UNIVERSIDAD DE QUINTANA ROO

División de Ciencias e Ingeniería

Análisis de desempeño y caracterización de parámetros de QoS en una red VoIP-H.323

Trabajo de Tesis

para obtener el grado de

Ingeniero en Redes

PRESENTA

Antonio Humberto Ríos Arreola

Director de Tesis

Dr. Homero Toral Cruz

Asesores

MTI. Vladimir Veniamín Cabañas Victoria

MSI. Laura Dávalos Castilla

Chetumal, Quintana Roo, México, Diciembre de 2012

UNIVERSIDAD DE QUINTANA ROO

División de Ciencias e Ingeniería

Trabajo de Tésis 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:

M.T.I. Vladimir Veniamin Cabañas Victoria

Asesor:

M.S.I. Laura Yésica Dávalos Castilla.

Chetumal, Quintana Roo, México, Diciembre de 2012.

Dedicatoria

Dedico este trabajo a todas aquellas personas que lo han hecho posible, a mis padres y mi hermano por su invaluable apoyo y paciencia que me han tenido en este proceso de estudios.

A mis maestros que me han compartido sus años de experiencia en conocimientos para poder realizar este trabajo.

A mis mejores amigos, Karen Carranza, Adriana Rosado, Jacim rodríguez, Martín Calderón, por siempre recordarme que es lo correcto y que gracias a ellos no me he desviado del camino correcto.

No sería lo que soy ni estaría donde me encuentro de no ser por ustedes.

Agradecimientos

A mis maestros, que también considero mis amigos. No tengo palabras para expresar lo agradecido que estoy con ustedes por compartirme mucho de su tiempo y sus conocimientos. Son pieza clave de lo que soy y que lograré más adelante.

Las siguientes generaciones de ingenieros en redes tienen mucha suerte de tener personas como ustedes en su formación académica.

Dr. Homero Toral Cruz

MTI. Vladimir Veniamín Cabañas Victoria

M.S.I. Laura Yésica Dávalos Castilla

Ing. Rubén Enrique González Elixavide

MTI. Melissa Blanqueto Estrada

Índice

Índice de tablas

Índice de figuras

Lista de abreviaciones

ARP	Protocolo de Resolución de Direcciones (Address Resolution Protocol)
BCH	Bose - Chaudhuri - Hocquengham
CIF	Formato Intermedio Común (Common Intermediate format)
CODEC	COdificador-DECodificador (COder-DECoder)
ICMP	Protocolo de Control de Mensajes de Internet (Internet Control Message Protocolo)
IETF	Fuerza de Trabajo de Ingeniería de Internet (Internet Engineering Task Force)
IP	Protocolo de Internet (Internet Protocolo)
ISDN	Red Digital de Servicios Integrados (Integrated Service Digital Network)
ITU	Unión Internacional de Telecomunicaciones (International Telecommunication Union)
LAN	Red de área local (Local Area Network)
MC	Controlador Multipunto (Multipoint Controller)
MCS	Sistema de Comunicaciones Multipunto (Multipoint Communications System)
MCU	Unidad de Control Multipunto (Multipoint Control Unit)
MOS	Calificación de Opinión Media (Mean Opinion Score)
MP	Procesador Multipunto (Multipoint Processor)
MTU	Unidad de Transmisión MáXima (Maximum Transmission Unit)
OWD	Retardo Extremo a Extremo (One Way Delay)
PCM	Modulación por Codificación de Pulsos (Pulse Code Modulation)
PLR	Tasa de Pérdida de Paquetes (Packet Loss Rate)
PSTN	Red Telefónica Pública Conmutada (Public Switch Telephone Network)
QoS	Calidad de Servicio (Quality of Service)
QCIF	Cuarto de CIF (Quarter CIF)
RAS	Registro, Admisión y Estado (Registration, Admission and Status)
RRJ	Rechazo de Registro (Registration Reject)
RRQ	Petición de Registro (Registration Request)
RTCP	Protocolo de Control de Transporte en Tiempo Real (Real-time Transport Control protocol)
RTP	Protocolo de Transporte en Tiempo Real (Real-time Transport Protocol)
SIP	Protocolo de Inicio de Sesión (Session Initiation protocol)
SNR	Relación Señal a Ruido (Signal Noise Ratio)
TCP	Protocolo de Control de Transporte (Transport Control Protocolo)
TTL	Tiempo de Vida (Time To Live)
TSAP	Punto de Acceso al Servicio de Capa de Transporte (Transport layer Service Access Point)
UAC	Agente de Usuario Cliente (User Agent Client)

UAS	Agente de Usuario Servidor (User Agent Server)
UDP	Protocolo de Datagrama de Usuario (User Datagram Protocolo)
VoIP	Voz sobre el Protocolo de Internet (Voice over Internet Protocol)

Resumen

La voz sobre el protocolo de Internet es una de las tecnologías más populares e innovadores que está comenzando a abrirse paso a través de muchos ámbitos, tales como: sociales, comerciales, tecnológicos, académicos, investigación, etc.

Se trata de una tecnología de menor costo en comparación con la telefonía convencional, debido a que utiliza Internet en lugar de la red tradicional de conmutación por circuitos para transportar la información.

Internet es una de las redes más importantes en el área de las telecomunicaciones y últimamente ha evolucionado a un ritmo muy acelerado; a tal grado de convertirse en la red convergente, sobre la cual, múltiples aplicaciones transmiten información de diversas naturalezas. Sin embargo, Internet ofrece un servicio de mejor esfuerzo y no garantiza calidad de servicio y la voz sobre el protocolo de Internet es una tecnología que demanda cierto nivel de calidad de servicio. La calidad de servicio en un sistema de voz sobre el protocolo de Internet está en función de un conjunto de parámetros: retardo extremo a extremo, jitter, pérdida de paquetes, ancho de banda, tipo de codificador, tamaño de paquete, tamaño de de-jitter buffer, etc.; sin embargo los que tiene mayor impacto son: retardo extremo a extremo, jitter y pérdida de paquetes.

La calidad de servicio o desempeño de una aplicación de voz sobre el protocolo de Internet, puede ser evaluado mediante técnicas subjetivas y/o objetivas, tales como el MOS y/o el Modelo E, respectivamente. El MOS requiere que algunas personas evalúen la calidad general de muestras de voz bajo las siguientes calificaciones: 1 (mala), 2 (pobre), 3 (aceptable), 4 (buena) y 5 (excelente). El Modelo E es un método computacional que combina los efectos del retardo, pérdida de paquetes, tipo de codificador y otros desperfectos en un solo valor llamado factor R, el cual, toma un rango de 0 (calidad mala) a 100 (calidad excelente) para evaluar la calidad de la voz. En este trabajo, se evaluó la calidad mediante el cálculo del factor R y finalmente se realizó un mapeo o relación entre este factor y valores de MOS.

En la actualidad existe un gran interés por diseñar aplicaciones y redes de voz sobre Internet que garanticen cierto nivel de calidad al usuario final. Para lograr esta tarea, es necesario realizar mediante mediciones de red, un estudio de los principales parámetros que determinan la calidad de servicio.

Derivado de los puntos mencionados con anterioridad, en este trabajo de tesis se realizó lo siguiente: a) generación de tráfico de voz real, mediante el establecimiento de un conjunto de llamadas de prueba sobre un escenario de red H.323; b) implementación de un escenario representativo donde se puedan realizar llamadas de larga distancia sobre Internet; c) captura de patrones de tráfico mediante mediciones de red, haciendo uso de un analizador de protocolos de red; d) caracterización de los principales parámetros que determinan la calidad de servicio a partir de la mediciones realizadas; y e) evaluación del desempeño del conjunto de llamadas establecidas.

Capítulo 1
Introducción

Hoy en día la telefonía juega un papel muy importante en las redes de telecomunicaciones. Principalmente existen tres redes de telefonía: la red telefónica pública conmutada (PSTN), la red de telefonía celular y recientemente la red de voz sobre el protocolo de Internet (VoIP).

Debido a que la red telefónica pública conmutada está basada en la tecnología de conmutación de circuitos (recursos dedicados), provee una calidad de voz óptima, sin embargo, esto se logra a un costo elevado. Por otro lado, la tecnología VoIP, basada en la conmutación de paquetes está comenzando a resolver el problema del costo elevado, sacrificando la calidad de servicio (QoS), puesto que ofrece la capacidad de poder hacer llamadas telefónicas usando una conexión a Internet, donde los recursos son compartidos.

Una gran cantidad de aplicaciones están siendo implementadas sobre Internet, una de las más importantes es VoIP; sin embargo, Internet ofrece un servicio de mejor esfuerzo y no garantiza calidad de servicio y VoIP es una aplicación que demanda cierto nivel de calidad de servicio. La QoS en un sistema VoIP está en función de un conjunto de parámetros: retardo extremo a extremo (OWD), jitter, pérdida de paquetes (PLR), ancho de banda, tipo de CODEC, tamaño de paquete, tamaño de de-jitter buffer, etc.

Para evaluar la calidad de servicio o desempeño de una aplicación VoIP existen técnicas subjetivas y objetivas, tales como el MOS y el Modelo E, respectivamente. El MOS requiere que algunas personas evalúen la calidad general de muestras de voz bajo las siguientes calificaciones: 1 (mala), 2 (pobre), 3 (aceptable), 4 (buena) y 5 (excelente). El Modelo E es un método computacional que combina los efectos del retardo, pérdida de paquetes, tipo de CODEC y otros

desperfectos en un solo valor llamado factor R, el cual, toma un rango de 0 (calidad mala) a 100 (calidad excelente) para evaluar la calidad de la voz.
En la actualidad existe un gran interés por diseñar aplicaciones y redes VoIP que garanticen cierto nivel de QoS al usuario final. Para lograr esta tarea, es necesario realizar un estudio de las principales métricas de QoS mediante mediciones de red.

Este trabajo tiene como objetivo central caracterizar los principales parámetros de QoS y evaluar el desempeño en un conjunto de llamadas entre dos redes LAN interconectadas por el backbone de Internet.

1.1 PLANTEAMIENTO DEL PROBLEMA

Internet no fue diseñado para manejar tráfico en tiempo real, tal como tráfico de voz. El tráfico de voz generado por una llamada telefónica realizada a través de Internet, demanda bajos niveles de retardo extremo a extremo, pérdida de paquetes y jitter, para poder garantizar cierto nivel de QoS al usuario final.

Los problemas que enfrentan los paquetes de voz a lo largo de su recorrido por Internet para llegar a su destino, se derivan de la naturaleza del servicio de entrega de “mejor esfuerzo”. El mecanismo de entrega de mejor esfuerzo envía paquetes sin ningún mecanismo que garantice la entrega correcta de los mismos. Debido a que en Internet hay millones de paquetes transitando y compitiendo por los recursos de la red, la probabilidad de pérdidas, colisiones, retardos y jitter es alta, de aquí que la QoS no esté garantizada

La calidad de servicio puede evaluarse mediante una gran variedad de métodos. Dos de los más usados son el MOS y el Modelo E. El MOS es un método subjetivo, en el cual, dos personas especialmente capacitadas son colocadas en habitaciones diferentes, a prueba de ruido, ambas personas realizan una llamada telefónica y evalúan la calidad en una escala de cinco valores, que va de malo a

excelente pasando por pobre, aceptable y bueno. Por otro lado, el Modelo E es un método objetivo, el cual, utiliza ciertos parámetros de la red IP y de la aplicación VoIP (retardo, pérdida de paquetes y tipo de CODEC) para obtener una calificación de la calidad de la voz en una escala de 0 a 100; donde esta escala puede ser mapeada a los cinco valores de MOS.

En esta tesis se evaluó la calidad de servicio de un conjunto de llamadas H.323 y se realizó la caracterización de los principales parámetros de desempeño. Estos resultados pueden ser usados para construir modelos representativos de estos parámetros y ayudar en la implementación y diseño adecuado de redes y aplicaciones VoIP.

1.2 JUSTIFICACIÓN

Dentro de la constante tendencia de evolución de las tecnologías de información y redes de comunicación, Internet es una de las redes de mayor impacto en la vida cotidiana de millones de personas y es de singular importancia mencionar el gran crecimiento que han tenido en estos últimos años. Internet, puede considerarse como base para el desarrollo de nuevos servicios y aplicaciones no restringidos no sólo al transporte de datos, sino a la integración de múltiples medios (voz, datos, video, fax, etc.) sobre una misma infraestructura de red. Esta integración de múltiples medios es mejor conocida como convergencia de redes de comunicación.

Una de las piezas clave que ha permitido dicha convergencia ha sido el desarrollo de la tecnología VoIP. La cual es una de las aplicaciones más importantes sobre Internet, derivado de su disponibilidad multiplataforma que permite a cualquier usuario hacer uso de esta tecnología independientemente del sistema operativo que utilice, ofrece servicios más atractivos de comunicación a un menor costo a diferencia de la red telefónica pública conmutada y en la actualidad su infraestructura de red permite la integración de múltiples servicios de

comunicación. Sin embargo, una de las debilidades que presenta es que la calidad de servicio no está completamente garantizada. Por tal motivo es de gran importancia conocer el funcionamiento de esta emergente aplicación y realizar un estudio de los principales parámetros que determinan la calidad de servicio.

1.3 OBJETIVOS

1.3.1 OBJETIVO GENERAL

Realizar un estudio y caracterización de los principales parámetros que afectan la calidad de servicio y evaluar el desempeño de un conjunto de llamadas en una red VoIP implementada entre dos redes LAN (UQROO Chetumal - CINVESTAV Guadalajara) interconectadas por el Backbone de Internet.

1.3.2 OBJETIVOS ESPECÍFICOS

- Conocer el conjunto de recomendaciones y protocolos que hacen posible la existencia de la tecnología VoIP.
- Implementar escenarios reales y representativos para generar tráfico de voz sobre el backbone de Internet, mediante un conjunto de llamadas de prueba.
- Realizar la captura del tráfico VoIP generado para colectar un conjunto de patrones de tráfico (series de tiempo o trazas).
- Realizar un estudio del comportamiento de los principales parámetros de QoS, mediante el análisis de las series de tiempo colectadas.
- Evaluar el desempeño de las llamadas de prueba.

1.4 CONTRIBUCIONES DE LA TESIS

- Se crearon escenarios reales y representativos para realizar llamadas telefónicas sobre Internet, mediante la implementación de una red VoIP-H.323.
- Se colectaron un conjunto de trazas de tráfico VoIP mediante la captura de los paquetes IP correspondientes a las llamadas establecidas.
- Se realizó un estudio y caracterización de los principales parámetros de QoS que tienen mayor impacto en una llamada telefónica sobre Internet.
- Se realizó la evaluación de desempeño mediante el Modelo E y el MOS al conjunto de llamadas de prueba establecidas.
- Se establecieron relaciones entre los principales parámetros de QoS.

Capítulo 2
Redes de Telecomunicaciones

El servicio de telefonía ha sido uno de los principales servicios durante toda la historia de evolución de las redes de telecomunicaciones.

La telefonía tradicional utiliza la tecnología de conmutación de circuitos, es decir, un canal dedicado para transportar la voz entre los usuarios finales, debido a esta característica, la calidad de la voz está completamente garantizada, sin embargo, los costos de instalación y tarificación de este servicio no siempre resulta ser accesible para muchos usuarios.

Últimamente, se está popularizando el servicio de telefonía a través de las redes IP (telefonía IP), debido a que las redes IP hacen uso de la tecnología de conmutación de paquetes, los recursos son compartidos y por tal motivo la calidad de la voz no está garantizada.

Uno de los principales retos es la unificación de redes, es decir que todos los servicios de comunicaciones sean transportados sobre una red convergente que garantice cierto nivel de QoS.

2.1 Redes de conmutación por circuito

Las comunicaciones mediante redes de conmutación por circuitos, permite que dos terminales (emisor y receptor) se comuniquen a través de un circuito dedicado. Por lo tanto los recursos están reservados exclusivamente para el intercambio de información entre terminales origen y destino. Este tipo de comunicación involucra tres fases: establecimiento del circuito, transferencia de datos y desconexión del circuito.

Es necesario antes de la comunicación establecer una ruta física entre las terminales. Esto quiere decir que la capacidad del enlace debe estar reservado entre ambos nodos y cada nodo debe tener capacidad interna de conmutación para manejar la petición de comunicación [1].

Una vez que el circuito se ha establecido, los recursos asociados a él no pueden ser usados por otra conexión hasta que el circuito es desconectado [1].

Los ejemplos más comunes de redes por conmutación de circuito son la Red Telefónica Publica Conmutada (PSTN) y la Red Digital de Servicios Integrados (ISDN) [1].

2.1.1 Red Telefónica Pública Conmutada (PSTN)

La red telefónica pública conmutada ha evolucionado desde que Alexander Graham Bell realizó su primera transmisión de voz a través de un cable en 1876 [3].

En la PSTN, el enlace de comunicación entre dos partes en una llamada es llevado a cabo a través de un conjunto de circuitos. Algunas de estas conexiones entre circuitos, son semipermanentes, son modificadas por operadores de red en lapsos de tiempo de meses o años conforme a sus suscripciones y patrones de tráfico [4].

Los circuitos pueden usar tecnología de transmisión de señales analógicas o digitales. Los estándares de circuitos de voz analógicos fueron diseñados para ajustarse a una banda de frecuencia de 4000 Hz, aunque únicamente 3400 Hz están disponibles para el usuario [4].

La primera transmisión de voz, enviada por Alexander Graham Bell, tuvo lugar en 1876 a través de lo que se llamó un circuito ring-down. Un circuito ring-down

significa que no hay marcación de número, en su lugar, un cable conectaba físicamente dos dispositivos. Básicamente una persona descolgaba el teléfono y otra se encontraba en el otro extremo (no había tonos de llamada). Con el tiempo, este diseño básico evolucionó desde una transmisión de voz de un único sentido, en la que sólo podía hablar un usuario, hasta una transmisión de voz bidireccional, en la que ambos usuarios podían hablar. Para mover las voces por el cable se necesitaba un micrófono de carbón, una batería, un electroimán y un diafragma de hierro. También se necesitaba un cable físico entre cada ubicación a la que el usuario quería llamar. Sin embargo, en ese tiempo todavía no existía el concepto de marcar un número para alcanzar un destino [3].

Para determinar cuántas líneas se necesitan en una casa hay que pensar en cada persona a la que se llama como un valor de N y utilizar la siguiente ecuación: N x (N-1) / 2. De esta manera, si se quiere llamar a 4 personas, se necesitan 6 pares de líneas en una casa (ver Figura 2 - 1).

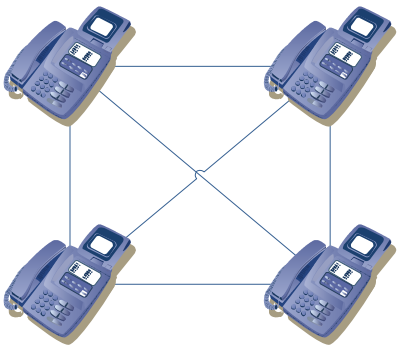


Figura 2 - 1. Ejemplo de cuántas líneas se necesitan para 4 Teléfonos

Debido al costo y a la imposibilidad de poner un cable físico entre todas las personas que quieren acceder a un teléfono en la Tierra, se ha desarrollado un dispositivo que puede asociar un teléfono con otro teléfono. Con este dispositivo, llamado conmutador, los usuarios de teléfono sólo necesitan un cable que vaya a la oficina del conmutador, y este se encarga de interconectarlo con los demás usuarios [3].

