

**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA**



**“CALIDAD DE SERVICIO EN REDES MULTISERVICIO”**

**INFORME DE SUFICIENCIA**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

**INGENIERO ELECTRÓNICO**

**PRESENTADO POR:**

**JORGE LUIS COVEÑAS CHIROQUE**

**PROMOCIÓN**

**1997 – 2**

**LIMA – PERÚ**

**2002**

*Dedicatoria:*

*Con mucha gratitud y aprecio a  
mis padres, hermanos, a todas  
las personas que me estiman y a  
mi Alma Mater, la UNI.*

# **CALIDAD DE SERVICIO EN REDES MULTISERVICIO**

## SUMARIO

El objetivo de este trabajo es presentar las herramientas tecnológicas actuales para que los servicios que van sobre una red de datos, especialmente la voz telefónica y video, tengan un nivel de calidad conforme a los requerimientos de los usuarios.

Desde hace tiempo, los responsables de comunicaciones de las empresas tienen en mente la posibilidad de utilizar su infraestructura de datos, para el transporte del tráfico de voz interno.

Después de haber constatado que desde una PC con elementos multimedia, es posible realizar llamadas telefónicas a través de Internet, podemos pensar que la telefonía en IP es poco más que un juguete, pues la calidad de voz que obtenemos a través de Internet es muy pobre. No obstante, si en nuestra empresa disponemos de una red de datos que tenga un ancho de banda bastante grande, también podemos pensar en la utilización de esta red para el tráfico de voz entre las distintas sucursales de la empresa. Las ventajas que obtendríamos al utilizar nuestra red para transmitir tanto la voz como los datos son evidentes

## V

- Ahorro de costos de comunicaciones pues las llamadas entre las distintas sucursales de la empresa saldrían gratis.
- Integración de servicios y unificación de estructura.

En los dos primeros capítulos (I y II) se describen las aplicaciones y los requerimientos de estas para que sean transmitidas sobre una red de datos.

En el capítulo III se habla de las técnicas de compresión de voz y se dan las fórmulas para el cálculo de ancho de banda y opciones para saber escoger los codecs adecuados.

Luego, en el capítulo IV se describen los protocolos y estándares desarrollados para voz y multimedia.

En el capítulo V se describe con cierto detalle las diversas herramientas de calidad de servicio desarrollados en la actualidad.

Finalmente, en el capítulo VI se muestra una aplicación sencilla de algunas herramientas de calidad de servicio. Se muestran los resultados de la misma y algunas consideraciones para su implementación.

# ÍNDICE

INTRODUCCIÓN .....	1
<b>CAPITULO I - APLICACIONES MULTIMEDIA .....</b>	<b>2</b>
1.1. Clasificación de las aplicaciones .....	2
1.2. Requerimientos de las aplicaciones .....	4
1.3. Sensibilidad de las aplicaciones.....	9
1.4. Clases de servicio .....	11
1.5. Multimedia.....	11
<b>CAPITULO II - RETARDO EN REDES DE PAQUETES .....</b>	<b>17</b>
2.1. Generalidades .....	17
2.2. Retardos de transmisión.....	19
2.3. Retardos de propagación y conmutación.....	24
2.4. Retardo de colas.....	25
2.5. Pérdida de paquetes .....	26
2.6. Errores en los medios.....	27
<b>CAPITULO III - TÉCNICAS DE COMPRESIÓN DE VOZ .....</b>	<b>29</b>
3.1. Estándares de codificación de voz – CODECS .....	30
3.1.1. G.711 .....	30

<b>CAPITULO V - CALIDAD DE SERVICIO .....</b>	<b>66</b>
5.1. Funciones de edge y core.....	67
5.2. Calidad de servicio relativa y absoluta .....	68
5.3. Tipos de servicio con calidad de servicio .....	69
5.4. Estándares de calidad de servicio .....	69
5.4.1 Resource reservation protocol. – RSVP (RFC 2205).....	69
5.4.2 IP RTP reserve.....	71
5.4.3 Multi-protocol Over ATM – MPOA .....	73
5.4.4 Multi-Protocol Label Switching – MPLS.....	73
5.4.5 Precedencia IP .....	75
5.4.6 Differentiated services .....	77
5.4.7 Policy – based routing – PBR.....	78
5.4.8 Servidor de control de QoS.....	80
5.4.9 802.1p & 802.1Q .....	80
5.4.10 Traffic shaping & policing.....	81
5.4.11 Control de flujo.....	86
5.5. Clases de servicio – COS.....	88
5.6. Service Level Agreement – SLA.....	89
<b>CAPITULO VI - IMPLEMENTACIÓN USANDO ROUTER CISCO ....</b>	<b>91</b>
6.1. Escenario de pruebas .....	91
6.2. Configuración del router Cisco.....	93
6.3. Explicación de la configuración .....	95
6.4. Resultados obtenidos .....	96

## VII

3.1.2.	G.726 .....	30
3.1.3.	G.728 .....	31
3.1.4.	G.729 .....	31
3.1.5.	G.723.1 .....	32
3.2.	MOS – Mean Opinion Score (P.800) .....	32
3.3.	PSQM – Perceptual Speech Quality Measurement (P.861) .....	34
3.4.	Retardos de extremo a extremo .....	36
3.5.	Ancho de banda para voz sobre una red de paquetes .....	38
3.6.	Voice Activity Detection – VAD .....	41
3.7.	Requerimiento de ancho de banda para video .....	42
<b>CAPITULO IV - STACK DE PROTOCOLOS PARA VOZ .....</b>		<b>44</b>
4.1.	Protocolos de señalización.....	45
4.1.1.	ITU –T H.323 .....	46
4.1.2.	SGCP (Simple Gateway Control Protocol) .....	46
4.1.3.	PDC (IP Device Control de Level 3).....	47
4.1.4.	MGCP (Multimedia Gateway Control Protocol).....	47
4.1.5.	SIP (Session Initiation Protocol) .....	48
4.2.	Señalización por canal común N° 7 .....	51
4.3.	Protocolos de transporte .....	53
4.4.	Estándar H.32X.....	58
4.5.	Videoconferencia.....	61
4.6.	Multimedia sobre IP .....	63
4.7.	Integración de redes.....	65



## IX

6.4.1	Saturación del ancho de banda .....	96
6.4.2	Prueba con llamada.....	98
CONCLUSIONES .....		102
BIBLIOGRAFÍA.....		104

## INTRODUCCIÓN

La evolución de las redes de datos ha contribuido a la integración de servicios. En estos días se tiene que las redes de voz y datos se han integrado y para los usuarios es muy importante la calidad del servicio que recibe de la empresa prestadora de servicio. En vista de esto se han desarrollado diversas herramientas y tecnologías que van a hacer que sea posible tener integración de servicios en una sola red sin degradación de la calidad del servicio

Para crear un diseño de red apropiado para las aplicaciones, es importante conocer todas las advertencias y el trabajo interno de la tecnología de internetworking. En este trabajo se explican muchos de los problemas que se encuentran al diseñar una red de datos para tráfico de voz sobre IP y cómo se pueden enfrentar para darle solución.

# CAPÍTULO I

## APLICACIONES MULTIMEDIA

Una aplicación es un programa o grupo de programas (software) que ejecutan una tarea determinada. Es un medio por el cual se usan recursos computacionales para combinar la tecnología con el trabajo o la vida diaria. Por ejemplo una hoja de cálculo o un procesador de texto son aplicaciones, así como un programa contable es también una aplicación. Entre otras tenemos la transferencia de archivos, video conferencia, llamadas telefónicas, correo electrónico, chat, etc,

### 1.1. Clasificación de las aplicaciones

Las aplicaciones se pueden agrupar por su naturaleza como: *Non real – time* y *Real – time*.

a) **Aplicaciones non real – time:** Este tipo de aplicaciones admite retransmisión de paquetes y control de flujo.

Como algunos ejemplos de este tipo de aplicaciones podemos citar los siguientes:

- Teleproceso, telecomando, telesupervisión.
- POS (Point of Sale).

- Correo electrónico (e-mail), Chat.
- Transferencia de archivos, Aplicaciones cliente – servidor.
- Browsing:
  - Bases de datos, Web (hay una tendencia a multimedia con contenido real-time).
- Fax (fax-relay o store and forward).

b) **Aplicaciones real – time:** No admiten retransmisión de paquetes ni control de flujo. A su vez se puede subdividir de la siguiente manera: *Interactivas y No interactivas.*

**b.1) Aplicaciones interactivas:** No admiten gran retardo, ni su variación (JITTER). Hay una respuesta entre dos o más entidades que se estén comunicando.

Ejemplos:

- Voz
  - Telefonía: Si hubiese algún retardo en la transmisión de la voz la conversación se va a ver afectada, llegando en algunos casos a ser ininteligible.
- Video
  - Videoconferencia: Para este caso los recursos de red necesarios para la transmisión de video son mayores.

**b.2) Aplicaciones no interactivas:** Admiten retardo pero no su variación: (JITTER).

Ejemplos:

- Música.
- Video-streaming point – to – point o multicast.
- Broadcasting de voz, música o video.

## 1.2. Requerimientos de las aplicaciones

Como se aprecia en el párrafo anterior las aplicaciones tienen ciertos requerimientos según su naturaleza. Para esto se tienen diversos parámetros que hay que tomar en cuenta cuando se va a diseñar o especificar una red multiservicio para que el usuario no tenga problemas al momento de hacer uso de dichas aplicaciones.

a) **Ancho de banda (BW):** El primer problema de mayor preocupación cuando se diseña una red es la limitación en el recurso de ancho de banda. El ancho banda mide la velocidad de transmisión de una línea de comunicaciones. Esta velocidad, ritmo o flujo de transmisión se mide por la cantidad de datos que se pueden enviar/recibir por unidad de tiempo. Dado que los datos están compuestos por bits, la forma más habitual para medir esta velocidad de transmisión es el número de bits por segundo transmitidos (bps).

El ancho de banda necesario para video depende de la cantidad de: pixels, colores, cuadro por segundo y mecanismo de compresión.

Para el caso de tráfico de voz depende del codec y de las muestras de voz por paquete. Asimismo, el ancho de banda por llamada depende del medio de enlace de datos y del protocolo usado.

- b) **Retardo (delay):** El retardo es la cantidad de tiempo que toma la señal para salir del origen y llegar al destino. El retardo no afecta a la calidad, pero sí a la interacción.

El retardo siempre está presente, sólo que en una conversación telefónica convencional es tan pequeña que el oído humano no lo aprecia.

En los sistemas telefónicos convencionales, el retardo se debe principalmente a la demora en la propagación de la señal, evidente en transmisión vía satélite (250 milisegundos para un salto satelital a un satélite geoestacionario ubicado a 36,000 Km. de altitud).

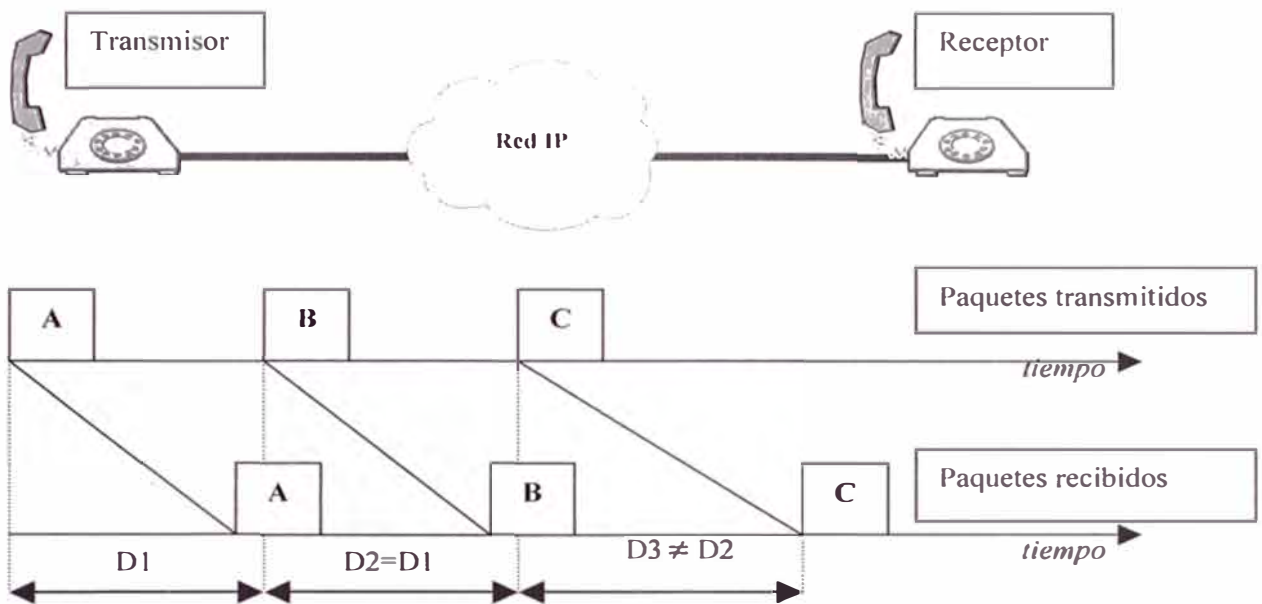
Todo sistema digital tiene retardo debido al tiempo de procesamiento de las muestras digitales, especialmente si el tratamiento implica una compresión y descompresión digital hecha por procesadores digitales de señales (DSPs).

La voz requiere bajo retardo y poca variación del mismo. El tráfico SNA requiere retardo mínimo para evitar el time – out de las sesiones.

- c) **Jitter (variación del retardo):** El jitter es la variación del tiempo de llegada entre paquetes. Es un problema que existe sólo en las redes de paquetes. En un ambiente de voz paquetizada, se espera que el origen envíe los paquetes a intervalos regulares (por ejemplo: cada 20 ms). Estos paquetes de voz pueden verse retardados a través de la red de paquetes y no llegarán en los mismos intervalos regulares de tiempo a la estación receptora. La diferencia entre cuando el paquete es esperado y cuando ha sido recibido es el jitter.

En la figura N° 1 se ilustra el concepto de jitter. Se puede ver que la cantidad de tiempo para enviar los paquetes A y B son iguales ( $D_1=D_2$ ). El paquete C encuentra retardo en la red y por lo tanto es recibido después del tiempo en que es esperado. Nótese que el jitter y el retardo total no es lo mismo, aunque teniendo bastante jitter aumenta el valor del retardo total de la red. Si la red está bien diseñada y se toman las precauciones el jitter no representa ningún problema.

El jitter afecta a la calidad de la voz y el video, es producto del “store and forward” en cada nodo y CPE (Customer Premise Equipment).



**FIGURA N° 1- JITTER**

- d) **Tasa de error de bits:** Los errores de bits ocurren en los sistemas de transmisión digital debido a fallas de cableado, malfuncionamiento de equipos, niveles de señal débil, o varias otras razones. Los errores de bits se caracterizan porque son cambios ocasionales y no periódicos de la señal transmitida (por ejemplo un “0” reemplazado por un “1”) y son usualmente medidos en términos de la tasa de error de bits (BER). Si el BER de un circuito es  $10^{-6}$ , significa que hay un bit errado de cada millón de bits transmitidos.

En una red de paquetes un único bit errado daña todo un paquete. En situaciones con alta tasa de pérdida de bits es muy importante reducir este valor ya que esto afecta a la aplicación. Los errores de bits son detectados, generalmente, con algoritmos de CRC (Cyclic Redundance Check) que confirman la integridad de los datos de un paquete. Por ejemplo, en IPX (Novell) se requiere baja tasa de

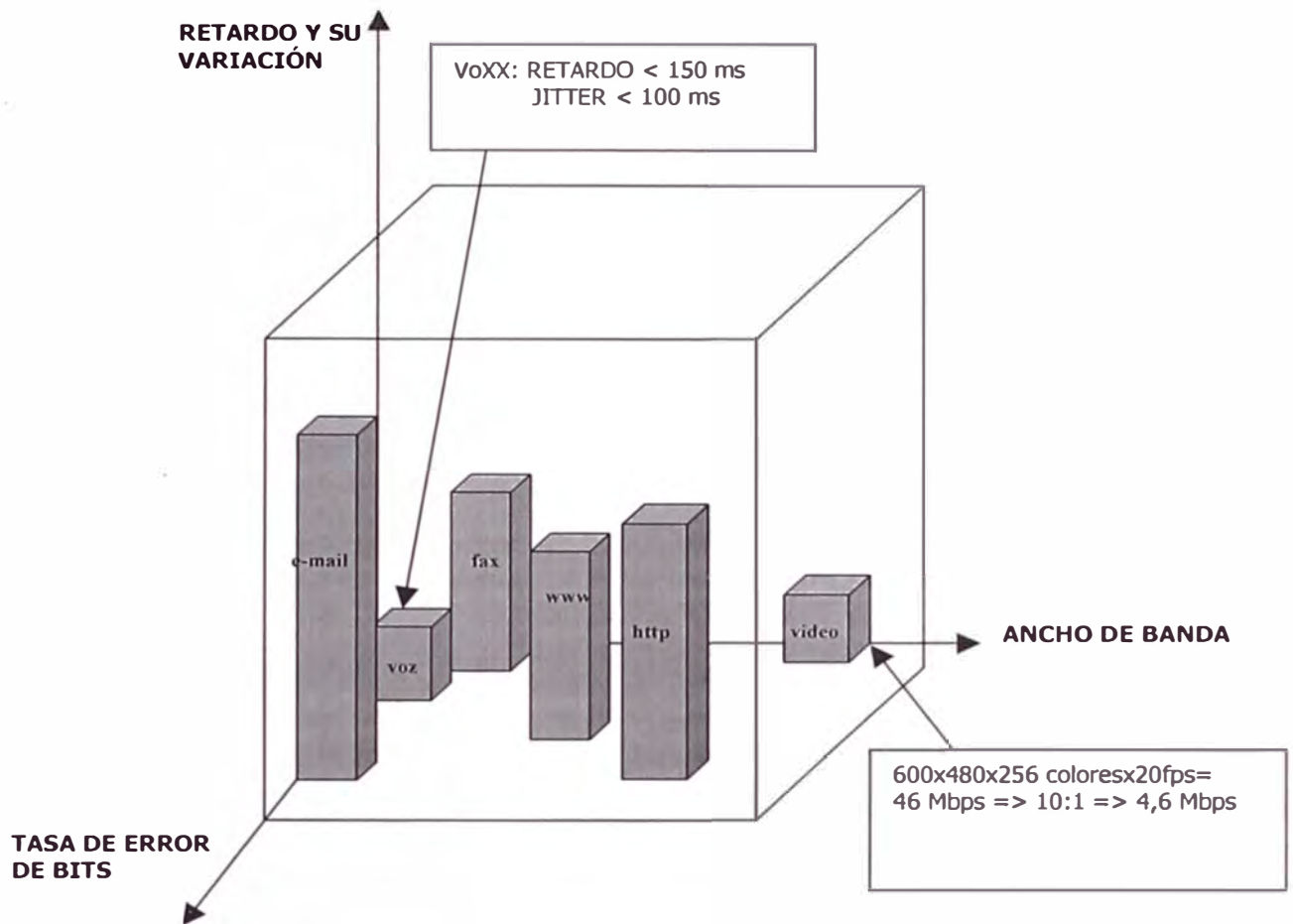


error y para TCP/IP se descarta la ventana (window) completa si pierde un segmento.

Para redes de voz sobre IP, el campo CRC de la cabecera UDP es responsable de detectar errores de bit que pueden producirse por procesamiento de software en la estación receptora. El campo CRC en la cabecera UDP no es generalmente requerida para corregir errores de bit que ocurren a través de una línea de transmisión, debido a que la cabecera de la capa de enlace de datos usualmente incluye un campo CRC para este propósito.

- e) **Prioridad:** La prioridad puede cambiar a lo largo del día, para cada aplicación y para cada empresa en particular, es relativa al manejo del negocio. Para proveer la priorización apropiada sobre una red de datos congestionada, esta debe tener conocimiento de las características de las aplicaciones.
- f) **Disponibilidad:** Depende del contrato de nivel de servicio (SLA – Service Level Agreement). Por ejemplo, 10 minutos de pérdida de servicio: por año = , por mes = 99.997%.

En la figura N° 2 se muestran los requerimientos de las aplicaciones en función de tres parámetros que determinan su comportamiento en la red: tasa de error, ancho de banda y el retardo y su variación.



**FIGURA N° 2 – REQUERIMIENTOS DE LAS APLICACIONES**

### 1.3. Sensibilidad de las aplicaciones

Debido a que las aplicaciones pueden tener diversos requerimientos hay que saber cual de ellos son los que van a afectar su performance. En la tabla N° 1 se tienen los valores de algunos parámetros que afectan a las aplicaciones. Se aprecia que para aplicaciones real – time la sensibilidad al delay, jitter y errores es mayor, lo mismo que los recursos de ancho de banda. Sin embargo las aplicaciones non real – time son insensibles al jitter..

El tráfico multimedia integra varias aplicaciones y requiere que la red respete la sensibilidad de cada una de ellas.

