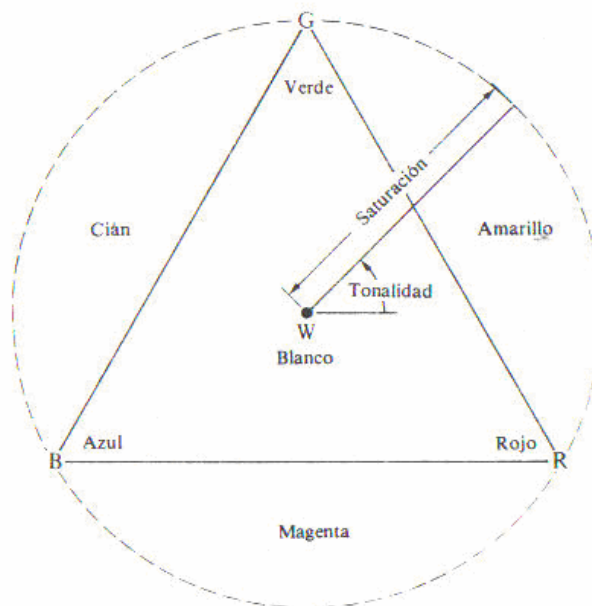


Figura 1.3 Triángulo de color que muestra la tonalidad y saturación<sup>6</sup>

Técnicamente, la señal de video está compuesta por la luminancia o brillantez (Y), la crominancia (Cr) y las señales de sincronismo. La luminancia produce la sensación de claridad proveniente o reflejada de un objeto, la crominancia contiene información del color definida a su vez por la saturación y la tonalidad, las señales de sincronismo determinan períodos de barrido de la imagen adecuados para asegurar la continuidad y estabilidad de la imagen. La luminancia y crominancia se pueden definir en función de los colores primarios: rojo, verde y azul, en la mayoría de literatura se emplea RGB por sus siglas en inglés (Red, Green, Blue).

La información del color se puede establecer en términos de coordenadas circulares Figura 1.3, en donde la saturación varía con la distancia radial y el ángulo representa la tonalidad o el color. Todos estos principios fueron empleados en el desarrollo de los sistemas de televisión que a continuación se describen brevemente.

<sup>6</sup> Imagen digitalizada de la referencia bibliográfica. STREMLER, Ferrel. SISTEMAS DE COMUNICACIÓN Segunda Edición, Editorial Alfaomega, México 1982, p 652.

## 1.2 Sistemas de Televisión

Los sistemas de televisión (TV)<sup>7</sup> son los primeros sistemas de video analógico, permiten la transmisión y recepción de audio y video, para su transmisión emplea ondas electromagnéticas (refiérase al Anexo A.1 para establecer las frecuencias empleadas en cada canal de TV) y actualmente se transmite televisión, a través de redes especializadas.

A nivel global los sistemas de televisión se han estandarizado y dependiendo del sistema empleado, la codificación de la señal de video cambia.

### NTSC<sup>8</sup> (National Television System Committee)

SISTEMA	NTSC
Líneas	525
Campos	60
Frecuencia vertical	60 Hz
Frecuencia horizontal	15.750 KHz
Frecuencia de la suportadora de color	3.579545 MHz
Portadora de audio	4.5 MHz
Ancho de banda del video	4.2 MHz

Tabla 1.1 Características de NTSC

El comité nacional de sistemas de televisión (NTSC), es el estándar empleado actualmente en Canadá, Estados Unidos, México, Japón y países de América del Sur, entre ellos Ecuador.

En la Tabla 1.1 se tiene las principales características de NTSC: número de líneas,

<sup>7</sup> [http://www.ebu.ch/dvb\\_articles/dvb\\_tv-history.htm](http://www.ebu.ch/dvb_articles/dvb_tv-history.htm), Anexo A.

<sup>8</sup> STREMLER, Ferrel, SISTEMAS DE COMUNICACIÓN, Segunda Edición, Editorial Alfaomega, México 1982.

frecuencia vertical, frecuencia horizontal, frecuencia de la portadora de video y frecuencia de la portadora de audio.

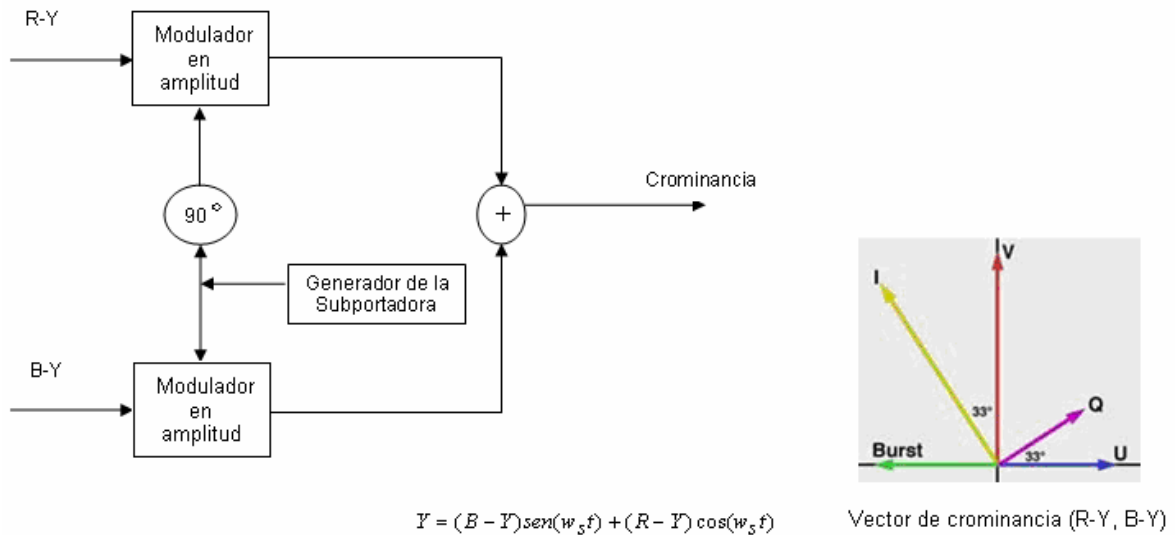


Figura 1.4 Diagrama de la modulación en NTSC y su diagrama fasorial<sup>9</sup>

NTSC emplea barrido entrelazado, que divide la pantalla en dos campos, las líneas horizontales pares e impares, el número de líneas de exploración horizontal de un campo es la mitad del total de 525 líneas de un cuadro completo, esto da por resultado 262.5 líneas horizontales para cada campo.

En un segundo pasan 60 campos, el tiempo que corresponde a un campo es 1/60 s o 16.666 ms, cada campo contiene 262.5 líneas, el tiempo por línea es 16.666 ms dividido entre 262.5 líneas, que es igual a 63.5 us, de ahí la frecuencia de exploración horizontal de 15.750 KHz.

Toda la información necesaria de una imagen a blanco y negro está contenida en la señal de luminancia, al restar la luminancia de las componentes RGB se obtienen la señales de croma en rojo ( $Cr_R-Y$ ) conocida como **V**, la señal de croma en azul ( $Cr_B-Y$ ) conocida como **U** y la señal de croma en verde ( $Cr_G-Y$ ).

<sup>9</sup> [http://www.des.udc.es/~adriana/Perifericos/T1.trabajos2006/Trabajo\\_3\\_Digitalizacion.pdf](http://www.des.udc.es/~adriana/Perifericos/T1.trabajos2006/Trabajo_3_Digitalizacion.pdf)

Para transmitir señales de TV, el sistema NTSC envía la señal de color sobre una subportadora empleando modulación en cuadratura, Figura 1.4. En amplitud se modula la saturación y en fase la tonalidad estableciendo un vector cuyo módulo es la saturación y el ángulo la tonalidad.

$$Y = (Cr_R - Y) \cos \omega_s t + (Cr_B - Y) \text{sen} \omega_s t \tag{Ecuación 1.1}$$

En el receptor se detecta: 
$$\begin{cases} Cr_R - Y = V \\ Cr_B - Y = U \\ Y \end{cases} \tag{Ecuación 1.2}$$

La componente de croma en verde  $Cr_G$  se obtiene empleando:

$$(Cr_G - Y) = -\frac{0.30}{0.59} (Cr_R - Y) - \frac{0.11}{0.59} (Cr_B - Y) \tag{Ecuación 1.3}$$

La obtención de  $(Cr_G - Y)$  se logra mediante la adecuada valoración y suma (conocido como "matriciado") de las componentes de croma en rojo y azul.

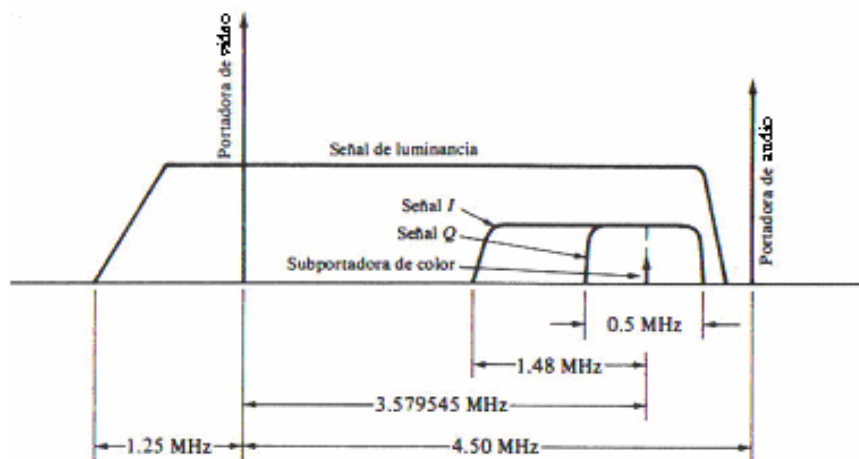


Figura 1.5 Diagrama espectral de la modulación NTSC<sup>10</sup>

<sup>10</sup> Imagen digitalizada de la referencia bibliográfica. STREMLER, Ferrel, SISTEMAS DE COMUNICACIÓN, Segunda Edición, Editorial Alfaomega, México 1982, p 657.



En el espacio de 6 MHz se localiza: la componente de video, la componente de audio, la subportadora de color y dos espacios que funcionan como banda de guarda entre canales adyacentes. En la Figura 1.5 y 1.6 se observa el diagrama espectral de la modulación de una señal de TV en el sistema NTSC.

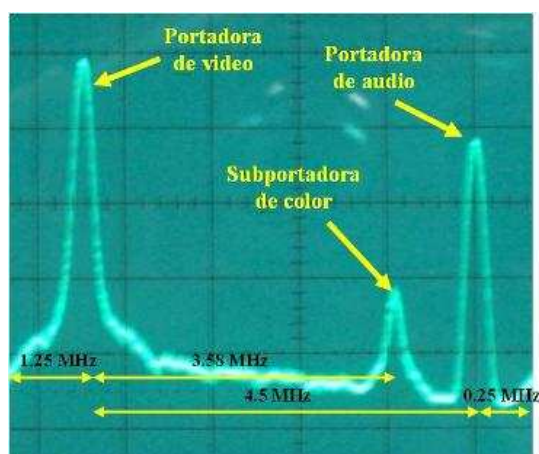


Figura 1.6 Diagrama en un analizador de espectros<sup>11</sup>

### **PAL (Phase Alternating Line)**

El sistema de línea alternada en fase (PAL) fue desarrollado por Telefunken, compañía alemana que intentaba mejorar la calidad de NTSC; sin embargo, incrementa el ancho de banda empleado. Actualmente se emplea en Europa y en países de medio oriente; en la Tabla 1.2 se tiene sus principales características.

### **SECAM (Séquential Couleur á Mémoire)**

El sistema de color secuencial con memoria (SECAM) fue desarrollado en Francia y es usado por Rusia, algunos países de África y países de Europa Oriental, es compatible con sistemas a blanco y negro. En la Tabla 1.3 se tiene las diferentes variaciones de SECAM y sus principales parámetros de operación.

<sup>11</sup> <http://www.cinit.org.mx/articulos.php>

SISTEMA	PAL B,G,H	PAL I	PAL D	PAL N	PAL M
Líneas	625	625	625	625	525
Campos	50	50	50	50	60
Frecuencia vertical	50 Hz	50 Hz	50 Hz	50 Hz	50 Hz
Frecuencia horizontal	15.625 KHz	15.625 KHz	15.625 KHz	15.62 KHz	15.750 KHz
Frecuencia de la portadora color	4.433618 MHz	4.433618 MHz	4.4336 MHz	3.582 MHz	3.57561MHz
Frecuencia de la portadora de audio	5.5 MHz	6.0 MHz	6.5 MHz	4.5 MHz	4.5 MHz
Ancho de banda de video	5.0 MHz	5.5 MHz	6.0 MHz	4.2 MHz	4.2 MHz

Tabla 1.2 Características de PAL

SISTEMA	SECAM B,G,H	SECAM D,K,K1,L
Líneas	625	625
Campos	50	50
Frecuencia vertical	50 Hz	50 Hz
Frecuencia horizontal	15.625 KHz	15.625 KHz
Frecuencia de la subportadora de color	4.4 MHz	4.4 MHz
Frecuencia de la portadora de audio	5.5 MHz	6.5 MHz
Ancho de banda de video	5.0 MHz	6.0 MHz

Tabla 1.3 Características de SECAM

Los sistemas NTSC, PAL y SECAM se han distribuido a lo largo de regiones a nivel mundial Figura 1.7.

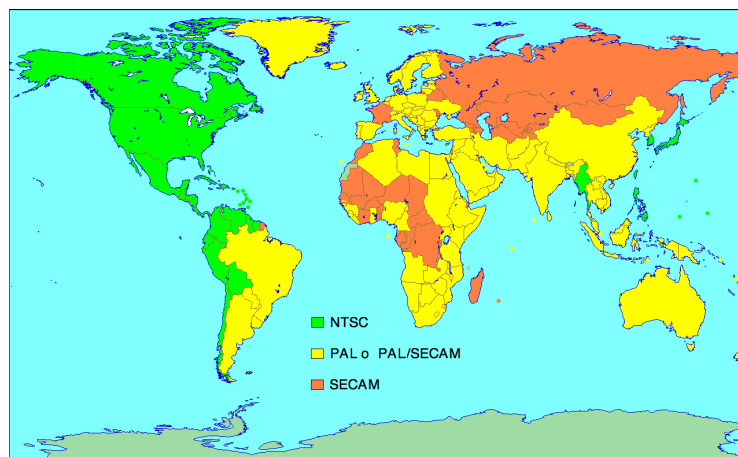


Figura 1.7 Distribución de los sistemas de televisión por regiones<sup>12</sup>

## 1.2.1 Generación de la imagen

La generación de imágenes en movimiento (video), comienza cuando la luz se proyecta por medios ópticos en un sensor de imagen dispuesto en el interior de la cámara de video. Las variaciones en la amplitud de las señales son proporcionales a la intensidad de luz acumulada sobre el sensor.

### 1.2.1.1 Sensor de imagen<sup>13</sup>

El sensor de imagen es un dispositivo que transforma la energía lumínica (fotones) en energía eléctrica<sup>14</sup>(electrones), consiste en una superficie sensible a la luz formando una matriz de puntos, cada punto aporta en la formación de la imagen. En esta parte se hace referencia la analogía entre los miles de nervios ópticos del ojo humano, frente a los limitados puntos contenidos en la matriz de un sensor de imagen y la sensibilidad a las frecuencias del rojo, verde y azul Figura 1.8.

<sup>12</sup> [http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo\\_3\\_Digitalizacion.pdf](http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo_3_Digitalizacion.pdf)

<sup>13</sup> [http://www.axis.com/files/brochure/bc\\_techguide\\_26556\\_es\\_0508\\_lo.pdf](http://www.axis.com/files/brochure/bc_techguide_26556_es_0508_lo.pdf)

<sup>14</sup> A estos dispositivos se los conoce como transductores, transforman un tipo de energía en otra.

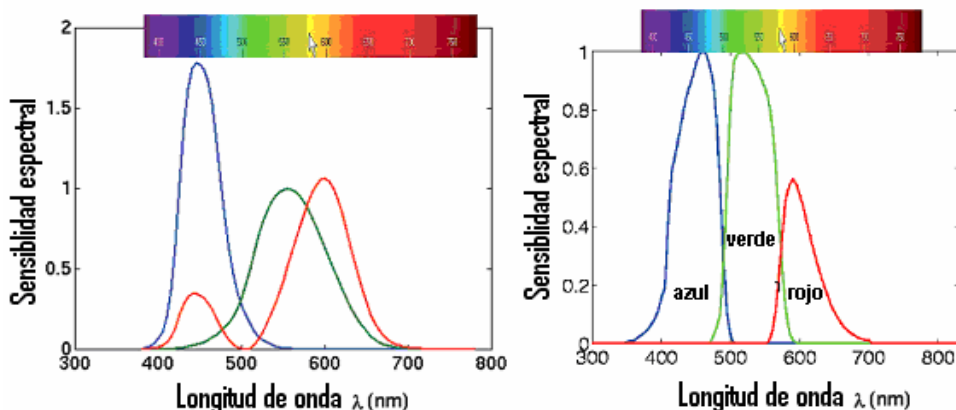


Figura 1.8 Sensibilidad del ojo humano (izquierda), sensibilidad de una cámara de video (derecha)<sup>15</sup>

En principio los sensores de imagen consistían en tubos al vacío, en su interior una superficie fotosensible y el generador de un haz de electrones encargado de barrer la imagen, esto hacía que las cámaras de video sean grandes y difíciles de portar. Podemos mencionar desarrollos como: el iconoscopio (antecesor de las cámaras de televisión de tubos), el orthicón y el vidicón Anexo A.2.

Actualmente, los nuevos sensores de imagen presentan dos tecnologías: el dispositivo de acoplamiento de carga (CCD, Charge Coupled Device) y el semiconductor complementario de óxido metálico (CMOS Complementary Metal Oxide Semiconductor).

### 1.2.1.2 Tecnología CCD

El CCD consiste en el sensor de imagen junto a un circuito integrado que realiza funciones de almacenamiento y exploración (scanning) de la imagen; consiste en un arreglo de celdas de dos dimensiones, cada celda acumula carga eléctrica proporcional a la luz que incide sobre ella, esta carga puede ser movida a una celda adyacente ya sea horizontal o verticalmente y puede transferir el contenido de filas, columnas o incluso toda la imagen conformada en la matriz para su procesamiento.

<sup>15</sup> <http://www.ugr.es/~jnieves/Textos/Tema2Psicofisica%20VisionColor%20I%20CONt%20Gelstalt.pdf>

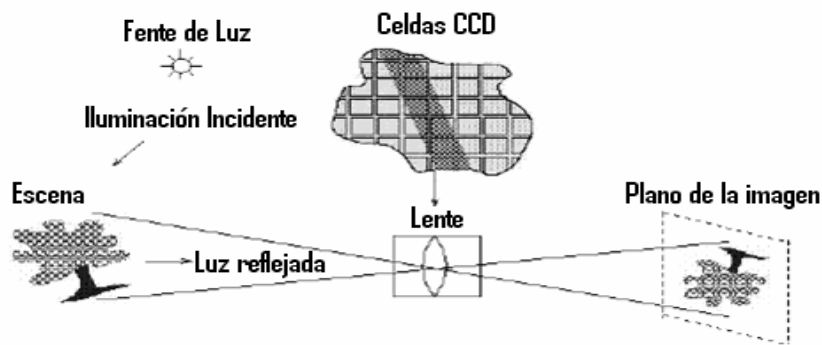


Figura 1.9 Proceso de captación de una imagen<sup>16</sup>

Los modos de almacenamiento y de transferencia del CCD permiten captar y almacenar la imagen. En la Figura 1.9 se muestra como la iluminación producida por la fuente incide sobre el objeto de la escena, en donde el lente (en este caso celdas CCD) “absorbe” la luz reflejada y la convierte en la señal de video.



Figura 1.10 Sensor CCD

La variación de luz obtenida pertenece a la parte activa de la señal de video, se agrega pulsos de sincronía al final de cada línea barrida y al inicio de una nueva línea horizontal. Los sensores CCD se fabrican usando una tecnología desarrollada específicamente para la industria de cámaras, poseen mayor sensibilidad a la luz que se traduce en mejores imágenes en situaciones de luz escasa; sin embargo, los sensores CCD son caros ya que están fabricados siguiendo un proceso no estandarizado y más complejo para ser incorporados en una cámara.

<sup>16</sup> <http://redescolar.ilce.edu.mx/redescolar/cursos/sepacomputo/pdiauvi.pdf>

### 1.2.1.3 Tecnología CMOS

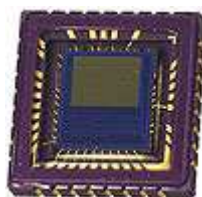


Figura 1.11 Sensor CMOS

Los sensores CMOS se basan en una tecnología estándar ampliamente utilizada en los chips de memoria Figura 1.11. Son inadecuados para cámaras donde se exige mayor calidad de la imagen. Los sensores CMOS proporcionan soluciones de cámaras más económicas y hacen posible la producción de cámaras de menor tamaño.

En un sensor de imagen, la sensibilidad es una característica muy importante y se refiere a la capacidad de captar imágenes en condiciones de poca luz. En cámaras a color la sensibilidad mínima normalmente está entre 3 y 0,5 lux<sup>17</sup>, mientras que en cámaras blanco y negro la sensibilidad mínima tiene valores normales en el entorno de 1 y 0,05 lux.

### 1.2.1.4 Resolución

Entre la resolución digital y analógica existen algunas diferencias, el video analógico está compuesto por imágenes que se forman por líneas horizontales, el video digital se compone por imágenes que se forman por cientos de píxeles.

Cada píxel puede representar uno o varios colores y dependiendo de la profundidad a la que se trabaje, se emplean n bits para representarlos, por ejemplo: con 8 bits se tiene 256 colores, con 16 bits se obtienen 65.536 colores, con 24 bits se tienen 16.777.216 colores y

---

<sup>17</sup> El lux es la unidad de la intensidad de luminosa  
[www.fotomundo.com/tecnic/varios/terminolo6.shtml](http://www.fotomundo.com/tecnic/varios/terminolo6.shtml)

con 32 bits se obtienen 429.4967.296 colores.

### Resolución NTSC y PAL<sup>18</sup>

Hace referencia al tamaño del cuadro y las dimensiones físicas que posee el video, cuando el video es digitalizado, la cantidad máxima de píxeles a obtenerse dependerá del número de líneas de TV disponibles. En NTSC, el tamaño máximo de las imágenes digitalizadas es de 704x480 píxeles; en PAL, el tamaño es de 704x576 píxeles, se ha estandarizado a CIF<sup>19</sup> como un formato de video digital, así se tiene:

FORMATO	RESOLUCIÓN NTSC/PAL(Píxeles)
QCIF	176 x 120/144
CIF	352 x 240/288
2CIF	704 x 240/288
4CIF	704 x 480/576

Tabla 1.4 Formatos velocidad y aplicaciones

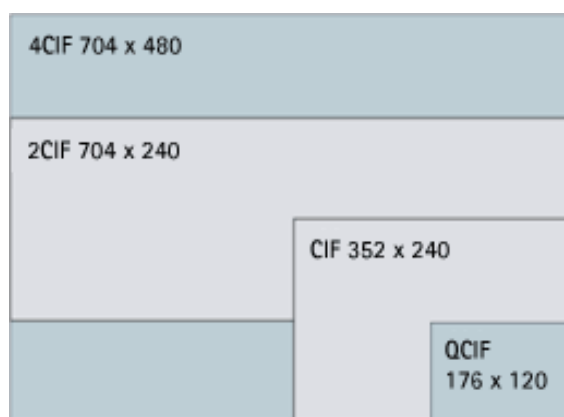


Figura 1.12 Resolución NTSC

<sup>18</sup> [http://www.axis.com/files/brochure/bc\\_techguide\\_26556\\_es\\_0508\\_lo.pdf](http://www.axis.com/files/brochure/bc_techguide_26556_es_0508_lo.pdf)

<sup>19</sup> (CIF, Common Intermediate Format) formato de video utilizado en sistemas de video conferencia que fácilmente soporta señales NTSC y PAL, es parte del estándar de videoconferencia ITU H.261. <http://www.alegs.com.ar/Dic/cif.php>  
<http://es.wikipedia.org/wiki/CIF>

La resolución 2CIF es 704x240 (NTSC) ó 704x288 (PAL) píxeles, lo que significa dividir el número de líneas horizontales para dos.

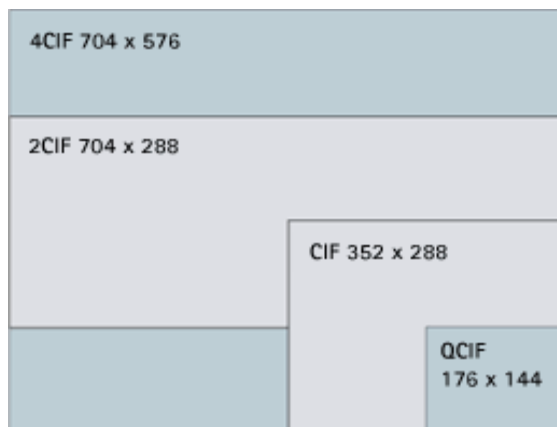


Figura 1.13 Resolución PAL.

### Resolución VGA

Las nuevas cámaras digitales han introducido mejor resolución empleando el sistema de exposición gráficos para PC (VGA, Video Graphics Array) desarrollado originalmente por IBM, la resolución VGA es adecuada para las cámaras digitales ya que el video en la mayoría de casos se mostrará en pantallas de ordenador que emplean resoluciones VGA o múltiplos de VGA (640 x 480), QVGA (320 x 240), SVGA (1.024x768 píxeles) y de 1.280x960 píxeles cuatro veces VGA que ofrecen una resolución mega píxel.



Figura 1.14 Resoluciones VGA



Mientras más alta sea la resolución, más detalles pueden observarse en una imagen; la resolución máxima en NTSC y PAL digitalizada es de 400.000 píxeles (704x576 = 405.504) equivale a 0,4 mega píxeles.

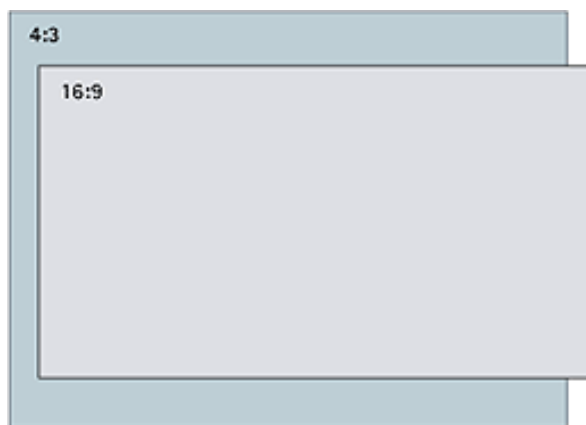


Figura 1.15 Resolución mega píxel

Un formato mega píxel es 1.280x1.024, ofrece una resolución de 1,3 mega píxeles, tres veces la resolución de las cámaras analógicas; se dispone además, de cámaras de 2 a 6 mega píxeles y se esperan resoluciones superiores en el futuro.

### 1.2.2 Medios ópticos<sup>20</sup>

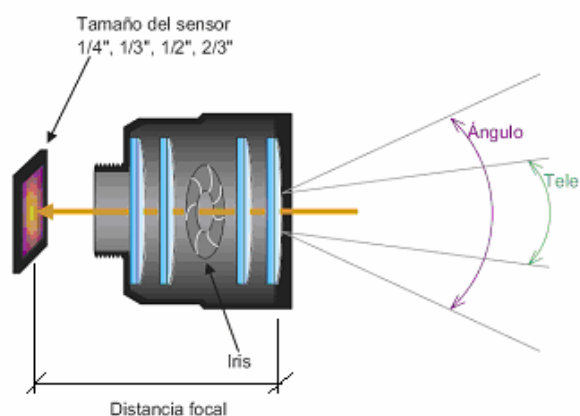


Figura 1.16 Longitud focal, iris, sensor de imagen

<sup>20</sup> [http://www.axis.com/files/brochure/bc\\_techguide\\_26556\\_es\\_0508\\_lo.pdf](http://www.axis.com/files/brochure/bc_techguide_26556_es_0508_lo.pdf)

El empleo de medios ópticos es de vital importancia y parte complementaria de las cámaras de video por lo que es necesario definir términos importantes.

**Distancia focal:** es la distancia en milímetros del lente a la superficie del sensor de imagen, cuanto menor sea la distancia más ancho será el ángulo de visión y cuanto mayor sea la distancia, más estrecho será el ángulo de visión.

**Iris:** controla la cantidad de luz que alcanza al sensor de imagen, de acuerdo a la su apertura.

**Objetivos:** son aditamentos ópticos que permiten enfocar de mejor manera la imagen en el sensor. Normalmente los objetivos son activos, es decir poseen un sistema de control automático del iris.



Figura 1.18 Lente fija



Figura 1.19 Lente varifocal

El objetivo fijo mantiene la distancia focal; el objetivo varifocal permite variar manualmente la distancia focal; el objetivo zoom motorizado varia la distancia focal mediante un sistema automático, acercándonos o alejándonos convenientemente del objeto a captar. Figura 1.7 y 1.8.

### 1.2.3 Barrido entrelazado frente al barrido progresivo<sup>21</sup>

En un video, la sensación de movimiento se logra por imágenes fijas continuas, el número de imágenes por segundo que pasan, debe ser el adecuado para obtener movimiento, la forma de presentar las imágenes en una pantalla o monitor, manteniendo un número de imágenes por segundo emplea dos métodos, el barrido entrelazado y el barrido progresivo.