2.1.2 RED DIGITAL DE SERVICIOS INTEGRADOS (ISDN)

La red digital de servicios integrados (ISDN) 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 ISDN, 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 [5].

Las ISDN soportan aplicaciones diversas, entre las cuales están las conexiones conmutadas y no conmutadas. Las conexiones conmutadas en una ISDN comprenden conexiones con conmutación de circuitos, conexiones con conmutación de paquetes, y sus concatenaciones [5].

En la medida en que sea posible en la práctica, los nuevos servicios que se introduzcan en una ISDN deberán disponerse de modo que sean compatibles con las conexiones digitales conmutadas a 64 kbit/s [5].

Una ISDN 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 una inteligencia compatible en los terminales 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 torcidos 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 [8].
ISDN, es un sistema 100% digital de telecomunicaciones [6].

2.2 RED DE CONMUTACIÓN POR PAQUETES

La red de conmutación por paquetes fue diseñada para cumplir con los requerimientos de transmitir tráfico de naturaleza rafagoza. En este tipo de comunicación, la información es fragmentada en bloques de tamaño moderado denominados paquetes. Estos paquetes son capaces de moverse a través de la red gracias a un encabezado que contiene las direcciones de origen y destino. Cuando un router recibe un paquete, examina el encabezado del mismo y lo reenvía al router apropiado. Esta técnica se denomina almacenamiento y envío (store-and-forward) y se hace uso de ella hasta que el paquete llega al segmento de red adecuado o se pierda [1].

En este tipo de comunicación, el nodo puede enviar paquetes de distintos usuarios usando los mismos recursos, es decir, varios tipos de comunicaciones comparten los mismos recursos. Por esta razón la multiplexación de diferentes conexiones compartiendo recursos puede causar retardo y pérdida de paquetes [1].

En las redes de conmutación de paquetes se hace distinción de dos modos de operación: modo orientado a la conexión y modo no orientado a la 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 no orientado a la conexión, cada paquete es tratado de manera independiente. Los paquetes se envían con su 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 [1].

Hay muchas características que deben ser consideradas para realizar una comparación entre redes de conmutación de circuitos y redes de conmutación de paquetes, la Tabla 2 - 1 muestra las más importantes.

Tabla 2 - 1. Comparación de Conmutación de Circuitos y de Paquetes

Conmutación de circuitos	Conmutación de paquetes
Canal de comunicación dedicado	No hay canal de comunicación dedicado
Transmisión de datos continua	Transmisión de paquetes a ráfagas
Mensajes no son almacenados	Los paquetes se almacenan hasta que son enviados
La ruta es establecida para toda la conversación	Se establece una ruta por paquete
Retardo de transmisión constante e insignificante	Retardo en la transmisión de paquetes en función del estado de la red
Señal de ocupado si el medio está ocupado	El emisor será notificado si el paquete no es enviado
La sobrecarga implica ausencia de una ruta física y por tanto no hay establecimiento de llamada	La sobrecarga implica aumento en retardos
Nodos de conmutación electromecánicos o digitales	Nodos de conmutación digitales

Ancho de banda de transmisión fijo	Ancho de banda de transmisión dinámico
Sin encabezados de bits	Encabezados de bits en cada paquete

2.2.1 INTERNET

Definir hoy en día Internet es algo más problemático de lo que era años atrás. Su definición varía de persona a persona. Por ejemplo, podríamos definirlo como una colección de redes de computadoras basadas en un conjunto específico de estándares de red, los TCP/IP. Otros usuarios, podrían definir Internet como una colección global de diversos recursos o como una comunidad electrónica de gente. Incluso otros, cuya única experiencia con Internet es el uso de la World Wide Web (www), podrían decir que Internet y la www son sinónimos y que, por consiguiente, Internet es la World Wide Web. En consecuencia, definir Internet depende de la perspectiva con la que se vea. Independientemente de la definición o perspectiva, Internet conecta redes de computadoras individuales, autónomas, heterogéneas y les permite funcionar y parecer una sola red global [8].

Internet es una red de cientos de miles de computadoras interconectados entre sí, que ofrecen acceso y comparten información a través de un lenguaje común. Desde inicio del año 2006 ya era la red de computadoras más grande que existe en el mundo, con unos 1000 millones de usuarios que se conectan por las redes telefónicas fijas y/o móviles, por cable, fibra óptica o por satélite y transmiten toda clase de información [7].

La palabra Internet es el resultado de la unión de dos términos: Inter, que hace referencia a enlace o conexión y Net (Network) red, que significa interconexión de redes. Es decir, Internet no es otra cosa que una conexión integrada de redes de computadoras o redes interconectadas. Por medio de un conjunto de componentes de hardware y software se crearon y continúan desarrollándose

numerosos servicios y aplicaciones que son aprovechados para diferentes fines [7].

Gracias a la arquitectura TCP/IP, para el usuario final, esa gigantesca red heterogénea aparece como una única red. La arquitectura TCP/IP está dividida en cuatro capas, iniciando por la capa física, la cual se encarga de realizar la conexión con el medio físico, seguida de la capa Internet, cuyo protocolo principal es el IP (Internet Protocol) que se ejecuta tanto en los routers, gateways y estaciones de trabajo. Al mismo nivel, encontramos otros protocolos complementarios como el ARP (Address Resolution Protocol) o el ICMP (Internet Control Messaging Protocol). Por encima del nivel de Internet se encuentra la capa de transporte que ofrece dos tipos de servicios a las capas del nivel superior: fiable y no fiable. Del primer tipo de servicio se encarga el protocolo TCP (Transmission Control Protocol), mientras que el segundo o proporciona el protocolo UDP (User Datagram Protocol). Finalmente sobre la capa de transporte están las aplicaciones finales que utilizan los usuarios [7]. La Figura 2 - 2 y la Tabla 2 - 2 resumen la arquitectura TCP/IP.

<table>
<tr><td colspan="2">RTP
RTCP</td><td colspan="2">BGP, FTP, HTTP, SMTP, TELNET</td><td>Aplicación</td></tr>
<tr><td colspan="2">UDP</td><td colspan="2">TCP</td><td>Transporte</td></tr>
<tr><td>IP</td><td colspan="2">ARP</td><td>ICMP</td><td>Internet</td></tr>
<tr><td colspan="4">Redes locales</td><td>Física</td></tr>
</table>

Figura 2 - 2. Capas del Modelo TCP/IP y sus Protocolos

La mayor ventaja de esta arquitectura es su independencia respecto de la red física subyacente, de manera que los paquetes IP son capaces de viajar por cualquier tipo de red. Precisamente, esta flexibilidad ha convertido a TCP/IP en el principal factor que facilitó la convergencia de redes y servicios [7].

Tabla 2 - 2. Resumen de las Capas y Funciones TCP/IP

Aplicación
Sirve como interfaz de comunicación proporcionando servicios de aplicación específicos. Ejemplos: correo electrónico, terminal virtual, transferencia de archivos, www
Transporte
Definido por dos protocolos: TCP y UDP
Internet
El centro y fundamento es el Protocolo de Internet. El usuario de transferencias envía mensajes del host origen al host destino. Se trata de un servicio de Datagrama sin conexión. La selección de la ruta se basa en alguna métrica. Usa direcciones IP como mapa de caminos para localizar un host dentro de Internet. Parte integral es el Protocolo de Control de Mensajes de Internet (ICMP), que usa un datagrama IP para llevar mensajes respecto al estado del ambiente de las comunicaciones
Física
Conecta un nodo al hardware de red local. Realiza una conexión con el medio físico. Usa un protocolo específico para tener acceso al medio. Coloca datos en bloques.

En los inicios de Internet y hasta la década de los 80, los equipos conectados a la red se identifican a través de una dirección IP, una larga serie de números que los usuarios tienen que recordar para cada sitio de Internet que desean visitar [7].

En 1984, los estadounidenses Jon Postel, Paul Mockapetris y Craig Patrige introducen el Sistema de Nombres de Dominio, que simplifica considerablemente el uso de Internet, sustituyendo las direcciones IP por nombres de dominio. Por ejemplo:

La dirección IP 74.125.227.82 es la dirección de www.google.com

En la actualidad, se sigue utilizando este sistema y los nombres de dominio han adquirido para cualquier empresa un valor muy similar al de una marca comercial.

2.2.1.1 EL PROTOCOLO IP

El protocolo IP (Internet Protocol) es el protocolo de nivel de red y ofrece un servicio de garantía de tipo best effort (mejor esfuerzo) [7].
Las redes IP son redes tipo datagrama, es decir, que la información se divide en fragmentos más pequeños denominados datagramas o paquetes que se envían de manera independiente por la red. Además, ofrece un servicio sin conexión caracterizado porque la información se envía sin un diálogo previo entre los extremos que garantice que la comunicación tendrá recursos suficientes para llevarse a cabo [7].

El hecho de que cada paquete sea tratado de manera independiente en cada uno de los nodos de la red tiene importantes implicaciones, algunas de las cuales resultan críticas a la hora de transportar tráfico en tiempo real como es el caso de la voz. En primer lugar, como cada paquete sigue, a priori, un camino diferente a través de la red es posible que lleguen en un orden distinto con el que se generaron, por lo que en el destino habrá que reordenarlos adecuadamente.

También es posible que algún paquete se pierda porque alguno de los nodos o enlaces que atraviese esté fuera de funcionamiento o congestionado [7].

El paquete es la unidad mínima de información que se trabaja a nivel IP. Dado que los paquetes viajan de manera independiente por la red, en todos y cada uno de ellos se incluye información de direccionamiento que permitirá a cada nodo enviar por el camino óptimo en función de su destino. Esta información útil en el encaminamiento de los paquetes son las direcciones origen y destino [7].

Por otro lado, los paquetes resultan de la fragmentación de los datos de usuario originales. Cada fragmento o paquete viajará por la red por un camino diferente y llegarán al destino desordenados. Sin embargo, el protocolo IP tiene la capacidad de reconstruir la secuencia original de los paquetes en el receptor [7].

Uno de los problemas de los protocolos de enrutamiento, son la posible creación de lazos o bucles, es decir, que un paquete viaje indefinidamente. Este peligro, si acontece frecuentemente, podría llegar a provocar la congestión de la red. La solución es incluir un contador de saltos (TTL, Time To Live) que marcará el límite de retransmisiones de los paquetes por la red. Este contador, con el valor inicial de 255, se va decrementando en cada nodo de la red hasta alcanzar el valor cero cuando será descartado por el nodo correspondiente [7].

2.2.1.2 PROTOCOLOS DE TRANSPORTE: TCP Y UDP

Los servicios proporcionados por la capa de transporte consisten, básicamente, en el transporte de bits entre dos aplicaciones de red, de ahí su nombre. A este nivel, la topología de la red y los problemas que puedan acontecer en aspectos de enrutamiento o congestión, son totalmente transparentes y la capa de red se encarga de solucionarlos [7].
El nivel de transporte en Internet ofrece dos tipos de servicios:

- Si las aplicaciones requieren la entrega garantizada de los datos al destinatario, sin errores, pérdidas ni datos duplicados, deberá utilizar un servicio orientado a la conexión que en redes IP es ofrecido por TCP (Transport Control Protocol).
- Por el contrario, si las aplicaciones aceptan un servicio menos fiable caracterizado por el envío independiente de mensajes, para ello se emplea el servicio no orientado a la conexión que proporciona UDP (User Datagram Protocol).

El protocolo TCP proporciona un servicio fiable orientado a la conexión garantizando un flujo ordenado de bytes extremo a extremo con independencia del tipo y del número de redes que atraviesen los paquetes. TCP identifica a cada proceso de usuario mediante una dirección IP y un número de puerto, cuya

asignación es estática para los servidores y dinámica para los clientes. A este conjunto de dirección IP y número de puerto se le llama dirección de transporte o TSAP (Transport Service Access Point). El número de puerto es un entero comprendido entre 0 y 65535 y se agrupan en dos categorías: los puertos bien conocidos, entre el 0 y el 1023 (la Tabla 2 - 3 muestra los puertos TCP bien conocidos más utilizados) y los puertos de usuarios mayores del 1023 [7].

Tabla 2 - 3. Puertos TCP bien conocidos más utilizados

Puerto	Aplicación	Descripción
20	FTP-Data	Transferencia de datos FTP
21	FTP	Diálogo en transferencia de ficheros
23	TELNET	Login remoto
25	SMTP	Correo electrónico
110	POP3	Servidor de correo
119	NNTP	News

El protocolo TCP lleva a cabo un control de errores y de flujo extremo a extremo, además de un control preventivo de la congestión que consiste en modificar, en función de las condiciones de la red, el tamaño de las ventanas deslizantes utilizadas en el control de errores y en el control de flujo. Todas estas características van a influir en el formato del paquete TCP [7].

Las principales características de TCP se resumen de la siguiente manera [8]:

- Protocolo orientado a la conexión.
- Proporciona una transmisión confiable de datos mediante detección y corrección de errores extremo a extremo.
- Garantiza que los datos sean transferidos a través de una red de manera exacta y en el orden apropiado.
- Retransmite cualesquiera datos no recibidos por el nodo destino.
- Ofrece garantía contra duplicación de datos entre los nodos emisor y receptor.

Por otro lado, el protocolo UDP ofrece un servicio no fiable, no orientado a la conexión, a las aplicaciones de red. Se utiliza en aplicaciones sencillas que no

requieran una alta fiabilidad o en entornos locales de alta fiabilidad con el fin de reducir los retardos asociados a TCP. La identificación de los procesos UDP es similar a la de los procesos TCP: dirección IP y número de puerto [7]. En la tabla 2 - 4 se muestran los puertos UDP más utilizados.

Tabla 2 - 4. Puertos UDP más utilizados

Puerto	Servicio	Descripción
7	Echo	Devuelve el paquete al emisor
9	Discard	Descarta el paquete
123	NTP	Network Time Protocol (Sincronización de relojes)
161	SNMP	Usado para recibir consultas de gestión de la red

Debido a la simplicidad del protocolo, el formato del paquete UDP es muy sencillo tal y como se muestra en la Figura 2 - 3.

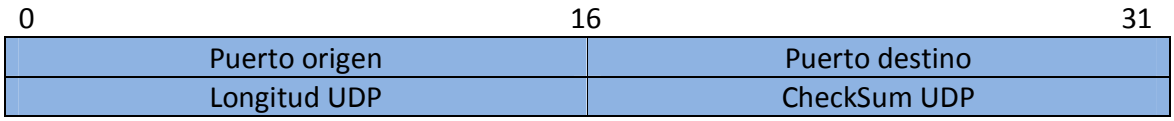


Figura 2 - 3. Formato de Paquete UDP

Las principales características de UDP se resumen de la siguiente manera [8]:

- Protocolo no orientado a la conexión.
- Proporciona servicio de datagrama no confiable (ninguna detección o corrección de errores extremo a extremo).
- No retransmite ningún dato que no haya sido recibido.
- Requiere poca sobrecarga.

2.2.1.3 PROTOCOLOS RTP/RTCP

Las transmisiones de VoIP, por definición, usan el protocolo IP, aunque no es muy adecuado para transmisiones de voz. Aplicaciones en tiempo real como la voz y el video requieren garantía de conexión con características de retardo aceptable.

Para proveer realimentación en la calidad del enlace de transmisión, los protocolos RTP/RTCP, desarrollados por la IETF, son los indicados. El protocolo de

transporte en tiempo real (RTP) transporta muestras digitalizadas de información en tiempo real, y el protocolo de control de tiempo real (RTCP) proveen los mecanismos de retroalimentación de la calidad.

El protocolo de transporte en tiempo real provee funciones de transporte de red extremo a extremo para aplicaciones que transmiten paquetes de audio o video en tiempo real sobre redes unicast o multicast. Fue desarrollado por la IETF y es usado con los protocolos H.225 de la recomendación H.323 para proveer una comunicación confiable. Por sí solo, RTP no reserva recursos ni garantiza QoS para servicios en tiempo real. El transporte de paquetes es acompañado por un protocolo de control (RTCP) que monitorea la entrega de datos de una manera escalable. RTP define los siguientes elementos [10]:

Carga útil RTP: La carga del medio es transportada por RTP en paquetes, tales como muestras de audio o datos de video [10].
Paquete RTP: un paquete consiste de una cabecera RTP fija, una posible lista vacía de fuentes y los datos, como se muestra en la Figura 4-5 [10].
Paquete RTCP: Un paquete de control RTP consiste en una cabecera similar al paquete RTP, seguido de elementos estructurados que dependen del tipo de paquete RTCP [10].

0							7
V	P	X	Cuenta CSRC				
M			Tipo de carga útil				
			Número de Secuencia (2 bytes)				
			Estampa de tiempo (4 bytes)				
			SSRC (4 bytes)				
			CSRC (0 – 60 bytes)				

Figura 2 - 4. Cabecera RTP

Los campos de la cabecera RTP desempeñan ciertas funciones [10].

V (versión): Identifica la versión de RTP

P (Relleno): Cuando se activa, el paquete o más rellenos de octetos adicionales al final que no son parte de la carga útil.

X (extensión del bit): Cuando se activa la cabecera es seguida por, exactamente una extensión con un formato definido.

Cuenta CSRC: Contiene el número de identificadores CSRC que siguen la cabecera.

M (Marcador): La definición del marcador es definida por un perfil. Está destinado a permitir eventos significantes como los límites de la trama que se marcarán en el stream de paquetes.

Tipo de carga útil: Identifica el formato de la carga RTP y determina su interpretación por la aplicación. Un perfil especifica un mapeo estático por defecto de os códigos de tipo de carga para los formatos de carga.

Son cinco los tipos de paquetes RTCP definidos en la especificación RTP; informe de receptor (RR), reporte del emisor (SR), descripción de origen (SDES), manejo de membresías (BYE), y definido por aplicación (APP). Estos paquetes tienen una estructura en común que se ilustra a continuación [14].

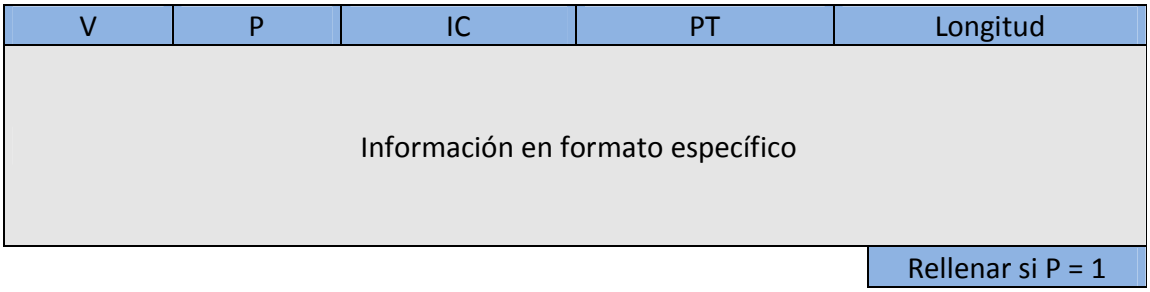


Figura 2 - 5. Cabecera RTCP

V (Versión): Este número siempre es 2 para la versión actual de RTP. No hay planes para introducir una nueva versión, y las versiones anteriores ya no están en uso [14].