	Non real time		Real time		
	<ul style="list-style-type: none"> <li>• Teleproceso</li> <li>• POS</li> <li>• Chat</li> <li>• Cajeros automáticos</li> <li>• Telecomando</li> <li>• Telecontrol</li> </ul>	<ul style="list-style-type: none"> <li>• LAN to LAN</li> <li>• Internet</li> <li>• e - mail</li> <li>• Browsing</li> <li>• File Transfer</li> <li>• Client/Server</li> </ul>	Voz	Video Conferencia	Video
Throughput	1.2 a 9.6 Kbps	1.2 Kbps a 2 Mbps	4 a 16 Kbps	128 Kbps	33.6 Kbps a 34 Mbps
Delay	0.5 a 2 seg.	0.5 a 60 seg.	< 0.5 seg.	< 0.5 seg.	0.5 a 5 seg.
Jitter	insensible	Insensible	sensible	sensible	sensible
Errores	insensible	Insensible	tolerante	sensible	sensible

**TABLA N° 1 – REQUERIMIENTOS DE LAS APLICACIONES.**

El throughput es el número de bits, caracteres o bloques pasando a través de un sistema de comunicaciones de datos o por una parte de este sistema. Puede variar hasta su valor teórico máximo. Se expresa en unidades de datos por unidad de tiempo. Es, también, la máxima capacidad de un canal o sistema de comunicaciones.

#### 1.4. Clases de servicio

En la tabla N° 2 se tiene una clasificación referencial de servicio por niveles. Este es un ejemplo y puede variar según el proveedor del servicio. El precio de los servicios son mayores a medida que las especificaciones son más estrictas.

Servicio/QoS	Gold/High	Silver/Medium	Bronze/Low
Web browsing	Incoming>6Kbps Latency<500ms	Incoming<=6Kbps or Latency<500ms	Incoming<=2Kbps or Latency<800ms
File transfer	Bandwidth>20Kbps	Bandwidth<=20Kbps	Bandwidth<=5Kbps
e-mail	Bandwidth>100Kbps	Bandwidth<=100Kbps	Bandwidth<=2Kbps
VoIP	Bi-directional>5Kbps Packet Loss<5% Latency<150ms	Bi-directional<=5Kbps or Packet Loss<25% or Latency<700ms	Bi-directional<=2Kbps or Packet Loss>25% or Latency>=700ms
Video	Bi-directional>48Kbps Packet Loss<5% Latency<300ms	Bi-directional<=48Kbps or Packet Loss<=25% or Latency<500ms	Bi-directional<=4Kbps or Packet Loss>25% or Latency>=500ms

**TABLA N° 2 – CLASES DE SERVICIO**

#### 1.5. Multimedia

Como su nombre implica Multimedia es la integración de múltiples formas de media. Multimedia es cualquier combinación de texto, arte gráfico, sonido, animación y vídeo que llega a nosotros por computadora u otros medios electrónicos. Es un tema presentado con lujos de detalles. Cuando conjuga los elementos de multimedia - fotografías y animación deslumbrantes, mezclando sonido, vídeo clips y textos informativos - puede electrizar a su auditorio; y si además le da control interactivo del proceso, quedarán encantado. La multimedia se compone, como ya se describió,

de combinaciones entrelazadas de elementos de texto, arte gráfico, sonido, animación y vídeo.

**a) Formatos de audio**

- **WAV (Waveform Audio):** Es el formato de archivos de audio de Windows (\*.wav). Es audio pre-generado. Los archivos .wav pueden grabarse a 11, 22 y 44 KHz con 8 o 16 bits monoaural o estéreo.
- **MIDI (Musical Instrument Digital Interface):** Interface digital para instrumentos musicales. Es un estándar para transferencia de información entre instrumentos musicales electrónicos y computadoras.
- **MP3 (Moving Picture Expert Group, MPEG-1 Audio Layer 3):** Es un estándar de compresión de audio digital, permite tener archivos pequeños con calidad CD con compresión de 10:1. Esto significa que en lugar de tener 72 minutos de audio en un CD se puede grabar hasta 12 horas de audio (740 minutos).
- **RealAudio (Real Networks):** Es un estándar de facto para audio streaming sobre la World Wide Web (archivos \*.wma). También se usa para video streaming en formato comprimido. Fue desarrollado por Real Networks (Seattle, Washington) y soporta calidad de sonido FM estéreo. Está incluido en las versiones actuales de Netscape Navigator (hasta 7) y Microsoft Internet Explorer (hasta 4.01).

**b) Formatos gráficos**

- **BMP (Bit Mapped):** Formato de archivo de imagen digital de mapa de bits usado en ambiente Windows (archivos \*.bmp)
- **GIF (Graphics Interchange Format):** Es un formato de archivos comprimidos de 8 bits, típicamente para una resolución de 256 colores.
- **JPEG (Joint Photographic Expert Group):** Es un formato de archivo y un algoritmo de compresión. Tiene un formato de compresión de 24 bits.
- **DIB (Device Independent Bitmap):** Es un formato de archivo diseñado para asegurar que los gráficos de mapa de bits creados usando una aplicación puedan ser cargados y visualizados por otra aplicación exactamente en la misma forma en que aparece en la aplicación original.
- **TIF (Tagged Information Format):** Es usado para el intercambio de documentos entre aplicaciones y plataformas de computadoras.
- **Vector Oriented:** Gráficos que consisten de objetos cada uno de los cuales puede ser manipulado separadamente. Los componentes gráficos son calculados, por lo tanto se pueden cambiar de tamaño sin distorsión. Los formatos de archivo incluye AutoCAD DXF, CBM, EPS, HGL, PIC, DRW, WPG.
- **WMF (Windows Metafile Format):** Es un vector gráfico y es editable con muchas aplicaciones basadas en Windows tales como Office.
- **VSD (Visio Standard Drawing):** Es un formato basado en vectores que es editable con el programa Visio.

- **PDF (Portable Document Format):** Es un formato de archivos inventado por Adobe. Combina varios los elementos en un diseño (texto, gráfico, etc). Es independiente de la plataforma, sistema operativo y aplicación, que permite la visualización, impresión y edición de documentos basados en PostScript sin requerir el uso de la aplicación nativa.

Los siguientes son términos que permiten especificar un formato gráfico:

- **Graphic Format:** pixel o vector imagen.
- **Pixel:** Bloque básico para la construcción de una imagen o mapa de bits.
- **Image Resolution:** Es el número de pixels por unidad de área usado en una imagen.
- **Pixel Depth:** El número de bits por imagen. La profundidad de pixels controla el número de tonos de gris que contiene una imagen.

## c) Formatos de video, videoconferencia y fax

### c.1) Video:

- **MPEG-1:** Es un esquema de compresión para video en movimiento total. Su ancho de banda va de 1 a 1.5 Mbps. Produce video de calidad ligeramente inferior a la calidad convencional de VHS. MPEG-1 fue diseñado para proveer una resolución de 352 por 240 pixels a 30 cuadros por segundo.
- **MPEG-2:** Es un estándar de compresión de video seleccionado por MPEG. Esta diseñado para proveer una resolución de 720 por 480 pixels

a 30 cuadros por segundo y 1280 por 720 a 60 cuadros por segundo con audio de calidad CD. Esto es suficiente para la mayoría de los estándares de TV incluyendo NTSC y HDTV. Se usa en DVD-ROMs Puede ocupar un ancho de banda de 6 a 20 Mbps. Puede comprimir dos horas de video en unos pocos Gigabytes

- **MPEG-4:** Integración de video digital, gráficos interactivos y multimedia. El cuerpo de estándar ISO está trabajando actualmente en este formato, no hay MPEG-3. MPEG-4 será basado en el formato de archivos Quick Time.
- **AVI (Audio Video Interleave):** Es un formato de multimedia para Windows, de Microsoft. El audio y video codificado aparecen en segmentos alternados. Este tipo de archivos tiene extensión \* avi. Los clips de video de la World Wide Web están disponibles en este tipo de formato y Quick Time.
- **Indeo 3.2:** Es hardware de compresión de video para computadoras fabricado por Intel. El chip i750 captura video a 32 por 240 pixel a 15 cuadros por segundo y 160 por 120 pixel a 30 cuadros por segundo. Pueden ocupar un ancho de banda de 200 a 300 Kbps.
- **RealVideo (Realnetworks):** Es una tecnología de video streaming, desarrollada por Real Networks, para transmitir video en vivo sobre Internet. Usa una variedad de técnicas de compresión de datos y trabaja con conexiones IP normal asi como también con conexiones IP Multicast. El proceso es espasmódico y algunas veces lento debido a que requiere



gran cantidad de ancho de banda y frecuentemente nos es disponible con comodidad.

- **Quick Time:** Es una tecnología desarrollada por Apple Computers que permite unificación de multimedia (texto, sonido, animación) en un archivo. Es un tipo de formato de reproducción de video y sonido sobre una computadora.

**c.2) Videoconferencia:**

- Netmeeting, H.323, H.324, RTSP

**c.3) Fax:** G3, G4, T.30, T.38

## CAPÍTULO II

### RETARDO EN REDES DE PAQUETES

#### 2.1. Generalidades

El retardo total para transmitir un paquete sobre una red de paquetes depende de diversos factores. Hay tres tipos de retardo: retardo de propagación, retardo de serialización y retardo de manipulación de los datos. El retardo de propagación es causado por la velocidad del medio de transmisión (fibra óptica, cobre, etc.). El retardo de manipulación – también conocido como retardo de procesamiento – define muchas causas diferentes de retardo (paquetización, compresión y conmutación de paquetes) y es causado por los dispositivos que transmiten las tramas sobre la red. El retardo de serialización es la cantidad de tiempo que se toma para colocar un bit o byte en una interface de red para que sea transmitido.

Por lo tanto, el retardo total está dado por la siguiente fórmula:

$$T_{\text{tiempo total}} = T_{\text{transmisión}} + T_{\text{propagación}} + T_{\text{conmutación}} + T_{\text{colas}}$$

Donde:

- $T_{\text{transmisión}}$  = es función de velocidad y tamaño de trama, paquete o celda.

- $T_{propagación}$  = es función del medio de transmisión.
- $T_{conmutación}$  = depende de la tecnología (store & forward o cut-through).
- $T_{colas}$  = es función de la ocupación, prioridad, tamaño y velocidad.

La recomendación ITU-T G.114 especifica que para una buena calidad de voz no debe haber retardo extremo a extremo mayor que 150 ms.

### **Recomendación G.114 de la ITU – T**

“TIEMPO DE TRANSMISIÓN EN UN SENTIDO DENTRO DE LAS CARACTERÍSTICAS GENERALES DE LAS CONEXIONES Y CIRCUITOS TELEFÓNICOS INTERNACIONALES, SISTEMAS Y MEDIOS DE TRANSMISIÓN”, se recomienda como **límites para el tiempo de transmisión extremo a extremo:**

- De 0 a 150 ms: aceptable para la mayoría de las aplicaciones de usuario.
- 150 a 400 ms: aceptable siempre y cuando se conozca su influencia del tiempo de transmisión en la calidad de las aplicaciones de usuario (así, se consideran aceptables conexiones internacionales con saltos de satélite cuyos tiempos de transmisión sean inferiores a 400 ms).
- Por encima de 400 ms: inaceptable a efectos de planificación general de la red; se acepta, sin embargo, que este límite pueda ser rebasado en ciertos casos excepcionales (como excepción, se puede citar por ejemplo: los saltos de satélite dobles inevitables, la utilización de satélites para restaurar rutas terrenas, las interconexiones del servicio fijo por satélite y el servicio celular digital, la videoconferencia por circuitos de satélite o las conexiones

internacionales de gran longitud con dos sistemas celulares digitales conectados mediante facilidades terrenales de larga distancia).

Es decir, el término “tiempo real” se establecería para telefonía, según las recomendaciones de la ITU – T como aquel que no sobrepase los 400 milisegundos de retraso, mas allá de esto no es recomendable. Por debajo de los 150 milisegundos de retraso tenemos una calidad muy buena.

## **2.2. Retardos de transmisión**

En la tabla N° 3 se tienen cálculos de tiempo de retardo para un paquete de datos y uno de voz típicos a diferentes velocidades. Nótese que a bajas velocidades se requiere fragmentar la trama de datos, ya que el retardo de estos es mucho mayor que el de un paquete de voz.

### **2.2.1 Fragmentación**

Fragmentar significa romper los paquetes grandes en otros más pequeños. Se puede llevar a cabo tanto en capa 2 como en capa 3 del modelo de referencia OSI.

La fragmentación es necesaria sobre las tramas de datos en presencia de tramas de voz. Los paquetes de voz son siempre de tamaño pequeño, del orden de 74 bytes. Por ejemplo, si se tuviese una transmisión de paquetes de datos de 1500 bytes y a la vez se realiza una llamada telefónica con paquetes de 74 bytes, el espaciamiento en el

tiempo entre los paquetes de voz sería muy grande y variable. Esto va a afectar notablemente la calidad de la conversación. La fragmentación reduce el jitter a valores tolerables.

	Trama (Bytes)	
	Ethernet MTU 1500	Paquete de voz 74
Velocidad (Kbps)	Tiempo (ms)	Tiempo (ms)
64	<b>187.50</b>	4.00
128	<b>93.75</b>	2.00
256	<b>46.88</b>	1.00
512	<b>23.44</b>	0.50
1024	11.70	0.25
2048	5.88	0.13
34000	0.35	0.01
155000	0.08	0.00
622000	0.02	0.00

**TABLA N° 3 – RETARDOS DE TRANSMISIÓN SEGÚN EL ANCHO DE BANDA Y TAMAÑO DEL PAQUETE**

Como se puede apreciar en la figura N° 3, si se fragmentan los paquetes de datos y se establece una política para la transmisión de la voz en presencia de datos se van a tener resultados satisfactorios para estos últimos. Esto significa que la fragmentación misma no es suficiente para remover los problemas de latencia en circuitos de baja velocidad. El router también debe ser hábil para el manejo de colas basado en los fragmentos en lugar del paquete original.

La fragmentación es necesaria en enlaces de baja velocidad, menores a 512 Kbps. Para enlaces mayores a 512 Kbps no es necesario realizar la fragmentación. Si el

extremo remoto es de baja velocidad, debe fragmentarse igual. La red pasa las tramas sin modificar la longitud. El que fragmenta es el CPE y debe hacerlo en ambos extremos de la red.

En Frame Relay se usan FRF.11/12 y en HDLC se usa MLPPP. La fragmentación debe ser dinámica cuando hay paquetes de voz. En las interfaces que no soportan FRF.12 debe configurarse el MTU (Maximum Transmit Unit) del CPE menor que el MTU de la red. Los paquetes provenientes de Internet no pueden fragmentarse.

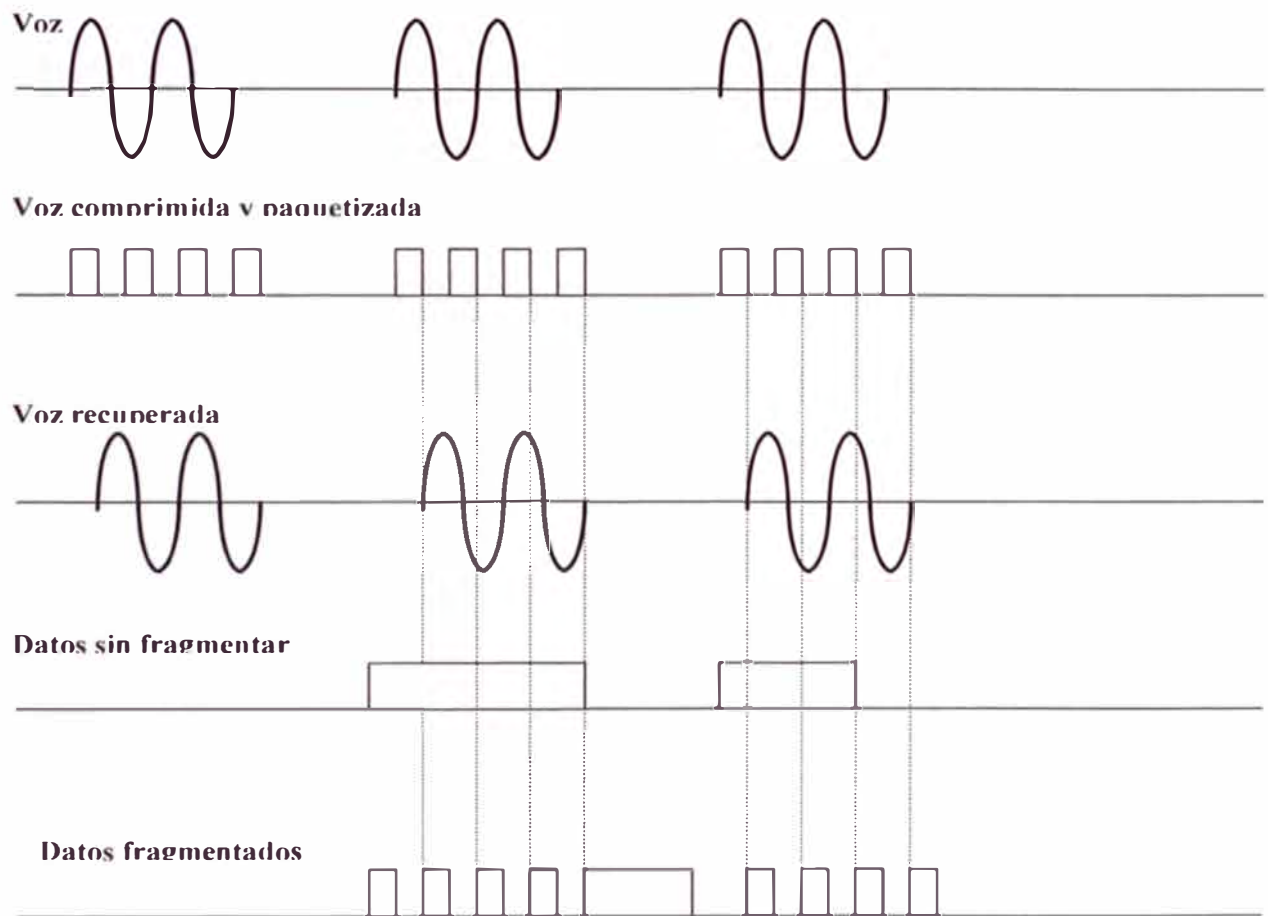


FIGURA N° 3 - FRAGMENTACIÓN DE PAQUETES

### **2.2.1.1 FRF.11 – Voice over Frame Relay**

Transporte de voz comprimida en el payload de una trama Frame Relay. Soporta una variedad de algoritmos de compresión de voz. Permite utilización efectiva de conexiones de baja velocidad, multiplexado de hasta de 255 sub-canales en un DLCI, soporte de múltiples voice payloads dentro de una trama y soporte de sub-canales de datos en DLCIs multiplexados.

### **2.2.1.2 FRF.12 – Frame Relay Fragmentation**

Es una especificación del Frame Relay Forum Technical Committee. Permite compartir datos real-time y non-real-time en el mismo enlace Frame Relay. La fragmentación ocurre en la capa de enlace de datos, por lo que las capas superiores no se enteran de la fragmentación. Permite la fragmentación de tramas de todos los formatos. Define un procedimiento de fragmentación que puede ser usado por otros protocolos, tales como FRF.11. Define tres modelos de fragmentación, los que comparten un procedimiento común de fragmentación.

FRF.12 también especifica la interoperabilidad entre Frame Relay y ATM. Esto permite que los paquetes fragmentados por FRF.12 sobre el lado Frame Relay del circuito sean reensamblados en el lado ATM del circuito.

### 2.2.1.3 Fragmentación, compresión y encriptación

La encriptación DES (Data Encryption Standard) usa bloques de 64 bits (8 bytes). La voz no puede encriptarse con DES. No puede fragmentarse algo encriptado. Asimismo, los datos a comprimir no deben estar comprimidos. Por lo tanto, debe comprimirse antes de encriptarse.

La compresión no ayuda con tráfico “liviano” o interactivo con largos tiempos muertos. La compresión de la cabecera IP impide rutear o priorizar por lo que debe descomprimirse en cada salto.

El ATM hace SAR (Segmentation And Reassembling), por lo que los routers necesitan rearmar el paquete para poder rutearlo.

La fragmentación produce el efecto “pipelining” reduciendo el tiempo total de transmisión en paquetes grandes y causando que la performance de la estación receptora se degrade ya que debe manipular muchos paquetes pequeños en un lugar de un solo paquete grande. Por ejemplo, si se tiene un paquete de 1500 bytes con una cabecera de 40 bytes y se fragmenta usando tramas de 100 bytes, se obtiene 14 paquetes de 100 bytes y uno de 60. Ahora, se necesita colocar la cabecera de 40 bytes a cada fragmento, por lo tanto la cabecera se incrementa a 600 bytes.



### 2.3. Retardos de propagación y conmutación

Cuando una señal digital pasa a través de un circuito, siempre experimenta un retardo temporal llamado tiempo de retardo de propagación. Este tiempo es muy importante porque limita la frecuencia máxima a la que es posible trabajar. Los retardos de propagación dependen del medio de transmisión. No pueden reducirse ni eliminarse ya que dependen de la naturaleza del medio. A continuación se tienen valores típicos de retardo de propagación para algunos medios de transmisión:

- Satélite (geoestacionario): 250 ms por salto.
- Cable: 4 ms por Km.
- Fibra: 5 us por Km.

Los tiempos de conmutación depende de la implementación (hardware o software), pero en los switches modernos están por debajo del milisegundo.