#### Barrido entrelazado



Figura 1.20 Barrido entrelazado

Un monitor de televisión está constituido por líneas horizontales distribuidas a lo ancho de la pantalla, el barrido entrelazado divide las líneas en campos, una imagen se forma por dos campos: el campo de líneas pares y el campo de líneas impares, las cuales se

<sup>21</sup> [http://www.axis.com/files/brochure/bc\\_techguide\\_26556\\_es\\_0508\\_lo.pdf](http://www.axis.com/files/brochure/bc_techguide_26556_es_0508_lo.pdf)

actualizan a 30 o 25 imágenes por segundo de acuerdo al sistema empleando (NTSC o PAL). Figura 1.20.

### Barrido progresivo

Actualmente, la tecnología de la pantalla ha cambiado y con la llegada de las pantallas de cristal líquido (LCD) los monitores basados en transistores de película delgada (TFT<sup>22</sup>), se ha establecido un método alternativo para barrer la imagen en la pantalla, conocido como barrido progresivo.

El barrido progresivo a diferencia del entrelazado, escanea la imagen entera, línea a línea, las imágenes captadas no se dividen en campos, como ocurre con el barrido entrelazado.



Figura 1.21 Esquema del barrido progresivo

## 1.2.4 Dispositivos de video

### Tubo de rayos catódicos

El tubo de imagen de TV está formado básicamente por una ampolla de cristal, a la que se le ha practicado el vacío; la parte frontal del tubo de imagen es una pantalla de cristal transparente cuya superficie interior está recubierta por un material luminiscente.

---

<sup>22</sup> TFT transistor de película delgada, la tecnología TFT brinda la mejor resolución y calidad de imagen.

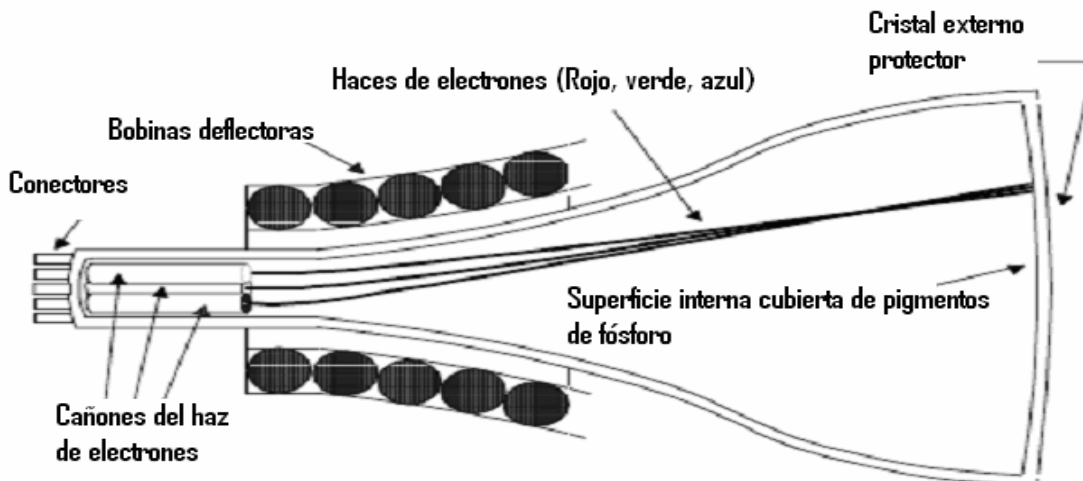


Figura 1.22 Esquema del tubo de rayos catódicos<sup>23</sup>

En el cuello del tubo está situado un cañón electrónico que inyecta un fino haz de electrones que al impactar sobre la pantalla recubierta con material luminiscente, emite energía en forma de luz, en el caso de un monitor a color se necesitan tres haces de electrones y tres partículas de fósforo para representar un punto. Figura 1.22.

## LCD<sup>24</sup>

Basa su funcionamiento en la propiedad de ciertos materiales de permanecer en estado isotrópico (estado híbrido, sólido/líquido), el material (cristal líquido) se sitúa entre dos capas conductoras transparentes y orienta la luz a su paso mediante impulsos eléctricos; los rayos atravesando el material representa un píxel encendido, los rayos que no atraviesan representan un píxel apagado.

Las pantallas se clasifican dependiendo de la capacidad de generar luz o reflejar luz de una fuente exterior

- **Matriz pasiva LCD:** disposición de electrodos en forma de enrejado, la luz se

<sup>23</sup> [www.mtas.es/insht/ntp/ntp\\_678.htm](http://www.mtas.es/insht/ntp/ntp_678.htm)

<sup>24</sup> <http://www.plasma.com/>

genera globalmente y la matriz modifica la luz.

- **Matriz activa (TFT, Thin Film Transistor):** matriz de transistores foto-emisores FET, cada célula tiene luz propia y permite obtener mejor resolución y contraste de la imagen. Figura 1.23.

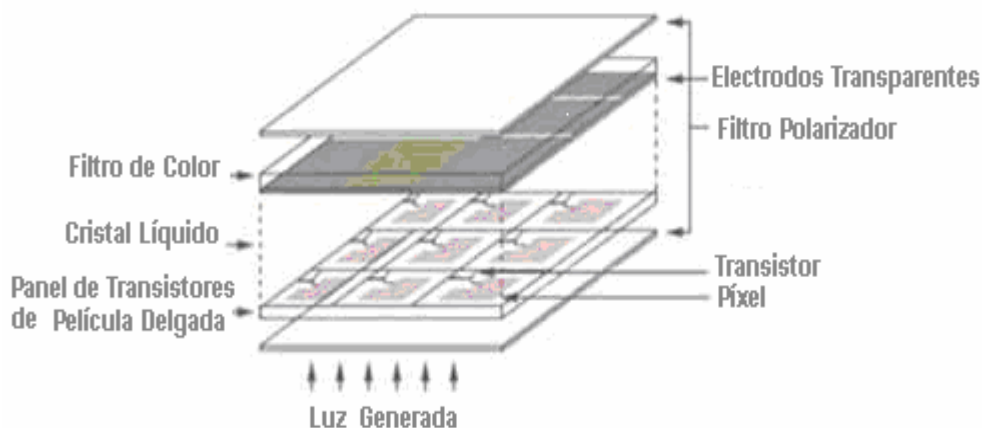


Figura 1.23 Estructura del **TFT-LCD** (Pantalla de cristal líquido-Transistor de película delgada)<sup>25</sup>

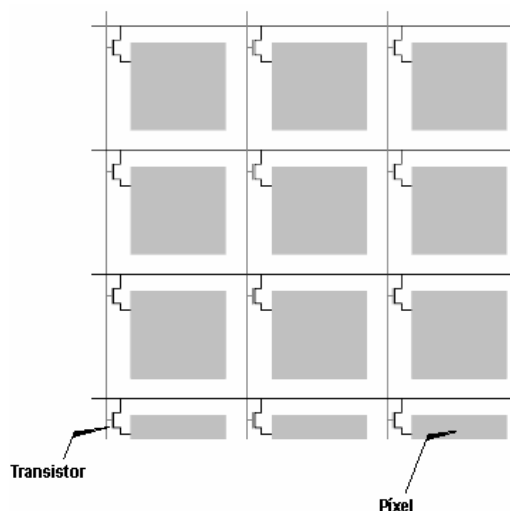


Figura1.24 Diagrama ampliado del arreglo de transistores de efecto de campo (FET Field Effect Transistor)<sup>26</sup>

<sup>25</sup> [http://www.plasma.com/classroom/what\\_is\\_tft\\_lcd.htm](http://www.plasma.com/classroom/what_is_tft_lcd.htm)

<sup>26</sup> [http://upload.wikimedia.org/wikipedia/commons/b/b1/TFT\\_Matrix.svg](http://upload.wikimedia.org/wikipedia/commons/b/b1/TFT_Matrix.svg)

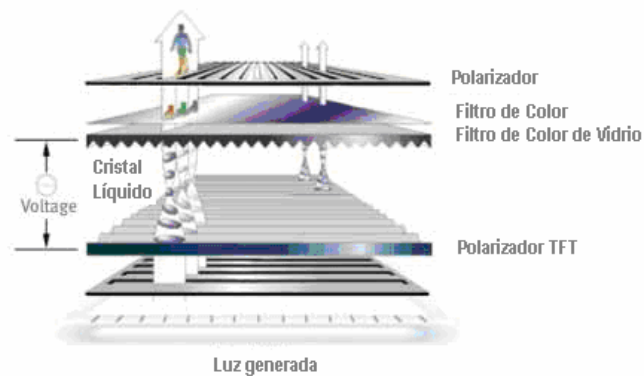


Figura 1.25 Diagrama ampliado de TFT-LCD<sup>27</sup>

## Plasma

Emplean una mezcla de xenón y argón dentro de cada píxel, que está compuesto a su vez por tres celdas separadas, cada una con fósforo de diferentes colores: rojo, verde y azul. Estos colores se mezclan para crear el color final del píxel. Mediante electrodos se aplican pulsos de corriente que fluyen creando diferentes combinaciones de rojo, verde y azul y dan forma a la imagen.

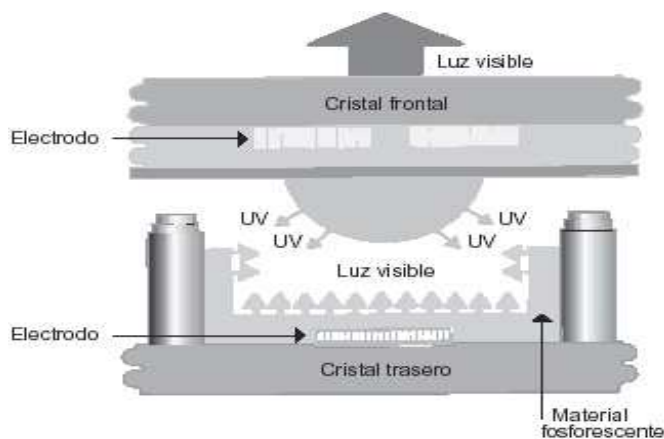


Figura 1.26 Esquema de funcionamiento de una celda de pantalla de plasma<sup>28</sup>

<sup>27</sup> [http://www.plasma.com/classroom/what\\_is\\_tft\\_lcd.htm](http://www.plasma.com/classroom/what_is_tft_lcd.htm)

<sup>28</sup> [www.mtas.es/insht/ntp/ntp\\_678.htm](http://www.mtas.es/insht/ntp/ntp_678.htm)

## 1.2.5 Tipos de señales de video

### Video compuesto

La imagen se trasmite en una sola señal, emplea un conector amarillo RCA presente en casi todos los reproductores de video y video cámaras, la señal de luminancia y crominancia están codificadas en una única señal.

### Súper video (S-video)

Utiliza el conector de S-VHS o Din de 4 pines, la señal posee una calidad superior a la señal de video compuesto, su diferencia radica en que las señales de luminancia (Y) y cromina (Cr) van separadas en la Tabla 1.5, se tiene la asignación de pines y en la Figura 1.28, el diagrama del conector de súper video.

PIN	NOMBRE
1	GND Luminancia
2	GND Crominancia
3	Y Luminancia
4	Cr Crominancia

Tabla 1.5 Diagrama de pines del conector Súper Video.

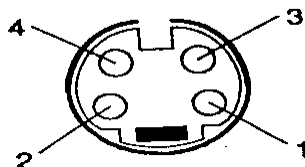


Figura 1.27 Conector S Video<sup>29</sup>

<sup>29</sup> <http://www.electronicafacil.net/docs/docesp4.html>



## Señal Y, UV

Emplea las señales de crominancia en rojo  $Cr_R$  (**V**), crominancia en azul  $Cr_B$  (**U**) y la señal de luminancia (**Y**). La señal de croma en verde se calcula por fórmulas matemáticas, los equipos emplean un procedimiento llamado matriciado.

## RGB (Red, Green, Blue)

Esta señal se compone por tres señales que contienen información de los tres colores, RGB (rojo, verde, y azul) que proporciona una mejor calidad de imagen que la señal de súper video y video compuesto.

### 1.3 Digitalización de Video

La digitalización de la información (voz, datos y video) ha dado lugar a la multimedia, el video su principal componente se ha desarrollado de tal manera que aplicaciones como: video conferencias, educación a distancia, video-vigilancia se han popularizado en nuestro medio.

El proceso de digitalización de la señal de vídeo consiste en el muestreo, cuantificación y codificación de la señal, la Unión Internacional de Telecomunicaciones (UIT) define una norma de video digital que no aplica ningún tipo de compresión, razón por la cual los formatos derivados de esta, se han mantenido como estándares digitales para el intercambio de información entre equipos y terminales en estudios de grabación y producción de video.

### 1.3.1 Recomendación ITU 601

El primer paso para la digitalización del video, es la obtención de las señales de luminancia y crominancia, en base a las señales RGB, posteriormente se muestrea la señal de luminancia (Y) a 13.5 MHz y la crominancia a 6.75MHz.

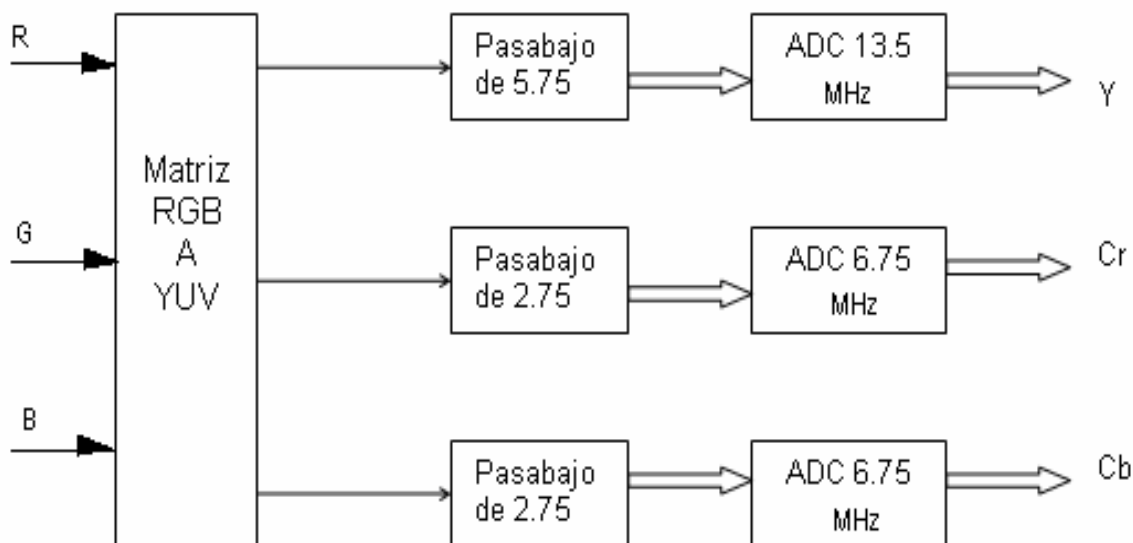


Figura 1.28 Esquema de obtención Y, Cr y Cb en base a RGB

La luminancia y crominancia se relacionan con RGB, mediante las siguientes relaciones:

$$Y = 0,299R + 0.58G + 0.114B \tag{Ecuación 1.4}$$

$$Cr = 0.713(R - Y) \tag{Ecuación 1.5}$$

$$Cb = 0.564(B - Y) \tag{Ecuación 1.6}$$

Si se emplea en el muestreo 8 bits.

$$\begin{aligned} Y &= 8 * 13.5 = 108Mbps \\ Cr &= 8 * 6.75 = 54Mbps \\ Cb &= 8 * 6.75 = 54Mbps \end{aligned} \tag{Ecuación 1.7}$$

El flujo de datos que se obtiene es de 216 Mbps empleando el muestreo de las componentes de luminancia y crominancia, se han normalizado varios formatos de acuerdo a una estructura (Y,Cr,Cb), así se tiene:

- 4:4:4
- 4:2:2
- 4:2:0
- 4:1:1
- 4:0:0

### Digitalización 4:4:4

Por cada cuatro muestras de luminancia existen 4 muestras de croma en rojo (**V**) y 4 de croma en azul (**U**) respectivamente Figura 1.29, cada fila está formada por cuadrados que representan los píxeles y contienen una muestra de luminancia y dos de crominancia, una en rojo y una en azul.

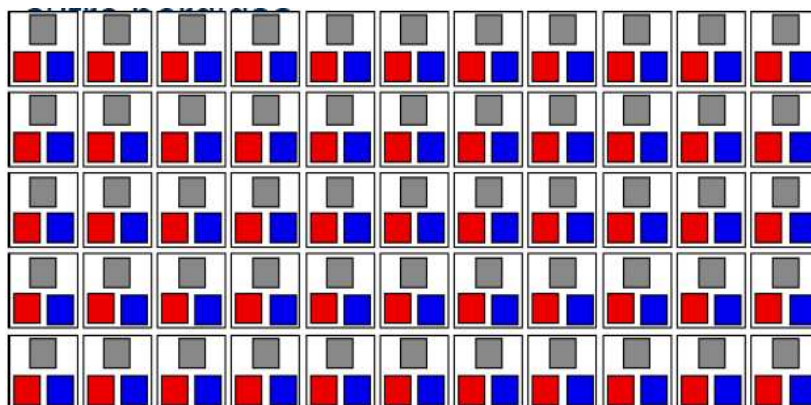


Figura 1.29 Representación del muestreo 4:4:4<sup>30</sup>

<sup>30</sup> [http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo\\_3\\_Digitalizacion.pdf](http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo_3_Digitalizacion.pdf)

Este método guarda toda la información y la imagen no sufre pérdidas.

**Digitalización 4:2:2**

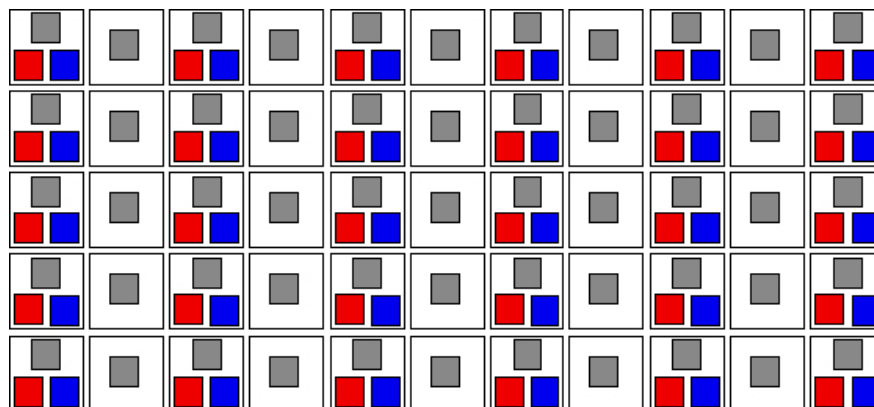


Figura 1.30 . Representación del muestreo 4:2:2<sup>31</sup>

Este esquema de digitalización posee una calidad de video de estudio y aprovecha la deficiencia del ojo humano que es más sensible al brillo que al color, por lo que reduce la información cromática a la mitad; el color tiene la mitad de resolución (en horizontal) que el brillo, por cada cuatro muestras de luminancia Y, hay dos de crominancia. Figura 1.30.

**Digitalización 4:2:0**

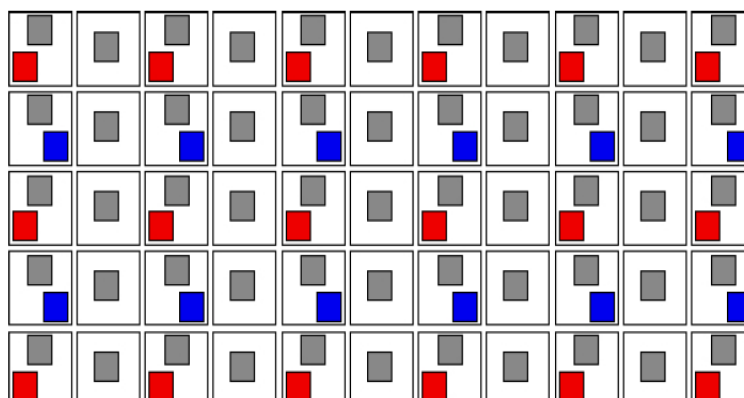


Figura 1.31 Representación del muestreo 4:2:0<sup>32</sup>

31 [http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo\\_3\\_Digitalizacion.pdf](http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo_3_Digitalizacion.pdf)

Se explora los colores en líneas alternas, de manera que se muestrea una línea en rojo y la siguiente el azul, refiérase a la Figura 1.31. La información de color se reduce a la cuarta parte en un factor de dos en ambas direcciones: horizontal y vertical.

### Digitalización 4:1:1

En este esquema de digitalización, las muestras de crominancia (**U y V**) son tomadas una cada cuatro muestras horizontales de luminancia (Y) Figura 1.32.

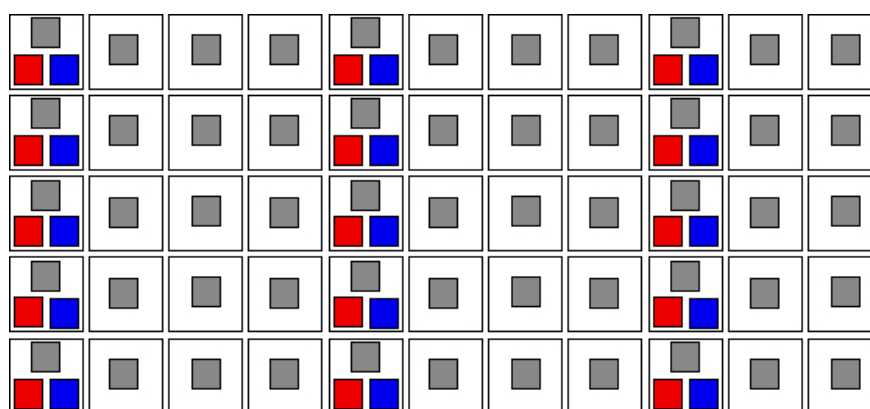


Figura 1.32 Muestreo 4:1:1

### Digitalización 4:0:0

El formato elimina por completo la información de color, la imagen queda en blanco y negro.

#### 1.3.2 Necesidad de comprimir

Para la transmisión de video digital se emplea un proceso de reducción de la tasa de bits (compresión) sin perder la calidad de las imágenes. Sin compresión no tendría sentido poner imágenes, audio o video en Internet, la calidad de las comunicaciones celulares no

<sup>32</sup> [http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo\\_3\\_Digitalizacion.pdf](http://www.des.udc.es/~adriana/Perifericos/Trabajos2006/Trabajo_3_Digitalizacion.pdf)

sería la misma y desde luego la TV digital no sería posible.

La información digital debe pasar necesariamente por un proceso de compresión para obtener información que no requiera canales de comunicación de gran capacidad para transmitirse.

### 1.3.2.1 Principios de compresión<sup>33</sup>

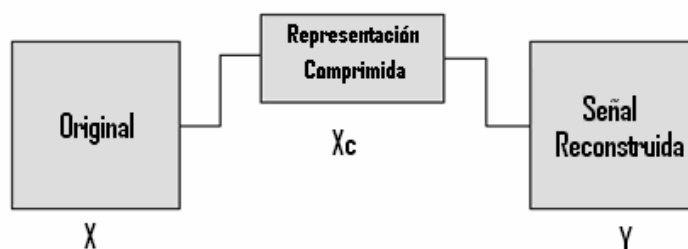


Figura 1.33 Esquema básico de compresión y descompresión

Un algoritmo de compresión toma una entrada  $X$  y genera una representación  $X_c$ , que necesita menos bits, el algoritmo de reconstrucción toma  $X_c$  y genera la señal original  $X$ .

La compresión de video se consigue explotando las semejanzas o redundancias contenidas en la información de video, información que resulta irrelevante al ojo humano. Un algoritmo de compresión debería transmitir solamente información relevante para el ojo humano y a partir de esta, reconstruir la señal completa; sin embargo, esto requiere de períodos de tiempo que en aplicaciones en tiempo real son inconcebibles.

Los algoritmos de compresión se clasifican en dos grupos:

---

<sup>33</sup> Referencia bibliográfica  
<http://trappist.elis.ugent.be/~wheirman/compression/>  
<http://www.cs.sfu.ca>  
<http://ict.ewi.tudelft.nl/index.php?Itemid=124>

- **Algoritmos sin pérdida<sup>34</sup>:** son aquellos en los que la entrada al codificador X y la salida del decodificador Y coinciden. Se emplean en aplicaciones donde no se permiten diferencias entre los datos originales y los reconstruidos.
- **Algoritmos con pérdidas:** son aquellos que transmiten solamente información importante para el ojo humano, el decodificador debe calcular la redundancia y recuperar la información, los datos originales no pueden ser recuperados exactamente; estos datos toleran pérdidas en la compresión y pueden ser la voz o el video.

La efectividad de la técnica de compresión, viene dada por la relación del tamaño de la información de la imagen original dividido para el tamaño que ocupa después de la compresión, a mayor valor, menos ancho de banda se emplea.

### 1.3.2.2 Redundancia

La redundancia es la información que no es perceptible para el ojo humano; se tiene tres tipos de redundancia: espacial, temporal y psicovisual.

Los imágenes consecutivas de una secuencia de video pueden exhibir redundancia temporal cuando, los objetos en una escena mantienen su posición entre una imagen y la siguiente, la redundancia espacial esta dada por la similitud de píxeles cercanos y la redundancia psicovisual tiene que ver con las limitaciones físicas del ojo humano, para fijarse en los detalles de los objetos en cambios rápidos escena.

---

<sup>34</sup> Los algoritmos sin pérdida se los conoce como lossless y los algoritmos con pérdida se los conoce como lossy

### 1.3.2.3 Técnicas de compresión

Llamadas también tecnologías posibilitadoras<sup>35</sup>, consisten en algoritmos estandarizados internacionalmente, que permiten comprimir secuencias de video, los principales estándares son: JPEG, Motion JPEG, H.261, MJPEG, MPEG-1. MPEG-2, H.263, MPEG-4, H.264/AVC.

Los estándares son adoptados por fabricantes de software y hardware, la adopción de los estándares asegura la funcionalidad de equipos y sistemas de video.

## JPEG

JPEG realiza la compresión de imágenes individuales, aprovecha la redundancia espacial es decir la similitud entre píxeles vecinos, JPEG explota esta característica dividiendo la imagen en bloques de 8x8 píxeles y aplicando la transformada discreta del coseno que representa la imagen en el dominio de frecuencia, eliminando los componentes menor frecuencia, mientras que los componentes de mayor frecuencia se codifican empleando codificación Huffman, en una imagen a color se realiza primero una conversión de la imagen RGB al espacio luminancia y crominancia en donde la capacidad de percepción del ojo humano se explota de mejor forma.

La compresión de imágenes en movimiento es un caso particular de la compresión de una imagen fija y explota la redundancia entre las imágenes, esta forma de compresión se le conoce como codificación *interframe*. Las diferencias entre imágenes sucesivas pueden deberse principalmente a los objetos individuales que se mueven en donde la mayor parte de la imagen es idéntica. Las técnicas de codificación interframe eliminan esta redundancia basándose en la información de una imagen previa para obtener la siguiente;

---

<sup>35</sup> El término también empleado en diferentes bibliografías es *enabling technologies*



sin embargo, cuando hay un cambio de escena se debe tratar una nueva imagen y se recurre al uso de codificación de imágenes simples.

## **Motion JPEG**

Comprime imágenes individuales y no aprovecha la redundancia temporal, lo que hace es encadenar una serie de imágenes comprimidas con JPEG, una de las ventajas de Motion JPEG, es que se puede acceder a cualquier imagen de una secuencia de video lo que resulta útil en la edición y producción del video.

## **H.261**

Es un estándar de compresión empleado en la video-conferencia, vigilancia, monitoreo, telemedicina y otros servicios audiovisuales, trabaja sobre la red digital de servicios integrados (RDSI); su calidad no es lo más destacable, consume un ancho de banda múltiplo de 64 kbps y cuenta con un mecanismo para controlar la calidad en función del movimiento de la secuencia, cuanto mayor sea el movimiento de las secuencias, menor será su calidad, H.261 soporta dos resoluciones:

- CIF (Common Interchange Format, 352x288 pixels)
- QCIF (Quarter Common Interchange Format, 176x144 pixels).

El H.261 es un sistema que aprovecha tanto la redundancia espacial como la redundancia temporal y ha sido la base para los formatos MPEG que se han convertido en estándares ampliamente empleados.

## MPEG

Grupo de expertos de fotografía en movimiento MPEG (Motion Picture Expert Group) es uno de los más populares estándares de compresión de video. Para obtener imágenes en movimiento, se captan imágenes en el tiempo, las imágenes intermedias (imágenes-I) serán calculadas a partir de estas, que se intercalan con imágenes B.

- **Imágenes-I:** para decodificar una imagen de este tipo no hace falta otras imágenes, hay una cada 10 o 15 imágenes, se las llama imágenes clave y siempre son un punto de acceso al flujo de bits de video.
- **Imágenes-P:** codificada como predicción de la imagen I anterior usa un mecanismo de compensación de movimiento. El proceso de codificación explota tanto la redundancia espacial como la temporal.
- **Imágenes-B:** se codifica utilizando una imagen I ó P anterior y la imagen I ó P siguiente sirve como referencia para la compensación y estimación de movimiento, estas imágenes consiguen los niveles de compresión más elevados y por tanto son las más pequeñas.

Existen otro tipo de imágenes llamadas imágenes intraframe (o imágenes D) que son de las mismas características que las imágenes I, pero con menos resolución. Se usan en aplicaciones que no necesitan gran calidad.

### Grupos de imágenes

Se forman grupos de imágenes consecutivas que empiezan con imágenes I y contienen imágenes P y B mezcladas, el grupo de imágenes (GOP, Group Of Pictures) está formado por una imagen I, luego una o varias imágenes P y varias imágenes B intercaladas.