Relleno: Este campo indica que el paquete ha sido rellenado más allá de su tamaño natural. Si esta opción es activada, uno o más bytes de relleno serán

añadidos al final de este paquete, y el último byte contiene un conteo de números de paquetes de relleno añadidos [14].

IC (Conteo de objetos): El campo de conteo de objetos es usado para indicar el número de objetos incluidos en el paquete. Hasta 31 objetos pueden incluirse en un paquete RTCP. Un IC igual a cero indica que la lista de objetos está vacía (esto no implica que el paquete en realidad esté vacío). Tipos de paquetes que no necesiten un conteo de objetos usarán este campo con otros propósitos.
PT (Tipo de paquete): Aquí se identifica el tipo de información dentro del paquete. Cinco paquetes estándar son definidos en la especificación RTP; otros tipos son especificados; otros quizá serán definidos en un futuro próximo.

Longitud (Length): En este campo se define la longitud del contenido del paquete siguiendo las cabeceras comunes. Es medido en unidades de 32 bits ya que todos los paquetes RTCP son múltiplos de 3 bits en longitud, entonces el conteo de octetos podría permitir la posibilidad de inconsistencias. Cero es un valor valido, indicando que el paquete consiste de una cabecera de solo cuatro octetos.

2.3 Red de Voz sobre el Protocolo de Internet

Las redes de voz sobre IP, utilizan las redes de datos para transportar voz, por tanto, el primer paso será, la digitalización de la señal de voz y su conversión en paquetes. Si, además, se desea prestar el servicio de telefonía será necesario ofrecer todas las funciones propias de una red telefónica, tales como la señalización de llamada, entre otras funciones avanzadas [7].

Las redes de VoIP operan bajo el principio de funcionamiento de una red de datos, con algunos componentes adicionales que soportan funcionalidades de telefonía. Así, además de routers y switches en una red podemos encontrar teléfonos IP, gateways y equipos encargados del control de llamadas y señalización [7].

Los teléfonos IP son los terminales de usuario que realizan la función de digitalización, paquetización y señalización en una llamada telefónica. Un teléfono IP tiene el mismo aspecto que un teléfono convencional aunque puede incluir elementos adicionales como una pequeña pantalla y un teclado para navegar por la web. Otra característica que lo diferencia de un teléfono convencional es su conexión a la red. En efecto, los teléfonos convencionales disponen de un conector RJ-11 a través del que se conectan a la red telefónica o la ISDN. Sin embargo, los teléfonos IP son equipos de datos y, por ello, su conector es del tipo RJ-45, el tipo de una red de datos Ethernet. Un tipo especial de teléfonos son los softphones. Se trata de un software especial que se ejecuta en una computadora y que permite al usuario utilizar la computadora como si de un teléfono IP se tratara [7].

2.4 Redes Convergentes

La industria de las telecomunicaciones está en constante cambio, esto se debe al surgimiento de nuevas tecnologías que proporcionan nuevos y más atractivos servicios. Acompañado de este constante cambio tecnológico, la tendencia en las redes de telecomunicaciones está orientada hacia la convergencia de redes.

La convergencia de redes tiene como objetivo central, lograr que viejos y nuevos servicios puedan ser integrados en una misma infraestructura de red, minimizando el número de protocolos en la capa de red y combinando el transporte de diversos tipos de tráfico a través de un simple núcleo de red multiservicio común, como se muestra en la Figura 2-6. De esta manera, para el usuario será transparente comunicarse de una red PSTN, a una red de telefonía celular o a una conexión de red de banda ancha [4].

Son muchos los beneficios que proporciona el uso de una red convergente, los cuales van desde reducción de costos por conceptos de administración,

mantenimiento e infraestructura hasta ofrecer servicios más atractivos como la transmisión de múltiples medios (voz, video, datos).

En la actualidad la red que resulta de esta convergencia es Internet, con las capacidades suficientes para cubrir los beneficios mencionados con anterioridad. Sin embargo, Internet proporciona un servicio de mejor esfuerzo y por tanto, la calidad de servicio no está garantizada. Por otro lado, las aplicaciones multimedia demandan cierto nivel de QoS.

Derivado de los puntos mencionados con anterioridad, existe una gran preocupación por el desarrollo de mecanismos que garanticen la QoS en redes convergentes.

Figura 2 - 6. Esquema de Red Convergente o Multiservicio

Capítulo 3
Tecnologías de Voz sobre IP

En la década de los 90, un grupo de personas perteneciente al entorno de la investigación, tanto de instituciones educativas como empresariales, comenzaron a mostrar un cierto interés por transportar voz y video sobre redes IP, especialmente a través de Intranets corporativas e Internet. Esta tecnología es conocida hoy en día como VoIP y es el proceso de dividir el audio en pequeños fragmentos a través de una red IP y re ensamblar esos fragmentos en el destino final permitiendo de esta manera que los usuarios pueda comunicarse [2].

La idea de transmitir voz sobre el protocolo IP no es nueva, existen patentes y publicaciones de investigación que datan de varias décadas.

En 1995, una pequeña compañía llamada Vacoltec anunció el lanzamiento del primer teléfono software para Internet. Este software era únicamente útil para entablar una comunicación de PC a PC y para ello necesitaba hacer uso de diversos requisitos hardware tales como micrófono, altavoces, tarjeta de sonido y módem. Sin embargo, esta alternativa a la comunicación telefónica tradicional fue comercialmente un fracaso ya que las conexiones a Internet que se disponían ofrecían un ancho de banda muy escaso [2].

Durante los años siguientes, la tecnología asociada a las redes de datos y las comunicaciones continuó mejorando, en 1998 se dieron definitivamente los primeros pasos desde un punto de vista comercial. En este año diversas compañías lanzaron al mercado adaptadores que permitían hacer uso de los teléfonos tradicionales en un entorno VoIP. Este hecho, facilitó el acercamiento de los clientes al uso de la tecnología VoIP, por lo que algunas empresas importantes se lanzaron al mercado ofreciendo productos y servicios relacionados con esta

tecnología. Durante el año 1998 la tecnología VoIP alcanzaba ya el 1% del tráfico total de voz: su carrera había comenzado [2].

En 1999, compañías dedicadas a las redes de datos tales como Cisco crearon las primeras plataformas destinadas a empresas capaces de tratar con tráfico VoIP. Esto supuso un nuevo impulso a la VoIP ya que comenzó a implantarse en muchas empresas. La consecuencia directa fue que la VoIP alcanzara en el año 2000 más del 3% del tráfico total de voz [2].

Las redes de datos siguieron mejorando en años venideros, y alrededor del año 2005 ya era fácil para cualquier persona de países desarrollados conseguir una conexión a Internet que cumpliera los requisitos mínimos para ofrecer una buena calidad de voz y una comunicación fiable a través de VoIP, reduciendo al mínimo las posibles interrupciones que se pudieran producir durante la conversación [2].

Esto supuso otro gran impulso a la VoIP y provocó que al día de hoy existan muchas soluciones que hacen uso de esta tecnología. Un ejemplo claro es Asterisk, una central telefónica de software libre que se distribuye bajo licencia GPL. Este producto, soportado comercialmente por Digium, se ha convertido en pocos años en una de las soluciones IP más extendidas en diversos ámbitos, tales como empresarial o educativo. Otro ejemplo es Skype, que fue creado por dos jóvenes universitarios en el año 2003. A diferencia de Asterisk, Skype hace uso de un protocolo privado que no está basado en un estándar, lo que a largo plazo se piensa que limitará a sus usuarios. En la actualidad, Skype se puede emplear en multitud de plataformas y su uso se encuentra también ampliamente extendido [2].

Son muchos los beneficios que proporciona la transmisión de voz sobre el protocolo IP, dos de los más importantes son: la reducción de costos de las llamadas larga distancia y las nuevas funcionalidades o servicios más atractivos que proporciona al usuario [2].

3.1 SISTEMAS DE VOZ SOBRE IP

Las implementaciones actuales de VoIP tienen principalmente dos tipos arquitecturas, basadas en H.323 y en el protocolo de inicio de sesión (SIP). H.323 fue ratificado por la unión internacional de telecomunicaciones (ITU-T), y está conformado por un conjunto de protocolos para soportar voz, video y datos sobre una red conmutada por paquetes. SIP (RFC 3261) es un protocolo de control de señalización de la capa de aplicación para crear, modificar y terminar sesiones con uno o más participantes [1].

Ambas arquitecturas consisten de tres componentes lógicos principales; una terminal, servidor de señalización y un gateway. Las actuales implementaciones de H.323 y SIP no proveen calidad de servicio (QoS) , sin embargo, VoIP es uno de los servicios más sensibles a la QoS y demanda estrictos niveles del mismo. El nivel de calidad de servicio en VoIP depende de varios parámetros, tales como: ancho de banda, retardo extremo a extremo (One Way Delay, OWD), jitter, pérdida de paquetes (Packet Loss Rate, PLR), tipo de CODEC, longitud de los datos de voz y tamaño de “de-jitter buffer”. En particular, OWD, jitter y PLR tienen un impacto importante [1].

Todo sistema VoIP cuenta con tres componentes principales: emisor, receptor y medio o canal.

- *Emisor:* Muestrea la señal de voz y le asigna un número de bits por cada muestra, creando un flujo constante de bits. Finalmente, a un determinado número de muestras le agrega un encabezado IP para generar un conjunto de paquetes que serán enviados a través de la red IP.

- *Receptor:* El encabezado de los paquetes son removidos y las muestras de voz son extraídas. El conjunto de muestras de voz deben ser presentadas

al decodificador de tal que sean convertidas a su formato original en una señal audible.

- *Medio:* El medio es el canal de comunicación, por el cual el flujo de paquetes de voz se transportaran para alcanzar su destino final. En una comunicación de voz sobre el protocolo de Internet, el medio más utilizado es Internet.

3.2 CODIFICACIÓN DE LA VOZ

El proceso de convertir señales analógicas a información digital e información digital a señales analógicas se realiza mediante un codificador y decodificador (CODEC), respectivamente; además de realizar la conversión analógico-digital (A/D) y digital-analógico (D/A), comprime la secuencia de datos, como se ilustra en la Figura 3 - 1.

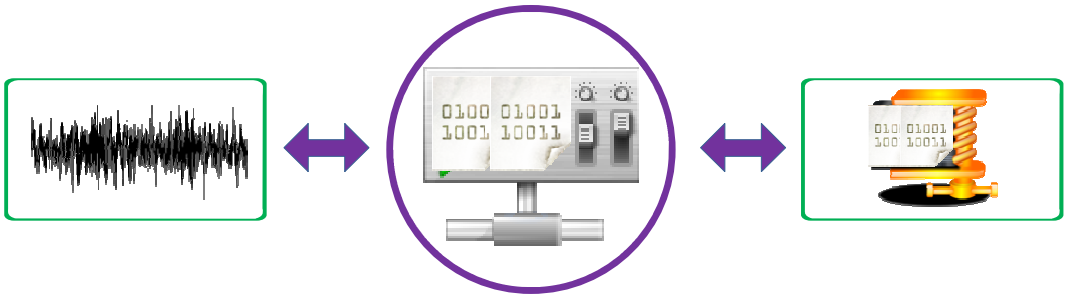


Figura 3 - 1. Codificador-Decodificador de Voz

El proceso de codificación se lleva a cabo mediante tres etapas fundamentales: muestreo, cuantificación y codificación, estos procesos se discutirán a detalle en las sub-secciones 3.2.1, 3.2.2 y 3.2.3, respectivamente.

Cada CODEC tiene principalmente dos atributos: tamaño de muestra (sample size) e intervalo de muestra (sample interval). El primero determina el número de bits de la muestra. El segundo atributo indica el intervalo de tiempo o cada cuanto tiempo se genera una muestra (frecuencia de muestreo). Conocido el tamaño de

cada muestra codificada y la frecuencia de muestreo, es posible calcular el ancho de banda que utiliza el CODEC [2].

Los CODECs de forma de onda son utilizados a altas tasas de datos y proporcionan muy buena calidad de voz. Los paramétricos operan a tasas de datos muy bajas pero proporcionan baja calidad en la voz. Los híbridos usan técnicas de los dos anteriores y proporcionan buena calidad a una tasa de datos intermedia [1].

Por lo tanto un parámetro a considerar a la hora de elegir un CODEC es la tasa de transmisión. Si la tasa es alta, la compresión del CODEC será baja y por tanto se espera una buena calidad en la voz y se requiere un mayor ancho de banda para transmitirlo. Por otro lado, si la tasa de transmisión es baja, la compresión es baja y no se espera baja calidad en la voz y se requiere bajo ancho de banda para poder ser enviada por la red. Por tanto existe un compromiso a la hora de elegir un CODEC que proporcione mayor o menor calidad de voz, ya que no siempre es tan importante un alto grado de calidad. Por ejemplo, en el caso de los humanos nuestros oídos tienen unos límites a partir de los cuales no percibe mejoras en la calidad de la voz y por consiguiente, es más apropiado elegir un códec de tasa de compresión media para mantener una calidad de voz aceptable y hacer uso más optimo del ancho de banda, y de esta manera, permitir un mayor número de llamadas de VoIP simultáneamente [2].

La Tabla 3 - 1 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, además de que se requiere mayor procesamiento [7].

Tabla 3 - 1. Algunos de los CODECs más comunes para Telefonía

CODEC	Ancho de banda (kHz)	Frecuencia de Muestreo (kHz)	Duración muestras/ tramas (ms)	MOS	Aplicación
G711 (PCM)	64	8	0.125	4.4	Telefonía
G721 (ADPCM)	32	8	0.125	4.2	Telefonía
G722 (SB-ADPCM)	48 / 56 / 64	16	0.250		Videoconferencia
G728 (LD-CELP)	16	8	0.625	4.2	Telefonía / Videoconferencia
G729 (CS-ACELP)	8	8	10	4.1	Telefonía
G723.1 (MP-MLQ)	6.3	8	30	3.9	Telefonía Internet
G723.1 (ACELP)	5.3		30	3.6	Telefonía Internet

Algunos esquemas de compresión, tales como G729 y el G723.1, proporcionan una tasa baja de bits, reduciendo considerablemente la necesidad de ancho de banda, por lo que resultan apropiados para transmitir voz sobre Internet. Por ejemplo, el G723.1 incluye un sistema de compresión basado en la supresión de silencios y detección de actividad e voz (VAD-Voice Activity Detection).

Estos esquemas de codificación intentan reproducir la voz de manera subjetiva de la señal, más que la forma de onda original, pero son muy sensibles a la pérdida de paquetes o jitter, por lo que emplean, técnicas de interpolación para reducir al mínimo los efectos mencionados [7]. *El proceso de conversión A/D y D/A se realiza mediante tres técnicas principales: por codificación de forma de onda, por codificación paramétrica (basada en modelos matemáticos) o vocoders y modelos híbridos que combinan ambas técnicas anteriores,* estas técnicas se discutirán a detalle en las sub-secciones 3.2.4, 3.2.5 y 3.2.6, respectivamente [7].

3.2.1 MUESTREO

El muestreo o discretización de una señal analógica consiste en elegir o seleccionar valores que toma dicha señal a lo largo del tiempo, únicamente en instantes de tiempo discreto, como se ilustra en la Figura 3 - 2. Estos instantes de tiempo o intervalo de muestreo debe seleccionarse con cuidado para que la pérdida de información que supone el muestreo no represente una pérdida importante [7].

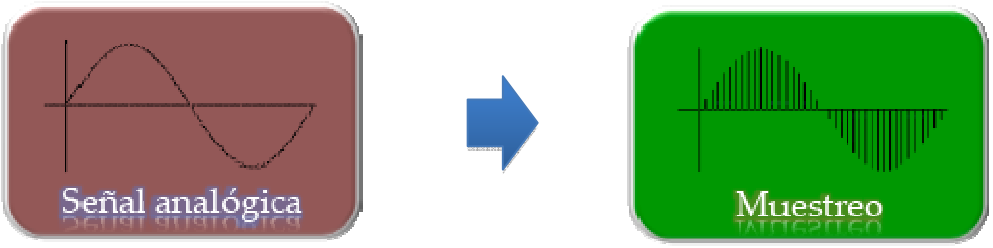


Figura 3 - 2. Muestreo

3.2.2 CUANTIFICACIÓN

El resultado del muestreo es un conjunto de valores de la señal analógica tomados en ciertos instantes de tiempo discretos, pero la señal sigue siendo continua en amplitud y es necesario discretizarla también en este dominio. La cuantificación consiste en asignar a cada una de las muestras continuas un valor discreto de uno de los M posibles y mantener ese valor de la señal hasta el siguiente instante de muestreo, tal como se muestra en la Figura 3 - 3. La diferencia entre la muestra sin cuantificar y la salida cuantificada se denomina error de cuantificación (ruido de cuantificación) [7]. A la salida cuantificada se le conoce como señal digital.

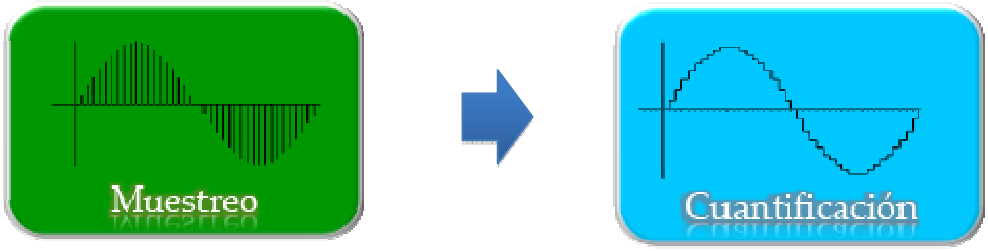


Figura 3 - 3. Cuantificación

3.2.3 CODIFICACIÓN

Una vez que la señal ya se presenta en 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 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 ($256=2^8$) [7].

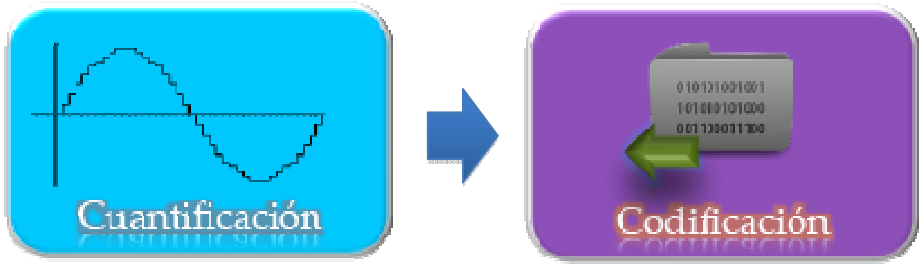


Figura 3 - 4. Codificación

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 (8000 x 8) que coincide con la velocidad asignada a un canal básico de ISDN. En Estados Unidos, al codificar con 7 bits en lugar de 8, resultan 56 kbps, la velocidad empleada para un canal digital básico [7].