Salvo en satélite, los retardos de propagación son despreciables. En la transmisión satelital se tiene 250 ms para alcanzar el satélite y otros 250 ms para el retorno a tierra, lo cual hace un tiempo total de 500 ms. Aunque, en este caso, se pase la recomendación G.114 del ITU-T se pueden establecer comunicaciones de voz sobre este medio, pero los usuarios aceptan su uso.

## 2.4. Retardo de colas

Una red basada en paquetes puede experimentar retardo por otras razones. Dos de estas son el tiempo necesario para mover un paquete a la cola de salida (conmutación de paquetes) y retardo de encolamiento (queueing delay). Cuando los paquetes se mantienen en una cola debido a congestión sobre una interface de salida el resultado es retardo de colas.

El retardo de colas ocurre cuando se envían mas paquetes que los que una interface puede manipular sobre un intervalo dado. Las colas de espera se distribuyen en cada nodo de acceso y en los nodos internos de la red. Las colas de mayor prioridad tienen menor retardo.

El tiempo de retardo de colas (ver figura N° 4), es dependiente de la ocupación del recurso, la distribución del tráfico, la velocidad del enlace y el tamaño de las celdas, tramas o paquetes.

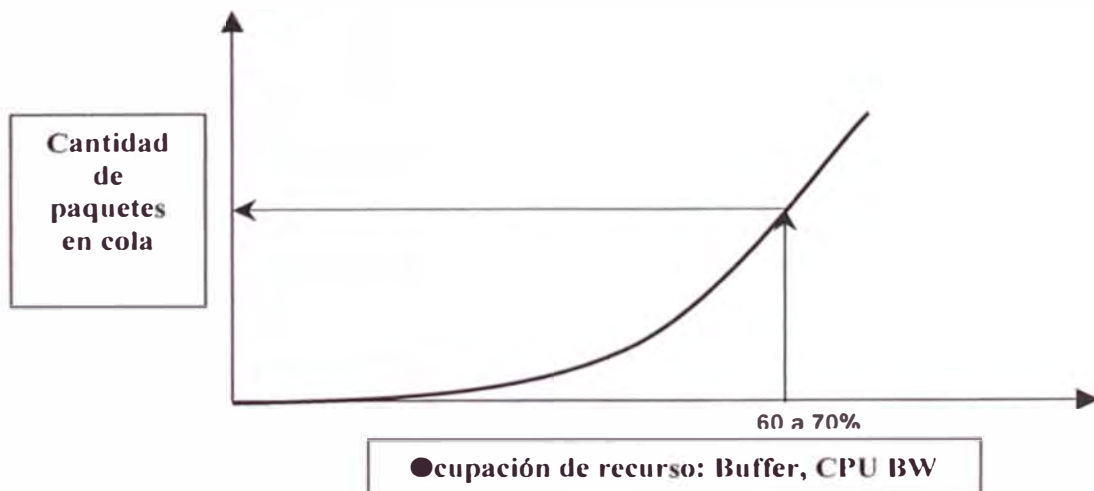


FIGURA N° 4 - RETARDO DE COLAS

## 2.5. Pérdida de paquetes

La pérdida de paquetes en una red de datos es algo común y esperado. Muchos protocolos de datos usan la pérdida de paquetes para conocer las condiciones de la red y pueden reducir los paquetes que se envían.

Cuando se tiene tráfico crucial sobre una red de datos es importante conocer la cantidad de pérdida de paquetes en esta red.

Con protocolos tales como SNA que no toleran pérdida de paquetes se necesita una red con muy buen diseño que pueda priorizar los datos sensibles al tiempo delante de los datos que pueden manipular el retardo y pérdida de paquetes.

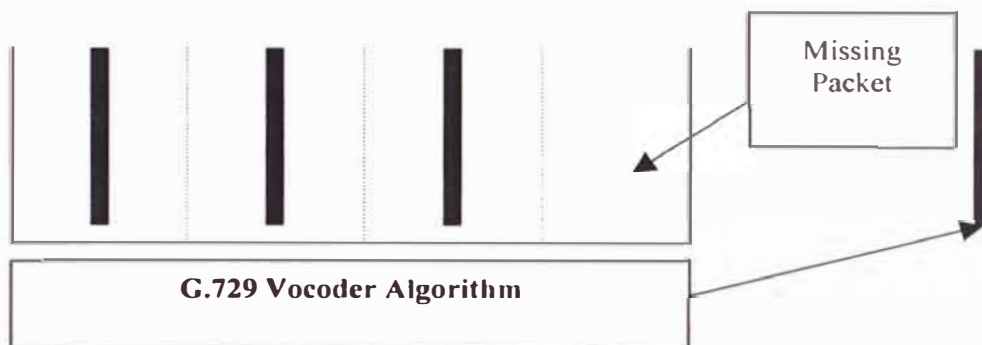
Cuando se transmite voz en una red de datos es importante construir una red que pueda transportar exitosamente la voz en una manera confiable en el tiempo. Además, es útil usar un mecanismo para hacer que la voz pueda soportar un poco de pérdida de paquetes. En todo caso, se debe mantener la pérdida de paquetes a un valor mínimo.

Por ejemplo, la implementación de Voz sobre IP de Cisco (figura N° 5) hace lo siguiente: si un paquete de voz no es recibido durante el periodo de tiempo esperado (jitter buffer) se asume que se ha perdido y el paquete se reemplaza con el último recibido y se evita un hueco de silencio. Debido a que se pierde solo 20 ms de voz, el

oyente promedio no va a notar la diferencia en la calidad de voz. G.729 tolera hasta 5% de pérdida de paquetes. Esto hay que tomar en cuenta en el diseño de la red.

La pérdida de paquetes se produce por varias causas:

- **Errores de transmisión** (poco probable).
- **Descarte por congestión** (evitable con QoS).
- **Descarte por time – out** (delay excesivo), afecta fundamentalmente al tráfico real time y las aplicaciones críticas (evitable con QoS, diseño y control de la red).



**FIGURA N° 5 - PÉRDIDA DE PAQUETES**

## 2.6. Errores en los medios

La tasa de error de cada medio condiciona al protocolo que se utiliza para transportar los datos.

- Línea telefónica conmutada.
- Celular.
- Radio.
- Línea digital (cobre)

- Satélite.
- Cable coaxial.
- Fibra óptica.

El tráfico real-time no admite retransmisión, pero sí corrección de errores.

A continuación se muestra una tabla comparativa de los medios de transmisión.

<b>Medio de transmisión</b>	<b>Velocidad máxima</b>	<b>Retardo (ms)</b>	<b>Tasa de error de bits</b>	<b>Disponibilidad (%)</b>
Línea conmutada	56Kbps	0	$1 * 10^{-5}$	99.5
Línea digital	1920Kbps	10	$1 * 10^{-7}$	99.7
Satélite	2Mbps	240	$1 * 10^{-7}$	99.7
Radio digital	155Mbps	0	$1 * 10^{-8}$	99.9
Coaxial	155Mbps	0	$1 * 10^{-8}$	99.99
Fibra óptica	2.4Gbps	0	$1 * 10^{-8}$	99.99

**TABLA N° 4 – COMPARACIÓN DE LOS MEDIOS DE TRANSMISIÓN**

Se puede apreciar el retardo natural para los enlaces satelitales de 240 milisegundos para un salto satelital y la gran velocidad de propagación de las señales por fibra óptica. Asimismo, vemos que la disponibilidad de una línea conmutada es baja.

## CAPÍTULO III

### TÉCNICAS DE COMPRESIÓN DE VOZ

Dos variantes básicas del PCM de 64 Kbps se usan comúnmente: ley  $\mu$  y ley A. Los métodos son similares en que ellos usan compresión logarítmica para tener calidad de PCM lineal de 12 o 13 bits con 8 bits, pero son diferentes en detalles de compresión relativamente menores (ley  $\mu$  tiene una ligera ventaja en bajo nivel, performance señal a ruido).

Otro método de compresión usado frecuentemente es el ADPCM (Adaptive Differential Pulse Code Modulation). Una instancia comúnmente usada de ADPCM es el ITU-T G.726, que codifica usando muestras de 4 bits dando una tasa de transmisión de 32 Kbps. En contraste con PCM los 4 bits no codifican directamente la amplitud de la señal sino las diferencias en amplitud así como la razón de cambio, empleando algunos métodos de predicción lineal rudimentarias.

PCM y ADPCM son ejemplos de codificación de formas de onda – técnicas de compresión que explotan características redundantes de la forma de onda misma. Se han desarrollado nuevas técnicas de compresión en los últimos 10 o 15 años que explotan el conocimiento de las características fuente de la generación de voz. Estas técnicas emplean procedimientos de procesamiento de señales que comprimen la voz

enviando sólo información paramétrica simplificada acerca de la excitación de voz original requiriendo menos ancho de banda para transmitir esta información.

Estas técnicas pueden ser agrupadas generalmente como codecs e incluyen variaciones tales como LPC (Linear Predictive Coding), CELP (Codec Excited Linear Predictive Compression) y MP-MLQ (Multipulse, Multilevel Quantization).

### **3.1. Estándares de codificación de voz – CODECS**

El ITU-T estandariza los esquemas de codificación CELP, MP-MLQ PCM y ADPCM en las recomendaciones serie G. Los estándares de codificación de voz más populares para telefonía y voz paquetizada son:

#### **3.1.1. G.711**

Describe la técnica de codificación de voz PCM de 64 Kbps. La voz codificada con G.711 está en el formato correcto para envío de voz digital en la red telefónica pública o a través de centrales telefónicas – PBXs (Private Branch eXchanges).

#### **3.1.2. G.726**

Describe codificación ADPCM a 40, 32, 24 y 16 Kbps; también se puede intercambiar voz codificada en ADPCM con la red pública o redes de PBX.

### 3.1.3. G.728

Describe una variación de la técnica de compresión CELP de bajo retardo a una tasa de bits de 16Kbps.

### 3.1.4. G.729

Es el estándar más utilizado en las redes de voz paquetizada. Provee buena calidad de voz como el ADPCM de 32 Kbps. Usa el algoritmo CS-ACELP (Conjugated Structure-Algebraic Code Excited Linear Prediction). Toma 80 muestras de PCM lineal de 16 bits (10ms). Las codifica en 10 bytes. Agrega 5 ms de “look – ahead”, lo cual da un retardo total de  $10+5 = 15$  ms. Genera un paquete de 20 bytes cada 20 ms (50 pps). Lo que da una tasa de 8 Kbps (codec bit rate). A esto hay que agregarle el encabezado del protocolo de transporte.

Anexo A: Algoritmo de complejidad reducida (ocupa menos CPU)

Anexo B: Esquemas de compresión de silencios compatible con V.70

Anexo C: Punto flotante

Anexo D: CS-ACELP a 6.4 Kbps

Anexo E: CS-ACELP a 11.8 kbps



### 3.1.5. G.723.1

Describe una técnica de compresión que se puede usar para comprimir voz u otra señal de audio componente de un servicio multimedia a una baja tasa de bits, como parte de la familia H.324.

Se tienen dos variantes para este codec:

#### **MP-MLQ. Multipulse, Multilevel Quantization**

- 6.3 Kbps payload de 24 bytes, provee mayor calidad de voz.

#### **ACELP: Algebraic Code Excited Linear Prediction**

- 5.3 kbps payload de 20 bytes, provee buena calidad de voz y

Toma 240 muestras de PCM lineal de 16 bit 8 Khz. (30ms). Agrega un “look ahead” de 7.5 ms, retardo total:  $30+7.5=37.5$ ms. Genera una trama de 24/20 bytes de cada 30 ms (33.33pps). Pasa el DTMF mejor que el G.279.

Anexo A: Esquemas de compresión de silencios

Anexo B: Punto flotante

Anexo C: Wireless.

### 3.2. MOS – Mean Opinion Score (P.800)

Se puede probar la calidad de voz en dos formas: subjetivamente y objetivamente.

Los seres humanos ejecutan prueba de calidad en forma subjetiva, mientras que las computadoras hacen la prueba de calidad de voz en forma objetiva.

Los codecs son desarrollados y sintonizados basados en mediciones subjetivas de calidad de voz. Las mediciones estándares de calidad, tales como la distorsión armónica total, relación señal a ruido, no correlaciona bien con la calidad de voz percibida por el oído humano, el cual es en el fondo la meta de las técnicas de compresión de voz.

Una referencia subjetiva común para cuantificar la performance de los codecs de voz es el MOS (Mean Opinion Score). Es el resultado de una prueba de clasificación categórica absoluta (ACR – Absolute Category Ranking). Las pruebas de MOS se hacen a un grupo de individuos receptores. Debido a que la calidad de voz y sonido, en general, es subjetivo para los oyentes, es importante tener un amplio rango de oyentes y material de muestra cuando se hace una prueba de MOS. Los oyentes califican la muestra en un rango de 1 (mala) a 5 (excelente). Los resultados son promediados para obtener el valor del MOS.

Las pruebas de MOS también se usan para comparar como trabaja un codec particular bajo circunstancias variables, incluyendo diferentes niveles de ruido de fondo, múltiples codificaciones y decodificaciones y así.

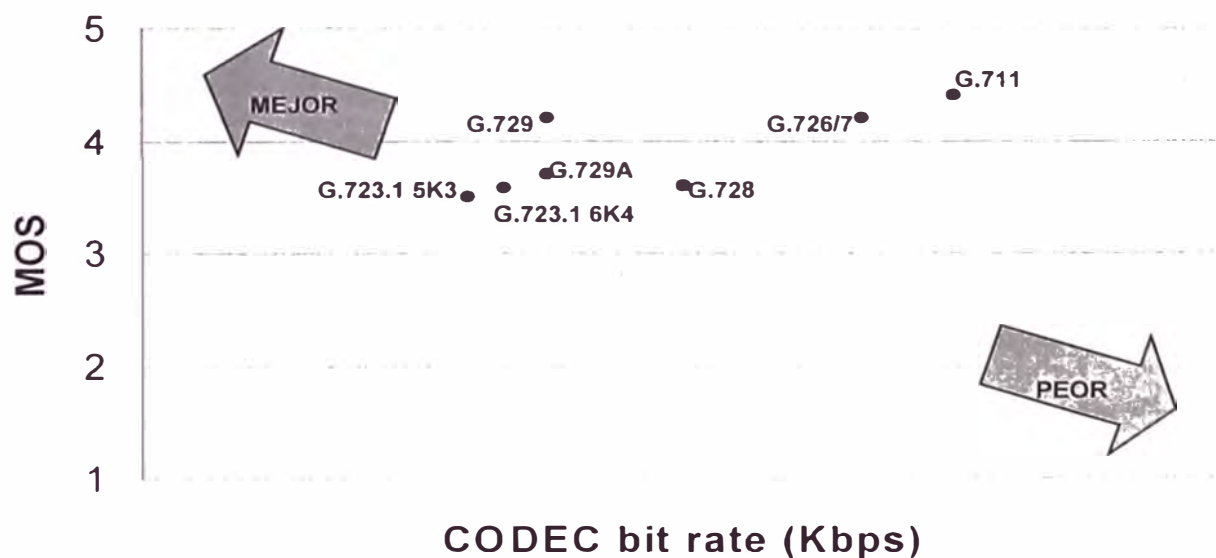
El MOS es un valor subjetivo, no indica una calidad en términos absolutos. Depende de la implementación.

En la tabla N° 5 se tienen valores de MOS para varios codecs ITU-T. Esta tabla muestra la relación entre varios codificadores de baja tasa de bits y el estándar PCM.

Método de compresión	Bit rate (Kbps)	Tamaño de la muestra (ms)	MOS	Retardo CODEC (ms)	Uso de CPU	Observaciones
G.711 PCM	64	0.125	4.1	0.75	-	Telefonía digital
G.726 ADPCM	32	0.125	3.85	1	Bajo	Telefonía digital comprimida
G.728 LD-CELP	15	0.625	3.61	bajo	muy alto	
G.729 CS-ACELP	8	10	3.92	15	alto	VoIP/FR/ATM Netmeeting
G.729a CS-ACELP	8	10	3.7	15	alto	VoIP/FR/ATM Netmeeting
G.723.1 MP-MLQ	6.3	30	3.9	37.5	moderado	VoIP/FR/ATM Netmeeting
G.723.1 ACELP	5.3	30	3.65	37.5	moderado	VoIP/FR/ATM Netmeeting

**TABLA N° 5 – MEAN OPINION SCORE**

En la figura N° 6 se tiene una comparación gráfica de los codecs en función del MOS y el codec bit rate.



**FIGURA N° 6 – COMPARACIÓN GRÁFICA DE LOS CODECS EN FUNCIÓN DEL CODEC BIT RATE Y MOS**

### 3.3. PSQM – Perceptual Speech Quality Measurement (P.861)

Aunque el MOS es un método subjetivo para determinar la calidad de voz, no es el único método para hacerlo. El ITU-T emitió la recomendación P.861, mediante el cual se puede determinar objetivamente la calidad de voz usando PSQM (Perceptual Speech Quality Measurement). Intenta automatizar el proceso de MOS.

PSQM tiene muchos inconvenientes cuando se usa con codificadores de voz (vocoders). Un inconveniente es que el oído de la “máquina” o PSQM no percibe el audio como lo hace el oído humano. El modelo PSQM simula las propiedades físicas del oído humano. El algoritmo PSQM añade ruido durante su procesamiento para simular un ambiente telefónico normal. En términos de leyes humanas, una persona puede decir que es lo que está percibiendo como alta calidad de voz, lo cual no lo puede hacer una computadora. Además, el PSQM se desarrolló para “escuchar” los deterioros causados por la compresión y descompresión y no la pérdida de paquetes o jitter.

Produce resultados consistentes con una prueba subjetiva de un codec bajo ciertas condiciones, pero no se verifica para muchas condiciones presentes en una red viva. El uso de PSQM es conveniente para medición de performance de un codec en un laboratorio, y puede ser sintonizado para análisis de calidad de voz en una red VoIP (luego de capturar algunas muestras de sonido), pero es difícil de implementar para evaluaciones de una red en tiempo real.

El ITU – T está trabajando en el borrador de la recomendación P.862, que especifica PESQ (Perceptual Evaluation of Speech Quality). PESQ realiza buena predicción de calidad de voz en una amplia variedad de aplicaciones, incluyendo voz sobre IP en redes reales. Cuando se estandarice el P.862 reemplazará al P.861.

Se hicieron cálculos del ancho de banda para algunos CODECs más usados teniendo en cuenta el medio de transmisión. Los resultados están en la tabla N° 6.

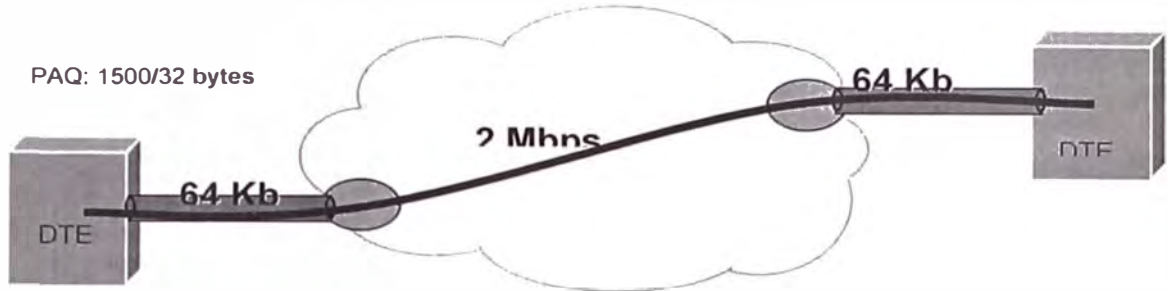
Needed BW (Kbps)		VoIP								VoFR	VoATM
		RTP				CRTP					
Link layer		Ethernet	PPP	FR	ATM	Ethernet	PPP	FR	ATM		
Link overhead		14	6	6	5	14	6	6	5	6	5
Header IP		20	20	20	20	2 a 4				0	0
Header UP		8	8	8	8						
Header RTP		12	12	12	12						
CODEC	Payload										
G.723.1 @ 5,3Kbps	20	19,6	17,5	17,5	28,1	10,1	7,95	7,95	13	6,69	14,0
G.723.1 @ 6,4Kbps	24	20,8	18,7	18,7	28,3	11,2	9,07	9,07	14,1	8	14,1
G-729 @ 8Kbps	30	29,8	26,4	26,4	42,2	15,2	12	12	21,2	10,4	21,2

**TABLA N° 6 - COMPARACIÓN DEL OVERHEAD**

### 3.4. Retardos de extremo a extremo

En la figura N° 7 se ilustran los diversos tipos de retardo que sufre un paquete de datos y uno de voz al ser transmitidos por una red de datos, de extremo a extremo. En la parte de acceso de la red la velocidad es de 64 Kbps y en el core se dispone de 2Mbps. Se considera un paquete de datos de 1500 bytes de longitud, el cual es el

valor máximo que puede transferirse por una interface Ethernet. Asimismo, se ha considerado un paquete de voz de 32 bytes y el codec para voz es el G.723.



