



**UNIVERSIDAD DE QUINTANA ROO**  
**DIVISIÓN DE CIENCIAS E INGENIERÍA**

---

**Evaluación de la Calidad de Servicio de un Sistema de Voz sobre  
una red Inalámbrica Ad Hoc**

---

**TESIS**

Para obtener el grado de  
**Ingeniero en Redes**

**PRESENTA**

**José Efraín Beltrán Ávila**

**DIRECTOR DE TESIS**

**Dr. Homero Toral Cruz**

**ASESORES**

**Dr. Freddy Ignacio Chan Puc**

**Dr. Jaime Silverio Ortegón Aguilar**

**Dr. Gliserio Romeli Barbosa Pool**



**UNIVERSIDAD DE QUINTANA ROO**  
**DIVISIÓN DE CIENCIAS E INGENIERÍA**

Trabajo de Tesis elaborado bajo supervisión del Comité de asesoría y aprobada  
como requisito parcial para obtener el grado de:

**INGENIERO EN REDES**

**Comité de Trabajo de Tesis**

**Director:**

  
\_\_\_\_\_  
**Dr. Homero Toral Cruz**

**Asesor:**

  
\_\_\_\_\_  
**Dr. Freddy Ignacio Chan Puc**

**Asesor:**

  
\_\_\_\_\_  
**Dr. Jaime Silverio Ortega Aguilar**



Chetumal Quintana Roo, México, Noviembre de 2015



## Agradecimientos

Agradezco a mis padres, la señora Elsy Noemí Ávila Herrera y al señor José Luis Beltrán Medina, primero por haberme regalado el don de la vida, por haberme apoyado incondicionalmente durante toda mi carrera, haberme educado e inculcado valores, llenarme de consejos y enseñanzas, por no poner un pretexto de por medio ante cualquier situación, si no, ayudándome a resolver y encontrar solución.

Agradezco a mis hermanos y hermanas por haberme alentado a llegar hasta donde estoy, por ponerme como ejemplo a seguir y por ese cariño mutuo e incondicional que siempre me han ofrecido.

A mis abuelos por ser un ejemplo de vida, por esos jalones de orejas y muchos consejos, que al final sirvieron para lograr una meta más en mi vida, a mis tíos que durante mi carrera me apoyaron y que siempre me proporcionaban experiencias y consejos.

A la persona especial que siempre está a mi lado y me apoya en mis decisiones y me alienta a seguir superándome.

A todos mis familiares que siempre me apoyaron moralmente y que gracias a ellos tuve un motivo para poder sobresalir.

Y sin lugar a duda agradecerles a mis maestros de la universidad que fueron los que me guiaron hacia este camino, los que me enseñaron y los que me apoyaron en toda la carrera.

En especial al Dr. Homero Toral Cruz, tutor durante mi formación académica y asesor de tesis, que gracias a él fui capaz de realizar y concluir esta tesis.

## Dedicatoria

Con mucho cariño a mis padres, mi mamá Elsy y mi papá José, les dedico esta tesis que con mucho sacrificio y esfuerzo hemos logrado juntos con el apoyo que me han brindado, y sobre todo por la oportunidad que me otorgaron de forjar la puerta de mi futuro, gracias por estar orgullosos de mí, los quiero mucho, siempre ocupan el lugar más importante en mi vida, gracias por todo.

A toda mi familia en general y mis seres queridos, les dedico y agradezco que estuvieron siempre pendientes a mí y me apoyaron en todo momento, siempre habrá un lugar para ustedes en mi corazón.

Al Dr. Homero Toral Cruz, por apoyarme, por brindarme sabios consejos y por estar al pendiente y motivarme a lograr la realización de este trabajo.

# Contenido

Capítulo 1. INTRODUCCIÓN.....	1
1.1 Planteamiento del problema .....	1
1.2 Justificación .....	3
1.3 Objetivo general.....	4
1.4 Objetivos particulares .....	4
1.5 Metodología .....	5
Capítulo 2. REDES DE TELECOMUNICACIONES .....	7
2.1 Redes de conmutación por circuito .....	7
2.1.1 Red telefónica publica conmutada (PSTN).....	8
2.1.2 Red digital de servicios integrados (ISDN).....	10
2.2 Red de conmutación por paquetes.....	11
2.2.1 Internet.....	13
2.2.2 Protocolos de nivel de transporte.....	14
Capítulo 3. TRANSMISIÓN DE VOZ SOBRE REDES DE DATOS.....	26
3.1 Ventajas de VoIP .....	26
3.1.1 Funcionalidad .....	27
3.1.2 Movilidad.....	27
3.1.3 Servicios de valor añadido.....	28
3.1.4 Mantenimiento y gestión integrada.....	28
3.2 La señal de la voz.....	29
3.2.1 Muestreo .....	30
3.2.2 Cuantificación.....	31
3.2.3 Codificación de la voz.....	32
3.2.4 Codecs (codificador-decodificador) .....	35
3.3 Protocolos de señalización .....	37
3.3.1 H.323 .....	39
3.3.2 SIP .....	40
Capítulo 4. ARQUITECTURA H.323 .....	45
4.1 Entidades H.323 .....	46
4.2 Funcionamiento .....	50
4.2.1 Descubrimiento de Gatekeeper y registro de la terminal.....	51

4.2.2 Fase 1: La llamada .....	53
4.2.3 Fase 2: Inicio de la comunicación e intercambio de capacidades.....	54
4.2.4 Fase 3: Establecimiento de la comunicación audio visual .....	55
4.2.5 Fase 4: Servicios de llamada.....	55
4.2.6 Fase 5: Terminación de la llamada .....	56
Capítulo 5. CALIDAD DE SERVICIO EN VOZ SOBRE PROTOCOLO IP (QoS).....	59
5.1 Concepto de calidad de la voz.....	59
5.2 Factores que influyen en la calidad.....	61
5.2.1 Disponibilidad.....	62
5.2.2 Jitter.....	63
5.2.3 Pérdidas.....	64
5.2.4 Retardo.....	66
5.2.5 Ancho de banda.....	69
5.2.6 Correlación entre paquetes.....	71
5.3 Medida de la calidad de la voz .....	73
5.3.1 ITU-T P.800 (Escala MOS) .....	75
5.3.2 Modelado perceptual de la voz.....	78
5.3.3 Modelo E .....	82
5.3.4 VQMon .....	84
Capítulo 6. MEDICIÓN DE PARÁMETROS DE QOS EN UNA RED VOIP H.323 .....	86
6.1 Escenarios de medición.....	86
6.1.1 Descripción de los equipos.....	88
6.1.2 Configuración de los equipos .....	94
6.1.3 Software empleado .....	97
6.2 Características de la transmisión de flujos de voz (llamadas de pruebas).....	98
6.3 Capturas .....	100
Capítulo 7. ANÁLISIS DE DESEMPEÑO Y CARACTERIZACIÓN DE PARÁMETROS DE QoS.....	108
Capítulo 8. CONCLUSIONES.....	116
REFERENCIAS.....	118

## Índice de Figuras

Figura 1. Conmutación por circuitos .....	8
Figura 2. Conmutación por paquetes.....	12
Figura 3. Protocolo TCP/IP .....	15
Figura 4. Protocolo UDP .....	17
Figura 5. Protocolo TCP .....	19
Figura 6. Tráfico RTP y RTCP a través de la red.....	22
Figura 7. Encapsulamiento de los paquetes de voz.....	23
Figura 8. Digitalización de la señal de voz.....	29
Figura 9. Muestreo de la voz.....	30
Figura 10. Cuantificación de la voz.....	32
Figura 11. Codificación de la voz .....	33
Figura 12. Protocolos de la familia H.323 .....	45
Figura 13. Arquitectura H.323.....	47
Figura 14. Llamada de terminal directa y control de señalización.....	49
Figura 15. Búsqueda de Gatekeeper y registro de la terminal.....	52
Figura 16. Establecimiento de la llamada con señalización directa.....	53
Figura 17. Inicio de la comunicación e intercambio de capacidades con señalización directa .....	54
Figura 18. Establecimiento de comunicación audio visual con señalización directa.....	55
Figura 19. Terminación de llamada con señalización directa .....	56
Figura 20. Señalización completa de una llamada H.323.....	57
Figura 21. Relación entre la inteligibilidad y la calidad de la codificación de la voz .....	61
Figura 22. Disponibilidad.....	63
Figura 23. Jitter.....	64
Figura 24. Pérdidas de paquetes.....	66
Figura 25. Relación entre el retardo en un sólo sentido y la calidad de la voz.....	67
Figura 26. Puntos donde se puede activar el VAD .....	70
Figura 27. Medidas no pasivas .....	74
Figura 28. Comparación entre estándares de la calidad de la voz.....	75
Figura 29. Evaluación MOS.....	76
Figura 30. Modelado perceptual de la voz.....	78
Figura 31. Evaluación de la calidad de la voz a partir del Modelo E .....	83
Figura 32. Escenario de red de medición H.323.....	87
Figura 33. Configuración captura wireshark .....	100
Figura 34. Decodificar paquetes UDP.....	101
Figura 35. Pantalla Decode As.....	102
Figura 36. Paquetes RTP.....	102
Figura 37. Opción Show All Streams .....	103
Figura 38. Flujos RTP .....	104
Figura 39. Opción Analyze en flujos RTP .....	104
Figura 40. Análisis de flujo RTP y guardado de información en archivo CSV .....	105
Figura 41. Conjunto de TXT obtenidos al procesar archivos CSV.....	106

Figura 42. Valores de H con relación a las horas de medición en 5 Mhz .....	108
Figura 43. Valores de Jitter promedio obtenidos en las mediciones de 5 Mhz .....	109
Figura 44. Valores de pérdidas de paquetes en 5 Mhz .....	110
Figura 45. Valores obtenidos de H en frecuencia 10 Mhz .....	111
Figura 46. Valores obtenidos del Jitter promedio de frecuencia de 10 Mhz .....	111
Figura 47. Pérdida de paquetes en la frecuencia de 10 Mhz .....	112
Figura 48. Valores H con relación a las horas de medición de 20 Mhz .....	113
Figura 49. Valores de Jitter promedio obtenidos en las mediciones de 20 Mhz .....	113
Figura 50. Pérdidas de paquetes en las mediciones de 20 Mhz .....	114

## Índice de Tablas

Tabla 1. Comparación de conmutación por circuitos y conmutación por paquetes .....	13
Tabla 2. Audio digital sin compresión .....	34
Tabla 3. Codecs más comunes de la telefonía .....	36
Tabla 4. Comparación entre H.323 y SIP .....	39
Tabla 5. Descripción de la relación entre el retardo y calidad de la voz.....	67
Tabla 6. Características de algunos codecs .....	68
Tabla 7. Escala MOS utilizada para medir la calidad de la voz .....	76
Tabla 8. Escala MOS para la medida del esfuerzo de interpretación del mensaje .....	77
Tabla 9. Escala MOS de calidad para los codecs más comunes .....	77
Tabla 10. Configuraciones básicas de los VIP-480FS de la zona WANETCTIC .....	96
Tabla 11. Configuraciones básicas de los VIP-480FS de la zona WANETDCS .....	96
Tabla 12. Configuraciones parte 1 de las llamadas .....	98
Tabla 13. Configuraciones parte 2 de las llamadas.....	99

## Abreviaturas

ACF	ADMISSION CONFIRM
ACR	ABSOLUTE CATEGORY RATING
ARJ	ADMISSION REJECT
ARQ	ADMISSION REQUEST
ATM	ASYNCHRONOUS TRANSFER MODE
BW	BANDWIDTH
DCF	DISENGAGE CONFIRM
DRQ	DISENGAGE REQUEST
DTMF	DUAL TONE MULTI-FREQUENCY
FTP	FILE TRANSFER PROTOCOL
GCF	GATEKEEPER CONFIRM
GK	GATEKEEPER
GNU	GNU'S NOT UNIX
GRJ	GATEKEEPER REJECT
GRQ	GATEKEEPER REQUEST
GSM	GLOBAL SYSTEM FOR MOBILE
HTTP	HYPertext TRANSFER PROTOCOL
IETF	INTERNET ENGINEERING TASK FORCE
IP	INTERNET PROTOCOL
ISDN	INTEGRATED SERVICES DIGITAL NETWORK
ITU-T	INTERNATIONAL TELECOMMUNICATION UNION
MCU	MULTIPOINT CONTROL UNITS
MHZ	MEGAHERTZ
MS	MILISEGUNDOS

PCM	PHASE-CHANGE MEMORY
PESQ	PERCEPTUAL EVALUATION OF SPEECH QUALITY
PoE	POWER OVER ETHERNET
PSQM	PERCEPTUAL SPEECH QUALITY MEASURE
PSTN	PUBLIC SWITCHED TELEPHONE NETWORK
QoS	QUALITY OF SERVICE
RAS	REGISTRATION, ADMISSION AND STATUS
RDSI	RED DIGITAL DE SERVICIOS INTEGRADOS
RTCP	REAL TIME CONTROL PROTOCOL
RTP	REAL TIME TRANSPORT PROTOCOL
SID	SILENCE INSERTION DESCRIPTOR
SSID	SERVICE SET IDENTIFIER
TCP	TRANSMISSION CONTROL PROTOCOL
UA	USER AGENT
AUC	USER AGENT CLIENT
UAS	USER AGENT SERVER
UBR	UNSPECIFIED BIT RATE
UDP	USER DATAGRAM PROTOCOL
VAD	VOICE ACTIVITY DETECTION
VIP	ID TERMINAL
VoIP	VOICE OVER INTERNET PROTOCOL
WWW	WORLD WIDE WEB

## Resumen

La tecnología de voz sobre el protocolo de Internet conocida como VoIP, está en constantes cambios y evolución, sin embargo hasta ahora el protocolo de Internet para el uso de la tecnología de VoIP no nos otorga garantías de calidad en términos de ancho de banda, jitter, retardos, Y pérdida de paquetes.

Estos factores son importantes ya que influyen y determinan de manera directa la calidad de servicio en la tecnología de VoIP.

Existen métodos para contra restar los impactos de dichos factores. Es por eso que en esta tesis se describe y desarrolla el proceso de caracterización de los parámetros mediante mediciones en un ambiente inalámbrico, donde se transmite tráfico de VoIP.

En esta tesis se realizó la medición y análisis así como la caracterización de los principales parámetros que influyen en la calidad de servicio (QoS) del tráfico de voz sobre IP en un escenario de enlace punto a punto inalámbrico.

Se generó tráfico de llamadas de pruebas entre dos zonas de H.323 mediante dispositivos en hardware (VIP-480FS) y sus respectivos Gatekeeper implementados con el software GNU Gatekeeper.

El tráfico generado fue capturado el analizador de protocolos de red Wireshark. De las capturas se obtuvieron datos estadísticos los cuales se analizaron y se obtuvo una caracterización de las principales métricas de desempeño de un conjunto de llamadas de prueba.

# CAPÍTULO 1

## Capítulo 1. INTRODUCCIÓN

El incremento en el número de usuarios con acceso a Internet, el despegue de la banda ancha y la evolución de nuevas tecnología de transmisión (alámbricas e inalámbricas), han incrementado la demanda de nuevos servicios convergentes, y en consecuencia la aparición de un considerable porcentaje de tráfico de voz sobre el protocolo de Internet (VoIP) dentro del tráfico de voz internacional.

Por otro lado, el surgimiento de múltiples protocolos de señalización de llamada, para la transmisión de voz sobre las redes de paquetes, ha apoyado considerablemente esta convergencia. H.323 y SIP son considerados como los protocolos de señalización más importantes que se encuentran presentes en las principales implementaciones de esta tecnología. Cabe mencionar que estos protocolos no proveen garantía de calidad de servicio (QoS).

QoS es un factor importante a considerar en el diseño e implementación de una red de servicios convergentes. En la red convergente (Internet) los recursos son compartidos y proporciona un servicio de “mejor esfuerzo”; por consiguiente, la calidad de servicio no está garantizada.

Por tal motivo muchos investigadores se han centrado en el estudio de las principales métricas de desempeño que determinan la calidad de servicio en aplicaciones multimedia, tales como VoIP.

### 1.1 Planteamiento del problema

El Internet de nuestros días proporciona un servicio llamado de “mejor esfuerzo” (“best-effort”) y no garantiza calidad de servicio a aplicaciones multimedia en tiempo real, tal como voz sobre el protocolo de Internet. En los últimos años VoIP ha llegado a ser una de las aplicaciones de mayor crecimiento vistas sobre Internet y de mayor demanda de recursos y servicios

de red para proporcionar un nivel de QoS aceptable por los usuarios. La QoS de aplicaciones VoIP depende de varias métricas de desempeño, tales como, pérdida de paquetes, jitter, ancho de banda, tipo de CODEC, tamaño de paquete, tamaño de de-jitter buffer, entre otros. Sin embargo, las métricas que tienen mayor impacto son: jitter y pérdida de paquetes.

Debido a que en el futuro se espera la migración de las comunicaciones de voz a Internet, las expectativas por mejorar el desempeño de los sistemas VoIP han aumentado considerablemente, por lo cual es esencial conocer el comportamiento de los factores que afectan el desempeño y funcionamiento del mismo, con el objetivo de definir mejores modalidades de prestación de servicios, traducidas en aceptable nivel de QoS.

En esta tesis se realizará, la medición y evaluación de la calidad de servicio de un sistema de voz sobre IP en un escenario de dos redes inalámbricas Ad Hoc "WANETs" (WANETCTIC y WANETDCS). Las pruebas a efectuar consistirán en el establecimiento de llamadas entre un conjunto de terminales (terminales Vip1, Vip2..., Vipn, localizado en Red WANETCTIC y terminales Vip5, Vip6..., Vipn, localizado en Red WANETDCS). Para establecer el conjunto de llamadas de prueba se utilizará una aplicación VoIP en hardware (Internet Telephony Gateway-FXS). Dentro del conjunto de llamadas a establecer se utilizarán diferentes configuraciones de algunos parámetros, tales como tipo de CODEC, y habilitación/des-habilitación de detección de actividad de voz (VAD), con el objetivo de evaluar su desempeño bajo diferentes configuraciones.

Paralelamente al establecimiento de llamadas se efectuará la captura de patrones de tráfico (trazas de pérdida de paquetes y trazas de jitter) mediante

un analizador de protocolos y finalmente se realizará el procesamiento de las trazas capturadas para evaluar la calidad de servicio de las llamadas de prueba mediante la caracterización de parámetros tales como: jitter, pérdida de paquetes y parámetro de Hurst.

## 1.2 Justificación

Son muchos los beneficios que proporciona el uso de VoIP sobre la red de datos, los más importantes se mencionan a continuación:

- Convergencia de las redes de voz y datos en una sola red
- Servicios más atractivos de comunicación
- Reducción de los costos de comunicaciones
- Uso más eficiente de la infraestructura de datos
- Reducción de costos por concepto de gestión y operación de infraestructura

Por otro lado, el uso de las redes inalámbricas Ad Hoc como sistemas de comunicaciones emergentes, juegan un papel muy importante en situaciones de desastres naturales.

Partiendo de los beneficios antes mencionados, surge el interés de implementar un sistema de voz sobre una red inalámbrica Ad Hoc, medir las principales métricas de desempeño y evaluar la calidad de servicio del sistema de voz mediante la caracterización de parámetros de QoS, tales como: jitter, pérdida de paquetes y parámetro Hurst.

### 1.3 Objetivo general

Evaluar la calidad de servicio de un sistema de voz sobre una red inalámbrica Ad Hoc mediante la caracterización de parámetros tales como, jitter, pérdida de paquetes y parámetro Hurst.

### 1.4 Objetivos particulares

Realizar la configuración de 2 GNU Gatekeepers (GK1 y GK2) en modo “Neighbors”.

- Configurar y registrar los terminales Vip1, Vip2...., Vipn en GK2 y Vip5, Vip6...., Vipn en GK1
- Realizar un conjunto de llamadas bajo diferentes configuraciones
- Capturar el tráfico de VoIP generado, mediante un analizador de protocolos (Wireshark)
- Realizar el procesamiento de los patrones de tráfico mediante scripts en Matlab

## 1.5 Metodología

1. Configurar 2 GNU Gatekeepers (GK1 y GK2) en modo “Neighbors”. Donde, GK estará ubicado en la Red WANETDCS y GK2 en la Red WANETCTIC.
2. Configurar y registrar los terminales Vip1, Vip2,....., Vipn en el gatekeeper GK2 y los terminales Vip5, Vip6,....., Vipn en el gatekeeper GK1. Para realizar de manera más práctica y centralizada las mediciones, se utilizará una configuración entre los terminales y los GKs, en la cual, tanto la señalización como el tráfico RTP sea transportado vía los GKs.
3. Establecer un conjunto de llamadas de prueba bajo diferentes configuraciones de tipo de CODEC (G.711 y G.729) y la habilitación/deshabilitación de la detección de actividad de voz (VAD).
4. Configurar el Wireshark y aplicar los filtros necesarios, de tal manera que capture el tráfico de voz entrante y saliente de las tarjetas de red de cada uno de los “Gatekeeper”.
5. Procesar los patrones de tráfico capturados mediante scripts en Matlab, con el objetivo de obtener series de tiempo, las cuales serán útiles para obtener estadísticas básicas.
6. Caracterizar los patrones de tráfico para evaluar la calidad de servicio de las llamadas de prueba.
7. Escritura del documento de tesis.

# CAPÍTULO 2

## Capítulo 2. REDES DE TELECOMUNICACIONES

Las redes de comunicación de datos, constituyen en la actualidad un apoyo de vital importancia para todas las empresas cuyo éxito depende del buen manejo de la gran cantidad de información que generan.

Diferentes aspectos tecnológicos de las redes sirven para hacer una clasificación de ellas. Por ejemplo, si emplean o no conmutación y qué tipo de ésta utilizan. Otro parámetro es el tipo de procesamiento que realizan, su cobertura, topología, velocidad.

En una red conmutada, el enlace entre dos terminales de usuario se puede establecer a través de centros de conmutación sobre demanda, es decir, haciendo una solicitud de enlace, como en la red telefónica clásica. En redes no conmutadas estos enlaces se suministran de manera exclusiva, es decir, se alquilan a la empresa pública o privada para utilizarse de manera permanente para el envío de datos de la organización.

### 2.1 Redes de conmutación por circuito

Las redes conmutadas pueden utilizar dos técnicas distintas para el establecimiento de los enlaces: la conmutación de circuitos o la conmutación de paquetes.

En las redes con conmutación por circuitos (RCC) interesa establecer, sobre demanda, una conexión entre dos cualesquiera equipos terminales. Cada conexión se establecerá durante el tiempo de la sesión (diálogo entre terminales para el intercambio de información) mediante una trayectoria física bien definida de transmisión. Al final de la llamada, la conexión se libera para poder atender nuevas solicitudes de conexión. Éste es el principio de la red telefónica pública conmutada.

El beneficio de las redes conmutadas por circuito es el gran rango de destinos disponibles para las diferentes llamadas, lo que representa gran utilidad para las redes de datos. Un ejemplo de este tipo de red de datos es la red telex; una vez que se establece una conexión telex, se transmiten

caracteres entre dos terminales telex a la velocidad de línea de 50 bauds. La red digital de servicios integrados (ISDN) es otro ejemplo de una red digital que ofrece posibilidades de conmutación por circuitos [3].

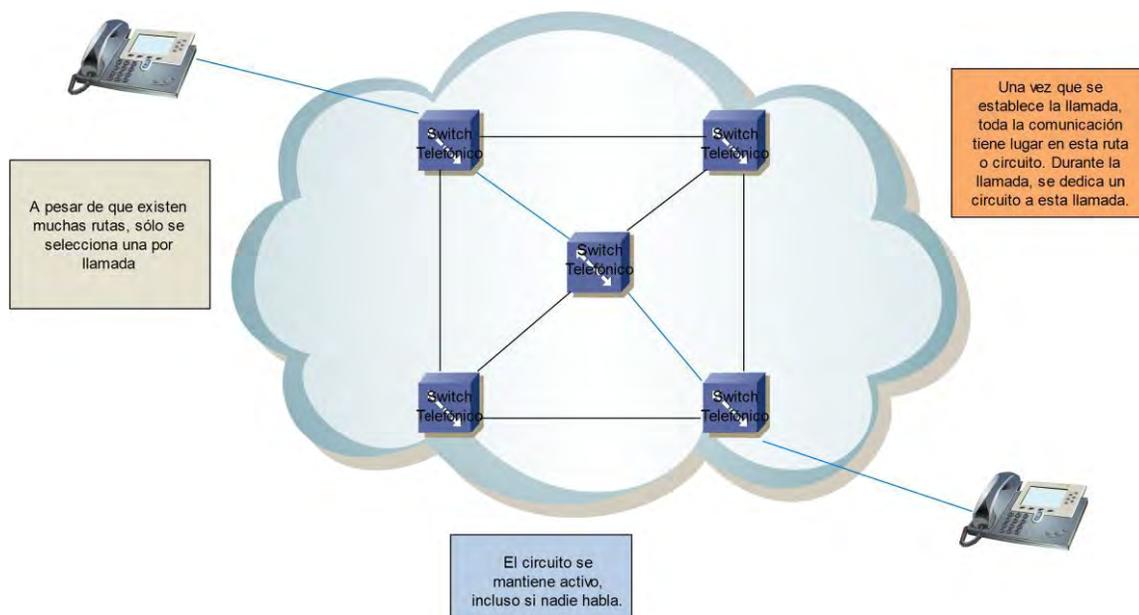


Figura 1. Conmutación por circuitos

### 2.1.1 Red telefónica pública conmutada (PSTN)

Los sistemas que permiten las comunicaciones telefónicas fueron inventados hace ya más de un siglo y constituyen uno de los pilares en que se sustenta la vida y la economía mundial.

La Red Telefónica Pública Conmutada (PSTN, Public Switched Telephone Network) está funcionando en todos los países del mundo, con alrededor de 800 millones de abonados.

Esta red, que permite que los seres humanos se puedan comunicar a grandes distancias a través de la voz, ha significado una verdadera revolución en el concepto de comunicaciones que existía en el momento de su invención. Desde su inicio hasta nuestros días, la red se mantiene con las lógicas variantes que ha deparado el avance tecnológico y el constante aumento de las personas con posibilidades de conectarse a través de ella.

La red telefónica está compuesta básicamente por los siguientes elementos:

**Nodos de comunicación:** Existen dos tipos de nodos, los que se denominan centrales locales, a los que se conectan los usuarios de la red y los que se denominan centrales de tránsito que son los encargados de encaminar las comunicaciones para llegar a los destinos solicitados.

**Enlaces:** Los enlaces se usan en la PSTN para interconectar las centrales locales entre sí, y con las centrales de tránsito, se suelen denominar enlaces troncales por la cantidad de voz que suelen transportar.

**Enlace de Abonado:** Es el vínculo de comunicación que conecta al equipo, o los equipos terminales del usuario, con la central local.

**Equipos Terminales:** Son los equipos que cada usuario podrá conectar a la red de acuerdo con las prestaciones que la misma pueda brindar, el primero y más importante que debe considerarse en esta red e precisamente el teléfono.

La red telefónica pública conmutada presenta las siguientes características particulares:

- La red está siempre casi en todas partes. Resulta muy difícil que no se encuentre disponible un equipo terminal de la red telefónica pública conmutada en algún punto geográfico del planeta.
- El sistema telefónico es un sistema abierto. Esta condición permite que el uso de la red no necesite autorizaciones especiales, ni requerimientos adicionales ante las administraciones.
- La obtención del hardware necesario no presenta complicaciones técnicas importantes. Los equipos necesarios son de fácil obtención en el mercado tele informático [4].

### 2.1.2 Red digital de servicios integrados (ISDN)

El concepto de redes digitales de servicios integrados se caracteriza esencialmente por el hecho de que permite una amplia gama de aplicaciones vocales y no vocales en la misma red. Un elemento clave para la integración de servicios en una RDSI, es la prestación de una gama de servicios mediante el empleo de un conjunto limitado de tipos de conexión y configuraciones de interfaz polivalente usuario-red.

Las RDSI Soportan diversas aplicaciones, entre la cuales están las conexiones conmutadas y no conmutadas. Las conexiones conmutadas en un RDSI comprenden conexiones con conmutación de circuitos, conexión con conmutación de paquetes y sus concatenaciones. En medida en que sea posible en la práctica, los nuevos servicios que se introduzcan en una RDSI deberán disponerse de modo que sean compatibles con las conexiones digitales conmutadas a 64 kbit/s.

Una RDSI contendrá inteligencia para asegurar las características de servicio, y las funciones de mantenimiento y gestión de la red. Es posible que esta inteligencia no sea suficiente para algunos nuevos servicios y sea necesario suplementarla mediante inteligencia adicional dentro de la propia red o, lo que también es posible, mediante inteligencia compatible en los términos de usuario [5].

La ventaja que la ISDN ofrece sobre otros servicios, es la capacidad de soportar sobre una sola conexión ISDN la transmisión de diferentes medios, tales como, voz, fax, y video. Esto es, la ISDN integra todos esos servicios en uno solo. El servicio ISDN es completamente digital de extremo a extremo. Además, la tecnología ISDN permite que el alambrado estándar de pares trenzados lleve datos digitales por medio de circuitos o paquetes conmutados. También proporciona una estrategia de costo efectivo para el trabajo entre redes. En vez de pagar por líneas dedicadas alquiladas, los sitios remotos pueden interconectarse con otros sitios vía enlaces telefónicos [6].

## 2.2 Red de conmutación por paquetes.

Una limitante muy importante de las Red de conmutación por circuitos cuando se emplea para datos es su incapacidad para proporcionar conexiones con ancho de banda variable. Esto provoca el uso ineficiente de recursos cuando se requiere sólo un ancho de banda estrecho (baja velocidad de bits) o la aparición de retardos en la transmisión de datos cuando se necesita ráfagas cortas de ancho de banda grande. Un método más eficiente de envío de datos, la conmutación de paquetes, apareció a principios de la década de 1970 para constituir el principio de las redes de conmutación de paquetes.

Se llama conmutación de paquetes porque el mensaje del usuario se descompone en cierto número de fragmentos, llamados paquetes, cada uno de los cuales se envía por separado. En cada paquete de datos se agrega la dirección de su destino y varios campos de control adicionales antes de ser enviado. El extremo receptor se encarga de re-ensamblar los paquetes en el orden apropiado con la ayuda de los números de secuencia.

La conmutación por paquetes provee buena confiabilidad de extremo a extremo; con conmutación y redes bien diseñadas es posible evadir las fallas de red aun durante el progreso de una llamada. La conmutación por paquetes es eficiente también en el empleo de enlaces y recursos de la red, compartiéndolos entre varias llamadas incrementado, por consecuencia, su utilización. Como en conmutación de paquetes la longitud de estos es corta, existe mucho menor bloqueo o congestión de datos, por lo que las demoras son comúnmente pequeñas, el trabajo interactivo entre usuarios es posible [3].