3.2.4 CODECS DE FORMA DE ONDA

La característica principal de este tipo de codificador es preservar la forma general de la señal de entrada. Su desempeño puede ser medido de manera efectiva en términos de su relación señal a ruido (SNR). El primer estándar mundial para la codificación de la voz fue el CODEC G.711 de forma de onda, basado en el método de codificación PCM.

El método de codificación PCM (Pulse Code Modulation) es el método más usado para codificar una señal de voz analógica a una cadena de bits (0s y 1s) en formato digital. Todas las técnicas de muestreo hacen uso del teorema de Nyquist, que básicamente establece que si se realiza un muestreo al doble del ancho de banda de una señal de voz, lograrás una transmisión de voz de buena calidad [3].
El proceso de PCM se realiza de la siguiente manera:

- Formas de onda analógicas son sometidas a un filtro de frecuencia de voz para filtrar todo lo que exceda los 4000 Hz. Esto se hace para limitar la cantidad de ruido que existe en la red de voz.
- De acuerdo al ya mencionado “Teorema de Nyquist”, se necesitaría muestrear a 8000 muestras por segundo para lograr una aceptable calidad en la transmisión de voz.
- Después de que la onda es muestreada, es convertida a una forma digital discreta. Esta muestra es representada por un código que indica la amplitud de la onda en el instante en que la muestra fue tomada. PCM usa ocho bits para el código y un método de compresión logarítmico que asigna más bits para señales de baja amplitud.

Si se multiplica esos 8 bits por las 8000 muestras por segundo, obtendrás 64000 bits por segundo (bps). Lo básico para la infraestructura telefónica es 64000 bps (64 kbps).

3.2.5 CODECs PARAMÉTRICOS O VOCODERS

Los codificadores dentro de esta categoría asumen que la señal de voz se genera bajo un modelo, el cual es controlado por algunos parámetros. Durante la codificación, los parámetros del modelo son estimados de la señal de entrada, siendo dichos parámetros los datos que se codifican y transmiten. Este tipo de codificación no realiza algún intento de preservar la forma original de la señal de entrada, es por eso que SNR es una métrica de calidad de poca utilidad. La calidad percibida de la voz decodificada está directamente relacionada con la precisión y la complejidad del modelo. Debido a esta limitante, el funcionamiento del codificador es específico a un tipo de señal, teniendo un pobre desempeño con señales que no son de voz. Los codificadores más conocidos de este tipo son los codificadores por predicción lineal (LPC)
Los CODECs paramétricos producen señales de muy baja tasa de bits, pero tienen un potencial limitado de calidad de voz. Se han usado mucho sobre todo en aplicaciones de comunicación militar segura.

3.2.6 CODECs HÍBRIDOS

Los CODECs híbridos proporcionan mayor calidad de conversación que los CODECs paramétricos, con proporciones de bits más bajos que los CODECs de forma de onda. Para cumplir este rendimiento, los CODECs híbridos usan una combinación de modelado de la señal de voz y del análisis de forma de onda. Estos algoritmos tienden a ser bastante complejos [9].

Los CODECs híbridos más comunes operan en el dominio de tiempo usando técnicas de predicción lineal de análisis-por-síntesis (LPAS). Igual que los CODECs paramétricos, los CODECs híbridos modelan una señal de estímulo y un filtro. El componente del filtro es similar al modelado en los CODEC paramétricos,

pero la codificación de la señal de estímulo es más sofisticada. Hay tres estrategias principales para codificar la señal de estímulo:

- Estímulo multi-impulso (MPE)
- Estímulo de impulso regular (RPE)
- Predicción lineal de código estimulado (CELP)

Cada una de estas técnicas genera la señal de estímulo de varios modos, pero todas ellas procesan una variedad de señales de estímulo a través del filtro para ver qué estímulo produce la mejor coincidencia con la forma de onda original. Una vez obtenida la mejor coincidencia, el CODEC transmite las variables del filtro y la información sobre la señal de estímulo. La representación de la señal de estímulo es diferente para las estrategias MPE, RPE y CELP [9].

3.3 EMPAQUETADO DE LOS DATOS

Después de la compresión y codificación la trama es empaquetada. El proceso de empaquetado es implementado para que un conjunto de datos de voz pueda ser transmitido y se le añaden la información necesaria para el enrutamiento y el manejo de los paquetes a través de la red IP. Si encapsulamos los paquetes de voz en tramas Ethernet, los encabezados MAC, IP, UDP, RTP, FCS, preámbulo e IPG son necesarios para su transmisión [1]. La longitud de los datos de voz de un paquete IP usualmente depende del algoritmo de codificación utilizado. Por ejemplo, paquetes de voz de 80 bytes corresponden a G.711 mientras que de 20 bytes a G.729 en comunicaciones convencionales de VoIP.

3.4 PROTOCOLOS DE SEÑALIZACIÓN

Las implementaciones actuales de VoIP tienen principalmente dos tipos arquitecturas, basadas en H.323 y en el protocolo de inicio de sesión (SIP). Ambas

arquitecturas consisten de tres componentes lógicos principales; una terminal, servidor de señalización y un gateway.

3.4.1 H.323

Conjunto de protocolos que está designado a operar por encima de la capa de transporte. El estándar H.323 define un conjunto de terminales que proporcionan servicios de comunicación multimedia por redes de conmutación de paquetes que no necesariamente garanticen calidad de servicio.
Los principales componentes definidos en la arquitectura H.323 son los siguientes:

Terminal: Dispositivo final de una red el cual provee comunicación en tiempo real con otra terminal H.323, gateway o MCU.

Gateway: Permite la comunicación entre redes IP y redes conmutadas por circuito.

Gatekeeper: Es una entidad que centraliza la comunicación en una red VoIP H.323. Proporciona servicios de traducción de direcciones y control de acceso a la red, también realiza tareas de administración de ancho de banda y descubrimiento de gateways.

MCU: Permite que tres o más terminales y gateways participen en una videoconferencia. Consiste de dos partes, un MC y un MP. El MC controla tres o más terminales que participan en una conferencia multipunto y es capaz de negociar con todas las terminales para lograr niveles comunes de comunicación. El MP provee procesamiento centralizado para flujos de audio, video y/o datos en una conferencia.

H.323 es una recomendación que depende de varios estándares para permitir comunicación multimedia en tiempo real. Los principales son [1]:

Señalización de llamada y control: señalización RAS (registro, admisión y estado) protocolo de control de llamada (H.225), protocolo de control de medio (H.245), seguridad (H.265), señalización digital de suscriptor (Q.931), protocolo genérico funcional para soportar servicios suplementarios en H.323 (H.450.1), características suplementarias (H.450.2-H.450.11).

Anexos H.323: Fax en tiempo real sobre H.323 (Anexo D), protocolo cableado para transporte de llamadas señalizadas por multiplexación (Anexo E), tipo de dispositivo final simple SET (AnexoF), conversación de texto y texto SET (Anexo G), seguridad para Anexo F (Anexo J), protocolo de transferencia de hipertexto (HTTP) basado en servicios de control de transporte de canales (Anexo K), protocolo de control de estímulos (Anexo L), protocolo de señalización de túneles (Anexo M),

CODECs de audio: modulación de pulsos codificados (PCM), CODEC de audio 56/64 kbps (G.711), CODEC de audio para 7 kHz a 48/56/64 kbps (G.722), CODECspeech para 5.3 y 6.4 (G.723), CODECspeech para 16 kbps (G.728) y CODECspeech para menos de 64 kbps (H.263).
CODECs de videos: CODEC de video para más de 64 kbps (H.261) y para menos de 64 kbps (H.263).

3.4.2 SIP

SIP es un protocolo de control de capa de aplicación capaz de establecer, modificar y terminar sesiones multimedia o llamadas. Soporta mapeo de nombre y servicios de redirección permitiendo implementaciones ISDN y servicios inteligentes de red de abonado de telefonía. Soporta protocolos tales como el protocolo de reservación de servicios (RSVP), protocolo de transporte en tiempo real (RTP), protocolo de streaming de tiempo real (RTSP), protocolo de anuncio de sesión (SAP) y el protocolo de descripción de sesión (SDP). Los dos componentes

principales de la arquitectura SIP son el agente de usuario (UA) y los servidores de red (servidor de registro, servidor de localización, servidor proxy y servidor de redireccionamiento [1].

- Agente de usuario: es una aplicación de punto final y contiene dos componentes agente de usuario cliente (UAC) y agente de usuario servidor (UAS). El UAC envía peticiones SIP mientras que el UAS las recibe y responde al usuario.
- Servidor de registro: acepta únicamente peticiones de registro usadas por el usuario con el propósito de actualizar una base de datos de localización con información de contacto.
- Servidor proxy: actúa como servidor para los agentes de usuario reenviando peticiones SIP y también actúa como cliente a otros servidores SIP presentando peticiones enviadas a él en nombre de agentes de usuario o servidores proxy.
- Servidor de redireccionamiento: ayuda a localizar agentes de usuario mediante localizaciones alternativas donde el usuario puede ser alcanzado. Proporciona servicio de mapeo de direcciones. Responde a una petición SIP destinada a una dirección con una lista de direcciones.

Existen dos tipos de mensajes SIP; peticiones y respuestas [1].

Respuestas SIP: cada respuesta tiene un código que indica el estado de la transacción. Los códigos de estado son números enteros dentro del rango de 100 a 699 y agrupados en seis clases. Una respuesta con un código de estado de 100 a 199 es considerada provisional y de 200 a 699 es considerada respuesta final [1].

a) 1xx Informativo: petición recibida, continuando con el proceso de petición.
b) 2xx satisfactorio: la acción fue recibida satisfactoriamente, entendida y aceptada.

c) 3xx redireccionamiento: se deben adoptar nuevas medidas para completar la petición. El cliente debe terminar cualquier búsqueda existente e iniciar una nueva.
d) 4xx error de cliente: la petición contenía un error de sintaxis o no pudo ser completada por el servidor.
e) 5xx error de servidor: la petición no pudo ser completada en el servidor debido a un error del mismo. El cliente debe tratar en otro servidor.
f) 6xx falla global: la petición es inválida en cualquier servidor. El cliente debe abandonar la búsqueda.

El primer dígito del código de estado define la clase de respuesta. Los últimos dos no tienen importancia [1].

Peticiones SIP: hay seis tipos. Cada petición SIP contiene un campo, llamado método, el cual define su propósito.

- INVITE (invitación): invita usuarios a participar en una sesión. SIP solo puede manejar invitaciones al usuario y los usuarios aceptan la invitación. SIP puede invitar a usuarios a cualquier tipo de sesión.
- ACK: se usa para notificar la recepción de una respuesta final para una invitación
- CANCEL: cancela transacciones pendientes.
- BYE: abandona sesiones.
- REGISTER: informa a un servidor acerca de la localización actual.
- OPTIONS: consulta a un servidor a cerca de las capacidades, incluyendo cual, método y que protocolo de descripción de sesión puede soportar.

Capítulo 4
Arquitectura H.323

La presente arquitectura describe un conjunto de entidades y terminales que proporcionan servicios de comunicaciones multimedios por redes de paquetes que no necesariamente garantizan QoS.

Las entidades H.323 pueden proporcionar comunicaciones de audio, video y/o datos en tiempo real. El soporte del audio es obligatorio, mientras que el de datos y video es opcional, pero si se soportan es necesario poder utilizar un modo de funcionamiento común especificado, para que puedan interfuncionar todos las terminales que soporten ese tipo de medios.

La red por paquetes por la cual se comunican las entidades H.323, puede ser una conexión punto a punto, un segmento de red único o una interred que tenga múltiples sistemas con topologías complejas.

Las entidades H.323 pueden utilizarse en configuraciones punto a punto, multipunto o de difusión. Pueden interfuncionar con terminales H.310 por la ISDN de banda ancha, con terminales H.320 por la ISDN de banda angosta, con terminales H.321 por la ISDN de banda ancha, con terminales H.322 en redes LAN de calidad de servicio garantizada, con terminales H.324 por la PSTN y redes inalámbricas, con terminales V.70 por la PSTN, y con terminales vocales por la PSTN o por la ISDN utilizando gateways [13].

Las entidades H.323 pueden estar integradas en computadores personales o implementadas en dispositivos autónomos como son los videoteléfonos.

4.1 FLUJOS DE INFORMACIÓN

Los componentes videotelefónicos se comunican mediante la transmisión de flujos de información. Dichos flujos de información se clasifican en flujos de video, audio, datos, control de las comunicaciones y control de la llamada de la siguiente manera:

Señales de audio que contienen señales vocales digitalizadas y codificadas. Para reducir la velocidad binaria media de las señales de audio, se puede proporcionar activación por la voz. La señal de audio va acompañada por una señal de control de audio [13].

Señales de video que contienen video en movimiento digitalizado y codificado. El video se transmite a una velocidad no superior a la seleccionada como resultado del intercambio de capacidades. La señal de video va acompañada por una señal de control de video.

Señales de datos que incluyen imágenes fijas, facsímil, documentos, archivos de computadoras y otros trenes de datos.

Señales de control de medios de las comunicaciones que transfieren datos de control entre elementos funcionales que se comportan como distantes y se utilizan para el intercambio de capacidad, apertura y cierre de canales lógicos, control de modo y otras funciones que forman parte del control de las comunicaciones [13].

Señales de control de la llamada que se utilizan para el establecimiento de comunicaciones, la desconexión de las mismas y otras funciones del control de la llamada.

Los flujos de información descritos anteriormente son acondicionados y enviados a la interfaz de red.

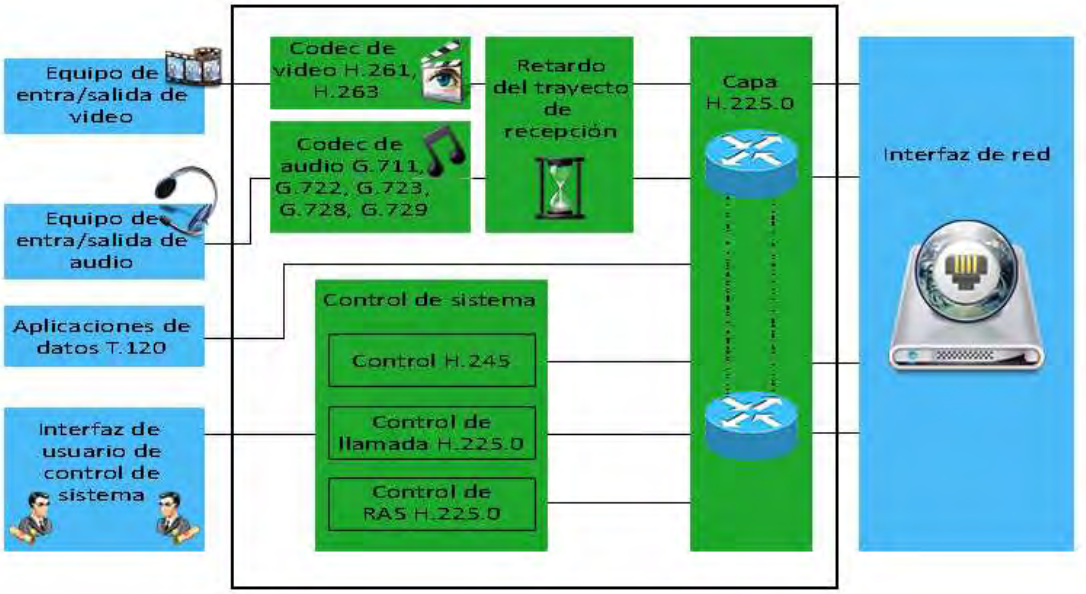


Figura 4 - 1. Equipo Terminal H.323

4.2 ELEMENTOS DEL TERMINAL DENTRO DEL ALCANCE DE LA PRESENTE RECOMENDACIÓN

Los siguientes elementos de la terminal quedan dentro del alcance de la presente Recomendación y, por consiguiente, son objeto de normalización y se definen en la misma [13]:

- El códec de vídeo (H.261, H.263, etc.), que codifica el vídeo a partir de la fuente de vídeo para transmisión y decodifica el código de vídeo recibido, que es la salida hacia una presentación visual del vídeo.
- El códec de audio (G.711, G.729, etc.), que codifica la señal de audio de entrada al micrófono para transmisión y decodifica el código de audio recibido que es la salida hacia el altavoz.
- El canal de datos, que soporta aplicaciones telemáticas tales como pizarras electrónicas, transferencia de imágenes fijas, intercambio de archivos,

acceso a bases de datos, conferencias audiográficas, etc. La aplicación de datos normalizada para conferencia audiográfica en tiempo real es la Recomendación UIT-T T.120. Se pueden utilizar también otras aplicaciones y protocolos mediante la negociación H.245.

- La unidad de control del sistema (H.245, H.225.0), que proporciona la señalización para un funcionamiento adecuado de la terminal H.323. Permite el control de la llamada, el intercambio de capacidad, la señalización de instrucciones e indicaciones y facilita mensajes de apertura y descripción completa del contenido de los canales lógicos.
- La capa H.225.0 (H.225.0), que da formato a los trenes de vídeo, audio, datos y control transmitidos en mensajes de salida hacia la interfaz de la red y recupera los trenes de vídeo, audio, datos y control recibidos de los mensajes que han sido introducidos desde la interfaz de la red. Además, lleva a cabo la alineación de trama lógica, la numeración secuencial, la detección de errores y la corrección de los mismos según conviene a cada tipo de medio.

4.2.1 CODEC DE VIDEO

El códec de video es opcional. Si se dispone de capacidad de video, se hará con arreglo a las exigencias de la presente Recomendación. Todas las terminales H.323 que proporcionen comunicaciones de video deberán ser capaces de codificar y decodificar video de acuerdo con QCIF H.261.

Opcionalmente, una terminal también puede ser capaz de codificar y decodificar video de acuerdo con los otros modos H.261 o H.263. Si un terminal soporta H.263 con CIF o con una resolución mayor, deberá también soportar CIF H.261. Todas las terminales que soporten H.263 deberán soportar QCIF H.263. Los códecs H.261 y H.263 de la red serán utilizados sin corrección de errores BCH y sin alineación de trama en la corrección de errores [13].

Se pueden utilizar también otros códecs de video y otros formatos de imagen mediante la negociación H.245. Más de un canal de video puede ser transmitido y/o recibido de acuerdo con lo negociado a través del canal de control H.245. La terminal H.323 puede, opcionalmente, enviar más de un canal de video al mismo tiempo, por ejemplo, para llevar la imagen del conferenciante y una segunda fuente de video. La terminal H.323 puede recibir, opcionalmente, más de un canal de video al mismo tiempo, por ejemplo, para visualizar los múltiples participantes en una conferencia multipunto distribuida [13].

La velocidad binaria de video, el formato de imagen y las opciones de algoritmo que pueden ser aceptados por el decodificador se definen durante el intercambio de capacidades utilizando H.245. El codificador tiene la libertad de transmitir cualquier cosa que se halle dentro del conjunto de capacidades del decodificador. El decodificador debería tener la posibilidad de generar peticiones de modos determinados vía H.245, pero el codificador está autorizado a ignorar simplemente estas peticiones si no son modos obligatorios. Los decodificadores que indican capacidad para una determinada opción de algoritmo deberán también ser capaces de aceptar trenes binarios de video que no utilicen esa opción [13].