T trans 190/4 6/0,13 190/4  
 T cola(2 paq) 380/8 12/0,26 380/8  
 Codec (G.723) 37,5  
 T total paquete de 1500 bytes: 772 ms  
 T total paquete de 32 bytes: 53,9 ms

**FIGURA N° 7 – COMPARACIÓN DE RETARDO EXTREMO A EXTREMO EN UNA RED DE PAQUETES PARA VOZ Y DATOS**

En la tabla N° 7 se tiene la distribución de los retardos para un paquete de voz usando dos tipos de CODEC (G.723.1 y G.729). Se puede apreciar que con G.729 se obtiene menor retardo total, el cual es debido básicamente al retardo de codificación y se está dando un margen razonable para el retardo en la red.

Causa del retardo	G.723.1	G.729
Codificación	37.5	15
Cola CPE origen	1 a 10	1 a 10
Transmisión	1 a 4	1 a 4
Red	20 a 200	20 a 200
Transmisión	1 a 4	1 a 4
Cola CPE destino	1 a 10	1 a 10
Buffer Jitter	10 a 100	10 a 100
Decodificación	10	10
<b>Retardo total</b>	<b>81.5 a 271.5</b>	<b>59 a 253</b>

**TABLA N° 7 – COMPARACIÓN DE RETARDOS PARA DOS TIPOS DE CODEC**

### 3.5. Ancho de banda para voz sobre una red de paquetes

El ancho de banda para una llamada de voz en una red de paquetes se puede calcular con la siguiente fórmula:

$$\text{Bandwidth} = (\text{Voice packet size}) * \text{pps}$$

Donde:

- $\text{Voice packet size} = (\text{layer 2 MLPPP or FRF.12 header}) + (\text{IP/UDP/RTP Header}) + (\text{voice payload})$
- $\text{Packets per second (pps)} = \text{codec bit rate} / \text{voice payload size}$

#### 3.5.1. Ejemplo de cálculo de ancho de banda

Para una llamada de voz sobre IP usando el CODEC G729r8 (8Kbps de CODEC bit rate), con su payload por defecto de 20 bytes (160 bits) y en medio Ethernet, tenemos:

- $\text{voice packet size (bytes)} = (\text{layer 2 header: 14 bytes}) + (\text{IP/UDP/RTP header: 40 bytes}) + (\text{voice payload: 20 bytes}) = 74 \text{ bytes.}$
- $\text{voice packet size (bits)} = (74 \text{ bytes}) * 8 \text{ bits per byte} = 592.$
- $\text{voice packets per second (pps)} = (8\text{Kbps codec Bit Rate}) / (160 \text{ bits}) = 50 \text{ pps.}$
- $\text{bandwidth per call} = \text{voice packet size}(592 \text{ bits}) * 50 \text{ pps.}$

Por lo tanto: *Ancho de Banda por llamada = 29,6 Kbps.*

### 3.5.2 Técnica de Bundling

Debido al tamaño de la cabecera de los paquetes de voz sobre IP (54 bytes) en comparación al payload por defecto (20 bytes) es necesario optimizar la transmisión de dichos paquetes. Esto se hace posible mediante la técnica de *Bundling*, el cual consiste en aumentar el tamaño del *payload*. El *bundling* puede mejorar un 40% de la eficiencia.

En la tabla N° 8 se tiene un cálculo del ancho de banda luego de agrupar varios paquetes de voz en una sola trama a enviar para el codec G.729. Hay que notar que se tiene que llegar a un compromiso entre el retardo y el ahorro de ancho de banda, ya que la cantidad de paquetes por segundo aumenta a medida que se hace el *bundling*. La desventaja de aumentar el tamaño del payload es que se incrementa el retardo de la data, haciendo que la conversación pueda verse afectada. Por tal motivo, se debe elegir un valor que permita un equilibrio entre eficiencia y retardo.

CODEC		Kbps	PAYLOAD	PPS	BW (bps)
G.729	1	8	20	50	29600
G.729	2	8	40	25	18800
G.729	3	8	60	16.7	15200
G.729	4	8	80	12.5	13400

**TABLA N° 8 – CÁLCULO TEÓRICO DE APLICACIÓN DE LA TÉCNICA DE BUNDLING**

Por ejemplo, para el caso de una red con enlace satelital, estas tienen un retardo inherente de 500 ms. Las pérdidas de paquetes son altamente controladas debido al



costo del ancho de banda. En este tipo de red es conveniente hacer uso del bundling y se recomienda usar el codec G.729 con payload size de 80 bytes.

Para configurar el bundling en un router cisco se aplica la siguiente línea bajo el dial-peer tipo VoIP:

```
Router#confiugte terminal
```

```
Router(config)#dial-peer 10 voip
```

```
Router(config-dial-peer)#codec g729r8 bytes [payload size]
```

```
Router(config-dial-peer)#end
```

### 3.5.3 CODECS y su ancho de banda

Cada codec tiene diferentes requerimientos de ancho de banda según el tipo de red sobre el que se transmiten los paquetes y según el tipo de enlace de datos (ver figura N° 8).

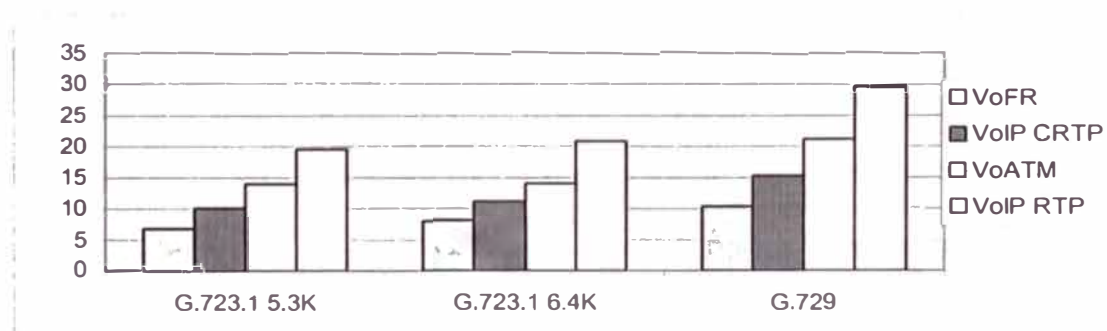


FIGURA N° 8 – ANCHO DE BANDA DE CODECS SEGÚN EL TIPO DE RED

Puede mejorarse el overhead haciendo bundling donde cada trama puede llevar dos o más paquetes de voz. Con la detección de silencio se mejora un 50% la eficiencia respecto del PCM.

### 3.6. Voice Activity Detection – VAD

En conversaciones de voz normalmente algunas personas hablan y otras escuchan. Las redes de comunicación de voz son bidireccionales y tienen un ancho de banda sin importar quien de los interlocutores está hablando. Esto significa que en una conversación normal, al menos el 50% del ancho de banda total se gasta. La cantidad de ancho de banda gastado puede ser mas alto si se toma en cuenta un muestreo estadístico de pausas e interrupciones en el patrón de voz normal de una persona.

Cuando se usa voz sobre IP, se puede utilizar este ancho de banda “gastado” para otros propósitos cuando se habilita la detección de actividad de voz (VAD – Voice Activity Detection). VAD trabaja detectando la magnitud de la señal de voz en decibelios (dB) y decide cuando cortar la transmisión de la voz (ver figura N° 9). Típicamente, cuando el VAD detecta una caída en la amplitud de la señal de voz, espera una cantidad de tiempo fijo antes de detener la colocación de tramas de voz en los paquetes a ser transmitidos. Esta cantidad de tiempo se conoce como “*hangover*” y su valor típico es de 200ms.

Otros problemas inherentes del VAD es detectar cuando empieza la voz. Típicamente el inicio de una palabra es cortada. Este fenómeno es conocido como

“*front-end speech clipping*”. Usualmente, la persona que escucha esta conversación no nota el fenómeno de “*front-end speech clipping*”.

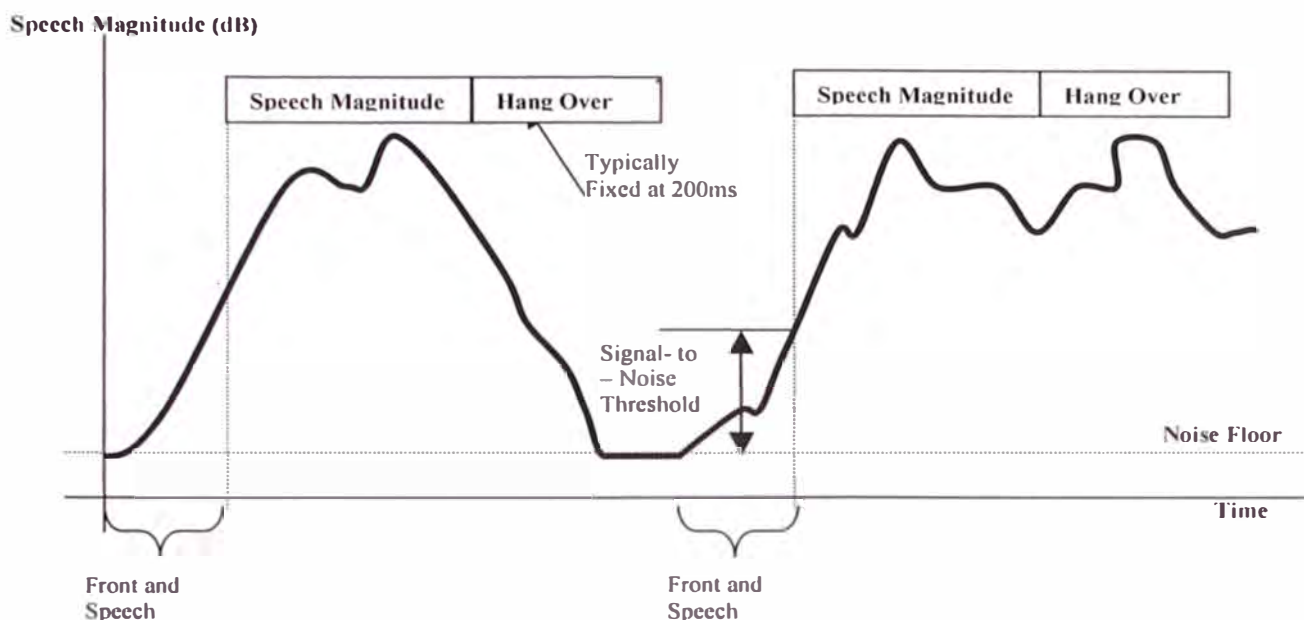


FIGURA N° 9 - VOICE ACTIVITY DETECTION

### 3.7. Requerimiento de ancho de banda para video

En la tabla N° 9 se tiene el resultado de el cálculo de ancho de banda necesario en función del formato en la pantalla tomando como constantes las tramas por segundo y la compresión.

Para este cálculo se ha usado la siguiente fórmula:

$$BW_{Video} = ancho * alto * colores * tramas/seg / Compresión$$

Ancho	Alto	Bit Color	Colores	Trams/seg	BW 10:  Mbps
640	480	4	16	20	2.46
800	600	4	16	20	3.84
640	480	8	256	20	4.92
1024	768	4	16	20	6.29
800	600	8	256	20	7.68
640	480	16	65536	20	9.83
1024	768	8	256	20	12.58
640	480	24	16M	20	14.75
800	600	16	65536	20	15.36
800	600	24	16M	20	23.04
1024	768	16	65536	20	25.17
1024	768	24	16M	20	37.75

**TABLA N° 9 – CÁLCULO DE ANCHO DE BANDA PARA VIDEO SOBRE UNA RED DE DATOS**

## TELEVISIÓN BROADCAST

### Formato PAL-N:

- Cuadros por segundo: 25.
- Líneas por cuadro: 625
- Relación de Aspecto (Ancho/Alto): 4/3
- BW digital equivalente no comprimido: 216 Mbps
- Comprimido 10:1, 21.6 Mbps

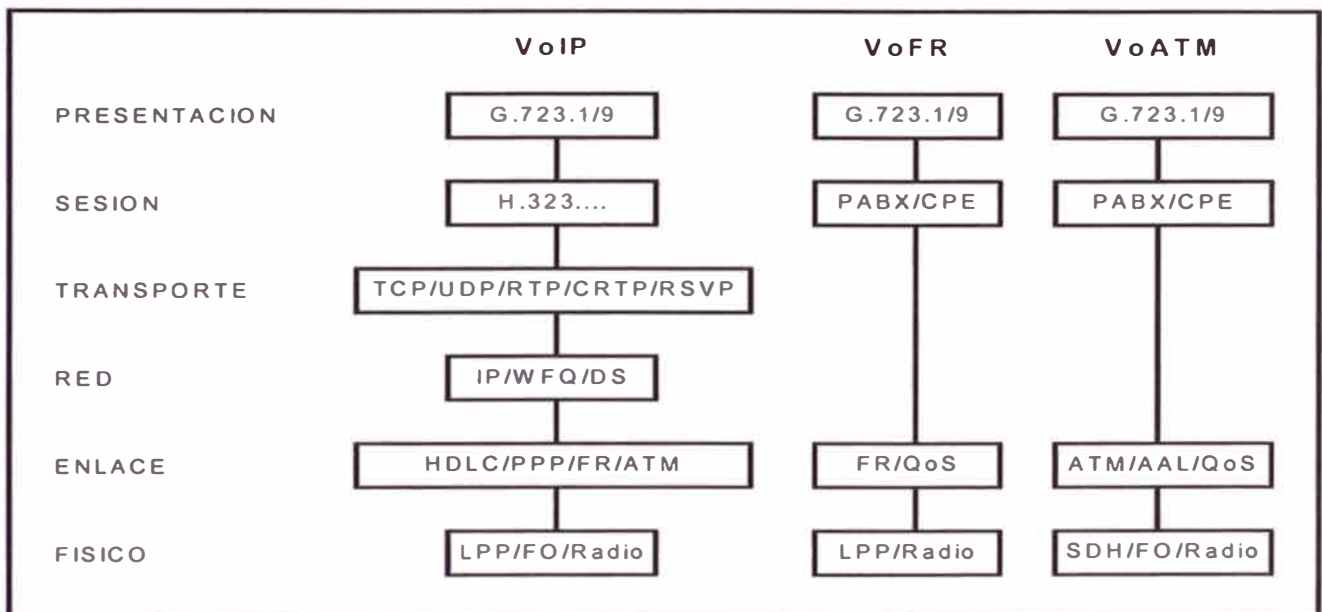
**HDTV:** de 4.2 a 20 Mbps dependiendo de la norma, compatible con MPEG-2

**VIDEO POR DIAL – UP:** Con 33.6 kbps. Usando solo blanco y negro. Bajando el tamaño de la pantalla a 1/8. Hay que usar 17 tramas por segundo. El H.324 es un estándar para video-conferencia por dial-up. Hay otros sistemas propietarios para vigilancia con control de la cámara (PTZ).

## CAPÍTULO IV

### STACK DE PROTOCOLOS PARA VOZ

La figura N° 10 muestra la relación entre el modelo de referencia OSI y los protocolos para transporte de voz sobre una red de datos.



**FIGURA N° 10 – STACK DE PROTOCOLOS PARA VOZ**

La capa de red provee funciones de enrutamiento y direccionamiento lógico de red. Los protocolos de la capa de enlace de datos proveen control y dirigen el control de la información sobre el medio físico.

El principal factor para escoger entre VoIP (Voice over IP) y las alternativas en capa 2 VoFR (Voice over Frame Relay) o VoATM (Voice over ATM), es la interoperabilidad con otras aplicaciones de voz o multimedia. Generalmente hablando, VoFR y VoATM tienen tecnologías de transporte WAN con uso eficiente del ancho de banda, pero VoFR y VoATM no pueden ser desarrolladas sobre LAN o en el desktop. VoIP permite interoperabilidad multivendor, que es más difícil de lograr con VoFR o VoATM debido a que los estándares de estas soluciones han emergido recientemente.

#### **4.1. Protocolos de señalización**

El control de llamada es el proceso mediante el cual se toma una decisión acerca de adonde necesita ser enrutada una llamada y de algún modo hacer la llamada. En la PSTN estas decisiones son llevadas a cabo por el Sistema de Señalización por canal común N° 7 (SS7).

En el modelo de separar los portadores de la capa de control de llamada y separando la capa de control de llamada de los servicios, es necesario asegurar que se usen los protocolos basados en estándares. La tecnología de transmisión de datos es única en el hecho de que múltiples protocolos pueden coexistir en una red y pueden arreglarse a la medida de las necesidades particulares de las aplicaciones.

#### 4.1.1. ITU –T H.323

El protocolo de control de llamada más desarrollado es el **H.323**. Es una recomendación del ITU – T que especifica como transportar el tráfico multimedia sobre una red de paquetes. Utiliza estándares existentes (por ejemplo: Q.931) para lograr sus objetivos. Fue creado para permitir que las aplicaciones multimedia puedan correr sobre redes de datos “no confiables”. El tráfico de voz es sólo una de las aplicaciones H.323. Mucho del trabajo inicial en esta área está enfocada sobre aplicaciones multimedia, con video y datos compartiendo gran parte de los protocolos.

Es la recomendación ITU –T con la base instalada más grande, simplemente porque ninguna otra opción existía antes de H.323. El H.323 no es escalable a implementaciones globales, pero si es apto para VPNs, por ejemplo para llevar a cabo una transferencia de llamada requiere una especificación separada (H.450.2).

El Netmeeting de Microsoft es una implementación de H.323 de Pictoretel.

#### 4.1.2. SGCP (Simple Gateway Control Protocol)

Cisco y Bellcore (Telcordia) impulsaron el **SGCP** (Simple Gateway Control Protocol). Fue desarrollado en los principios de 1998 para reducir los costos de los puntos finales (Media Gateways) teniendo control de llamada

inteligente en una plataforma centralizada (Media Gateway Controller – MGC o soft – switch). Los MG convierten la voz TDM a voz paquetizada.

SGCP se usa para establecer, mantener y desconectar llamadas a través de una red IP. Esto se lleva a cabo controlando las conexiones requeridas entre los correspondientes end points. La autorización de llamadas y conexiones está fuera del alcance de este protocolo. SGCP no contiene mecanismos de seguridad para interferencias o establecimiento de llamadas no autorizadas. La especificación dice, sin embargo, que se espera que todas las transacciones se lleven a cabo sobre conexiones seguras.

#### **4.1.3. PDC (IP Device Control de Level 3)**

Es muy similar al SGCP, pero este tiene muchos otros mecanismos de operación, administración, mantenimiento y provisionamiento (OAM&P) que SGCP. OAM&P es muy importante para los portadores debido a que cubre como el soporte de la red.

#### **4.1.4. MGCP (Multimedia Gateway Control Protocol)**

En los finales de 1998 el IETF impulsó la fusión de IPDC y SGCP para dar origen al MGCP (Multimedia Gateway Control Protocol) que sumado al MDCP (Media Device Control Protocol) da origen al MEGACO. El MGCP ya ha sido adoptado por los cables (DOCSIS).



MGCP controla VoIP a través de elementos externos de control de llamada. MGCP utiliza el mismo modelo de conexión que SGCP, donde las bases de construcción son los end points y las conexiones. Los end points pueden ser físicos o lógicos, y las conexiones pueden ser punto a punto o multipunto. Sin embargo, MGCP permite que las conexiones se establecen sobre varios tipos de redes portadoras, tales como redes IP, ATM. Los mensajes MGCP se transmiten sobre UDP en una red de paquetes. Permite que varios mensajes sean enviados a un mismo gateway en un único paquete UDP. Estos mensajes se procesan como si se hubiesen recibido varios mensajes simultáneamente. Se pueden desarrollar aplicaciones a través del uso de APIs (Application Programming Interface) basados en estándares que interactúen con MGCP y ofrecer funcionalidades adicionales tales como llamadas en espera.

#### **4.1.5. SIP (Session Initiation Protocol)**

El IETF creó el suite del protocolo SIP (Session Initiation Protocol). Es un protocolo basado en media que permite a los dispositivos (endpoints o gateways) ser más inteligentes, así como dar servicios avanzados bajo la capa de control de llamada.

Está descrito en el RFC 2543, que indica que este es un protocolo de control de la capa de aplicación para creación, modificación y terminación de sesiones con uno o más participantes. Es parte de la arquitectura multimedia

del IETF. Es importante notar que SIP puede interoperar con otros protocolos de señalización, tales como H.323. SIP tiene algunas ventajas sobre H.323, tales como tiempo de establecimiento de llamada más corto y menos complejo, implementación con una arquitectura modular conteniendo funciones que residen en protocolos separados.

Estas sesiones multimedia incluyen audio, video y datos y pueden tener múltiples compañeros. Permite que los participantes sean invitados a una conferencia de repente. Estas sesiones multimedia se pueden comunicar a través de multicast, unicast o como una combinación de ambos mecanismos de envío, así como llamadas punto a punto y multipunto. Se puede establecer y terminar sesiones usando las siguientes cinco facetas de SIP: ubicación del usuario, capacidad del usuario, disponibilidad del usuario, establecimiento de llamada, y manipulación de llamada.

Hay muy pocas implementaciones de SIP actualmente, aunque muchos vendedores y clientes están interesados en usar SIP para desarrollar servicios avanzados. Todos estos protocolos manejan la forma de establecer la sesión.

Es fundamental la interoperabilidad con SS7 para facilitar la migración e intercomunicación entre redes y mantener las prestaciones.

Hay una interfaz de desarrollo del IETF: **TALI** (Transport Adaptation Layer Interface).