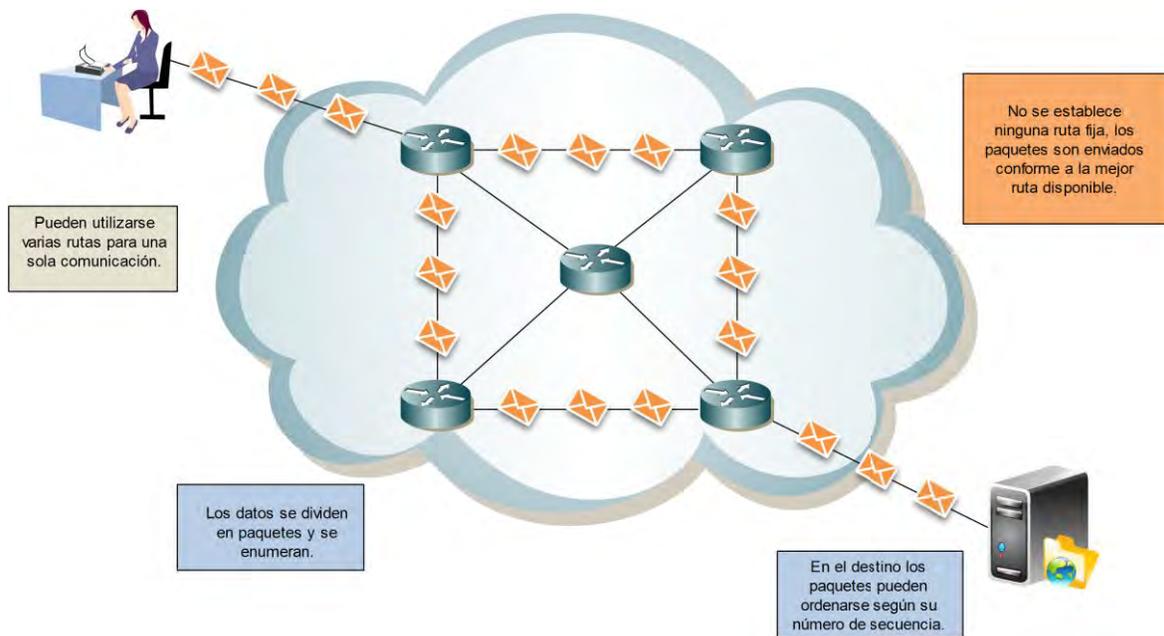


Figura 2. Conmutación por paquetes

En las redes de conmutación de paquetes se hace distinción de dos modos de operación: modo orientado a la conexión y modo orientado a la no conexión. En el modo orientado a la conexión, una ruta es establecida antes de que los paquetes puedan ser enviados; esta ruta se denomina circuito virtual. Se realiza un intercambio de paquetes de señalización inicial para reservar recursos y establecer la ruta. El usuario establece la conexión, la usa y por último se desconecta.

En el modo orientado a la no conexión, cada paquete es tratado de manera independiente. Los paquetes se envían con sus respectivas direcciones de origen y destino y las decisiones de hacia dónde deben ser enrutados se toman en cada nodo de la red. El primer paquete en ser enviado no es siempre el primero en llegar. Internet es la principal red que utiliza este modo de conexión y ha evolucionado rápidamente, al punto de poder soportar aplicaciones multimedia (voz, video, datos), sin embargo, no garantiza la calidad del servicio, debido a que ofrece un servicio llamado servicio de entrega de mejor esfuerzo, es decir, los paquetes se envían y no hay ningún mecanismo que garantice la correcta recepción de los mismos [6].

Existen características importantes entre la conmutación por circuitos y la conmutación por paquetes, algunas de ellas son las que muestra la siguiente tabla:

Conmutación por circuitos	Conmutación por paquetes
El canal de comunicación es dedicado	Los canales de comunicación no son dedicados
La ruta se establece para toda la comunicación	Se establece una ruta por paquete
No tiene encabezado de bits	Cuenta con encabezados de bits en cada paquete
Envía una señal de ocupado si el medio está ocupado	Se notifica al emisor si el paquete no es enviado
La transmisión de datos es continua	La transmisión de paquetes es por ráfagas
El ancho de banda de transmisión es fijo	El ancho de banda de transmisión es dinámico
Los mensajes no son almacenados	Los paquetes se almacenan hasta que son enviados

*Tabla 1. Comparación de conmutación por circuitos y conmutación por paquetes*

### 2.2.1 Internet

A principios de la década de los 90, solo los profesores, estudiantes universitarios y gente muy interesada en nuevas tecnologías conocían la palabra “Internet” y, en apenas siete años, este medio ha pasado a ser uno de los temas centrales de la cultura y de la economía.

En diciembre de 1969 había, en total, cuatro nodos formando una red repartida por Estados Unidos y con un patrocinador público fuera del pentágono: fue el nacimiento de ARPANET. Gracias a ARPANET, los científicos e investigadores podían utilizar a distancia las instalaciones informáticas de otros centros académicos.

Internet es un concepto inseparable de los términos TCP/IP e IP. Se trata del protocolo de transmisión de Internet que regula el intercambio de datos

entre hosts. Para un funcionamiento sin problemas de una estructura tan compleja como la de Internet es necesaria la identificación de unas directivas muy claras y una clara denominación de los dispositivos (hosts, impresoras, cámaras de Internet) que se encuentran en la estructura de Internet. Cada recurso de red posee un número unívoco: la dirección IP.

Este número tiene la forma aaa.bbb.ccc.ddd y siempre está compuesto de cuatro números. Cada uno de ellos tiene un máximo de tres cifras y unos valores entre 0 y 255, como, por ejemplo, 194.167.25.23. Como estos números son difíciles de recordar, los servidores FTP y WWW disponen de un nombre más expresivo. Puede acceder a la página inicial de Microsoft, por ejemplo, a través de la dirección <http://www.microsoft.com>, mucho más fácil de retener que uno de esos números antes mencionados [7].

### 2.2.2 Protocolos de nivel de transporte

La red Internet utiliza el protocolo IP para que las estaciones envíen paquetes. Este protocolo es orientado a la no conexión, o sea que no tiene como objetivo hacer llegar correctamente los paquetes, ni que lleguen en orden. De hecho no garantiza ni siquiera que lleguen. De todo esto se encargan los protocolos de transporte.

En Internet se definieron básicamente dos protocolos de transporte: UDP (User Datagram Protocol) y TCP (Transmission Control Protocol). UDP sólo garantiza la entrada libre de errores. No preserva la secuencia ni garantiza la entrega, mientras TCP sí preserva el orden en que han sido transmitidos los paquetes y garantiza su entrega. En la siguiente imagen podemos observar la situación de los protocolos de transporte en la pila TCP/IP.

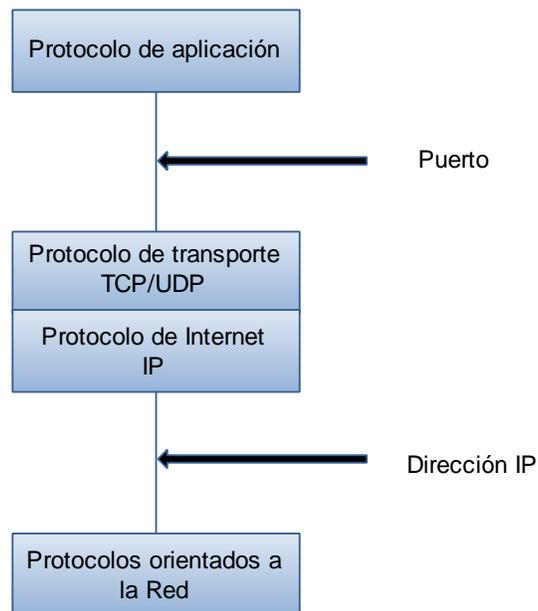


Figura 3. Protocolo TCP/IP

En TCP/IP la comunicación se establece entre aplicaciones, no sólo entre máquinas. Por lo tanto, para comunicarnos con una aplicación remota necesitamos saber no sólo en IP de la máquina, sino también el número de aplicación. Este número se conoce como puerto. Pero el puerto se define a nivel transporte, no a nivel de aplicación. Por eso, aparecerá en la cabecera de TCP (o UDP), de la misma manera que en la cabecera de IP aparece un identificador relacionado con el protocolo de transporte que hay en aquel paquete.

#### 2.2.2.1 UDP

UDP es un protocolo orientado a la no conexión, de manera que no proporciona ningún tipo de control de errores ni de flujo, aunque utiliza mecanismos de detección de errores. En caso de detectar un error, UDP no entrega el datagrama a la aplicación, sino que lo descarta.

La simplicidad del UDP hace que sea ideal para aplicaciones que requieren poco retardos (por ejemplo, aplicaciones en tiempo real como pueden ser aplicaciones de voz y vídeo). UDP también es ideal para aquellos dispositivos que no pueden implementar un sistema tan complejo como el TCP. Otro uso interesante de UDP es en aplicaciones que trabajan en modo multicast o broadcast. En este caso, se envía información a muchos receptores sin esperar una respuesta de todos, de manera que es ideal disponer de un protocolo de transporte simple y sencillo orientado a la no conexión como UDP. Algunas de las características más importantes de UDP son las siguientes:

- No garantizan la fiabilidad; es decir, no se tiene la seguridad de que cada datagrama UDP transmitido llegue a su destino; es un protocolo best-effort: el UDP hace todo lo posible para transferir los datagramas de su aplicación, pero no garantiza su entrega.
- No preserva la secuencia de la información que le proporciona la aplicación. Como está en modo datagrama y utiliza un protocolo por debajo como IP, que también está en modo datagrama, la aplicación puede recibir la información desordenada. La aplicación debe estar preparada para que haya datagramas que se pierdan, lleguen o se hayan desordenado.

La figura siguiente muestra la unidad de datos del protocolo UDP y su encapsulamiento en un datagrama IP. Cada operación de salida de un datagrama UDP provoca la generación de un datagrama IP.

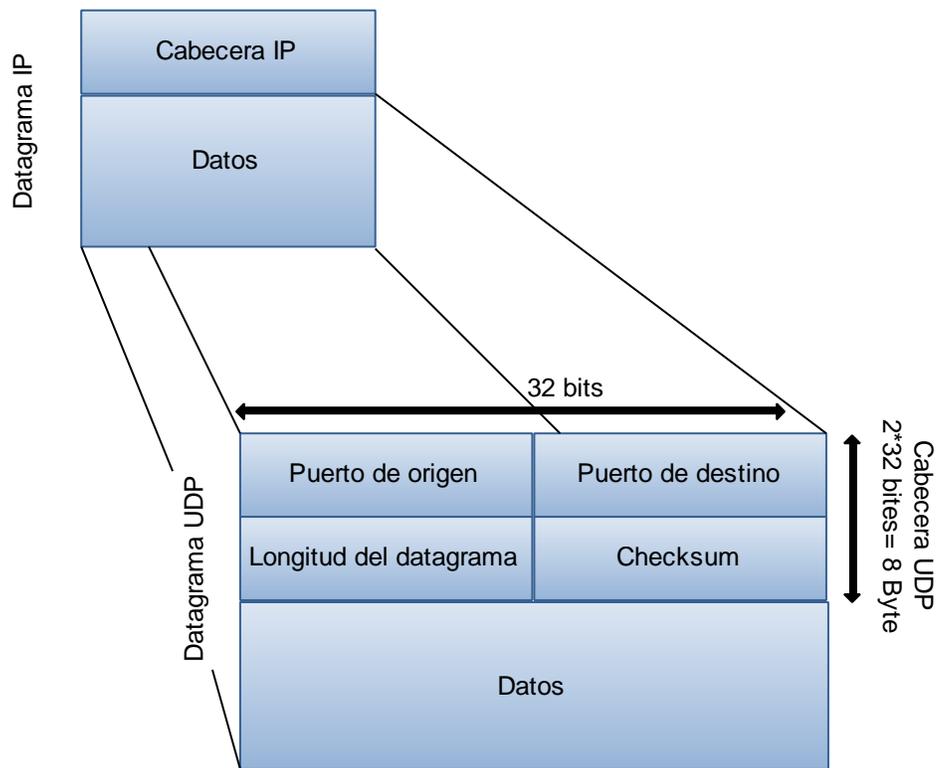


Figura 4. Protocolo UDP

El datagrama UDP consta de una cabecera y un cuerpo para encapsular los datos. La cabecera consta de los siguientes elementos:

**Los campos de puerto origen y puerto destino**, que identifican las aplicaciones en los terminales de origen y de destino.

**El campo Longitud** indica la longitud, en bytes, del datagrama UDP incluyendo la cabecera UDP (es la diferencia de la longitud del datagrama IP menos la cabecera de IP).

**El campo Checksum** es opcional y protege tanto la cabecera como los datos UDP. Cuando el UDP recibe un datagrama y determina que hay errores, lo descarta y no lo entrega a ninguna aplicación.

### 2.2.2.2 TCP

UDP no garantiza la entrega de la información, tampoco reordena la información en caso de que llegue en orden diferente de aquél en que se ha transmitido. Existen aplicaciones que no pueden tolerar dichas limitaciones. Para superarlas, el nivel de transporte proporciona un protocolo fiable extremo a extremo llamado TCP. TCP proporciona fiabilidad a la aplicación; es decir garantiza la entrega de toda la información en el mismo orden en que se ha sido transmitida por la aplicación de origen. Para conseguir esta fiabilidad, TCP proporciona un servicio orientado a la conexión con un control de flujo y errores. Para proporcionar un servicio fiable a la aplicación, TCP se basa en los principios siguientes:

1. Transmisión libre de errores. TCP debe entregar a la aplicación de destino exactamente la misma información que le entregó la aplicación de origen.
2. Garantía de entrega de la información. TCP garantiza que toda la información transmitida por la aplicación de origen se entregue a la aplicación destino.
3. Garantía de mantenimiento de la secuencia de transmisión. TCP garantiza la entrega de flujo de información en el mismo orden en que le fue entregado por la aplicación de origen.
4. Eliminación de duplicados. TCP garantiza que sólo entregará una copia de la información transmitida a la aplicación de destino. En caso de que reciba copias a causa del funcionamiento de la red o de los protocolos que se implementan por debajo del nivel de transporte, TCP las eliminará.

La unidad de información del protocolo TCP se llama segmento TCP y su formato el siguiente:

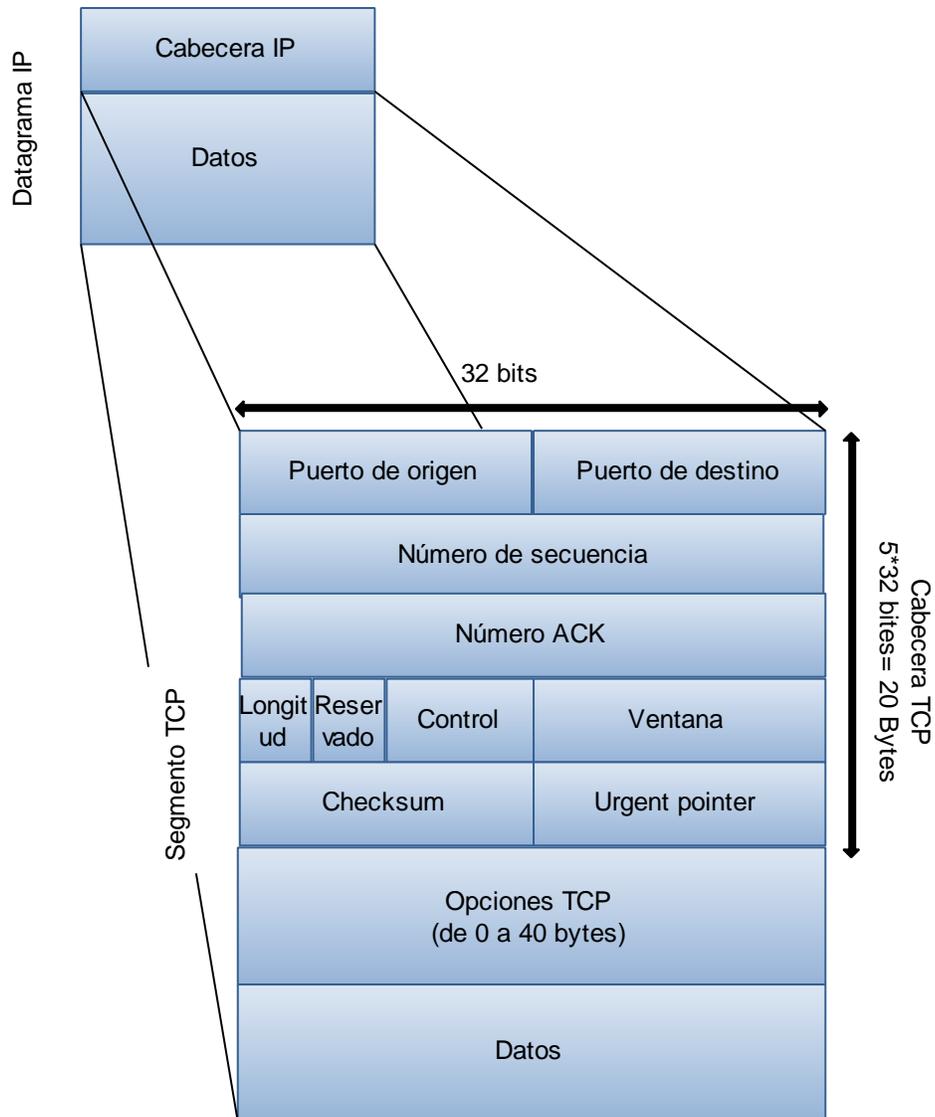


Figura 5. Protocolo TCP

El segmento TCP consta de una cabecera y un cuerpo para encapsular datos. La cabecera consta de los campos siguientes:

El campo **puerto de origen** identifica la aplicación en el terminal de origen.

El campo de **puerto destino** identifica la aplicación en el terminal de destino.

El campo **número de secuencia** identifica el primer byte del campo de datos. En TCP no se enumeran segmentos, sino bytes. Por tanto, el número de secuencia identifica el primer byte de los datos que envía el segmento.

El campo **número ACK**. El TCP reconoce datos por medio de la técnica de *piggybacking*. Al activar un bit de la cabecera (el bit ACK), TCP tiene en cuenta el número de secuencia ACK que indica al otro extremo TCP el próximo byte que está dispuesto a recibir.

El campo **longitud de cabecera** indica la longitud de la cabecera, que puede ser variable. La longitud típica es de 20 bytes; sin embargo, si TCP utiliza el campo de opciones, puede llegar a una longitud máxima de 60 bytes.

El campo **reservado**, tal como su nombre lo indica, ésta reservado y se inicializa con ceros.

El campo **control** está formado por ocho indicadores independientes, cada uno de los cuales señala una función específica cuando está activo:

**FIN:** indica que el transmisor ha acabado la conexión.

**SYN:** se utiliza para iniciar una conexión y también sirve para resincronizar los números de secuencia.

**RST:** realiza un *reset* de la conexión.

**PSH:** invoca la función *push* en el protocolo. Esta función dice al receptor que entregue a la aplicación todos los datos que tenga disponibles en la memoria intermedia de recepción sin esperar a completarlos con datos adicionales.

**ACK:** cuando este bit está activo, el campo número ACK indica al byte siguiente que espera recibir la conexión TCP. Si este bit no está activo, el campo número ACK no tiene ningún significado para el TCP.

**URG:** indica que hay datos urgentes (y el campo *Urgent pointer* indica la cantidad de datos urgentes existentes en el segmento).

**ECE, CWR:** estos dos bits se utilizan si se implementa un control de congestión.

El campo **ventana** indica cuántos bytes componen la ventana de transmisión del protocolo de control de flujo por ventana deslizante. A diferencia de los protocolos del nivel de enlace, en que la ventana era constante y contaba tramas, en el TCP la ventana es variable y cuenta bytes.

El campo **Checksum** se utiliza para detectar errores.

El campo **Urgent pointer** tiene sentido cuando el bit de control URG está activo, indica que los datos que envía el origen son urgentes e identifica el último byte del campo de datos que también lo es.

El campo **Opciones TCP** permite añadir campos a la cabecera para realizar operaciones siguientes:

Marcar el tiempo (*timestamp*) en que se transmitió el segmento y de este modo poder monitorizar los retrasos que experimentan los segmentos desde el origen hasta el destino.

Aumentar el tamaño de la ventana.

Indicar el tamaño máximo del segmento (mss) que el origen está preparando para recibir [8].

### 2.2.2.3 RTP/RTCP

El objetivo de cualquier comunicación es el intercambio de información entre los extremos de dicha comunicación, es decir, el transporte de información entre el origen y el o los destinos. Este transporte no sólo incluye el trasladar los paquetes de información de un lado a otro sino que, además, habrá que fragmentar y reensamblar los paquetes y proveer los mecanismos necesarios para reducir el impacto de las pérdidas, el jitter, el retardo, entre otras.

Los protocolos empleados en comunicaciones de audio y video en tiempo real vienen definidas en la RFC 1889 y son RTP (*Real Time Protocol*) para el intercambio de la información y RTCP (*Real Time Control Protocol*) para

el control de dicho intercambio. Aunque la norma no lo indica explícitamente, tanto RTP como RTCP suelen emplearse sobre UDP y que el servicio fiable que proporciona TCP daría lugar a un retardo demasiado elevado como consecuencia de las retransmisiones.

En una comunicación de voz sobre paquetes, cada canal de comunicaciones está compuesto por un flujo RTP y un flujo RTCP cuyos puertos UDP se eligen independientemente en cada extremo de la comunicación, si bien deben notificárselo mutuamente utilizando algún mecanismo de señalización. El único requisito es que el puerto UDP asociado al flujo RTP sea par y el puerto asociado al flujo RTCP sea impar inmediatamente superior al flujo RTP.

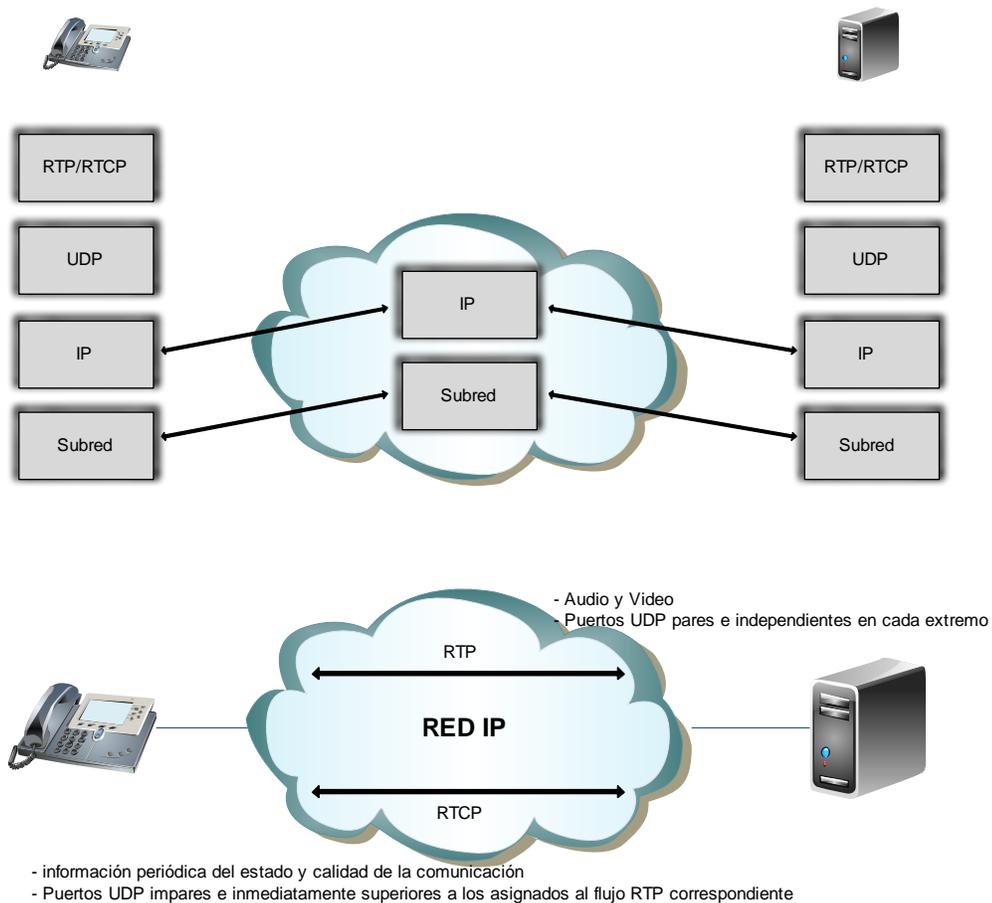


Figura 6. Tráfico RTP y RTCP a través de la red.

RTP (*Real Time Protocol*) es el estándar que define las comunicaciones de audio y vídeo en tiempo real sobre redes IP, asumiendo, por tanto, la existencia de pérdidas y retardos y la posibilidad de variación dinámica de las características de la red en el transcurso de la comunicación. Suministra funciones de transporte extremo a extremo y ofrece servicios tales como identificación del tipo de carga, numeración de secuencia, *timestamping*. No garantiza la entrega de tráfico en tiempo real pero sí suministra los recursos para que éste se entregue de manera sincronizada. Como su nombre indica, está orientado a la transmisión de información en tiempo real, como la voz o el vídeo.

Los paquetes RTP están divididos en dos grandes bloques. Por un lado, encontramos la cabecera que contiene la información necesaria para reconstruir el flujo de bits generado por el códec del emisor y, por otro, la carga útil, es decir, el propio flujo de bits.

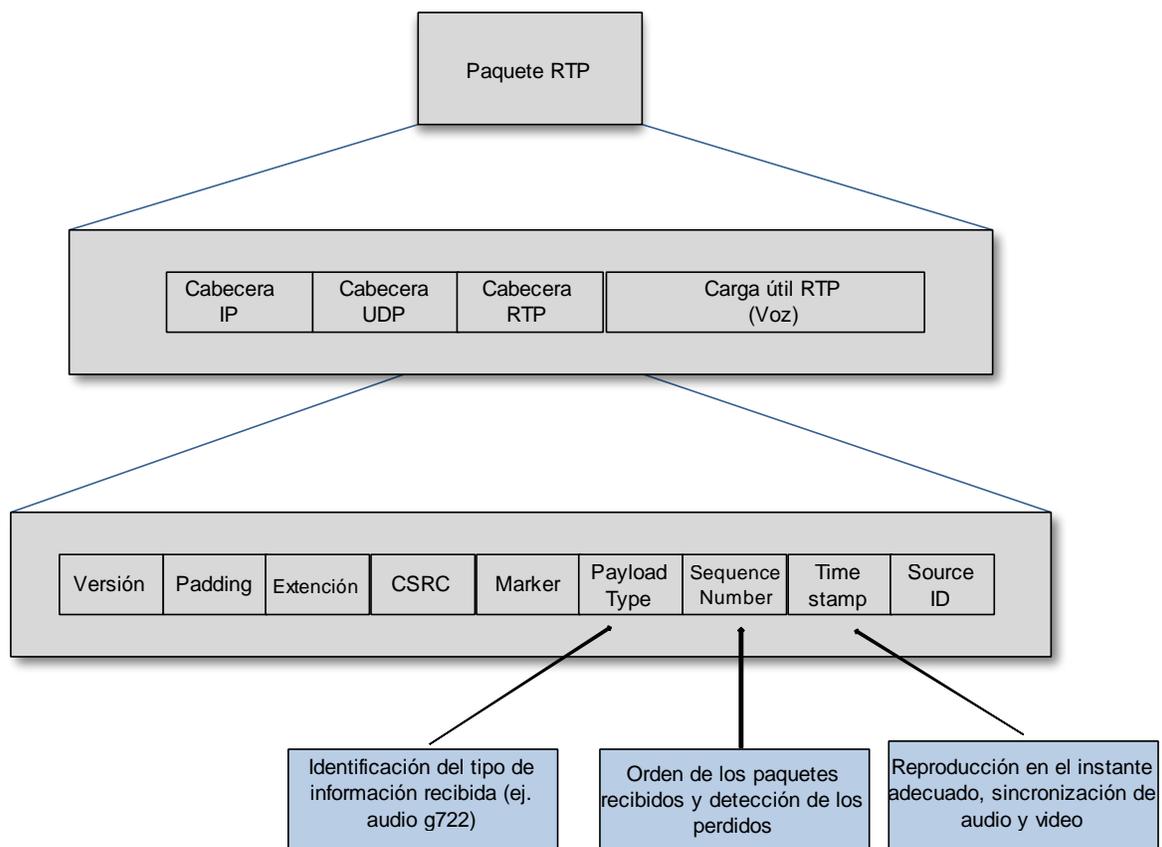


Figura 7. Encapsulamiento de los paquetes de voz.

RTP permite identificar el tipo de información transportada, añadir marcas de tiempo (*time stamp*) y números de secuencia y controlar la llegada de los paquetes. Toda esta información es utilizada por los receptores para reconstruir el flujo de paquetes que generó el receptor eliminado, en la medida de lo posible, los efectos de las pérdidas, el retardo y el *jitter*. En este sentido, hay que destacar que RTP no lleva a cabo ninguna reserva de recursos, sino que ofrece al receptor los mecanismos necesarios para recobrar la información en presencia de éstos.

En algunos casos, para aumentar la eficiencia del transporte se ofrece la posibilidad de comprimir las cabeceras utilizando una versión comprimida de RTP denominada cRTP (*compressed RTP*). El principal problema de cRTP es que introduce un retardo adicional debido al proceso de compresión.

RTP dispone de medios para, en presencia de pérdidas, jitter o retardo, poder continuar con la reproducción del flujo de paquetes. Sin embargo, no especifica ningún medio de estimar los valores de dichos parámetros. De esto, precisamente, se encarga su protocolo asociado, RTCP (*Real Time Control Protocol*), definido en la RFC 1889.

RTCP describe el intercambio de mensajes de control relacionados, fundamentalmente, con la calidad de servicio (retardo, jitter, tasa de pérdidas). Aunque opcional, su utilización es recomendable porque proporciona información de estado de la comunicación con el fin de detectar, por ejemplo, situaciones en las que la calidad de la transmisión no es suficiente aunque no provee de los mecanismos necesarios para mejorar las prestaciones de la red (reserva de ancho de banda, control de congestión).

Además de la información relacionada con la calidad de servicio, RTCP proporciona otras funciones adicionales como la identificación de la fuente, la sincronización entre medios o control de la sesión [9].

# CAPÍTULO 3

## Capítulo 3. TRANSMISIÓN DE VOZ SOBRE REDES DE DATOS

Voz sobre el protocolo de Internet es una tecnología que permite transmitir voz a través de una red de datos bajo la pila de protocolos TCP/IP.

La tecnología VoIP funciona de la siguiente manera: en el transmisor, los flujos de voz se digitalizan y comprimen por medio de un codificador de audio y se hace llegar hasta el receptor en forma de paquetes, a través de una red IP. Una vez realizado el recorrido de los paquetes, en el receptor, se suprimen los encabezados IP y mediante un decodificador de audio se convierten los paquetes al flujo de voz original [10].