Las terminales H.323 habrán de poder funcionar con velocidades binarias de video, velocidades de trama y, si se soporta más de una resolución de imagen, resoluciones de imagen que pueden ser asimétricas. Esto permitirá, por ejemplo, que una terminal con capacidad de CIF transmita QCIF mientras recibe imágenes CIF.

Cuando se abre uno de los canales lógicos de video, se señaliza al receptor el modo de funcionamiento seleccionado que se ha de utilizar en ese canal en el mensaje openLogicalChannel (apertura de canal lógico) H.245. El encabezamiento dentro del canal lógico de video indica qué modo se utiliza realmente para cada imagen dentro de la capacidad indicada [13].

4.2.2 CODEC DE AUDIO

Todos las terminales H.323 tendrán un CODEC de audio y serán capaces de codificar y decodificar señales vocales de conformidad con la Recomendación G.711. Todos las terminales transmitirán y recibirán ley A y ley μ. Una terminal puede, opcionalmente, ser capaz de codificar y decodificar señales vocales utilizando otros CODECs de audio que se pueden señalizar mediante negociación H.245. El algoritmo de audio empleado por el codificador se obtendrá durante el intercambio de capacidades utilizando H.245. La terminal H.323 debería tener la posibilidad de funcionamiento asimétrico para todas las capacidades de audio que haya declarado dentro del mismo conjunto de capacidades; por ejemplo, debería poder enviar G.711 y recibir G.728 si es capaz de ambas cosas.

La terminal H.323 puede, opcionalmente, enviar más de un canal de audio al mismo tiempo, por ejemplo, para transportar las señales de dos idiomas.

Los paquetes de audio deberán ser entregados a la capa de transporte periódicamente, con un intervalo determinado por la Recomendación de CODEC de audio que se utilice (intervalo de trama de audio). La entrega de cada uno de los paquetes de audio tendrá lugar no más tarde de 5 ms después de un múltiplo completo del intervalo de trama de audio, medido desde la entrega de la primera trama de audio (jitter del audio). Los codificadores de audio capaces de limitar más aun su jitter del audio pueden indicarlo utilizando el parámetro maximumDelayJitter H.245 de la estructura h2250Capability (capacidad h2250) contenida en un mensaje del conjunto de capacidades de terminal, de tal manera que los receptores puedan reducir, opcionalmente, sus memorias intermedias de jitter. Esto no es lo mismo que el jitter de arribo del campo RTCP [13].

4.3 SEÑALIZACIÓN DE CANAL LÓGICO

Cada canal lógico lleva información de un transmisor a uno o más receptores y se identifica mediante un número de canal lógico que es único en cada sentido de la transmisión.

Los canales lógicos se abren y cierran utilizando los mensajes openLogicalChannel y closeLogicalChannel (cerrar canal lógico) y los procedimientos de la Recomendación H.245. Cuando se abre un canal lógico, el mensaje openLogicalChannel describe totalmente el contenido del canal lógico, incluyendo el tipo de medios, el algoritmo utilizado, cualesquiera opciones y cualquier otra información que necesite el receptor para interpretar dicho contenido. Los canales lógicos se pueden cerrar cuando ya no se necesiten. Los canales lógicos abiertos pueden estar inactivos si la fuente de información no tiene nada que enviar [13].

Los canales lógicos se abrirán utilizando el procedimiento siguiente [13]:

La terminal llamante enviará un mensaje *openLogicalChannel*. Si el canal lógico ha de transportar un tipo de medios que utiliza RTP (audio o video), el mensaje *openLogicalChannel* incluirá el parámetro *mediaControlChannel* que contiene la dirección de transporte para el canal RTCP inverso.

La terminal llamado responderá con un mensaje *openLogicalChannelAck*. Si el canal lógico ha de transportar un tipo de medios que utiliza RTP, el mensaje *openLogicalChannelAck* incluirá el parámetro *mediaChannel* que contiene la dirección de transporte RTP para el canal de medios y el parámetro *mediaControlChannel* que contiene la dirección de transporte para el canal RTCP hacia adelante.

Los tipos de medios (tales como datos T.120) que no utilizan RTP/RTCP omitirán los parámetros *mediaControlChannel.*

Si se abre un canal inverso correspondiente para una determinada sesión RTP existente (identificada por el sessionID (ID de sesión) del RTP), las direcciones de transporte del *mediaControlChannel* intercambiadas por el proceso *openLogicalChannel* serán idénticas a las utilizadas para el canal directo. Los valores de *sessionID* 1, 2 y 3 están pre asignados a sesiones principales de audio, video y datos, respectivamente. Incluso el punto extremo subordinado puede abrir canales lógicos para estas sesiones primarias sin negociar el valor de *sessionID* con el punto extremo principal.

El punto extremo principal puede abrir sesiones adicionales con valores particulares de *sessionID* mayores que 3. El punto extremo subordinado puede abrir las correspondientes sesiones con el *sessionID* dado. De lo contrario, el punto extremo subordinado podría abrir sesiones adicionales con *sessionID=0* en el mensaje *openLogicalChannel*, aunque deberá adquirir el valor real de *sessionID* del mensaje *openLogicalChannelAck* del punto extremo principal. Si se produce una colisión cuando ambos extremos tratan de establecer sesiones RTP contradictorias en el mismo momento, la terminal principal rechazará el intento de conflicto como se describe en la Recomendación H.245. El intento *openLogicalChannel* rechazado se puede repetir en un momento posterior.

Salvo que se especifique otra cosa para un tipo de datos en particular, los canales de datos fiables son canales bidireccionales y, como tales, contendrán los elementos *forwardLogicalChannelParameters* y *reverseLogicalChannelParameters* sin los elementos *mediaChannel.* La terminal que acepta el canal devolverá el elemento *reverseLogicalChannelParameters* y estará preparado para aceptar la conexión fiable de la terminal solicitante antes de devolver el mensaje *OpenLogicalChannelAck*.

Una terminal que acepte un canal fiable bidireccional deberá estar preparado para aceptar una conexión fiable de la terminal solicitante antes de devolver el mensaje *OpenLogicalChannelAck.*

4.4 GATEWAY

El gateway proporcionará la conversión adecuada entre formatos de transmisión (por ejemplo, H.225.0 a/de H.221) y entre procedimientos de comunicaciones (por ejemplo, H.245 a/de H.242). El gateway llevará a cabo además el establecimiento y la liberación de la llamada en el lado de red de paquetes y en el lado de la PSTN, como se muestra en la Figura 4-2. La conversión entre formatos de video, audio y datos también puede efectuarse en el gateway. Por lo general, la finalidad del gateway (cuando no funciona como una MCU), consiste en reflejar las características de una terminal de red a una terminal de la PSTN, y a la inversa, de manera transparente.

Una terminal H.323 puede comunicar con otro terminal H.323 de la misma red directamente y sin que participe en ello un gateway. Se puede prescindir del Gateway si no se requieren comunicaciones con terminales de la PSTN (terminales no en la red). También es posible que una terminal de un segmento de la red llame al exterior a través de un gateway y de nuevo a la red a través de otro gateway para evitar un router o un enlace de banda ancha reducida [13].

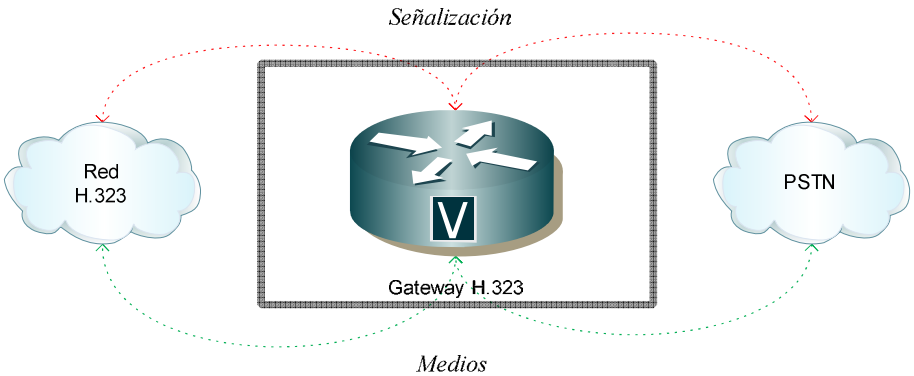


Figura 4 - 2. Flujos de Señalización y Medios de un Gateway H.323

El gateway tiene las características de una terminal H.323 o una MCU de la red de paquetes y de la terminal de la PSTN o una MCU de la red PSTN. La elección entre terminal o MCU se deja a criterios del fabricante. El gateway proporciona la conversión necesaria entre los diferentes tipos de terminal. Hay que tener en cuenta que el gateway puede funcionar al principio como una terminal, pero utilizando más tarde la señalización H.245 empieza a funcionar como una MCU para la misma llamada que inicialmente era punto a punto. Los controladores de acceso saben qué terminales son gateways ya que esto es algo que se indica cuando la terminal/el gateway se registra en el gatekeeper.

Un gateway que transfiere datos T.120 entre la PSTN y la red puede contener un proveedor de MCS T.120 que conecta a los proveedores de MCS T.120 de la red con los proveedores de MCS T.120 de la PSTN [13].

4.5 GATEKEEPER

El gatekeeper, que es opcional en un sistema H.323, presta servicios de control de llamada a los puntos extremos H.323. Puede haber más de un gatekeeper que se comunican entre sí de una manera no especificada. El gatekeeper está separado lógicamente de los puntos extremos. Sin embargo, su implementación física puede coexistir con una terminal, MCU, gateway, MC u otro dispositivo de red no H.323.

En cada zona y cada momento sólo puede haber un gatekeeper, aunque pueden existir muchos dispositivos que proporcionen la función de gatekeeper en una zona. Los dispositivos que proporcionan la función de señalización RAS para el gatekeeper se denominan gatekeepers alternos. Cada gatekeeper alterno puede aparecer ante los puntos extremos como un gatekeeper acceso distinto.

Cuando esté presente en un sistema, el gatekeeper deberá prestar los siguientes servicios [13]:

- Conversión de dirección: El gatekeeper efectuará la conversión de dirección alias a dirección de transporte.
- Control de admisiones: El gatekeeper autorizará el acceso a la red utilizando mensajes ARQ/ACF/ARJ H.225.0. La autorización del acceso puede basarse en la autorización de la llamada, en el ancho de banda o en algún otro criterio que se deja a decisión del fabricante. También puede ser una función nula que admita todas las peticiones.
- Control de ancho de banda: El gatekeeper soportará mensajes BRQ/BRJ/BCF. Esto puede basarse en la gestión del ancho de banda. También puede ser una función nula que acepte todas las peticiones de cambio de anchura de banda.
- Gestión de zona: El gatekeeper proporcionará las funciones anteriores para terminales, MCU y gateways.

El gatekeeper también puede efectuar otras funciones opcionales, tales como [13]:

- Señalización de control de llamada: El gatekeeper puede optar por completar la señalización de la llamada con los puntos extremos y puede procesar él mismo la señalización de la llamada. De manera alternativa, el gatekeeper puede encaminar los puntos extremos para que conecten directamente entre ellos el canal de señalización de llamada. De esta manera, el gatekeeper puede evitar el tratamiento de señales de control de llamada H.225.0.
- Autorización de llamada: Utilizando la señalización H.225.0, el gatekeeper puede rechazar llamadas procedentes de una terminal por ausencia de autorización. Pueden ser motivos de rechazo, entre otros, el acceso restringido hacia/desde terminales o gateways particulares y el acceso restringido durante determinados periodos de tiempo.
- Gestión de ancho de banda: Control del número de terminales H.323 a los que se permite el acceso simultáneo a la red. Utilizando la señalización H.225.0, el gatekeeper puede rechazar llamadas procedentes de una

terminal debido a limitaciones de ancho de banda. Tal cosa puede ocurrir si el gatekeeper determina que no hay suficiente ancho de banda disponible en la red para soportar la llamada. Téngase en cuenta que esta función puede ser una función nula, es decir, que a todos los terminales se les permita el acceso. Esta función actúa también durante una llamada activa, cuando una terminal pide ancho de banda adicional.

- Gestión de llamada: Por ejemplo, el gatekeeper puede mantener una lista de llamadas H.323 en curso. Esta información puede ser necesaria para indicar que una terminal llamado está ocupado y proporcionar información para la función de gestión de ancho de banda.
- Modificación del alias de dirección: El gatekeeper puede devolver un alias de dirección modificado. Si el gatekeeper devuelve un alias de dirección en una ACF, la terminal utilizará el alias de dirección en el establecimiento de la conexión.
- Conversión de los dígitos marcados: El gatekeeper puede convertir los dígitos marcados en un número E.164 o en un número de red privada.

4.6 CONTROLADOR MULTIPUNTO

El controlador multipunto (MC) proporciona funciones de control para soportar conferencias entre tres o más puntos extremos de una conferencia multipunto. El MC lleva a cabo el intercambio de capacidades con cada uno de los puntos extremos de una conferencia multipunto y envía un conjunto de capacidades a los puntos extremos de la conferencia indicando los modos de funcionamiento en los que pueden transmitir. El MC puede revisar el conjunto de capacidades que envía a las terminales como consecuencia de la incorporación de terminales a la conferencia o el abandono de terminales de la misma, o por otros motivos [13].

4.7 Procesador Multipunto

El MP recibe trenes de audio, video y/o datos de los puntos extremos que participan en una conferencia multipunto centralizada o híbrida. El MP procesa estos trenes de medios y los devuelve a los puntos extremos.

Un MP que procese video deberá proporcionar conmutación o mezcla de video. La conmutación de video es el proceso de selección del video que el MP envía como salida hacia las terminales desde una fuente a otra. Los criterios utilizados para efectuar la conmutación pueden determinarse mediante la detección de un cambio en el conferenciante (percibido por el nivel de audio asociados) o mediante el control H.245. La mezcla de video es el proceso de creación de un formato correspondiente a más de una fuente de video en el flujo de video que el MP envía como salida hacia las terminales [13].

4.8 Unidad de Control Multipunto

La MCU es una terminal que da soporte a conferencias multipunto y deberá estar formada por un MC y cero o más MP. La MCU utiliza los mensajes y procedimientos H.245 para implementar características similares a las que figuran en la Recomendación H.243.

Una MCU típica, que soporta conferencias multipunto centralizadas, consta de un MC y de un MP de audio, video y datos. Una MCU típica, que soporta conferencias multipunto descentralizadas, consta de un MC y de un MP de datos que soporte la Recomendación T.120. Se basa en el procesamiento descentralizado de audio y video.
El lado de red de un gateway puede ser una MCU. Un gatekeeper puede incluir también una MCU. En uno y otro caso, se trata de funciones independientes que casualmente están coubicadas [13].

4.9 DIRECCIÓN ALIAS

Una terminal puede tener también una o más direcciones alias asociadas al mismo. Una dirección alias puede representar la terminal o puede representar conferencias que la terminal está acogiendo. Las direcciones alias proporcionan un método alternativo de direccionamiento de la terminal. Dichas direcciones incluyen direcciones de dialledDigits o de *partyNumber* (incluyendo números telefónicos privados y números E.164 públicos), identidades H.323 (cadenas alfanuméricas que representan nombres, direcciones similares a las del correo electrónico, etc.) y cualesquiera otras direcciones definidas en la Recomendación H.225.0. Las direcciones alias deberán ser únicas dentro de una zona. Los controladores de acceso, los MC y los MP no tendrán direcciones alias [13].

4.10 REGISTRO DE LA TERMINAL

El registro es el proceso por el cual una terminal se incorpora a una zona y comunica al gatekeeper sus direcciones de transporte y sus direcciones alias. Como parte de su proceso de configuración, todos los puntos extremos se registrarán en gatekeeper identificado mediante el proceso de descubrimiento. El registro deberá tener lugar antes de que se intente cualquier llamada y podrá producirse periódicamente, según se necesite.

Un gateway o una MCU pueden registrar una sola dirección de transporte o múltiples direcciones de transporte como su dirección de señalización de llamada, y puede registrar una sola dirección de transporte o múltiples direcciones de transporte como su dirección de RAS. La utilización de múltiples direcciones de transporte indicará una lista de direcciones priorizada para intentar cuando se comunica con una determinada terminal a través de su canal RAS o de señalización de llamada.

Una terminal debería enviar una petición de registro (RRQ, registration request) al gatekeeper. La petición se enviará a la dirección de transporte de canal RAS del gatekeeper. La terminal tiene la dirección de red del gatekeeper desde el proceso de descubrimiento de aquel y utiliza el identificador TSAP de canal RAS conocido. El gatekeeper responderá con una confirmación de registro (RCF, registration confirmation) o un rechazo de registro (RRJ, registration reject). Un punto extremo se registrará en un único gatekeeper [13].

La provisión de la comunicación se efectúa siguiendo los pasos que a continuación se indican [13]:

- Fase A: Establecimiento de la llamada (call Setup)
- Fase B: Comunicación inicial e intercambio de capacidad
- Fase C: Establecimiento de comunicación audiovisual
- Fase D: Servicios de la llamada
- Fase E: Terminación de la llamada

4.10.1 ESTABLECIMIENTO DE LA COMUNICACIÓN

El establecimiento de la comunicación se efectúa utilizando los mensajes de control de llamada definidos en la Recomendación H.225.0, de acuerdo con los procedimientos de control de llamada definidos más abajo. Las peticiones de reserva de ancho de banda deberán efectuarse lo antes posible.
Si se especifican la dirección alias y la dirección de transporte, se preferirá la dirección alias.

No hay ninguna sincronización explícita ni enganche entre dos terminales durante el procedimiento de establecimiento de la comunicación. Esto significa que la terminal A puede enviar un mensaje *Setup* al terminal B exactamente al mismo tiempo que la terminal B envía un mensaje *Setup* a la terminal A. Corresponde a la aplicación de terminal determinar si sólo se desea una llamada y ejercer la acción

apropiada. Esta acción puede ser para una terminal indicar que está ocupado siempre que tiene un mensaje *Setup* pendiente. Si una terminal puede soportar más de una llamada simultánea, debe indicar que está ocupado siempre que recibe un mensaje *Setup* de la misma terminal con el cual tiene un mensaje *Setup* pendiente.

Una terminal será capaz de enviar el mensaje *Alerting*. *Alerting* tiene el significado de que la parte llamada (usuario) ha sido avisada de una llamada entrante. "*Alerting"* será sólo originado por la última terminal llamado y por tanto sólo cuando ha sido avisado el usuario. En el caso de interfuncionamiento a través de un gateway, el gateway enviará *Alerting* cuando reciba una indicación de llamada de la PSTN. Si una terminal puede responder a un mensaje *Setup* con un mensaje *Connect*, *call proceeding*, o *release complete* en el plazo de 4 segundos, no es necesario enviar el mensaje *Alerting*. Una terminal que envía el mensaje *Setup* puede esperar recibir un mensaje *Alerting, Connect, Call proceeding o Release complete* en un plazo de 4 segundos después de su transmisión con éxito.
El mensaje *Connect* debe enviarse sólo si se está seguro de que el intercambio de capacidades H.245 concluirá con éxito y puede existir un nivel mínimo de comunicaciones, con el objeto de mantener la coherencia del significado del mensaje Conexión entre redes de paquetes y redes con conmutación de circuitos [13].