El Grupo MEGACO del IETF y el grupo 16 de ITU-T acordaron impulsar el estándar H.gcp (H.323 + MEGACO) (Santiago, Chile 5/1999) hoy llamada la H.248 de ITU – T.

En la figura N° 11 se ilustra la evolución de los protocolos de servicios de voz, algunos de los cuales se han descrito líneas arriba.

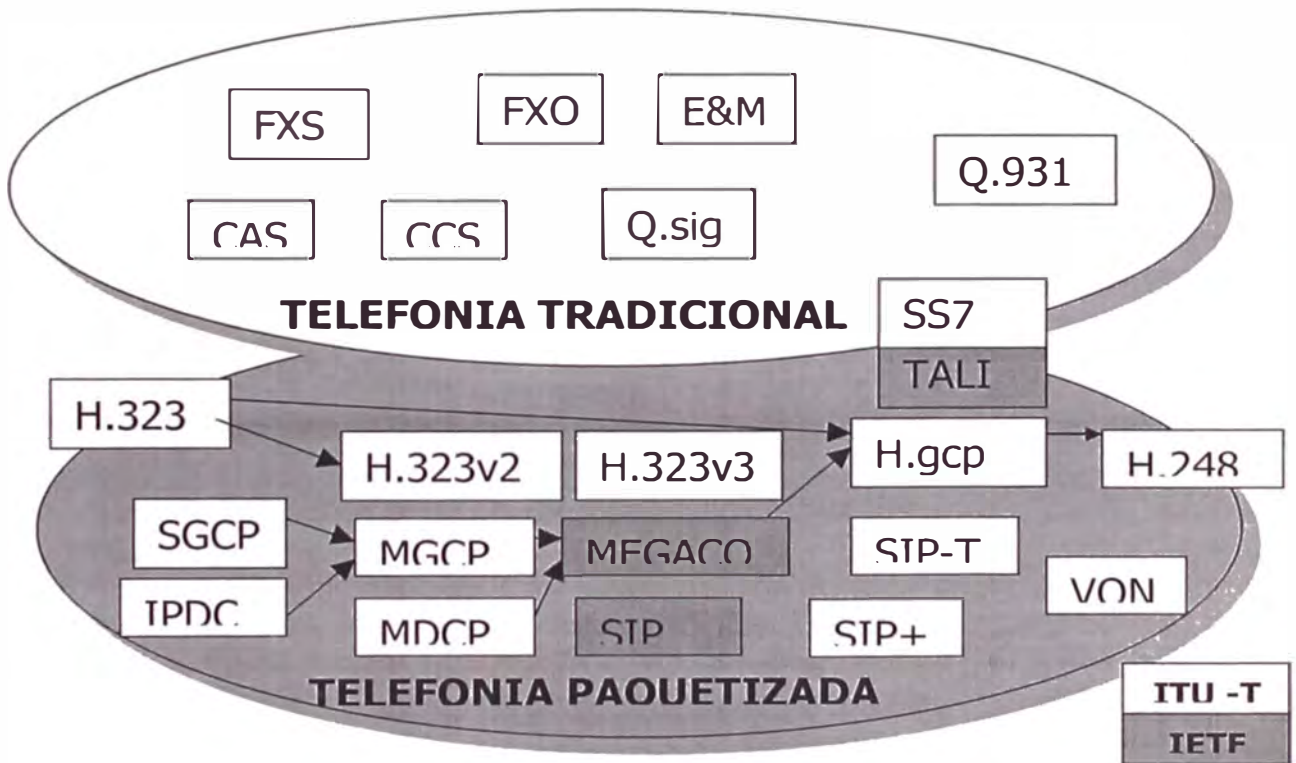


FIGURA N° 11 EVOLUCIÓN DE LOS PROTOCOLOS DE SERVICIO DE VOZ

#### 4.2. Señalización por canal común N° 7

El SS7 es un estándar de señalización por canal común desarrollado en los finales de 1970 por el ITU –T. SS7 fue derivado de SS6, el cual se desarrolló a finales de 1960 y fue la primera generación de señalización por canal común. SS7 fue desarrollado inicialmente para aplicaciones de control de llamada en telefonía. Las aplicaciones de SS7 se han expandido ampliamente desde su aparición, las funcionalidades actuales de SS7 incluyen mensajes de establecimiento y terminación de llamadas, consulta a base de datos, transacciones, operaciones de red, y la red digital de servicios integrados (ISDN).

Es una arquitectura usada para soportar señalización de llamadas fuera de banda en una red de paquetes a una velocidad de 56/64 Kbps. Soporta el establecimiento de llamadas, intercambio de información, routing, operaciones, tarificación y soporte para servicios de redes inteligentes (Intelligent Networks – IN). Es un estándar internacional para señalización de circuitos conmutados entre proveedores de servicio telefónico, aunque su estructura permite variaciones de país a país. Las redes SS7 no transportan datos del cliente (excepto en el caso de los mensajes Q.931 generados por los CPE).

SS7 es importante para voz sobre IP (VoIP) y la forma en que este interactúa con la PSTN. Esta interacción es crítica para la aceptación y el éxito de las soluciones de VoIP en las redes telefónicas actuales.

#### **4.2.1. ISDN User Part (ISUP)**

ISUP define los mensajes y protocolos usados en el establecimiento y corte de llamadas de voz y datos sobre la red pública conmutada y administra la red troncal sobre la que se apoya.

A pesar de su nombre, se usa tanto para las llamadas ISDN como para non-ISDN. También se usa en redes celulares o móviles para conexiones de troncales. Está ampliamente adoptada internacionalmente, aunque existen variaciones en algunos países.

#### **4.2.2. Transaction Capabilities Application Part (TCAP)**

El TCAP define los mensajes y protocolos usados para comunicar las aplicaciones (sub-sistemas) en los nodos. Es usado para servicios de bases de datos tales como “calling card”, números 800 y ANI (Automatic Number ID) como también servicios de consulta y respuesta “switch to switch” incluyendo discado repetido y retorno de llamada. Las redes inteligentes usan el TCAP para invocar características en oficinas remotas, cuentan con los servicios de TCAP para permitir un punto final de señalización para acceder a características en un punto final de señalización remoto.

Debido a que los mensajes TCAP deben ser enviados a aplicaciones individuales dentro de los nodos que direccionan, es que usan el SCCP para el transporte.

#### **4.2.3. Operations, Maintenance and Administration Part (OMAP)**

OMAP define mensajes y protocolos diseñados para asistir a los administradores de la red SS7. A la fecha, las capacidades más integralmente implementadas y desarrolladas son procedimientos para validar tablas de ruteo y diagnósticos para problemas de enlace. OMAP incluye mensajes que usan ambos, el MTP y el SCCP para el ruteo.

#### **4.3. Protocolos de transporte**

Debido a la naturaleza del tráfico de voz (sensible en el tiempo), la opción lógica para transportarla es UDP/IP. Sin embargo, se necesita más información de los paquetes que la ofrecida por UDP. La cabecera UDP sólo tiene cuatro campos: puerto fuente, puerto destino, longitud y checksum (opcional).

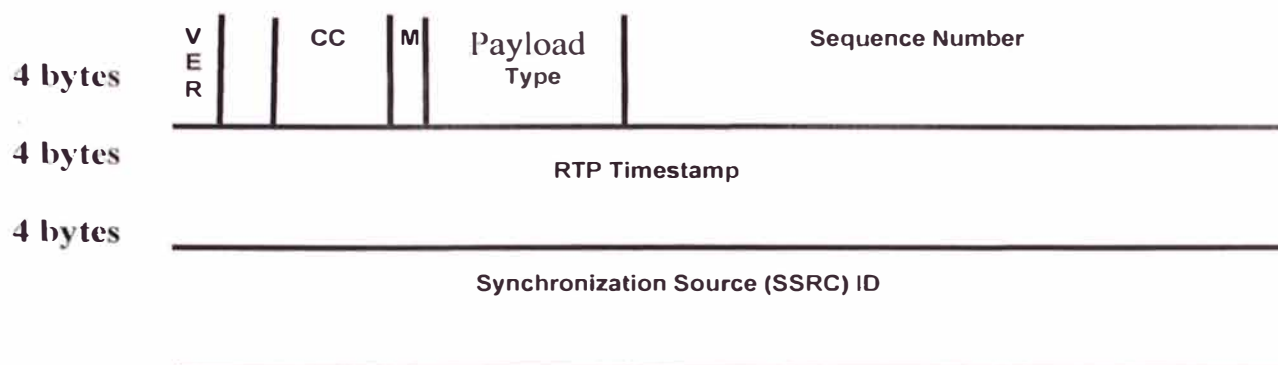
No se usa TCP debido a que el control de flujo y retransmisión de paquetes de voz son innecesarios. Si se usara TCP para VoIP, la latencia incurrida esperando por acknowledgments y retransmisiones puede dar calidad de voz inaceptable. Para VoIP y otras aplicaciones de tiempo real el control de la latencia es más importante que

asegurar el envío confiable de cada paquete. TCP se usa para el establecimiento de la llamada en muchos protocolos de señalización de VoIP.

Para tráfico de tiempo real o sensible al retardo, el IETF (Internet Engineering Task Force) adoptó el RTP (Real Time Transport Protocol). La voz sobre IP reside sobre RTP, que a su vez está sobre UDP. Entonces, un paquete de voz es llevado con un encabezado IP/UDP/RTP.

#### **4.3.1. Real-time Transport Protocol (RTP)**

Es un estándar para transmisión de tráfico sensible al retardo a través de una red basada en paquetes. RTP corre sobre UDP/IP. Está especificado en los RFC 1889 y RFC 1890. Provee función de transporte en la red extremo a extremo. Entrega el tráfico del tiempo real con información de timing y feedback de calidad de recepción, que no está en la trama UDP (campos “sequence number” y “timestamp”), tal como se aprecia en la figura N° 12. RTP usa el campo “sequence number” para determinar si los paquetes están llegando en orden y la información de “timestamp” para determinar el tiempo de llegada entre paquetes (jitter), si es que estos están en orden o si fueron recibidos cuando eran esperados. Esta información sirve de ayuda a la estación receptora para determinar como sintonizar sus parámetros para enfrentar los problemas de la red, tales como el retardo, jitter y pérdida de paquetes.



**FIGURA N° 12 – ESTRUCTURA DE LA CABECERA RTP**

Se puede usar para servicios interactivos tales como telefonía por Internet. A la fecha, todos los protocolos de señalización de voz sobre IP usan IP/UDP/RTP como mecanismo de transporte para tráfico de voz. RTP consiste de una parte de datos y una parte de control (RTCP – Real Time Control Protocol).

La parte de datos de RTP es un protocolo que provee soporte para aplicaciones con propiedades de tiempo real, tales como audio y video, incluyendo reconstrucción de timing, detección de pérdidas e identificación de contenido. Usa el RTCP para controlar la calidad de la sesión.

Un paquete de voz tiene un encabezado compuesto de 20, 8 y 12 bytes respectivamente para IP, UDP y RTP. Esto añade una cabecera de 40 bytes, el cual es el doble de la carga útil cuando se usa el CODEC G.729 con dos muestras de voz (20ms). En la figura N° 13 se tiene la estructura de un paquete de voz sobre IP en una red Ethernet.



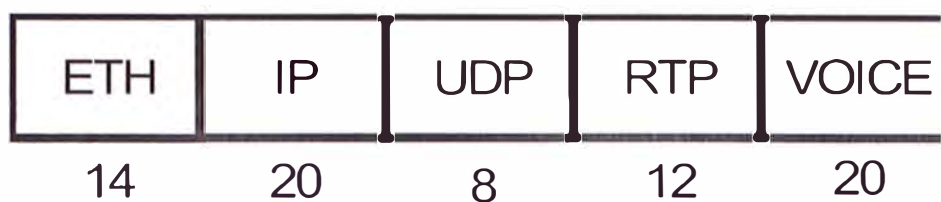


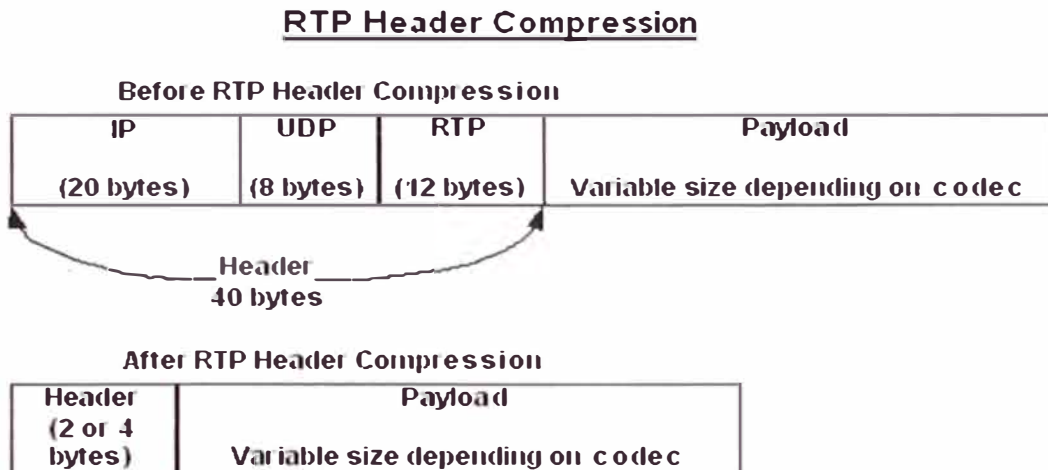
FIGURA N° 13 – ESTRUCTURA DE UN PAQUETE DE VOZ SOBRE IP

#### 4.3.2. CRTP: Compressed RTP

Para reducir la cantidad de ancho de banda consumida por una llamada de voz, se puede usar el CRTP (ver figura N° 14). Comprime la cabecera IP/UDP/RTP de 40 a 2 bytes (sin checksum) ó 4 bytes (con checksum). Usa las mismas técnicas de compresión que TCP/IP. Por ejemplo, para una llamada con G.729 sin CRTP en un enlace Frame Relay se puede consumir hasta 24 Kbps, pero si usamos CRTP se puede obtener un consumo de hasta 11.2 Kbps. Entonces, si un enlace de 64 Kbps puede llevar normalmente dos llamadas de VoIP, con CRTP se puede tener cuatro llamadas de VoIP.

A velocidades mayores a 512 Kbps, el CRTP compromete la ocupación de porcentaje de CPU si la compresión es por software. El CRTP requiere descomprimir en el primer hop para recuperar el header IP. Una buena regla es mantener el uso del CPU en no más de 60 o 70%. Se debe ponderar las ventajas (bajo consumo de ancho de banda) contra las desventajas (alto consumo de CPU) al momento de aplicar esta técnica.

Se debe usar el CRTP sobre cualquier enlace WAN donde el ancho de banda es un recurso limitado y se tiene alto tráfico RTP.



**FIGURA N° 14 – COMPRESIÓN DE LA CABECERA RTP**

Las siguientes líneas configuración pertenece a un router Cisco en el que se muestra como se debe activar el CRTP, para los casos de una línea dedicada y de un circuito con enlace frame relay:

#### **Leased line**

!

interface serial 0

ip address 192.168.121.18 255.255.255.248

no ip mroute-cache

**ip rtp header-compression**

encapsulation ppp

!

#### **Frame Relay**

!

interface serial 0/0

```

ip address 192.168.120.1 255.255.255.0
encapsulation frame-relay
no ip route-cache
no ip mroute-cache
frame-relay ip rtp header-compression

```

!

#### 4.4. Estándar H.32X

H.323 es una especificación del ITU-T para transmisión de audio, video y datos a través de una red IP, incluyendo el Internet. H.323 dirige la señalización y control de llamadas, transporte y control de multimedia, y control de ancho de banda para conferencias punto a punto y punto a multipunto.

Las recomendaciones serie H también especifican H.320 para ISDN y H.324 para POTS como mecanismo de transporte (ver figura N° 15).

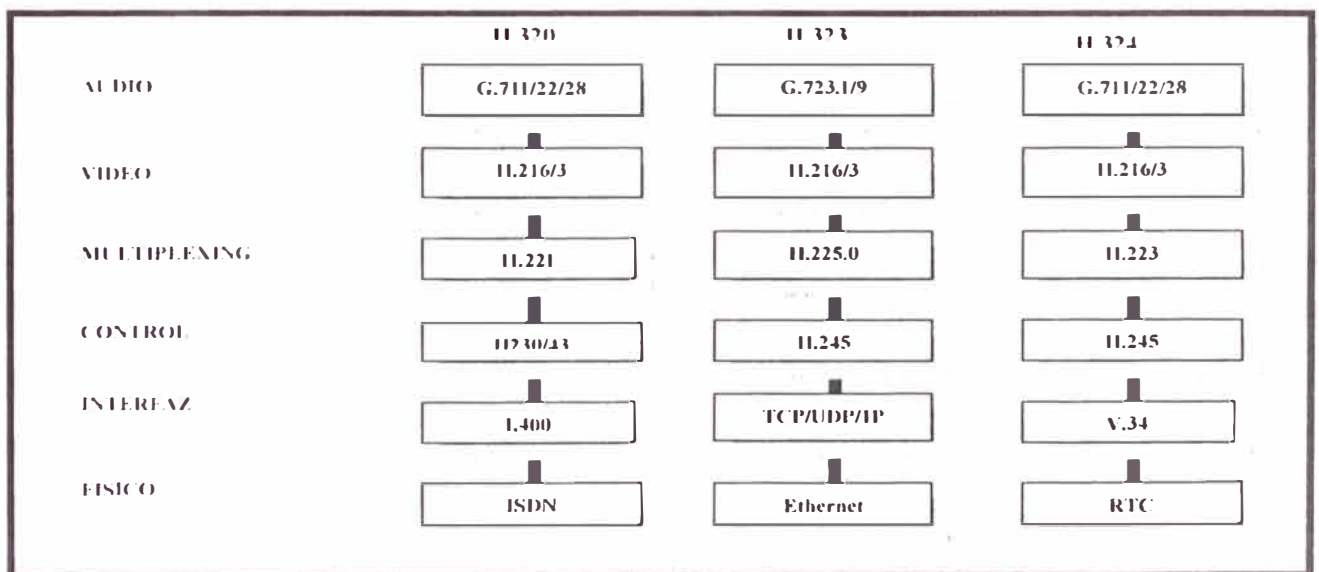


FIGURA N° 15 – SUITE DE PROTOCOLOS DE LOS ESTÁNDARES H.320, H.323 Y H.324

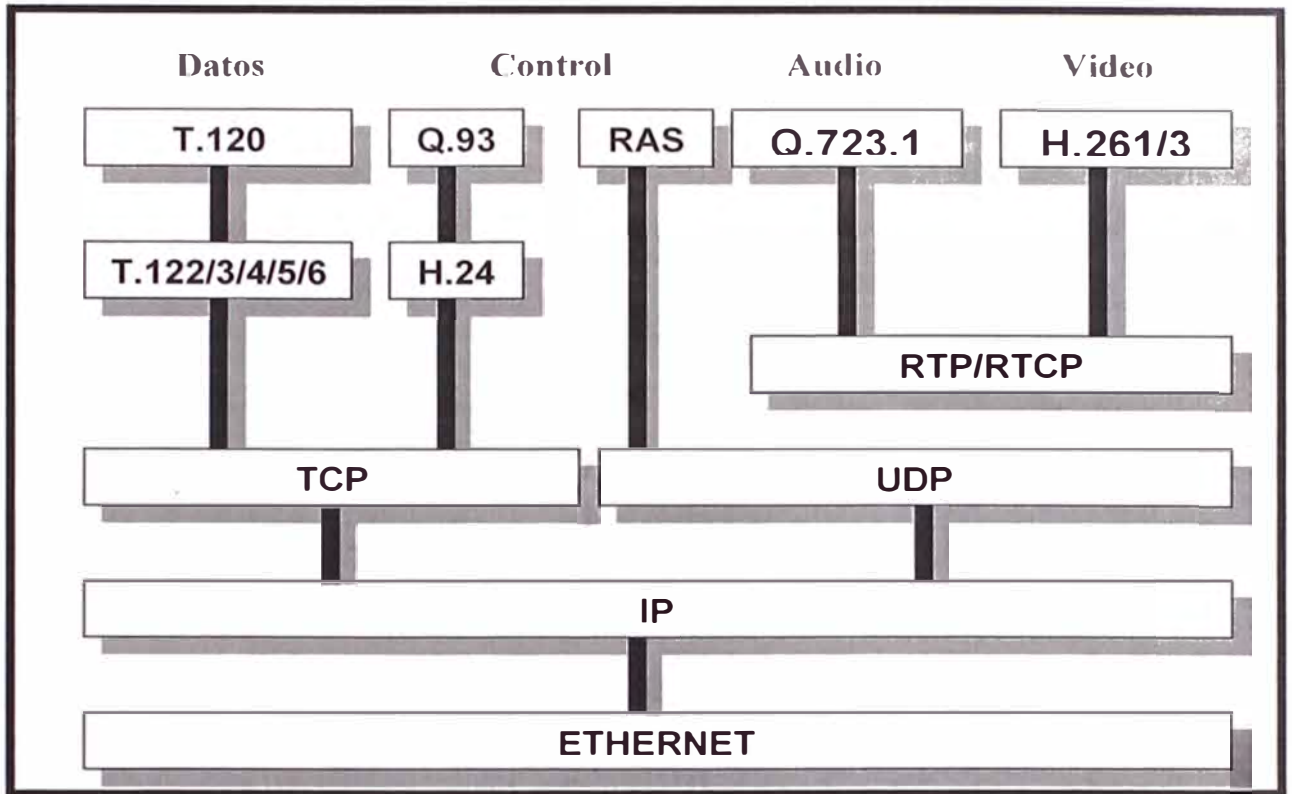
#### 4.4.1. Estándar H.323

El estándar H.323 consiste de los siguientes componentes y protocolos:

<b>Feature</b>	<b>Protocolo</b>
Señalización de llamada	H.225
Control de medio	H.245
CODECs de audio	G.711, G.722, G.723, G.728, G.729
CODECs de video	H.261, H.263
Datos	T.120
Transporte	RTP/RTCP

En la figura N° 16 se tiene un cuadro del suite de protocolos del estándar H.323. H.323 es un sistema híbrido construido de gatekeeper centralizados inteligentes, MCUs y al menos puntos finales inteligentes. Aunque el estándar H.323 es más completo en versiones recientes, tiene algunos problemas, tales como largo tiempo de establecimiento de llamada, cabecera de un protocolo de conferencia con todas sus facilidades, demasiadas funciones requeridas en cada gatekeeper y consideraciones de escalabilidad para implementaciones de ruteo de llamadas en gatekeepers.