Usualmente en un GOP encontramos 12 imágenes, las cuales el codificador puede cambiar dinámicamente. En el caso de detectar un cambio de escena forzará una imagen I y dará inicio a un nuevo GOP.

En la siguiente tabla se ve el orden de las imágenes, junto con el orden en el que son creados y enviados:

Orden de Visualización												
I	B	B	P	B	B	P	B	B	P	B	B	I
1	2	3	4	5	6	7	8	9	10	11	12	13

Orden de transmisión												
I	P	B	B	P	B	B	P	B	B	I	B	B
1	2	3	4	5	6	7	8	9	10	11	12	13

Figura 1.34 Orden de fotogramas en un GOP

El principio básico de MPEG es comparar entre dos imágenes para que puedan ser transmitidas a través de la red y usar la primera imagen como imagen de referencia (Imagen-I), enviando tan solo partes de las siguientes imágenes (Imágenes B y P). En el lado del receptor se reconstruirá todas las imágenes basándose en la imagen de referencia y en los datos diferentes contenidos en las imágenes B y P.

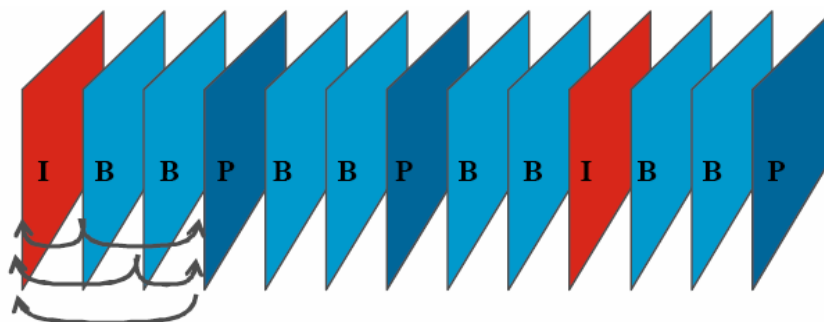


Figura 1.35 Secuencias de imágenes<sup>36</sup>

<sup>36</sup> [www.axis.com/es/documentacion/compresion\\_video\\_es.pdf](http://www.axis.com/es/documentacion/compresion_video_es.pdf)

Se debe tener en cuenta que una imagen-P, puede solo referenciar a una imagen I o P anterior, mientras que una imagen-B puede referenciar tanto a una imagen I o P anterior y posterior.

Aunque con mayor complejidad el resultado de aplicar la compresión de video MPEG es la cantidad de datos transmitidos a través de la red que relación a Motion JPEG es menor. MPEG es de hecho bastante más complejo que lo indicado anteriormente, e incluye parámetros como la predicción de movimiento en una escena y la identificación de objetos que son técnicas o herramientas de MPEG, se definen tres estándares:

### **MPEG-1**

El estándar MPEG-1 está dirigido a aplicaciones de almacenamiento de video digital en CD's. Por esta circunstancia la mayoría de los codificadores y decodificadores MPEG-1 precisan un ancho de banda de aproximadamente 1.5 Mbps a resolución CIF. MPEG-1 mantiene el consumo de ancho de banda relativamente constante aunque varíe la calidad de la imagen que es comparable a la calidad del video VHS. El número de imágenes por segundo en MPEG-1 es de 25 en PAL y 30 en NTSC.

### **MPEG-2**

MPEG-2 fue diseñado para video digital de alta calidad, TV digital de alta definición (HDTV), difusión de video digital (Digital Video Broadcasting, DVB) y televisión por cable. MPEG-2 se centró en ampliar la técnica de compresión MPEG-1 para cubrir imágenes más grandes y de mayor calidad, pero con un nivel de compresión menor y un consumo de ancho de banda mayor, proporciona herramientas adicionales para mejorar la calidad del video, con lo que se produce imágenes de muy alta calidad cuando lo comparamos con otras tecnologías de compresión.

## **MPEG-4**

El estándar MPEG-4 es uno de los desarrollos principales de MPEG-2, posee perfiles que se adaptan a diferentes aplicaciones.

Muchos de estos perfiles representan una excelente opción para ser implementados en redes inalámbricas, perfiles como: 'Escalable Simple', 'Tiempo Real Avanzado', 'Simple Avanzado' pueden representar una mejor elección respecto al Simple, en términos de desempeño y adaptabilidad a las características variantes del ambiente inalámbrico; sin embargo, siempre el mayor desempeño de un algoritmo de compresión viene asociado a un aumento en la complejidad.

El perfil Simple, a pesar de ser el perfil más básico, cuenta con las características necesarias para ser utilizado en aplicaciones inalámbricas y dispositivos móviles; su baja complejidad lo convierte en la elección predilecta de la industria.

### **MPEG-4 short header y long header**

El soporte de MPEG-4 se lo denomina "MPEG-4 long header", que indica el empleo herramientas de compresión de MPEG-4.

MPEG-4 short header no aprovecha ninguna de las herramientas adicionales especificadas en el estándar MPEG-4; está solo especificado para asegurar compatibilidad con equipos antiguos que emplean la recomendación H.263; de hecho no es más que un transmisor de video H.263, encapsulado con cabeceras MPEG-4.

### **MPEG-4 Visual**

Cuando se habla de MPEG-4, generalmente se refiere a MPEG-4 Parte 2 el cual fue especialmente desarrollado para tecnología multimedia; MPEG-4 introduce un nuevo

concepto en codificación de información visual: codificación basada en objetos, en lugar de codificación basada en imágenes, MPEG-4 Visual permite la codificación individual de los diferentes objetos que conforman la escena.

MPEG-4 Visual no fue creado específicamente para aplicaciones de streaming, pero se adecua extremadamente bien a este tipo de aplicaciones. Otra característica del estándar es que no dice nada sobre como deben llegar los streams al usuario final, así que estos pueden llegar a través de diferentes redes y tecnologías. Esto hace que MPEG-4 sea el estándar ideal por su tecnología, para implementar aplicaciones multimedia sobre canales inalámbricos (PCS, GSM y WLAN)<sup>37</sup>.

### **MPEG-4 Parte 10 (AVC, Control de Video Avanzado)**

MPEG-4 AVC, al que también se refiere como H.264, es un desarrollo posterior en el que se tiene un conjunto completamente nuevo de herramientas, que incorpora técnicas más avanzadas de compresión que reducen aun más el consumo de ancho de banda.

El H.264 de la ITU-T y el MPEG-4 Parte 10 de la ISO son mantenidos conjuntamente para que sean idénticos.

Proporciona suficiente flexibilidad para permitir al estándar ser aplicado en una gran variedad de aplicaciones.

### **Tasa de bits constante (CBR, Constant Bit Rate) y tasa de bits variable (VBR Variable Bit Rate)**

Otro aspecto importante de MPEG es el modo en el que se usa el ancho de banda

---

<sup>37</sup> PCS Sistema de comunicaciones personales, GSM Sistema de comunicaciones Globales Móviles, WLAN red local Inalámbricas.

disponible. En la mayoría de los sistemas MPEG es posible seleccionar la cantidad de bits por segundo y emplear el modo CBR (constante) o VBR (variable). La selección óptima depende de la aplicación y de la infraestructura de red disponible.

El ancho de banda disponible es el único inconveniente; el modo preferido es normalmente CBR dado que este modo consume un ancho de banda constante en la transmisión.

## **H.263**

Estándar de compresión de video para aplicaciones sobre redes móviles y la red de telefonía pública. Para obtener un considerable nivel de compresión se utiliza los tipos de imágenes-I, imágenes-P, imágenes-B empleando un sistema de corrección de errores que proporciona velocidades inferiores a 64 kbps.

Es empleado en video-conferencias, videotelefonía y video por Internet, especialmente por sus bajas tasas de transmisión con buenos niveles de calidad.

### **1.3.3 Formatos de video digital y reproductores**

Los sistemas digitales de video, la televisión de alta definición, la televisión por Internet, ha dado pie a que se emplee diferentes tipos de formatos y codecs.

Las arquitecturas para multimedia incluyen distintos módulos de software que permiten la creación, almacenamiento, distribución y reproducción de medios digitales, cada arquitectura tiene sus formatos de archivos específicos, para el almacenamiento de la información y de módulos para visualizar distintos tipos de formato. Algunas están optimizadas para determinadas plataformas de hardware, hay las que están orientadas a la distribución por la Web y otras a los soportes CD/DVD. No existe una arquitectura ideal, y la elección de una u otra dependerá del tipo de aplicación a desarrollar. A continuación

se detalla algunas arquitecturas:

## QuickTime (QT)

Es la arquitectura de Apple para la gestión de datos multimedia, es empleado en aplicaciones informáticas, cámaras digitales, CD's de audio con suplementos multimedia y en la distribución de contenidos multimedia a través de la Web.



Figura 1.36 Reproductor Quick Time<sup>38</sup>

## Características

**Flujo de datos (streaming)**, consiste en el envío de datos multimedia a través de una red para su reproducción en tiempo real a medida que van llegando al cliente. El origen de los datos puede ser una señal en directo (cámara de video, emisora de radio) o un archivo pregrabado.

**Descarga progresiva (fast start)**, es un método para empezar a reproducir un medio continuo, sin tener que esperar a que termine la descarga. No se trata de streaming, ya

---

<sup>38</sup> [http://fermat.movimage.com/eside/tim/data/TIM\\_06\\_Video.pdf](http://fermat.movimage.com/eside/tim/data/TIM_06_Video.pdf)



que se utilizan los protocolos como el protocolo de transferencia de hipertexto HTTP<sup>39</sup> o el protocolo de transferencia de archivo FTP<sup>40</sup> y el usuario se queda con una copia del fichero en su disco duro. Si la conexión es suficientemente rápida, la descarga progresiva es para el usuario como un streaming.

Además permite el uso de imágenes de realidad virtual y la modificación interactiva de la dirección de visualización en una escena.

## RealMedia



Figura 1.37 Reproductor Real Player<sup>41</sup>

Es la arquitectura de la compañía RealNetworks para la distribución de audio y video en la Web. Existe un reproductor que se puede instalar como una aplicación cliente o como plugin del navegador.

<sup>39</sup> Protocolo de transferencia de hipertexto

<sup>40</sup> Protocolo de transferencia de archivos

<sup>41</sup> [http://fermat.movimage.com/eside/tim/data/TIM\\_06\\_Video.pdf](http://fermat.movimage.com/eside/tim/data/TIM_06_Video.pdf)

## DivX

Se basa en MPEG-4 pero tarda menos en comprimir reproduce video de mejor calidad, se puede combinar con sonido MP3 y permite almacenar una película de cine de calidad DVD en un solo CD.



Figura 1.38 Reproductor DIVX<sup>42</sup>

Los formatos de archivos pueden estar asociados a una o más arquitecturas multimedia, y cada arquitectura admite numerosos formatos de archivos diferentes.

Algunos de los formatos de archivos más habituales para audio y video son:

- **.MOV:** son los archivos de la arquitectura QuickTime de Apple.
- **.AVI:** Está pensado para la reproducción de secuencias de video a 15 imágenes por segundo. La información de audio y de video está alternada en el archivo y cada una posee su propia sincronización. Es uno de los formatos más usados en las plataformas Windows.
- **.ASF/.WMV.** Son formatos de Windows Media para video, se usan respectivamente para streaming y para la descarga convencional.
- **.RM/.RA/.RAM.** Son los archivos de la plataforma de RealNetworks orientados a la

<sup>42</sup> [http://fermat.movimage.com/eside/tim/data/TIM\\_06\\_Video.pdf](http://fermat.movimage.com/eside/tim/data/TIM_06_Video.pdf)

distribución de audio y video por la Web.

- **.MPG/.MPEG/.MP2/.MP3.** Son una serie de estándares de audio y video, ampliamente implementados.

## CAPÍTULO II

### 2 VIDEO SOBRE REDES IP

#### 2.1 Introducción

En los últimos años se ha experimentado un gran desarrollo y despliegue de aplicaciones de audio y video sobre Internet, como por ejemplo: una video-conferencia. Las nuevas aplicaciones de redes llamadas aplicaciones de medios continuos (**transmisión de video, telefonía IP, radio Internet, video-conferencias, juegos interactivos, mundos virtuales, aprendizaje a distancia y muchas más**), exigen requisitos de servicio que difieren significativamente de aplicaciones flexibles (**correo electrónico, Web, acceso remoto, compartición de archivos**). En particular, estas aplicaciones multimedia son altamente sensibles a retardos entre sistemas finales y pueden tolerar ocasionalmente la pérdida de datos, la Internet no proporciona garantías al respecto, por lo tanto en su diseño es necesario proporcionar soporte explícito a estas nuevas aplicaciones.

##### 2.1.1 Evolución necesaria de la Internet

La evolución de la Internet es necesaria para acomodar mejor el tráfico multimedia, con sus rígidas restricciones temporales las aplicaciones deben reservar explícitamente ancho de banda entre los sistemas finales, si un usuario quiere realizar una llamada de telefónica sobre Internet, desde el host A al host B, la aplicación de telefonía del usuario debería reservar ancho de banda en un enlace entre los hosts.

Esto implica grandes cambios: **Primero**, se necesita un protocolo que reserve ancho de banda entre los extremos. **Segundo**, se deben modificar las políticas de planificación en las colas de los routers (esto a nivel de red), con las nuevas políticas de planificación, no todos los paquetes obtienen el mismo tratamiento y solamente aquellos que reservan

recursos los obtienen. **Tercero**, las reservas de ancho de banda se basan en la diferenciación del tráfico presente en la red. **Cuarto**, la red debe tener una forma de determinar si tiene suficiente ancho de banda para soportar cualquier nueva petición de reserva.

### 2.1.2 Requerimientos para la transmisión de multimedia<sup>43</sup>

En general los requerimientos para la transmisión de audio y video tienen que ver en consideraciones de:

**Ancho de banda.** La información de video actualmente se trata casi exclusivamente de forma comprimida. El ancho de banda dependerá del tipo de compresión y calidad con que se quiera transmitir.

**Retardo de transmisión.** Estos requerimientos son más estrictos que los de ancho de banda. La experiencia con los sistemas de conferencia multimedia y los estándares ITU<sup>44</sup> sugieren un retardo máximo de 150 ms en las aplicaciones de video interactivas.

**Tiempos de compresión y descompresión de las imágenes de video.** El retardo en la compresión y descomposición de las secuencias de video en la fuente y el retardo en el destino, son también importantes factores en este tipo de aplicaciones.

**Fiabilidad.** Las redes tradicionales proporcionan una comunicación fiable entre emisor y receptor. Los protocolos de transmisión tienen sistemas de control de errores y de retransmisión de paquetes que aseguran que esta fiabilidad sea transparente a los niveles superiores.

---

<sup>43</sup> E. Orallo, "Transmisión de datos en Tiempo Real Síntesis de protocolos y redes para transmisión en tiempo real, Síntesis Protocolos" Jul23, 98

<sup>44</sup> Unión Internacional de Telecomunicaciones.

Para aplicaciones en tiempo real, la gestión de errores puede ser negativa debido al retardo que produciría la retransmisión de un paquete. Para evitar este problema, se plantea un tratamiento y gestión de los errores a niveles superiores.

**Sincronización de canales.** Cuando el audio, video y otros datos vienen por distintos canales, se necesitan mecanismos para la sincronización de los distintos flujos en el destino. Esto se puede conseguir usando una combinación de asignación de tiempos y almacenamiento antes de visualización. Esto en general no afecta a la red y es problema del destino.

**Pérdida de paquetes.** La pérdida de paquetes es la principal causa de degradación de la calidad, en el caso de MPEG si se tratase de una imagen I, la degradación de esta provocaría la degradación de imágenes subsecuentes hasta la recepción de una nueva imagen I.



Figura 2.1 MPEG-2 y MPEG-4<sup>45</sup>

En la Figura 2.1, se tiene la distorsión producida por la pérdida de paquetes para un mismo video empleando codificación MPEG-1 y MPEG-4.

---

<sup>45</sup> <http://iie.fing.edu.uy/investigacion/grupos/artes/pqos/pqos.pdf>

### 2.1.3 Calidad de Servicio

La calidad de servicio que se ofrece a una aplicación, se refiere a la disponibilidad de los recursos de red que permiten obtener una calidad de la información aceptable para el usuario. Los requerimientos están relacionados casi exclusivamente con el tiempo de entrega de los paquetes de información, la tasa de pérdida de información, ancho de banda empleado.

Se pueden dividir las aplicaciones en: aplicaciones en tiempo real, aplicaciones en tiempo real adaptativas y aplicaciones elásticas Figura 2.2.



Figura 2.2 Sensibilidad a la pérdida versus sensibilidad al retardo <sup>46</sup>

La duración del retardo y el retardo individual de los paquetes pueden hacer que la calidad de información decaiga, si se excede un tiempo determinado, en este caso la aplicación

<sup>46</sup> <http://www.disca.upv.es/enheror/pdf/Doctorado2Creditos.PDF>

puede retrasar la reproducción de la secuencia, introduciendo distorsión o bien pérdida de información.

**Aplicaciones intolerantes.** Estas aplicaciones no pueden aceptar retardos de la información; por lo tanto esta información requiere de un trato diferente.

**Aplicaciones adaptivas.** Estas aplicaciones pueden tolerar retardo en la llegada de los paquetes y aceptar la disminución de la calidad ya que se incrementa el uso de los recursos de red y proporciona un límite en el retardo de los paquetes, el tráfico tiene que caracterizarse y debe emplear algún algoritmo de control que asegure que una petición de recurso pueda ser aceptada.

**Aplicaciones elásticas.** Este tipo de aplicaciones no requieren ninguna caracterización del servicio para funcionar. Ejemplos de estas aplicaciones son las transferencias (FTP).

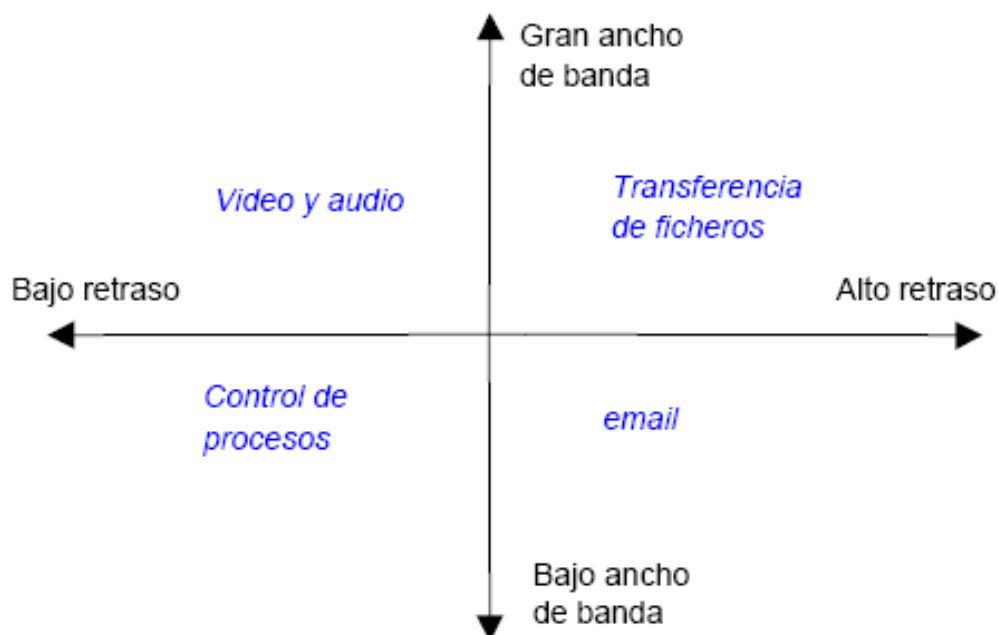


Figura 2.3 Retardo de Información vs Ancho de Banda empleado<sup>47</sup>

<sup>47</sup> <http://www.disca.upv.es/enheror/pdf/Doctorado2Creditos.PDF>



INFORMACIÓN	RETARDO MÁXIMO (s)	ANCHO DE BANDA(Mbps)	TASA DE PÉRDIDA DE PAQUETES
Voz	0.25	0.064	$10^{-1}$
Video	0.25	100	$10^{-3}$
Imagen	1	2 a10	$10^{-9}$

Tabla 2.1 Tipos del tráfico y sus parámetros más importantes.

Los distintos tipos de servicio poseen características y requerimientos particulares, por ejemplo: en telefonía la calidad de servicio será tener un canal de 64Kbps y una disponibilidad del circuito empleado del 99,999 %. En el caso de la Internet, las características heterogéneas de los distintos servicios que soporta, hacen del tema de calidad de servicio un problema mayor. Es difícil identificar en cada caso cuales son los parámetros de rendimiento que aseguran niveles aceptables de calidad.

### Codificación en tiempo real versus pre-codificación

En muchas aplicaciones el contenido de video es codificado previamente y almacenado para su visualización local o remota. Un ejemplo de almacenamiento local es el DVD y video CD y de almacenamiento remoto es el video bajo demanda (VOD: video on demand) y el video streaming.

Las aplicaciones interactivas como la video-conferencia, video-vigilancia o los juegos interactivos son ejemplos de aplicaciones que requieren codificación en tiempo real, aunque otras no interactivas también pueden requerirlo (ejemplo: Transmisión de eventos deportivos).

El video codificado previamente tiene la ventaja de no poseer restricciones de tiempo real, esto permite la implementación de una codificación más eficiente, por otro lado la

flexibilidad es limitada, ya que el video previamente codificado no puede adaptarse a los canales de tasa de bits variable.

Algunos canales soportan una tasa de bits constante (CBR, Bit Rate Constant) como ISDN y otros una tasa de bits variable (VBR, Variable Bit Rate) como el almacenamiento en DVD y la comunicación sobre redes de conmutación de paquetes.

Por otro lado las secuencias de video tienen complejidad variable en el tiempo, por lo que codificar una secuencia con una calidad fija en el tiempo requiere una tasa de bits variable y codificar con una tasa de bits constante requiere una calidad variable, los requerimientos de calidad serán definidos por el tipo de aplicación.

## **2.2 Redes IP**

Una red IP utiliza los protocolos TCP/IP para su funcionamiento, actualmente permiten la integración de todo tipo de información incluyendo transmisión de video en tiempo real.

IP transmite información en datagramas (paquetes de información), desde un origen a un destino, posiblemente atravesando varias redes, está implementado en cada sistema final e intermedio (routers) y proporciona conectividad, direccionamiento y fragmentación. IP fue diseñado como un sistema del mejor esfuerzo, en el que no se asegura ni la ruta, ni el tiempo de transmisión de un datagrama, es por eso que no es adecuado para la transmisión de información en tiempo real.

### **TCP/IP (Transport Control Protocol / Internet Protocol)**

El modelo de referencia TCP/IP y su pila de protocolos, apilan la posible comunicación entre dos computadoras, consta de cuatro capas: aplicación, transporte, Internet, acceso a red Figura 2.4.

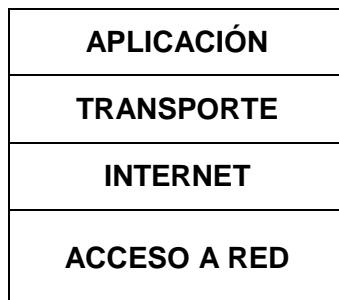


Figura 2.4 Capas del modelo de referencia TCP/IP

A nivel de transporte se tiene el protocolo de datagrama de usuario (UDP, User Datagram Protocol) y el protocolo de control de transmisión (TCP, Transmission Control Protocol), que brindan servicios específicos para soportar diversas aplicaciones.

Tomando en cuenta el modelo TCP/IP, la información de la capa aplicación se fragmenta en datagramas de 64 KB. Cada datagrama es transmitido a través de la red, posiblemente siendo fragmentado en unidades más pequeñas mientras viaja; cuando todas las piezas finalmente llegan al destino son ensambladas para formar el mensaje original.

En la capa de red, tres elementos son principales: el protocolo de red, el protocolo de ruteo y el protocolo de control de errores. Figura 2.5.

**Protocolo de la capa de red:** define el direccionamiento y las acciones tomadas a nivel de red por los routers y sistemas finales.

**Protocolo de ruteo:** determina la ruta que habrá de seguir el datagrama desde la fuente al destino.

**Protocolo de control de errores:** es un mecanismo de información de errores que responde a diversas solicitudes de información de la capa de red.

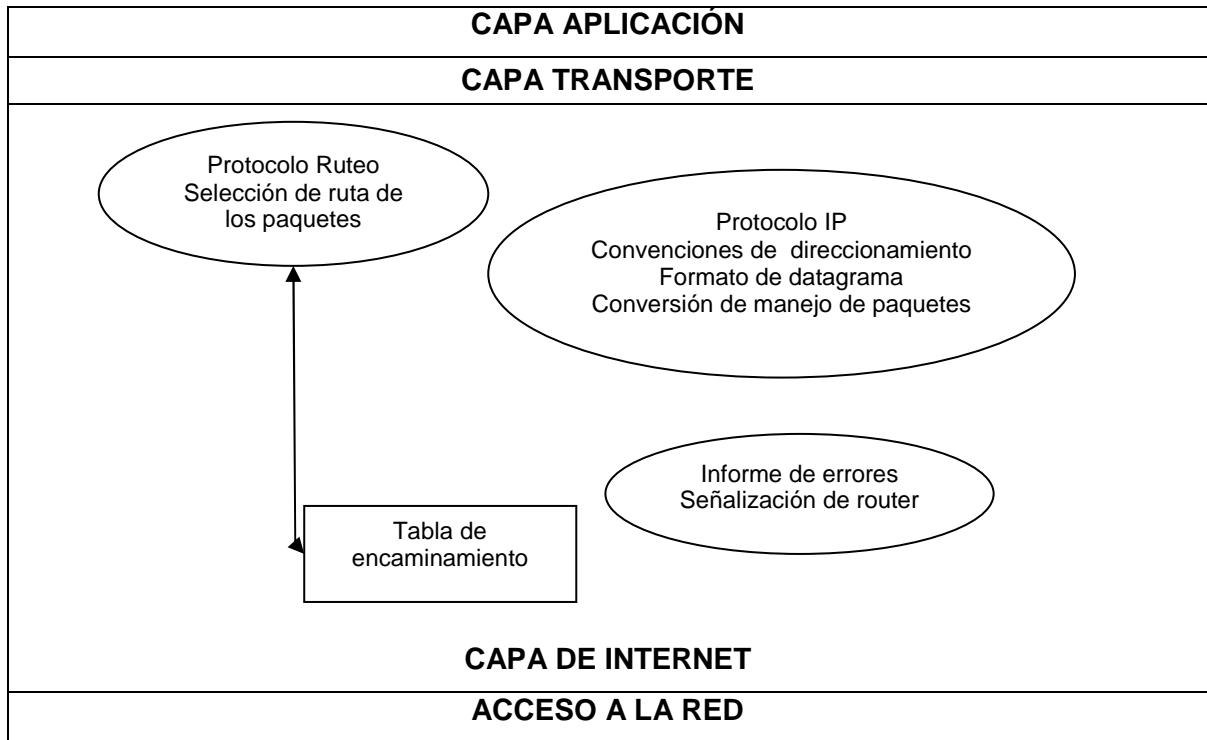


Figura 2.5 Componentes de la capa de red

### 2.2.1 Direccionamiento IPv4

Técnicamente la dirección IP es una característica asociada al interfaz de cada host y tiene un tamaño de 32 bits, se escribirse como 4 dígitos decimales separados por puntos.

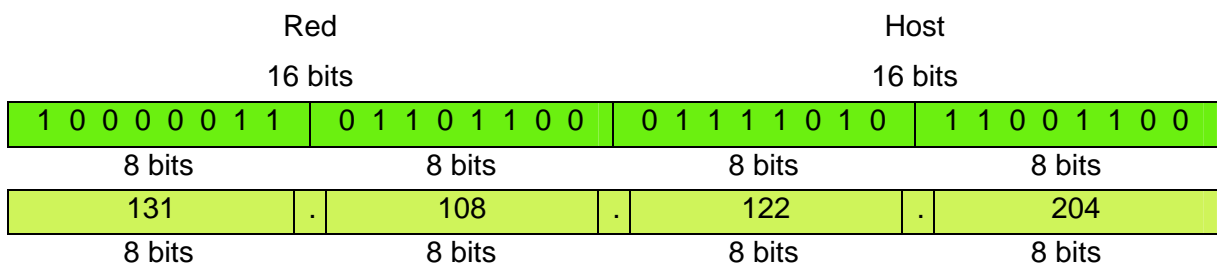


Figura 2.6 Formato de la dirección IPV4

Una dirección IPV4 proporciona la ubicación de un sistema, indicando la subred donde se encuentra.

CLASE DE DIRECCIÓN	NÚMERO DE REDES	NÚMERO DE HOST POR RED
A	126*	16777216
B	16384	65535
C	2097152	254
D (Multidifusión)	-----	-----

Tabla 2.2 Tipos de redes, número de redes y hosts en cada tipo.

**Direcciones clase A.** Dan soporte a redes grandes y emplea el primer octeto (referido de izquierda a derecha) para representar la dirección de red, los tres octetos restantes se emplean para representar los host.

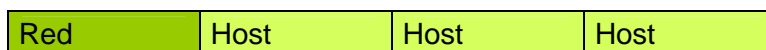


Figura 2.7 Formato de la dirección de red clase A

El primer bit del primer octeto, siempre es cero, puede representar entre 0 (00000000 binario) y 127(01111111 binario), las direcciones 0 y 127 están reservadas, siendo útiles las direcciones de 1 a 126.

**Direcciones clase B.** Dan soporte a las necesidades de redes de tamaño moderado a grande, utiliza dos de los cuatro octetos para las direcciones de red, los dos primeros bits del primer octeto son: 1 y 0 las direcciones va desde 128 (10000000 binario) hasta 191(10111111).