4.10.2 COMUNICACIÓN INICIAL E INTERCAMBIO DE CAPACIDAD

Una vez que ambos lados han intercambiado los mensajes de establecimiento de comunicación de la fase A, las terminales, si proyectan emplear H.245, establecerán el canal de control H.245.

Las capacidades de las terminales se intercambian mediante la transmisión del mensaje *terminalCapabilitySet* H.245. Este mensaje de capacidad será el primer mensaje H.245 enviado a menos que la terminal indique que comprende el campo

parallelH245Control. Si antes de la terminación satisfactoria de intercambio de capacidad terminal, cualquier otro procedimiento presenta un fallo (es decir, rechazado, no comprendido, no soportado) la terminal de origen se ha de iniciar y completar satisfactoriamente el intercambio de capacidad terminal antes de intentar cualquier otro procedimiento. Una terminal que recibe un mensaje terminalCapabilitySet de una entidad par antes de iniciar el intercambio de capacidades responderá conforme a lo requerido, e iniciará y completará satisfactoriamente el intercambio de capacidades con esa entidad par antes de iniciar cualquier otro procedimiento [13].

4.10.3 ESTABLECIMIENTO DE COMUNICACIÓN AUDIOVISUAL

Después del intercambio de capacidades y la determinación de principal-subordinado, se utilizarán los procedimientos de la Recomendación H.245 para abrir canales lógicos para los diversos flujos de información. Los flujos de audio y video y comunicaciones de datos que se transmiten por los canales lógicos establecidos en H.245, se transportan en identificadores TSAP dinámicos utilizando un protocolo no fiable [13].

4.10.4 SERVICIOS DE LA LLAMADA

Cambios de ancho de banda: El ancho de banda de la llamada la establece y aprueba inicialmente el gatekeeper, durante el intercambio de admisiones. Una terminal deberá asegurar que la suma correspondiente a todos los canales transmitidos y recibidos de audio y de video excluidos cualesquiera encabezados RTP, encabezados de carga útil RTP, encabezados de red, etc., se halla dentro de ese ancho de banda. Los canales de datos y de control no se incluyen en ese límite [13].

4.10.4.1 ESTADO

Para determinar si una terminal se ha desconectado o ha pasado a un modo fallo, el gatekeeper puede utilizar la secuencia de mensajes de petición de información (IRQ, information request) o de respuesta a petición de información (IRR, information request response), a fin de sondear secuencialmente las terminales con un intervalo establecido por el fabricante. El gatekeeper puede solicitar información para una única llamada o para todas las llamadas activas. Excepto cuando se soliciten segmentos IRR adicionales, el intervalo de sondeo secuencial para solicitar información de una llamada en particular o de todas las llamadas deberá ser superior a 10 segundos. Sin embargo, el gatekeeper puede enviar mensajes IRQ que contengan valores de *callReferenceValue* únicos sin tener en cuenta el periodo de sondeo [13].

4.10.4.2 AMPLIACIÓN DE UNA CONFERENCIA AD HOC

Los siguientes procedimientos son opcionales para las terminales y gateways, y obligatorios para los MC.

Cuando un usuario efectúa una llamada, el punto extremo llamante a menudo desconoce el propósito de la llamada. El usuario puede desear simplemente crear una conferencia para él mismo y el punto extremo llamado, el usuario puede desear incorporarse a una conferencia en la entidad llamada, o el usuario puede desear obtener una lista de conferencias que la entidad llamada puede proporcionar. Utilizando los procedimientos de esta cláusula, las conferencias pueden ampliarse de conferencias punto a punto a conferencias multipunto ad hoc [13].

4.10.5 TERMINACIÓN DE LA LLAMADA

Cualquier terminal o entidad intermedia de señalización de llamada puede terminar la llamada. La terminación de llamada deberá realizarse de acuerdo con cualquiera de los procedimientos A o B siguientes [13]:

Procedimiento A [13]:

- Interrumpir la transmisión de video al final de una imagen completa, en su caso.
- Interrumpir la transmisión de datos, en su caso.
- Interrumpir la transmisión de audio, en su caso.
- Transmitir el mensaje Liberación Completa y cerrar el canal de señalización de llamada H.225.0 y, de estar abierto por separado, el canal de control H.245 sin enviar ningún mensaje H.245. Obsérvese que se supone implícito el cierre de los canales de medios.

Procedimiento B [13]:

- Interrumpir la transmisión de video al final de una imagen completa y a continuación cerrar todos los canales lógicos de video, en su caso.
- Interrumpir la transmisión de datos y a continuación cerrar todos los canales lógicos de datos, en su caso.
- Interrumpir la transmisión de audio y a continuación cerrar todos los canales lógicos de audio, en su caso.
- Transmitir el mensaje H.245 endSessionCommand por el canal de control H.245, para indicar al extremo distante que desea desconectarse de la llamada, e interrumpir a continuación la transmisión del mensaje H.245.
- Esperar a recibir el mensaje endSessionCommand del otro punto extremo y cerrar a continuación el canal de control H.245.

- Transmitir un mensaje Liberación Completa y cerrar el canal de señalización de llamada H.225.0.

Capítulo 5
Calidad de la voz

Las aplicaciones distribuidas tradicionales están orientadas al tráfico de paquetes, mientras que las aplicaciones tradicionales de voz están orientadas al tráfico de circuitos, cuando ambas son integradas a una red convergente, ambas demandan diferentes recursos de red y exigencias de calidad, por ejemplo, la transferencia de archivos es tolerante a los retardos, sensible a las pérdidas y 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, no es muy sensible a las pérdidas, pero tiene una baja tolerancia al retardo y jitter. Es decir, que aunque los dos tipos de tráfico pueden transportarse por la misma red, no es posible manejarlos de la misma manera [7].

La calidad de la voz es un aspecto en el que influyen gran cantidad de factores, tanto subjetivos como objetivos. Definir claramente que se entiende por calidad de la voz no es, una tarea sencilla puesto que el concepto puede contemplarse desde distintos puntos de vista.

Desde el punto de vista del usuario que, finalmente será quien decidirá sobre la bondad o no de la solución tecnológica, la calidad de la voz está en función de la fidelidad con la que se escucha la voz en el otro extremo (claridad de la voz) y la capacidad de la red para soportar el flujo de la conversación.
Por otro lado, desde una perspectiva más centrada en la ingeniería de red, la calidad está en función de la capacidad de la red para satisfacer las demandas de un tráfico en tiempo real (como es el caso de la voz) en términos de diferentes parámetros. Estos parámetros son el ancho de banda, pérdida de paquetes, retardo y jitter.

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 emplee el término.

Los aspectos que mayor influencia ejercen sobre la percepción de la calidad de servicio de la telefonía sobre redes de paquetes, desde el punto de vista de los usuarios son [7]:

Tasa de conectividad: Hace referencia a la posibilidad con la que la red dispondrá de recursos para soportar 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 ese 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 cuánta 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 si ejerce gran influencia) y del eco, puesto que éste este último 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 el siguiente paso es codificar la voz, transmitirla a través de la red y evaluar qué tal se escucha. El resultado será una medida de la bondad del esquema de codificación empleado. La calidad de la voz y la inteligibilidad están relacionadas entre sí y ambas dependen de la tasa binaria y de la tasa de error, como se muestra en la Figura 5 - 1. Cuando mayor es la tasa binaria, es más probable obtener una buena calidad de voz. Por otro lado, 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.

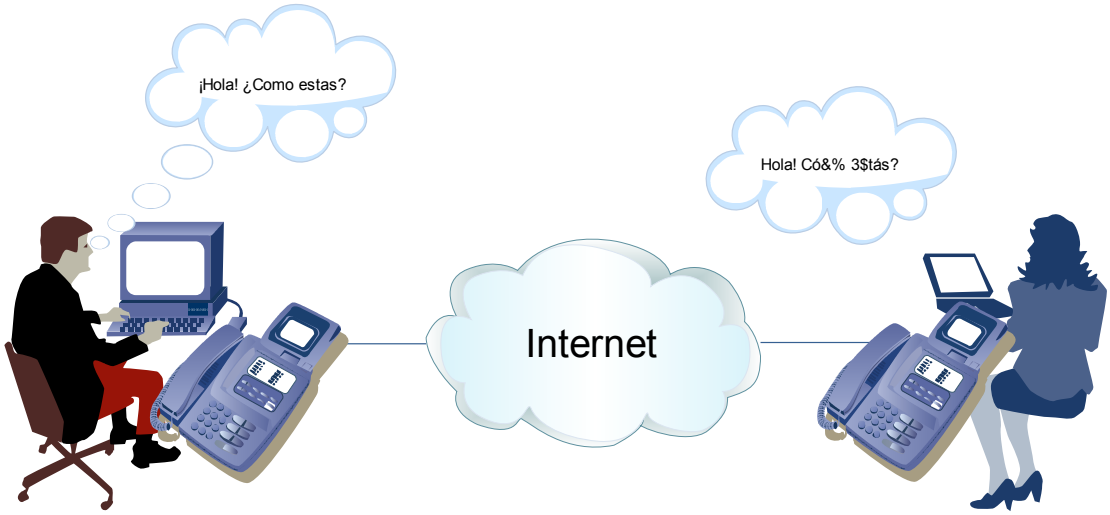


Figura 5 - 1. Calidad de Servicio en VoIP

5.1 FACTORES QUE INFLUYEN EN LA CALIDAD DE LA VOZ

Para los usuarios de las redes de VoIP, las diferencias tecnológicas existentes entre las redes de conmutación de circuitos 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 de 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: jitter, pérdida de paquetes (PLR), retardo extremo a extremo (OWD) y ancho de banda [7].

5.1.1 PÉRDIDA DE PAQUETES

Las pérdidas de paquetes (*PLR*) son el resultado del descarte de paquetes que se producen en los nodos de la red como consecuencia de la congestión de dichos nodos (ver Figura 5 - 2). Puesto que en la redes de conmutación de paquetes no se produce una reserva de recursos previa al envío de la información del 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 [7].

La solución más viable al problema de las pérdidas es la mejora de la arquitectura de la red y la implementación de mecanismos que ayuden a atenuar los efectos de dichas pérdidas. En efecto, ya que las pérdidas son, básicamente, una cuestión de capacidad, si se sustituyen los enlaces 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 [7].

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, puesto que esta es más sensible a los retardos. Son necesarias, por tanto, otro tipo de mecanismos que atenúen los efectos de las pérdidas. Para este fin, se han desarrollado principalmente tres técnicas o mecanismos:

Corrección de errores (FEC, Forward Error Control): En este tipo de técnicas, junto con los paquetes, se incluye información de redundancia que permite recuperar el valor del paquete perdido a partir del valor de los paquetes perdidos. Su principal inconveniente es el retardo de procesamiento introducido y el incremento en el uso del ancho de banda ya que para decodificar un paquete son necesarios paquetes vecinos.

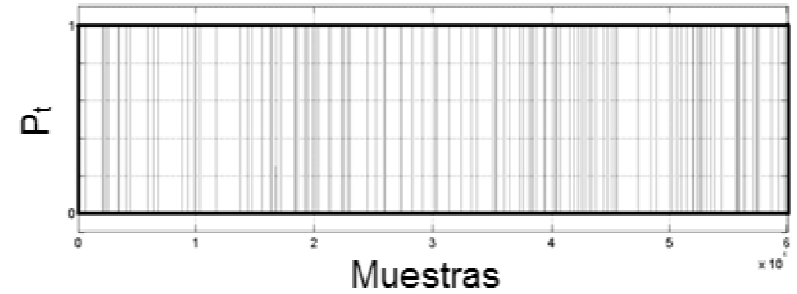
Interleaving: Consiste en aleatorizar o cambiar el orden de los paquetes en el momento de envío, para disminuir el impacto de las pérdidas por el efecto de las perdidas en ráfagas. El inconveniente es el retardo adicional que se introducen por procesamiento en el momento de cambiar el orden en el transmisor y reordenar en el receptor.

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 costo de procesamiento y mayor retardo introducido.

Figura 5 - 2. Pérdida de Paquetes

La pérdida de paquetes se representa comúnmente como un número porcentual, sin embargo, este valor no proporciona información de cómo se presentaron las perdidas en función del tiempo. En Internet las pérdidas se pueden presentar de dos formas: de manera no consecutiva o independiente y de manera consecutiva o dependiente (a ráfagas). Una forma de saber cómo se presentaron las pérdidas en una comunicación de voz sobre redes IP, es mediante una representación binaria (vectores de pérdida).Sea $P=\{P_t : t=1,...,N\}$, donde $P_t=1$ representa un paquete perdido, $P_t=0$, representa un paquete recibido y N es la longitud de los paquetes enviados desde el transmisor.

De esta manera, si graficamos los vectores de perdida en cada comunicación, podremos conocer la información referente a cómo se presentaron las perdidas en función del tiempo. La Figura 5-X muestra la representación gráfica de vectores de perdida correspondiente a una comunicación donde se presento la perdida de manera no consecutiva (Figura 5-3 (a)) y donde se presento de manera consecutiva o a ráfagas (Figura 5-3 (b)).



(a) Pérdidas no Consecutivas

(b) Pérdidas Consecutivas

Figura 5 - 3. Vectores de Pérdida de Paquetes

5.1.2 RETARDO

El retardo o latencia es el tiempo invertido por la señal de voz en su viaje desde el origen hasta el destino, como se ilustra en la Figura 5 - 4 [7].

También podríamos definir la latencia como la cantidad de tiempo que le toma a una trama de voz desde que sale de la boca del emisor hasta que llega a los oídos del receptor (ver Figura 5 - 4) [3].

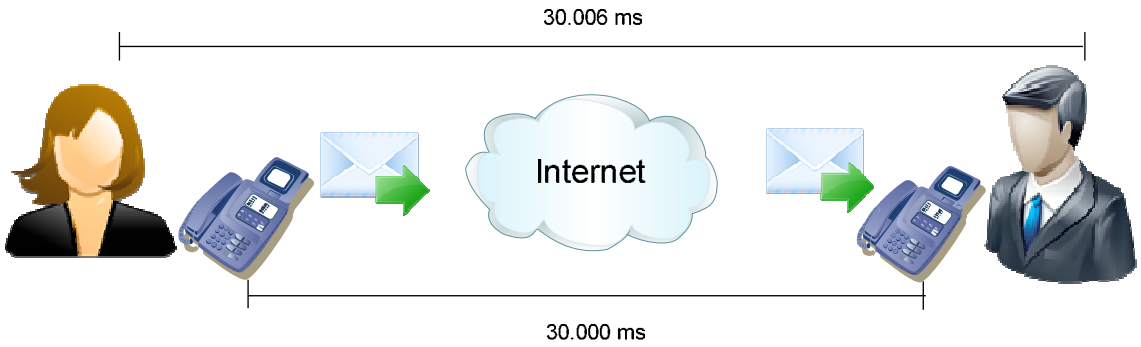


Figura 5 - 4. Retardo Extremo a Extremo (OWD)

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 [7]. Al estudiar la influencia del retardo en las comunicaciones de voz sobre IP son dos los aspectos más importantes a determinar: 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 (ver Figura 5 - 5 y Tabla 5 - 1). La recomendación G.144 de la ITU-T establece este umbral entorno a los 400 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 soportarse retardos de hasta 400 ms.

Tabla 5 - 1. Calidad de la Voz

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

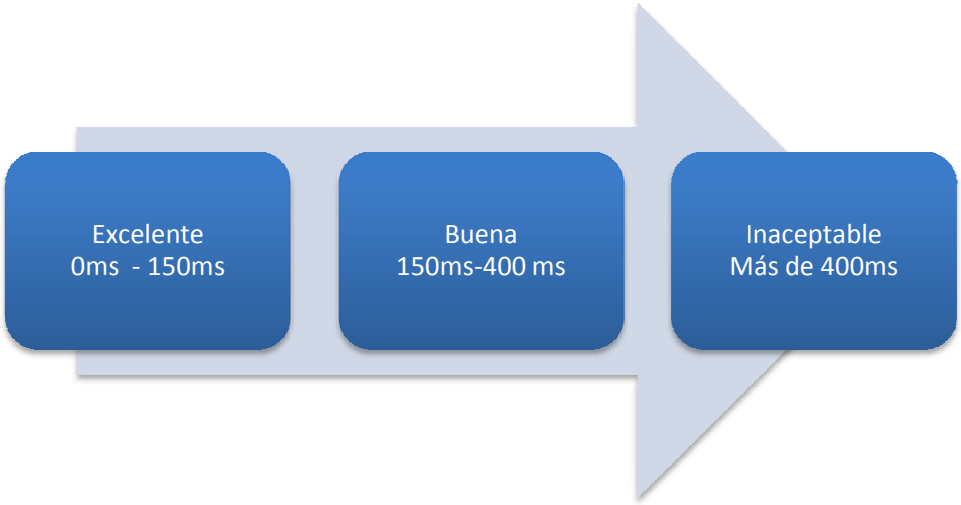


Figura 5 - 5. Relación entre el Retardo Extremo a Extremo y la Calidad de la Voz

El retardo extremo a extremo está formado por distintas fuentes de retardo. Para conocer cada componente, 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 en el emisor 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 CODECs realizan además, una compresión que reduce el ancho de banda consumido por la comunicación vocal.

Los principales retardos introducidos en el proceso de codificación de algunos CODECs más usados en VoIP son: retardo complejidad, tiempo entre paquetes, retardo de paquetización y retardo del buffer de supresión de jitter, La Tabla 5 - 2 resume dichos retardos. 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 [7].

Tabla 5 - 2. Características de algunos CODECs

	G711	G729	G723.1
Tasa binaria (kbps)	64	8	6,3/5,3
Complejidad (MIPS)	0,1	22	16/18
Retardo codificador (ms)	0,125	15	37,5
Tiempo entre paquetes (ms)	20	20	30
Retardo de empaquetamiento (ms)	1,5	15	37,5
Tamaño del buffer de supresión de Jitter (ms)	40	40	60
Calidad (MOS)	4,4	4,1	3,5-3,9

Una vez que los paquetes llegan al gateway, éste invertirá un cierto tiempo en retransmitirlos 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. [7].

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 μs/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 pertenecen dichos paquetes, el retardo de encolado varía mucho de un paquete a otro. En cualquier caso, generalmente, el retardo de la red se encuentra comprometido entre 70ms y 100 ms y es, por tanto, una de las contribuciones más importantes al retardo total.

El retardo OWD puede ser medido en función de estampas de tiempo y tiempos de arribo de los paquetes transmitidos en una comunicación de voz sobre redes

IP. Sea S_K la estampa de tiempo RTP para un paquete K de tamaño L, y R_K es el tiempo de arribo en unidades de estampas de tiempo RTP del paquete K de tamaño L. Entonces el retardo extremo a extremo para dos paquetes K y $K-1$ quedará determinado mediante la siguiente ecuación:

$$OWD^{K}(L) = (R_K - S_K)$$

$$OWD^{K-1}(L) = (R_{K-1} - S_{K-1}) \qquad 5.1$$

5.1.3 JITTER

En general, a la hora de analizar las prestaciones de una red se habla del retardo en términos de 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 eficiente.