Se han desarrollado alternativas para necesidades de gateways de alta densidad para conexión con la PSTN tales como SGCP y MGCP. Estos sistemas de control de llamada proveen soluciones más efectivas y escalables para satisfacer implementaciones carrier-class.



**FIGURA N° 16 – SUITE DE PROTOCOLOS DEL ESTÁNDAR H.323.**

En la figura N° 17 se ilustran los elementos de un sistema H.323. Frecuentemente referidos como puntos finales (endpoints), los terminales proveen conferencia punto a punto y punto a multipunto para audio y, opcionalmente, video y datos. Los gateways interconectan el sistema con la red pública PSTN o con una red ISDN para interoperabilidad de los terminales H.323. Los gatekeeper proveen control de admisión y servicios de traducción de direcciones para terminales o gateways. Los MCU son dispositivos que permiten conferencia con sesiones de audio o video a dos o más terminales o gateways

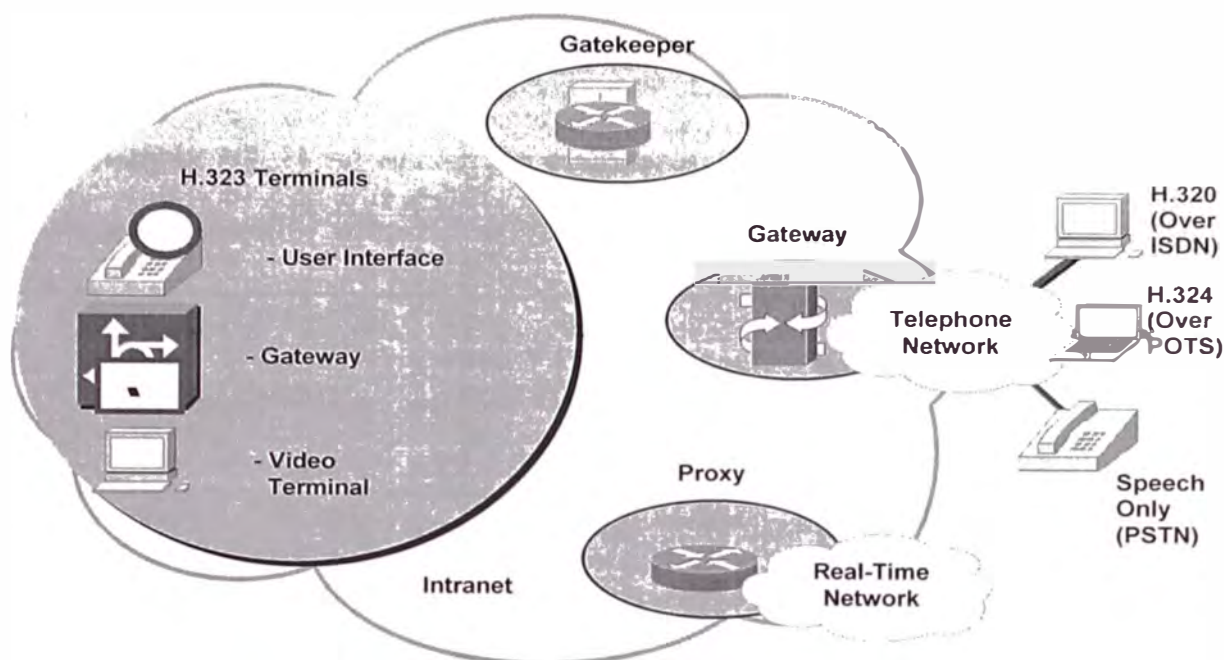


FIGURA N° 17 – INFRAESTRUCTURA DE UN SISTEMA H.323.

#### 4.5. Videoconferencia

Todo terminal H.32X con video debe soportar el codec H.261 (ver tabla N° 10), y opcionalmente el H.263 para bajas velocidades. El H.261 se usa con canales con velocidades de  $n \times 64$  kbps.

Formato	Tamaño de pixels	H.261	H.263
Sub – QCIF	128x96	no especificado	requerido
QCIF	176x144	Requerido	requerido
CIF	352x288	Opcional	opcional
4CIF	704x576	N/A	opcional
16CIF	1408x1152	N/A	opcional

(CIF: Common Intermediate Format)

TABLA N° 10 – ESPECIFICACIONES PARA VIDEOCONFERENCIA

#### 4.5.1. Estándares ITU para Video Tele Conferencia (VTC)

En la tabla N° 11 se tienen los estándares ITU para VTC en diferentes tipos y aplicaciones.

	<b>Narrowband VTC (H.320)</b>	<b>Low bit rate VTC (H.324)</b>	<b>Iso Ethernet VTC (H.323)</b>	<b>Ethernet VTC (H.321)</b>	<b>ATM VTC (H.321)</b>	<b>High resolution ATM VTC (H.310)</b>
Video	H.261	H.261	H.261	H.261	H.261	MPEG-2
		H.263		H.263		H.261
Audio	G.711	G.723	G.711	G.711	G.711	MPEG
	G.722		G.722	G.722	G.722	
	G.728		G.728	G.728	G.728	
				G.723.1 G.729		
Data	T.120	T.120 T.434 T.84 Others	T.120	T.120	T.120	T.120
Multiplex	H.221	H.223	H.221	H.225	H.221	H.222.1 H.221
Signalling	H.230	H.245	H.230	H.345	H.242	H.245
	H.243		H.242			
Multipoint	H.243	N/A	H.243	N/A	H.243	N/A
Encryption	(In draft revision)	H.233 (adapted in H.324)	(By reference to H.320)	TBD	H.233 H.234	N/A
	H.233					
	H.234	H.234				

**TABLA N° 11 – ESTÁNDARES ITU – T PARA VTC.**

## 4.6. Multimedia sobre IP

A continuación se tiene un esquema (figura N° 18) del transporte de multimedia sobre IP usando el protocolo RTP para el payload. Los router hacen reservación de recursos para priorizar el tráfico multimedia mediante el protocolo RSVP. El protocolo RTCP se encarga del control del tráfico RTP.

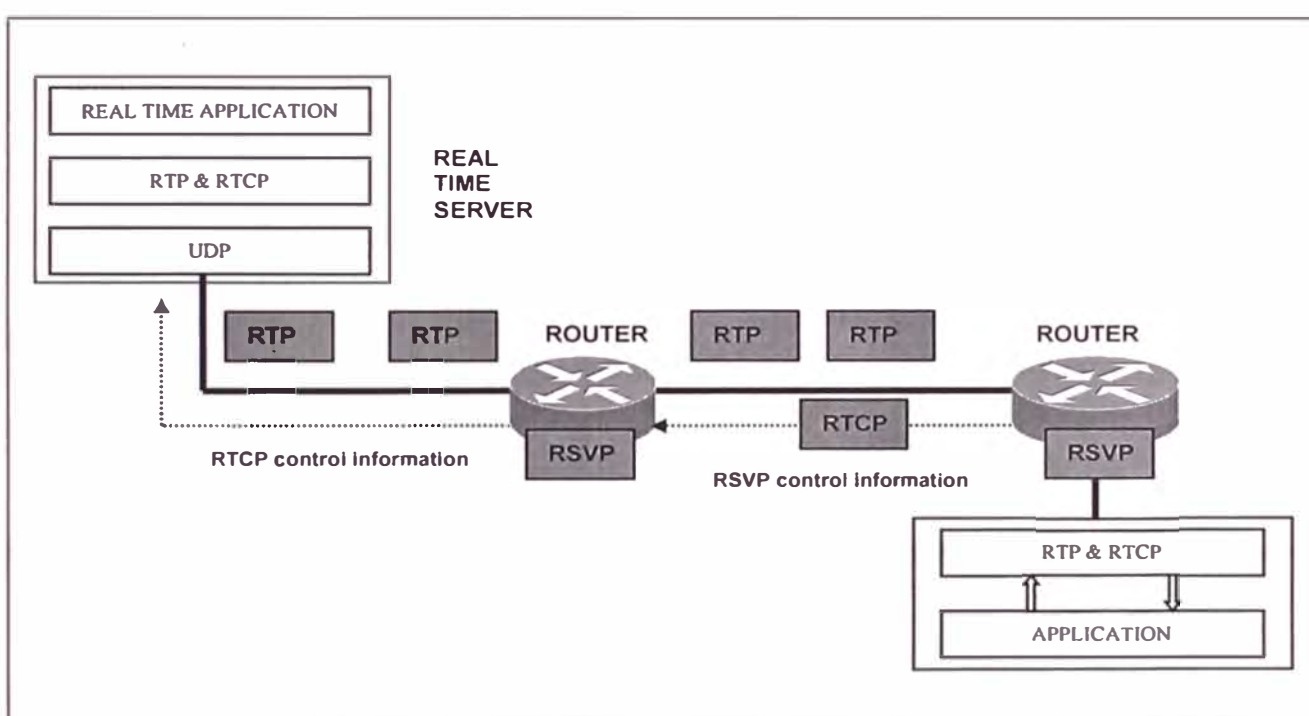


FIGURA N° 18 – MULTIMEDIA SOBRE IP.

### 4.6.1. RTSP (Real Time Streaming Protocol)

Está especificado en el RFC 2326. Provee audio y video on-demand para aplicaciones como broadcasting. El receptor decodifica la información de tiempo real rápidamente (on the fly).



#### 4.6.2. Stack de protocolos

En la figura N° 19 se ilustra la pila de protocolos que se usan para el transporte de tráfico multimedia.

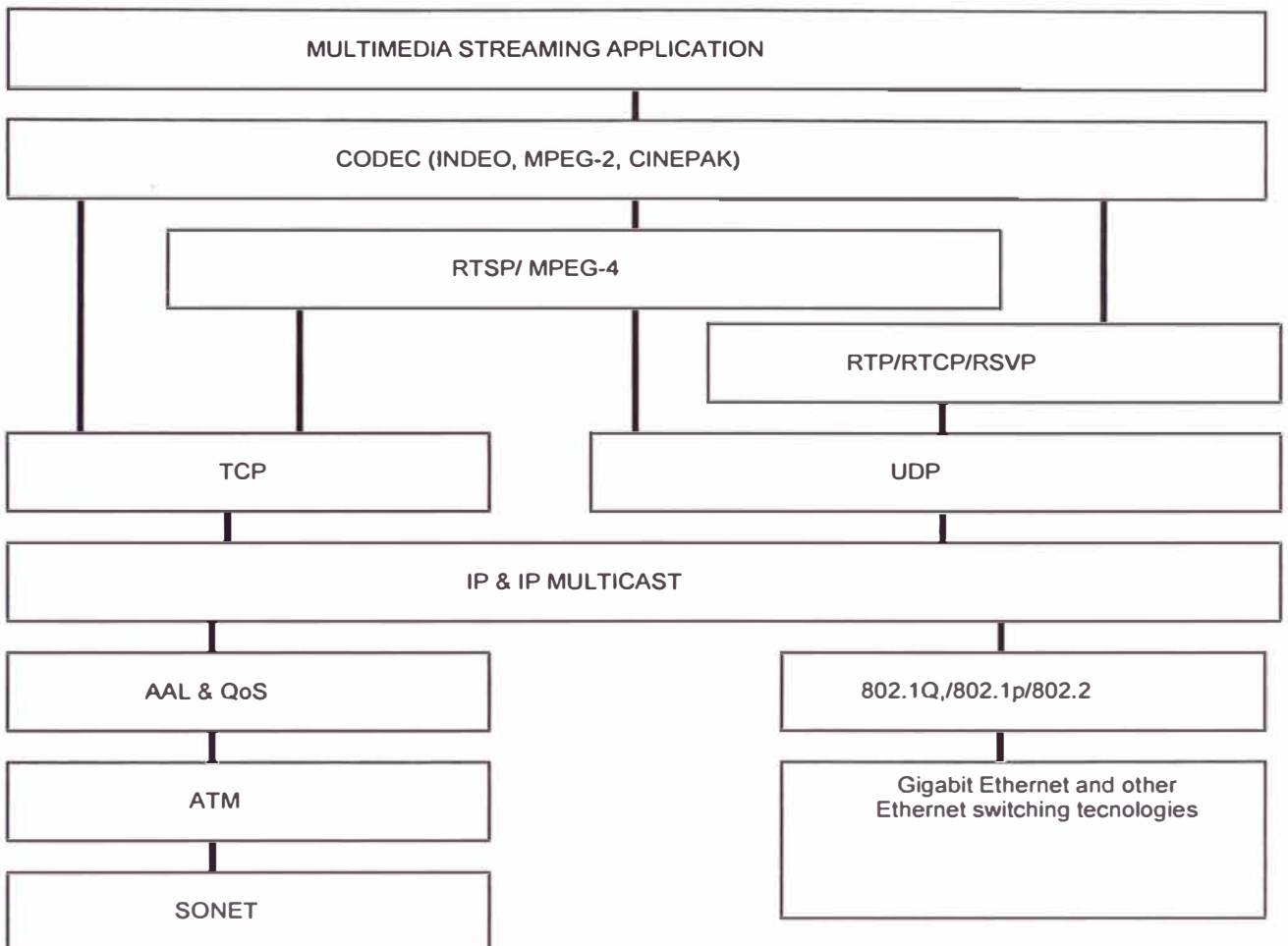


FIGURA N° 19 – STACK DE PROTOCOLOS PARA MULTIMEDIA

#### 4.7. Integración de redes

Uno de los aspectos por el que se combinan las redes de voz y datos es el ahorro de costos. Entonces, integrando las redes de voz y datos los costos se reducen debido a una única estructura de soporte para todos los servicios.

A continuación se tiene un cuadro comparativo (tabla N° 12) del costo de servicios de voz, datos y video al ser transportados sobre una red de voz, de datos y de video.

Se han tomado valores referenciales.

Servicio	Voz	Datos	Video
Llamada internacional de 10 minutos	4.00	0.14	0.001
Transferencia de archivo de 10 MB	66.27	2.40	0.02
Videoconferencia de 2 horas	12288.00	432.00	3.00

**TABLA N° 12 – ¿CUANTO CUESTAN LOS BITS?**

Tanto en la LAN (Local Area Network) como en la WAN (Wide Area Network) se produce integración por consolidación del tráfico. Es fundamental garantizar la calidad de servicio (QoS – Quality of Service) requerida por cada aplicación. En tráfico orientado a conexión, puede hacerse, ya que las conexiones tienen parámetros de QoS. En tráfico no orientado a conexión se requieren otros mecanismos, algunos de ellos pueden ser: RTP, RTCP, RSVP, RTSP, WFQ, y Over-engineering de la red (caños holgados). La tendencia es TCP/IP sobre FR ATM para dar QoS.

La QoS extremo a extremo requiere ingeniería de detalle.

## **CAPÍTULO V**

### **CALIDAD DE SERVICIO**

Calidad de servicio (QoS) es un término usado frecuentemente y mal empleado a la vez, con una diversidad de significados. En este trabajo nos vamos a referir a calidad de servicio como clase de servicio (CoS) y tipo de servicio (ToS). La meta fundamental de CoS y ToS es dar el ancho de banda y latencia necesaria para una aplicación en particular.

CoS permite a un administrador de red agrupar diferentes flujos de paquetes, cada uno de ellos teniendo distintos requerimientos de latencia y ancho de banda. ToS es un campo en el encabezado del protocolo IP que hace posible el CoS. El campo ToS usa tres bits, que permiten una escala de ocho tipos de tráfico o CoS (0-7). Nuevos RFC permiten el uso de seis bits para el campo ToS, es decir mas CoS.

Se tienen varias herramientas disponibles para llevar a cabo el necesario nivel de QoS para una aplicación dada.

Es importante notar que las herramientas para implementar estos servicios no son tan importantes como el resultado final. En otras palabras, no hay que mirar solo una

herramienta de QoS para resolver los problemas. Por el contrario hay que ver la red como un todo para determinar que herramientas nos van a ser útiles, ya que cada red tiene sus problemas individuales.

Algunos de los problemas que no se pueden resolver son el retardo de propagación, retardo de codificación, retardo de muestreo y digitalización. Los problemas que podemos resolver son las pérdidas de paquetes, jitter, retardo de manipulación.

Cisco ofrece muchas herramientas para implementar QoS. En algunos escenarios, no se usa ninguna herramienta de QoS y se logra el QoS necesario para la aplicación dada. En general, cada red tiene problemas individuales que se pueden resolver usando una o más herramientas de QoS.

### 5.1. Funciones de edge y core

En una red de paquetes bien diseñada se debe tener cuidado de separar las funciones que ocurren sobre los bordes de la red y las funciones que deben ocurrir en el core o backbone (figura N° 20). Es importante separar estas funciones para lograr la mejor calidad de servicio posible.

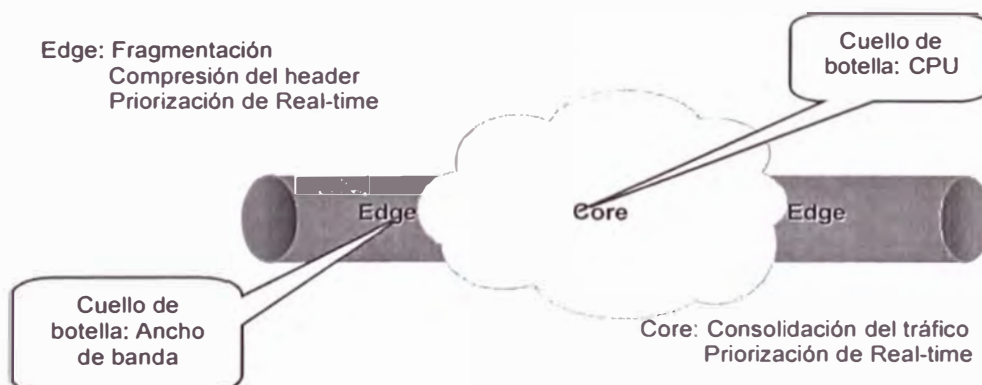


FIGURA N° 20 – FUNCIONES DE EDGE Y CORE

En los bordes los cuellos de botella se deben principalmente al ancho de banda. Entonces, se deben aplicar herramientas como fragmentación, compresión, priorización de tráfico, control de admisión, clasificación de datos, garantía de ancho de banda, etc. En el core o backbone los recursos de ancho de banda no son escasos. El problema acá es el procesamiento de CPU, memoria.

## **5.2. Calidad de servicio relativa y absoluta**

### **5.2.1 Relativa:**

Se basa en la priorización del tráfico (“alta/media/baja”), esto es, dice quien es mejor, pero no cuanto. Es inútil si todos tienen la misma prioridad o si la mayoría de los usuarios tienen QoS alta. Se involucra varias redes (LAN1-WAN1-WAN2-LAN2). Se requieren beneficios cuantificables (tráfico de tiempo real).

### **5.2.2 Absoluta:**

Define diferentes clases de servicio (CoS). Cada CoS tiene sus parámetros de QoS. No sólo es prioridad. Define como administrar el tráfico cuantificadamente.

Por ejemplo en Frame Relay mediante CIR, FTD, FRL, y en ATM: SIR, CTD, CDV, CLR. Los SLAs (Service Level Agreement) definen los límites. Por ejemplo, si se permite un máximo de 10 minutos de pérdida de servicio al

año, entonces se especifica la disponibilidad del servicio en 99.998%. Los TCA (Traffic Conditioning Agreement) fuerzan el cumplimiento.

La calidad de servicio absoluta asegura parámetros:

- Throughput.
- Delay.
- Loss rate.

Cada paquete tiene sus CoS & QoS asignado.

### **5.3. Tipos de servicio con calidad de servicio**

#### **5.3.1 QoS Absoluto**

- End to end, bandwidth, throughput, delay garantizado

#### **5.3.2 QoS Relativo**

- Prioriza tráfico relativo a otro tráfico

#### **5.3.3 Servicio Best – effort**

- No garantiza nada. Por ejemplo: Tráfico IP tradicional.

### **5.4. Estándares de calidad de servicio**

#### **5.4.1 Resource reservation protocol. – RSVP (RFC 2205)**

Permite a los puntos extremos señalar la red con el tipo de QoS necesario para una aplicación particular. Reserva ancho de banda de cada nodo para

cada conexión en forma dinámica, con señalización end to end y para cada flujo de tráfico.