Figura 2.8 Formato de la dirección de red clase B

**Direcciones clase C.** Se crearon para dar soporte a redes pequeñas con muchos host, el primer octeto empieza con 1, 1, 0 por lo tanto las direcciones de red van desde 192 (11000000) a 223(11011111).



Figura 2.9 Formato de la dirección de red Clase C

**Direcciones clase D.** Se crearon para habilitar la difusión broadcast, la dirección de difusión es una dirección de red única, direcciona paquetes a grupos de direcciones IP predefinidas, por tanto una sola dirección puede transmitir simultáneamente un mismo flujo a varios receptores. Los primeros cuatro bits del primer octeto son 1, 1, 1, 0 por tanto pueden tener valores de 224 (11100000 binario) a 239(11101111 binario).



Figura 2.10 Formato de la dirección de red clase D

**Direcciones clase E.** El IETF<sup>48</sup> reserva estas direcciones, para investigaciones y no pueden emplearse a la Internet.



Figura 2.11 Formato de la dirección de red clase E

Los primeros cuatro bits del primer octeto son fijados a 1, las direcciones están entre 240 (11110000) y 255(11111111).

### 2.2.2 Direccionamiento IPV6 <sup>49</sup>

El motivo del esfuerzo e impulso a IPV6 fue el hecho de que las direcciones IP de 32 bits estaban comenzando a agotarse, se pensó entonces en IP de próxima generación (**IP Next Generation**), el resultado IPV6. Los cambios más significativos introducidos se evidencian en el formato de los datagramas.

<sup>48</sup> Grupo de Tareas de Ingeniería de Internet (IETF, Internet Engineering Task Force)

<sup>49</sup> KUROSE, James, REDES DE COMPUTADORAS UN ENFOQUE ESCENDENTE BASADO EN INTERNET. Segunda edición, Editorial Pearson Addison Wesley, Madrid España 2004, p287.

Las direcciones IPv6 son de 128 bits en relación a 32 bits de IPv4, la forma de representar las direcciones IPv6 es empleando valores hexadecimales. Se introducen a más de direcciones unicast y multicast, direcciones anycast que permiten dirigir un datagrama a una dirección de un host dentro de un grupo de hosts.

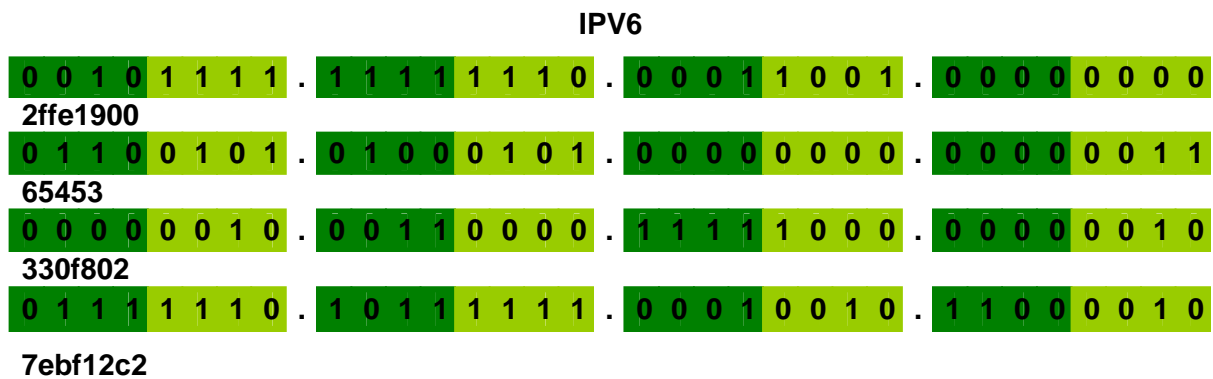


Figura 2.12 Formato de una dirección IPv6

## 2.3 Video sobre IP<sup>50</sup>

Video sobre IP y el IP streaming son las tecnologías que permiten a las señales de video ser capturadas, digitalizadas, secuenciadas y administradas sobre redes IP.

Las presentaciones de video pueden agruparse en tres categorías: video broadcasting, video bajo demanda (Video on Demand) y video-conferencia; de las tres, solo la video-conferencia es full duplex, las otras son esencialmente transmisiones unidireccionales.

<sup>50</sup> Referencias bibliográficas, brindan una introducción a los sistemas de transmisión de video sobre IP  
<http://www.siemon.com/la/>  
<http://www.erg.abdn.ac.uk/users/gorry/level2dp.pdf>  
[http://www.erg.abdn.ac.uk/public\\_html/research/future-net/digital-video/index.html](http://www.erg.abdn.ac.uk/public_html/research/future-net/digital-video/index.html)  
<http://mpeg.telecomitalialab.com/>  
<http://www.openh323.org>

### 2.3.1 Broadcast sobre IP

La forma de emisión broadcast es aquella en la que el contenido se difunde en forma homogénea a todos los espectadores. Esta forma de emisión se ha empleado básicamente en la televisión desde sus orígenes y en la actualidad su uso se ha extendido a otros sistemas.

Video broadcast sobre IP es una transmisión unidireccional de un archivo con contenido de video. Los puntos terminales son meramente visualizadores pasivos. El video broadcast puede ser unicast o multicast, en una configuración unicast, el servidor hace un replica de la transmisión para cada terminal, el efecto que tiene sobre los recursos de la red es el consumo acumulativo de ancho de banda. Cada usuario que se conecta a una transmisión multimedia consume tantos kbps como la codificación del contenido lo permita. En una configuración multicast la misma señal es enviada sobre la red como una sola transmisión pero hacia varios puntos terminales o simplemente hacia un grupo de usuarios.

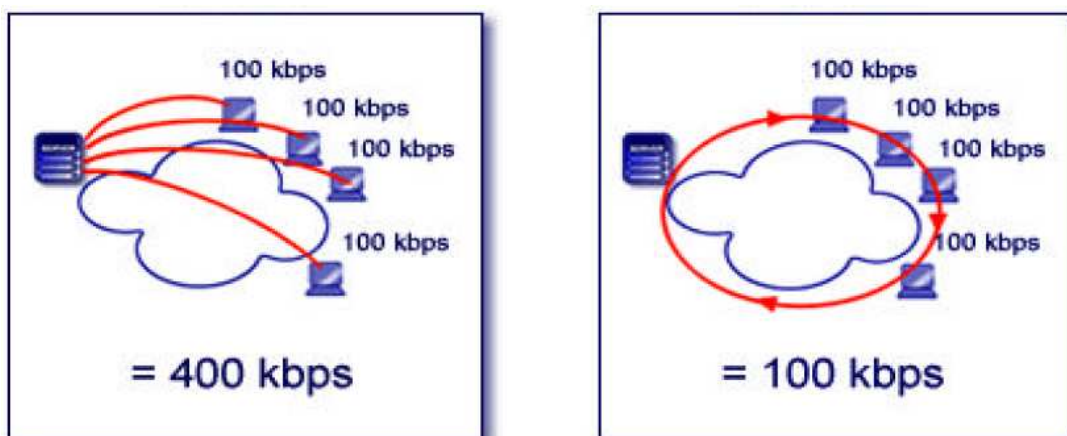


Figura 2.13 Esquemas y consumos en modos unicast y multicast<sup>51</sup>

<sup>51</sup> [www.tdx.cesca.es/TESIS\\_UAB/AVAILABLE/TDX-1124104-162550//fcp1de4.pdf](http://www.tdx.cesca.es/TESIS_UAB/AVAILABLE/TDX-1124104-162550//fcp1de4.pdf)



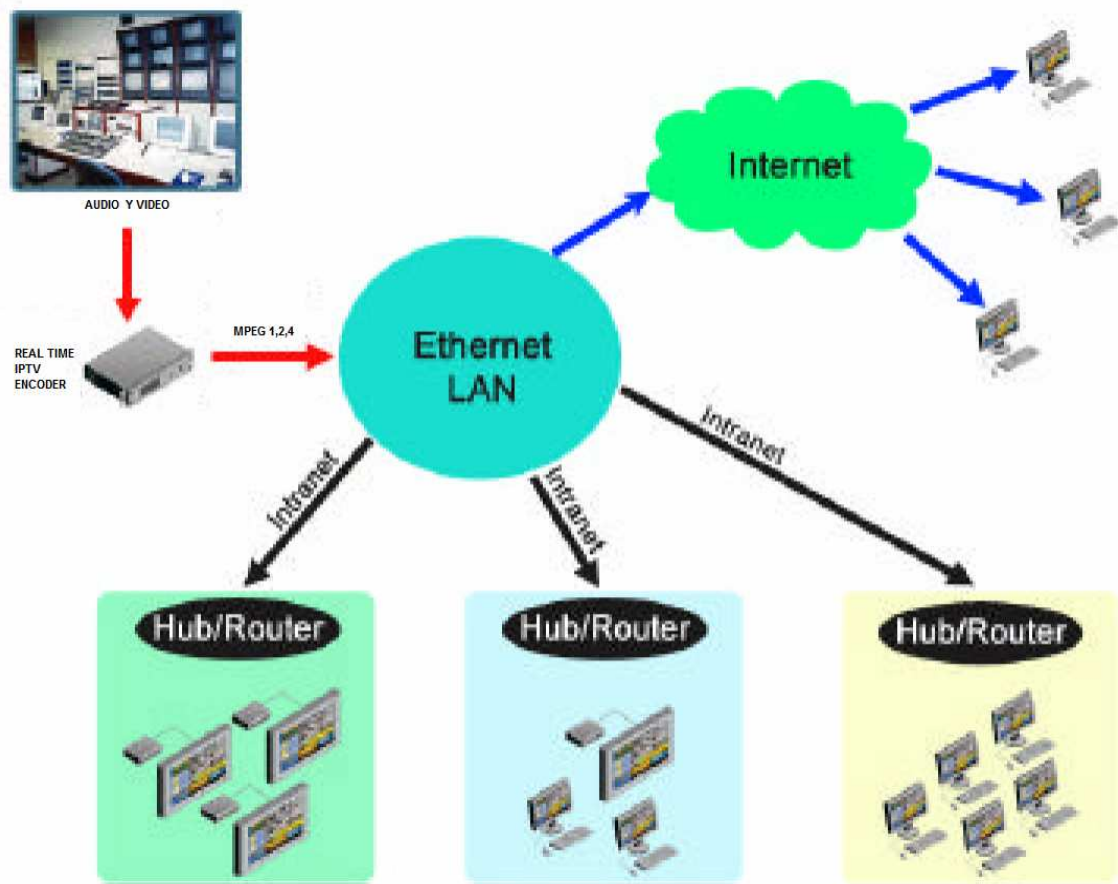


Figura 2.14 Ejemplo de un sistema broadcast IP TV<sup>52</sup>

Con el método *multicast* el consumo del ancho de banda en una red es equivalente al de un único usuario, independientemente de si se conectan a la transmisión cinco, mil, o el número de receptores simultáneamente, también podría implantarse en las redes públicas de los proveedores de acceso a Internet, permitiendo el desarrollo de la televisión sobre Internet. Las diferentes aplicaciones del streaming de video tienen diversos alcances.

<sup>52</sup> [www.etsit.ulpgc.es/cat\\_telefonica/Proyectos\\_Tesis/PI+D/PI%20I+D.htm](http://www.etsit.ulpgc.es/cat_telefonica/Proyectos_Tesis/PI+D/PI%20I+D.htm)

### 2.3.2 Streaming

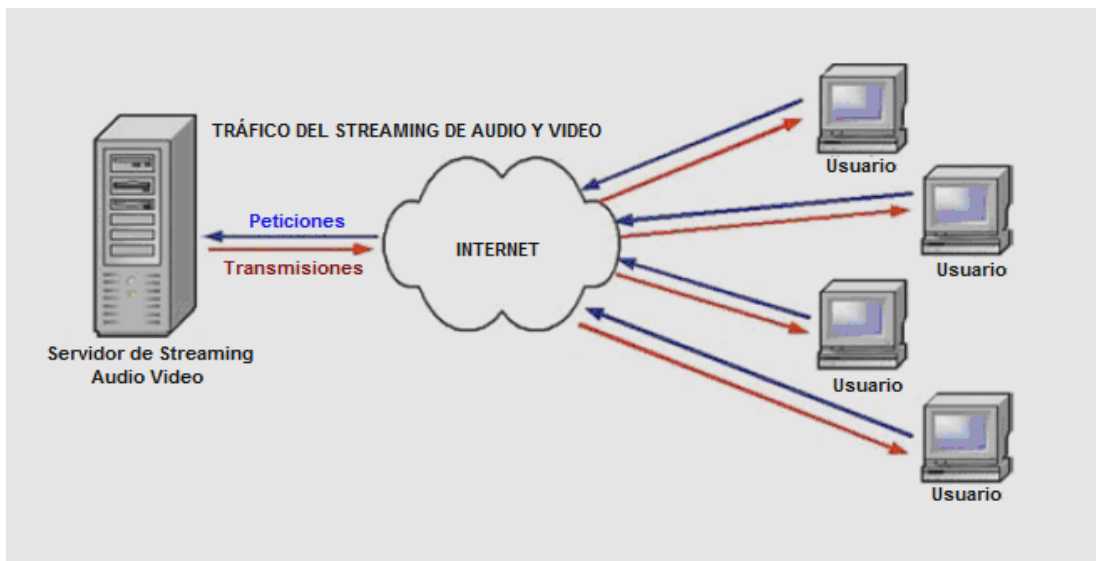


Figura 2.15 Esquema básico del sistema Streaming<sup>53</sup>

Los archivos de audio y video han sido y seguirán siendo el tráfico más conflictivo en una red. Los tamaños de los videos o de cualquier elemento multimedia pueden superar fácilmente los MB, es por ese motivo que se pensó en una manera de transmitir este tipo de información y el resultado es la tecnología streaming.

El streaming es una tecnología que permite la recepción y reproducción instantánea, de información desde un servidor, gracias a los avances en técnicas de compresión, podemos transportar ahora las señales compuestas de video y audio sobre redes típicas como: una LAN, una WAN<sup>54</sup> o la Internet.

#### Redes que soportan streaming

Un atributo fundamental que afecta el diseño del sistema de streaming es la elección del tipo de red ya sea una red de conmutación de circuitos como: la red de telefonía pública

<sup>53</sup> [www.etsit.upgc.es/cat\\_telefonica/Proyectos\\_Tesis/PI+D/PI%20I+D.htm](http://www.etsit.upgc.es/cat_telefonica/Proyectos_Tesis/PI+D/PI%20I+D.htm)

<sup>54</sup> LAN (Red de Área Local), WAN (Red de Área ampliada)

(PSTN) o la red de servicios integrados (ISDN) o una red de conmutación de paquetes como: ethernet y la Internet, cualquiera que se emplee requiere cumplir con requerimientos de ancho de banda, disponibilidad y calidad de servicio.

### **Video bajo demanda**

El video bajo demanda (VoD) hace referencia a los servicios en los que el usuario es capaz de acceder a contenidos multimedia cuando lo desee. Esta tecnología es de vital importancia para diversas aplicaciones como por ejemplo: aprendizaje a distancia, bibliotecas digitales, video-conferencias en Internet, televisión IP, etc.

### **Componentes:**

**Terminal:** los terminales suelen ser de dos tipos: ordenadores PC y Set Top Box, que soportan sesiones tanto unicast (para el video bajo demanda), como multicast (para el video vivo).

**Servidor de video:** es el ordenador que almacena los contenidos audiovisuales, incluye un software que gestiona los recursos y monitorea: el ancho de banda, CPU y capacidad de almacenamiento. El servidor de video debe asegurar que el flujo de datos de video sea continuo, estableciendo una calidad adecuada además debe permitir el control en tiempo real de las secuencias, por parte del usuario empleando el protocolo RTSP (Real Time Streaming Protocol).

**Servidor de aplicaciones:** debe incluir dos módulos básicos: la aplicación de usuario y la aplicación de gestión; la aplicación del usuario garantiza la conectividad de los usuarios ya sea con un Set Top Box<sup>55</sup> o con un PC. Permite la autenticación del usuario y le permite

---

<sup>55</sup> Se llama así al tradicional decodificador

explorar los contenidos para los que está autorizado, una vez que el usuario ha elegido el contenido, la aplicación de usuario controla al servidor de video, para que el contenido llegue al terminal, así mismo se envía los datos de las sesiones establecidas para su tratamiento.

La aplicación de gestión de un sistema de video bajo demanda incluye alta clasificación, modificación y borrado de los contenidos, además de la facturación mediante exploración y borrado de datos de cada usuario.

### 2.3.3 Stack de protocolos TCP/IP para streaming

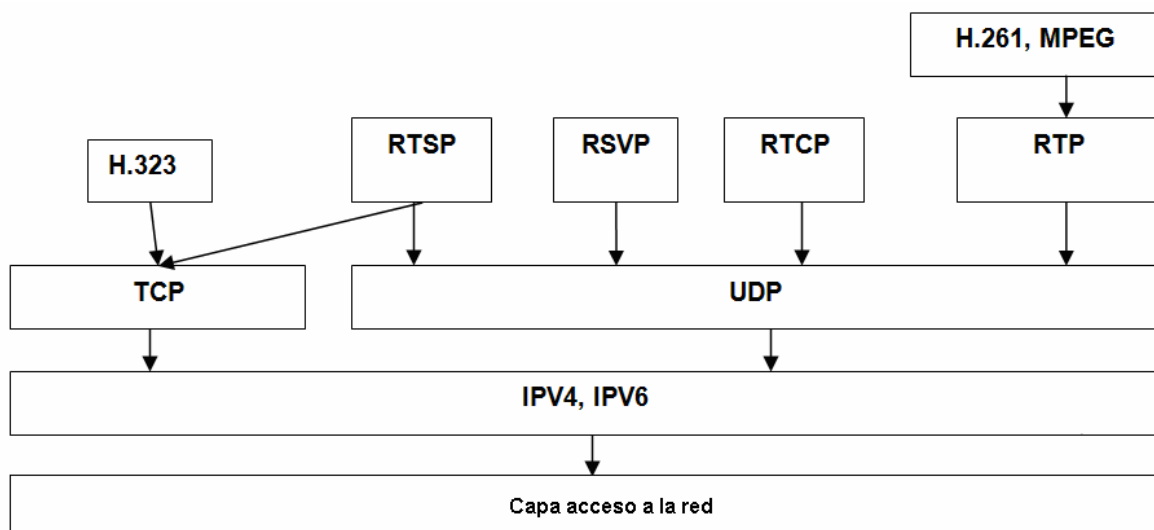


Figura 2.16 Protocolos de soporte para aplicaciones en tiempo real<sup>56</sup>

Las características del tráfico multimedia difieren notablemente de las del tráfico normal; entonces requiere el uso de protocolos adicionales, estas aplicaciones de audio y video no requieren la complejidad del protocolo TCP, ya que los mecanismos de retransmisión de paquetes pueden interferir las emisiones, causando una degradación del servicio demandado.

<sup>56</sup> <http://www.tdx.cesca.es/TESISUAB/AVAILABLE/TDX-1124104-162550/fcp1de4.pdf>

Se han establecido protocolos especializados para las aplicaciones de voz y video, así tenemos: RSVP (Resource Reservation Protocol), RTP (Real Time Protocol), RTCP (Real Time Control Protocol), RTSP (Real Time Streaming Protocol) Figura 2.16.

### 2.3.3.1 Protocolo RSVP

El protocolo de reserva de recursos (RSVP, Resource Reservation Protocol) se ha diseñado para permitir sesiones entre emisores, receptores y routers, mediante el establecimiento de rutas que puedan soportar calidad de servicio, RSVP no es un protocolo de encaminamiento se usa únicamente para reservar recursos a través de la ruta que se establezca por cualquiera de los protocolos de niveles inferiores, se encuentra encima de la capa Internet dentro de la especificación TCP/IP.

RSVP identifica una sesión por medio de una dirección de destino, un tipo de protocolo de transporte y un número de puerto de destino., su diseño cumple con los siguientes objetivos:

- Establecer variaciones de miembros en grupos multicast, así como la conexión o desconexión de miembros de un grupo.
- Permitir a los usuarios especificar sus necesidades a nivel de aplicación, para que los recursos reservados para un grupo multicast puedan reflejar con precisión los recursos necesitados.
- Controlar la sobrecarga que produce el protocolo en la red para que el tráfico no crezca con el número de participantes.

### Principios de diseño de RSVP

**Reserva iniciada por el receptor.** Los receptores escogen el nivel de servicio requerido y son responsables de iniciar y mantener la reserva activa.

**Filtro de paquetes.** La reserva de recursos en un router, asigna ciertos recursos a la entidad que hace la reserva, no determina que paquetes pueden usar estos recursos. La función llamada filtro de paquetes, selecciona los paquetes que pueden usar estos recursos, estos filtros puede ser: filtro libre, filtro fijo y filtro dinámico. Por medio de los filtros de paquetes se pueden definir diferentes modelos de reserva.

- **Filtro libre:** el modelo indica que cualquier paquete con destino al grupo multicast puede utilizar los recursos reservados.
- **Filtro fijo:** indica que mientras dure la conexión, el receptor recibirá paquetes de las fuentes indicadas en la petición de reserva original.
- **Filtro dinámico:** durante la conexión permite modificar la función del filtro y seleccionar un canal entre distintas fuentes, requiere que se asignen recursos para el peor de los casos que es cuando todos los receptores pidan diferente fuentes.

**Control del estado de red.** Durante una comunicación larga es posible que nuevos miembros se unan al grupo mientras otros lo dejen, las rutas puedan cambiar debido a cambios en la red. Por esto, RSVP debe mantener un estado de la red, mediante mensajes que se envían periódicamente.

**Control de sobrecarga del protocolo.** La sobrecarga de RSVP se determina por tres factores: el número de mensajes RSVP enviados, el tamaño de estos mensajes y las frecuencias de refresco de ruta y reserva, para reducir la sobrecarga RSVP estos dos mensajes se unen mientras atraviesan la red.

### **Mensajes de establecimiento de ruta**

Los mensajes primarios usados por RSVP son el mensaje Path, que tiene su origen en el emisor y el mensaje Resv, que tiene su origen en el receptor:

**Mensaje path:** su objetivo es determinar el estado del encaminamiento inverso a través

de la ruta, además de proporcionar a los receptores información del tráfico.

**Mensaje resv:** realiza las peticiones de reserva a los *routers* a lo largo del árbol de distribución entre receptores y emisores.

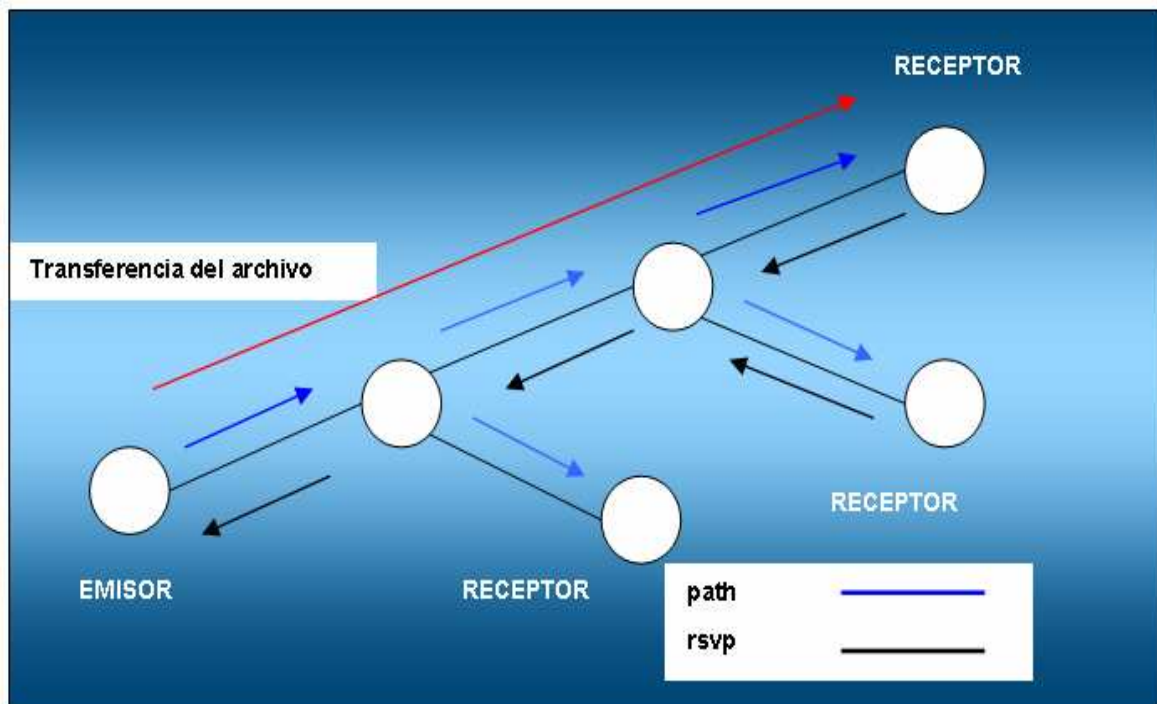


Figura 2.17 Mensajes de establecimiento de ruta

### Tipos de encaminamiento:

Encontrar una ruta que soporte la reserva de recursos. Se puede optar por dos formas de encontrar esta ruta. Una podría ser la de modificar los protocolos de encaminamiento y gestionarlos de acuerdo a un mecanismo de control del tráfico y el rediseño del protocolo para proporcionar múltiples rutas alternativas.

- **Adaptarse a un fallo de ruta:** cuando un nodo falla, el encaminamiento adaptivo encontrará una ruta mediante el refresco periódico de RSVP; aunque la nueva

reserva puede fallar porque no hay suficiente capacidad disponible.

- **Adaptarse a un cambio de ruta:** los cambios de ruta pueden ocurrir sin que se produzcan fallos. Aunque RSVP podría usar las mismas técnicas de reparación mencionadas, esta solución podría producir un deterioro de la calidad de servicio. Podría ocurrir que, si el control de admisión falló en la nueva ruta, el usuario vería una degradación del servicio innecesaria ya que la ruta original está todavía funcional. Se sugiere un mecanismo de fijado de rutas que las mantenga fijas mientras sean viables.

RSVP está actualmente diseñado para trabajar con cualquier protocolo de encaminamiento disponible. Esto puede provocar que se produzcan ciertas degradaciones en la calidad de servicio al no cumplirse los anteriores requerimientos. Se espera que las futuras generaciones de protocolos de encaminamiento se incluyan mecanismos de RSVP.

### 2.3.3.2 Protocolo RTP

RTP proporciona funciones de transporte extremo a extremo, para aplicaciones de transmisión de datos en tiempo real. Trabaja a nivel de aplicación y no garantiza calidad de servicio ni reserva de recursos para los servicios de tiempo real.

El lado emisor en una aplicación multimedia, añade campos de cabecera de audio y video antes de pasarlas a la capa de transporte, dentro del paquete RTP se transporta información de formatos de audio como: PCM, GSM y MP3 y MPEG, H.263 para video, es ampliamente implementado en cientos de productos y prototipos de investigación; SIP y H.323 lo utilizan.



## Paquete RTP

20 Bytes	8 Bytes	12 Bytes	Variable
<b>Cabecera IP</b>	<b>Cabecera UDP</b>	<b>Cabecera RTP</b>	<b>Datos de Audio y Video digital</b>

Figura 2.18 Estructura del empaquetamiento de datos

Los paquetes RTP encapsulan la información de audio y video este a su vez se encapsula en un segmento UDP y después pasa el segmento a IP. El lado receptor extrae el paquete RTP del segmento UDP y después pasa al reproductor de medios para su decodificación y procesamiento.

## La cabecera RTP

El paquete RPT está formado por la información de audio, video y la cabecera que posee cuatro campos principales: el tipo de carga, número de secuencia, marca de tiempo, e identificador de fuente.

7 bits	16 bites	32 bits	bits	Variable
<b>Tipo de carga</b>	<b>Número de secuencia</b>	<b>Marca de tiempo</b>	<b>Identificador de la fuente de sincronización</b>	<b>Varios Campos</b>

Figura 2.19 Campos de la cabecera del paquete RTP

**Tipo de carga en el paquete RTP:** se utiliza para indicar la codificación de audio o video empleada. Se citan algunos de los tipos de material que en la actualidad están soportados por RTP para audio y video.

NÚMERO DE TIPO DE CARGA	FORMATO DE AUDIO	FRECUENCIA DE MUESTREO	TASA
0	PCM, u Law	8kHz	64 kbps
1	1016	8kHz	4.8 kbps
3	GSM	8kHz	13 kbps
7	LPC	8kHz	2.4 kbps
9	G.722	16kHz	48-64 kbps
14	MPEG Audio	90kHz	-----
15	G.723	8kHz	16 kbps

Tabla 2.3 Valores de estándares de audio soportados por RTP

NÚMERO DE TIPO DE CARGA	FORMATO DE VIDEO
26	Motion JPEG
31	H.261
32	MPEG 1 video
33	MPEG 2 video

Tabla 2.4 Valores de estándares de video soportados por RTP

**Número de secuencia:** es de 16 bits, se incrementa en uno para cada paquete RTP que es enviado y puede ser utilizado por el receptor para detectar pérdida de secuencias en los paquetes. Por ejemplo: si el lado receptor de una aplicación, recibe un flujo de paquetes RTP con un espacio entre los números de secuencia 86 y 89 el receptor sabe que los paquetes 87 y 88 se han perdido. El receptor puede, entonces, intentar encubrir los datos perdidos.

**Marca de tiempo:** es de 32 bits y refleja el instante de muestreo del primer byte del paquete de datos RTP, el receptor utiliza las marcas de tiempo para eliminar la fluctuación de paquetes introducida en la red.