### 3.1 Ventajas de VoIP

En general, el servicio de telefonía vía VoIP es gratuito o cuesta muchísimo menos que el servicio tradicional y similar a la alternativa que los proveedores del servicio de la Red Pública Conmutada (PSTN) ofrecen.

Algunos ahorros en el costo son debidos a que se utiliza una misma red para llevar voz y datos, especialmente cuando los usuarios no utilizan toda la capacidad de una red ya existente la cual puede usar VoIP sin costo adicional y tiene asociada una tarifa plana. La tarifa de llamadas entre usuarios VoIP entre cualquier proveedor está incluida en la tarifa plana del pago del servicio de Internet, en contraste con las llamadas entre usuarios VoIP a usuarios de la PSTN, generalmente tienen asociadas una tarifa.

Hay dos tipos de servicio de PSTN a VoIP: llamadas locales directas (*Direct Inward Dialling, DID*) y números de acceso. DID conecta a quien hace la llamada directamente al usuario VoIP mientras que los números de acceso requieren que este introduzca el número de extensión del usuario VoIP. Los números de acceso son usualmente cobrados como una llamada local para quien hizo la llamada desde PSTN y sin costo para el usuario de VoIP.

### 3.1.1 Funcionalidad

La tecnología de voz sobre el protocolo IP puede facilitar tareas que serían más difíciles de realizar usando las redes telefónicas tradicionales. Las llamadas telefónicas locales pueden ser automáticamente enrutadas a tu teléfono VoIP, sin importar en donde estés conectado a la red.

Algunos paquetes de VoIP incluyen los servicios extras por los que la PSTN (Red Telefónica Conmutada) normalmente cobra un cargo extra, o que no se encuentran disponibles en algunos países, como son las llamadas 3 a la vez, retorno de llamada, remarcación automática, o identificación de llamadas.

### 3.1.2 Movilidad

Los usuarios de VoIP pueden viajar a cualquier lugar del mundo y seguir haciendo y recibiendo llamadas de la siguiente forma:

Los subscriptores de los servicios de las líneas telefónicas pueden hacer y recibir llamadas locales fuera de su localidad. Por ejemplo, si un usuario tiene un número telefónico de la ciudad de Nueva York y está viajando por Europa y alguien llama a su número telefónico, esta se recibirá en Europa. Además si una llamada es hecha de Europa a Nueva York, esta será cobrada como llamada local, por supuesto, debe haber una conexión a Internet disponible, como WiFi para hacer esto posible.

Los usuarios de mensajería instantánea basada en servicios de VoIP pueden también viajar a cualquier lugar del mundo y hacer y recibir llamadas telefónicas. Los teléfonos VoIP pueden integrarse con otros servicios disponibles en Internet, incluyendo video llamadas, intercambio de datos y mensajes con otros servicios en paralelo con la conversación, audio conferencias, administración de libros de direcciones e intercambio de información con otros [11].

### 3.1.3 Servicios de valor añadido

El protocolo IP permite crear nuevos servicios o nuevas funcionalidades de una forma muy fácil. Al ser un protocolo pensado para la transmisión de datos, basta añadirle algunos datos a la transmisión de voz para obtener un servicio nuevo. Por ejemplo, durante la conversación de voz se podrían intercambiar archivos (fotografías, gráficos, documentos entre otros). O bien, utilizar una pizarra virtual o compartir aplicaciones (de presentaciones, hojas de cálculo o archivos de texto).

En un entorno corporativo se podrían combinar distintos tipos de servicios. Por ejemplo, al recibir una llamada se podría ver automáticamente la ficha del compañero que nos llama para conocer exactamente con quién vamos a hablar, área en la que trabaja, posición en la jerarquía, su foto, etc. De hecho, nos podría informar hasta de los asuntos pendientes con esa persona.

### 3.1.4 Mantenimiento y gestión integrada

Con la VoIP se puede llegar a disponer de toda la información de facturación y estadísticas en tiempo real y desde cualquier lugar. Si se desea conocer cómo va el consumo por cada área de la empresa, qué empleados hacen un uso más intensivo o qué terminal no se utiliza nunca, basta con conectarse al sistema de gestión.

La integración de los servicios de voz y datos en una misma infraestructura aporta unos beneficios adicionales, ya que se pueden emplear una misma base de datos de usuarios para gestionar todos los servicios de comunicaciones y realizar informes. Además, al utilizar la red de voz y de datos la misma tecnología IP, no será necesario contar con personal de mantenimiento diferenciado. Aunque la VoIP supone de adquirir unos conocimientos particulares, el personal de mantenimiento con experiencia en IP no suele tener dificultades para ello.

Las redes locales de datos (redes corporativas) suelen estar utilizadas al 30% de su potencial. Esto quiere decir que, cuando se introducen los servicios de voz sobre IP, no suele ser necesario ampliar esta red [12].

### 3.2 La señal de la voz

En un estudio de las redes de voz sobre paquetes no podemos pasar por alto el tema de la transmisión de una señal de voz analógica sobre sistema digital. La señal de la voz es una señal analógica, es decir, continua tanto en el tiempo como en la amplitud. Sin embargo, este tipo de señales no pueden transmitirse por sistemas digitales por lo que habrá que cambiarle el formato y convertirla en una señal digital. Este cambio de formato implica la discretización temporal y la posterior cuantificación de las muestras obtenidas (discretización en amplitud). Posteriormente, las muestras de voz cuantificadas habrán de ser codificadas adecuadamente para su transmisión por el canal de comunicaciones.

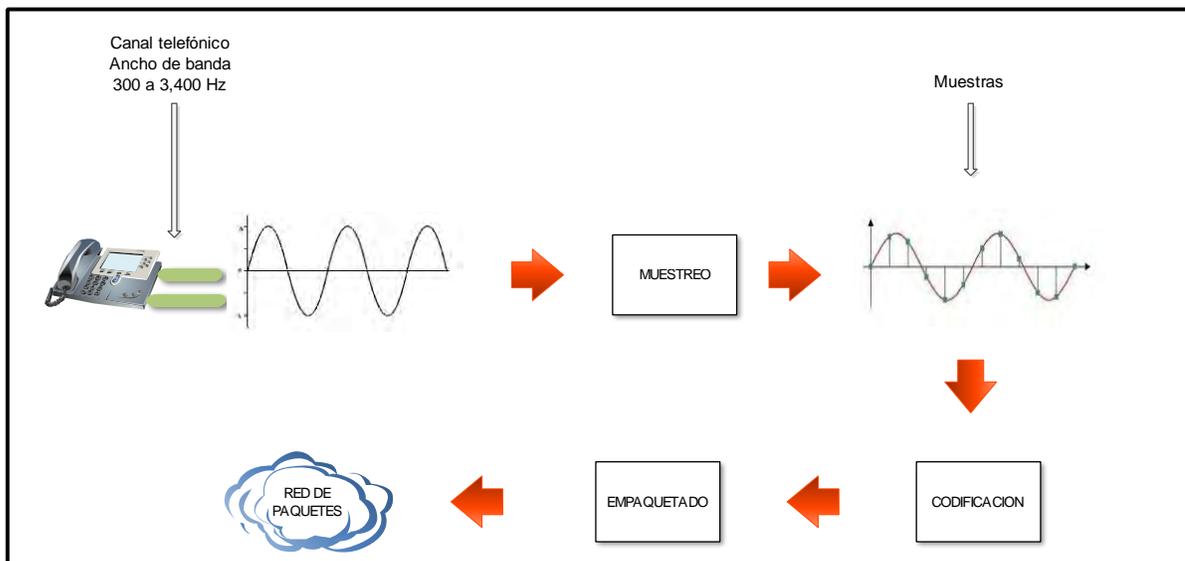


Figura 8. Digitalización de la señal de voz

### 3.2.1 Muestreo

El muestreo o discretización, de una señal consiste en seleccionar, de todos los valores que toma dicha señal a lo largo del tiempo, únicamente los de ciertos instantes concretos. Estos instantes deben seleccionarse con cuidado para que la pérdida de información que supone el muestreo no represente una pérdida importante.

Una de las características de señal que va a marcar el proceso de muestreo es su limitación de banda. Según el Teorema de Nyquist, la frecuencia con que han de recogerse las muestras de una señal debe ser, al menos del doble de su ancho de banda:

$$f_s \geq 2 * BW_{señal}$$

Así, en el caso de un canal telefónico: 4 kHz con las frecuencias de guarda, el muestreo debe ser a 8,000 Hz, es decir, una muestra cada 0.125 milisegundos.

Para frecuencias de muestreo menores que la frecuencia de Nyquist, resulta complejo recuperar la señal original a partir de sus muestras porque se produce una dispersión energética que distorsiona la señal resultante. En efecto, el muestreo periódico de la señal puede verse, en el dominio de la frecuencia, como una repetición de dicha señal cada  $f_s$ . Si el valor de la frecuencia de muestreo es menor que el doble del ancho de banda de la señal, se produce un solapamiento entre versiones desplazadas del espectro que hace imposible la recuperación de la señal original. Sin embargo, para valores mayores esto no ocurre.

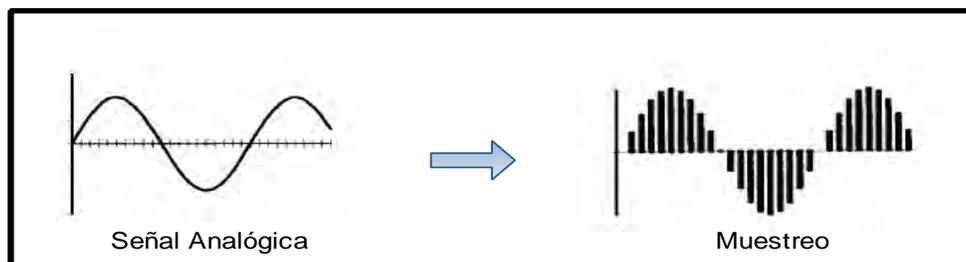


Figura 9. Muestreo de la voz

### 3.2.2 Cuantificación

El resultado del muestreo es un conjunto de valores de la señal tomados en ciertos instantes de tiempo, pero la señal sigue siendo continua en amplitud y es necesario discretizarla también en este dominio, es decir, asignar a cada una de las muestras un valor de uno de los  $M$  posibles y mantener ese valor de la señal hasta el siguiente instante de muestreo (señal en escalones). Este proceso recibe el nombre de cuantificación y supone la introducción de un error (ruido de cuantificación) que debe controlarse y que es fruto de que ciertos valores de la señal se eliminan o ajustan al conjunto de valores posibles.

El objetivo de la cuantificación es, conseguir que la señal quede representada en un número finito de bits  $N$ , con el que podrán representarse hasta  $2^n - 1$  valores diferentes. El resultado será la representación digital de la señal. Para cuantificar la señal, se divide su rango dinámico (conjunto de valores que puede tomar la señal), de manera que el valor cuantificado vendrá dado según una cierta distribución.

La cuantificación uniforme (en el que el paso del cuantificador es constante) supone que todas las amplitudes de la señal acontecen con la misma probabilidad y requiere de 12 bits por muestra, resultando 96 kbps.

Sin embargo, éste no es el caso de la señal de la voz, en la que los valores más probables son los de amplitudes medias. Por ello, sería deseable emplear más bits con estos niveles y menos bits en los que aparecerán con menor probabilidad. Esto se consigue haciendo que el paso del cuantificador varíe dentro del rango dinámico de la señal tratándose, entonces, de una cuantificación no uniforme (logarítmica).

La cuantificación logarítmica equivale a una compresión de la señal, una cuantificación uniforme de la señal comprimida y una posterior expansión en el otro extremo, para la que se utiliza la misma ley empleada en transmisión. Esta ley recibe el nombre de ley de compansión (compresión-expansión).

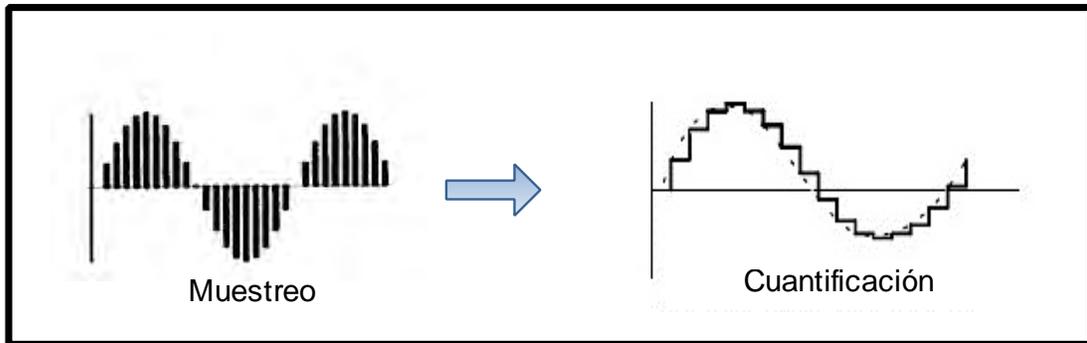


Figura 10. Cuantificación de la voz

### 3.2.3 Codificación de la voz

Una vez que la señal ya presenta un formato digital el paso siguiente es codificarla, es decir, adaptarla para que sus características sean las idóneas a la hora de transmitirla por un canal de comunicaciones concreto. La codificación (aunque también se entiende por tal proceso completo de digitalización y compresión), en este caso consiste en asignar un código binario a cada uno de los valores discretos de la señal (con  $k$  bits codifico  $M$  valores, siendo  $M = 2^k$ ). En el ejemplo del canal telefónico  $M = 256$  y  $k = 8$  ( $256 = 2^8$ ).

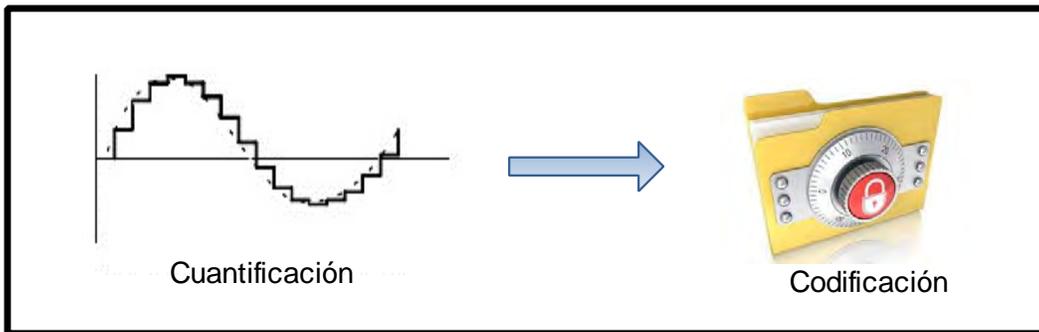


Figura 11. Codificación de la voz

Tradicionalmente, en entornos telefónicos se ha venido utilizando la modulación por codificación de pulsos o PCM (*Pulse Code Modulation*), en la que cada muestra de voz se representa por 8 bits, resultando un flujo de 64 kbps que coincide con la velocidad asignada a un canal básico RDSI.

Si la frecuencia de muestreo se escoge respetando el Teorema de Nyquist, la codificación recibe el nombre de G.711, por la norma ITU-T que la recoge. Este codificador o códec es el que mayor calidad de voz consigue, sin embargo, el precio que hay que pagar es un mayor consumo de ancho de banda, pues requiere de 64 kbps.

El ancho de banda (*bandwidth*) de un decodificador, también llamado régimen binario de salida, es función de la frecuencia de muestreo ( $f_s$ ) y del número de bits empleados para codificar cada muestra ( $N$ ):

$$R = N * f_s$$

Para el códec G.711, el ancho de banda de salida será, por tanto, de 64 kbps. En cualquier red, el ancho de banda es un recurso limitado, no sólo por sus características físicas sino porque hay que pagar por él. Por tanto, son necesarias otras soluciones de codificación de la voz que, ofreciendo unos niveles de calidad menores, sean más efectivas en el aprovechamiento del ancho de banda.

Una manera de reducir el ancho de banda consumido es disminuir el número de bits empleados en la codificación. Esto puede hacerse, por ejemplo, codificando la diferencia entre muestras sucesivas (DPCM, *Differential PCM*) o adaptando el tamaño del paso de cuantificación varíe a lo largo del rango dinámico de la señal (ADPCM, *Adaptive Differential PCM*). Codificadores de este tipo son el G.721 y G.726 que, junto con G.711, pertenecen a los codificadores de forma de onda, es decir, los que codifican directamente los valores que toma la señal en el dominio temporal.

Tipo	Frec. De Muestreo (kHz)	Ancho de banda (kHz)	Bits por muestra	Relación S/R (dB)	Canales	Caudal
Sonido Telefónico (G.711)	8	4	8	48	1	64 kbps (RDSI)
CD-DA (Compact Disk Digital Audio)	44.1	22.05	16	96	2	1.411 Mbps (CD-ROM 1x)
DVD-Audio	192 (Máx.)	96	24 (Máx.)	144 (Máx.)	2	9.216 Mbps

Tabla 2. Audio digital sin compresión

Otra manera de disminuir el consumo de ancho de banda es emplear frecuencias de muestre menores que la de Nyquist. El problema, entonces, es la distorsión energética que sufre la señal durante el muestreo. Para resolverlo y poder así recuperar una copia de la señal original, deben emplearse otros mecanismos más complejos como la interpolación o la predicción lineal que tratan de estimar tramas anteriores y posteriores a la actual mediante diversas técnicas. Ejemplos de este tipo de codecs son el G.723 y G.729. Todos estos codificadores, especializados en la codificación de la voz, reciben el nombre de codificadores de voz o *vocoders*, que no resultan adecuados para música u otros sonidos.

Sin embargo, conseguir la disminución del ancho de banda supone pagar un precio nada despreciable en algunos casos, es decir, un aumento del retardo y una disminución de la calidad. En efecto, la compresión implica un procedimiento adicional de las muestras de voz que tiene por objetivo eliminar o, al menos, reducir, la información de redundancia presente en la señal de voz. Este procedimiento consume ciclos de CPU del DSP del codificador e introduce un retardo adicional que, en tráfico de tiempo real como el tráfico de voz, puede producir problemas si llega a alcanzar valores excesivamente altos.

Una vez codificada la señal, habrá que empaquetarla y transmitirla por la red, bien sea por una red IP privada o pública, o la propia Internet.

#### 3.2.4 Codecs (codificador-decodificador)

La comunicación de voz es analógica, mientras que la red de datos es digital. El proceso de convertir ondas analógicas a información digital, se hace con un único equipo: el codificador/decodificador (códec), que además de llevar a cabo la conversión de analógico/digital, comprime la secuencia de datos, y proporciona la cancelación del eco.

Existen muchas maneras de transformar una señal de voz analógica a digital. El proceso de la conversión es complejo y la codificación de la voz, entendida como la digitalización y la compresión de la voz, puede ser realizada mediante tres técnicas principales: por codificación de forma de onda, por codificación basada en modelos matemáticos sobre la producción de la voz y en modelos híbridos que combinan ambas técnicas.

El ahorro de ancho de banda para los servicios de voz puede lograrse de varias maneras y trabajar en diversos niveles. Por ejemplo, la compresión analógica puede ser parte del esquema de codificación (algoritmo) y no necesita la compresión digital adicional de las capas de trabajo más altas. Otra manera de ahorrar ancho de banda es el uso de supresión del silencio, que es el proceso de no enviar los paquetes de la voz entre silencios en conversaciones humanas.

La siguiente tabla muestra algunos de los estándares de codificación más importantes cubiertos por la unión internacional de telecomunicaciones (ITU) y, como se puede ver, la calidad es menor cuanto mayor es la compresión.

Códec	Ancho de banda (kHz)	Frecuencia de Muestreo kHz	Duración muestras/tramas (ms)	MOS	Aplicación
G.711 (PCM)	64	8	0.125	4.4	Telefonía
G.721 (ADPCM)	32	8	0.125	4.2	Telefonía
G.722 (SB-ADPCM)	48/56/64	16	0.250		Videoconferencia
G.728 (LD-CELP)	16	8	0.625	4.2	Telefonía / Videoconferencia
G.729 (CS-ACELP)	8	8	10	4.1	Telefonía
G.723.1 (MP-MLQ)	6.3	8	30	3.9	Telefonía Internet
G.723.1 (ACELP)	5.3		30	3.6	Telefonía Internet

Tabla 3. Codecs más comunes de la telefonía

Algunos esquemas de compresión, tales como el G.729 y el G.723.1, consiguen una baja tasa de bits, reduciendo considerablemente la necesidad de ancho de banda, por lo que resultan apropiadas para Internet. Por ejemplo, el G.723.1, aparecido en noviembre de 1988, incluye un esquema de compresión basado en la supresión de silencios y VAD (*Voice Activity*

*Detection*). Estos esquemas de codificación intentan reproducir el sonido subjetivo de la señal, más que la forma de onda, pero son muy sensibles a la pérdida de paquetes o al *jitter*, por lo que emplean, por ejemplo, las técnicas de interpolación para reducir al mínimo estos efectos.

Utilizar la compresión y/o supresión del silencio puede dar lugar a un ahorro importante del ancho de banda necesario para la transmisión; sin embargo, hay algunos usos que se podrían ver afectados por ella, como, por ejemplo, el impacto en los usuarios que utilizan modem, puesto que los esquemas de compresión pueden interferir con su funcionamiento y el resultado podría ser que nunca se sincronicen o que tuviesen un rendimiento de procesamiento muy bajo. Para evitarlo algunos gateways implementan una cierta inteligencia que si detecta el uso de módems la inhabilita.

Los codecs que ofrecen tasa constante, por ejemplo, G.711 y G.729, son más adecuados para servicios orientados a conexión, como RSVP o circuitos CBR de ATM, mientras que los de tasa variable, por ejemplo, G.723.1, son los que mejor se adaptan a redes sin reserva de tasa constante, como el modelo DiffServ de Internet o los servicios UBR o ABR de ATM. La compresión MPEG es la más eficiente y da mayor calidad, pero consume mucho CPU e introduce mucho retardo por lo que no puede emplearse en aplicaciones interactivas (telefonía o videoconferencias).

La salida del códec es una secuencia de datos, que se convierte en paquetes IP y se transporta a través de la red IP hasta su destino, que debe utilizar los mismos estándares, así como parámetros comunes, para poder realizar el proceso inverso, pues si no, el resultado es una comunicación ininteligible [9].

### 3.3 Protocolos de señalización

Los protocolos de señalización en redes de voz sobre paquetes realizan muchas funciones similares a las que llevan a cabo sus homólogos en redes telefónicas (establecimiento de la sesión, señales de progreso de llamadas, gestión de los participantes en la llamada). Sin embargo, dadas las características particulares de las redes de paquetes deben encargarse, junto

con las anteriores, de otras tareas específicas. Por otra parte, las expectativas de calidad del usuario exigen una red de señalización de altas prestaciones, pues la disponibilidad de la red debe ser similar a la de la PSTN. Ésta es la razón la que la fiabilidad no sólo deba residir en los elementos de la red sino también en la arquitectura de señalización empleada.

Al igual que ocurría en las redes telefónicas convencionales en las que es posible distinguir entre señalización de usuario y señalización de redes, en redes de voz sobre paquetes también existen dos tipos de protocolos de señalización dependiendo del ámbito de ésta. Así, tenemos protocolos de señalización entre terminales y protocolos de señalización en la red IP.

En realidad, los protocolos de señalización entre terminales son comunes a cualquier tipo de comunicaciones multimedia (voz, vídeo y audio) a través de las redes de paquetes. Aplicadas a la voz sobre paquetes, tienen como objetivo mantener la interfaz con el usuario típica de las redes telefónicas, es decir, generar los tonos y señales necesarios para que el usuario no perciba que la tecnología de soporte de las llamadas telefónicas ha cambiado. Los protocolos más destacados, son el H.323 de la ITU-T y el SIP del IETF.

	<b>H.323</b>	<b>SIP</b>
<b>Organismo de estandarización</b>	ITU	IETF
<b>Arquitectura</b>	Distribuida	Distribuida
<b>Versión actual</b>	H.323v5 (julio 2003)	RFC 3261-3266
<b>Responsable del control de llamadas</b>	Gatekeeper	Servidor proxy o servidor de desvío
<b>Puntos finales</b>	Gateway, terminal	Agente de usuario
<b>Señalización</b>	TCP o UDP	TCP o UDP
<b>Soporte multimedia</b>	Sí	Sí
<b>DTMF-relay</b>	H.245 (señalización) o RFC 2833 (datos)	INFO (señalización) p RFC 2833 (datos)
<b>Fax-relay</b>	T.38	T.38

<b>Servicios suplementarios</b>	Proporcionados por los puntos finales o el responsable del control de llamadas	Proporcionados por los puntos finales o el responsable del control de llamadas
-------------------------------------	--	--

Tabla 4. Comparación entre H.323 y SIP

Los protocolos de señalización en la red de paquetes, por su parte, surgen como respuesta a la necesidad de nuevas redes de señalización más atractivas que las basadas en la conmutación por circuitos tradicional. El mayor coste económico asociado a la señalización procede del precio de las centrales de conmutación, que deben tener cierta inteligencia y capacidad de proceso puesto que las funciones de señalización se encuentran distribuidas entre todos los conmutadores de la red. A la hora de diseñar los protocolos de señalización en la red de paquetes se ha tenido en cuenta este hecho y, por ello, las funciones de señalización se centralizan en un controlador que establece un diálogo con los gateways y las controla remotamente.

### 3.3.1 H.323

H.323 es, en realidad, un conjunto de protocolos que definen los componentes y los medios de interacción entre los mismos que deben cumplirse para soportar comunicaciones multimedia sobre redes de paquetes sin conexión ni garantía de calidad de servicio, como es el caso de las redes IP. Originalmente fue desarrollado en 1996, bajo la protección de la ITU, para soportar conferencias multimedia sobre las redes LAN, aunque con posterioridad se ha aplicado también a la voz sobre paquetes. Hoy día, la gran parte de *Gateways* y *Gatekeepers* que se encuentran en el mercado lo soportan.

Una red H.323 está compuesta por cuatro tipos de elementos:

**Terminales:** Son los equipos utilizados por los usuarios finales y abarcan desde teléfonos tradicionales hasta teléfonos IP pasando por hosts y sistemas de grupo o sala.

**Gateways:** Se encargan de la interconexión entre redes heterogéneas, por ejemplo de una red H.323 con la PSTN. Sus funciones básicas son la traducción de protocolos de establecimiento y liberación de llamadas y la conversión de los formatos de la información entre diferentes tipos de redes.

**Gatekeepers:** Es el componente más crítico de una red H.323, cuando está presente, puesto que se encarga de las tareas de control en su zona de influencia, como, la gestión de la zona, el control de admisión, la gestión del ancho de banda y la traducción de direcciones (IP-números de teléfono).

**MCU:** Se emplean en comunicaciones simultáneas entre más de dos usuarios. Se encargan de mezclar los flujos de audio y vídeo y distribuir dichos flujos entre los participantes.

El diálogo entre terminales H.323 y el *gatekeeper* se desarrolla mediante la especificación H.225 RAS (*Registration, Admission and Status*) y la comunicación entre ambas entidades tiene como finalidad el registro, la admisión y el control del estado de un terminal de usuario. Para cada llamada se establece un canal de señalización entre el terminal (*endpoint*) y el *gatekeeper* RAS y un canal de señalización entre los terminales (señalización Q.931) y un canal lógico de control entre los terminales (canal H.245).

H.323 emplea UDP como protocolo de transporte de la voz y el vídeo. Ambos flujos de información se codifican respectivamente según las especificaciones G.7xx y H.26x. Dentro de H.323, complementado a UDP, nos encontramos los protocolos RTP y RTCP.

### 3.3.2 SIP

Como alternativa a H.323, el IETF publicó en marzo del año 1999 el protocolo SIP (*Session Initiation Protocol*) bajo la RFC 2543. Se trata de un protocolo de control de la capa de aplicación que define cómo establecer, modificar o finalizar una sesión entre dos o más extremos, independientemente de tipo de sesión de que se trate.

A diferencia de lo que sucede en H.323, en SIP sólo se definen los elementos que participan en un entorno SIP y el sistema de mensajes que

intercambian estos. Estos mensajes están basados en HTTP y se emplean esencialmente en procedimientos de registro y para establecer entre que direcciones IP y puertos TCP/UDP intercambiarán datos los usuarios. En este sentido, su sencillez es altamente valorada por los desarrolladores de aplicaciones y dispositivos.

La arquitectura SIP es muy similar a la de HTTP. Las solicitudes del cliente son enviadas a un servidor, éste las procesa y envía una respuesta al cliente. El estándar distingue entre agentes de usuario y servidores de red.

Los agentes de usuario (UA, *User Agent*) son las terminales de los que parten las solicitudes de iniciar una nueva llamada o de terminar una llamada en curso. Los UA se dividen, a nivel lógico, en dos entidades: el agente de usuario cliente (UAC, *User Agent Client*), encargada de iniciar las sesiones SIP, y el agente de usuario servidor (UAS, *User Agent Server*), responsable de aceptar las peticiones de establecimiento de sesión recibidas.

Los servidores o servicios de red, por su parte, actúan como intermediarios en las comunicaciones entre los agentes de usuario y existen cuatro tipos, que son:

- **Servidor *Proxy***. Es una entidad intermedia que actúa como servidor y como cliente y que realiza peticiones de parte de otros clientes.
- **Servidor de localización**: Proporciona información acerca de la localización del usuario. Si un usuario A desea comunicarse con un usuario B, en primer lugar A necesita descubrir la localización actual de B en la red, con el fin de que la petición de establecimiento de sesión pueda llegar.
- **Servidor de redirección**: Acepta una petición SIP, mapea la dirección en cero o más direcciones nuevas y devuelve estas direcciones al cliente. A diferencia de los servidores *proxy*, los servidores de redirección no inician su propia petición SIP.

Tampoco pueden aceptar o terminar llamadas como ocurre con los UAS.

- **Servidor de registro:** Acepta peticiones de registros de los UAC y actualiza la información relativa a cada uno de ellos en una base de datos de localización.

El diálogo entre los clientes y los servidores SIP se basa en el intercambio de mensajes de texto. Existen dos tipos de mensajes SIP, los mensajes de petición, que son enviados por las entidades cliente a las entidades servidor y los mensajes de respuesta que son enviados por las entidades servidor a las entidades cliente.

Las peticiones SIP son las siguientes:

**INVITE:** Mensaje inicial de invitación enviado por el extremo llamante.

**ACK:** Respuesta de agente llamante ante el mensaje de aceptación de la llamada por parte del destino.

**BYE:** Señal de terminación de la sesión por parte de uno de sus participantes.

**CANCEL:** Cancela una petición pendiente.

**REGISTER:** Empleado por los usuarios para registrar su dirección de contacto actual.

**OPTIONS:** Consulta a un agente de usuario acerca de sus capacidades.

**INFO:** Contiene información fuera de banda, como dígitos DTMF.

Las respuestas SIP se relacionan con un código que indica el estado de la transacción:

**1xx:** Mensaje de información.

**2xx:** Éxito, la acción fue recibida satisfactoriamente.

**3xx:** Mensaje de desvío.

**4xx:** Error en la petición, la petición contenía error y no pudo ser completada por el servidor

**5xx:** Error en el servidor, el servidor presento un error.

**6xx:** Error general.

Los seis pasos básicos implicados en el manejo de sesiones SIP son, registro, inicio y localización; descripción de la sesión multimedia que se pretende establecer; aceptación de la petición de conexión del otro extremo; establecimiento de la llamada; comunicación y terminación de la llamada [9].