Los administradores de red pueden usar RSVP como una lista de acceso dinámica. Esto significa que los administradores de red no deben preocuparse por los números de puerto de los flujos de paquetes IP debido a que RSVP señala esta información durante su requerimiento original.

Es un protocolo de señalización de extremo a extremo, fuera de banda que requiere una cierta cantidad de ancho de banda y latencia en cada salto de red que lo soporte. Si un nodo (router) no soporta RSVP, RSVP no se activará para el siguiente nodo. Es decir, requiere que todos los nodos lo soporten.

RSVP trabaja como una ambulancia que va limpiando el tráfico, sencillamente los paquetes favorecidos van a seguir a la ambulancia.

En la implementación de Cisco, cada interface en la que se desee tener RSVP debe ser configurado explícitamente. Asimismo, el administrador de la red debe configurar el ancho de banda reservado para RSVP en cada interface.

Las aplicaciones reciben una realimentación para saber si su QoS requerido fue aprobado o denegado. Algunas aplicaciones transmiten su data a cualquiera sin preocuparse por su QoS; sin embargo, otras aplicaciones inteligentes escogen no transmitir o escogen una ruta alternativa.

RSVP tiene sus desventajas que son: escalabilidad (no es escalable a más de 10 saltos), control de admisión y el tiempo que toma para establecer la reservación de extremo a extremo.

El tema de la escalabilidad que rodea a RSVP lo relega a los bordes de la red y fuerza al uso de otras herramientas en el backbone. El IETF está trabajando para mejorar el factor de escalabilidad de RSVP.

Periódicamente refresca la reserva (30 segundos). Si falta el refrescado, la reservación se resetea por tiempo muerto. No es aplicable para enlaces de alta velocidad (por ejemplo: OC-48). Tiene un complejo handshaking (hop for hop) para establecer sesión a nivel 3. No negocia reducir lo requerido y una vez reservado si no se usa se pierde

RSVP no es un protocolo de ruteo y no puede modificar la tabla de ruteo basado en los tipos de tráfico. Los protocolos de ruteo RIP o OSPF buscan las rutas mas cortas, los más óptimos, no las mas descargadas o las que puedan lograr la reserva, por lo tanto esta ruta no puede ser la más ideal para QoS.

#### **5.4.2 IP RTP reserve**

Cisco tiene otro mecanismo para ponderar el tráfico basado en el rango de los puertos UDP que usa. Se puede comparar IP RTP Reserve a una RSVP



estático. Cuando se usa IP RTP Reserve no se debe usar la precedencia IP ni RSVP.

Aunque IP RTP Reserve clasifica los paquetes basado en el rango de los puertos UDP también se puede especificar la cantidad de ancho de banda que se permitirá al tráfico a priorizar. Cuando no haya tráfico a ser priorizado el ancho de banda total está disponible para el resto de tráfico.

La configuración de IP RTP Reserve en un router Cisco será como sigue:

```
!  
interface virtual-template 1  
    ppp multilink  
    encapsulation ppp  
    ppp multilink interleave  
    ppp multilink fragment-delay 20  
    ip rtp reserve 16384 100 64
```

```
!
```

Donde: Lowest-UDP-port (16384), range of ports (100), maximum bandwidth (64).

IP RTP Reserve no es la solución más escalable disponible. Sólo permite un rango de 100 puertos UDP y no tiene control de admisión.

### 5.4.3 Multi-protocol Over ATM – MPOA

Evolución: LANE Classical IP over ATM.

Classical IP + NHRP Multi-protocol Over ATM.

Estándares del ATM Forum que especifica como múltiples protocolo de red pueden operar sobre un substrato de ATM

Se basan en SVCs para mapear tráfico IP del tipo Best effort. Es muy compleja la interacción cuando se requiere QoS en el IP. Cursa los flujos frecuentes por SVCs de ATM (> 10 pps).

Luego de un periodo de inactividad libera el SVC.

### 5.4.4 Multi-Protocol Label Switching – MPLS

Es un desarrollo clave en tecnologías de internet que asisten en la adición de un número de capacidades esenciales para las redes IP best effort de hoy, incluyendo: Ingeniería de tráfico, provisión de tráfico con clase de servicio y calidad de servicio diferenciada cualitativamente, y soporte de VPN basada en IP.

Facilita el forwarding de paquetes en los routers bajando el retardo. Rutea en base de labels de largo fijo en vez de direcciones IP. El label se inserta entre la cabecera de nivel 2 y de nivel 3 (4 bytes). Permite brindar QoS usando FC

(Forwarding Classes) para cada tipo de tráfico, incorpora 3 bits de CoS donde mapea el “IP precedence”.

Aún no es un estándar, hay un foro: <http://www.mplsforum.com>. El Draft especifica que toda implementación debe interoperar con RSVP. Ha sido definido para montar IP sobre: PPP, FR o ATM

El bit S permite apilar labels para extender la capacidad. El MPLS permite escalar redes IP, en conjunto con RSVP brinda QoS absoluta para redes reducidas ya que el RSVP no es escalable. Encapsula el paquete con IP con un header 32 bits (figura N° 21), es decir, mas cabecera.



**FIGURA N° 21 – ESTRUCTURA DE UN PAQUETE MPLS**

Todos los routers deben mantener una tabla de flujo de labels. No agrega valor alguno en redes en estrella. Mejora el retardo pero carga la CPU y la RAM. MPLS es excelente en el núcleo donde la velocidad es critica. Puede ser tan bueno el “edge” donde se toman las decisiones de ruteo.

### 5.4.5 Precedencia IP

La precedencia IP se refiere a los tres bits en el campo ToS (Type of Service) en la cabecera del paquete IP versión 4. Estos tres bits permiten ocho tipos diferentes de CoS (ver figura N° 22).

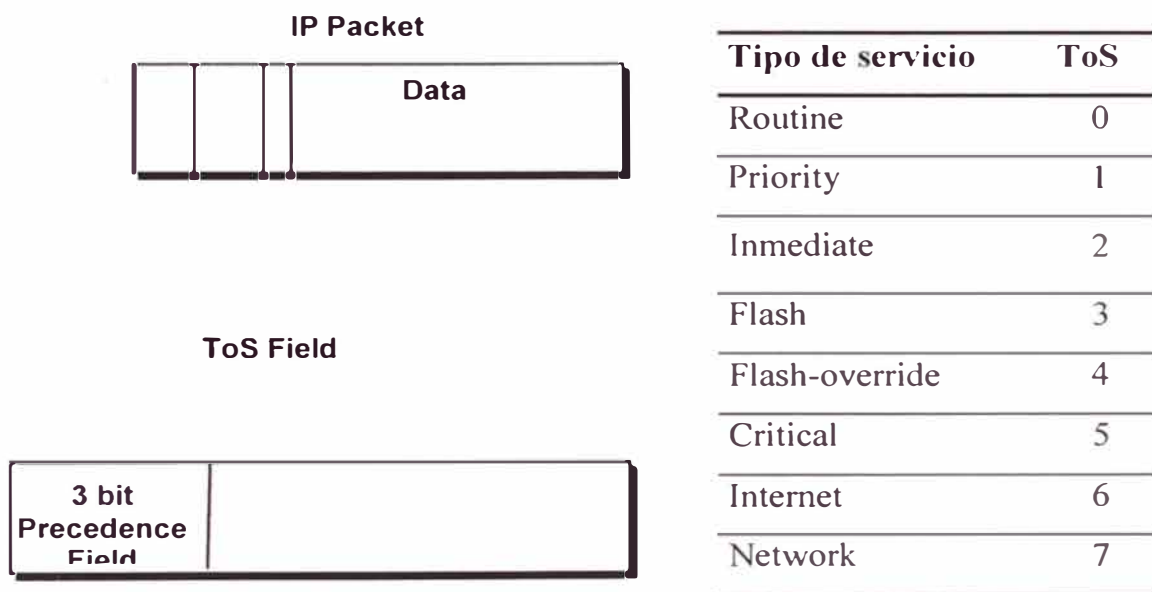


FIGURA N° 22 – PRECEDENCIA IP

Los valores de precedencia 6 y 7 están reservados para información de red (routing updates, hello packets, etc). Esto nos deja seis niveles para el flujo de tráfico IP normal. Para el tráfico de voz se usa el valor de precedencia en 5. En el caso de tráfico de videoconferencia el valor recomendado es 4. La precedencia IP permite a un router agrupar los flujos de tráfico basado en su valor correspondiente.

También se puede considerar a la precedencia IP como un mecanismo de QoS en banda. No se involucra señalización adicional, no se agrega cabecera al paquete existente. Dados estos beneficios, la precedencia IP es uno de los mecanismos de QoS que se usa frecuentemente.

Para asignar la precedencia IP en un router Cisco se puede hacer de diversas formas. Con el diseño de voz sobre IP de Cisco se puede asignar la precedencia IP basado en el destino (el número llamado). La configuración en el router Cisco debe hacerse manualmente por el administrador de la red y sería de la siguiente forma:

```
!  
dial-peer voice 10 voip  
    destination-pattern 456  
    ip precedence 5  
    session target ipv4:172.20.10.1  
!
```

En este caso los paquetes que corresponden a las llamadas al anexo 456 se van enviar vía voz sobre IP al router con dirección IP 172.20.10.1 y el campo de ToS será marcado con 15, es decir prioridad crítica. Asimismo, se les puede asignar la precedencia a los diferentes tipos de tráfico en función de listas de acceso.

#### 5.4.6 Differentiated services

Diff serv, usa ToS a nivel 3 (IP). Está orientado a operar en los bordes de la red y en el CPE. El ToS fue redefinido para su utilización en Diff Serv (RFC 2474, 2475) mediante la utilización de los primeros seis bits en lugar de sólo tres de los ocho bits (ver figura N° 23). Con esto se logra tener mayor granularidad para la clasificación de los flujos de tráfico ya que se cuenta con 64 ( $2^6$ ) niveles. El DS byte especifica el DSCP (Diff Serv Code Point) y el PHB (per-hop-behaviour). El PHB indica como debe tratarse al paquete en cada nodo. Está implementado en los router Motorola Vanguard y para el caso de tráfico de voz sobre IP se asigna el valor de 46. Se tiene dos opciones de clasificaciones: MF: Multifield o BA: Behaviour Aggregate

Un nodo Diffserv hace.

- Re-marking del DS (asignacion o cambio del TOS)
- Shaping (retardo del tráfico para conformar un perfil)
- Policing (descarte de cierto trafico para asegurar al resto)

Provee propiedades con mayor granularidad que el MPLS

Estos nodos están condicionados por el TCA (Traffic Conditioning Agreement)

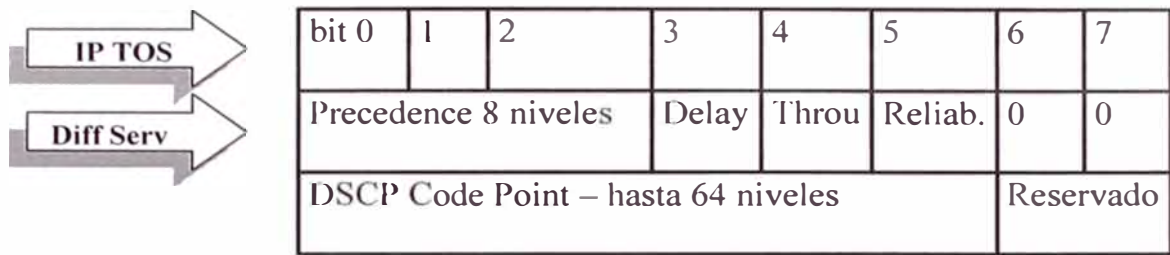


FIGURA N° 23 – DIFFERENTIATED SERVICE

#### 5.4.7 Policy – based routing – PBR

También conocido como “Layer – 4 switching”. Provee políticas de ruteo para los flujos de tráfico basadas en definiciones del administrador de red. Convierte a las políticas en el comportamiento de la red. Se especifica en la interface que reciben los paquetes, no en la interface que envía los paquetes. También permite configurar el campo de precedencia IP de modo que la red pueda utilizar diferentes clases de servicio.

Permite:

- Balance de carga
- Políticas horarias
- Políticas de seguridad
- Políticas tarifarias
- Políticas de backup.

Las políticas se pueden basar en las direcciones IP, número de puerto, protocolo o tamaño de los paquetes. Entonces, se cuenta con diversos descriptores con los que podemos seleccionar el tráfico y crear políticas sencillas o hacer una combinación de varias de ellas para crear políticas un poco más complejas.

Todos los paquetes recibidos en una interface con PBR son pasados a través de un filtro avanzado conocido como “route – maps”. Los “route – maps” indicarán a donde deben ser enviados los paquetes.

También se pueden usar listas de acceso estándar o extendidas para establecer los criterios de correspondencia, para los que se puede usar como items de clasificación las direcciones IP fuente y/o destino, aplicación, tipo de protocolo, ToS y precedencia IP.

Las cláusulas de correspondencia también pueden incluir un rango de tamaño de paquetes, con lo que el administrador de la red puede distinguir el tráfico interactivo del tráfico voluminoso (este tipo de tráfico generalmente tiene paquetes de longitud grande, como por ejemplo para el caso de tráfico FTP o HTTP). Por defecto el tráfico originado en un router no es enviado directamente a un “route – map”. Para que esto suceda debe ser configurado. Se debe tener cuidado al momento de escoger las políticas a usar con el PBR. Priorizar en el desktop o en el CPE elimina este problema.



#### 5.4.8 Servidor de control de QoS

Administra en forma centralizada las políticas de QoS. Resulta complejo administrar todo configurando cada equipo, por lo que se centraliza en este servidor. Las políticas definen que tráfico es prioritario y esto puede ser dinámico, cambiando para distintas horas del día por ejemplo. El servidor genera los comandos necesarios para cada CPE.

COPS: Common Open Policy Service del IETF, per-flow protocol. Es un protocolo cliente servidor para comunicar información de políticas de tráfico de red.

Usan SNMP (Simple Network Management Protocol) y/o LDAP ( Lightweight Directory Access Protocol ) para distribuir las políticas.

Su limitación es que está restringido a redes corporativas. No es escalable a un backbone público.

#### 5.4.9 802.1p & 802.1Q

Protocolos de IEEE para dar QoS en Ethernet.

802.1Q: marker tag

- 12 bits para identificar VLAN (VLAN ID)
- 3 bits en el header para indicar prioridad, no usados

802.1p: spin/off del 802.1Q

- Priorización del tráfico (Best Effort)

- 3 bits => 8 niveles de prioridad en la MAC, Layer 2
- Transporta cualquier tráfico: IP, IPX, Apple Talk, SNA, etc.

Brindan QoS relativa, pero en general el ancho de banda es importante: 10/100/1000 Mbps.

#### 5.4.10 Traffic shaping & policing

Es necesario regular o limitar la cantidad de tráfico que se permiten a las aplicaciones para que envíen a través de varias interfaces o redes. El policing verifica que el tráfico cumpla los compromisos antes de ingresar a la red y descarta el exceso.

Cisco tiene unas herramientas para definir cuanto ancho de banda puede usar cierta aplicación. Se tiene dos formas de controlar el uso de ancho de banda: *rate limiting* (tales como CAR – Committed Access rate) y *traffic shaping* (por ejemplo GTS – Generic Traffic Shaping.).

La diferencia principal entre estas dos herramientas de regulación de tráfico es que las primeras descartan el tráfico basado en políticas y las otras almacenan el tráfico en exceso en un buffer mientras espera un nuevo intervalo de tiempo para poder transmitir, con lo que trata de evitar el descarte de los paquetes. Ambos identifican el exceso basado en los umbrales definidos por el administrador de la red.

Frecuentemente ambas herramientas se usan juntas. Traffic shaping se usa en los bordes de la red (CPE) para asegurar que el cliente está utilizando el ancho de banda necesario. Rate limit es usado frecuentemente por los proveedores de servicio de red para asegurar que los abonados no excedan la cantidad de ancho de banda contratado.

El mecanismo de rate limit de CAR permite a los usuarios:

- Controlar la máxima razón de tráfico transmitido o recibido en una interface.
- Proporcionar control granular a capa 3, que permite a una red IP exhibir cualidades de una red TDM.

Se puede usar como items de selección: la precedencia IP, número MAC, dirección IP, u otros parámetros. También se pueden usar listas de acceso.

Es muy importante tener en cuenta que esta técnica no almacena los paquetes en un buffer, por lo que se recomienda su uso en ambientes de alta velocidad, ya que el queueing no agrega retardo.

Para configurar CAR en un router Cisco se usa la siguiente línea de comando en la interface correspondiente:

```
rate-limit {input|output} bps burst-normal burst-max conform-action action  
exceed-action action
```

- Con las opciones *input|output* se elige la dirección de los paquetes: entrantes o salientes.
- *bps*, indica el promedio de la tasa de bits para la interface.
- *burst-normal* y *burst-max*, se determina cuanto tráfico en ráfaga se permite.
- Con las opciones *action* se le indica al router que se debe hacer con los paquetes que conforman el tráfico y los que exceden.

Para el parámetro *action* se tienen las opciones que se mencionan en la tabla N° 13.

Acción	Descripción
<b>continue</b>	Evalúa el siguiente comando rate-limit.
<b>drop</b>	Descarta el paquete.
<b>set-prec-continue new-prec</b>	Asigna la precedencia IP y evalúa el siguiente comando rate-limit.
<b>set-prec-transmit new-prec</b>	Asigna la precedencia IP y transmite el paquete.
<b>transmit</b>	Transmite el paquete.

**TABLA N° 13 – RATE LIMIT ACCES RATE**

CAR sólo funciona para tráfico IP. No es aplicable si se tiene una red multiprotocolo (por ejemplo: IP/IPX). Asimismo, no es soportado en

interfaces Fast Ether Channel, Tunnel, PRI (Primary Rate Interface) y cualquier otra interface que no soporte CEF (Cisco Express Forwarding).

En la implementación de Cisco hay dos tipos de traffic shaping: GTS (Generic TS) y FRTS (Frame Relay TS). Ambos métodos son similares en implementación, aunque las líneas de comando difieren en el tipo de colas que se usan para almacenar el tráfico.

En el caso de GTS se usa WFQ, y para el caso de FRTS se usa PQ o CQ. GTS es configurable por interface o sub-interface. FRTS se basa en cada DLCI (Data Link Connection Identifier).

TS permite controlar el tráfico saliente de la interface. Es requerido cuando un extremo de un circuito tiene mucha mayor velocidad que el remoto. Requiere de suficiente buffer para amortiguar las ráfagas de tráfico y distribuirlas en el tiempo.

El uso de TS es primariamente para:

- Controlar el uso de ancho de banda disponible.
- Establecer políticas para el tráfico.
- Regular el flujo de tráfico para eliminar la congestión.

Se puede usar TS en las siguientes situaciones:

- Si se tiene una red con diferentes tasas de bit para el acceso.

- Si se ofrece un servicio de sub-rate. En este caso TS permite hacer una partición de un circuito, un E1 por ejemplo, en otros de menor ancho de banda.

TS previene la pérdida de paquetes. Es especialmente importante usar TS en redes Frame Relay debido a que los Switch de Frame Relay no pueden determinar que paquetes tiene precedencia y por lo tanto que paquetes deben ser descartados cuando hay congestión.

Para configurar TS en una interface se usa el siguiente comando:

```
traffic-shape rate bps [burst.size [excess-burst.size]]
```

Donde:

- *bps*, es la tasa de bits promedio en bps en la interface. Este el ancho de banda contratado con el proveedor de servicio.
- *burst-size*, (opcional) es el número de bits que se puede transmitir en forma sostenida. Su valor por defecto es bps dividido entre ocho.
- *excess-burst-size*, (opcional) es el máximo número de bits que se puede exceder en el primer intervalo cuando se produce la congestión. Su valor por defecto es igual al de burst-size.

Para usar TS basado en listas de acceso se usa el comando: **traffic-shape group**.

#### 5.4.11 Control de flujo

Para evitar el descarte de tramas es necesario el control de flujo en situaciones de congestión.

El Frame Relay utiliza los bits de notificación explícita: BECN y FECN. El IWF entre FR y ATM convierte al BECN en GFC y viceversa. El CPE debe reducir su flujo de datos ante la recepción de BECN, esto generalmente se hace gradualmente.

El TCP/IP tiene como control el flujo un sistema de ventanas, el que es estimulado con el RED (Random Early Detection) y/o WRED (Weighted RED).

RED es un mecanismo de control de congestión que es potencialmente útil, especialmente en redes de tránsito de alta velocidad. La teoría detrás de WRED es, simplemente, que muchos transportes de datos son, de algún modo, sensible a las pérdidas y al menos momentáneamente lentos si una parte de su tráfico es descartado.

Para indicar a una estación TCP que deje de transmitir sencillamente se descarta parte del tráfico que envía. WRED combina las capacidades del algoritmo RED con la precedencia IP. Esta combinación provee manipulación

preferencial para el tráfico de alta prioridad. Puede descartar selectivamente tráfico de baja prioridad cuando la interface empieza a congestionarse.