**Identificador de la fuente de sincronización (SSRC):** es de 32 bits, identifica la fuente del flujo RTP, cada flujo en una sesión RTP tiene un SSRC distinto. El SSRC no es la dirección IP del emisor sino que es un número que asigna la fuente aleatoriamente cuando comienza un nuevo flujo.

A modo de ejemplo: considérese el uso de RTP para transportar voz. Supóngase que la fuente de voz está codificada en PCM (muestreada, cuantizada y digitalizada) a 64 kbps. Además, supóngase que la aplicación recoge los datos codificados en porciones de 20 mseg, esto es, 160 bytes en una porción. En el lado del emisor cada porción de los datos de audio va precedida de una cabecera RTP que incluye un tipo de codificación de audio, un número de secuencia y una marca de tiempo; la cabecera RTP es normalmente de 12 bytes. La porción de audio y la cabecera RTP forman el paquete RTP. La aplicación extrae la porción de audio del paquete RTP y utiliza los campos de cabecera para decodificar y reproducir adecuadamente el audio.

RTP es utilizado junto con los estándares de telefonía Internet y no garantiza la entrega o tiempo de entrega de los paquetes. De hecho el encapsulamiento RTP sólo se percibe en los sistemas finales.

### 2.3.3.3 Protocolo RTCP

Protocolo de control en tiempo real (RTCP, Real Time Control Protocol) está diseñado para proveer realimentación sobre la calidad de servicio a los participantes de la sesión de la sesión RPT, trabaja junto con RTP y provee los siguientes servicios:

- **Realimentación de QoS:** es la función principal del RTCP, la información se envía a través de reportes de remitente y reportes de receptor.
- **Identificación del participante:** la fuente puede ser identificada por el campo SSRC en la cabecera RTP.

Los paquetes RTCP llevan información de control, calidad de los datos transmitidos, control de flujo, congestión e informes estadísticos y anchos de banda adecuados y pueden ser:

**SR (Sender Report):** ofrece estadísticas de transmisión y recepción de los participantes que son emisores activos.

**RR (Receiver Report):** ofrece estadísticas de recepción de los participantes que no son emisores activos.

**SDES (Source Description):** lo utilizan los emisores para anunciarse, estos paquetes contienen información del usuario: teléfono, e-mail y otros.

**BYE:** indica el final de la participación.

Con la información de RTCP, los emisores pueden ajustar el flujo de datos según el estado de la red.

#### 2.3.3.4 Protocolo RTSP

Protocolo de streaming en tiempo real (RTSP, Real Time Streaming Protocol), trabaja a nivel de aplicación y controla la de sesión para la realización de streaming de medios sobre Internet. Una de las funciones principales de RTSP es el soporte de funciones como: parada, pausa, resumir, avance rápido y retroceso rápido. RTSP funcionan tanto en difusión punto a punto como en multidifusión, permite controlar múltiples sesiones y escoger protocolos de transporte a utilizar como UDP o TCP.

El protocolo soporta las siguientes operaciones:

**Petición de medios:** el cliente pide una descripción de presentación vía HTTP u otro método. Si la presentación es multicast, la descripción contiene las direcciones multicast y

los puertos que pueden ser usados.

**Invitación a un servidor de medios para una conferencia:** un servidor puede ser invitado a unirse a una conferencia existente, bien como participante o simplemente para grabar parte de la conferencia, útil para las aplicaciones de enseñanza distribuida.

**Adición de medios a una presentación existente:** particularmente en presentaciones en directo, el servidor avisa al cliente de que nuevos medios están disponibles. Actualmente existen varias compañías que emplean este protocolo entre ellas: RealMedia y Microsoft Windows Media.

#### 2.3.4 SIP

Protocolo de iniciación de sesiones (SIP, Session Initiation Protocol) permite el establecimiento, la liberación y la modificación de sesiones multimedia; empleando el modelo cliente/servidor.

Está formado por cuatro entidades lógicas, los siguientes son cuatro tipos de entidades lógicas:

##### **Agente de Usuario (AU)**

El Agente de Usuario, es una entidad final que inicia o termina las sesiones empleando señalización; se define el Agente de Usuario como una aplicación que contiene tanto al agente de Usuario Cliente (AUC) y al Agente Usuario Servidor (AUS).

- **Agente de Usuario Cliente (UAC):** es una aplicación cliente que inicia los requerimientos SIP.
- **El Agente de Usuario Servidor (UAS):** es una aplicación servidor que responde los requerimientos SIP de los usuarios.

### **Servidor Proxy SIP**

Es la entidad intermedia que actúa tanto como cliente como servidor, con el propósito de hacer requerimientos a nombre de otros clientes.

### **Servidor Remitente SIP**

Acepta los requerimientos SIP, mapas de direcciones SIP y no pasa los requerimientos a otros servidores.

### **Servidor de Registro SIP**

Acepta requerimientos del registro con el propósito de actualizar la base de datos con la información del usuario contenida en el requerimiento.

### **Mensajes SIP**

Existen dos tipos de mensajes: los requerimientos que envían los usuarios al servidor y las respuestas que envía el servidor a los clientes.

**Requerimientos:** Los requerimientos contienen mensajes para iniciar, mantener y cancelar una sesión Tabla 2.5.

**Respuestas:** Los mensajes contienen respuestas en códigos numéricos Tabla 2.6.

- **Respuesta provisional (Clase 1xx):** son usadas por el servidor para indicar el progreso de la comunicación, pero no terminan los procesos de llamada SIP.
- **Respuesta final (Clases 2xx, 3xx, 4xx, 5xx, 6xx):** son respuestas finales que terminan transacciones SIP.

MÉTODO	DESCRIPCIÓN
INVITE	Inicializa y cambia (re-INVITE) los parámetros de la llamada.
ACK	Es la respuesta final a la invitación.
BYE	Finaliza una sesión.
CANCEL	Cancela una llamada a establecerse
OPTIONS	Pregunta las capacidades y estado del otro extremo.
REGISTER	Registra la situación del servicio.
INFO	Envía información en medio de una sesión, pero no la afecta.

Tabla 2.5 Mensajes empleados del usuario al servidor

CLASE	DESCRIPCIÓN
1xx	Información provisional, requerimiento en progreso pero no terminado.
2xx	Requerimiento completado satisfactoriamente.
3xx	Petición de redirección.
4xx	Error del cliente (error en la petición).
5xx	Error de servidor.
6xx	Falla global.

Tabla 2.6 Respuestas del servidor al cliente

### **Establecimiento de una llamada con una dirección IP conocida**

A manera de ejemplo, Alicia está en su PC y quiere llamar a Roberto, que también está trabajando en su PC. Tanto el PC de Alicia como el de Roberto están equipados con

software basado en SIP para hacer y recibir llamadas de teléfono. Se asume que Alicia conoce la dirección IP del PC de Roberto.

La sesión SIP comienza cuando Alicia envía a Roberto un mensaje INVITE. Este mensaje INVITE se envía sobre UDP. (Los mensajes SIP también pueden ser enviados sobre TCP) El mensaje INVITE incluye: el identificador de Roberto roberto@193.64.210.89, una indicación que Alicia desea recibir audio codificado en PCM ley u y encapsulado en RTP y una indicación de que quiere recibir los paquetes RTP en el puerto 38.060. Después de recibir el mensaje de invitación de Alicia, Roberto envía un mensaje SIP de respuesta. La respuesta de Roberto incluye un 200 OK, una indicación de su dirección IP, la codificación y empaquetamiento para la recepción y el número de puerto al que deberían enviarse los paquetes de audio.

Obsérvese que en este ejemplo, Alicia y Roberto van a utilizar diferentes mecanismos de codificación de audio; Alicia debe codificar su audio con GSM<sup>57</sup>, mientras que a Roberto codificará su audio con PCM U-law.

Después de recibir la respuesta de Roberto, Alicia envía un mensaje SIP de reconocimiento. Después de esta transacción SIP, Roberto y Alicia pueden hablar.

SIP es un protocolo fuera de banda, los mensajes SIP son enviados y recibidos sobre puertos distintos a aquellos utilizados para enviar y recibir datos; los mensajes SIP son ASCII y se parecen a los mensajes http, SIP requiere que todos los mensajes sean reconocidos, por lo que puede ejecutarse sobre UDP o TCP.

---

<sup>57</sup> Es un popular codificador de audio que emplea técnicas de compresión de la tasa de bits del flujo, trabaja a 13 Kbps



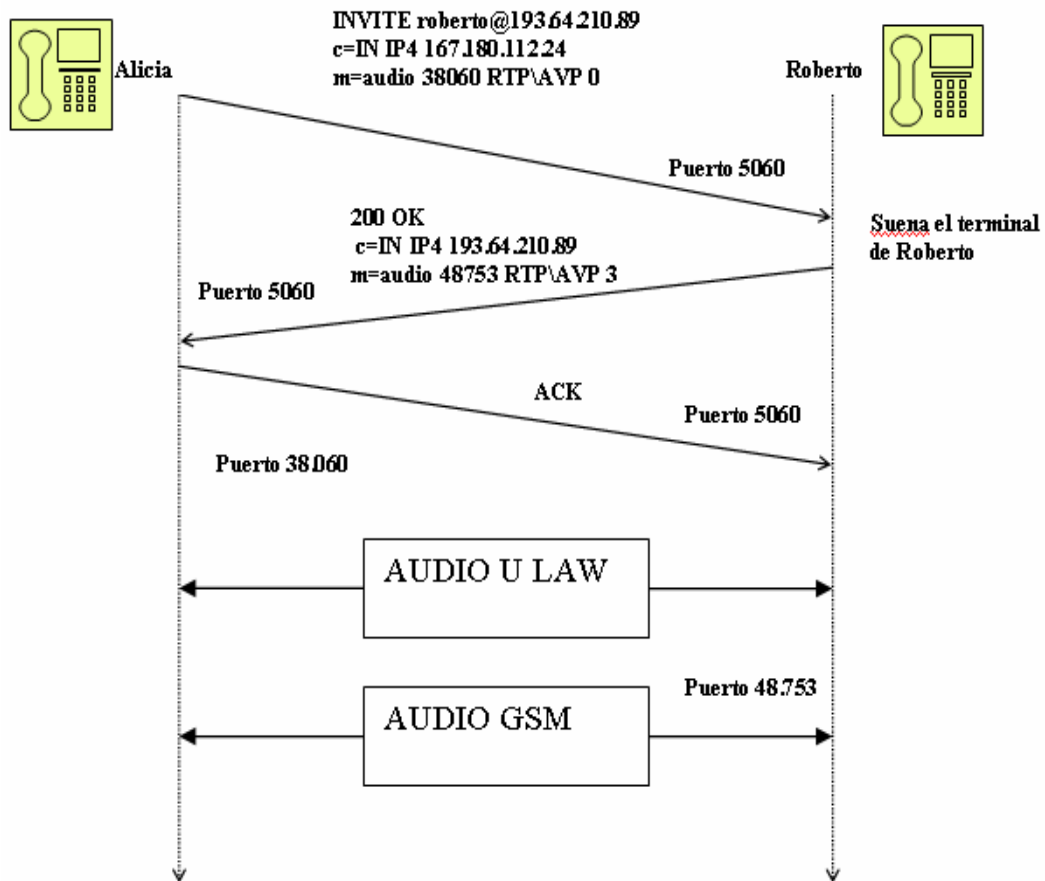


Figura 2.20 Establecimiento de una llamada empleando SIP

En este ejemplo, considérese qué ocurriría si Roberto no tuviese un codificador PCM u-law para el audio. En este caso, en vez de responder con 200 OK, respondería posiblemente con un 600 Not Acceptable, y mostraría en el mensaje todos los codificadores que puede utilizar. Alicia elegiría entonces uno de los codificadores de la lista y enviaría otro mensaje INVITE, esta vez anunciando el codificador elegido. Roberto también podría rechazar la llamada enviando uno de los múltiples códigos de rechazo. Tabla 2.6.

**Direcciones SIP**

Una dirección SIP es roberto@193.64.210.89. Sin embargo, se quiere que las direcciones SIP se parezcan a las direcciones de correo electrónico. Por ejemplo, la dirección anterior

puede ser roberto@domain.com. Cuando el dispositivo SIP de Alicia envía el mensaje INVITE, éste incluiría esta dirección tipo correo; la infraestructura SIP enrutaría entonces el mensaje al dispositivo IP que Roberto está utilizando, otras formas posibles para las direcciones SIP podrían ser el número de teléfono o el nombre y apellidos (asumiendo que son únicos).

Una característica importante de las direcciones SIP es que pueden ser incluidas en páginas Web, de la misma forma que las direcciones de correo electrónico se incluyen en las páginas Web.

### **Traducción de nombres y localización de usuarios**

Un usuario podría tener múltiples dispositivos IP (por ejemplo: para su casa, trabajo y coche) y si se conoce únicamente la dirección de correo electrónico del usuario roberto@domain.com y esa misma dirección es usada para las llamadas SIP, se necesita entonces obtener la dirección IP del dispositivo que está utilizando en la actualidad, para descubrirlo, el usuario envía un mensaje INVITE y envía este mensaje a un proxy SIP, la respuesta SIP puede contener la dirección del dispositivo que está usando.

#### **2.3.5 H.323<sup>58</sup>**

Es la alternativa a SIP, H.323 es un protocolo de señalización que se encarga de la comunicación entre sistemas finales sobre redes que no ofrecen calidad de servicio (QoS) como: la Internet y redes basadas en IP.

---

<sup>58</sup> J. Hidrovo M. "H.323. Multimedia sobre redes IP" , [http://www.ericsson.es/J\\_Manuel Huidobro.htm](http://www.ericsson.es/J_Manuel Huidobro.htm)

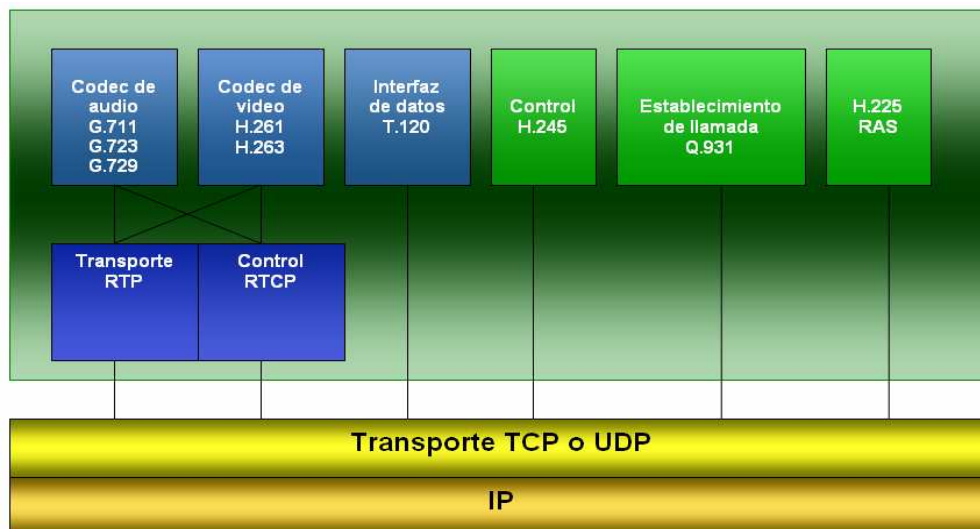


Figura 2.21 Arquitectura de H.323<sup>59</sup>

## Descripción de la arquitectura

### Direccionamiento

- **H.225:** RAS (Registration, Admission and Status), protocolo de comunicaciones que permite a una estación H323 localizar a otra estación H323, a través de una entidad llamada Gatekeeper.

### Señalización

- **Q.931:** se emplea para la señalización de llamada, es usado para configurar una conexión entre dos terminales.
- **H.245:** después del establecimiento Q.931, conecta dos participantes mediante el intercambio de información como: el tipo de mensaje (audio, video o datos) y formato.

H.323 utiliza como transporte el protocolo RTP (Real Time Transport Protocol), diseñado

<sup>59</sup> <http://217.116.8.23/publicac/publbit/bit109/quees.htm>

para manejar los requerimientos de audio y video en tiempo real por Internet y soporta distintos estándares de codificación de audio y video como:

- **H.261:** es un codec estándar de video de la UIT diseñado para transmitir video comprimido a una tasa de 64 kbps y una resolución de 176x144 píxeles (formato QCIF ó Quarter Common Interface format).
- **G.711:** es un codec estándar de audio de la UIT diseñado para transmitir audio PCM (Ley-A y Ley-E) a tasas de 48, 56 y 64 kbps. Opcionalmente, un cliente H.323 puede soportar codecs adicionales: H.263 y G.723.
- **H.263:** es un codec estándar de video de la UIT basado y compatible con H.261. Ofrece compresión mejorada a H.261 y transmite video a una resolución de 176x144 píxeles (QCIF).
- **G.723:** es un codec estándar de audio de la UIT diseñado para operar a tasas de bits muy bajas.

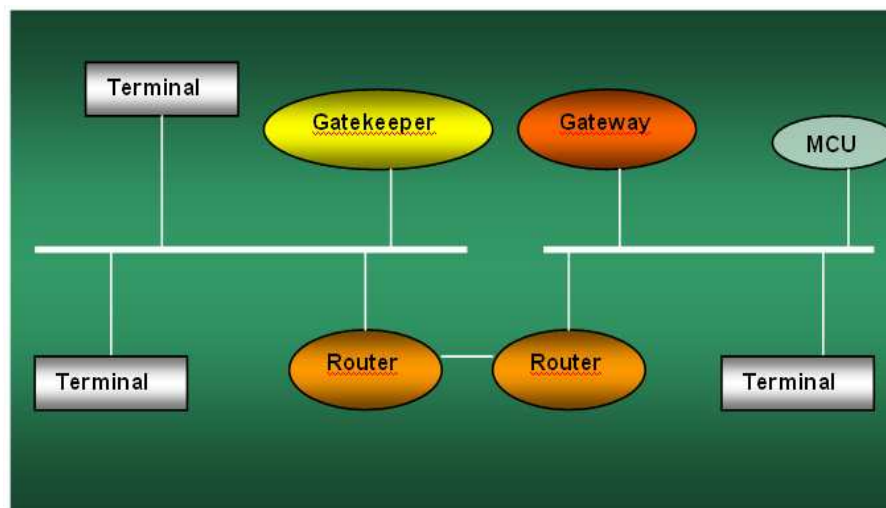


Figura 2.22 Zona de control y gestión H.323

Se definen cuatro componentes principales en un sistema basado en H.323: terminales, gateways, gatekeepers, unidades de control multipunto (MCUs: Multipoint Control Units).

Figura 2.22.

**Terminales:** son los clientes de la red; los terminales deben soportar comunicaciones de voz; en tanto que el soporte de video y datos es opcional.

**Gateway:** es un elemento opcional en una conferencia H.323. Los gateways tienden un puente entre las conferencias H.323 y otras redes, protocolos de comunicaciones y formatos multimedia.

**Gatekeepers:** traducen nombres LAN a direcciones IP y proveen búsqueda de direcciones cuando se necesite, administran el ancho de banda y limitan el número de conexiones de video-conferencia.

### **Unidad de control multipunto (MCU)**

Está diseñada para soportar una conferencia entre tres o más puntos y establece la negociación entre terminales. Un MCU consiste de un controlador multipunto (MC: Multipoint Controller) y ninguno o algunos procesadores multipunto (Multipoint Processors). El controlador MC realiza negociaciones H.245 entre todos los terminales para determinar capacidades de procesamiento de audio y video comunes, mientras que el procesador MP enruta audio, video y datos entre puntos terminales.

## **2.4 Aplicaciones de video digital**

La transmisión de video digital sobre redes ha provocado que las aplicaciones sobre Internet se multipliquen, a continuación se describen las aplicaciones más destacadas entre ellas la video-vigilancia.

### **2.4.1 Televisión IP**

La Televisión IP, mejor conocida como IPTV transformará la televisión actual en una

experiencia totalmente personalizada.

El reto actual es la migración a redes alta velocidad que integren varios servicios, empleando decodificadores que permiten conectarse a la Internet y al servicio de televisión digital, esto implica gran inversión en el tendido de red hasta la instalación de adaptadores IP.

Para que un proveedor de telecomunicaciones pueda ofrecer servicios IPTV deberá entregar múltiples canales de video simultáneos en la misma conexión de acceso. Si en cada hogar se tiene tres receptores de televisión entonces, el proveedor deberá entregar una conexión que soporte tres flujos de video hacia el usuario.

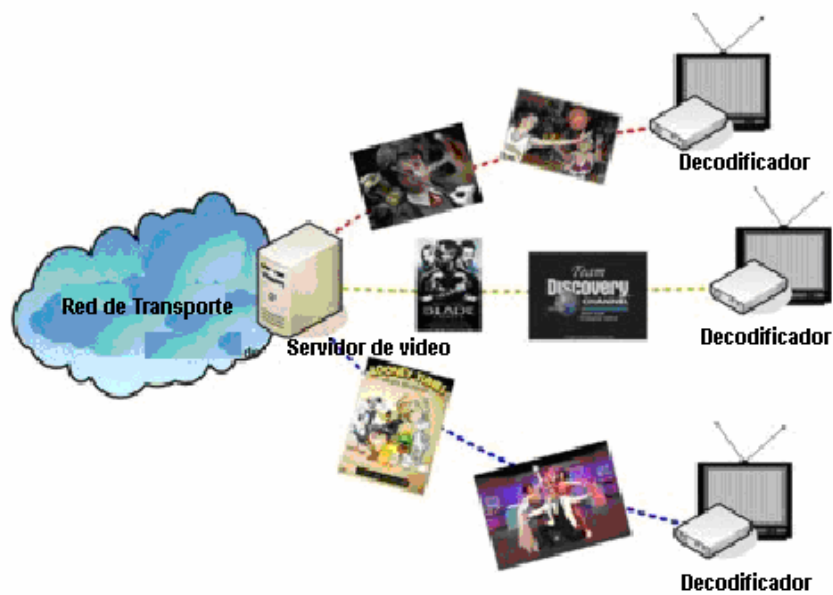


Figura 2.23 Distribución de TV sobre IP<sup>60</sup>

Con IPTV el contenido únicamente llegará al usuario cuando éste lo solicite.

<sup>60</sup> [www.uam.es/centros/eps/esp/posgrado/proyectos/walter.pdf](http://www.uam.es/centros/eps/esp/posgrado/proyectos/walter.pdf)



Figura 2.24 Software para TV en Internet<sup>61</sup>

La capacidad estimada para servicios IPTV se asume en 1.5 Mbps por cada canal estándar (SDTV), para tres canales simultáneos el ancho de banda empleado es de 4.5 Mbps y 11 Mbps se si incluye un canal de Alta Definición (HDTV) y emplean MPEG-4 para la codificación de la señal de video.

SERVICIO	TASA BÁSICA	CONEXIÓN A INTERNET	CAPACIDAD TOTAL REQUERIDA
3 canales SDTV	4.5 Mbps	1 Mbps	5.5 Mbps
2 canales SDTV + 1 canal HDTV	11 Mbps	1 Mbps	12 Mbps

Tabla 2.7 Capacidades y requerimientos de Televisión digital

<sup>61</sup> [www.uam.es/centros/eps/esp/posgrado/proyectos/walter.pdf](http://www.uam.es/centros/eps/esp/posgrado/proyectos/walter.pdf)

CATEGORÍA	FORMATO	RESOLUCIÓN (PIXELES)
SDTV	480i	704 x 480
EDTV	480p	704 x 480
HDTV	720p	1280 x 720
	1080i	1920 x 1080
	1080p	1920 x 1080

Tabla 2.8 Principales formatos de Televisión digital

p = escaneo progresivo, i = escaneo entrelazado

### 2.4.2 Video-Vigilancia<sup>62</sup>

La video-vigilancia es la aplicación del video digital que toma fuerza y actualmente emplean tecnologías alámbricas e inalámbricas basadas en IP.

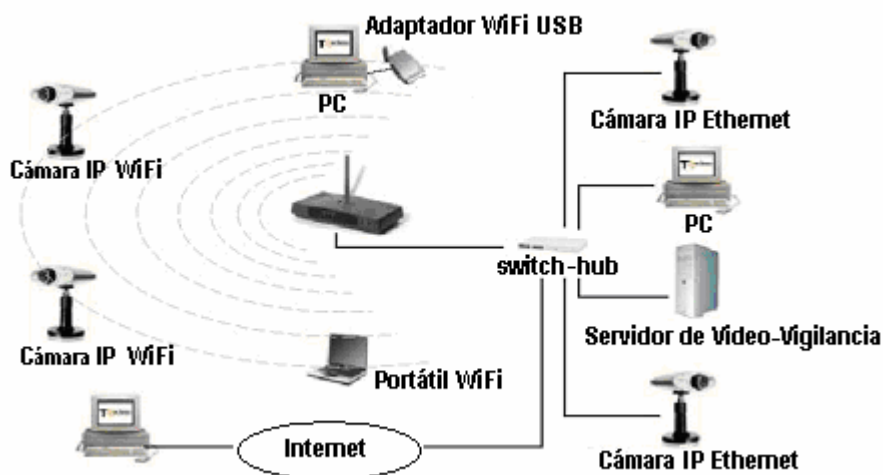


Figura 2.25 Esquema de un sistema de video-vigilancia IP<sup>63</sup>

<sup>62</sup> <http://www.idg.es/pcworldtech/index.asp?seccion=comunicaciones>



Un sistema de vigilancia se vuelve flexible pues empleando tecnología inalámbrica los cambios de la infraestructura no significará mayor problema.

Todos los componentes que actualmente conforman el sistema de video-vigilancia IP, poseen características digitales; la información digital por su naturaleza es más flexible, segura y fiable.

Los sistemas de video-vigilancia se tratan con más detalle en el capítulo siguiente.

---

<sup>63</sup> [www.axis.com/files/brochure/bc\\_techguide\\_26556\\_es\\_0605\\_lo.pdf](http://www.axis.com/files/brochure/bc_techguide_26556_es_0605_lo.pdf)

## CAPÍTULO III

### 3 DISEÑO DEL SISTEMA DE VIDEO-VIGILANCIA

#### 3.1 Introducción

En la actualidad las aplicaciones del video digital, van desde una video-conferencia hasta la televisión y los sistemas de video de alta resolución; el video digital ha evolucionado y sus aplicaciones como la video-vigilancia lo han hecho hasta alcanzar IP. La video-vigilancia IP es una de tantas aplicaciones, fruto del desarrollo de la información multimedia y las redes computacionales; actualmente la mayoría de equipos de video-vigilancia se basan en IP (Protocolo de Internet).



Figura 3.1 Centro de video-vigilancia<sup>64</sup>

Implementada en: aeropuertos, hoteles, hospitales, instituciones bancarias, centros comerciales; la video-vigilancia IP junto al factor humano conforman una infraestructura de seguridad a pequeña, mediana y gran escala. Los centros de monitoreo pueden ocupar pequeñas salas o extensas habitaciones de monitoreo. Figura 3.1.

---

<sup>64</sup> <http://www.siemon.com/la/CCTV> y [Vigilancia por Video sobre 10G ip™](http://www.siemon.com/la/white_papers/SD-03-08-CCTV.asp)  
[http://www.siemon.com/la/white\\_papers/SD-03-08-CCTV.asp](http://www.siemon.com/la/white_papers/SD-03-08-CCTV.asp)

## Sistemas de circuito cerrado de televisión <sup>65</sup>

Los circuitos cerrados de televisión (CCTV, Closed Circuit Television Systems) son los primeros sistemas de video-vigilancia, su origen se remonta a los años 50's, sus características analógicas encarecen su precio y no proporcionan: flexibilidad, escalabilidad, redundancia y tolerancia a fallas.

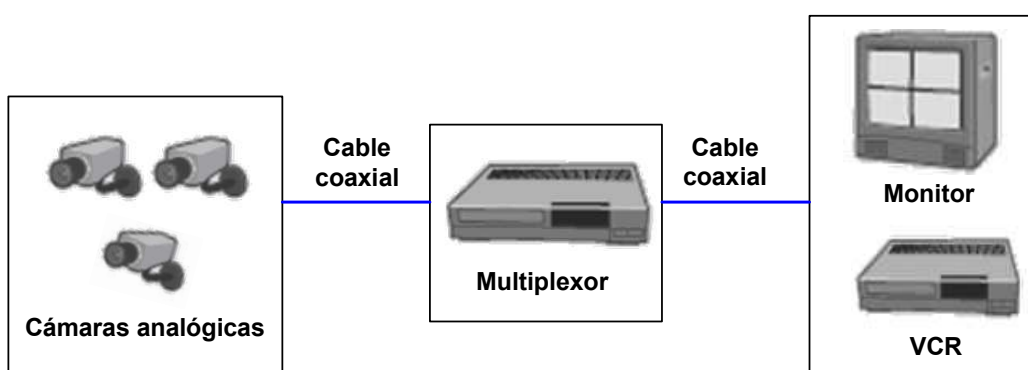


Figura 3.2 Esquema de la plataforma de video-vigilancia analógica

La Figura 3.2 muestra un sistema CCTV que emplea cable coaxial, las conexiones punto a punto desde las cámaras analógicas hasta el multiplexor (MUX) o concentrador de las señales de video, constituyen un reto al momento de implementar el cableado, las grabadoras de video (VCR, Video Camera Recorder) almacenan las secuencias en cintas magnéticas, se puede conectar un multiplexor entre la cámara y el VCR para grabar video proveniente de cualquier cámara.

### CCTV usando DVR

En estos sistemas CCTV, se usa un grabador de video digital (DVR, Digital Video Recorder) encargado de almacenar el video digitalizado de las cámaras analógicas. Figura 3.3. El disco duro reemplaza a las cintas magnéticas y es necesario comprimir el video

<sup>65</sup> [http://www.axis.com/files/brochure/bc\\_techguide\\_26556\\_es\\_0508\\_lo.pdf](http://www.axis.com/files/brochure/bc_techguide_26556_es_0508_lo.pdf)

para almacenar la máxima cantidad de imágenes en un día. En los primeros DVR el espacio del disco duro era limitado y representaba un inconveniente.