# CAPÍTULO 4

## Capítulo 4. ARQUITECTURA H.323

H.323 es una recomendación de la ITU-T basada en la familia de estándares H.320. La versión actual de la recomendación es la versión 7. Inicialmente, el protocolo (versión 1) fue diseñado para proporcionar señalización a sistemas de conferencia multimedia para entornos LAN sin proveer calidad en los servicios. Sin embargo, actualmente, se ha convertido en un conjunto de especificaciones que definen la arquitectura completa y el funcionamiento de un sistema de conferencia multimedia sobre redes de área amplia. En contraste con su ámbito de aplicación original, se ha convertido en una solución escalable que puede interactuar con redes administrativas a gran escala.

### Familia H.323

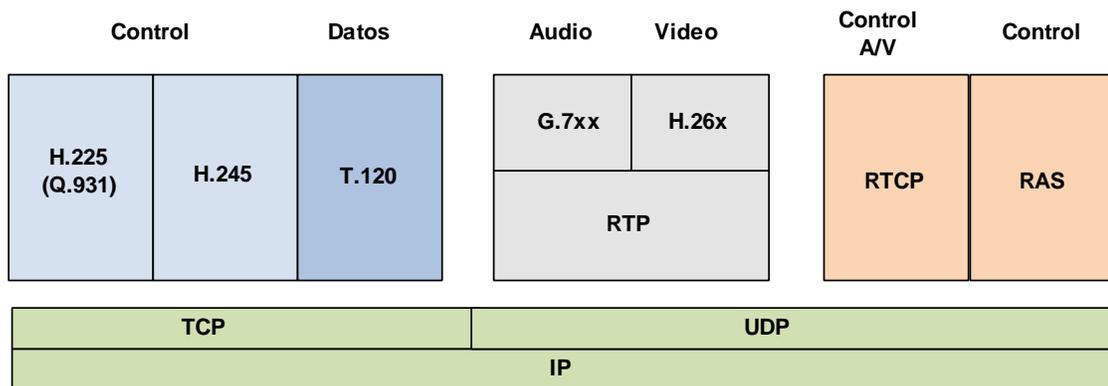


Figura 12. Protocolos de la familia H.323

Un sistema H.323 proporciona las operaciones de señalización y de control necesarios para llevar a cabo comunicaciones multimedia sobre una red de paquetes, que no proporciona una calidad de servicio garantizada. La interfaz de red actual, la red física y la de los protocolos de transporte utilizados en la red no están incluidas en el alcance de H.323. Un sistema H.323 se compone de las siguientes entidades: terminales, gatekeeper, gateways y MCU [13].

## 4.1 Entidades H.323

**Terminal:** es el cliente en el punto extremo y debe soportar 4 protocolos, H.245, Q.931, RAS y RTP. RAS tiene 4 funciones básicas: ayudar a la terminal a encontrar un Gatekeeper y registrarse con él y solicitarle permiso de admisión antes de comenzar una llamada. H.225 utiliza un subconjunto de mensajes y parámetros de Q.931, y realiza toda la señalización necesaria para establecer y mantener una conexión entre 2 terminales. Q.931 es transmitida sobre TCP. H.245 se utiliza para transmitir de extremo a extremo señalización de control de medios. Estos mensajes de control contienen información de las capacidades de cada terminal y definen procedimientos para el mapeo lógico de los canales. H.245 también define cuál de las terminales será “maestro” y cual será “esclavo”, lo cual es necesario para evitar conflictos que pueden ocurrir cuando 2 terminales inician un evento similar al mismo tiempo.

**Gatekeeper:** encargado del control de llamadas y la administración de las políticas por zonas para cada terminal. Todas las terminales deben estar registradas con un gatekeeper. Las principales funciones de un Gatekeeper son:

- *Traducción de direcciones Alias:* la traducción de una dirección alias a una dirección de transporte.
- *Control de Admisión:* el acceso a una red puede ser aceptada o rechazada basándose en una autorización de llamada, la dirección de origen o destino. Para hacer esto utiliza los mensajes RAS, pero determina como serán utilizados los recursos de red.
- *Administración y Control del Ancho de Banda:* controla el número de terminales que pueden existir en una red al mismo tiempo.
- *Administración de Zonas:* las terminales, Gateway y MCUs (Unidades Multipunto) son administrados por un solo Gatekeeper.

Además de las funciones anteriores, también brinda algunas funciones opcionales:

- *Autorización de la llamada:* puede aceptar o rechazar llamadas basándose en muchos factores, como la hora del día y restringir acceso a una terminal o Gateway en particular.

- *Señalización del control de la llamada:* puede decidir si va a procesar todas las llamadas de señalización asociadas a los terminales dentro de su zona o permitir que los terminales puedan intercambiar directamente mensajes de señalización entre ellos.
- *Administración de llamadas:* conserva información acerca del estado de las llamadas que están activas, como el ancho de banda utilizado. El redireccionamiento de las llamadas también está incluido en este servicio.

**Gateway:** es el dispositivo de red que proporciona comunicaciones bidireccionales en tiempo real, entre la red IP y la red telefónica tradicional.

**MCU:** provee soporte para conferencias entre 3 o más terminales. Está compuesto por un Controlador Multipunto (MC) y uno o más procesadores multipunto (MP). El primero brinda control de conferencia, y el segundo se encarga de mezclar, intercambiar y procesar las cadenas de medios. [14]

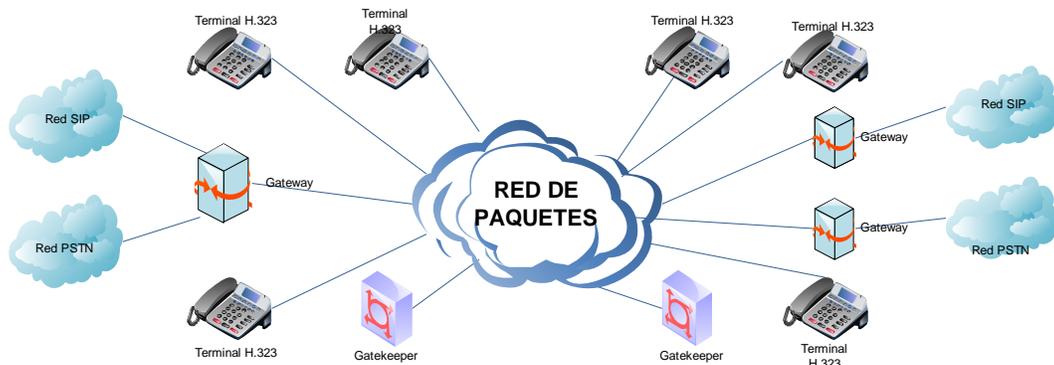


Figura 13. Arquitectura H.323

Un aspecto central de H.323 es la llamada H.323. Se define como la comunicación multimedia de punto a punto entre dos extremos. Si una terminal H.323 se comunica con una terminal que utiliza un protocolo de señalización diferente, entonces la llamada H.323 se define como el segmento

de la llamada H.323 y el Gateway es el que proporciona la interconexión con la red externa.

La llamada puede tener múltiples participantes. Comienza con el procedimiento de establecimiento de llamada y termina con el procedimiento de terminación de llamada. Los mensajes de señalización y de control pueden intercambiados directamente entre las dos terminales o a través de una o más entidades H.323, como el gatekeeper, gateways o controladores multipunto.

El sistema H.323 se divide en zonas. Cada zona está compuesta por el conjunto de todas las terminales, gateways y controladores de multipunto, todos gestionados por un gatekeeper. Una zona no está necesariamente debe estar dentro de un mismo segmento de red; ya que puede abarcar a través de la red múltiple segmentos que están interconectados.

El protocolo H.323 es una familia estrechamente unida de subprotocolos que deben de operar de manera ordenada para establecer con éxito una sesión de llamada. Los principales subprotocolos que se describen en las Recomendaciones de la ITU-T, son los siguientes:

**H.225:** Intercambia mensajes entre las terminales de H.323 para el establecimiento y finalización de las llamadas, así como mensajes de control.

**H.245:** Intercambia mensajes para la negociación de las capacidades de la sesión y creación de los canales lógicos.

**H.235:** Subprotocolo de seguridad y cifrado para terminales H.323.

**H.450:** Subprotocolo para servicios suplementarios, tales como transferencia de llamadas, llamada en espera.

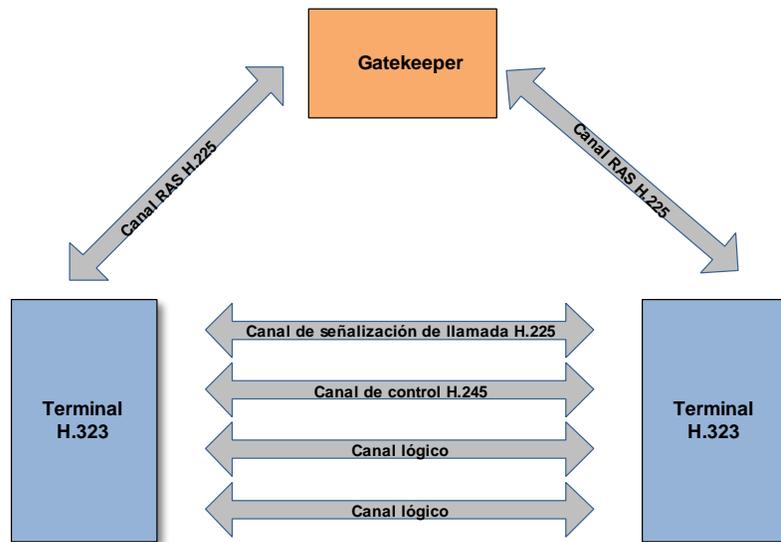


Figura 14. Llamada de terminal directa y control de señalización

El protocolo H.323 se implementa mediante el intercambio de mensajes entre las terminales y las entidades que intervengan. Estos mensajes se codifican utilizando ASN.1, que es un formato binario para definir la sintaxis de datos de información. A pesar de que H.323 es un protocolo basado en paquetes, permanece firmemente apegada con las normas tradicionales de telecomunicaciones. Como tal, se define diversos canales de comunicación, cada uno con su propio subprotocolo para la comunicación entre las distintas entidades H.323. En concreto, el protocolo distingue entre tres canales de comunicación:

1. **El canal RAS:** (*Registration, Admission and Status*) que es un canal entre la terminal H.323 y el gatekeeper. Se utiliza para el intercambio de mensajes de registro, admisión, ancho de banda y de estados. Los mensajes intercambiados en este canal se llevan a cabo mediante H.225.
2. **El canal de la señal de llamada:** es un canal entre las terminales H.323 (directamente o a través del gatekeeper). Se utiliza para realizar el establecimiento de llamada y la finalización. Los mensajes intercambiados en este canal se llevan a cabo mediante H.225.

3. **El canal de control H.245:** es un canal entre las terminales H.323 (directamente o a través del gatekeeper). Se utiliza para intercambiar los mensajes de control de llamada H.245. Los mensajes intercambiados en este canal se llevan a cabo mediante H.245.

La señalización de toda la sesión de llamada se realiza a través de mensajes intercambiados en estos tres canales. H.323 define un tipo adicional de canal; el canal de comunicación lógica. Esta canal lleva el contenido de los medios y cada sesión puede tener uno o más canales establecidos a través del canal de control H.245.

## 4.2 Funcionamiento

En H.323, una sesión de llamada consta de cinco fases:

1. Llamada
2. Inicio de comunicación e intercambio de capacidades
3. Establecimiento de comunicación audio visual
4. Servicios de llamada
5. Terminación de llamada

H.323 hace uso de tres canales de señalización diferentes con el fin de completar estas cinco fases. La secuencia de eventos que tienen lugar durante una sesión de llamada típica se indica de la siguiente manera. Inicialmente, una terminal H.323 transmite la información de registro a una entidad H.323, el gatekeeper, mediante H.225 se intercambian mensajes a través del canal RAS. El registro será realizado sólo una vez o periódicamente de acuerdo con las políticas del gatekeeper; de cualquier manera, el proceso de registro no está relacionado con una llamada particular.

Siempre que una terminal desea realizar una llamada, solicita el permiso al Gatekeeper a través del canal RAS. Si se concede el permiso, la terminal de alguna manera descubre la dirección de señalización del canal de transporte

de la terminal a la que desea llamar. Luego intenta configurar una sesión de llamada a través de este canal, utilizando mensajes a través del H.225.

Una vez que el canal de señalización de la llamada se ha establecido, las dos terminales proceden a configurar el canal de control de la señalización, utilizando mensajes H.245. Este es el canal que se utiliza por completo para la llamada y a través del cual se realizan todos los servicios de señalización. En este punto cualquiera de las terminales puede finalizar la llamada a través del canal de señalización H.225 ya que hayan cumplido su propósito.

Los dos puntos finales proceden a intercambiar y negociar capacidades de recursos y establecer los canales de medios. Por lo general, se establecen tres canales de medios; uno para audio, uno para vídeo y otro para datos. Las dos terminales ahora pueden establecer una conversación multimedia. A menos que ningún servicio de llamada sea solicitado durante la conversación (modificación de las características de llamada, establecimiento de canales de medios adicionales, agregar una terminal más) y no exista más tráfico en el canal de control H.245 antes de que la fase de terminación de llamadas actúe.

Cuando la conversación ha terminado, las terminales cierran todos los canales de medios que habían abierto y una de las terminales inicia la petición de finalización de llamada al canal de control de señalización H.245. Si canal de señalización H.225 sigue activo, entonces se procede a cerrarlo.

#### 4.2.1 Descubrimiento de Gatekeeper y registro de la terminal

Antes de realizar cualquier llamada con la ayuda del gatekeeper, cada terminal debe registrarse en su propio gatekeeper. Esto se logra a través de un proceso de registro corto. Sin embargo, con el fin de registrarse, la terminal debe encontrar de alguna manera al menos una dirección de un gatekeeper. Esta fase es llamada descubrimiento del Gatekeeper.

Durante el descubrimiento del gatekeeper, la terminal determina con que gatekeeper puede registrarse. Si esto no se realiza de forma manual (por ejemplo, la terminal está configurada con una dirección de un gatekeeper asociado), H.323 proporciona un método automático.

Demostramos el método de detección automática con un ejemplo en el que una terminal y tres gatekeeper están involucrados. Después de la que la fase de descubrimiento de gatekeeper está completa, la terminal debe registrarse en uno los gatekeeper que mejor le convenga. Una vez registrada la terminal puede proceder a realizar llamadas con la ayuda del gatekeeper.

Hasta este momento el gatekeeper está involucrado mínimamente en la traducción de los alias de las terminales para el transporte de direcciones y garantizar las solicitudes de una llamada.

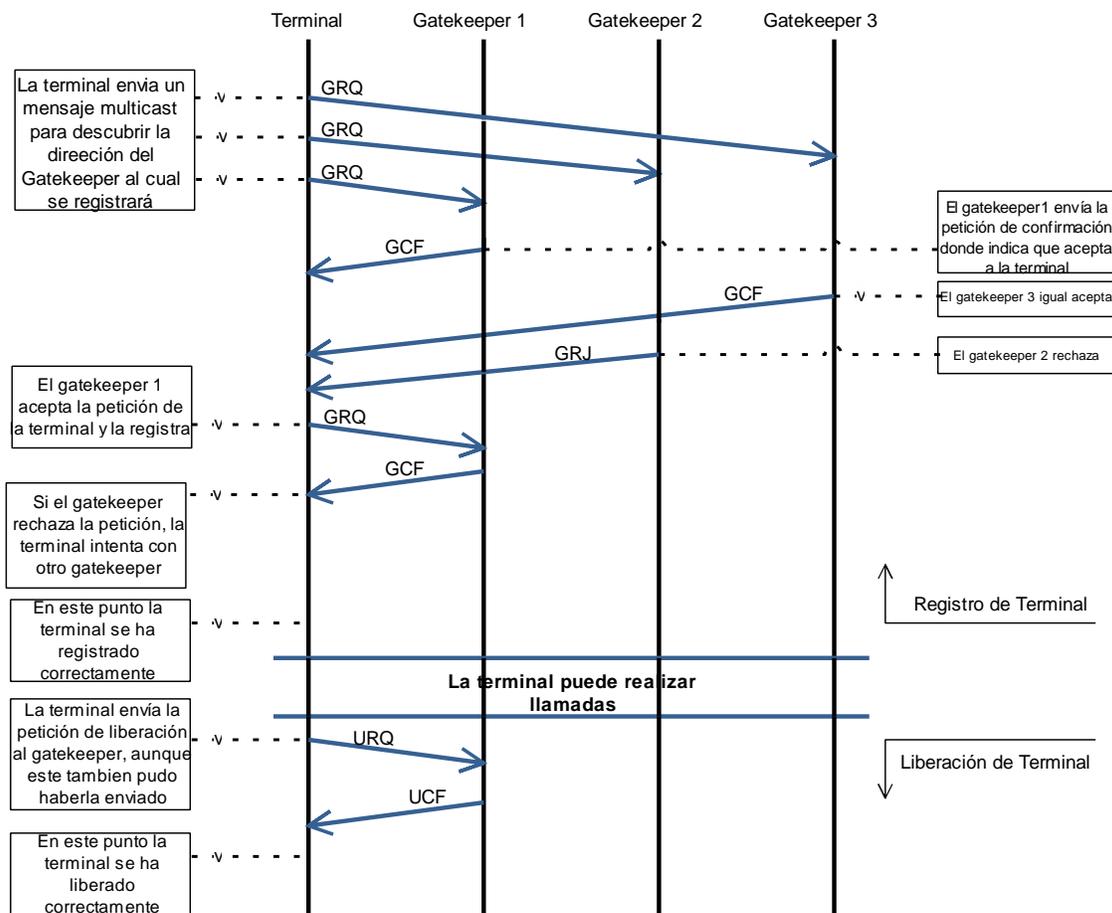


Figura 15. Búsqueda de Gatekeeper y registro de la terminal

#### 4.2.2 Fase 1: La llamada

El objetivo principal de esta fase es la de localizar a la terminal a la que se llama, establecer si el usuario acepta la llamada y continuar con la creación de del canal de señalización de llamada H.225. Luego, en la siguiente fase, el canal de señalización de llamada se utiliza para abrir el canal de control de llamada H.245, que es el protocolo que controla realmente la mayoría de los aspectos de la llamada.

Tenga en cuenta que si los puntos finales conocían las direcciones de transporte de unos a otros, podrían, en principio, omitir al gatekeeper y no solicitar el permiso para realizar la llamada. Sin embargo, la mayoría de las veces sólo se conoce el alias, y se necesitaría los servicios de traducción del gatekeeper.

Pero incluso en el caso en que no se requiera traducción del alias, una terminal puede ser obligada a utilizar el gatekeeper, por otros medios, como la colocación detrás de un firewall.

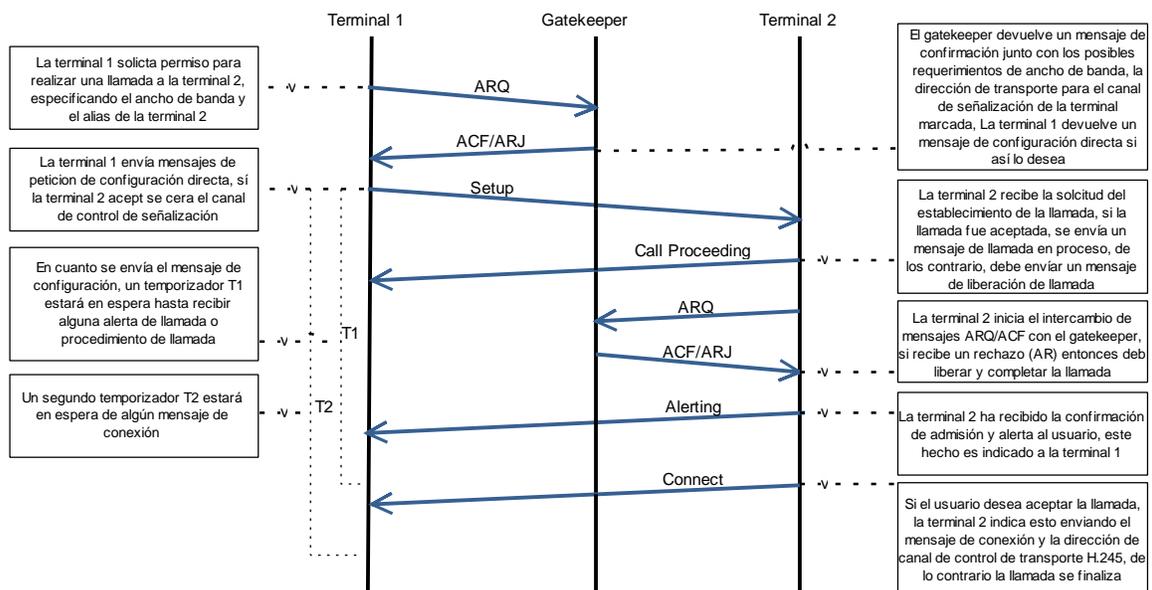


Figura 16. Establecimiento de la llamada con señalización directa

### 4.2.3 Fase 2: Inicio de la comunicación e intercambio de capacidades

Tras la instalación del canal de señalización de la llamada, las terminales procederán a establecer un canal de control H.245. Este es el canal más importante, ya que los mensajes intercambiados a través de él controlan todos los aspectos de la llamada desde este punto de la sesión llamada.

Después de que la terminal 1 ha recibido el mensaje conexión, se intercambian mensajes de control de canal H.245 donde indican la dirección de transporte donde la terminal 2 está escuchando. Por lo tanto, la terminal 1 establece el canal H.245 saliente desde su lado mediante el envío del primer mensaje, *TerminalCapabilitySet*. Este mensaje anuncia capacidades de recursos y contiene la dirección de transporte de la terminal 1 para el canal, que a su vez permite que la terminal 2 active su canal de salida.

Tenga en cuenta que el canal de control H.245 podría haber sido abierto tan pronto como cuando la terminal 2 mando su dirección de canal H.245. Esto puede ocurrir en los mensajes de procedimiento de la llamada, alertas o la conexión. El canal H.245 también puede haber sido establecido por la terminal 2 en el caso de que la terminal 1 haya enviado a su dirección de transporte H.245 la llamada.

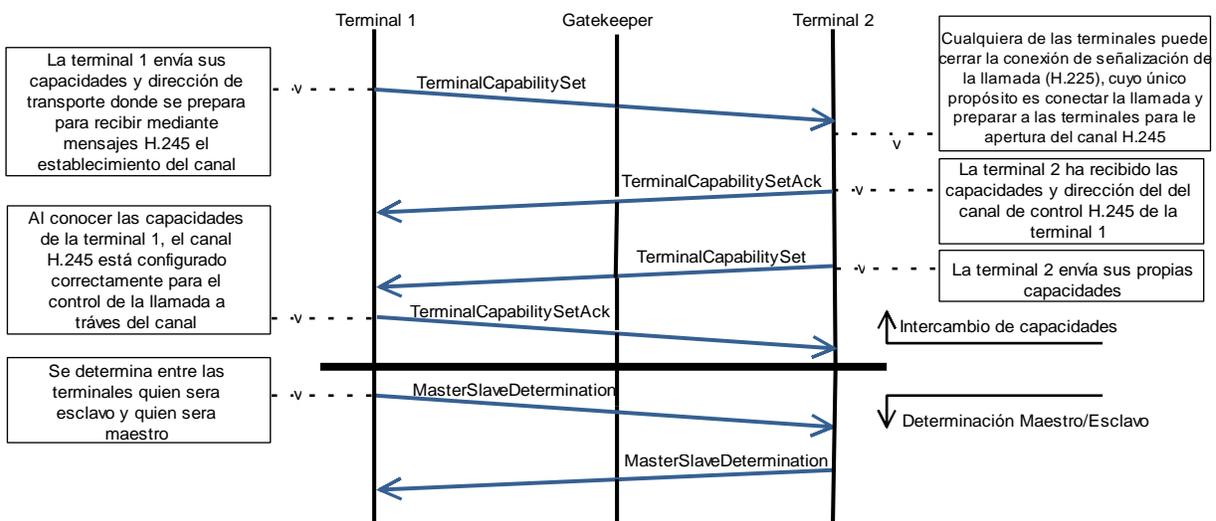


Figura 17. Inicio de la comunicación e intercambio de capacidades con señalización directa

#### 4.2.4 Fase 3: Establecimiento de la comunicación audio visual

En este punto, el canal de control H.245 se ha establecido, los criterios de valoración de capacidades se conocen entre sí y se han determinado las relaciones maestro / esclavo. Los criterios de valoración procederán a la apertura canales lógicos para el intercambio de flujos de medios

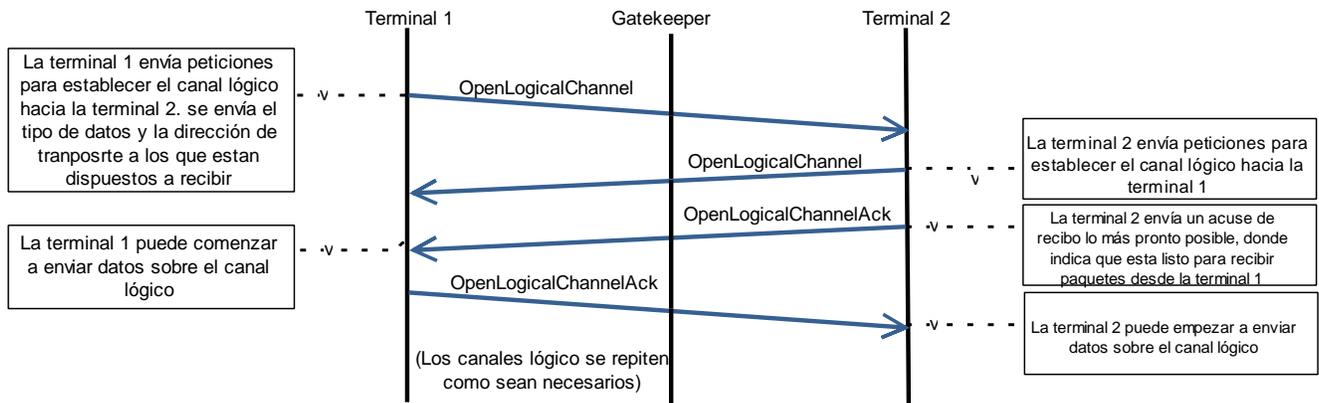


Figura 18. Establecimiento de comunicación audio visual con señalización directa

#### 4.2.5 Fase 4: Servicios de llamada

Todo está listo para la transferencia real de paquetes de medios, ya sean de audio, vídeo o datos. El protocolo en sí no está involucrado en el intercambio de los flujos reales de comunicación. Permite a otros protocolos, como RTP, para manejar la transferencia.

Sin embargo, para cualquier otro servicio durante la llamada, los mensajes H.245 se utilizan entre las terminales al igual que los mensajes RAS cuando se comunican con el gatekeeper. Estos servicios incluyen:

**Cambios de ancho de banda:** En cualquier momento durante la sesión de llamada, ya sea las terminales o el gatekeeper pueden solicitar aumentos o disminuciones de ancho de banda para los flujos de medios.

**Estado:** El gatekeeper tiene que consultar los criterios de valoración de las terminales de vez en cuando.

**Ampliación de la Conferencia:** Las terminales pueden añadir o retirar de la sesión de la conferencia en cualquier momento de la llamada.

**Servicios complementarios:** Estos servicios se describen en H.450 y son similares a los servicios proporcionados por el circuito tradicional de conmutación de redes, tales como llamada en espera, transferencia de llamadas, entre otros.

#### 4.2.6 Fase 5: Terminación de la llamada

En cualquier punto, y cualquier terminal puede terminar la llamada. Las terminales deben cerrar todos los canales lógicos que abrieron para el intercambio de medios, cerrar el canal H.245, cerrar el canal H.225 si sigue abierto e informar a sus respectivos gatekeeper sobre la finalización de la llamada.