Para habilitar WRED en un router Cisco se usa el siguiente comando en la interface:

**random-detect [*weighting*]**

Donde el parámetro *weighting* es opcional. Es una constante en el rango de 1 a 16 usado para determinar la tasa de paquetes descartados cuando ocurre la congestión; su valor por defecto es 10 (esto es, descarta 1 de  $2^{10}$  paquetes).

Los tráficos “real time” no admiten control de flujo.

Si el tráfico es solo TCP/IP non-real time:

- Reducción de ventanas por tiempo muerto.
- Para evitar la sincronización de los emisores en un nodo se usa RED.
- Para evitar descartar tráfico que no impacta se usa WRED.
- Ambos mecanismos no funcionan para otros tráficos como IPX.
- No hay comunicación entre estos mecanismos y los de nivel 2.
- IEEE 802.3
- Frame Relay
- ATM



## 5.5. Clases de servicio – COS

En la figura N° 23 se muestran algunos tipos de clase de servicio y la aplicación respectiva. Mas adelante se enumeran las clases de servicio para IP y Frame Relay.

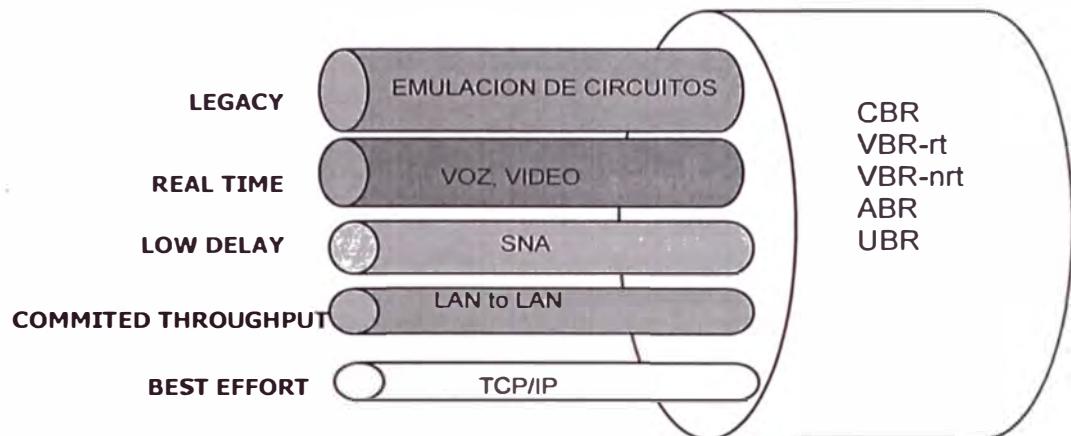


FIGURA N° 23 – CLASS OF SERVICE

### 5.5.1 IP – COS

- IP Precedence (prioridad, IP )
- Int serv (no fue escalable para DSCP)
- IP ToS (redefinido para DSVP)
- Diff Serv (da prioridad, QoS relativa)
- RSVP (reserva BW, no es escalable)
- MPLS y BGP/MPLS (aún no desarrollado)
- COPS (Common Open Policy Server) limitado a VPNs

Actualmente el Forum de voz sobre IP e IMTC (International Multimedia Teleconferencing Consortium). trabajan conjuntamente con un mismo objetivo

Común: establecer los estándares abiertos que satisfagan los requerimientos en tiempo real y alta calidad de servicio (QoS) para la telefonía sobre Internet y Redes Privadas de IP

### 5.5.2 FR – COS & QoS

Estándares: ITU X.144, X.146 y FRF.13

QoS:

- CIR: Committed Information Rate
- FTD: Frame Transfer Delay
- FLR: Frame Loss Rate

CoS: Class of Service

- Por prioridad (relativo)
- Por PVC (absoluto)

### 5.6. Service Level Agreement – SLA

Un Service Level Agreement (contrato de nivel de servicio) es una definición formal que existe entre dos organizaciones, usualmente entre un proveedor de servicios y sus clientes. El proveedor puede ser una organización IT que presta servicios de computación, un proveedor de intranets que ofrece servicios de conectividad de red, una compañía de telecomunicaciones que provee servicios de telefonía a sus clientes, una compañía de soporte de mantenimiento y administración de computadoras o un ISP (Internet Service Provider) que ofrece acceso a Internet a sus clientes. Los SLA

pueden ser definidos en el contexto industrial en el cual existe una relación proveedor – cliente.

El SLA provee el medio por el cual se define el nivel de servicio requerido por el cliente. Este especifica que es lo que el cliente quiere y que es lo que el proveedor puede ofrecer. Define los estándares para la calidad de servicio proveído y los objetivos de performance que el proveedor debe cumplir. También define los procedimientos y los reportes que deben ser dados para rastrear y asegurar el cumplimiento del SLA.

Frame Relay puede tener SLAs definidos para

- Cada VPN
- Cada puerta
- Cada PVC

Las redes IP solo tienen SLAs definidas para toda la red

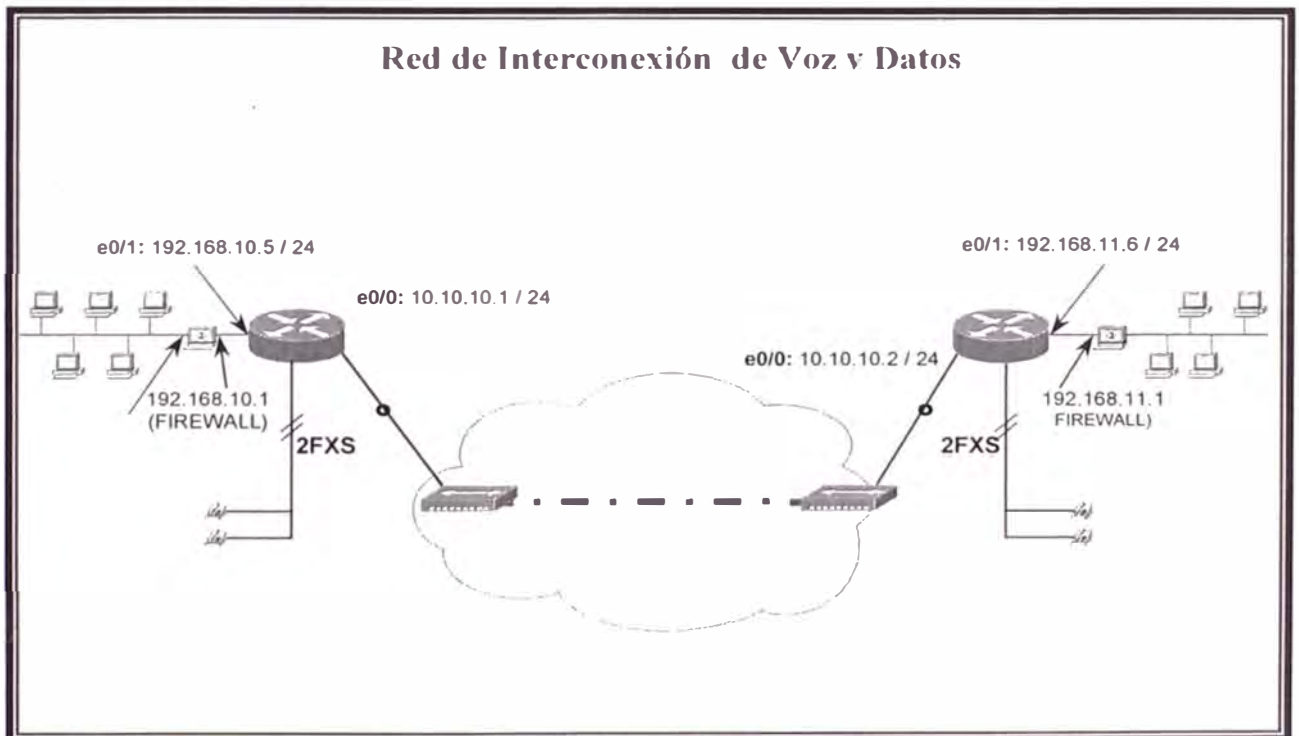
- Disponibilidad
- Round trip delay
- Call success rate.

## CAPÍTULO VI

### IMPLEMENTACIÓN USANDO ROUTER CISCO

#### 6.1. Escenario de pruebas

Se hizo una implementación práctica de las herramientas de calidad de servicio expuestas en este trabajo. El escenario de pruebas se muestra en la figura N° 24. Los CPE usados son dos router Cisco.2611, los cuales tiene 02 interfaces Ethernet cada uno.



**FIGURA N° 24 – ESCENARIO DE PRUEBAS PARA LA APLICACIÓN DE CALIDAD DE SERVICIO.**

Se trata de un enlace punto a punto a 64 Kbps para tráfico de voz y datos en una red IP vía un PVC (Circuito Virtual Permanente) en ATM usando switches Ethernet como equipo de acceso a la red con una interface ATM, los cuales hacen uso del RFC 1483 para interworking con ATM. En este caso la configuración en ambos routers es similar, sólo se diferencian en las direcciones IP y la numeración de los anexos. Para la comunicación de voz se usan interfaces FXS a la cual se conectan directamente los teléfonos analógicos. La solución que se presenta a continuación es aplicable a cualquier otro esquema de telefonía analógica y puede extenderse fácilmente a enlaces punto a multipunto. Es muy importante notar que cuando se usa otra tecnología tal como ATM o si el medio de enlace de datos cambia la solución se modifica ligeramente.

Bajo situaciones de alto tráfico de datos, llegando incluso a la saturación del ancho de banda de 64 Kbps, va a ser muy difícil realizar una comunicación de voz con buena calidad. Es decir, la voz se escuchará entrecortada, sólo se entenderán algunas palabras, va a causar molestias a los interlocutores. Se hace esta consideración ya que el enlace es usado primariamente para cursar tráfico de datos. Es fácil darse cuenta que en ausencia de cualquier otro tipo de tráfico o con tráfico “liviano” la conversación va a ser nítida.

La solución planteada es la siguiente: se marca el tráfico de interés, en este caso voz sobre IP, y se le asigna una prioridad sobre otras aplicaciones para el uso del ancho de banda. Se debe conocer las características del tráfico a priorizar. Esto se puede

saber haciendo uso de un analizador de protocolos. Para nuestro caso se usó el SNIFFER PRO LAN.

## 6.2. Configuración del router Cisco

```
!  
!  
! DEFINICION DE CLASES  
!  
class-map qos  
  match access-group 102  
!  
! DEFINICION DE POLITICA  
!  
policy-map wan  
  class qos  
    priority 30  
!  
! INTERFACE DE SALIDA  
!  
interface Ethernet0/0  
  description Sede Principal  
  ip address 10.10.10.1 255.255.255.0  
  no ip directed-broadcast  
  service-policy output wan  
  traffic-shape rate 64000 8000 8000 500  
  ip rsvp bandwidth 7500 7500  
!
```

**! LISTAS DE ACCESO PARA DEFINIR PAQUETES DE VOZ**

!

**access-list 102 permit ip any any precedence critical****access-list 102 permit tcp any any eq 1720****access-list 102 permit tcp any eq 1720 any****access-list 102 permit udp any any range 16384 32767**

!

**! ANEXOS LOCALES**

!

dial-peer voice 1 pots

destination-pattern 100

port 1/0/0

!

dial-peer voice 2 pots

destination-pattern 110

port 1/0/1

!

**! LLAMADAS SALIENTES**

!

dial-peer voice 10 voip

destination-pattern 2..

session target ipv4:10.10.10.2

**codec g729r8 bytes 80****ip precedence 5**

!

### 6.3. Explicación de la configuración

Con las listas de acceso 102 se definen los paquetes de voz. Los paquetes de voz sobre IP tienen las siguientes características cuando se implementa en un router Cisco (varía según el fabricante de los routers): precedencia crítica (5), usa el puerto TCP 1720 para establecimiento de llamada mediante H.323 y el tráfico de la conversación misma viaja sobre el protocolo UDP, usando puertos dinámicos en el rango de 16384 a 32767.

Se define el class-map “qos” en el cual se consideran todos los paquetes que cumplan con las condiciones del access-list 102.

Luego, se definen la política “wan”, en la cual se dice que para el tráfico de clase “qos” se le asigne un ancho de banda reservado con prioridad estricta. En este caso se está probando dos canales de voz de 14 Kbps usando codec G.729 con bundling a 80 bytes (cuatro paquetes en uno), por lo que el ancho de banda que se reserva es 30 Kbps (se dá un margen por seguridad).

Se aplica la política “wan” en la interface ethernet 0/0, en la dirección de salida hacia la red.

Se tiene los “dial-peer voice” tipo POTS 1 y 2 para definir los anexos 100 y 110, que corresponden a los teléfonos que se conectarán al router Cisco cuya configuración se



muestra. Asimismo, se define el “dial-peer voice” tipo VOIP número 10 para las llamadas salientes. Los anexos remotos son de la serie 2xx.

## 6.4. Resultados obtenidos

### 6.4.1 Saturación del ancho de banda

Solo se aplicó tráfico de datos para esta primera parte. Se trata de saturar el ancho de banda con el objeto de simular la peor condición para el tráfico de voz que se generará posteriormente. En primer lugar se saturó el ancho de banda de 64 Kbps con tráfico FTP en ambos sentidos. Este tipo de tráfico tiene la característica que se aplica sobre todo el ancho de banda disponible para su transmisión, es decir si se dispone de un ancho de banda de 128 Kbps este se usa completamente.

```
r-Agencia1#show policy-map interface e0/0  
Ethernet0/0
```

```
Service-policy output: wan
```

```
Class-map: voip (match-all)
```

```
157863 packets, 11763536 bytes
```

```
30 second offered rate 0 bps, drop rate 0 bps
```

```
Match: access-group 102
```

```
Weighted Fair Queueing
```

Strict Priority

Output Queue: Conversation 264

Bandwidth 30 (kbps) Burst 875 (Bytes)

Class-map: class-default (match-any)

36939 packets, 18270119 bytes

**30 second offered rate 64000 bps, drop rate 0 bps**

Match: any

r-Agencial#

**r-Agencial#show interface e0/0**

Ethernet0/0 is up, line protocol is up

Hardware is AmdP2, address is 0005.9bb4.4cc0 (bia 0005.9bb4.4cc0)

Internet address is 10.10.10.2/29

MTU 1500 bytes, BW 10000 Kbit, DLY 1000 usec,

reliability 255/255, txload 3/255, rxload 3/255

Encapsulation ARPA, loopback not set

Keepalive set (10 sec)

ARP type: ARPA, ARP Timeout 04:00:00

Last input 00:00:00, output 00:00:00, output hang never

Last clearing of "show interface" counters 1d01h

Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0

Queueing strategy: weighted fair

Output queue: 0/1000/64/0 (size/max total/threshold/drops)

Conversations 0/2/256 (active/max active/max total)

Reserved Conversations 0/1 (allocated/max allocated)

Available Bandwidth 7465 kilobits/sec

**30 second input rate 64000 bits/sec, 26 packets/sec**

**30 second output rate 63000 bits/sec, 25 packets/sec**

388728 packets input, 101128153 bytes, 0 no buffer

Received 1555 broadcasts, 0 runts, 0 giants, 0 throttles

0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

0 input packets with dribble condition detected

408249 packets output, 104362158 bytes, 0 underruns

0 output errors, 0 collisions, 5 interface resets

0 babbles, 0 late collision, 0 deferred

0 lost carrier, 0 no carrier

0 output buffer failures, 0 output buffers swapped out

#### 6.4.2 Prueba con llamada

Luego se hicieron pruebas con una llamada de VoIP. En la salida del comando "*show policy interface*" se aprecia el trabajo de las herramientas de QoS.

**rMain3.8BPX#show voice port summary**

PORT	CODEC	VAD	VTSP STATE	VPM STATE
2/0/0	g729r8	y	S_CONNECT	FXSLS_CONNECT
2/0/1	-	-	-	FXSLS_ONHOOK

**r-Agencia1#sh pol int e0/0**

Ethernet0/0

Service-policy output: wan

Class-map: voip (match-all)

163953 packets, 12217463 bytes

**30 second offered rate 14000 bps, drop rate 0 bps**

Match: access-group 102

Weighted Fair Queueing

Strict Priority

Output Queue: Conversation 264

Bandwidth 30 (kbps) Burst 875 (Bytes)

Class-map: class-default (match-any)

43103 packets, 21685963 bytes

**30 second offered rate 50000 bps, drop rate 0 bps**

Match: any

r-Agencia1#

**r-Agencia1#sh int e0/0**

Ethernet0/0 is up, line protocol is up

Hardware is AmdP2, address is 0005.9bb4.4cc0 (bia 0005.9bb4.4cc0)

Internet address is 10.10.10.2/29

MTU 1500 bytes, BW 10000 Kbit, DLY 1000 usec,

reliability 255/255, txload 3/255, rxload 3/255

Encapsulation ARPA, loopback not set

Keepalive set (10 sec)

ARP type: ARPA, ARP Timeout 04:00:00

Last input 00:00:00, output 00:00:00, output hang never

Last clearing of "show interface" counters 1d02h

Input queue: 1/75/0/0 (size/max/drops/flushes); Total output drops: 0

Queueing strategy: weighted fair

Output queue: 0/1000/64/0 (size/max total/threshold/drops)

Conversations 0/2/256 (active/max active/max total)

Reserved Conversations 0/1 (allocated/max allocated)

Available Bandwidth 7465 kilobits/sec

**30 second input rate 64000 bits/sec, 72 packets/sec**

**30 second output rate 65000 bits/sec, 72 packets/sec**

400084 packets input, 104675653 bytes, 0 no buffer

Received 1559 broadcasts, 0 runts, 0 giants, 0 throttles

0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

0 input packets with dribble condition detected

419667 packets output, 107991868 bytes, 0 underruns

0 output errors, 566 collisions, 0 interface resets

0 babbles, 0 late collision, 1530 deferred

0 lost carrier, 0 no carrier

0 output buffer failures, 0 output buffers swapped out

r-Agencial#

Se puede apreciar que aunque se tiene el ancho de banda de 64000 bps saturado, las herramientas aplicadas en la configuración del router Cisco separan 14000 bps para el tráfico de voz sobre IP del total de 30Kbps que se reservó. Si se hacen dos llamadas los algoritmos de queueing y QoS reservan un total de 28000 bps para ambas llamadas y reducen el ancho de banda para el tráfico restante a 36000 bps. Esto afecta la performance del tráfico de datos, pero permite tener un buen nivel de conversación. Hay que tener en cuenta que se debe llegar a un compromiso entre los recursos para las aplicaciones y dar la prioridad respectiva al tráfico de voz ya que es una aplicación en tiempo real. Se verificó que la calidad de la voz era muy buena y el usuario no percibía ningún problema como entrecortes, retardo, eco, etc...

## CONCLUSIONES

1. El concepto de recibir un paquete en el instante que el transmisor espera es muy complejo. Ciertamente, construir una red para correr tráfico de voz es complejo. El paso más importante para asegurar buena calidad de voz es escoger las herramientas adecuadas de QoS.
2. Los cuellos de la botella limitan la capacidad de dar QoS:
  - Ancho de banda, LAN, última milla, WAN, tramo internacional.
  - CPU y memoria: Workstation, CPE Edge devices, Service Server.
3. Se tienen soluciones para diversos escenarios, es decir voz sobre IP, voz sobre ATM, Videoconferencia, enlaces punto a punto, punto a multipunto, etc. Estas han sido probadas e implementadas en la práctica con éxito.
4. Las aplicaciones sobre IP requieren de configuración detallada extremo a extremo para brindar QoS.
  - Maximizar el tramo orientado a conexión (FR y/o ATM).
  - Distribuir procesos fuera de la red hacia el edge o CPE.
  - Codificar, comprimir y encriptar en hardware.
  - Configurar cada tramo en forma consistente.

5. Antes de implementar QoS en un borde de la red debe hacerse las siguientes preguntas:
  - ¿Se tiene un circuito WAN de bajo ancho de banda?  
Si la respuesta es si, use CRTP. Escoger un método de fragmentación. No se recomienda cambiar el parámetro MTU o IP MTU.
  - ¿El tráfico necesita ser priorizado en los circuitos WAN?  
Use algún método de queueing (LLQ, CB-WFQ).
  - Si se tiene una red Frame Relay con topología hub-and-spoke use Traffic Shaping.
  
6. Todos los medios para asegurar QoS sobre IP son válidos, pero no alcanzan.
  
7. Lograr la consistencia y QoS extremo a extremo dependen en gran medida de buena: ingeniería de red, configuración y control.



## BIBLIOGRAFÍA

- INTEGRATING VOICE AND DATA NETWORKS – Scott Keagy, Cisco Press 2000.
- VOICE OVER IP FUNDAMENTALS – Jonathan Davidson & James Peters, Cisco Press 2000.
- VOICE OVER IP – DESIGN IMPLEMENTATION GUIDE – Jonathan Davidson, White Paper Cisco Systems 1998
- CVOICE - CISCO VOICE OVER FRAME RELAY, ATM AND IP – Steve Mc Querry, Kelly McGrew & Stephen Foy, Cisco System 2000.
- SUPPORTING SERVICE LEVEL AGREEMENTS ON IP NETWORKS – Dinesh Verma, Macmillan Technical Publishing 1999.
- IP QUALITY OF SERVICE – Srinivas Vegesna, Cisco Press 2001
- CONFIGURING CISCO VOICE OVER IP – Elliot Lewis, Syngress Media, 2000.