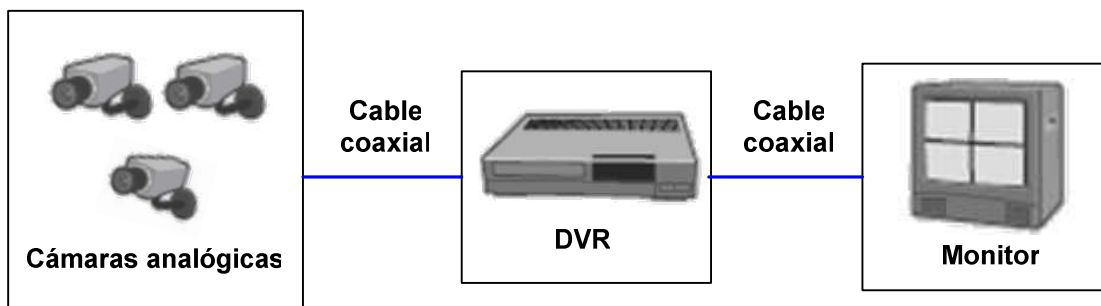


Figura 3.3 Sistema analógico que incluye un DVR para grabar y gestionar, la información de video digital

El DVR puede incluir características IP y mediante una conexión a Internet, el video digital se puede monitorizar remotamente.

### Sistemas de video IP que utilizan servidores de video

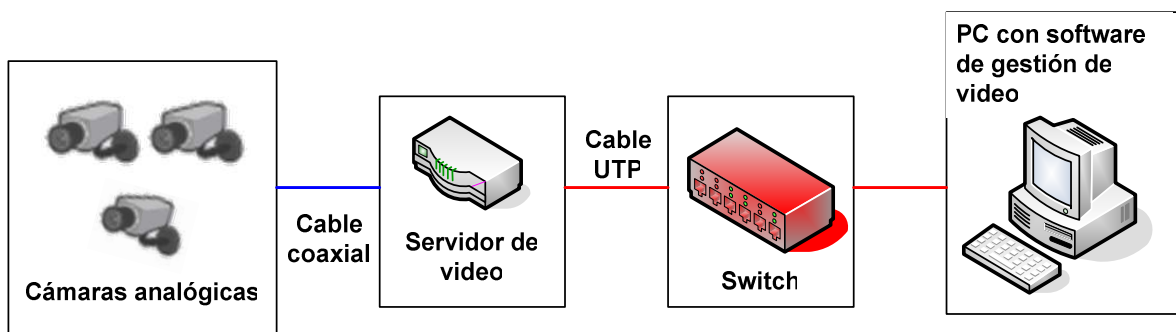


Figura 3.4 Sistema de video-vigilancia IP

El video digital y las redes de datos dan origen a la video-vigilancia IP, que permite migrar a sistemas y medios de transmisión más eficientes y flexibles. Un sistema de video-vigilancia IP puede emplear la infraestructura de una red de datos; el "servidor de video" como en el mercado se ha definido al encargado de digitalizar, comprimir y distribuir las secuencias de video en la red, es el nexo entre los sistemas de vigilancia analógicos y la video-vigilancia IP Figura 3.4.

## Sistemas de video-vigilancia IP<sup>66</sup>

Una cámara IP combina una cámara digital y un ordenador en una unidad, su función: digitalizar, comprimir y transmitir el video hacia una red de datos.

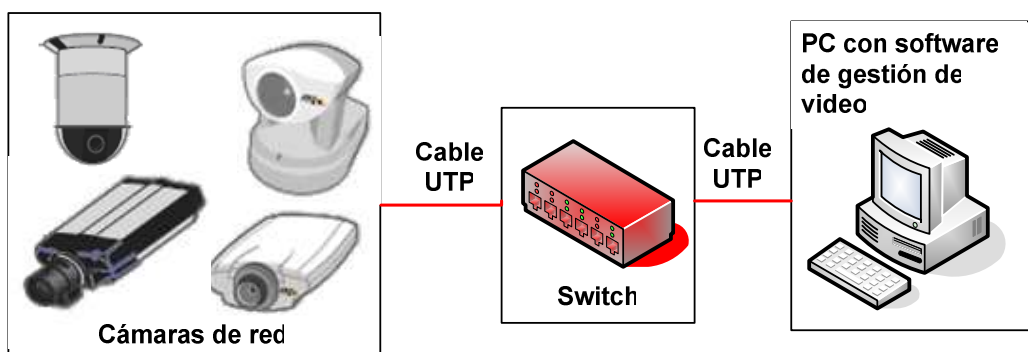


Figura 3.5 Sistema de vigilancia digital IP

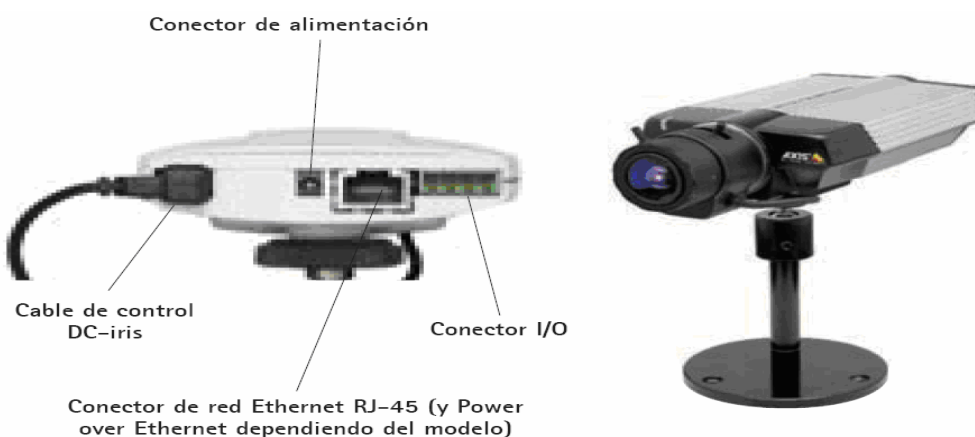


Figura 3.6 Cámara IP y sus partes posterior con sus partes

La cámara IP, completa la digitalización del sistema de video-vigilancia, brinda alta resolución, entradas y salidas para control de accesos y control de movimientos de la cámara entre otras características. Figura 3.6.

<sup>66</sup> Refiérase al anexo C

### 3.2 Descripción del sistema de video-vigilancia

El sistema de video-vigilancia consiste en una comunicación mediante video, con un sitio remoto a través de Internet. El esquema de la Figura 3.7 muestra a un supervisor que establece video-vigilancia desde la ubicación B, C o D con la ubicación A en donde se lleva a cabo un proceso.

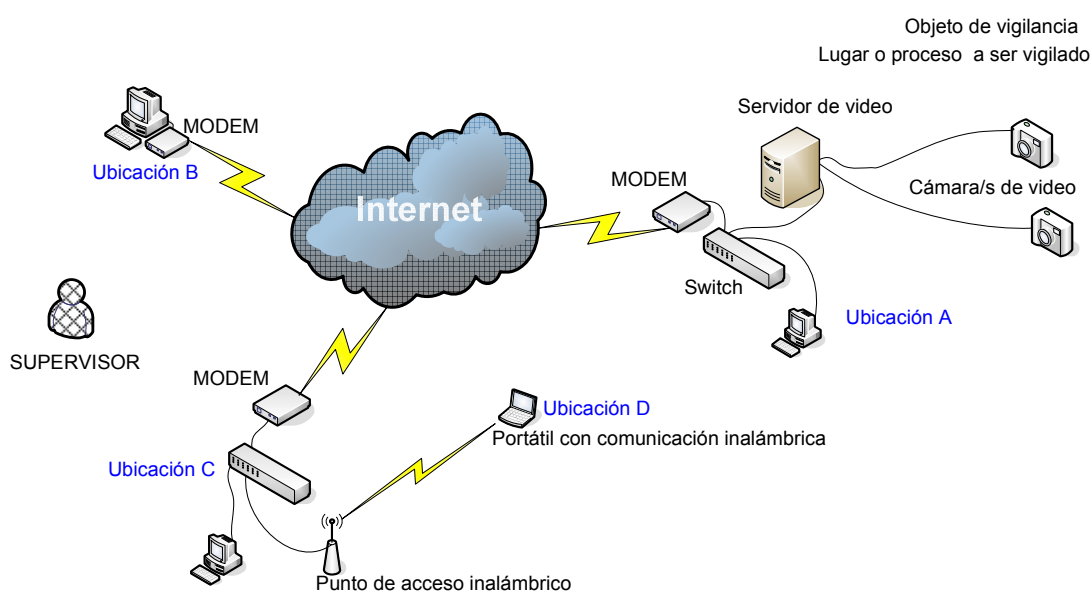


Figura 3.7 Esquema del sistema de video-vigilancia

#### Objeto de la vigilancia

El objeto de video-vigilancia se encuentra en el CICAM<sup>67</sup> (Centro de Investigación y Control Ambiental) de la Escuela Politécnica Nacional en donde se analiza y ejecuta proyectos encaminados a la remediación ambiental.

El proceso a monitorear mediante video-vigilancia consiste en la generación de microorganismos para el tratamiento de agua contaminada, mediante un reactor biológico Figura 3.9. Un reactor biológico genera elementos llamados biolodos, encargados de

<sup>67</sup> [www.cicamecuador.com](http://www.cicamecuador.com)

devorar la materia contaminante, su funcionamiento consiste en la difusión de materia orgánica en el agua contaminada mediante un agitador (compresor de aire) Figura 3.10, los biolodos resultantes se sedimentan y se emplean como punto de partida para la elaboración de estrategias de descontaminación de aguas.

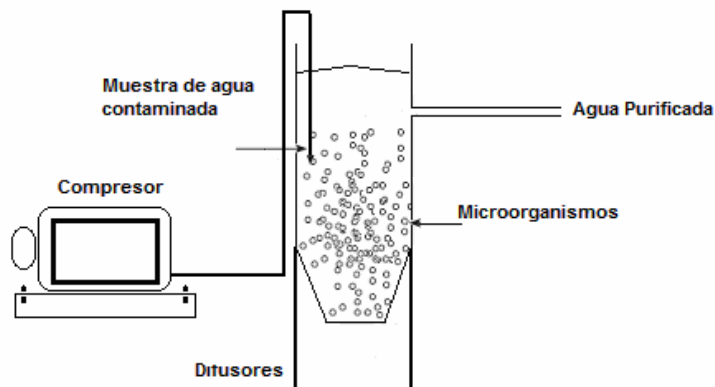


Figura 3.8 Esquema del reactor biológico



Figura 3.9 Reactor biológico del laboratorio del CICAM

De la difusión o agitación del agua depende en gran medida el rendimiento del proceso; entonces es fundamental mantener las condiciones de salubridad y concentración de los biolodos, aquí se emplea el sistema de video-vigilancia que permitirá tomar acciones en

caso de falta de aireación o la generación excesiva de biolodos. Figura 3.11.



Figura 3.10 Compresor empleado para la aireación el reactor biológico

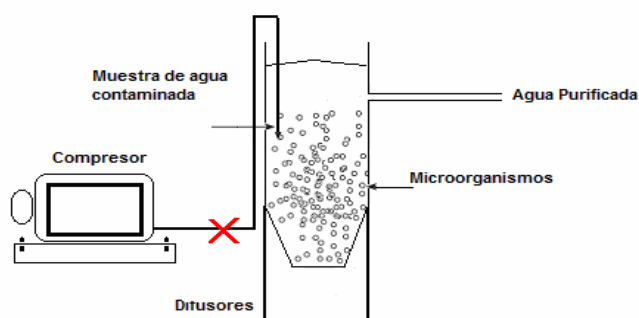


Figura 3.11 Falla del sistema de difusión o agitación

Se dispone de 24 horas para corregir cualquier problema en el proceso (tiempo promedio de vida de los microorganismos) hay que señalar que el proceso requiere de un promedio de 12 semanas para obtener los resultados deseados.

El reactor biológico es un módulo estático dispuesto en el cuarto de modelos del CICAM y requiere de su supervisión constante, incluso cuando el supervisor se encuentre fuera del laboratorio Figura 3.12.





Figura 3.12 Reactor biológico funcionando

### 3.3 Diseño del sistema

El diseño del sistema de video-vigilancia, consiste en delinear las características del equipo y la red sobre la que funciona, esto es: el empleo de Internet como red de video-vigilancia y el empleo de un sistema de video embebido, seguidamente se realiza el análisis del tráfico que genera el equipo de video y su configuración necesaria para acceder remotamente al sistema, finalmente se realizan las pruebas correspondientes.

#### 3.3.1 Internet como red de video-vigilancia

El sistema consiste en una transmisión unidireccional de video en tiempo real, para supervisar un proceso a través de Internet. Por tratarse de una transmisión de video en tiempo real sobre Internet, requiere de un nivel de calidad de servicio, partiendo de que la Internet es una red de conmutación de paquetes y no ofrece calidad de servicio es necesario emplear una estrategia o recurso adicional para obtenerlo. El soporte de calidad de servicio en Internet ha sido posible gracias a protocolos creados específicamente para este fin, así tenemos: RSVP (Resource Reservation Protocol), RTP (Real Time Protocol), RTCP (Real Time Control Protocol), RTSP (Real Time Streaming Protocol) además de

especificaciones de gestión y establecimiento de video-conferencias como H.323 y SIP.

El usuario del sistema requiere de cualquier forma de acceso a Internet, las tecnologías de acceso a Internet se pueden dividir en dos grupos: las que emplean medios guiados (medios alámbricos) y las que emplean medios no guiados (medios inalámbricos) de acuerdo a la infraestructura empleada por un ISP (Proveedor de servicio de Internet) se tiene varias alternativas que van desde un acceso telefónico, hasta un acceso inalámbrico, bajo este esquema el sistema y las redes de acceso a emplear requieren un equipo escalable, es decir la información de video debe transmitirse independiente de la red y del medio empleado, el sistema debe tener la capacidad de adaptarse a la red a la infraestructura donde opere, además de causar el menor impacto en la red.

La seguridad es otro punto importante, en vista de que una persona conectada a Internet es un potencial usuario del sistema, el sistema estará a disposición solamente de un supervisor y una persona encarga de tomar acciones en caso de algún problema, por lo que se asignará claves de acceso al supervisor y a la persona encargada.

### **3.3.2 Sistema de video embebido**

Un sistema embebido es un hardware, software y/o firmware encargado de procesar información, que se encuentra incrustado dentro un sistema más grande; opera en ambientes hostiles, consume poca energía por lo que su tamaño y costo son menores, algunos ejemplos de los sistemas embebidos son los dispositivos electrónicos: celulares, reproductores MP3/MP4, memorias USB, unidades de DVD, etc. Los sistemas embebidos son considerados las aplicaciones más importantes del área de tecnología de información para los años venideros.

En este proyecto se emplea un sistema de video embebido, referido al hardware que codifica el video y emplea algoritmos para transmitir video digital a través de la Internet.

En vista de las diferentes formas de establecer una sesión de video, que pueden ser una alternativa válida para realizar video-vigilancia, describimos las más básicas para comparar sus características técnicas con las de el sistema de video embebido.

### **Cámara Web**

La cámara Web es una cámara digital que funciona con un ordenador, permite establecer una video-conferencia empleando aplicaciones como: Yahoo Messenger, Messenger, NetMeeting entre otras, esta última trabaja tanto en una intranet como en la Internet, NetMeeting de Microsoft puede responder automáticamente a sesiones de video-conferencia, trabaja sobre H.323, emplea resoluciones CIF, QCIF y SQCIF y trabaja sobre canales de

- Modem de 14,4 – 28.8 Kbps
- xDSL<sup>68</sup> ó RDSI (128 Kbps-256 Kbps)
- Red de área local 100Mbps

Para trabajar en Internet requiere el registro del usuario en un directorio Microsoft, se deberá registrar un usuario remoto y un usuario que estará siempre activo en el laboratorio del CICAM.

### **Servidor de video**

Un servidor de video se encarga de atender las peticiones a requerimientos de contenidos de audio o video. En la video-vigilancia el término de servidor de video se aplica al equipo que a más de atender peticiones de video, procesa la información de video para su transmisión. Figura 3.13.

---

<sup>68</sup> Línea digital de subscritor (DSL)

Las señales de video son empaquetadas en IP y son distribuidas en una red de datos, existen muchas compañías dedicadas a producir sistemas de video-vigilancia IP.

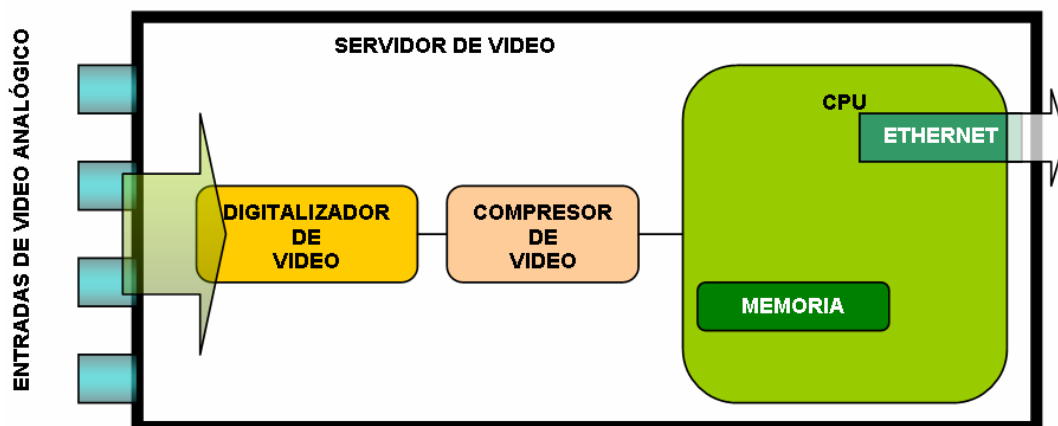


Figura 3.13 Esquema de un servidor de video

Un servidor de video se puede implementar en un ordenador, el servidor constituye en una aplicación que gestiona los recursos de procesamiento, memoria y almacenamiento, las plataformas más populares son RealN etworks y Microsoft Windows Media. Para transmitir video en tiempo real se emplean tarjetas digitalizadoras de video que se integran al ordenador y permiten digitalizar video analógico.

### Una cámara IP

Una cámara IP, digitaliza, comprime y transmite video digital sobre IP, incluye el sensor de imagen, su característica principal es que brinda mejor resolución. Debido a que en nuestro medio su implementación es nueva su costo suele ser todavía alto.

### Comparación con el sistema de video embebido

En las opciones revisadas se han empleado elementos que permiten procesar la señal de video para transmitirse local o remotamente, las tareas de digitalizar, comprimir y transmitir

el video se realiza mediante el ordenador en el caso de la cámara Web y un servidor de video en un ordenador, en cambio al emplear un servidor video-vigilancia o una cámara IP que se ajustan más a la descripción de nuestro equipo embebido, se tienen las unidades de digitalización, compresión y transmisión de video en un solo equipo.

### SLINGMEDIA PRO

El equipo a emplear pertenece a la compañía Sling Media, la cual ha logrado éxito y reconocimientos de la mano de IP, la compañía es la fusión de Texas Instruments y Microsoft Windows Media; el equipo Slingmedia-Pro permite transmitir video con calidad de televisión de alta definición, a través de Internet. Figura 3.14.



Figura 3.14 Vista frontal del equipo SLINGMEDIA PRO<sup>69</sup>



Figura 3.15 Vista posterior del equipo SLINGMEDIA, entradas de video analógico

<sup>69</sup> [www.slingcommunity.com](http://www.slingcommunity.com)

El sistema de video embebido emplea un procesador digital de señales, diseñado por Texas Instruments exclusivamente para la codificación de medios, compatible con los codificadores de Windows Media Series(9) de Microsoft, estas plataformas junto a la tecnología de optimización de streaming desarrollada y patentada por Sling Media conforman el sistema de video embebido.

**Estructura y funcionamiento**

El sistema de video embebido posee dos entradas de audio compuesto y dos de video una de video compuesto y una de s-video, además de una estrada de video de alta definición, cada una de ellas previamente habilitadas y configuradas pueden ser vistas por el cliente a través de Internet empleando la aplicación Sling Media, aplicación que se encuentra disponible en: [www.slingmedia.com](http://www.slingmedia.com).

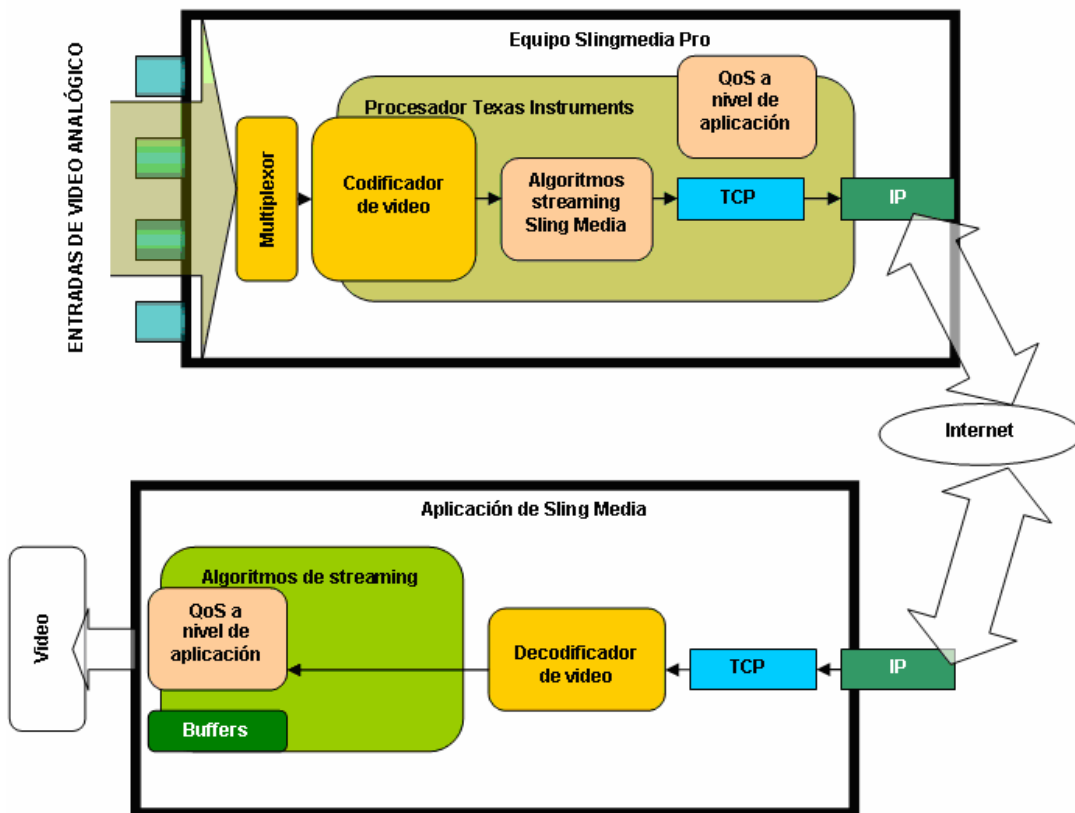


Figura 3.16 Esquema del sistema de video

Para que el equipo pueda ser encontrado por la aplicación en Internet, la aplicación debe apuntar hacia la dirección del equipo (dirección pública), éste genera un flujo de video (video stream) a través de Internet desde el servidor al cliente, que reproduce el flujo en tiempo real a medida que va recibiendo el video; la aplicación emplea recursos de memoria del equipo como buffers que permiten establecer una transmisión streaming de la información.

El equipo establece calidad de servicio a nivel de aplicación, es decir los sistemas finales se encargan de la calidad del video recibido, ya que a nivel de red no se establece diferenciación ni trato especial a la información de video.

Las señales de video analógicas se codifican en base al esquema de un codificador de Microsoft Windows Media, la codificación se realiza en base a hardware y la transmisión de video en tiempo real en base a algoritmos que transmiten empleando streaming. Figura 3.16.

Para la video vigilancia se elige una cámara a blanco y negro analógica; emplea un sensor de imagen CMOS que genera una señal de video analógica NTSC de 380 líneas de resolución; pesa 40 g y su tamaño es de 25x35x15 mm; su sensibilidad mínima es de 3 Lux; emplea 9V y consume menos de 640mW.



Figura 3.17 Cámara empleada para la video-vigilancia<sup>70</sup>

<sup>70</sup> <http://www.domolandia.com/tienda/camara-miniatura-espia-audio-color-carcasa-p-259.html>

## Tráfico generado

Para capturar y visualizar el tráfico generado por el equipo, se transmite video hacia un computador en una red de área local Figura 3.18, para esto se emplea Ethereal que es un software para analizar paquetes en una red de datos.

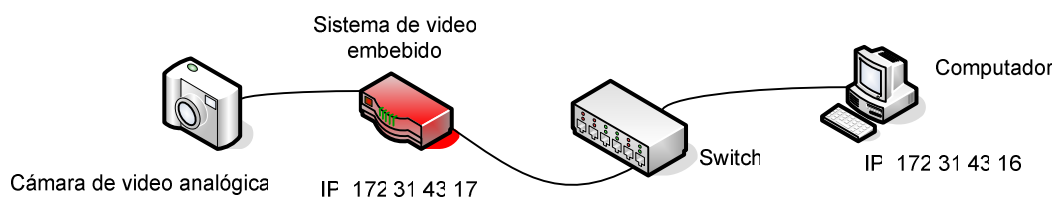


Figura 3.18 Esquema de transmisión de video para la captura de tráfico del equipo

Como se puede observar en la Figura 3.19, el paquete capturado generado por el equipo, muestra que trabaja con TCPI/IP, en la capa de transporte emplea TCP, en la capa aplicación emplea el formato de archivo *stream.asf*, el formato avanzado de secuencias (ASF, Advanced Streaming Format) que es un contenedor multimedia de audio y video digital, propiedad de Microsoft, diseñado especialmente para el streaming.

ASF está optimizado para enviar secuencias multimedia a través de una red y tiene la capacidad de adaptarse a condiciones variantes de una red, soporta una tasa de bits variable (VBR) en la compresión de audio y video, su formato es independiente del codec empleado. ASF define un nivel de intercambio de formato que es independiente del protocolo de la red, es decir no especifica como los datos están fragmentados dentro de los paquetes Figura 3.20.



No. .	Time	Source	Destination	Protocol	Info
8	3.828436	172.31.43.16	172.31.43.17	RTSP	Continuation
9	3.828450	172.31.43.16	172.31.43.17	RTSP	Continuation

```

Source: 00:e0:7d:93:c3:fa (Netronix_93:c3:fa)
Type: IP (0x0800)
Internet Protocol, Src Addr: 172.31.43.16 (172.31.43.16), Dst Addr: 172.31.43.17 (172.31.43.17)
  Version: 4
  Header length: 20 bytes
  Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
  Total Length: 134
  Identification: 0xde0e (56846)
  Flags: 0x04
    Fragment offset: 0
    Time to live: 128
    Protocol: TCP (0x06)
    Header checksum: 0x6e03 (correct)
    Source: 172.31.43.16 (172.31.43.16)
    Destination: 172.31.43.17 (172.31.43.17)
  Transmission Control Protocol, Src Port: 2787 (2787), Dst Port: ttcp (5001), Seq: 1, Ack: 1, Len: 94
    Source port: 2787 (2787)
    Destination port: ttcp (5001)
    Sequence number: 1
    Next sequence number: 95
    Acknowledgement number: 1
    Header length: 20 bytes
    Flags: 0x0018 (PSH, ACK)
    Window size: 65535
    Checksum: 0xacf3 (correct)
  Real Time Streaming Protocol
    GET /stream.asf HTTP/1.1\r\n
    Accept: */*\r\n
    Pragma: sling-Connection-Type=Control, session-Id=0\r\n
    \r\n
    
```

Figura 3.19 Paquete capturado

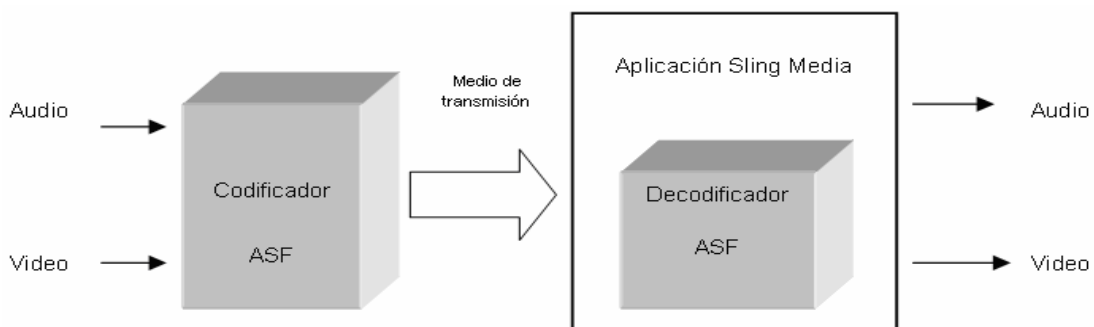


Figura 3.20 Esquema del funcionamiento de ASF

El Slingmedia-Pro además monitorea continuamente la velocidad de la red y ajusta dinámicamente el nivel de compresión de acuerdo con el ancho de banda disponible, aunque, la capacidad del canal define la calidad del video en el lado del cliente, en una

conexión banda ancha el usuario puede definir la calidad del video en función de los parámetros de codificación como: la resolución (160x120 píxeles, 256x192 píxeles, 320x240 píxeles, 640x240 píxeles, 640x480 píxeles), la tasa de bits del video, número de imágenes por segundo, el empleo de suavizado de tráfico (smoothness), la tasa de bits del audio y la tasa de bits de de audio.