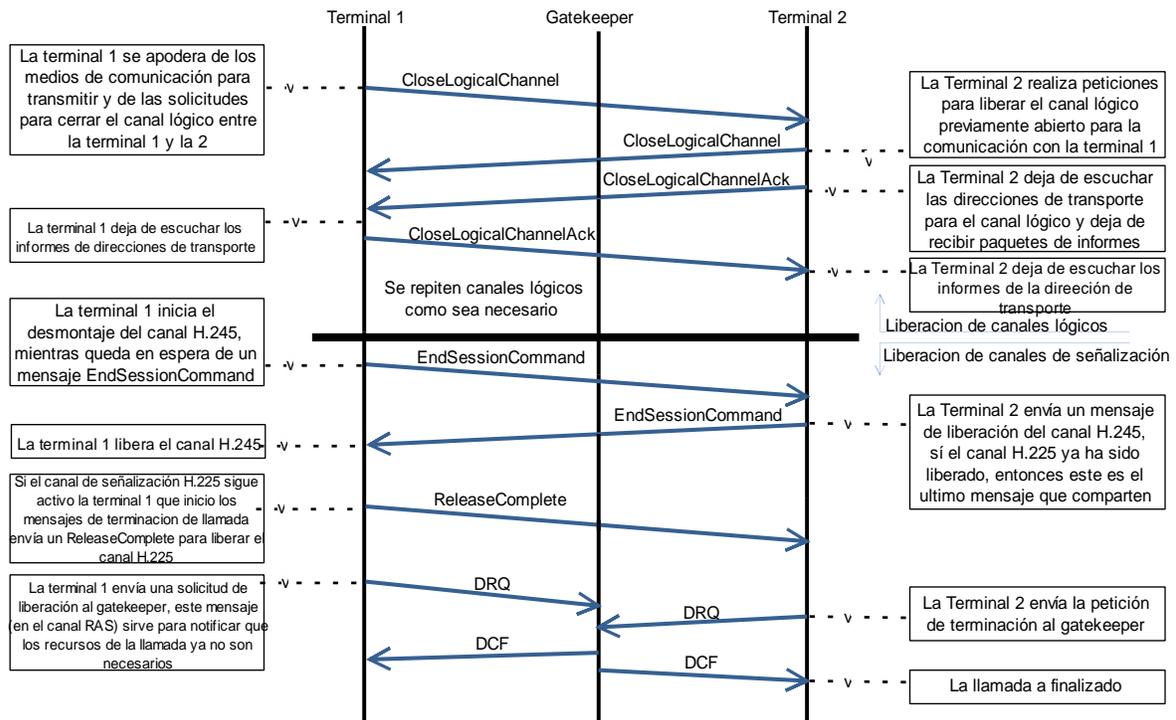


Figura 19. Terminación de llamada con señalización directa

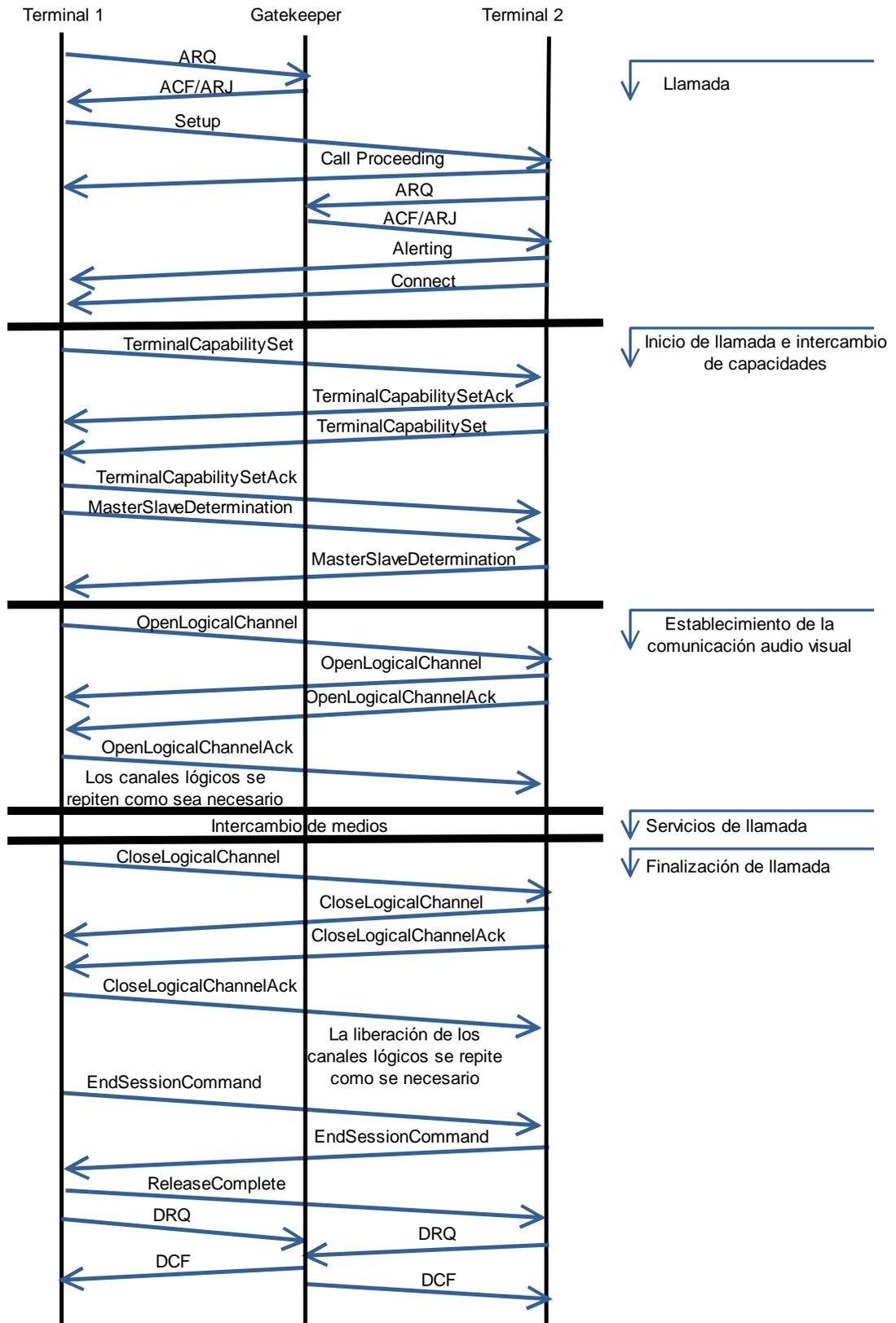


Figura 20. Señalización completa de una llamada H.323

# CAPÍTULO 5

## Capítulo 5. CALIDAD DE SERVICIO EN VOZ SOBRE PROTOCOLO IP (QoS)

Las aplicaciones distribuidas tradicionalmente están orientadas al tráfico de datos y, por tanto, presentan exigencias de calidad y prestaciones diferentes a las de las redes de voz sobre paquetes. Por ejemplo, la transferencia de un archivo consume gran cantidad de ancho de banda al requerirse que el envío se realice lo más rápido posible. Sin embargo, la voz es un tráfico en tiempo real que no consume demasiado ancho de banda, pero que tiene una baja tolerancia al retardo y al *jitter*. Es decir, que aunque los dos tipos de tráfico pueden transportarse por la misma red, no es posible manejarlos de la misma manera.

La calidad de voz es un aspecto en el que influyen gran cantidad de factores, tanto subjetivos como objetivos. Centrándonos en una visión más cercana al usuario que, finalmente, será el que decidirá sobre la bondad o no de la solución tecnológica, la calidad de voz se entendería como la fidelidad con la que se escucha la voz del otro extremo (claridad de la voz) y la capacidad de la red para soportar el flujo normal de la conversación.

Por otro lado, desde una perspectiva más centrada en la ingeniería de red, la calidad es la capacidad de la red para satisfacer las demandas de un tráfico de tiempo real (como es el caso de la voz) en términos de diferentes parámetros. Estos parámetros son el ancho de banda, las pérdidas, el retardo y el *jitter*.

### 5.1 Concepto de calidad de la voz.

El concepto de calidad de servicio o QoS (*Quality of Service*) es demasiado amplio y, por ello, su interpretación depende del contexto concreto en que se emplea el término. La calidad desde el punto de vista de la red, o lo que es lo mismo, los mecanismos que proporcionan las redes de telecomunicaciones para garantizar a un flujo de información unos valores concretos de ancho de banda, retardo, *jitter* y pérdidas.

Las consideraciones generales sobre la evaluación de la calidad del servicio telefónico se encuentran en la recomendación E.420 de la ITU-T. Esta recomendación subraya los aspectos que mayor influencia ejercen sobre la percepción de la calidad del servicio de telefonía por parte de los usuarios. Algunos son tan curiosos como la posibilidad de comunicarse con usuarios que se encuentren en otros países o que la facturación esté libre de errores. Otros aspectos importantes son:

**Tasa de conectividad:** Hace referencia a la probabilidad con la que la red dispondrá de los recursos para cursar un intento de llamada.

**Inteligibilidad de la voz:** Un requisito, previo a todos los demás, es que cada extremo sea capaz de entender claramente las palabras de su interlocutor. En este sentido, juega un papel fundamental la claridad de la voz.

La claridad de la voz es un parámetro subjetivo que puede definirse como la fidelidad con que la voz es percibida por el extremo remoto e indica cuanta información puede extraerse de las palabras del otro extremo. Depende de la distorsión introducida por los componentes de la red. Sin embargo, es independiente del retardo (aunque el *jitter* sí ejerce influencia) y del eco (puesto que éste es escuchado por el emisor y la claridad se evalúa en el receptor).

**Codificación de la voz:** Una vez que la llamada ha sido establecida y que la voz del otro extremo puede entenderse con claridad el siguiente paso es codificar la voz, transmitirla a través de la red y ver qué tal se escucha. El resultado será una medida de la magnitud del esquema de codificación empleado.

La calidad de la codificación y la inteligibilidad de la voz están relacionadas entre sí y ambas dependen de la tasa binaria y de la tasa de error.

Como puede verse, cuanto mayor es la tasa binaria, tanto más probable es obtener una buena calidad de la codificación. Por otra parte, el incremento de la tasa de error es mayor cuanto menor es la tasa binaria debido a la disminución en la información de redundancia por la compresión.

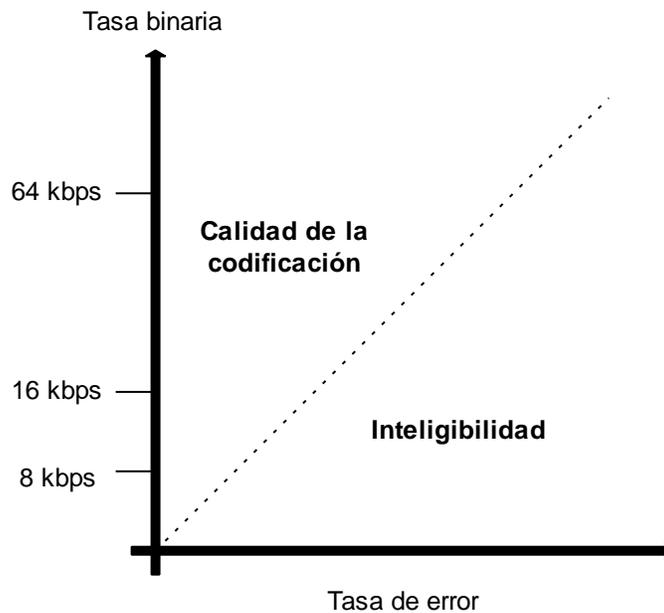


Figura 21. Relación entre la inteligibilidad y la calidad de la codificación de la voz

## 5.2 Factores que influyen en la calidad.

Para los usuarios de las redes de voz sobre paquetes las diferencias tecnológicas existentes entre las redes de conmutación de circuito y paquetes deben ser totalmente transparentes. Es decir, que de alguna manera hay que conseguir que las redes de conmutación de paquetes ofrezcan una calidad del servicio telefónico similar a la de las redes de conmutación de circuitos sin perder sus características propias. En general, los factores que determinan esta calidad son, por orden de importancia, la disponibilidad, el *jitter*, las pérdidas, el retardo y el ancho de banda.

### 5.2.1 Disponibilidad

La disponibilidad de un sistema es una medida de la probabilidad con que se encontrará en condiciones de funcionamiento, de manera que cuanto mayor es la disponibilidad mayor es dicha probabilidad.

Las redes telefónicas convencionales como PSTN o la RDSI presentan valores de disponibilidad muy elevados, en torno al 99.999%, es decir, 5.25 minutos fuera de servicio al año. Sin embargo, las redes de datos no alcanzan estos niveles de disponibilidad. Al contrario, sufren de pérdidas de servicio periódicas causadas, por ejemplo, por errores de usuario o del software.

A la hora de diseñar cualquier sistema debemos analizar, en primer lugar, el tiempo que el sistema puede dejar de estar operativo debido a fallos inesperados en el hardware y el coste que ello supone para contrastarlo con el coste de la inversión necesaria para prevenir dichos fallos. Para la VoIP los componentes críticos son los servidores, los *gateways* y las terminales de usuarios.

La clave de la tolerancia a fallos es la redundancia, cuyo principio es simple: cualquier parte del sistema que resulte crítica para su funcionamiento debe estar replicada, de tal modo, que el sistema de reserva reemplace al principal en caso de fallo de este último. Replicar absolutamente todos y cada uno de los componentes de la red no tiene sentido y, además, no es viable económicamente, por lo que se opta por duplicar única y exclusivamente aquellos que realmente son críticos para el funcionamiento de la red, generalmente, los servidores encargados del control de llamadas, la señalización y los *gateways*.

Por otra parte, se suele configurar las terminales de usuario para que si ocurriera que la llamada no puede cursarse por carecer de recursos o por estar la red fuera de servicio, dicha llamada se enrute por RTPC o la RDSI. Esto se le conoce como enrutamiento de *backup*.

Además, es recomendable utilizar sistemas de alimentación interrumpida (SAI) que reduce el impacto de los cortes de suministro eléctrico. Estos sistemas, si son de pequeña potencia, constan de una batería que

acumula energía y de un ondulator que transforma la corriente continua en alterna a 220 V/50 Hz para alimentar al equipo durante un cierto tiempo que dé lugar a poder cerrar aplicaciones y evitar la pérdida de datos. En caso de sistemas de gran potencia se tienen generadores con motores de gasolina, con lo que el sistema puede presentar una gran autonomía. En ambos casos, entran en funcionamiento ante los cortes de energía y también realizan un filtrado/estabilizado de la corriente, para eliminar los picos de tensión y los microcortes.

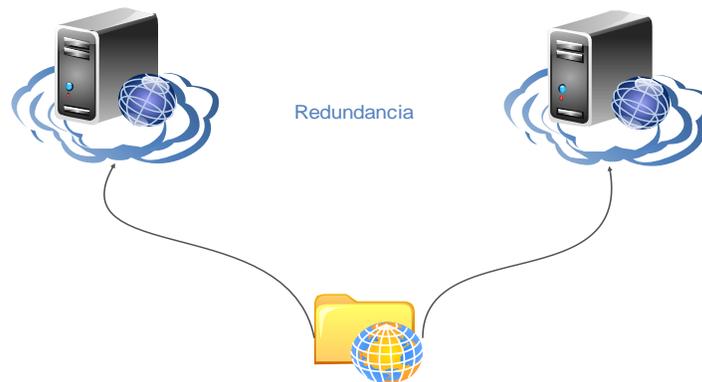


Figura 22. Disponibilidad.

### 5.2.2 Jitter

En general, a la hora de analizar las prestaciones de una red se habla del retardo en valores medios. Sin embargo, el tráfico de voz es muy sensible a las variaciones del retardo y, por ello, trabajar con valores medios no resulta suficiente.

En las redes IP, y en general en cualquier red de paquetes, no es posible garantizar que todos los paquetes de una misma comunicación sigan el mismo camino, al contrario de lo que ocurre en las redes de conmutación por circuitos. Como consecuencia, cada paquete llegará al destino atravesando un número distinto de nodos de la red y, por tanto, alcanzarán su objetivo con un retardo diferente. Esta variabilidad del retardo recibe el nombre de *Jitter*.

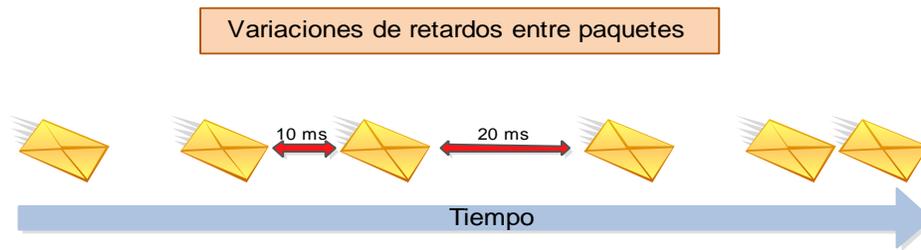


Figura 23. Jitter.

Los paquetes se generan en el origen con una cadena fija, por ejemplo, un paquete de voz cada 20 ms. Sin embargo, al llegar al destino esta cadena es variable debido a las diferencias en los retardos de encolado y propagación fundamentalmente.

Para absorber estas variaciones se utilizan los llamados *buffers* de supresión de *jitter*. La supresión consiste en el almacenamiento de los paquetes durante el tiempo suficiente para que los paquetes que han llegado fuera de secuencia puedan reordenarse y reproducirse en el orden correcto. Por tanto, cuanto mayor es el *jitter* de los paquetes, tanto mayor es el tamaño del *buffer* de supresión de *jitter* necesario para reducir su impacto en la calidad.

La supresión del *jitter* introduce un retardo adicional que puede afectar a la calidad de la voz resultante. Por ello, es necesario encontrar una solución de compromiso entre el tamaño del *buffer*, el retardo y las pérdidas. La situación ideal es aquella en la que el tamaño del *buffer* varía dinámicamente con las condiciones de la red durante su funcionamiento.

### 5.2.3 Pérdidas

Las pérdidas de paquetes son el resultado del descarte de paquetes que se producen en los nodos de la red como consecuencia de la congestión de

dichos nodos. Puesto que en las redes de conmutación de paquetes no se produce una reserva de recursos previa al envío de la información de usuario, las pérdidas son inevitables. El efecto de las pérdidas es una disminución de la calidad de la voz, puesto que faltan paquetes a la hora de reconstruir la señal vocal. Esta disminución de la calidad es tanto mayor cuanto mayor sea la tasa de compresión del códec.

La solución más inmediata al problema de las pérdidas es la mejora de la arquitectura de la red. En efecto, puesto que las pérdidas son, básicamente, una cuestión de capacidad, si se sustituyen las líneas y los *routers* por otros de mayor capacidad el problema queda aparentemente resuelto. Sin embargo, esta solución no es definitiva puesto que en cuanto aumente ligeramente el tráfico de la red, los efectos nocivos de las pérdidas volverán a aparecer.

La alternativa que puede parecer más obvia es solicitar la retransmisión de los paquetes perdidos. Sin embargo, esto introduciría un retardo adicional que todavía empeoraría más la calidad de la voz. Son necesarias, por tanto, otro tipo de técnicas que atenúen los efectos de las pérdidas. A este fin, se han desarrollado tres grupos de medidas:

- **Corrección de errores (FEC, *Foward Error Control*):** en este tipo de técnicas, junto con los paquetes, se incluyen información de redundancia que permiten recuperar el valor del paquete perdido a partir del valor de los paquetes perdidos. Su principal inconveniente es el retardo puesto que para decodificar un paquete son necesarios paquetes vecinos.
- **Distribución de errores:** consiste en aleatorizar las pérdidas para dispersar sus efectos. De nuevo, el inconveniente es el retardo adicional que introducen y que consumen un mayor ancho de banda.
- **Recuperación de errores (*Packet Loss Concealment*):** sustituyen el paquete perdido por otro. Esta situación puede ser tan simple como emplear un paquete perdido, un silencio o un ruido blanco, o tan compleja como el resultado de una técnica de predicción a partir de paquetes anteriores y posteriores. En este sentido, conviene tener en

cuenta que a mayor complejidad, mayor coste de procesamiento y mayor retardo introducido.



Figura 24. Pérdidas de paquetes

#### 5.2.4 Retardo

El retardo o latencia es el tiempo invertido por la señal de voz en su viaje desde el origen al destino. Una de las características más importantes de la voz es su temporalidad, no sólo porque el intervalo de pronunciación de dos sílabas determina su pertenencia a una misma palabra, sino porque la conversación entre dos interlocutores sigue un esquema temporal de escucha-respuesta cuya alteración puede convertir la conversación en ininteligible.

Por otra parte, uno de los problemas de las redes telefónicas es el eco, consecuencia de las reflexiones que sufre la señal en el otro extremo. Las redes telefónicas convencionales se diseñan para que el retardo no supere los 50 ms y, en estas circunstancias, el eco es enmascarado por la voz de los interlocutores.

Al estudiar la influencia del retardo en las comunicaciones de voz sobre paquetes son dos los aspectos que se determinan: el retardo máximo aceptable y las fuentes de retardo.

El retardo máximo aceptable marca un umbral por encima del cual la calidad de la voz resultante es inaceptable y la conversación resulta imposible. La recomendación G.144 de la ITU-T establece este umbral entorno a los 150 ms o 200 ms. Sin embargo, la influencia del retardo depende tanto de factores

objetivos como de factores subjetivos, por ello, en algunas ocasiones como es el caso de las comunicaciones vía satélite en que los usuarios se encuentran predispuestos a tolerar una menor calidad, podrían llegar a soportar retardos de unos 400 ms.

Rango (ms)	Descripción
0-150	<b>Excelente.</b> Muy válido para las aplicaciones más comunes
150-400	<b>Bueno-Pobre.</b> Aceptable, teniendo en cuenta que un administrador de red conozca las necesidades del usuario.
Sobre 400	<b>Inaceptable</b> para la mayoría de las aplicaciones de red; sin embargo, este límite puede ser excedido en algunos casos aislados.

Tabla 5. Descripción de la relación entre el retardo y calidad de la voz



Figura 25. Relación entre el retardo en un sólo sentido y la calidad de la voz

Una vez que ya se dispone de un objetivo de diseño en cuanto al retardo se refiere, el paso siguiente es estudiar las distintas fuentes de retardo con el fin de optimizar su comportamiento. Para ello, conviene analizar todo el proceso que sufre la señal de voz desde que es emitida por un extremo hasta que llega a su receptor.

En primer lugar, la voz del usuario debe digitalizarse ya que su formato natural es analógico y para su transmisión por la red de paquetes debe tener un formato digital. Junto con la digitalización, algunos códec realizan, además una compresión que reduce el ancho de banda consumido por la comunicación vocal. El paso siguiente es empaquetar las muestras de voz antes de su

transmisión por la red. El retardo introducido en todos estos procesos depende del códec.

La siguiente tabla muestra los parámetros de algunos de los codecs más empleados en VoIP, aunque algunos pueden variar ligeramente, dependiendo del fabricante.

	<b>G.711</b>	<b>G.729</b>	<b>G.723.1</b>
<b>Tasa binaria (kbps)</b>	64	8	6.3/5.3
<b>Complejidad (MIPS)</b>	0.1	22	16/18
<b>Retardo codificador (ms)</b>	0.125	15	37.5
<b>Tiempo entre paquetes (ms)</b>	20	20	30
<b>Retardo de empaquetamiento (ms)</b>	1.5	15	37.5
<b>Tamaño del <i>buffer</i> de supresión de <i>jitter</i> (ms)</b>	40	40	60
<b>Calidad (MOS)</b>	4.4	4.1	3.5-3.9

Tabla 6. Características de algunos codecs

Una vez que los paquetes llegan al Gateway, este invertirá un cierto tiempo en transmitirlos por una determinada línea. Este tiempo es lo que se conoce como retardo de serialización y depende de la velocidad de la línea y del tamaño de la trama. El retardo de serialización debe contabilizarse cada vez que el paquete atraviese un dispositivo *store-and-forward* como un *router* o un *switch*.

Los paquetes serializados viajarán por la red hasta llegar al destino. El tiempo invertido en este viaje deriva, fundamentalmente, de dos contribuciones, una fija y otra variable. La componente fija se corresponde con el retardo de propagación, que es el tiempo que tarda la señal en alcanzar su destino. Depende de las características del medio físico de transmisión y de la velocidad de la luz, por lo que suele ser muy pequeño (la recomendación G.114 aconseja un valor de  $6 \mu s/km$ ). Por otro lado, los paquetes son encolados en los nodos de la red un tiempo variable que depende de la carga de la misma y de la capacidad de dichos nodos. Puesto que el número de paquetes en espera

en la cola de transmisión depende de la caracterización estadística del tipo de tráfico al que pertenezcan dichos paquetes, el retardo de encolado varía mucho de un paquete a otro. En cualquier caso, generalmente el retardo de la red se encuentra comprendido entre 70 ms y 100 ms y es, por tanto, una de las contribuciones más importantes al retardo total.

### 5.2.5 Ancho de banda

El ancho de banda ( $BW$ , *BandWidth*) de una red puede definirse como la cantidad máxima de información que la red es capaz de transportar (por unidad de tiempo). El primer requisito que debe cumplir una red de voz sobre paquetes para ofrecer la calidad adecuada es disponer del ancho de banda suficiente para cursar las comunicaciones de voz. El ancho de banda medio de la red deber ser tal que:

$$BW = \frac{BW_{voz} + BW_{video} + BW_{datos}}{0.75}$$

De esta manera, nos dejamos un 25 % de margen para hacer frente a posibles picos de tráfico. Al dimensionar la red según este criterio, es decir, al garantizar que habrá ancho de banda suficiente para cursar las comunicaciones, se reduce la probabilidad de que el retardo, el *jitter* o las pérdidas tengan un impacto considerable. Sin embargo, esto no quiere decir, ni mucho menos, que sobredimensionando la red en cuanto ancho de banda se refiere se resuelvan definitivamente los problemas de calidad. En todo caso, deberá hacerse un estudio de cada uno de los factores para tomar las medidas adecuadas.

Una comunicación de voz sin comprimir, por cada sentido de la comunicación, consume 64 kbps. Sin embargo, puesto que el ancho de banda es un recurso escaso, sería deseable poder comprimir la voz y aumentar la eficiencia de utilización del ancho de banda cursando un mayor número de comunicaciones. Sin embargo, el precio a pagar es una disminución de la calidad. Aunque no existe ninguna relación directa entre el ancho de banda y

calidad de la voz, en general, cuanto mayor es la compresión, menor es la calidad de la voz ya que la señal es más sensible a las pérdidas y al retardo.

#### 5.2.5.1 Supresión de silencios

Es un mecanismo complementario al empleo de codecs compresores para reducir el ancho de banda. Se trata de aprovechar el hecho de que en una conversación normal aproximadamente el 60 % del tiempo lo ocupan los silencios debidos a las pausas para respirar y a la espera del turno en la comunicación. La idea es utilizar estos instantes en el que el canal está libre, para introducir tráfico de otras conversaciones. De esta manera, se pueden obtener reducciones de hasta 60 % en el flujo de paquetes. Estas técnicas reciben el nombre de detección de actividad, supresión de silencios o VAD (*Voice Activity Detection*).

La detección de actividad puede activarse en varios componentes de la red. Supongamos una red con una arquitectura *toll by pass*, una de las configuraciones más comunes. En este caso, existen seis puntos en los que podría activarse el VAD.

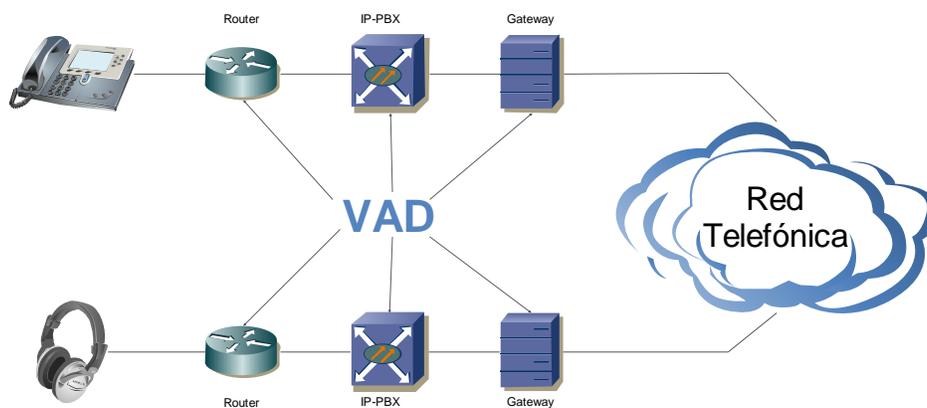


Figura 26. Puntos donde se puede activar el VAD

Para evitar que el interlocutor piense que se ha cortado la comunicación durante los intervalos de silencio se envían periódicamente paquetes de

silencio (SID, *Silence Insertion Description*) durante la pausa. Estos paquetes proporcionan una indicación del nivel del ruido que existe en el origen para que el receptor lo simule en la terminal remota mediante un generador de ruido.

Sin embargo, pese a que la supresión de silencios aporta beneficios en cuanto al ancho de banda empleado, también suele ser la responsable de un fenómeno conocido como *clipping* y que consiste en que la voz del interlocutor parece “recortada”. También las pérdidas, la latencia y el *jitter* pueden producir este fenómeno.

#### 5.2.6 Correlación entre paquetes

A partir de los años 80`s se han venido desarrollando modelos de tráfico que tienen en cuenta la correlación existente entre llegadas de paquetes a corto plazo. Es decir, al calcular la función de autocovarianza de las series de tiempo correspondientes a ciertas métricas de desempeño, se encontró que la velocidad a la que esta decae en el tiempo es exponencial. Sin embargo, mediciones y estudios recientes han demostrado que la función de autocorrelación en muchas trazas de tráfico real, decae más lentamente que una exponencial con el tiempo, es decir, de acuerdo a una ley de potencia.

En el clásico estudio presentado por Leland, Willinger, Taqqu y Wilson se mostró, tras la observación de exhaustivas mediciones realizadas sobre una red Ethernet en Bellcore, que el tráfico era de naturaleza “autosimilar” (self-similar). Dichos estudios muestran que

La autosimilitud describe el fenómeno en el que ciertas propiedades de un objeto o serie de tiempo se preservan sin importar el escalamiento en el tiempo o en el espacio. Una de las características atractivas de la este fenómeno es que el grado de autosimilitud puede ser cuantificado mediante un solo parámetro llamado, parámetro de Hurst ( $0 < H < 1$ ). Los estudios mencionados anteriormente, muestran que si el valor del parámetro H de cierta serie de tiempo que represente alguna métrica de desempeño está cerca de “1”, la calidad de servicio será baja.

**Autosimilitud de una serie agregada:** se dice que  $X$  es auto-similar si satisface la siguiente ecuación:

$$X_k^{(m)d} = m^{H-1} X$$

Dónde, cada término  $X_k^{(m)}$  se define como:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i; k = 1, 2, 3, \dots$$

dónde,  $m$  representa el nivel de agregación; es decir, cada nueva serie es obtenida partiendo la original en bloques disjuntos de tamaño  $m$  y promediando cada bloque para obtener los  $k$  valores de la nueva serie.

**Autosimilitud de Segundo Orden:** se dice que  $X$  es auto-similar de segundo orden con parámetro de Hurst ( $H$ ), si satisface la siguiente ecuación:

$$\text{var} \left( X_k^{(m)} \right) = \text{var}(X) \cdot m^{2H-2} = \frac{\text{var}(X)}{m^\beta}$$

Donde  $H = 1 - \frac{\beta}{2}$ .

Existen muchos métodos para evaluar el parámetro  $H$  en una serie de tiempo, sin embargo, uno de los más utilizados es el método de la varianza, el cual está basado en la definición de la autosimilitud de segundo orden.

Si aplicamos logaritmos a ambos miembros de la expresión anterior, obtenemos la siguiente ecuación:

$$\log \text{var} \left( X_k^{(m)} \right) = \log \text{var}(X) - \beta \log m$$

Si se grafican en un par de ejes los distintos puntos que surgen de ir escalando por  $m$ , la variable  $X$ , donde las abscisas están representadas por  $\log m$ , y las ordenadas por  $\log \text{var} \left( X_k^{(m)} \right)$ , se puede observar que los puntos se agrupan

en torno a una línea recta de pendiente  $-\beta$  y cuya ordenada al origen es la constante  $\log \text{var}(X)$ . Entonces si se realiza una regresión a la colección de puntos y se calcula el valor de  $\beta$ , se puede estimar el valor de  $H$  de la siguiente forma:

$$H = 1 - \frac{\beta}{2}$$

### 5.3 Medida de la calidad de la voz

Como veremos, la variedad de métodos y técnicas existentes que tiene como objetivo determinar la calidad de la voz es enorme. En general, todas ellas se suelen clasificar según dos criterios: su grado de intrusismo en la red y su objetividad.

El grado de intrusismo en la red hace referencia al modo en que el proceso de medida interacciona con dicha red. Según este criterio, podemos encontrar sistemas de medida intrusivos o activos y sistemas no intrusivos o pasivos.

Los sistemas intrusivos consisten en el envío de una señal conocida a través de la red (llamada prueba) y en la comparación entre la señal recibida y la transmitida para estudiar la degradación que introduce la red en la señal transmitida. Debido a la complejidad y al elevado coste de este tipo de técnicas, no permiten su utilización para medidas en tiempo real aunque son ideales para la medida de las prestaciones de un códec en el laboratorio. Además, durante el periodo de pruebas la red no transporta tráfico de usuario. Pertenecen a este grupo los estándares PSQM, PESQ y PAMS, entre otros.