En 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 (de hecho lo más probable es que no lo hagan), al contrario de lo que ocurre en las redes de conmutación de circuitos. Como consecuencia, cada paquete llegará al destino atravesando un número distinto de nodos de red y, por lo tanto, alcanzarán su objetivo con un retardo diferente, como se muestra en la Figura 5 - 6. Esta variabilidad del retardo recibe el nombre de jitter [7]. Existen una gran diversidad de jitter, los más utilizados en la evaluación de sistemas VoIP son:

- Jitter de OWD: diferencia de retardos extremo a extremo entre dos paquetes consecutivos.
- Jitter de Arribo: diferencia de tiempos de arribo entre dos paquetes consecutivos. El jitter es un problema que solo está presente en redes basadas en paquetes como las redes IP [3].

Los retardos se envían desde el transmisor a tasas constantes, por ejemplo, un paquete de voz cada 20 ms (tiempo de inter-partida). Sin embargo, al llegar al destino, el tiempo de inter-arribo o jitter de arribo es variable debido a las diferencias en los retardos de encolamiento y propagación fundamentalmente [7].
De igual forma como se mide el retardo OWD, el jitter puede ser medido en función de estampas de tiempo y tiempos de arribo de los paquetes transmitidos en una comunicación de voz sobre redes IP. Sea S_K la estampa de tiempo RTP para un paquete K de tamaño L, y R_K es el tiempo de arribo en unidades de estampas de tiempo RTP del paquete K de tamaño L. Entonces el jitter de OWD $J^K(L)$ para dos paquetes consecutivos K y $K-1$ quedará determinado mediante la siguiente ecuación:

$$J^K(L) = OWD^K(L) - OWD^{K-1}(L) \qquad 5.2$$

o

$$J^K(L) = (R_K - S_K) - (R_{K-1} - S_{K-1}) \qquad 5.3$$

reordenando términos, tenemos:

$$J^K(L) = (R_K - R_{K-1}) - (S_K - S_{K-1}) \qquad 5.4$$

donde, $IAT(K,K-1) = (R_K - R_{K-1})$ es el tiempo de inter-arribo y $IDT(K,K-1) = (S_K - S_{K-1})$ es el tiempo de inter-partida o jitter de arribo. Por tanto el jitter de arribo en función del jitter de OWD está determinado por:

$$IAT(K,K-1) = J^K(L) + IDT(K,K-1) \qquad 5.5$$

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 lo tanto, cuanto

mayor es el jitter de los paquetes, mayor es el tamaño del buffer de supresión de jitter necesario para reducir su impacto en la calidad [7].

La supresión de 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 (capacidad de absorción del jitter), o 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 [7].

Figura 5 - 6. Diferencias de Retardos

5.2 TECNOLOGÍAS DE EVALUACIÓN DE CALIDAD VOIP

Hoy en día las tecnologías de red conmutadas por circuito tienen muchas ventajas, motivo por el cual, siguen implementándose, algunas son: disponibilidad, capacidad, respuesta rápida y alto nivel de calidad. Este último es que más se tomará en cuenta en la presente sección. En redes por conmutación de paquetes no se establece ningún circuito para iniciar la comunicación ya que los paquetes son enviados del mismo origen al mismo destino pero por varias rutas alternas [1].

La transmisión de voz y datos en una sola red convergente es un logro muy importante en el mundo de las telecomunicaciones y es una interesante área de investigación ya que esta implementación proporciona servicios más atractivos e innovadores, provee servicios más avanzados que los sistemas de telefonía tradicional con grandes posibilidades de reducción de costos en llamadas telefónicas, menores requerimientos de ancho de banda, menores consumos de operación y administración, disponibilidad IP generalizada, entre otros. Sin

embargo, la red convergente es una red IP que no provee garantía de QoS. Muchas soluciones han sido propuestas para aliviar este problema [1].

Una estrategia que resuelve el problema antes mencionado consiste en reservar recursos a través del medio desde el emisor hasta el receptor, mediante un mecanismo llamado control de admisión de llamada (CAC), con el cual se puede determinar si hay suficiente ancho de banda para asignar y así mantener un nivel de calidad de servicio aceptable para aceptar o no una llamada entrante. Para poder implementar este mecanismo son necesarios los siguientes servicios: Protocolo de reservación de recursos (RSVP), servicios diferenciados (DiffServ), protocolo múltiple de etiquetado de conmutación (MPLS) y medición extremo a extremo basado en control de admisiones (EMBAC). La reservación de servicios es un mecanismo muy difícil y costos de implementar ya que requiere cambios en cada router del que se haga uso a través de las redes por las que cruce la llamada, lo cual lo hace muy complicado para aplicar en Internet [1].

La evaluación de la calidad de VoIP en redes convergentes mediante mediciones puede ayudar a determinar por qué algunas llamadas fallan o presentan problemas en la calidad de la voz [1].

La ITU (Unión Internacional de Telecomunicaciones) provee mecanismos especializados para la medición de calidad de voz de manera precisa. Un método subjetivo para evaluar la calidad de la voz puede ser encontrado en la recomendación P. 800 de la ITU-T, donde la calidad de la voz está expresada en términos de la puntuación media de opinión (MOS), la cual califica la calidad en un rango de valores de 1 a 5, tomando el 1 como la calidad más pobre y el 5 como calidad óptima [1].

Algunos métodos objetivos dependen de comparaciones entre señales recibidas y señales originales y de esta manera miden la calidad percibida en términos de MOS, estos métodos son denominados métodos intrusivos. El método más

reciente para la medición de calidad de voz de manera intrusiva es la evaluación perceptual de calidad de audio (PESQ). Otros métodos objetivos para la medición de la calidad de voz (los no intrusivos) miden la calidad de la señal recibida sin la necesidad de compararla con la señal original. En esta categoría existen dos métodos principales, los cuales son, la recomendación P.563 (ITU-T 2004) y el modelo-E definido en la recomendación G.107 (ITU-T 2009) [1].
Para seleccionar un método para medir la calidad de la voz en redes, es necesario tomar en cuenta las características de la red IP y llamadas de voz [1].

5.3 CATEGORÍAS DE TECNOLOGÍAS DE EVALUACIÓN DE CALIDAD DE VOIP

Los métodos para evaluación de calidad en VoIP se clasifican en dos: métodos subjetivos y métodos objetivos, éstos últimos a su vez se clasifican intrusivos y no intrusivos. Los métodos no intrusivos están divididos en métodos basados en señal y basados en parámetros. [1].

El primer criterio a tomar en cuenta para determinar calidad en voz y video es la calidad subjetiva, es decir, la calidad percibida por el usuario. Entre los factores que afectan la calidad subjetiva se encuentran: pérdida de paquetes, retardo, jitter, poco volumen, eco y distorsión de CODEC. El método más común para medir la calidad de manera subjetiva es métrica MOS. Sin embargo, otros métodos deben tomarse en cuenta para determinar la calidad de la señal, es necesario comparar la señal degradada con la señal original (método intrusivo) o también utilizando parámetros de calidad física en la señal recibida sin ninguna señal original como referencia (método no intrusivo) [1].

Entre las características deseadas para evaluación de calidad de la voz en VoIP se encuentran:

1. Automático: se debe proveer medición de calidad de la voz en línea mientras la red está trabajando.

2. No intrusivas: se debe proveer servicios de medición de calidad de la voz tomando en cuenta la señal recibida sin necesidad de la señal original.
3. Exacta: se debe proveer medición precisa de calidad de audio para reflejar como la calidad es percibida por el usuario final.
4. Debe ser adaptable a cambios nuevos y actualizaciones en la red.

5.3.1 EVALUACIÓN SUBJETIVA DE LA CALIDAD.

Evaluaciones subjetivas pueden llevarse a cabo mediante conversaciones . En el modo de conversaciones, dos sujetos comparten una conversación estando en dos cuartos aislados y separados para reportar sus respectivas opiniones mediante una escala y posteriormente se realiza un cálculo aritmético. Para llevar a cabo un experimento subjetivo como el explicado anteriormente se necesitan las siguientes condiciones estrictas en el laboratorio [1]:

- Un tamaño determinado del cuarto.
- Nivel de silencio aceptable.
- Cabina a prueba de sonidos con un volumen no menos de 20 m^3.
- En caso de que se trate de pruebas de escucha el cuarto debe tener un volumen entre 30m^3 y 120m^3 con una duración de eco de menos de 500 ms.
- Ruido de fondo menos a 30 dB.
- Señales de voz grabadas deben ser muestras simples, entendibles y de preferencia frases cortas, ordenadas aleatoriamente.
- Todo el material debe ser grabado con un micrófono a una distancia de 140-200 mm de la boca del emisor.

También se debe tomar en cuenta que los sujetos que participan en la prueba no deben estar ligados de ninguna manera al trabajo, red o empresa que se va a auditar, ni deben estar involucrados en labores de mantenimiento de circuitos

telefónicos o algo relacionado a codificación de voz, no deben haber participado en ninguna prueba subjetiva por al menos seis meses y en pruebas de conversaciones o escucha por al menos un año [1].

En una metodología de calificación por opinión el comportamiento del sistema es evaluado ya sea directamente (Evaluación por categoría absoluta, ACR) o de manera relativa mediante la calidad subjetiva de un sistema de referencia como en DCR (evaluación de categoría por degradación) ó mediante CCR (evaluación por categoría de comparación) [1].

La métrica más común en metodologías por opinión es el MOS, el cual es una métrica ACR con una escala de 5 valores:

1. Malo
2. Poco favorable
3. Pasable
4. Bueno
5. Excelente

Un valor MOS se obtiene mediante una operación aritmética utilizando varios valores MOS de varios sujetos [1].

En evaluaciones DCR se toman en cuenta dos muestras (A y B). La muestra A representa la muestra de referencia, como referencia de calidad, mientras que la muestra B representa una muestra degradada. Los sujetos son instruidos para comparar acústicamente dos muestras y medir la degradación de la muestra B en relación con la A de acuerdo a los siguientes niveles de degradación [1]:

1. Muy molesto
2. Molesto
3. Ligeramente molesto
4. Audible pero no molesto
5. Inaudible

Las muestras deben estar compuestas de dos períodos, separados por un silencio, primero la muestra A y posteriormente la B [1].
Los resultados son denominados MOS degradados (DMOS). Cada configuración es evaluada por medio de juicios de muestras de al menos cuatro emisores [1].
En el método CCR se usan dos muestras por cada señal. La referencia oculta es identificada y los sujetos son interrogados para calificar sus muestras contra las muestras de otros en la siguiente escala de 7 puntos [1]:

- (-3) Mucho peor
- (-2) Peor
- (-1) Muy malo
- Casi igual
- Ligeramente mejor
- Mejor
- Mucho mejor

El puntaje de la evaluación es denominado MOS de comparación (CMOS). Se recomienda también utilizar otros métodos con el afán de garantizar lo más que se pueda la calidad. En la prueba de diagnóstico de rima (DRT) se le asigna a un sujeto la tarea de reconocer una de dos posibles palabras en un conjunto de pares de rimas. En la medición de diagnóstico aceptable (DAM) los puntajes están basados evaluando la calidad de un sistema de comunicación basado en la aceptabilidad de voz percibida por alguien capacitado o certificado para escucharlo [1].

5.3.2 Evaluación objetiva de la calidad

La calidad de voz objetiva intenta simular mediante algún algoritmo las opiniones humanas o usar modelos computacionales para automáticamente evaluar la calidad de la voz transmitida sobre una red IP para no utilizar personal humano.

Tienen como objetivo principal predecir los resultados de la tabla de valores de MOS lo más acertadamente posible. Estos sistemas están basados en métricas objetivas de señales de audio o características de la red. Se puede hacer una clasificación de estos métodos: modelos intrusivos y no intrusivos [7].

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 la red de voz sobre paquetes y cuyo objetivo es predecir la calidad de la voz en función del retardo, Jitter, las pérdidas y otras características de la red [7].

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:

$$R = R_O - I_s - I_d - I_e + A \qquad 5.6$$

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

el término RoR hace referencia a la relación señal-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 y el uso de un CODEC (I_e); del retardo (I_d); y el margen de seguridad (A) [7]. Una vez que el códec a usar es bien conocido, solo es necesario las estadísticas de OWD y pérdida de paquetes para poder estimar la calidad mediante el factor R de acuerdo a la ecuación dada en [16] como sigue:

$$R = 93.2 - I_d\left(OWD\right) - I_e\left(CODEC, PLR\right) \qquad 5.7$$

donde:

$$I_d = 0.024\left(OWD\right) + 0.11\left(OWD - 177.3\right) y \left(OWD - 177.3\right) \qquad 5.8$$

$$y(x)=\begin{cases}0, & x<0\\ 1, & x>0\end{cases}$$

$$I_e(G.711)\sim 0+30\ln(1+15PLR) \qquad 5.9$$

$$I_e(G.729)\sim 11+40\ln(1+10PLR)$$

Por otro lado, la relación entre el factor R y el MOS, está dada por:

$$\begin{aligned} MOS &= 1; && R<0\\ MOS &= 1+0.035R+7\cdot 10^{-6}R(R-60)(100-R); && 0\le R\le 100\\ MOS &= 4.5; && R>100\end{aligned} \qquad 5.10$$

Típicamente, los valores del factor R están categorizados como se muestra en la Tabla 5 - 3.

Tabla 5 - 3. Rango de valores del Modelo E

Factor R	MOS	Calidad	Nivel de satisfacción
$90 \le R < 100$	4.34 - 4.50	Óptima	Muy satisfecho
$80 \le R < 90$	4.03 - 4.34	Alta	Satisfecho
$70 \le R < 80$	3.60 - 4.03	Media	Algunos usuarios están insatisfechos
$60 \le R < 70$	3.10 - 3.60	Baja	Muchos usuarios están insatisfechos
$0 \le R < 60$	1.00 - 3.10	Pobre	Casi todos los usuarios están insatisfechos

5.3.3. EVALUACIÓN OBJETIVA INTRUSIVA DE LA CALIDAD

Los métodos intrusivos basan su calidad en el resultado de comparar la señal original con la señal reconstruida por el decoder en el lado del receptor. Proveen un método acertado para medir la calidad de la voz. Algunos algoritmos intrusivos son usados en dominio de tiempo, otros de evaluación de calidad objetiva hacen

uso de un espectro de distorsión para evaluar el comportamiento de CODECs LBR. Pero ninguno de estos fue aprobado para ser adoptado por la ITU-T. Posteriormente, medidas de percepción de dominio fueron introducidas y estandarizadas.

Las medidas de percepción de dominio están basadas en modelos de percepción de auditoría humana. En estos modelos las dos señales, tanto la original como la degradada se transforman en una representación psicofísica que se aproxima a la percepción humana o simula la audición como una resolución espectro de banda crítica, frecuencia selectiva, curva de sonido igual e intensidad de volumen.

Entonces la diferencia perceptiva entre la señal original y la degradada es mapeada dentro de una estimación de diferencia de calidad perceptiva como la percibida por el que escucha. Estos modelos incluyen aplicaciones como: bloque de normalización de medición (MNB), sistema de medición de análisis perceptivo (PAMS), medición de calidad perceptual de audio (PSQM) y evaluación perceptiva de calidad de sonido (PESQ).

5.3.4. EVALUACIONES DE CALIDAD OBJETIVAS NO INTRUSIVAS

Mediciones no intrusivas hacen uso exclusivamente de la señal degradada, es decir, la señal recibida en el lado receptor. Proveen un conveniente mecanismo de monitoreo en redes en tiempo real. Esto es importante si lo que se pretende es realizar algún control de calidad dinámico. Es el método más apropiado para el monitoreo de la calidad de voz en redes de VoIP. Los métodos no intrusivos se dividen en dos categorías: los basados en señal y los basados en parámetros.

Los métodos basados en señales procesan la trama de audio que es codificada después de salir del buffer para extraer información relevante para la estimación de la calidad de voz. Los métodos basados en parámetros basan sus resultados en varias propiedades relevantes a parámetros de redes de comunicación como

por ejemplo, pérdida de paquetes, retardo y jitter. Esto hace que los métodos basados en parámetros sean más específicos para tipos particulares de redes de telecomunicaciones mediante la dependencia de predicción en los parámetros de la red.

Capítulo 6
Medición de parámetros de QoS en una red VoIP H.323

En el presente trabajo se evaluó el desempeño de una aplicación VoIP y se realizó la caracterización de las principales métricas de QoS. Para llevar a cabo esta tarea, se generó tráfico VoIP mediante el establecimiento de un conjunto de llamadas de prueba con una aplicación VoIP en hardware. Posteriormente se capturó el tráfico VoIP generado, mediante el analizador de protocolos de red Wireshark [15] para obtener un conjunto de patrones de tráfico o trazas (series de tiempo). El principal objetivo de estas mediciones fue colectar un conjunto de patrones de tráfico, tales como: jitter de arribo, pérdida de paquetes y retardo extremo a extremo.

Para colectar diversos conjuntos de trazas, las llamadas de prueba realizadas se llevaron a cabo bajo diferentes configuraciones en función de códec y tamaño de paquete de voz, como se muestra en la Tabla 6-1.

Tabla 6 - 1. Configuraciones de Llamadas de Prueba

Longitud de trazas (muestras)	Longitud de paquetes de voz (ms)	Longitud de paquetes de voz (Bytes)	
		G.711	G.729
180,000	20	160	20
90,000	40	320	40
60,000	60	480	60

Como se ilustra en la Tabla 6 - 1, las llamadas de prueba se realizaron bajo los dos esquemas de codificación de voz más importantes, G.711 y G.729; y diferentes tamaños de paquete de voz (G.711: 20ms/160bytes, 40ms/320bytes, 60ms/480bytes; G.729: 20ms/20bytes, 40ms/40bytes, 60ms/60bytes). Cada llamada de prueba tuvo duración de una hora, la primera columna de la Tabla 6-1

ilustra el número de muestras que corresponde a cada llamada con su configuración elegida.

6.1 Escenario de Medición

El escenario de medición en el cual se colecto el tráfico VoIP está formado por dos redes de área local (LAN), interconectadas mediante el Backbone de Internet como se muestra en la Figura 6 - 1. LAN “A”: Red de la Universidad de Quintana Roo-UQROO (30 MBPS) y LAN “B”: Red del Centro de Investigación y Estudios Avanzados del IPN-CINVESTAV Unidad Guadalajara (2MBPS).