Se emplea para la transmisión una resolución de 320 x 240 y se obtiene una calidad aceptable de la señal de video.



Figura 3.21 Imagen de video con resolución 320 x 240 píxeles

El tráfico que se genera Figura 3.22, alcanza un máximo de 240 Kbps, el tráfico en rojo es el que genera el equipo y el tráfico en negro es el tráfico que ingresa al equipo.

Los parámetros de codificación establecidos son:

- Tasa de bits de video: 50 Kbps

- Imágenes por segundo: 1
- Intervalo de la imagen-I : 30
- No se emplea audio

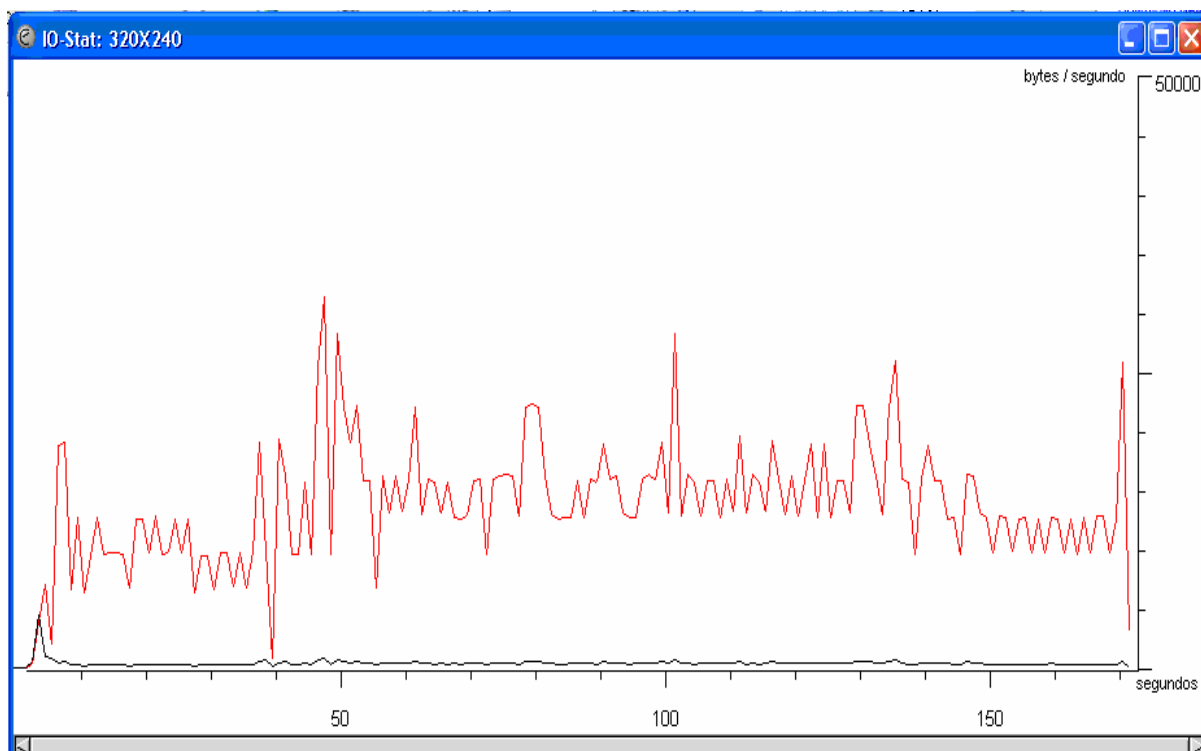


Figura 3.22 Tráfico generado empleando una resolución de 320\*240

### 3.4 Implementación

El reactor biológico funciona en el cuarto de modelos del edificio del CICAM, refiérase al Anexo D para el bosquejo de los planos arquitectónicos. El edificio del CICAM posee dos plantas: la planta baja en donde se encuentra el cuarto de modelos y la planta alta en donde se encuentra el rack de comunicaciones.

#### Descripción de la red

El edificio del CICAM dispone de una red Fast Ethernet, emplea un switch 3COM y el medio de transmisión UTP Cat 5e. La red de área local (LAN) del CICAM se conecta a la

POLIRED para acceder a los servicios informáticos de la misma como son: el servicio de Internet, correo electrónico, etc.



Figura 3.23 Rack de comunicaciones del edificio del CICAM

La red de la Escuela Politécnica Nacional (POLIRED), establece un diseño jerárquico compuesto por una capa de núcleo, capa de distribución y de acceso.

### **Capa de acceso**

Normalmente esta capa es una LAN o un grupo de ellas que proporcionan a los usuarios acceso de primera línea a los servicios de red. La capa de acceso es el punto en el que la mayoría de los hosts están conectados a la red. Se emplean switches Cisco 2950 y 2960 cuyas características se encuentran en el Anexo C.

A este nivel se implementan VLANs (Redes de Área Local Virtuales) para separar usuarios lógicamente, su creación puede ser temporal cuando se programan reuniones, congresos, conferencias dentro de la Escuela Politécnica Nacional o permanente para separar lógicamente a las facultades y profesores.

### **La capa de distribución**

Ofrece servicios de red a las distintas LANs; en esta capa se encuentra el backbone que habitualmente se basa en Fast Ethernet, Gigabit Ethernet o cualquier otra red de gran capacidad. A este nivel se emplean Switches Cisco 3560 y como medio de transmisión fibra óptica multimodo.

### **Capa de núcleo**

A este nivel se encuentran los enlaces WAN (Wide Area Network) de la POLIRED, Figura 3.24, mediante el router Cisco 3845 (refiérase al Anexo C para establecer sus características) la red se conecta a Internet.

En el router se establecen listas de acceso (ACL, Access List), como un nivel básico de seguridad de la red, además de conversión de direcciones de red (NAT Network Address Translation), mecanismo que permite conservar direcciones IP registradas en redes como Internet; según se enruta un paquete a través de un router se traduce la IP privada a una dirección IP pública. A este nivel se implementan los servidores de correo, el servidor Web, el servidor de la base de datos Figura 3.24.

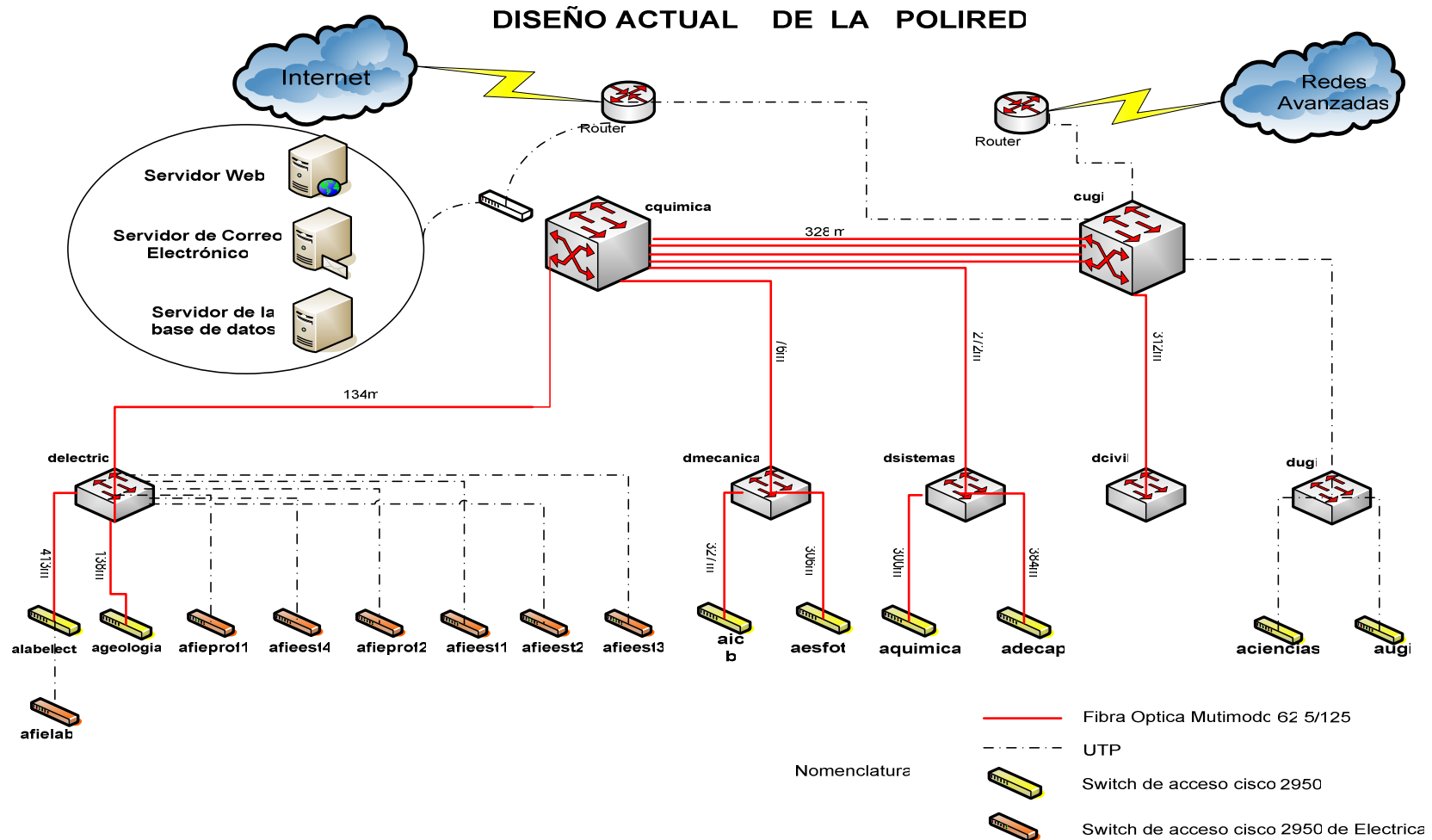


Figura 3.24 Esquema del diseño de la POLIRED a la fecha Agosto del 2007

En la Polired, el equipo se encuentra dentro de la VLAN del UGI (Unidad de Gestión de la Información); para habilitar el tráfico de video que genera el equipo se emplean listas de acceso (ACL) en el router Cisco 3845 y mediante NAT la dirección IP fija del sistema de video embebido se traduce a una dirección IP pública que le permite conectarse a Internet.

### 3.4.1 Configuración del sistema

Antes de realizar la configuración del equipo, se lo instala junto con la cámara de video Figura 3.25, en el cuarto de modelos del CICAM<sup>71</sup> en donde funciona el reactor biológico.

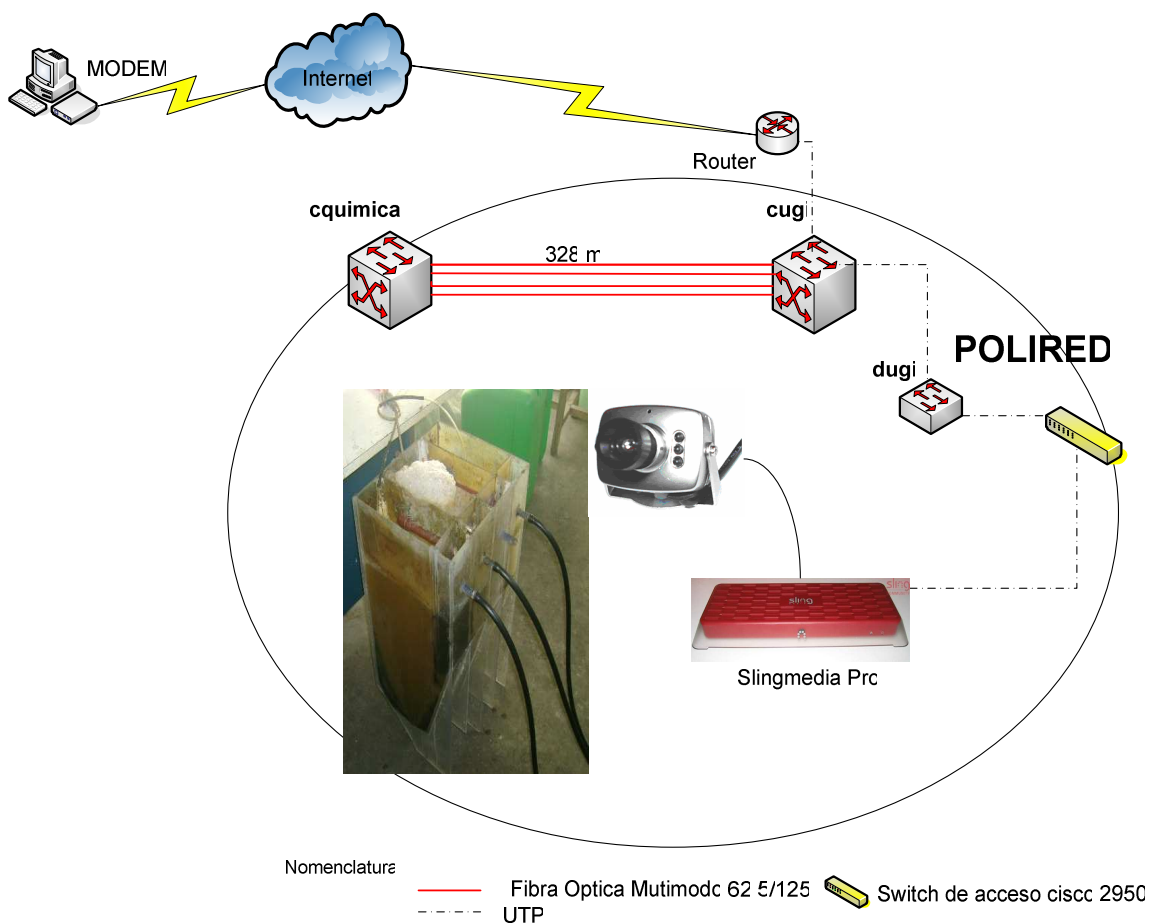


Figura 3.25 Esquema de la instalación del sistema

<sup>71</sup> Anexo D



Figura 3.26 Cámara y reactor biológico

La configuración del sistema de video se realiza empleando la aplicación Sling Media, instalada en un ordenador.

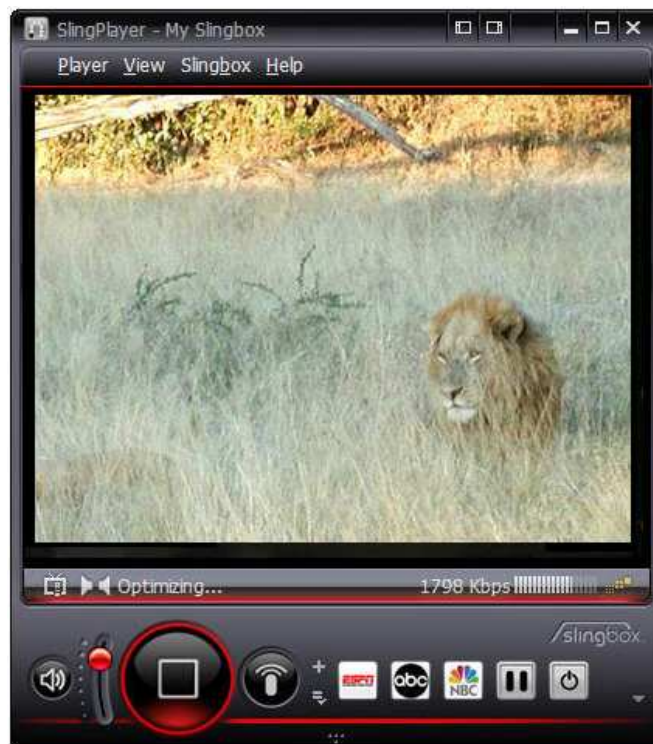


Figura 3.27 Interfaz básica de la aplicación



El sistema de video emplea la dirección IP por defecto 192.168.0.237, el ordenador con la aplicación Sling Media debe estar en la misma red del equipo (puede emplear la dirección 192.168.0.1).

En la interfaz básica de la aplicación, damos clic en **Slingbox**. Figura 3.27, ingresamos a **Slingbox Directory** si el icono del equipo se encuentra en celeste Figura 3.28, la aplicación ha reconocido al equipo y se puede empezar la configuración.

En **Edit Properties** en la cejilla **Slingbox Configuration** Figura 3.29, se comienza por determinar la contraseña de usuario y de administrador, seguidamente se configura y habilita las entradas de video y finalmente la dirección de red, la puerta de enlace predeterminada (default gateway) y el número de puerto de comunicación.

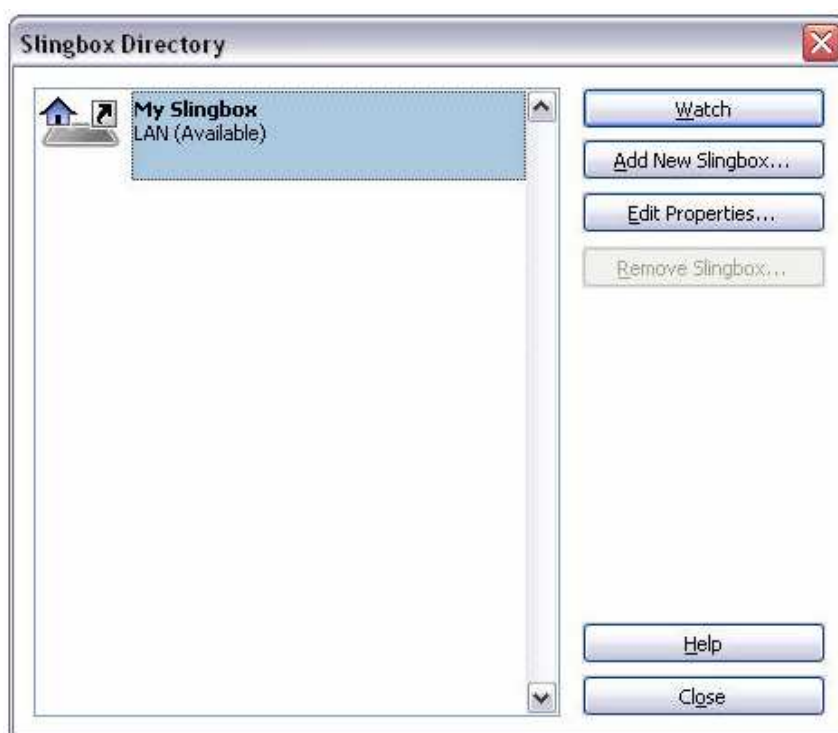


Figura 3.28 Equipo detectado

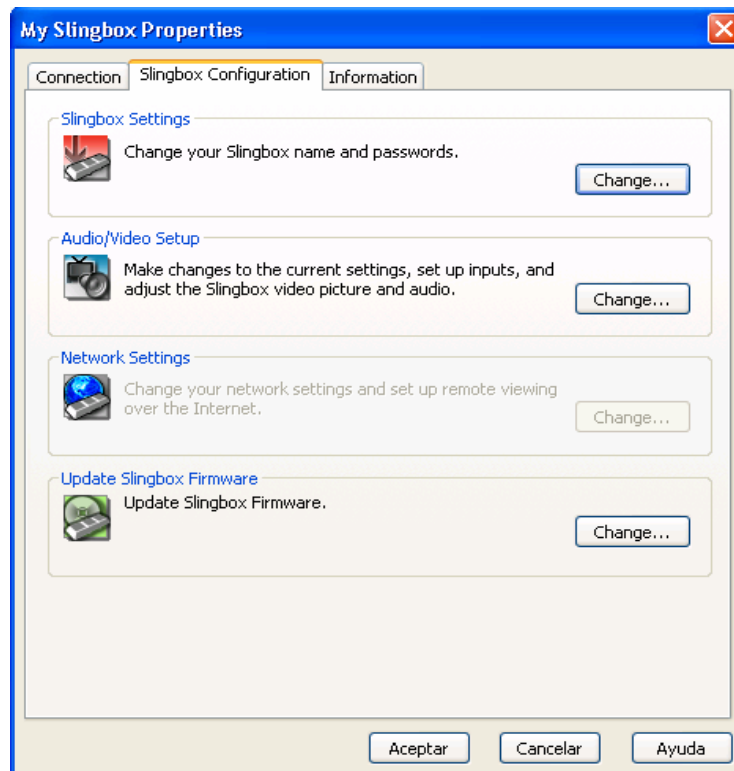


Figura 3.29 Ventana de configuración de parámetros

### Configuración de contraseñas



Figura 3.30 Configuración de contraseñas

La contraseña del administrador tiene preferencia respecto a la contraseña del usuario, una sesión de usuario puede ser finalizada por otro que ingrese como administrador. Figura 3.30. El proceso de descontaminación de aguas está a cargo de un supervisor que evalúa los resultados y una persona que vigila el proceso diariamente, por lo que serán necesarias dos contraseñas.

### Configuración del video

Se pueden habilitar las dos entradas de video y realizar los ajustes de brillo, contraste y saturación, además del volumen del audio.

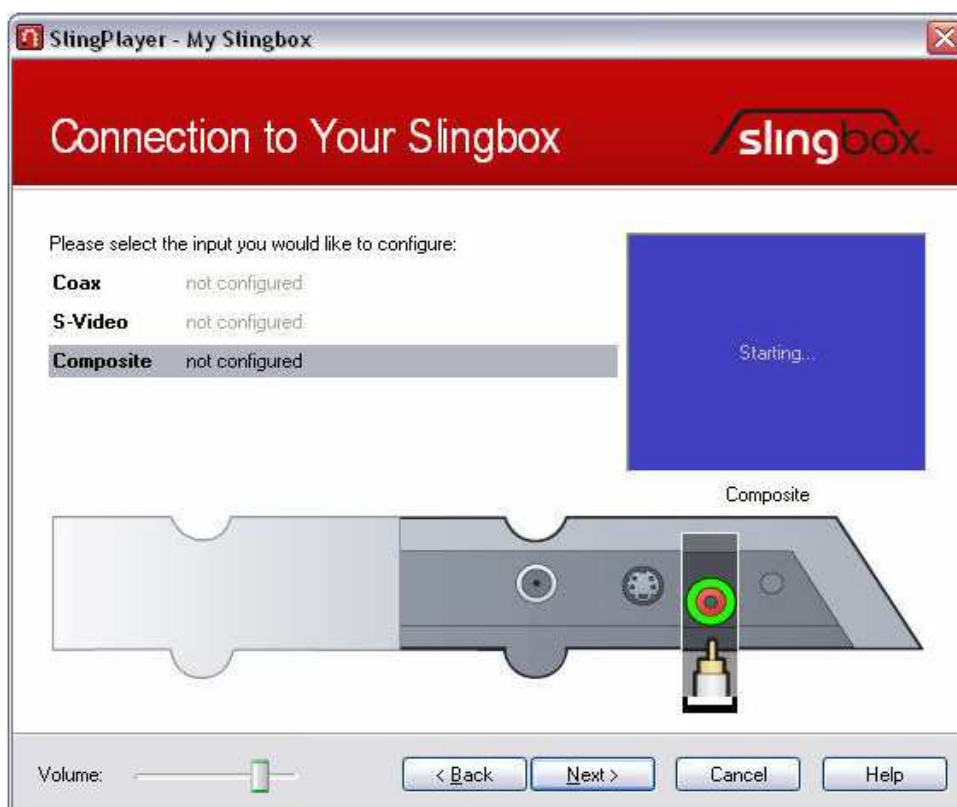


Figura 3.31 Configuración de las entradas de video

Las dos entradas de video del equipo pueden funcionar simultáneamente y el cliente remoto puede seleccionar la que desee ver.



Figura 3.32 Configuración de las características del audio y video

### Configuración de las direcciones de red

En **Network settings** se configura manualmente los parámetros de red Figura 3.33, la video-vigilancia remota se establece con la habilitación del tráfico TCP puerto 5001 y el mecanismo de conversión de direcciones IP NAT, en la POLIRED.

Dirección IP (IP address): 172.31.43.17

Máscara de subred (Subnet mask): 255.255.0.0

Puerta de enlace predeterminada (Default gateway): 172.31.43.1

Número de puerto (Port number): 5001



Figura 3.33 Configuración de características de la red

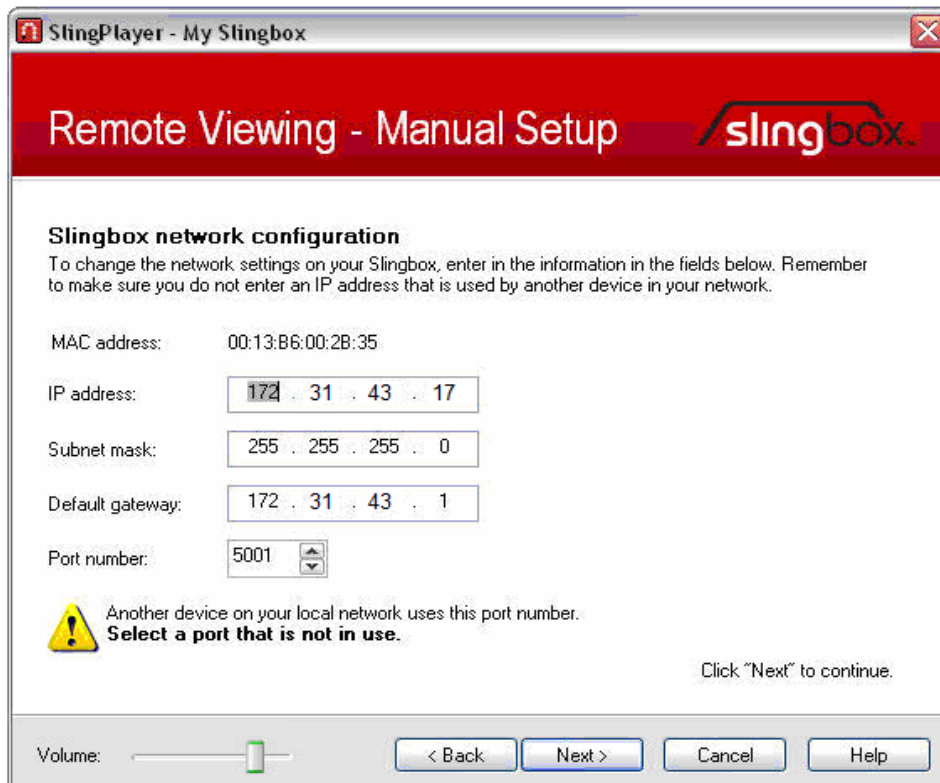


Figura 3.34 Configuración de la dirección IP

Una vez concluida la configuración y habilitado el tráfico, se puede acceder remotamente al equipo.

### Configuración de la aplicación cliente

Previa la instalación de la aplicación Sling Media y mediante un ordenador, se procede a ejecutar la aplicación; ya en la interfaz básica en **Slingbox Directory** ingresamos en **Direct Connection**, la dirección del equipo en este caso es: `vigilancia.epn.edu.ec` en cada sesión, se requiere ingresar la clave de usuario para acceder a la video vigilancia Figura 3.34.

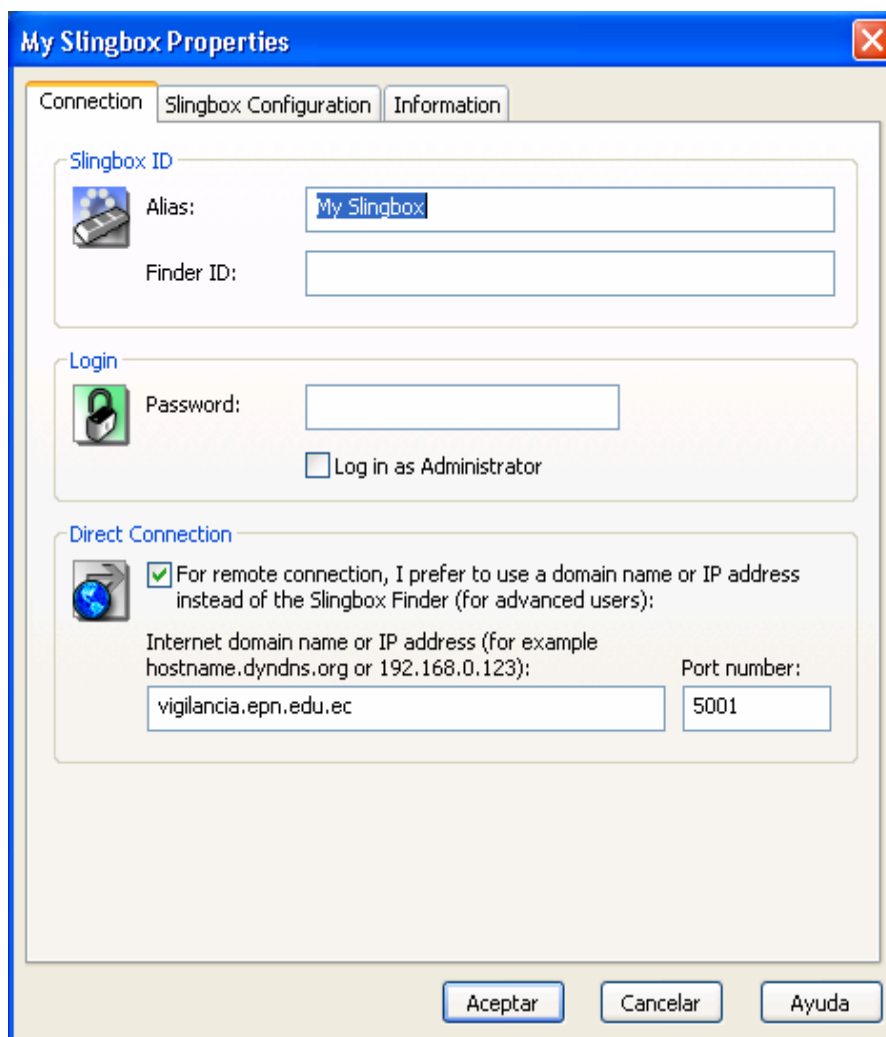


Figura 3.35 Acceso remoto

Cuando se establece una sesión de video, el usuario tiene el control del equipo mediante la configuración de los parámetros de codificación del video, en la Figura 3.36 se muestran las distintas opciones de optimización y codificación de video.