Los sistemas no intrusivos o pasivos, por su parte, efectúan medidas de tiempo real mientras el sistema sigue en explotación sin interferir en las llamadas existentes, y sin necesidad de señal de referencia. El inconveniente es que, por lo general, son menos exactas que las anteriores. Una variante muy útil de los sistemas de medición pasivos son los agentes embebidos que, por

ejemplo, pueden incorporarse directamente en el Gateway o en el propio teléfono IP. Medidas pasivas son las escalas MOS, el modelo E y VQMon.

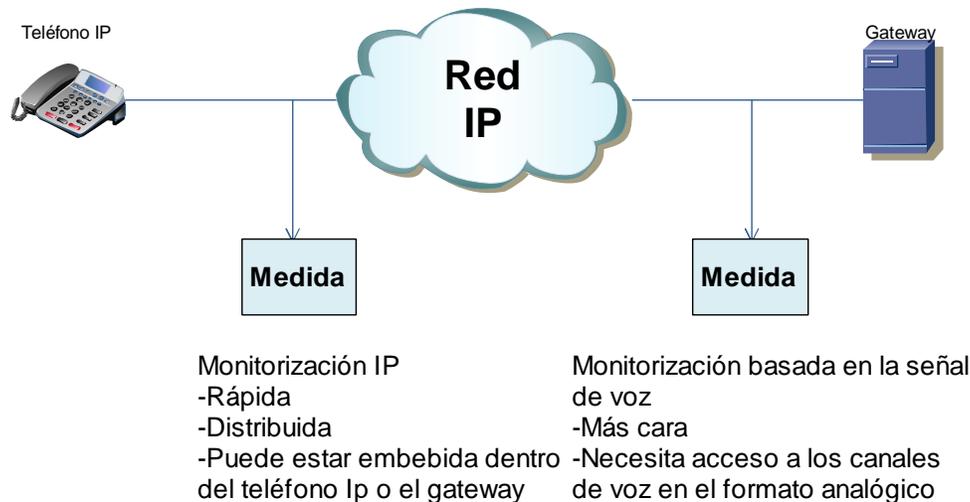
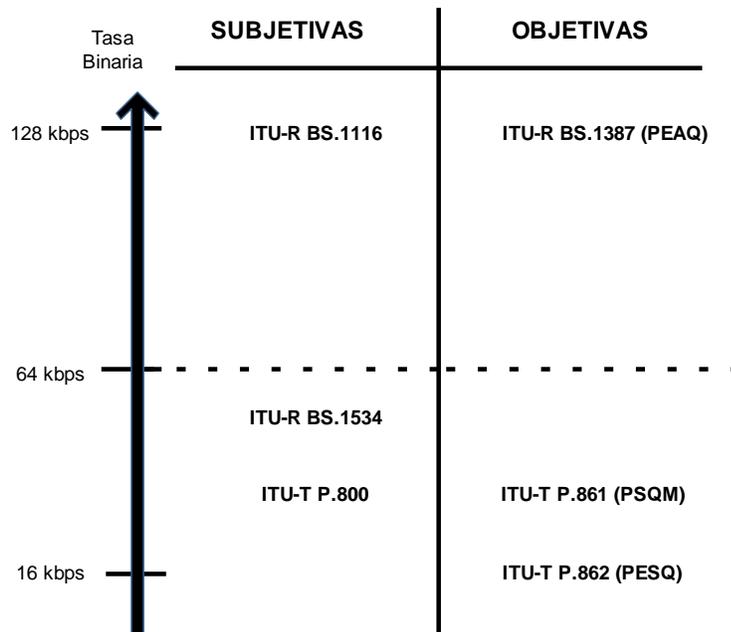


Figura 27. Medidas no pasivas

En cuanto a la objetividad de las medidas hace referencia cuán independiente de la opinión de los sujetos es la calidad de la voz medida en un determinado conjunto de pruebas. Las primeras que aparecieron fueron las medidas subjetivas consistentes simplemente en evaluar la calidad de voz en base a las opiniones de un grupo suficientemente extenso de usuarios como para que los resultados sean estadísticamente relevantes. Sin embargo, se necesitaba una aproximación más ingenieril al problema que, por otra parte, permitiera evaluar la calidad de la voz en el laboratorio. La solución son los modelos perceptuales de la voz que hacen posible la definición de modelos objetivos de evaluación de la calidad de la voz.

La siguiente figura muestra una comparación entre las medidas de la calidad de audio estandarizadas más extendidas. En las redes de telecomunicaciones en general y en las redes de voz sobre paquetes en particular, interesa utilizar un códec de reducida tasa binaria ya que el ancho de banda es un recurso escaso.



Figura

entre estándares de la calidad de la voz

28. Comparación

### 5.3.1 ITU-T P.800 (Escala MOS)

Los mecanismos tradicionalmente empleados para la evaluación de la calidad de la voz en las redes de telefonía fueron estandarizados por la ITU-T. Uno de ellos es la recomendación ITU-T P.800 que describe las escalas MOS y que se han venido utilizando para los codecs de voz desde 1993. Se trata de un conjunto de técnicas subjetivas de la medida de la calidad de la voz que reciben el nombre de test ACR (*Absolute Category Rating*) y que tienen el mismo esquema general: se reúne a una muestra de usuarios a los que se pide que opinen sobre la calidad que en algún concreto ofrece un determinado sistema de transmisión de la voz.

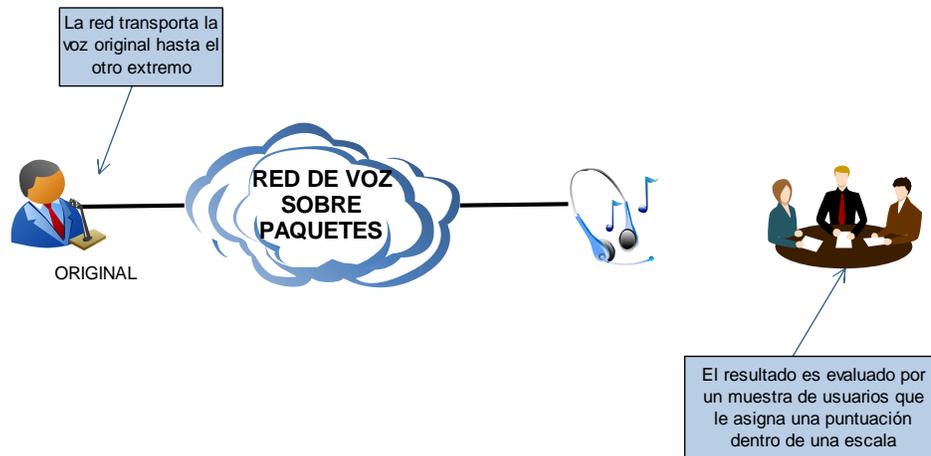


Figura 29.

Evaluación MOS

Los test ACR, a diferencia de lo que ocurre en el modelado perceptual, no utilizan la comparación con una señal de referencia. En realidad, esto es precisamente lo que ocurre en una conversación telefónica, en donde los extremos no disponen de la voz original del otro extremo. Sin embargo, es necesario ajustar las opiniones de los distintos usuarios a una escala absoluta y por ello, previamente, se presenta a los usuarios unos ejemplos predefinidos que les proporciona una base a la hora de evaluar la calidad de la voz. Estos ejemplos están regidos en la recomendación P.810 (MNRU, *Modulated Noise Reference Unit*).

Los aspectos que evalúan las escalas MOS son muy variados, sin embargo, los más comunes son, quizá, el de la calidad de la voz y el que evalúa el esfuerzo requerido para entender el significado del mensaje pronunciado por el otro extremo.

Puntuación	Calidad
5	Excelente
4	Buena
3	Aceptable
2	Pobre
1	Mala

Tabla 7. Escala MOS utilizada para medir la calidad de la voz

Puntuación	Esfuerzo
5	Relajación completa: no es necesario ningún esfuerzo
4	Necesario prestar atención: no se requiere esfuerzo apreciable
3	Esfuerzo moderado
2	Esfuerzo considerable
1	Imposible de entender

Tabla 8. Escala MOS para la medida del esfuerzo de interpretación del mensaje

La siguiente tabla muestra la puntuación en escala MOS de algunos de los estándares de codificación más habituales en voz sobre paquetes:

Estándar de codificación	Valor MOS (velocidad)
G.711	4.4 (64 kbps)
G.726	3.8 (32 kbps)
G.728	3.6 (16 kbps)
G.729	3.7 (8 kbps)
G.723.1	3.9 (6.3 kbps)
	3.6 (5.3 kbps)

Tabla 9. Escala MOS de calidad para los codecs más comunes

El principal inconveniente de las escalas MOS deriva de su propia concepción. En efecto, para que sean realmente significativas es imprescindible que la muestra de usuarios sea suficientemente grande y ello encarece el proceso de elaboración. Además, al ser subjetivas dependen de gran cantidad de factores como la predisposición y la actitud de los usuarios frente a la prueba o el nivel cultural de los individuos. Sin embargo, tienen la ventaja de que los patrones de comportamiento humanos son sobradamente conocidos y, en algunos casos, resulta sencillo determinar los efectos de las pérdidas y del retardo en una muestra de usuarios relativamente reducida.

### 5.3.2 Modelado perceptual de la voz

Las técnicas perceptuales o psicoacústicas se basan en el modelado de la respuesta del oído humano y la introducción a dicho del modelo de una señal de referencia y de la señal de entrada cuya calidad desea evaluarse y que proporciona una representación paramétrica de ambas señales. Seguidamente, se someten a un proceso de comparación del que se obtiene una estima de la diferencia audible. Ésta debe ser procesada de la misma manera en que lo haría el oído humano en un test con individuos y para ello se utiliza un modelo cognitivo que finalmente dará como resultado la calidad de la voz.

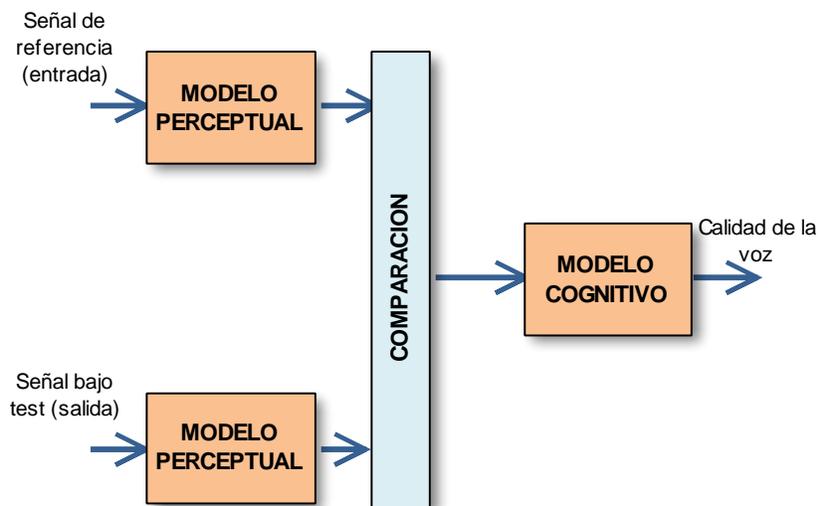


Figura 30. Modelado perceptual de la voz

#### Psioacústica

El comportamiento del sistema auditivo humano es altamente no lineal. En efecto, únicamente somos sensibles a aquellos sonidos comprendidos dentro de un rango de amplitudes y frecuencias muy concreto, siendo esta

sensibilidad dependiente de la frecuencia, de manera que se define un área de audición.

Por otra parte, el espectro audible se divide en zonas denominadas bandas críticas o bark, hasta un total de 25 bark. Cada uno de estos bark se caracteriza porque el nivel de sonoridad es constante en toda la banda de frecuencias que comprende. El ancho de cada banda crítica es aproximadamente constante e igual a 100 Hz por debajo de los 500 Hz y alrededor del 20 % de la frecuencia central, para valores por encima de los 500 Hz.

La percepción de un sonido es un fenómeno muy complejo y en el que se ven implicados gran cantidad de parámetros. El primero de ellos es el umbral de audición, que se define como el nivel de presión sonora mínimo para producir una sensación en el oído humano. Por debajo de este valor, el sonido pasará desapercibido.

El umbral de audición no es constante, ni siquiera para el mismo sujeto, sino que varía en función de la frecuencia y de la presencia o ausencia de otros sonidos. Esta posibilidad de que un sonido quede oculto por la ocurrencia de otros se conoce con el nombre de enmascaramiento y puede ser de tres tipos.

- **Enmascaramiento simultáneo:** las señales enmascarante y enmascarada ocurren el mismo instante de tiempo o con una diferencia tan pequeña que puede considerarse despreciable.
- **Preenmascaramiento:** el sonido enmascarado ocurre entre 5 y 20 ms antes que el enmascarante.
- **Postenmascaramiento:** el sonido enmascarado ocurre entre 50 y 200 ms después que el enmascarante.

Este fenómeno se aprovecha en algunos codificadores para comprimir, todavía más la señal, evitando transmitir las señales enmascaradas puesto que no se oirán. Tal es el caso de la codificación de audio MPEG.

Además de los parámetros anteriores, existen algunos otros que también se emplean en la caracterización subjetiva de la percepción sonora, los más importantes son los siguientes:

***Roughness y fluctuation strength:*** miden el índice de modulación de señal en amplitud o frecuencia.

***Sharpness:*** indica el contenido en alta frecuencia del sonido (es decir, su color), de tal manera que cuanto mayor es el *sharpness*, tanto mayor es aquél.

***Tonality:*** representa la riqueza tonal del sonido y suele ser proporcional a la percepción sonora.

#### *ITU-T P.861 (PSQM)*

En 1996 el grupo de estudio SG12 de la ITU acabó la recomendación P.861 para el análisis objetivo de los codecs de voz basado en un algoritmo denominado PSQM (*Perceptual Speech Quality Measure*). PSQM es una versión de otro algoritmo con carácter más general, el PAQM (*Perceptual Audio Quality Measure*), utilizado en señales telefónicas. A diferencia de este último, PSQM considera los efectos psicoacústicos en la percepción de la calidad del audio.

El cálculo de la calidad de la voz según el algoritmo empleado por PSQM, es la siguiente. En primer lugar, se convierte la representación temporal de las señales  $x$ ,  $y$  al dominio de la frecuencia dividiéndolas en bloques y obteniendo su FFT (*Fast Fourier Transform*). Seguidamente, las muestras en frecuencia se inventanan y la escala frecuencial se transforma a barks o bandas críticas, proceso que se conoce con el nombre de *frequency warping*. Una vez hecho esto, tanto la señal bajo test como la señal de referencia se filtran con un modelo del dispositivo receptor y se añade al resultado un ruido Hoth que simula el ambiente de oficina convencional. Restando las dos representaciones de la señal, se dispone de una estima del error audible en función del timbre y del tiempo.

Este algoritmo se utiliza para codecs con tasas binarias comprendidas entre los 8 Kbps y los 16 Kbps. Sin embargo, las medidas PSQM fueron diseñadas para analizar sólo los efectos de la compresión/descompresión

llevadas a cabo por el códec y no tiene en cuenta las degradaciones causadas por la red como consecuencia de las pérdidas o el *jitter* de paquetes.

#### ITU-T P.862 (PESQ)

Cuando se estandarizó PSQM como la recomendación P.861 los esfuerzos se centraban, sobre todo, en los codecs de voz empleados en comunicaciones móviles como GSM y su aplicación a VoIP se consideraba algo lejana. Sin embargo, las redes de próxima generación, como son las de voz sobre paquetes, han cambiado drásticamente las necesidades y la ITU se ha visto obligada a revisar los estándares de medición de la calidad, dado que en este tipo de redes el efecto predominante es el retardo y no la distorsión, como ocurre en los codecs GSM. El resultado fue el algoritmo PESQ (*Perceptual Evaluation of speech Quality*) y que fue incluido en 2001 dentro de la recomendación P.862.

PESQ, en realidad, es una evolución de otro algoritmo, el PSQM+. Este último resolvía las distorsiones producidas por las ráfagas de error, pero todavía presentaba problemas a la hora de compensar las variaciones del retardo. Por otro lado, BT desarrolló un algoritmo, el PAMS, que manejaba la variabilidad del retardo a la perfección. PESQ combina las ventajas de ambos aunque tiene el inconveniente de que no está concebido para aplicaciones de *streaming*. Sin embargo, las medidas PESQ son directamente trasladables a las escalas MOS con muy pocas manipulaciones. En realidad, las escalas PESQ están comprendidas entre 1.0 y 4.5, ya que los usuarios, en general, son bastantes cautos a la hora de asignar el 5, es decir, calidad excelente, incluso cuando no existe degradación alguna en la señal.

Aunque PESQ fue diseñada para aumentar las prestaciones de PSQM incluyendo factores como la distorsión, el filtrado y otras desigualdades del canal, no los analiza todos.

### 5.3.3 Modelo E

El modelo E es una aproximación matemática a la medida de la calidad de la voz basada en la evaluación de las características de transmisión de las redes de voz sobre paquete y cuyo objetivo es predecir la calidad de la voz en función del retardo, las pérdidas y el tipo de códec.

El modelo E está especificado en la recomendación ITU-T G.107 y estipula que la calidad de la voz puede evaluarse a través del parámetro R, definido como:

$$R = R_0 - I_s - I_d - I_e + A$$

El término  $R_0$  hace referencia a la relación señal-a-ruido mientras que  $I_s$  modela la degradación que sufre la señal como consecuencia de su conversión a un formato adecuado para su transmisión por la red. Los otros tres términos son el efecto de las pérdidas ( $I_e$ ), del retardo ( $I_d$ ) y el margen de seguridad ( $A$ ). En cualquier caso, la recomendación proporciona una expresión mucho más simple si se considera en la anterior los valores por defecto y que es:

$$R = 94.2 - I_d - I_e$$

El retardo es uno de los factores más importantes a considerar cuando se estudia la calidad de la voz. Su impacto en el modelo E está representado por el parámetro  $I_d$  que en redes IP es función del retardo extremo a extremo:

$$I_d = 0.024 * d + 0.11 * (d - 117.3) * H(d - 117.3)$$

Donde  $H(x)$  es la función de Heavyside:

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Si se representa gráficamente esta relación, se concluye que en torno a los 175 ms un aumento del retardo supone una disminución drástica de la calidad de voz, algo que concuerda con la recomendación ITU-T G.114 en la que se aconseja que los valores del retardo en un solo sentido no superen valores comprendidos entre 150 ms y 200 ms.

Otro de los factores que influyen en la calidad de la voz son las pérdidas cuyo impacto se modela a través del parámetro  $I_e$  (recomendación ITU-T G.133). En general, cuanto mayor es el valor de  $I_e$  tanto mayor es la disminución de la calidad.

Por fin, la salida del modelo E es el factor R, a partir del cual puede obtenerse un valor en escala MOS de la calidad de la voz.



Figura 31. Evaluación de la calidad de la voz a partir del Modelo E

En el caso concreto de redes IP, el factor R y la escala MOS están relacionados a través de la siguiente expresión:

$$MOS = \begin{cases} R < 0 \\ 1 + 0.035 * R + 7 * R * (R - 60) * (100 - R) * 10^{-6} & 0 < R < 100 \\ R > 100 \end{cases}$$

#### 5.3.4 VQMon

De todas las medidas vistas hasta ahora, ninguna satisface adecuadamente las necesidades de proveedores de servicio de VoIP. Gran parte de ellas se basan en el comportamiento en media de la red obtenido tras mediciones a largo plazo de parámetros como el *jitter*, las pérdidas y el retardo que, por otro lado, son difíciles de relacionar con la calidad que perciben los usuarios en llamadas individuales. Son necesarios, pues, otros tipos de medidas que superen estos inconvenientes.

Uno de los más recientes es VQMon (*Voice Quality Monitoring*), desarrollado por la empresa Telchemy. No tiene en cuenta directamente los aspectos de la codificación de la señal, pero analiza la degradación que introduce la red (*jitter*, pérdidas y retardos) y predice el impacto en la señal de voz reconstruida. VQMon permite medir, en tiempo real, la calidad de la voz para todas las llamadas que estén cursando en un determinado instante.

Generalmente, las medidas de calidad suelen hacerse en los puntos finales de la comunicación de VoIP, es decir, en los gateways en el caso de una arquitectura *toll-by-pass* o en los propios terminales de usuario si la VoIP está implantada extremo a extremo. Esta información resulta extremadamente importante, no sólo porque la calidad determina la percepción que los usuarios tienen del servicio que les ofrece la red, sino que también porque es un dato necesario a la hora de llevar a cabo el correcto dimensionamiento del sistema.

Una de las características más interesantes de VQMon es que emplea un modelo estadístico para el análisis de las degradaciones que introduce la red que considera la voz tal y como el usuario la percibe, en lugar de emplear valores medios en la red entera o en un segmento de la misma. Por ejemplo, VQMon no mide únicamente la tasa de pérdidas sino también su distribución. Varios estudios han concluido que para una cierta tasa de pérdidas, la degradación de la calidad de la voz es mucho mayor si las pérdidas se producen a ráfagas. Además, VQMon también explota la memoria a corto plazo de los seres humanos que nos lleva a conceder mayor importancia a las degradaciones más recientes frente a las que ocurrieron antes [9].

# CAPÍTULO 6

## Capítulo 6. MEDICIÓN DE PARÁMETROS DE QOS EN UNA RED VOIP H.323

En el presente capítulo se presenta una metodología para la medición del tráfico en una red VoIP inalámbrica. Con las mediciones se realizó el análisis y caracterización de tráfico en la red. El primer paso fue generar tráfico VoIP mediante el establecimiento de un conjunto de llamadas de prueba con aplicaciones VoIP en hardware. Posteriormente el tráfico VoIP generado fue capturado, mediante el analizador de protocolos de red Wireshark, con la intención de obtener un conjunto de trazas de tráfico y parámetros específicos. El objetivo primordial de las mediciones fue obtener una colección de trazas de jitter y las pérdidas de paquete, para posteriormente analizarlos y determinar el impacto en el desempeño de la red.

Para obtener esta colección de trazas, las llamadas de prueba fueron realizadas con diferentes configuraciones como, tipo de códec, supresión de silencios y diversos anchos de banda.

### 6.1 Escenarios de medición

El escenario principal de medición está conformado por dos redes WLAN, conectadas en modo punto a punto mediante antenas MIMO.

En la imagen siguiente podemos observar la WANETCTIC: que es la red de campus CTIC-UQROO y la WANETDCS: que es la red de campus DIVISION DE CIENCIAS DE LA SALUD, las cuales se conectan a través de una conexión inalámbrica mediante antenas UBIQUITI NanoStation M5.

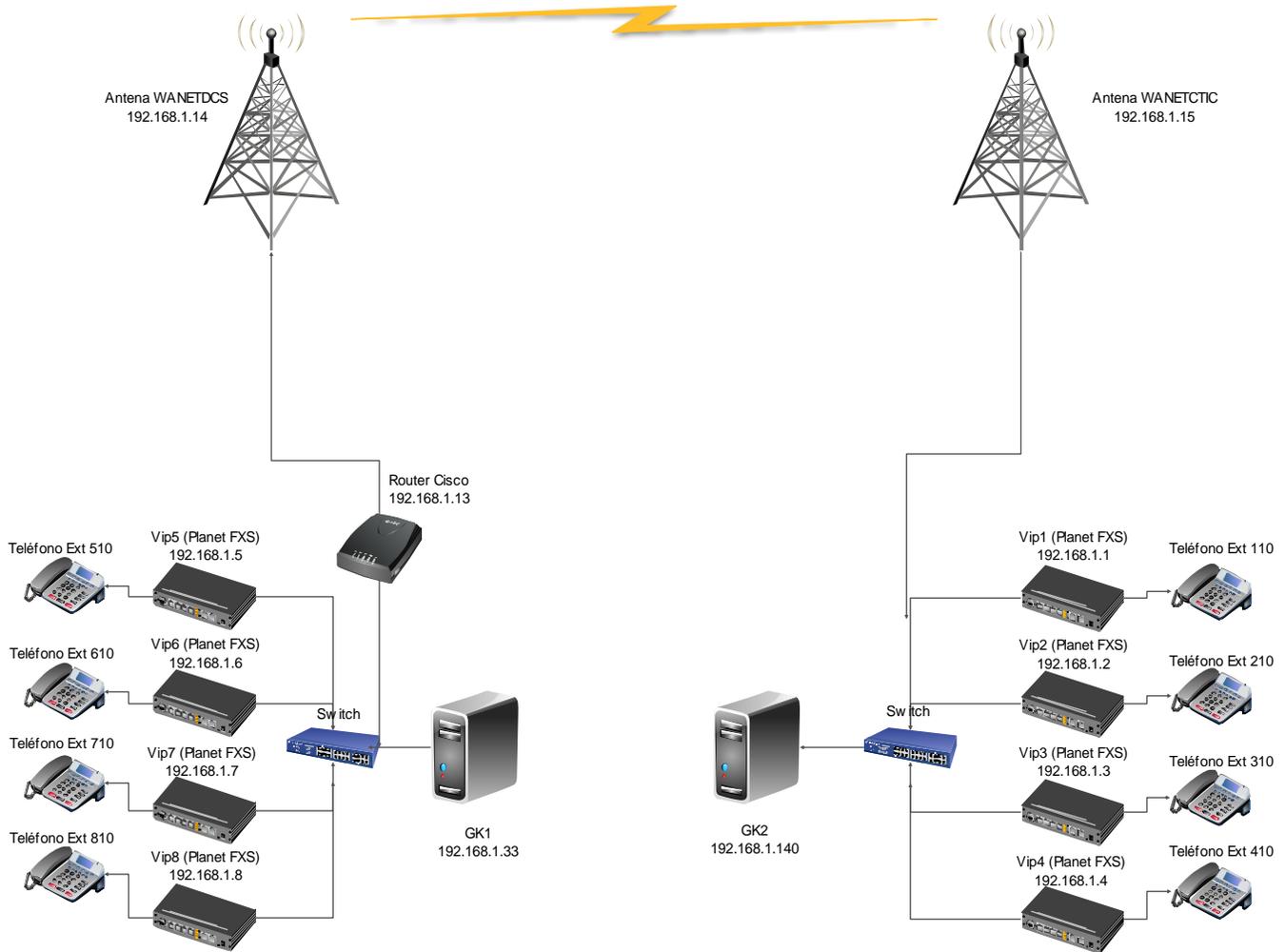


Figura 32. Escenario de red de medición H.323

Al igual se puede observar una arquitectura H.323 conformada por dos zonas H.323 interconectadas de forma inalámbrica en un enlace punto a punto. Cada zona contiene su propio Gatekeeper y sus propios dispositivos FXS y sus propias terminales, la zona WANETDCS está conectada a un router inalámbrico el cual a su vez está conectado a la antena MIMO que finalmente esta misma está conectada en la torre de comunicaciones de la División de Ciencias de la Salud, esta antena se comunica con la segunda antena MIMO la cual se encuentra en la torre de telecomunicaciones del Centro de Tecnologías

de la Información y la Comunicación esta zona H.323 la nombramos WANETCTIC.

La zona WANETDCS está conformada por las terminales Vip 5, Vip 6, Vip 7, Vip8, administradas por su propio Gatekeeper "GK1". La zona WANETCTIC está conformada por las terminales Vip 1, Vip 2, Vip 3, Vip4, administradas por su propio Gatekeeper "GK2".

Ambos Gatekeepers (GK1 y GK2) se implementaron mediante software, haciendo uso del proyecto GNU Gatekeeper de código abierto que implementa las características y funciones de un Gatekeeper H.323.

#### 6.1.1 Descripción de los equipos

La infraestructura de red utilizada para el escenario H.323 está conformada por los siguientes dispositivos:

2 Computadoras (donde se implementó el software GNU Gatekeeper, GK1 y GK2)

8 Dispositivos Planet VIP-480FS FXS

8 Teléfonos convencionales

1 Router Inalámbrico

2 Switches

2 Antenas MIMO

#### **HP Compaq 6000 Pro small form Factor [15]**

Se utilizaron para instalar y configurar el software GNU Gatekeeper, para que funcionen como los dos gatekeeper (Gk1, Gk2) de la arquitectura H.323

## **Características**

- Procesador: Pentium(R) Dual-Core CPU E5700 @ 3.00GHz
- Memoria (RAM): 4.00 GB
- Gráficos: Intel(R) Q45/Q43 Express Chipset (Microsoft Corporation - WDDM 1.1)
- Disco duro principal: 512 Gb
- Windows XP Profesional
- Fabricante: Hewlett-Packard
- Modelo: HP Compaq 6000 Pro SFF PC
- Tipo de sistema: Sistema operativo de 32 bits
- Número de procesadores principales: 2

## **Planet VIP-480FS FXS [16]**

El PLANET VIP-480FS es un Gateway VoIP de cuatro puertos FXS que cumple con los estándares SIP y H.323. Es una solución completa para la integración de redes voz/datos a las redes telefónicas analógicas. No solo provee comunicaciones de alta calidad sino también ofrece capacidades de compartir internet de modo seguro y confiable.

El VIP-480FS es capaz de manejar llamadas tanto SIP como H.323, cuenta con un switch de cuatro puertos y función de ruteador NAT, con estas funcionalidades los usuarios pueden disfrutar de llamadas de voz de alta calidad y acceso seguro a internet sin interferir con sus actividades de rutina.

Con sus clientes PPPoE/DHCP/DDNS, hasta cuatro conexiones concurrentes pueden ser establecidas en cualquier lugar del mundo. El VIP-480FS cuenta con una interfaz de usuario poderosa pero amigable (web/telnet), que simplifica la implementación y monitoreo de la red VoIP.

## **Características**

- 4 Ptos FXS + 4 Ptos Ethernet + WAN
- Soporte PPPoE, NAT, QoS
- Cliente DDNS para aplicación con IP dinámica
- Servidor virtual (a través de DDNS)

- Modo de comunicación dúal H.323v4/SIP 2.0
- Codecs de voz: G.711, G.729 AB, G.723
- Detección activa de voz, detección DTMF, supresor de eco
- Detección de silencios y Modo FAX
- Buffer adaptable para diversas condiciones de jitter
- Display de estatus de canales de voz

### **Teléfonos convencionales [17]**

Este exclusivo teléfono ha sido diseñado con los más altos estándares de calidad y fabricado con los mejores materiales, que harán de este el complemento perfecto para tu hogar y espacios, gracias a su sistema y novedosas funcionalidades.

#### **Características**

- Innovador teléfono alámbrico
- Tiene rediscado
- Timbre prendido/apagado
- Tiempo flash disponibles (100/300/600ms)
- Selección de pulso/tono
- Diseñado para montar en la pared

### **Router Inalámbrico Belkin N600 DB [18]**

La exclusiva tecnología de antena MultiBeam de Belkin ofrece la máxima velocidad de transferencia y minimiza los puntos muertos para un streaming de vídeo optimizado desde varios dispositivos en prácticamente cualquier punto.

Contiene velocidad de doble banda más rápida para el streaming de vídeo y los juegos en línea: hasta 300 Mbps (2,4GHz) + 300 Mbps (5 GHz).\*

Estándares de cifrado WPS/WPA2 y seguridad preconfigurada para utilizar el router con toda seguridad. Con el sistema de seguridad WPS (WiFi Protected Setup), podrás conectar varios ordenadores y dispositivos a la red con sólo pulsar un botón.

Mejora las actividades en línea gracias a las aplicaciones integradas de Belkin: SELF-HEALING, VIDEO MOVER, PRINT ZONE y MEMORY SAFE.

## Características

- Transmisión de doble banda (Dual Band)
- Hasta 600Mbps (hasta 300 Mbps (2,4GHz) + 300 Mbps (5 GHz))
- Tecnología MultiBeam: Cobertura total para varios dispositivos
- 1 puerto USB: Para la impresión o el almacenamiento inalámbrico
- 4 puertos LAN Ethernet: Para conexiones de red cableadas rápidas y fiables
- Diseño de la antena: Interna e integrada
- Adaptadores/cables incluidos: Fuente de alimentación y cable de red
- Compliant Standard(s): IEEE 802.11b, IEEE 802.11g y IEEE 802.11n, Mac y Wi-Fi
- Ancho de banda: radio dual concurrente 2,4 GHz y 5 GHz ISM
- Compliant Standard(s): IEEE 802.11b, IEEE 802.11g y IEEE 802.11n
- Conmutador integrado: Conmutador de 4 puertos
- Tipo de conector(s): USB de 4 patillas tipo A (1)
- Tipo de Interfaz: RJ-45 (LAN)
- Interfaz: Ethernet 10Base-T/100Base-TX (4)
- Nº máx. de conexiones WLAN: 16
- Navegadores compatibles para gestión remota: Firefox® y Safari®
- Protocolo de conmutación: Ethernet
- Protocolo de enrutamiento: Enrutamiento de IP estática
- Protocolo de gestión remota: HTTP y HTTPS
- Protocolo(s) de enlace de datos: Ethernet, Fast Ethernet, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g y IEEE 802.11n
- Protocolos de ISP compatibles: Estática (IPoA), PPPoA, PPPoE
- Velocidad de transferencia de datos: 300 Mbps
- Método(s) de autenticación: Identificación de conjunto de servicios de radio (SSID)
- Seguridad por pulsador WPS: Si
- Tipo(s) de encriptación compatible(s): WPA de 256 bits, WEP de 64 bits y WPA2
- Tipo(s) de firewall compatible(s): NAT y SPI

## Switches [19]

El conmutador GREENnet de 24 puertos a 10/100 Mbps proporciona un mejor ancho de banda, facilidad de uso y fiabilidad, y a la vez reduce el consumo de energía de hasta un 40% \*. La tecnología GREENnet ajusta de manera automática el voltaje de energía según la necesidad, lo que resulta en ahorros sustanciales de energía. La tecnología GREENnet integrada funciona desde el primer momento y no requiere ningún tipo de administración.

Aumente la eficiencia de su oficina y reduzca la congestión de la red con una capacidad de conmutación total de 4.8 Gbps con modo Full Dúplex.

### Características

- 24 x 10/100 Mbps Auto-Negotiation, Auto-MDIX Ethernet ports
- La tecnología GREENnet reduce el consume de energía cuando:
  - El puerto Ethernet no está en uso
  - El dispositivo conectado está apagado
  - El dispositivo conectado está en modo de ahorro de energía
  - Según la longitud del cable Ethernet
- Compatible con los estándares IEEE 802.3 y IEEE 802.3
- Capacidad de forwarding (reenvío) 4.8 Gbps
- Arquitectura de conmutación de almacenamiento y reenvío con rendimiento de velocidad por cable sin bloqueos
- Compatible con control de flujo IEEE 802.3x
- Ofrece entradas de dirección MAC de 8K
- 1.25 Mbits RAM data buffer
- Revestimiento de metal resistente
- LEDs de diagnóstico
- Plug & Play installation
- Standard 19" (1U) rackmount design

## **Antena Ubiquiti NanoStation M5 [20]**

El NanoStationM está diseñado para implementarse en conexiones punto a punto, por ejemplo, conectividad inalámbrica a un sistema de vigilancia de vídeo, o uso como un CPE airMAX.

Cuenta con salida PoE (PowerOverEthernet) y un puerto Ethernet secundario para alimentar un dispositivo externo, como una cámara de seguridad PoE, para una integración perfecta de vídeo IP.

La versátil Tecnología firmware airOS® de Ubiquiti permite un alto rendimiento. AirOS ofrece funciones como la configuración inalámbrica, configuración de puente o enrutamiento y servicios de administración de sistemas.

La Tecnología airMAX® de Ubiquiti está probado en millones de implementaciones en todo el mundo, exhibiendo un rendimiento excepcional en ambientes al aire libre. El protocolo TDMA AirMax permite una escalabilidad sin precedentes, de alto rendimiento y baja latencia en las redes multipunto.

### **Características**

- 5GHz Hi Power 2x2 MIMO AirMax TDMA Station
- Más de 150Mbps
- Alcance de 15Km
- 2x2 MIMO
- Antena de 16dBi con 500mW de potencia
- Trabaja con 5GHz de Frecuencia
- SISTEMA
- Procesador: Atheros MIPS 24KC, 400MHz
- Memoria: 32MB SDRAM, 8MB Flash
- Interfaz de Red: 2 x 10/100 BASE-TX (Cat. 5, RJ-45) Interface Ethernet
- Conformidad ROHS: SI
- Tamaño de la caja: 29.4 cm x 8 cm x 3cm
- Peso: 0.4kg
- Características de la caja: Exterior, Plástico UV estabilizado

- Kit de Montaje: Kit de montaje en mástil incluido
- Máximo consumo de energía: 8 Watts
- Alimentación: UBIQUITI POE-2412W Fuente de alimentación PoE incluido
- Tipo de Alimentación: PoE pasivo (pares 4,5+; 7,8 retorno)
- Temperatura Operación: -30C a +80C
- Humedad Operación: 5 a 95% de Condensación
- Rango de Frecuencia: 4.9-5.9 GHz
- Ganancia: 14.6-16.1dBi
- Polarización: Lineal Dual
- Aislamiento de Polaridad: 22dB Mínimo
- Frecuencia: 5470MHz-5825MHz

### 6.1.2 Configuración de los equipos

#### **Antenas Ubiquiti**

*Antena Instalada en WANETCTIC*

IP: 192.168.1.15

Wireless Mode: Estación WDS

Conexión: Punto a Punto

SSID: ubntuqroo

Login:

User: ubnt

Pass: redes2012

*Antena Instalada en WANETDCS*

IP: 192.168.1.14

Wireless Mode: Cliente WDS

Conexión: Punto a Punto

SSID: ubntuqroo

Login:

User: ubnt

Pass: redes2012

### **Router Belkin**

*Router Instalado en WANETDCS*

IP: 192.168.1.13

Modo de operación: Access Point

SSID: belkin.0

### **HP Compaq 6000 Pro**

*Gatekeeper WANETDCS*

IP: 192.168.1.33

ID: GK2

*Gatekeeper WANETCTIC*

IP: 192.168.1.140

ID: GK1

## Planet VIP-480FS FXS

En la siguiente tabla se observa los Planet VIP de la zona H.323 WANETCTIC, se describe su dirección IP, su identificador de registro H.323 con su respectivo gatekeeper, el alias e IP del gatekeeper que controla la zona y las extensiones asignadas a los teléfonos.

IP	H.323 ID	Gatekeeper	Ext Teléfono
192.168.1.1	vip1	Gk2 (192.168.1.140)	110
192.168.1.2	vip2	Gk2 (192.168.1.140)	210
192.168.1.3	vip3	Gk2 (192.168.1.140)	310
192.168.1.4	vip4	Gk2 (192.168.1.140)	410

*Tabla 10. Configuraciones básicas de los VIP-480FS de la zona WANETCTIC*

En la siguiente tabla se observa los Planet VIP de la zona H.323 WANETDCS, se describe su dirección IP, su identificador de registro H.323 con su respectivo gatekeeper, el alias e IP del gatekeeper que controla la zona y las extensiones asignadas a los teléfonos.

IP	H.323 ID	Gatekeeper	Ext Teléfono
192.168.1.5	Vip5	Gk1 (192.168.1.33)	510
192.168.1.6	Vip6	Gk1 (192.168.1.33)	610
192.168.1.7	Vip7	Gk1 (192.168.1.33)	710
192.168.1.8	Vip8	Gk1 (192.168.1.33)	810

*Tabla 11. Configuraciones básicas de los VIP-480FS de la zona WANETDCS*

### 6.1.3 Software empleado

**GNU Gatekeeper:** Se utilizó para habilitar ambos gatekeeper de cada zona y su función principal es administrar las llamadas realizadas entre las dos zonas a través de las terminales correspondientes.

**Wireshark:** Se utilizó para capturar todo el tráfico generado de las llamadas realizadas entre las terminales, para que posteriormente esta información pueda ser analizada.

**Vlc:** Se utilizó para reproducir y crear flujos de audios para las llamadas de prueba entre las terminales.

**Navegador Web:** se utilizó para acceder a las configuraciones y administrar vía web las antenas MIMO, los Planet VIP-480FS y el router Belkin.

**Sincronizador de Hora:** Se utilizó un script en ms-dos para poder sincronizar los dos gatekeeper

**Teamviewer:** Se utilizó para monitorear el status de las llamadas, funcionamiento de los gatekeeper, reproducción de los audios y monitoreo del software encargado de capturar los paquetes de red, todo esto remotamente mediante el software antes mencionado.

**Programador de tareas de Windows:** Se utilizó para programar las tareas en horarios específicos para poder automatizar y sincronizar las actividades como, apertura y cierre de los analizadores de tráfico para las capturas, apertura y cierre del vlc para la reproducción de los audios, apertura y cierre para del script encargado de sincronizar las horas en los gatekeeper.

## 6.2 Características de la transmisión de flujos de voz (llamadas de pruebas)

Las características de las llamadas de prueba realizadas entre ambas zonas H.323 (WANETCTIC y WANETDCS) se pueden observar en la siguiente tabla.

H323 ID Terminal Marcante	H323 ID Terminal Marcada	IP Terminal Marcante	IP Terminal Marcada	Gatekeeper de la Terminal Marcante	Gatekeeper de la Terminal Marcada
Vip5	Vip1	192.168.1.5	192.168.1.1	GK1 (192.168.1.33)	GK2 (192.168.1.140)
Vip6	Vip2	192.168.1.6	192.168.1.2	GK1 (192.168.1.33)	GK2 (192.168.1.140)
Vip7	Vip3	192.168.1.7	192.168.1.3	GK1 (192.168.1.33)	GK2 (192.168.1.140)
Vip8	Vip4	192.168.1.8	192.168.1.4	GK1 (192.168.1.33)	GK2 (192.168.1.140)

*Tabla 12. Configuraciones parte 1 de las llamadas*

Se puede observar que cada renglón de la tabla representa una llamada de terminal a terminal, en total fueron 4 llamadas a la par, entre las 8 terminales.

En la tabla anterior se especifica los identificadores H.323 con los cuales las terminales se registraron a su respectivo gatekeeper, las IP correspondientes de las 8 terminales, y los gatekeeper que administran cada zona H.323.

H323 ID Terminal Marcante	H323 ID Terminal Marcada	VAD	Smart QoS	Códec	Duración de la llamada
Vip5	Vip1	Activado	Desactivado	G.729	8 horas aprox.
Vip6	Vip2	Desactivado	Desactivado	G.729	8 horas aprox.
Vip7	Vip3	Activado	Desactivado	G.711	8 horas aprox.
Vip8	Vip4	Desactivado	Desactivado	G.711	8 horas aprox.

*Tabla 13. Configuraciones parte 2 de las llamadas*

En esta tabla se observa otras características de las configuraciones de las llamadas entre las terminales, se especifica el códec que se empleó, si se usó el VAD o no, la duración de cada llamada.

Como antes se mencionó cada renglón de la tabla representa una llamada diferente, 4 en total al mismo tiempo, estas 4 llamadas se realizaron con diferentes anchos de banda del enlace punto a punto:

- 4 llamadas con un ancho de banda de máximo 130 Mbps (20 MHz)
- 4 llamadas con un ancho de banda de máximo 65 Mbps (10 MHz)
- 4 llamadas con un ancho de banda de máximo 32 Mbps (5 MHz)

Esto con la intención de evaluar la calidad de la red en diferentes escenarios y diferentes parámetros como es el ancho de banda el cual es un factor importante en la calidad del servicio en la transmisión de voz sobre IP.

## 6.3 Capturas

La captura de los flujos de tráfico generados por las llamadas de prueba, fueron capturados con el analizador de protocolos de red wireshark, a continuación se describe el procedimiento que se siguió para realizar dichas capturas.

Se ejecuta el wireshark instalado en la computadora donde se está ejecutando el gatekeeper en este caso en ambas zonas (GK1 y GK2), seleccionamos opciones de captura y nos saldrá una ventana como la siguiente:

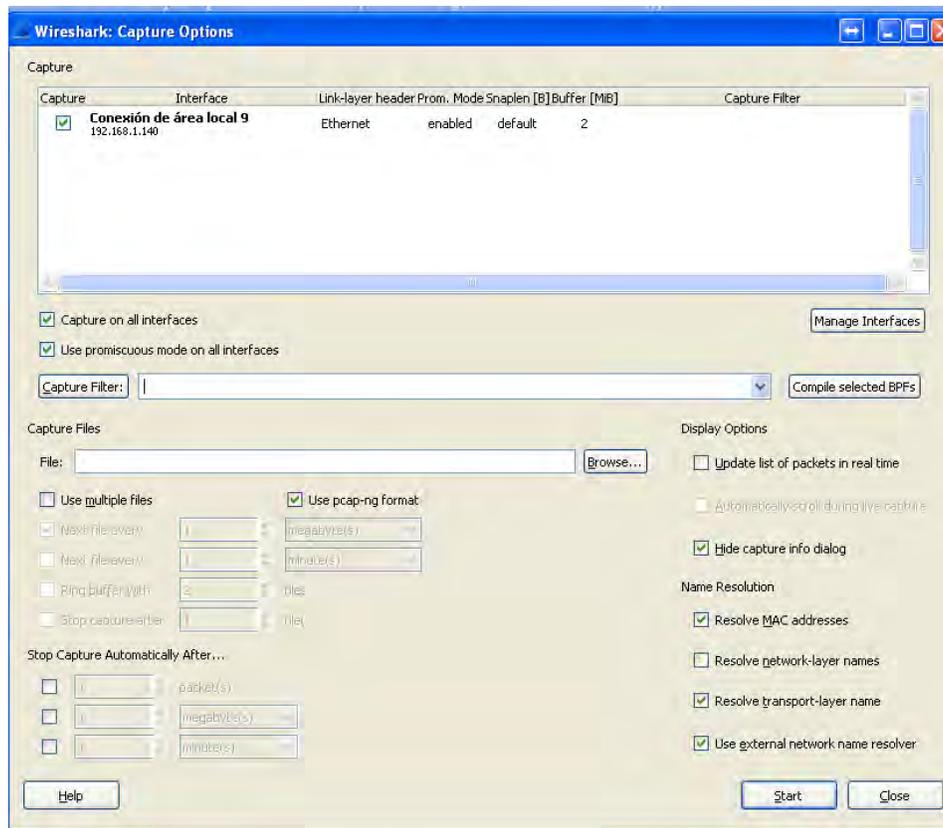


Figura 33. Configuración captura wireshark

En el campo capture filter se configuro de manera que capture los flujos de tráfico entre los dos gatekeeper y que detecte paquetes UDP, quedando de esta manera: host 192.168.1.140 and host 192.168.1.33 and UDP.



2.- Elegir la opción de RTP y dar en aceptar y se empezara a decodificar el paquete UDP a RTP, este proceso depende de las características del equipo en el cual se está realizando ya que demanda mucho procesamiento.

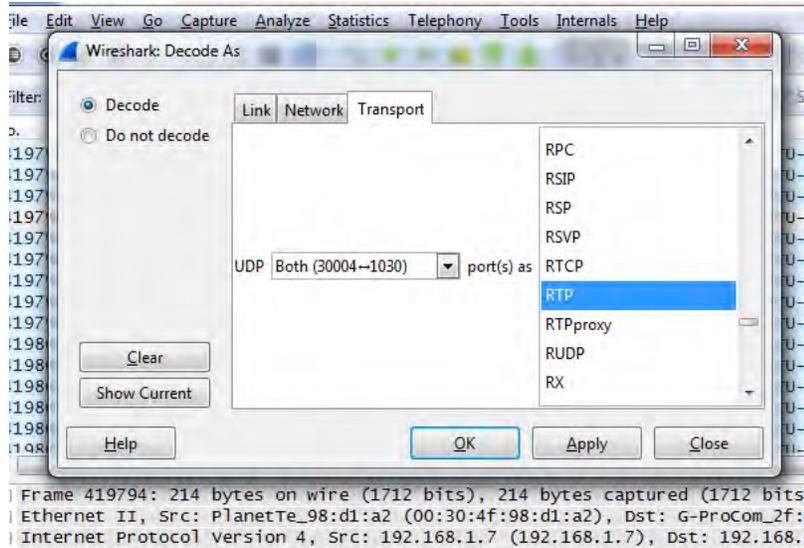


Figura 35. Pantalla Decode As

3.- Este proceso se repite hasta que ningún flujo UDP este visible y solo estén flujos RTP.

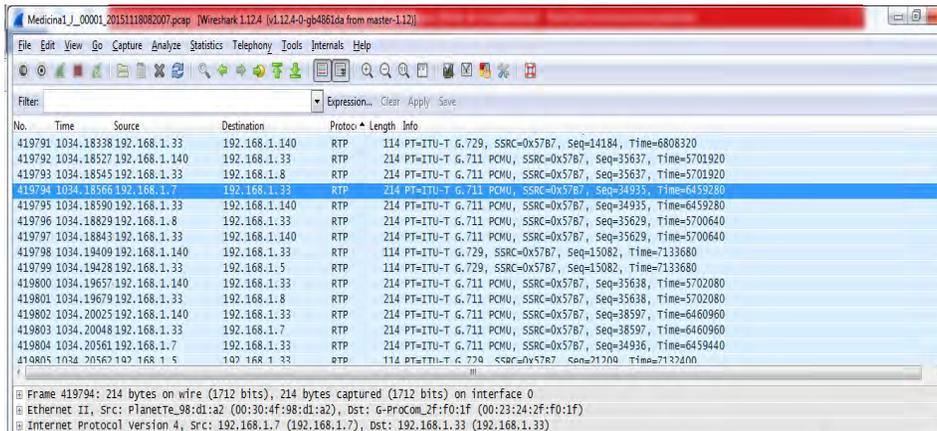


Figura 36. Paquetes RTP

4.- Cuando tenemos todos los flujos RTP seleccionamos en el menú “telephony” la opción de “RTP” y “Show All Streams”, para visualizar todos los flujos de las llamadas.

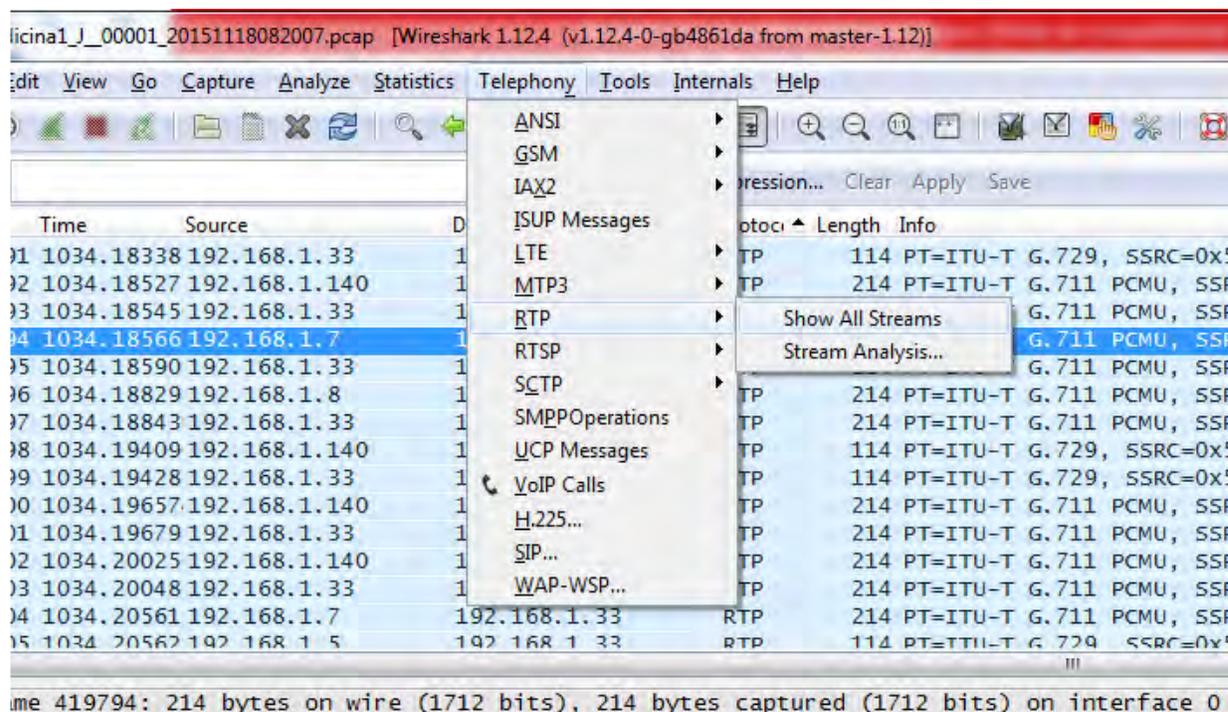


Figura 37. Opción Show All Streams

Nos aparecerá una pantalla como la de la siguiente imagen donde nos muestras los flujos de las llamadas, depende del número de llamadas realizadas son el número de flujos que nos aparecerán. En este caso son 4 llamadas y por cada llamada nos aparecerán 2 flujos uno de ida y el otro de vuelta en total serán 8 flujos, sin embargo, lo que nos interesa en este estudio es ver que impacto generan los parámetros de QoS, por lo tanto, nos interesa el flujo que recorre todo la red, es decir, el flujo de vuelta, este flujo lo identificamos mediante las direcciones IP de los gatekeeper y los puertos que se utilizan.

Src addr	Src port	Dst addr	Dst port	SSRC	Payload	Packets	Lost	Max Delta
192.168.1.33	1032	192.168.1.140	1034	0x57B7	g711U	163884	0 (0.0%)	338.310000
192.168.1.33	1030	192.168.1.140	1032	0x57B7	g711U	155996	0 (0.0%)	340.522000
192.168.1.33	1028	192.168.1.140	1028	0x57B7	g729	56934	2 (0.0%)	540.134000
192.168.1.33	1026	192.168.1.140	1026	0x57B7	g729	86544	0 (0.0%)	538.221000
192.168.1.140	1034	192.168.1.33	1032	0x57B7	g711U	157670	6219 (3.8%)	36182.1710
192.168.1.140	1032	192.168.1.33	1030	0x57B7	g711U	155383	6061 (3.8%)	36191.5300
192.168.1.140	1028	192.168.1.33	1028	0x57B7	g729	54860	2078 (3.6%)	36191.8240
192.168.1.140	1026	192.168.1.33	1026	0x57B7	g729	56669	2098 (3.6%)	36231.9740

Figura 38. Flujos RTP

Se toman los flujos de vuelta de la siguiente manera, que provengan de la dirección IP 192.168.1.40 (GK2) hacia la dirección IP 192.168.1.33 (GK1) ya que las llamadas se realizaron desde WANETDCS (GK1).

5.- Wireshark nos proporciona estadísticas de algunos parámetros de red, como, jitter, perdidas de paquetes, etc. Para obtener estas estadísticas es necesario realizar el proceso siguiente, en la ventana donde nos muestra los flujos de ida y vuelta, seleccionamos un flujo y con la opción “Analyze” nos proporciona una nueva ventana donde observamos que nos muestra datos como, el número de paquete, la secuencia, el delta, entre otros.

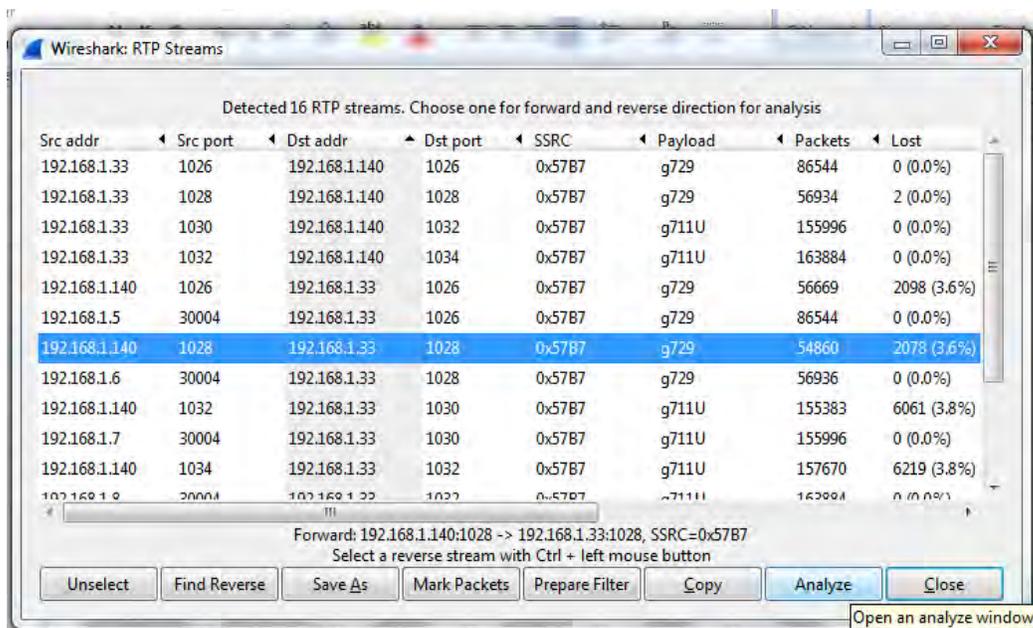
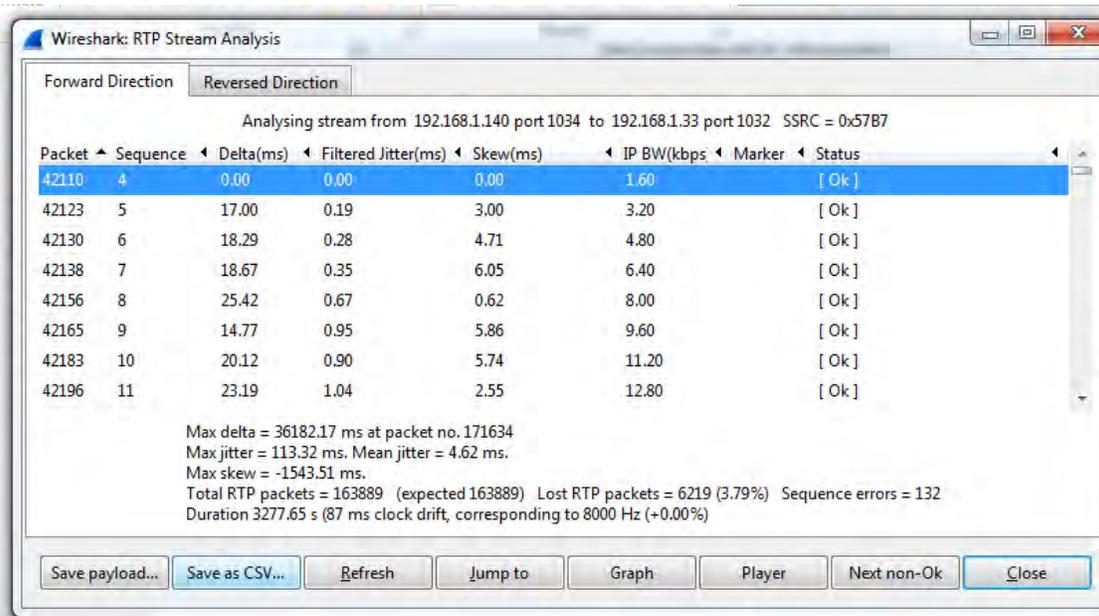


Figura 39. Opción Analyze en flujos RTP

Para poder utilizar estos datos posteriormente es necesario guardarlos mediante el botón “Save as CSV” lo cual nos permite guardar esta información en archivo con formato CSV de cada uno de los flujos RTP, este proceso se repite desde el paso número 5 con cada uno de los flujos deseados.

Figura 40. Análisis de flujo RTP y guardado de información en archivo CSV



Hasta el momento el proceso de filtrado se ha realizado completamente, sin embargo, esto se realizó para una hora de medición, por lo tanto, se deberá repetir el procedimiento para las demás horas restantes.

Una vez que se tengan todos los archivos CSV correspondientes, se extraen tres archivos TXT (traza de jitter, secuencia y delta), mediante un software en C++, estos TXT corresponden a cada uno de los flujos RTP capturados en una hora determinada.

Name	Date modified	Type	Size
 dcs_ctic_g711_novad_10mhz_9-10-sequency.txt	13/11/2015 10:32 a...	Text Document	838 KB
 dcs_ctic_g711_novad_10mhz_10-11-delta.txt	13/11/2015 10:16 a...	Text Document	1,113 KB
 dcs_ctic_g711_novad_10mhz_10-11-jitter.txt	13/11/2015 10:16 a...	Text Document	1,103 KB
 dcs_ctic_g711_novad_10mhz_10-11-sequency.txt	13/11/2015 10:16 a...	Text Document	1,189 KB
 dcs_ctic_g711_novad_10mhz_11-12-delta.txt	13/11/2015 10:18 a...	Text Document	1,118 KB
 dcs_ctic_g711_novad_10mhz_11-12-jitter.txt	13/11/2015 10:18 a...	Text Document	1,102 KB
 dcs_ctic_g711_novad_10mhz_11-12-sequency.txt	13/11/2015 10:18 a...	Text Document	1,198 KB
 dcs_ctic_g711_novad_10mhz_12-13-delta.txt	13/11/2015 10:19 a...	Text Document	1,103 KB
 dcs_ctic_g711_novad_10mhz_12-13-jitter.txt	13/11/2015 10:19 a...	Text Document	1,090 KB
 dcs_ctic_g711_novad_10mhz_12-13-sequency.txt	13/11/2015 10:19 a...	Text Document	1,180 KB
 dcs_ctic_g711_novad_10mhz_13-14-delta.txt	13/11/2015 10:21 a...	Text Document	1,085 KB
 dcs_ctic_g711_novad_10mhz_13-14-jitter.txt	13/11/2015 10:21 a...	Text Document	1,068 KB
 dcs_ctic_g711_novad_10mhz_13-14-sequency.txt	13/11/2015 10:21 a...	Text Document	1,176 KB
 dcs_ctic_g711_novad_10mhz_14-15-delta.txt	13/11/2015 10:22 a...	Text Document	1,114 KB
 dcs_ctic_g711_novad_10mhz_14-15-jitter.txt	13/11/2015 10:22 a...	Text Document	1,093 KB
 dcs_ctic_g711_novad_10mhz_14-15-sequency.txt	13/11/2015 10:22 a...	Text Document	1,198 KB

Figura 41. Conjunto de TXT obtenidos al procesar archivos CSV

# CAPÍTULO 7

## Capítulo 7. ANÁLISIS DE DESEMPEÑO Y CARACTERIZACIÓN DE PARÁMETROS DE QoS

En esta sección se realiza el análisis y caracterización del tráfico VoIP capturado en el escenario de red. El conjunto de mediciones analizadas en esta sección se mencionaron anteriormente haciendo referencia a las tablas 12 y 13.