En la interfaz básica, en el menú principal en **Player Options**, ingresamos a **Slingplayer options** y en la pestaña video se puede habilitar:

**SlingStream smoothing:** cuando el flujo de video es más uniforme, sin embargo se pierde detalle en las escenas de video.

**SlingStream optimizer:** realiza la optimización de la calidad del video, en base a los parámetros de codificación establecidos y los recursos de red disponibles.

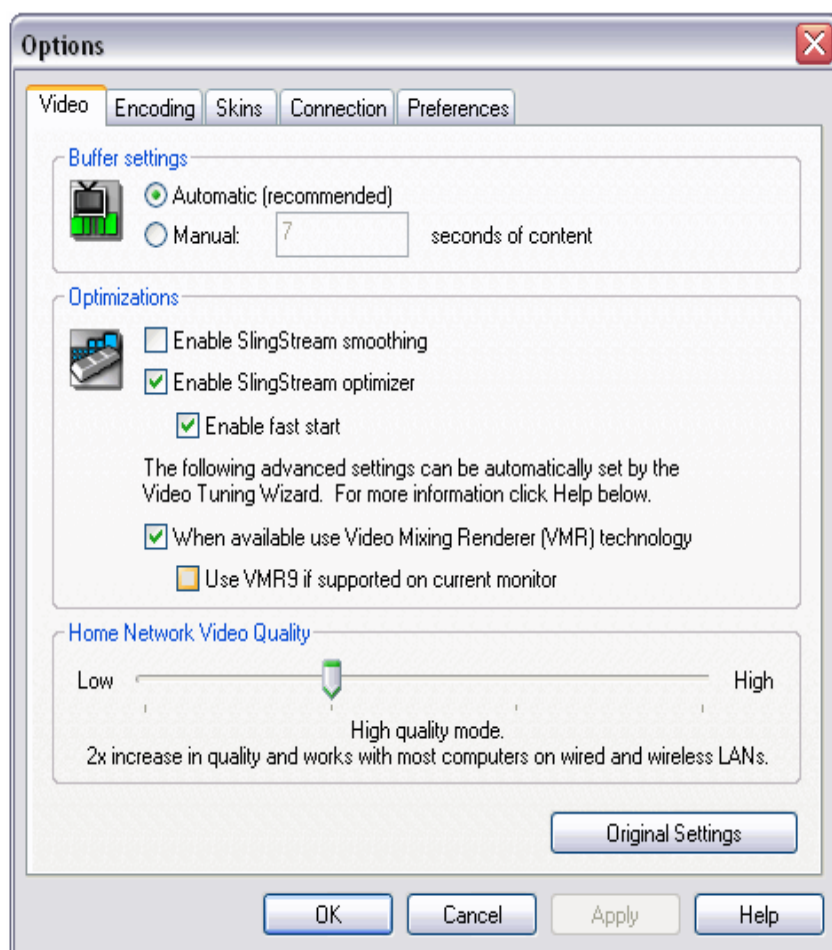


Figura 3.36 Configuración de las características de video

**Fast start:** permite la reproducción progresiva del video, mientras lo recibe.

**Representación de la mezcla de video (VMR, Video Mixing Renderer):** VMR es una tecnología de video que mejora el desempeño del ordenador cuando reproduce video (típicamente disponible en ordenadores con Windows XP) se lo puede emplear cuando se dispone de una tarjeta de video y sus controladores correctamente instalados, Sling Media proporciona buena calidad de video con o sin VMR.

**Calidad de video en una red de área local (Home Network Video Quality):** permite mejorar la calidad del video en una red de área local, posee cuatro niveles de calidad:

- **Calidad estándar (Standard Quality):** posee menores exigencias de CPU y ancho de banda. Se usa cuando la calidad de video es pobre y cuando sobre la red corren otras aplicaciones.
- **Calidad alta (High Quality Mode):** incrementa dos veces la calidad del video respecto a la calidad estándar en redes alámbricas o inalámbricas.
- **Máxima resolución, menos requerimientos de CPU (Full Resolution but Less CPU Intensive):** provee mayor resolución y pocos requerimientos de CPU, emplea una resolución de 640x480 píxeles y usa un ancho de banda entre 3 Mbps y 6 Mbps.
- **Máxima resolución y máxima tasa de bits de video (Full Resolution and Maximum Bit rate):** provee mayor resolución y mayor ancho de banda, con una resolución de 640x480 píxeles y un ancho de banda de 3 a 6 Mbps.

## Codificación

El equipo optimiza las señales de video en base a los recursos de la red y al movimiento contenido en ellas. En la Figura 3.37, se tiene tres opciones de codificación:



- **Baja acción (Low Action):** se usa cuando el video presenta poco movimiento en sus escenas.
- **Acción media (Medium Action):** se usa cuando el video posee movimiento moderado en sus escenas.
- **Acción alta (High Action):** se usa cuando el video posee mucha acción en sus escenas.

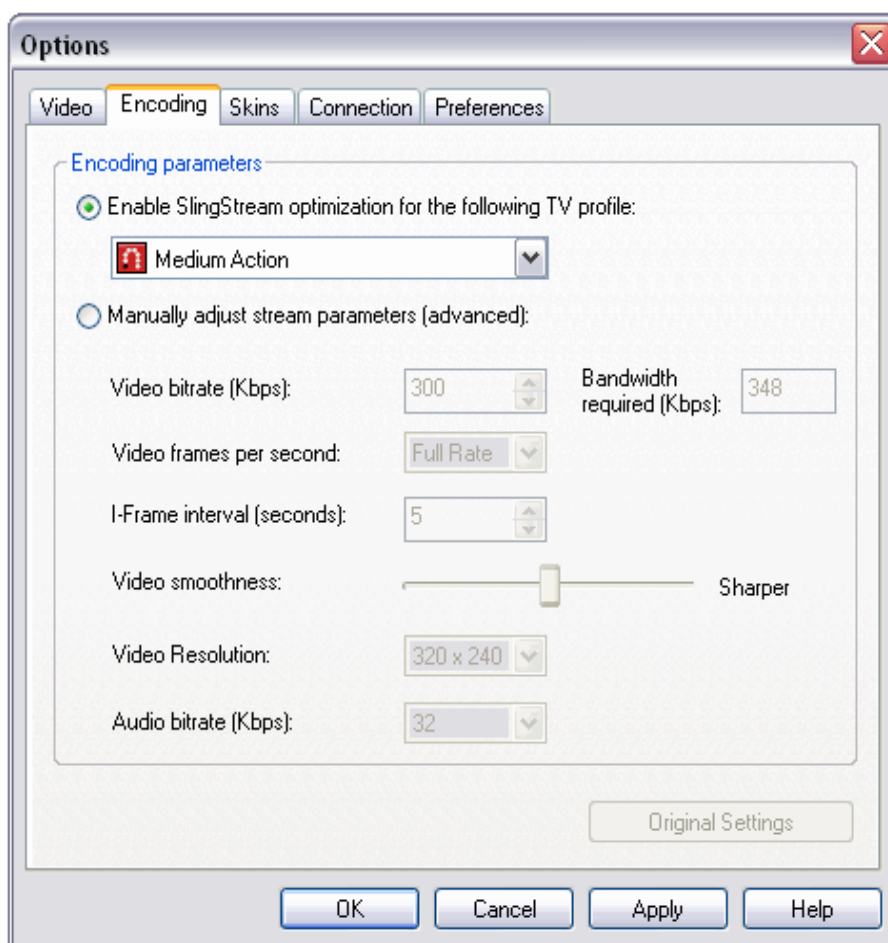


Figura 3.37 Configuración de las características de codificación de video

Los parámetros de codificación de la señal de video se puede predeterminar y optimizarlos para obtener mejor calidad en base a la conexión a Internet que se emplee:

- **Tasa de bits de video (Video bitrate):** la tasa de bits del video que generará el

equipo.

- **Imágenes por segundo (Video frames per second):** imágenes por segundo que transmite el equipo.
- **Intervalo entre imágenes I (I-Frame interval):** cuando se transmite video sobre la red, la tecnología de compresión transmite sólo el diferencial entre cada imagen. Cuando los intervalos son más cortos, la calidad video es más alta pero usa mayor ancho de banda.
- **Suavizado del video (Video smoothness):** cuando el video es más uniforme las escenas aparecen más naturales pero con menos detalles (smoothness), cuando el video posee más movimiento las escenas se vuelven inestables pero se preservan más detalles en cada escena (sharper).
- **Resolución de video (Video Resolution):** tamaño de la imagen.
- **Tasa de bits del audio.** Tasa de bits de audio.
- **Requerimiento de ancho de banda (Required bandwidth):** aquí se verá el ancho de banda requerido para transmitir el video.

Una buena calidad de imagen se obtiene empleando una resolución de: 320x240 píxeles, los demás parámetros de codificación se pueden fijar: a una tasa de bits de video de 50 Kbps, 1 imagen por segundo, 30 segundos de intervalo de imagen-I, empleando smoothness y sin emplear audio. Esta configuración requerirá de un ancho de banda de 69 Kbps. Figura 3.38.

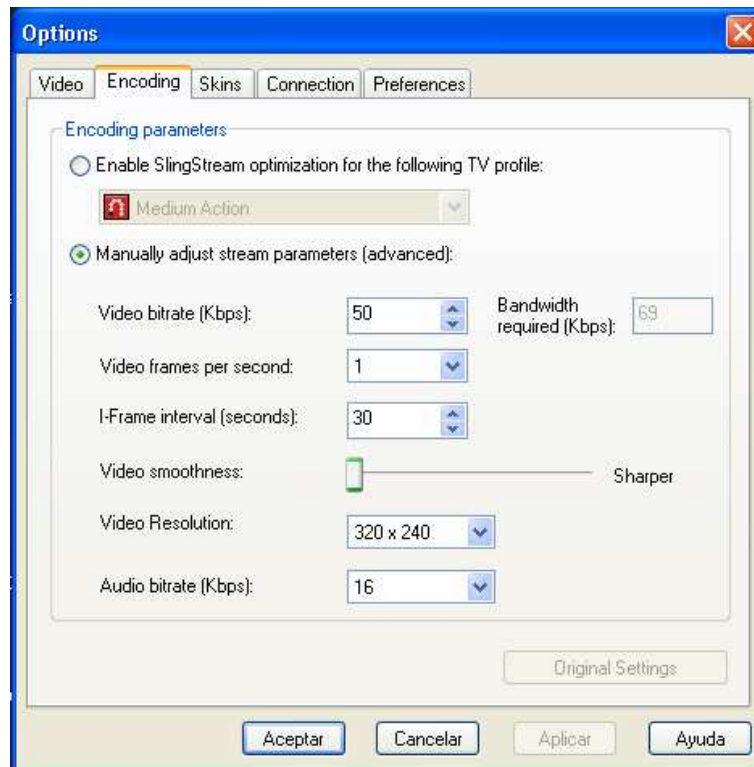


Figura 3.38 Configuración de los parámetros de codificación

### 3.4.2 Pruebas

```

C:\WINDOWS\system32\CMD.exe
Microsoft Windows XP [Versión 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\CASA>ping vigilancia.epn.edu.ec

Haciendo ping a vigilancia.epn.edu.ec [192.188.57.151] con 32 bytes de datos:

Respuesta desde 192.188.57.151: bytes=32 tiempo=395ms TTL=245
Respuesta desde 192.188.57.151: bytes=32 tiempo=405ms TTL=245
Respuesta desde 192.188.57.151: bytes=32 tiempo=388ms TTL=245
Respuesta desde 192.188.57.151: bytes=32 tiempo=385ms TTL=245

Estadísticas de ping para 192.188.57.151:
    Paquetes: enviados = 4, recibidos = 4, perdidos = 0
    (0% perdidos),
    Tiempos aproximados de ida y vuelta en milisegundos:
    Mínimo = 385ms, Máximo = 405ms, Media = 393ms

C:\Documents and Settings\CASA>_

```

Figura 3.39 Ping al equipo remoto

Antes de conectarse remotamente con el sistema de video, se verifica la conectividad, para esto en el modo DOS damos *ping a **vigilancia.epn.edu.ec*** en caso de no tener conectividad, es decir que el ping falle se puede emplear *tracert **vigilancia.epn.edu.ec*** para observar en que parte de la comunicación se pierde Figura 3.40.

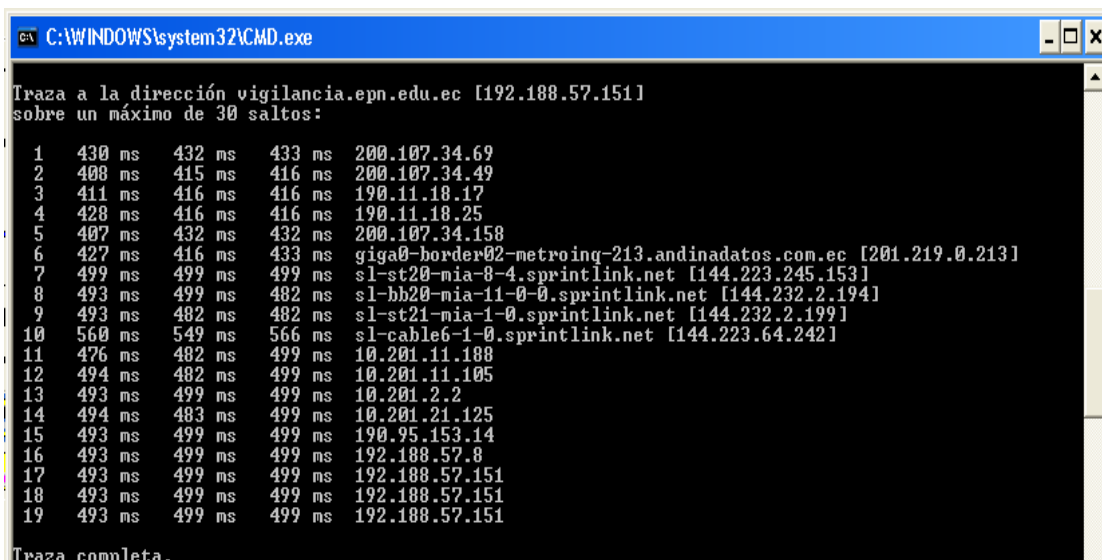


Figura 3.40 Tracert hasta el equipo remoto

El supervisor puede conectarse desde cualquier sitio donde se disponga de una conexión a Internet; si la red donde se encuentra el cliente remoto posee restricciones de tráfico se debe habilitar el tráfico TCP puerto 5001.

La primera conexión que se emplea es una conexión telefónica a 56 Kbps del proveedor Andinanet Figura 3.41, que no muestra una buena calidad de la señal y no permite mejorarla por el ancho de banda de la conexión.

Con una conexión banda ancha se puede establecer mejor respuesta del sistema y buena calidad del video Figura 3.41, esto trae consigo un aumento del tráfico en la conexión de la POLIRED a Internet, en esta prueba se ha empleado una conexión banda ancha ADSL línea digital de abonado asimétrica del proveedor Punto Net de (1024 / 256) Kbps.



Figura 3.41 Prueba con una conexión a 56Kbps



Figura 3.42 Prueba con una conexión ADSL



Figura 3.43 Prueba del sistema de visión nocturna

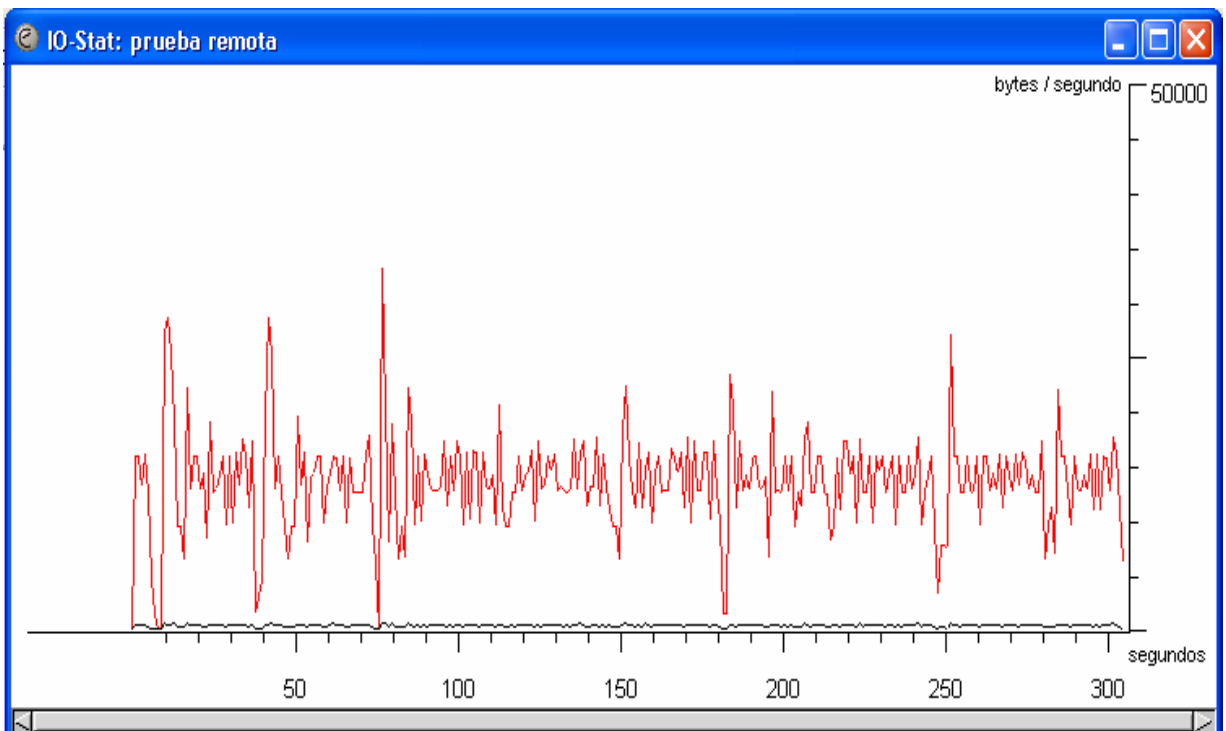


Figura 3.44 Tráfico empleando una conexión banda ancha remota

Las pruebas realizadas muestran que la señal de video se optimiza para alcanzar los parámetros de codificación establecidos y al emplear una conexión banda ancha el sistema puede generar un tráfico mayor, sin embargo esto trae consigo un aumento en el tráfico en los enlaces de conexión a Internet de la Polired, que es lo que se trata de evitar.

### Análisis de los costos

En un proyecto, un factor determinante puede ser el costo de los equipos e infraestructura empleada, en el siguiente análisis se establece una comparación de costos con los sistemas de video alternativos revisados en la sección 3.3.2.

**Cámara Web:** esta alternativa emplea una cámara Web y un ordenador de características: P4, 2.8Mhz, 256 MB de RAM.

EQUIPO	PRECIO
Computador <sup>72</sup>	Desde 450 \$
Cámara Web	Desde 15 \$
	465 \$

Tabla 3.1 Costos empleando una cámara Web

**Servidor de video:** requiere de una cámara de video analógica y el servidor de video, el cual se han tomado como referencia un servidor de marca Axis modelo 241S (Anexo B) que posee una entrada de video analógico.

Un servidor de video implementado con un ordenador que permite transmitir video en tiempo real, necesitará tarjetas para digitalizar, comprimir y almacenar las señales de video a más de un ordenador.

<sup>72</sup> [www.compu-tron.net/computron-contactenos.htm](http://www.compu-tron.net/computron-contactenos.htm)

EQUIPO	PRECIO
<b>Servidor de video</b>	<b>604 \$</b>
<b>Cámara analógica</b>	<b>35 \$</b>
	<b>639 \$</b>

Tabla 3.2 Costos empleando un servidor de video

### Equipo de video embebido Sling Media

EQUIPO	PRECIO
<b>Slingmedia PRO<sup>73</sup></b>	<b>380 (incluye impuestos) \$</b>
<b>Cámara de video analógica<sup>74</sup></b>	<b>35 \$</b>
<b>Canaletas, cable UPT Cat5</b>	<b>20 \$</b>
	<b>435 \$</b>

Tabla 3.3 Costos del sistema de video-vigilancia

Estableciendo una comparación de precios, se aprecia que no hay mayor diferencia de costos entre la solución que emplea la cámara Web y el equipo Slingmedia Pro, sin embargo y a pesar que se puede emplear un computador disponible en el laboratorio se debe destacar la capacidad del sistema de video así como su portabilidad, fácil instalación y mantenimiento; respecto al servidor de video-vigilancia su costo es alto y además posee una sola entrada de video, el sistema de video embebido posee otra entrada de video libre que puede emplearse para otro proceso adicional.

<sup>73</sup> [www.amazon.com](http://www.amazon.com)

<sup>74</sup> <http://www.domolandia.com/tienda/camara-miniatura-espia-audio-color-carcasa-p-259.html>



## CAPÍTULO IV

### 4 CONCLUSIONES Y RECOMENDACIONES

Los sistemas de video se han fusionado con IP (Internet Protocol) para dar origen a muchas aplicaciones, una de ellas la video-vigilancia IP. La convergencia tecnológica hacia IP constituye además una estrategia de las compañías para gozar de los beneficios y aceptación del protocolo.

Los sistemas de video-vigilancia poseen varios componentes encargados de: procesar, almacenar y transmitir video hasta un sitio requerido, con el empleo de IP esto tiene un impacto importante en la capacidad de la red empleada y los equipos que gestionan la información digital.

El sistema de video embebido empleado para la video-vigilancia, transmite los contenidos de video sobre redes IP y para mantener la calidad de la señal de video debe monitorear el ancho de banda de la conexión y establecer una codificación adecuada que se ajuste al ancho de banda disponible.

La video-vigilancia del proceso de curación de aguas contaminadas del CICAM es una aplicación que se le ha dado al sistema de video embebido, por sus características de trabajo sobre redes IP y a pesar de que su diseño que permite transmitir televisión sobre IP, el sistema de video funciona adecuadamente para nuestro propósito.

El tráfico que genera el equipo, no posee un trato especial por parte de la red, por lo que debe establecer mecanismos de control de transmisión a nivel de aplicación; el sistema de video embebido es totalmente escalable ya que permite acceder a la video-vigilancia empleando cualquier tipo de conexión a Internet, sin embargo permite establecer solamente una comunicación de video a la vez; en donde un usuario con la clave de

administrador puede dar por terminada cualquier sesión, por lo que se debe asignar convenientemente las claves de acceso.

Al establecer una sesión de video empleando una conexión banda ancha se recomienda determinar los parámetros de codificación de forma de causar el menor impacto en la Polired.

Se recomienda probar un dispositivo móvil con conexión a Internet en la video-vigilancia, Sling Media proporciona la aplicación para ciertos dispositivos y modelos móviles como: teléfonos, asistentes digitales (PDA) y modelos de Blackberry, la especificación de la conexión a Internet es una conexión banda ancha que próximamente estará disponible en nuestro medio.

## BIBLIOGRAFÍA

### Referencias bibliográficas

- [1] **OPPENHEIM, Alan, SEÑALES Y SISTEMAS, Segunda Edición, Editorial Prentice-Hall, México 1997.**
- [2] **STREMLER, Ferrel, SISTEMAS DE COMUNICACIÓN, Segunda Edición, Editorial Alfaomega, México 1982.**
- [3] **KUROSE, James, REDES DE COMPUTADORAS UN ENFOQUE ESCENDENTE BASADO EN INTERNET. Segunda edición, Editorial Pearson Addison Wesley, Madrid España 2004.**
- [4] **Academia de networking de cisco systems de segundo año CCNA 3 y 4 tercera edición, Pearson educación, España.**
- [5] **Academia de networking de cisco systems de primer año CCNA 1 y 2 tercera edición, Pearson educación, España.**
- [6] **DOUGLAS, Comer. REDES GLOBALES DE INFORMACIÓN CON INTERNET Y TCP/IP, Principios básicos protocolos y arquitecturas, Tercera edición, Editorial PERSON EDUCACION. México 1996.**
- [7] **DOUGLAS, Comer. INTERCONECTIVIDAD DE REDES CON TCP/IP, Diseño e implementación Vol. II, Tercera Edición, Editorial PERSON EDUCACION. México 1996.**
- [8] **BREMNER, Alservati, BIBLIA DE INTRANET, Editorial MCGRAW-HILL.**
- [9] **TANENBAUM, Andrew. REDES DE COMPUTADORAS, Cuarta Edición, Editorial PEARSON PRENTICE HALL.**
- [10] **GARCÍA, Tomas ALTA VELOCIDAD Y CALIDAD DE SERVICIO EN REDES IP, Editorial ALFAOMEGA RA-MA, España.**
- [11] **STALLINGS, William COMUNICACIONES Y REDES DE COMPUTADORAS, Séptima Edición, Editorial PEARSON PRENTICE HALL, Madrid 2004**
- [12] **SAGAMAR Deb, VIDEO DATA MANGEMENT AND INFORMATION RETREIVAL, University of Southern Queensland, Australia.**
- [13] **GRANJA, Eduardo, DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA DE TRANSMISIÓN DE AUDIO Y VIDEO POR INTERNET PARA LA EX-FACULTAD DE INGENIERÍA ELECTRICA, Tesis año 2004.**

Referencias de Internet

- [14] [http://www.ugr.es/~jnieves/Textos/Tema2Psicofisica%20VisionColor%20I%20CO  
Nt%20Gelstalt.pdf](http://www.ugr.es/~jnieves/Textos/Tema2Psicofisica%20VisionColor%20I%20CO<br/>Nt%20Gelstalt.pdf)
- [15] [http://www.rediris.es/media/gt/gt2004\\_2/CapturaVGA.ppt](http://www.rediris.es/media/gt/gt2004_2/CapturaVGA.ppt)
- [16] <http://www.video-computer.com/video.htm>
- [17] [http://www.ebu.ch/dvb\\_articles/dvb\\_tv-history.htm](http://www.ebu.ch/dvb_articles/dvb_tv-history.htm)
- [18] [www.whyfiles.org/163amd\\_eye](http://www.whyfiles.org/163amd_eye)
- [19] [www.fotomundo.com/tecnic/variost/terminolo6.shtml](http://www.fotomundo.com/tecnic/variost/terminolo6.shtml)
- [20] <http://www.alegsa.com.ar/Dic/cif.php>
- [21] <http://es.wikipedia.org/wiki/CIF>
- [22] [www.webelectronica.com.ar/news15/nota08.htm](http://www.webelectronica.com.ar/news15/nota08.htm)
- [23] <http://www.plasma.com/>
- [24] [http://www.plasma.com/classroom/what\\_is\\_tft\\_lcd.htm](http://www.plasma.com/classroom/what_is_tft_lcd.htm)
- [25] <http://www.siemon.com/la/>
- [26] <http://www.erg.abdn.ac.uk/users/gorry/level2dp.pdf> Godred Fairhurst: Digital  
Televisión:
- [27] [http://www.erg.abdn.ac.uk/public\\_html/research/future-net/digital-  
video/index.html](http://www.erg.abdn.ac.uk/public_html/research/future-net/digital-<br/>video/index.html) Godred Fairhurst: MPEG-2 Digital Video
- [28] <http://mpeg.telecomitalialab.com/> Página principal del proyecto OpenH323:
- [29] <http://www.openh323.org>. Interesante fuente de información sobre H.323,  
implementaciones y servicios relacionados (gateways, gatekeepers, sistemas  
de respuesta automatizada, etc.) para Linux y Windows.
- [30] <http://www.idg.es/pcworldtech/index.asp?seccion=comunicaciones>
- [31] [www.flexotex.com](http://www.flexotex.com)
- [32] <http://www.idg.es/pcworldtech/index.asp?seccion=comunicaciones>
- [33] [http://www.siemon.com/la/white\\_papers/SD-03-08-CCTV.asp](http://www.siemon.com/la/white_papers/SD-03-08-CCTV.asp)
- [34] [www.cicamecuador.com](http://www.cicamecuador.com)

## **ANEXOS**

### **ANEXO A:**

**A.1 BANDAS DE FRECUENCIAS EMPLEADAS EN LOS SISTEMAS DE TELEVISIÓN.**

**A.2 HISTORIA DE LA EVOLUCIÓN DE LA TELEVISIÓN.**

### **ANEXO B:**

**CATÁLOGO DE PRODUCTOS DE VIDEO-VIGILANCIA IP.**

### **ANEXO C:**

**HOJAS TÉCNICAS DE LOS EQUIPOS EMPLEADOS EN LA POLIRED.**

### **ANEXO D:**

**BOSQUEJO DE LOS PLANOS ARQUITECTÓNICOS DEL EDIFICIO DEL CICAM.**

**Nota: Por la cantidad de información de los anexos se adjunta un CD que contiene la información completa.**

## **ANEXO A.1**

# **BANDAS DE FRECUENCIAS EMPLEADAS EN LOS SISTEMAS DE TELEVISIÓN**

## **ANEXO A.2**

### **HISTORIA DE LA EVOLUCIÓN DE LA TELEVISIÓN**

## **ANEXO B**

### **CATÁLOGO DE PRODUCTOS DE VIDEO VIGILANCIA IP**



## **ANEXO C**

### **HOJAS TÉCNICAS DE LOS EQUIPOS EMPLEADOS EN LA POLIRED**

## **ANEXO D**

### **BOSQUEJO DE LOS PLANOS ARQUITECTÓNICOS DEL EDIFICIO DEL CICAM**