Se evaluó el desempeño de un conjunto de llamadas de prueba, mediante el comportamiento de los diferentes tipos de códec empleados así como la habilitación/des-habilitación del VAD y los principales parámetros de QoS, tales como: jitter, pérdidas de paquetes y parámetro de Hurst.

La Figura 42 muestra los valores del parámetro de Hurst correspondientes al conjunto de trazas de los 4 codecs y variantes utilizados.

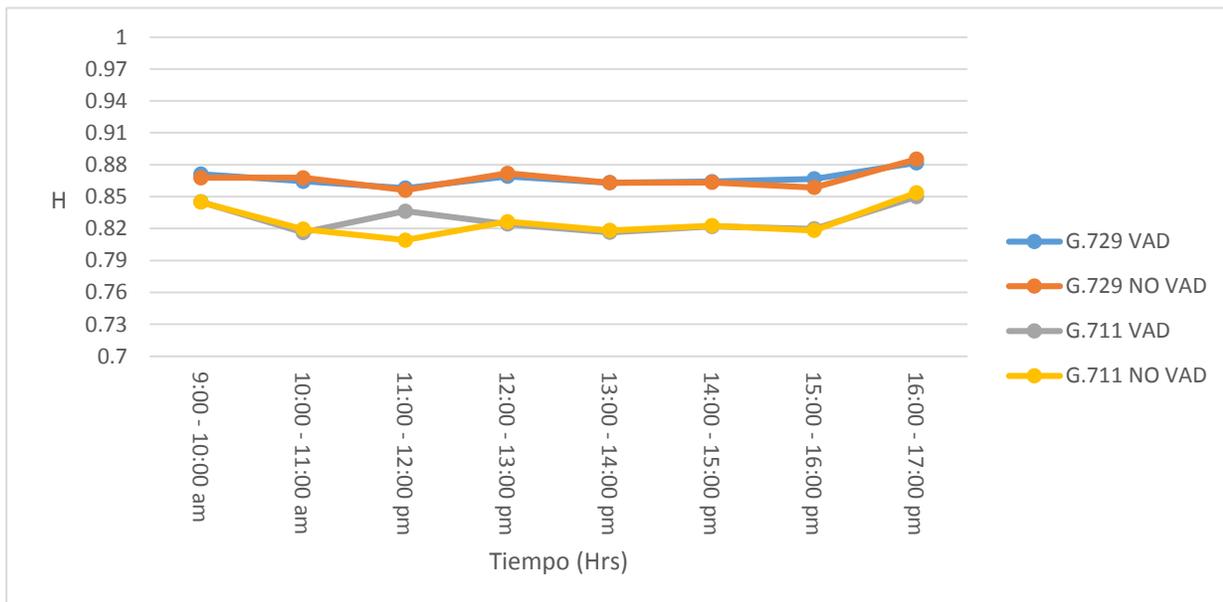


Figura 42. Valores de H con relación a las horas de medición en 5 Mhz

En esta figura analizamos los valores del parámetro H calculados por cada hora de medición, por cada diferente tipo de códec y utilizando el escenario donde la antena MIMO es configurada a una frecuencia de 5Mhz (32

Mbps). Podemos observar que el valor de H estimado en las series de tiempo de jitter del conjunto de llamadas realizadas mediante el CODEC G.729 presenta un valor mayor en el parámetro H respecto a las del conjunto de llamadas realizadas con el CODEC G.711. Como se mencionó anteriormente, mientras más alto sea el valor del parámetro H, la calidad de la llamada será más baja, es decir, a un valor más alto de H, corresponde un valor más bajo de la calidad.

Con el objetivo de obtener más información sobre que parámetros de QoS fue el responsable de las degradaciones de la voz presentadas en esta configuración (5 Mhz-32 Mbps), se analizaron los valores de jitter promedio del conjunto de llamadas de prueba como se muestra en la Figura 43.

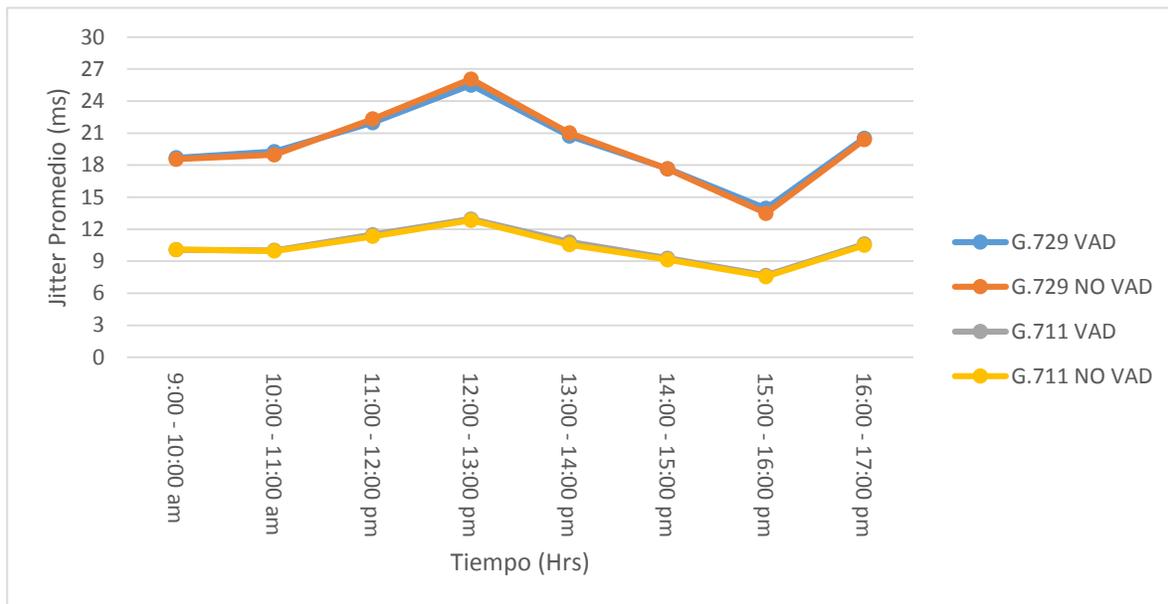


Figura 43. Valores de Jitter promedio obtenidos en las mediciones de 5 Mhz

En [21] se presenta la siguiente clasificación de la calidad de servicio en aplicaciones VoIP en función de los valores de jitter: para valores entre 0 y 20 ms la calidad es buena, para valores entre 20 y 50 ms la calidad es aceptables y para valores superiores a 50 ms, la calidad es mala.

Como se mencionó anteriormente, las llamadas realizadas mediante el códec G.711 presenta valores inferiores de H, respecto a las realizadas con G.729. En la Figura 43 de igual forma podemos observar y determinar que los valores

obtenidos del jitter promedio en el codec G.711 se encuentran entre 7.5ms y 12.8ms, mientras que en el códec G.729 se encuentran entre 13.5ms y 26ms. Esta Figura nos permite visualizar que los valores del jitter del códec G.711 se encuentran en el rango de calidad buena, mientras que los valores de jitter del códec G.729 se encuentran en los umbrales de calidad buena y aceptable. Este análisis nos confirma que un valor más grande de H implica una mayor degradación en la calidad de servicio percibida por el usuario final.

Otra de las métricas a analizar fueron las pérdidas. La Figura 44 muestra el comportamiento de esta métrica.

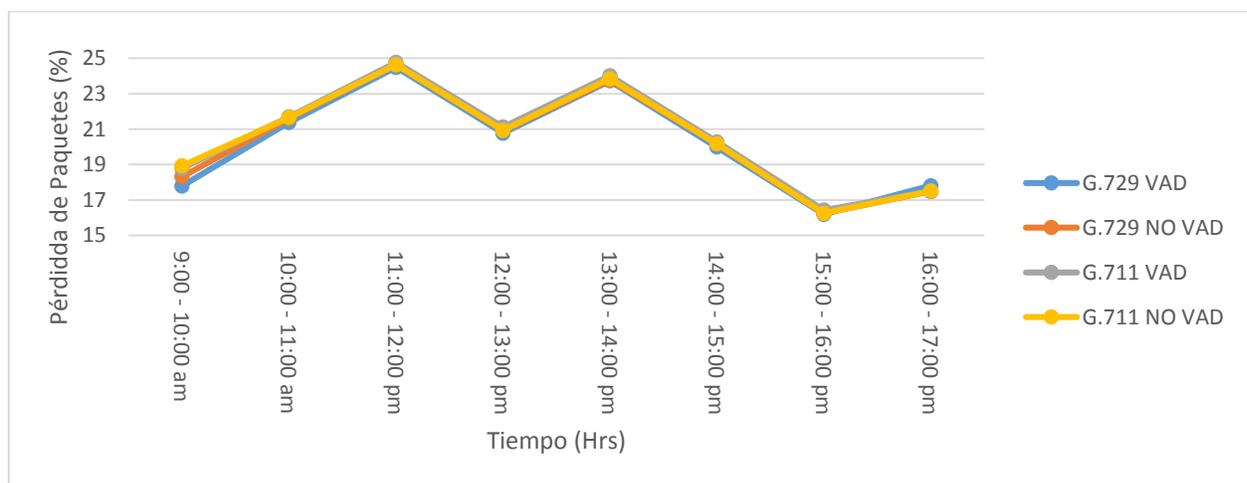


Figura 44. Valores de pérdidas de paquetes en 5 Mhz

En la Figura 44 se pueden observar que la tendencia de las pérdidas no está correlacionadas con los valores de jitter, es decir, no existe una relación proporcional entre el jitter y las pérdidas como se presentó en los valores de H. Sin embargo se puede observar que para este ancho de banda, las pérdidas fueron bastante altas.

El siguiente escenario bajo prueba fue con una configuración en la antena MIMO a una frecuencia de 10 Mhz que corresponden a un ancho de banda mayor (32-65 Mbps) que el escenario anterior. En este escenario se analizaron de igual forma, el parámetro H, jitter promedio y pérdida de paquetes. Las figuras 45, 46 y 47 nos muestran el comportamiento de estos parámetros.

En la Figura 45 se observa un comportamiento similar en los valores de H para las llamadas con G.711 Y G.729.

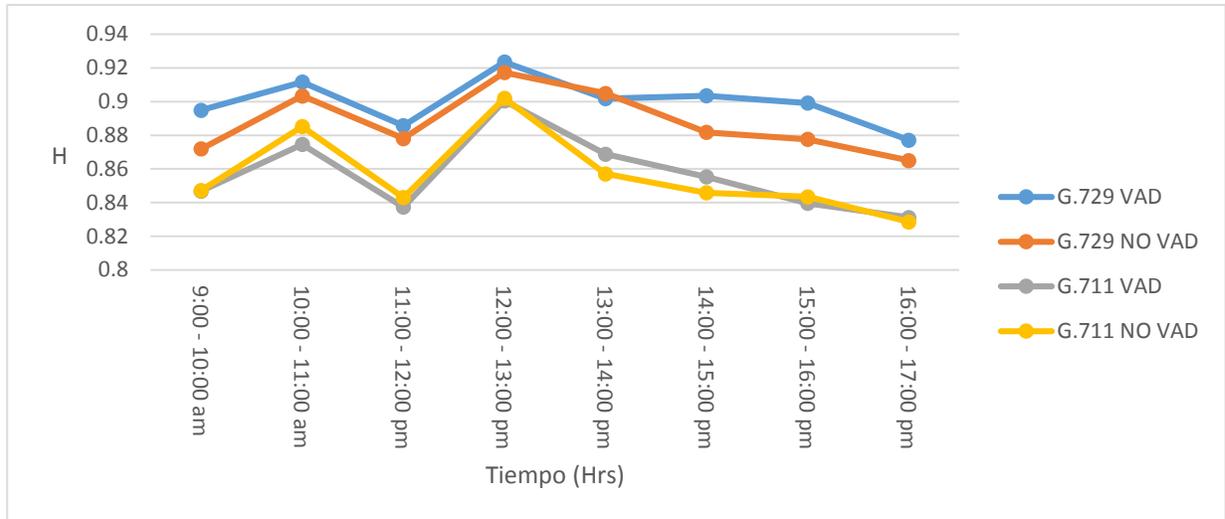


Figura 45. Valores obtenidos de H en frecuencia 10 Mhz

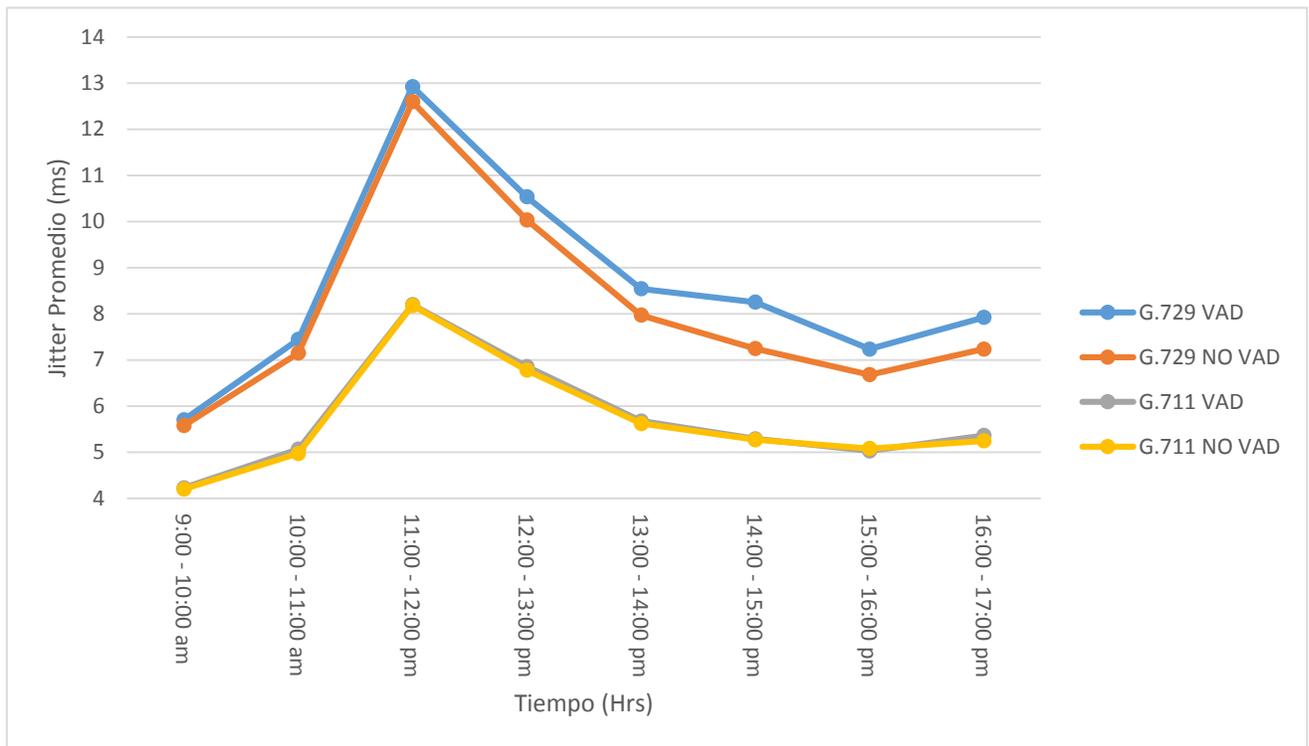


Figura 46. Valores obtenidos del Jitter promedio de frecuencia de 10 Mhz

La Figura 46 muestra los valores obtenidos del jitter promedio, se observan que son menores que en una frecuencia anterior y se encuentran en el rango

donde la calidad es buena. También se puede observar la correlación existente con los valores de H.

En la Figura 47 se observa que los valores de pérdidas no están correlacionados con los valores de jitter, es decir, no existe una relación proporcional entre el jitter y las pérdidas como se presentó en los valores de H. Sin embargo se puede observar que para este ancho de banda, las pérdidas fueron relativamente bajas.

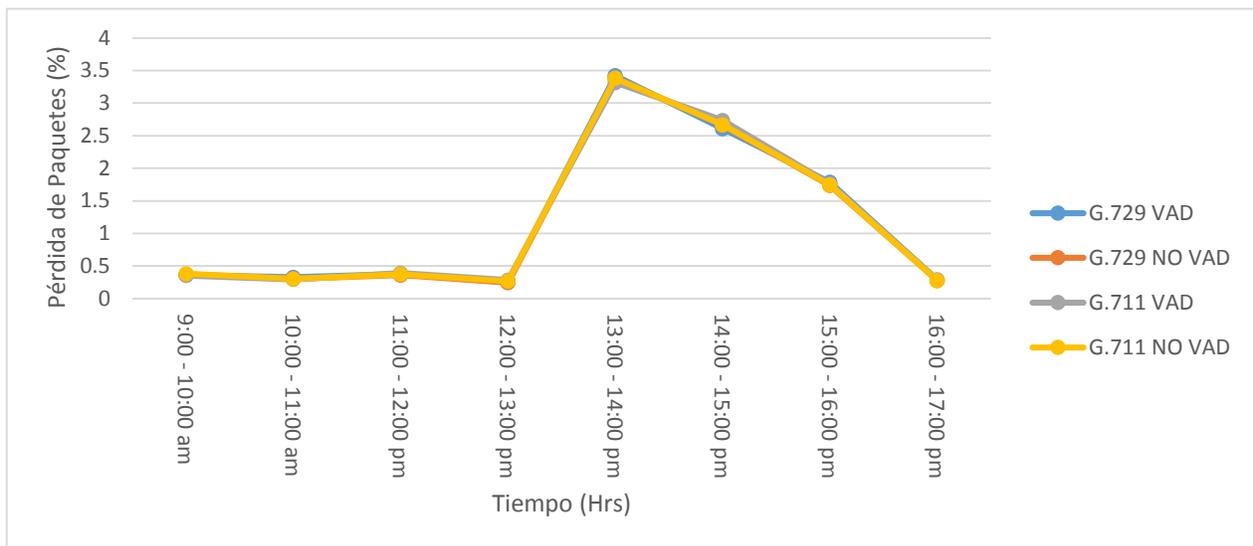


Figura 47. Pérdida de paquetes en la frecuencia de 10 Mhz

Por último, se analizaron las métricas anteriores en un escenario donde la antena MIMO fue configurada a una frecuencia de 20 Mhz que equivalen a un ancho de banda en el rango de 65-130 Mbps. Las Figuras 48,49 y 50 muestran los comportamientos de H, jitter promedio y pérdidas de paquetes.

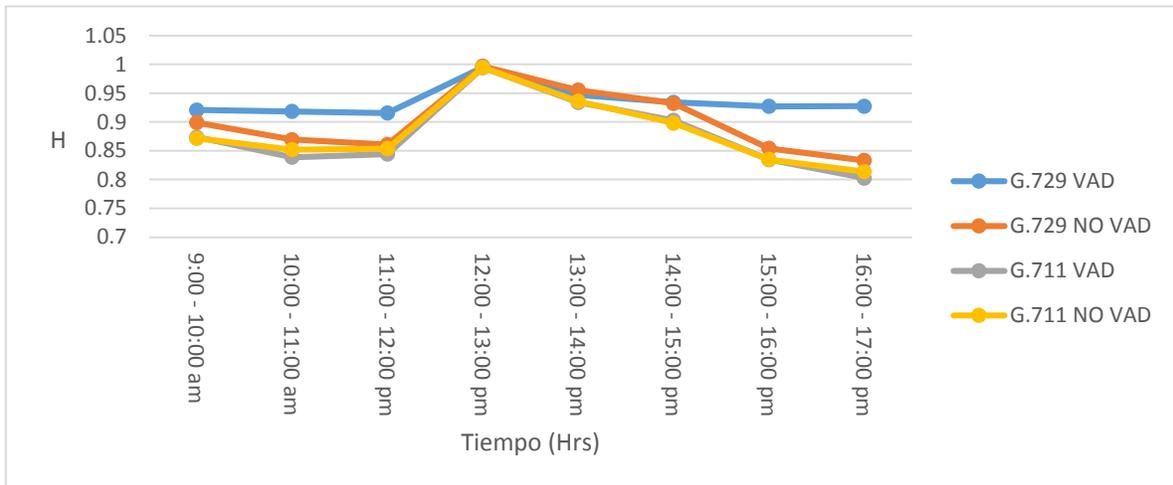


Figura 48. Valores H con relación a las horas de medición de 20 Mhz

En las Figuras 48 y 49 observamos que el comportamiento de H y jitter promedio es similar a los obtenidos a los escenarios de 5 Mhz y 10 Mhz, sin embargo los valores del jitter son menores a los encontrados en los escenarios anteriores.

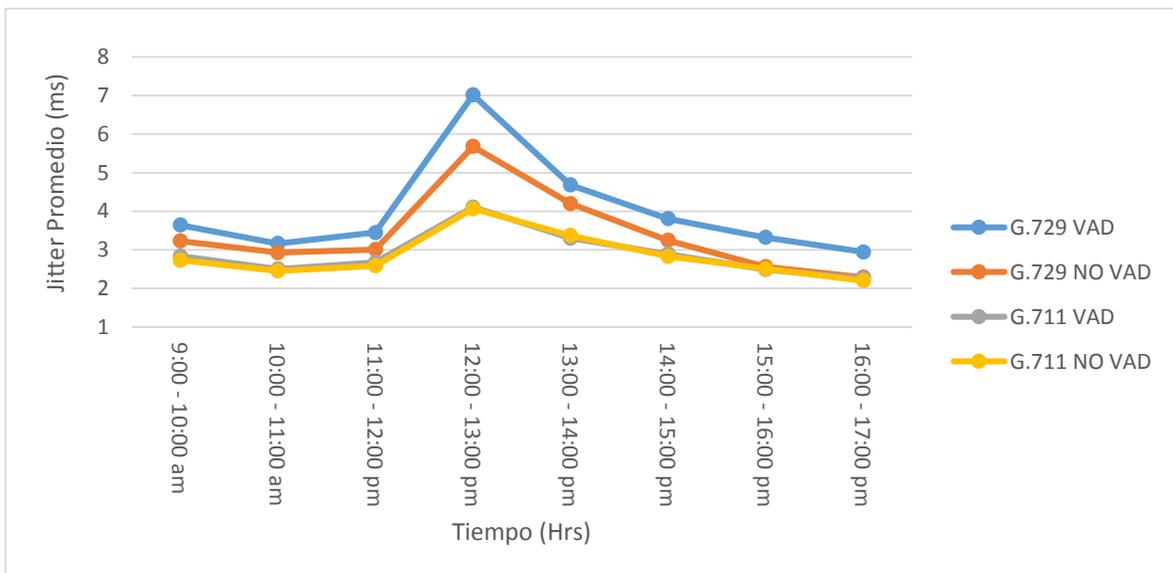


Figura 49. Valores de Jitter promedio obtenidos en las mediciones de 20 Mhz

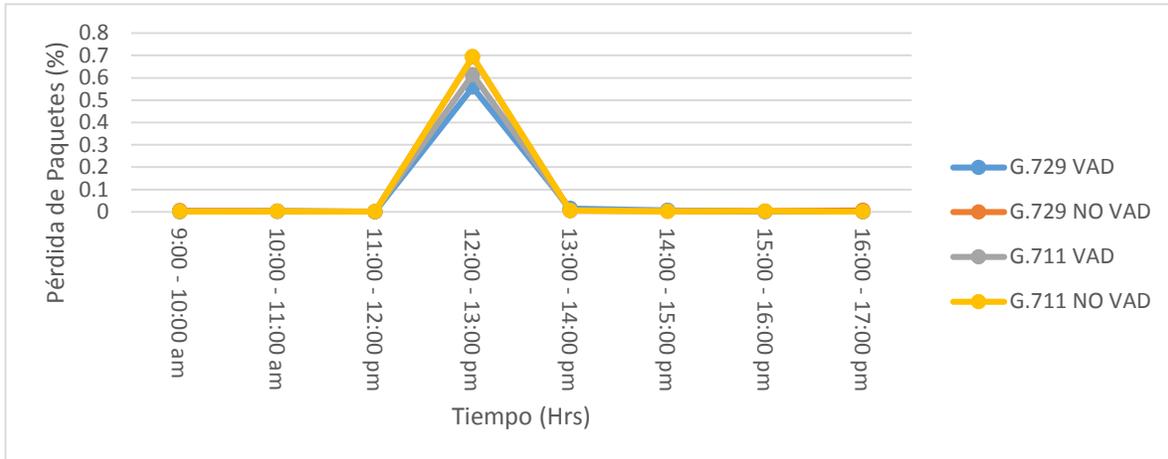


Figura 50. Pérdidas de paquetes en las mediciones de 20 Mhz

La Figura 50 nos muestra una ligera correlación entre el parámetro H, el jitter y la pérdida de paquetes. Además, se puede observar que para este ancho de banda, las pérdidas fueron muy bajas; por lo tanto, se puede concluir que el ancho de banda influye en gran medida en las pérdidas de paquetes en un canal de comunicación inalámbrico.

# CAPÍTULO 8

## Capítulo 8. CONCLUSIONES

Los sistemas de VoIP utilizan las redes IP para transportar voz, esto es posible gracias a la implementación de la digitalización de la señal de la voz mediante codecs, la voz es enviada en forma de paquetes y con la ayuda de los protocolos de señalización es posible realizar la comunicación entre los dispositivos (hardware o software) para que los usuarios finales pueden comunicarse.

H.323 y SIP son los protocolos más utilizados para la señalización de las llamadas de voz sobre redes IP. Estos protocolos carecen de calidad de servicio en sus implementaciones, el servicio de VoIP altamente sensible a la QoS. El nivel de calidad en un servicio VoIP es definido por diferentes parámetros, sin embargo, los de mayor impacto son el jitter, las pérdidas de paquetes y el ancho de banda.

Los resultados de nuestra investigación sobre el análisis y caracterización del tráfico de VoIP se pueden definir con los siguientes puntos:

- Se evaluó el desempeño de una red inalámbrica Ad Hoc donde los dispositivos de frontera para la comunicación son 2 antenas MIMO, dichas antenas se configuraron a 3 distintas frecuencias 5,10 y 20 Mhz para representar 3 escenarios distintos.
- Los valores de jitter promedio estuvo estrechamente relacionado en función del ancho de banda, mientras más elevado fue el ancho de banda, menores fueron los valores de jitter.
- El escenario de 5 Mhz fue el que presento mayores valores de jitter y pérdida de paquetes y en consecuencia las llamadas de prueba presentaron menor calidad de servicio.
- Los valores de pérdida en los tres escenarios de prueba no distingue entre tipo de codec, como sucede con el jitter y el parámetro H, en los cuales, las llamadas con el códec G.711 presentaron valores mayores respecto a las llamadas con G.729.
- El códec que presentó un mejor desempeño en los diferentes escenarios fue el G.711.

- En este estudio pudimos observar que los valores de jitter y pérdida de paquetes están altamente correlacionados con ancho de banda.
- Estos parámetros podrían entregarnos información muy valiosa si los integráramos al cálculo del MOS, ya que como recordaremos el MOS toma únicamente como parámetros al tipo de códec, delay y a las pérdidas.

## REFERENCIAS

- [1] José Antonio Carballar, VoIP: La telefonía de internet, 2007, España, Paraninfo, ISBN 978-84-283-2952-1.
- [2] M. Robles, “Voz sobre IP: Análisis del Servicio Instalado en Facultad de Telemática”, Tesis de Maestría, Universidad de Colima, Junio 2001.
- [3] E. Herrera, Tecnologías y redes de transmisión de datos, 2003, Limusa, ISBN 9681863836, 9789681863838.
- [4] A. Castro, R. Fusario, Teleinformática para ingenieros en sistemas de información II, 1999, Reverte, ISBN 8429143912, 9788429143911.
- [5] Recomendación UIT-T I.120, Redes Digitales de Servicios Integrados, 1993.
- [6] A. Ríos, “Análisis de desempeño y caracterización de parámetros de QoS en una red VoIP-H.323”, Tesis, Universidad de Quintana Roo, Diciembre 2012.
- [7] I. Lackerbauer, Internet, 2001, Marcombo, ISBN 8426713009, 9788426713001.
- [8] José Barceló, Protocolos y aplicaciones Internet, 2008, UOC, ISBN 8497887417, 9788497887410.
- [9] J. M. Moya, D. Martínez, Tecnología VoIP y Telefonía IP, Primera edición, 2006, Alfaomega, ISBN 84-96300-22-6.
- [10] G. Torres, “Desarrollo de un Generador de Trafico Sintético de VoIP”, Tesis, Universidad de Quintana Roo, Junio 2013.
- [11] M. García, O. Caselles, F. López, A. Sánchez, Servicios de Red. Una visión práctica, Segunda edición, 2010, Lulu.com, ISBN 978-1-4452-7134-7.
- [12] J. Andreu, Voz IP (Servicios en Red), 2011, Editex, ISBN 849003091X, 9788490030912.
- [13] Pavlos Papageorgiou, A Comparison of H.323 vs SIP, University of Maryland at College Park, 4 de Junio de 2001.
- [14] J. López, “Análisis y Caracterización de Tráfico de VoIP bajo un Ambiente Inalámbrico”, Tesis, Universidad de Quintana Roo, Diciembre 2012.

- [15] HP,  
[http://h20564.www2.hp.com/hpsc/doc/public/display?docId=emr\\_na-c01865794](http://h20564.www2.hp.com/hpsc/doc/public/display?docId=emr_na-c01865794), 2015.
- [16] WNI México Wireless Solution,  
[http://www.wni.mx/index.php?page=shop.product\\_details&flypage=flypage\\_new.tpl&product\\_id=234&category\\_id=44&option=com\\_virtuemart&Itemid=53](http://www.wni.mx/index.php?page=shop.product_details&flypage=flypage_new.tpl&product_id=234&category_id=44&option=com_virtuemart&Itemid=53), 2015.
- [17] Teléfono Alámbrico Vtech Clio 10b. <https://www.linio.com.co/p/telefono-ala-mbrico-vtech-clio-10b-negro-yn7gef#productDetails>. 2015.
- [18] Belkin Products,  
[http://www.belkin.com/es/IWCatProductPage.process?Product\\_Id=546435](http://www.belkin.com/es/IWCatProductPage.process?Product_Id=546435), 2015.
- [19] TrendNet Products  
[http://www.trendnet.com/langsp/products/proddetail.asp?status=view&prod=535\\_TE100-S24g#tabs-solution02](http://www.trendnet.com/langsp/products/proddetail.asp?status=view&prod=535_TE100-S24g#tabs-solution02), 2015.
- [20] Ubiquiti Networks, <https://www.ubnt.com/airmax/nanostationm/>, 2015.
- [21] *Performance Measurement and Analysis of H.323 Traffic*. University of Columbus.