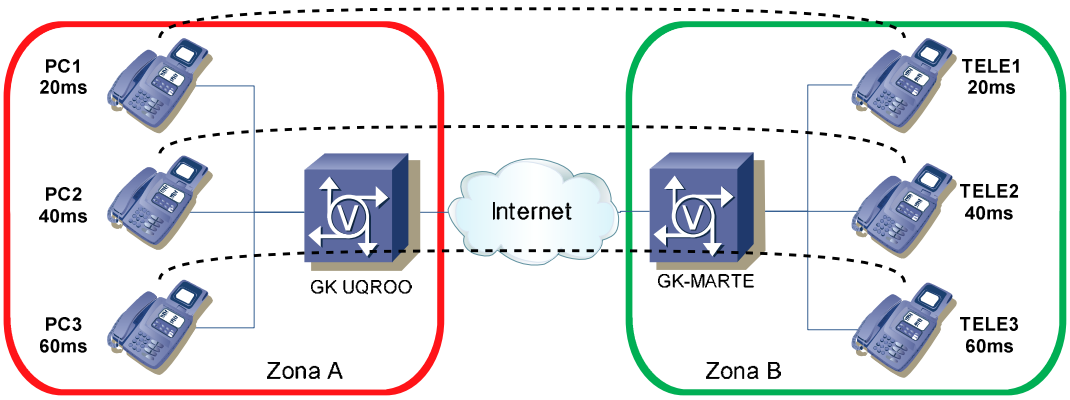


Figura 6 - 1. Escenario de Medición

La Figura 6 - 1 muestra una arquitectura H.323 formada por dos zonas interconectadas vía Internet. Cada zona está compuesta por un Gatekeeper (GK), y terminales H.323 (TE), interconectados vía una LAN. La zona “A” está conformada por las terminales PC1, PC2, y PC3, esta zona es administrada por el GK-UQROO; en esta zona se instaló el analizador de protocolos de red Wireshark para colectar las trazas. La zona “B” está compuesta por las terminales TELE1, TELE2, y TELE3, y el GK-MARTE.

Los gatekeepers utilizados en esta tesis fueron implementados mediante software, haciendo uso del proyecto GNU Gatekeeper [12] el cual es un proyecto de código abierto que implementa un Gatekeeper H.323. La tabla 6 - 2 muestra la configuración empleada en GK-UQROO y GK-MARTE.

SECCIÓN [GATEKEEPER::MAIN]

Fortytwo

Fortytwo=42

Default: N/A

Esta configuración permite probar la existencia del archivo de configuración (gatekeeper.ini o cualquiera que sea su nombre) con el que va a trabajar el gatekeeper. Un mensaje de advertencia se mostrará en caso de que no existiera dicho archivo [12].

Name

Name=OpenH323GK

Default: OpenH323GK

Identificador del gatekeeper. El gatekeeper responderá solamente a los mensajes GRQs con este ID y utilizará éste en los mensajes enviados a sus terminales [12].

Tabla 6 - 2. Configuración usada en GK-UQROO y GK-MARTE

[Gatekeeper::Main] Fortytwo=42 Name=GK1 *[RoutedMode]* GKRouted=1 H245Routed=0 RemoveH245AddressOnTunneling=0 AcceptNeighborsCalls=1 AcceptUnregisteredCalls=1 SupportNATedEndpoints=1 DropCallsByReleaseComplete=0 Q931PortRange=31000-31999 H245PortRange=30000-30999 *[Proxy]* Enable=1 InternalNetwork=192.168.9.0/24 T120PortRange=40000-40999 RTPPortRange=50000-59999 ProxyForNAT=1 ProxyForSameNAT=1	. . . *[RasSrv::ARQFeatures]* ArjReasonRouteCallToGatekeeper=1 RemoveTrailingChar=# RoundRobinGateways=1 *[RoutingPolicy]* default=explicit,internal,parent,neighbor *[RasSrv::RRQAuth]* default=confirm *[GkStatus::Auth]* rule=allow *[Gatekeeper::Auth]* default=allow *[CTI::Agents]* VirtualQueueAliases=CC RequestTimeout=100 *[CTI::MakeCall]* EndpointAlias=DialOut

[Endpoint] Gatekeeper=148.247.21.3 Type=Gateway H323ID=GDL . . .	UseH450=0 Interface=*:1724 Gatekeeper=192.168.9.71 *[Neighbors::GK2]* GatekeeperIdentifier=GK2 Host=192.168.11.35 SendPrefixes=80 AcceptPrefixes=* ForwardHopCount=5 ForwardLRQ=depends

- **SECCIÓN [ROUTEDMODE]**

Los mensajes de señalización de llamada pueden ser manejados de dos maneras. El primer método es Señalización de llamada en Modo Directo, en el cual los mensajes de señalización de llamada son intercambiados directamente entre las terminales. El segundo método es Señalización de llamada mediante Gatekeeper. En este método, los mensajes de señalización de llamada son enrutados a través del gatekeeper entre las terminales. La selección de cual método utilizar es realizada por el gatekeeper [12].

Cuando se utiliza la señalización de llamadas en Modo Gatekeeper, el gatekeeper puede seleccionar si enrutar o no el canal de control H.245 (H.245 control channel) y los canales lógicos (logical channels) [12].

- Caso I. El gatekeeper no enruta estos canales. El canal de control H.245 (H.245 control channel) y los canales lógicos (logical channels) se establecen directamente entre las terminales [12].
- Caso II. El canal de control H.245 (H.245 control channel) se enruta entre las terminales a través del gatekeeper, mientras que los canales lógicos (logical channels) se establecen directamente entre las terminales [12].
- Caso III. El gatekeeper enruta el canal de control H.245 (H.245 control channel), así como también los canales lógicos (logical channels),

incluyendo el RTP/RTCP para audio y video, y el canal T.120 para datos. En este caso, ningún tipo de tráfico es intercambiado directamente entre las terminales. Esto es usualmente llamado un Proxy H.323, el mismo que puede ser considerado también como un gateway H.323-H.323 [12].

GKRouted

GKRouted=1

Default: 0

Definir si se habilita o no el ruteo de la señalización en modo gatekeeper [12].

H245Routed

H245Routed=1

Default: 0

Definir si se enruta también el Canal de Control H.245 a través del gatekeeper. Solamente tendrá efecto si el parámetro GKRouted=1 y el tunneling H.245 está deshabilitado para una llamada. Incluso cuando esta opción está deshabilitada, si Proxy o ProxyForNAT tiene efecto, siempre será enrutado el canal H.245 a través del gatekeeper para las llamadas que están pasando por el proxy [12].

RemoveH245AddressOnTunneling

RemoveH245AddressOnTunneling=1

Default: 0

Si la opción es TRUE, el gatekeeper removerá las h245Address cuando el indicador h245Tunneling esté en TRUE. Esto forzará a que la parte remota (remote party) permanezca en modo tunnelling [12].

AcceptNeighborsCalls

AcceptNeighborsCalls=1

Default: 1

Con esta característica habilitada, el hilo de señalización de llamada aceptará llamadas sin un CallRec preexistente dentro del CallTable, de tal manera que una terminal correspondiente a una dirección de destino en un mensaje Setup pueda ser encontrado en la tabla de registro, y el emisor de la llamada es su vecino o su GK padre. El gatekeeper además utiliza su propia dirección de señalización de llamada dentro de LCF en respuesta a un LRQ. Eso significa, la señalización de llamada será ruteada hacia el GK2 en llamadas GK-GK. Como resultado, los CDRs en el GK2 pueden mostrar correctamente el tiempo de conexión en lugar de 'unconnected' [12].

AcceptUnregisteredCalls

AcceptUnregisteredCalls=1

Default: 0

Con esta característica habilitada, el gatekeeper aceptará llamadas desde cualquier terminal no registrado. Sin embargo, esto permite riesgos en la seguridad. [12].

SupportNATedEndpoints

SupportNATedEndpoints=1

Default: 0

Permite a una terminal detrás de una NAT registrarse con el gatekeeper. Si esto se permite, el gatekeeper traducirá las direcciones IP dentro del canal Q.931 y H.245 hacia la IP de la NAT [12].

DropCallsByReleaseComplete

DropCallsByReleaseComplete=1

Default: 0

De acuerdo con la recomendación H.323, el gatekeeper podrá colgar una llamada enviando un mensaje RAS Disengage Request hacia las terminales. Sin embargo, algunas terminales mal implementados ignoran este comando. Con esta opción establecida en "on", el gatekeeper enviará mensajes Q.931 Release Complete en

lugar de mensajes RAS DRQ hacia ambas terminales para forzar a que ellos terminen la llamada [12].

Q931PortRange

Q931PortRange=20000-20999

Default: N/A (permitir que el SO asigne los puertos)

Especificar el rango de números de puerto TCP para canales de señalización Q.931. Hay que presente que el tamaño del rango limita el número de llamadas concurrentes [12].

H245PortRange

H245PortRange=30000-30999

Default: N/A (permitir que el SO asigne los puertos)

Especificar el rango de puertos TCP para los canales de control H.245. Hay que tener en cuenta que el tamaño del rango podría limitar el número de llamadas concurrentes [12].

- **SECCIÓN [PROXY]**

Enable

Enable=1

Default: 0

Define si se habilita o no la función proxy. Se debe de habilitar el enrutamiento en modo gatekeeper primero. No se tiene que especificar el enrutamiento H.245. Este será automáticamente utilizado si se requiere [12].

InternalNetwork

InternalNetwork=10.0.1.0/24

Default: N/A

Aquí se definen las redes detrás del proxy. Se permite el uso de múltiples redes internas. El proxy enruta canales solamente de las comunicaciones entre una

terminal en la red interna y una externa. Si no se especifica esto, todas las llamadas serán llevadas a través del proxy. Si se utiliza un GnuGk detrás de una NAT y el parámetro ExternalIP de la sección [Gatekeeper::Main] está configurado, entonces no es obligatorio establecer éste, tal como es auto-detectado al inicio. Utilizar esta configuración simplemente sustituirá la configuración detectada por defecto [12].

T120PortRange

T120PortRange=40000-40999

Default: N/A (permitir que el SO asigne los puertos)

Especifica el rango de números de puertos TCP para los canales de datos T.120. Hay que tener en cuenta que el tamaño del rango podría limitar el número de llamadas concurrentes [12].

RTPPortRange

RTPPortRange=50000-59999

Default: 1024-65535

Especifica el rango de números de puertos UDP para los canales RTP/RTCP. Hay que tener en cuenta que el tamaño del rango podría limitar el número de llamadas concurrentes [12].

ProxyForNAT

ProxyForNAT=1

Default: 1

Si se activa, el gatekeeper pasará por el proxy aquellas llamadas en las cuales uno de las terminales participantes está detrás de una NAT. Esto asegura que el stream RTP/RTCP pueda penetrar dentro de la NAT sin modificar éste. Sin embargo, la terminal que se encuentra detrás de la NAT debe utilizar el mismo puerto para enviar y recibir el stream RTP/RTCP [12].

ProxyForSameNAT

ProxyForSameNAT=0

Default: 0

Define si se habilita el proxy para llamadas entre terminales desde la misma NAT [12].

- **SECCIÓN [ENDPOINT]**

El gatekeeper puede trabajar como una terminal registrándose con otro gatekeeper. Con esta característica, se puede construir fácilmente jerarquías de gatekeepers. Esta sección define las características de una terminal para el gatekeeper [12].

Gatekeeper:

Gatekeeper=10.0.1.1

Default: no

Define un gatekeeper padre (parent gatekeeper) para la terminal (el gatekeeper), con el cual se va a registrar [12].

Type

Type=Gateway

Default: Gateway

Definir el tipo de terminal. Los valores válidos son Gateway o Terminal [12].

H323ID:

H323ID=CitronProxy

Default: <Name>

Especificar el o los alias H.323 ID para la terminal. Múltiples alias pueden separarse con comas [12].

- **SECCIÓN [RASSRV::ARQFEATURES]**

ArjReasonRouteCallToGatekeeper

ArjReasonRouteCallToSCN=0

Default: 1

Si es "yes" (1), el gatekeeper rechaza un ARQ respondido sin que haya un pre-existente CallRec encontrado en la CallTable con razón routeCallToGatekeeper en modo de enrutamiento. La terminal liberará la llamada inmediatamente y reenviará un call Setup hacia el gatekeeper [12]

RemoveTrailingChar

RemoveTrailingChar=#

Default: N/A

Especificar el caracter de rastreo a ser removido en destinationInfo. Por ejemplo, si la terminal contiene incorrectamente el caracter de terminación como `#' dentro destinationInfo, se puede remover éste mediante esta opción [12].

RoundRobinGateways

RoundRobinGateways=0

Default: 1

Habilitar/Deshabilitar la selección de un gateway mediante round-robin, si más de un gateway se emparejan con un número marcado. Si se deshabilita, el primer gateway disponible será seleccionado. De otra manera, las llamadas subsiguientes serán enviadas hacia los gateways utilizando el modo round-robin [12].

- **SECCIÓN [ROUTINGPOLICY]**

Esta sección explica el funcionamiento de las diferentes políticas de enrutamiento del gatekeeper.

La siguiente es la configuración por defecto para las políticas de enrutamiento (routing policies):
default=explicit,internal,parent,neighbor
Si una política no se cumple, se tratará con la mencionada política [12].

- **SECCIÓN [RASSRV::RRQAUTH]**

Especifica la acción realizada sobre los mensajes de recepción RRQ (confirm or deny) para el módulo AliasAuth. El primer alias (éste será principalmente un H323ID) de la terminal a ser registrado es buscado en esta sección. Si un parámetro es encontrado, el valor será aplicado por regla general. Una regla consiste de condiciones separadas por "&". Un registro es aceptado cuando todas las condiciones coinciden [12].

- **SECCIÓN [GKSTATUS::AUTH]**

En esta sección se definen un número de reglas para determinar quienes están permitidos de conectarse al gatekeeper vía puerto de estado (vía telnet). Quien quiera que tenga acceso al puerto de estado tiene un control completo sobre el gatekeeper [12].

- **SECCIÓN [CTI::AGENTS]**

Esta sección permite la configuración de las llamadas colas virtuales (virtual queues) para permitir distribución de llamadas entrantes por una aplicación externa, mediante el puerto de estado. Una cola virtual tiene un alias H.323 que puede ser llamado como una terminal [12].

En el arribo de un mensaje ARQ a una cola virtual, el gatekeeper señala una RouteRequest en el puerto de estado y espera a que una aplicación externa responda ya sea con un RouteReject (entonces el ARQ será rechazado) o con un RouteToAlias/RouteToGateway el cual conduce al ARQ a ser reescrito de este

modo la llamada será ruteada hacia el alias (eg. call center agent) especificado por la aplicación externa.
Si no se recibe ninguna respuesta después de un período de tiempo, la llamada es finalizada [12].

VirtualQueueAliases

Default: none
Este parámetro define una lista de alias H.323 para las colas virtuales (Utilizado con la vqueue RoutingPolicy) [12].

RequestTimeout

Default: 10
Tiempo de espera en segundos para que la aplicación externa responda el RouteRequest. Si no se recibe ninguna respuesta durante este tiempo un ARJ será enviado hacia la terminal que llama (caller) [12].

- **SECCIÓN [CTI::MAKECALL]**

Esta sección contiene las características para el puerto de estado de comando MakeCall [12].

EndpointAlias

EndpointAlias=DialOut
Default: InternalMakeCallEP
Define el alias de la terminal utilizado para marcar [12].

UseH450

UseH450=1
Default: 0
Usa una transferencia H.250 en lugar de un mensaje de fondo para transferir la llamada desde una terminal hacia el destino [12].

Interface

Interface=192.168.1.1:1730

Default: *:1722

Interfaz y puerto a usar para la terminal [12].

Gatekeeper

Gatekeeper=192.168.1.2

Default: 127.0.0.1

Dirección IP del gatekeeper al cual se registrará la terminal [12].

- **SECCIÓN [NEIGHBORS::GK2]**

Las secciones que empiezan con [Neighbor:: son para las configuraciones especificas de un vecino.] [12].

GatekeeperIdentifier

Identificador del Gatekeeper para este vecino. Si esta opción no se especifica, el identificador es tomado de la segunda parte del nombre de esta sección Neighbor::.

Host

Host=192.168.1.1

Default: N/A

Una dirección IP para este vecino.

SendPrefixes

SendPrefixes=004,002:=1,001:=2

Default: N/A

Una lista de prefijos que este vecino espera recibir para los LRQs.

AcceptPrefixes

AcceptPrefixes=*

Default: *

Una lista de prefijos que el gatekeeper aceptará en los LRQs recibidos desde este vecino.

ForwardHopCount

ForwardHopCount=2

Default: N/A

Si el gatekeeper recibe un LRQ en el que el destino es desconocido, éste podría reenviar este mensaje hacia sus vecinos. Cuando el gatekeeper recibe un LRQ y decide que el mensaje debe ser reenviado hacia otro gatekeeper, éste primero decrementa el campo hopCount del LRQ. Si hopCount ha llegado a 0, el gatekeeper no reenviará el mensaje. Esta opción define el número de gatekeepers a través de los cuales un LRQ puede propagarse.

ForwardLRQ

ForwardLRQ=always | never | depends

Default: depends

Esta configuración determina si el LRQ recibido deberá o no ser reenviado. Si es *always*, reenvía LRQs de manera incondicional, si es *never*, bloquea los LRQ reenviados y si es depends le indica al gatekeeper que reenviará LRQ solamente si su contador de saltos (hop count) es mayor a 1. Este ajuste puede ser superpuesto o deshabilitado con la configuración de un vecino particular.

6.2 Conjuntos de Trazas Colectadas

La Tabla 6 - 3 muestra la configuración usada en las terminales de la Figura 6-1 para efectuar las mediciones.

Tabla 6 - 3. Configuración de Terminales

Conjunto de Trazas (CT)	PC1/TELE1	PC2/TELE2	PC3/TELE3
CT1, CT5, CT6, CT7, CT8	G.711-20ms	G.711-40ms	G.711-60ms
CT2, CT3, CT4	G.729-20ms	G.729-40ms	G.729-60ms

Las mediciones correspondientes a los conjuntos de trazas ilustrados en la Tabla 6 - 3, fueron colectadas de la siguiente manera:

- Tres llamadas de prueba fueron establecidas de manera simultáneas entre las terminales PC1/TELE1, PC2/TELE2 y PC3/TELE3, ver Figura 6 - 1.
- Las configuraciones usadas en las llamadas de prueba están basadas en dos parámetros: tipo de códec (G.711 y G.729) y tamaño de paquete de voz (20ms, 40ms y 60ms).
- Los periodos de medición fueron de 60 minutos por cada llamada de prueba (duración de llamada).
- Por cada periodo de medición (una hora), tres trazas de jitter de arribo, OWD y PLR fueron obtenidas.
- Los ocho conjuntos de trazas seleccionados contienen 23.76 millones de paquetes RTP, correspondiente a 216 trazas de jitter de arribo, 216 trazas de OWD y 216 trazas de PLR, medidos en horas típicas de trabajo.

Tabla 6 - 4. Descripción de los Conjuntos de Trazas Colectados

Conjunto de trazas (CT)	Período de medición	No. Total de trazas	No. Total de paquetes	No. Total de bytes
1	10/Julio/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	937,980,000
2	24/Julio/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	296,460,000
3	26/Julio/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	296,460,000
4	27/Julio/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	296,460,000
5	01/Agosto/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	937,980,000
6	02/Agosto/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	937,980,000
7	03/Agosto/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	937,980,000
8	28/Agosto/2012 11:00 – 19:00 horas	27 trazas Jitter arribo 27 trazas OWD 27 trazas PLR	2,970,000	937,980,000

La Tabla 6 - 4 muestra la descripción detallada de los conjuntos de trazas colectadas durante el proceso de medición.

Para llevar a cabo la captura de las trazas, se utilizó el analizador de protocolos de red Wireshark. El Wireshark se instaló y ejecutó en el gatekeeper GK-UQROO como se ilustra en la Figura 6 - 1.

Para iniciar la captura tenemos abrir Wireshark y elegir la opción “Capture options”, como se muestra en la Figura 6 - 2.

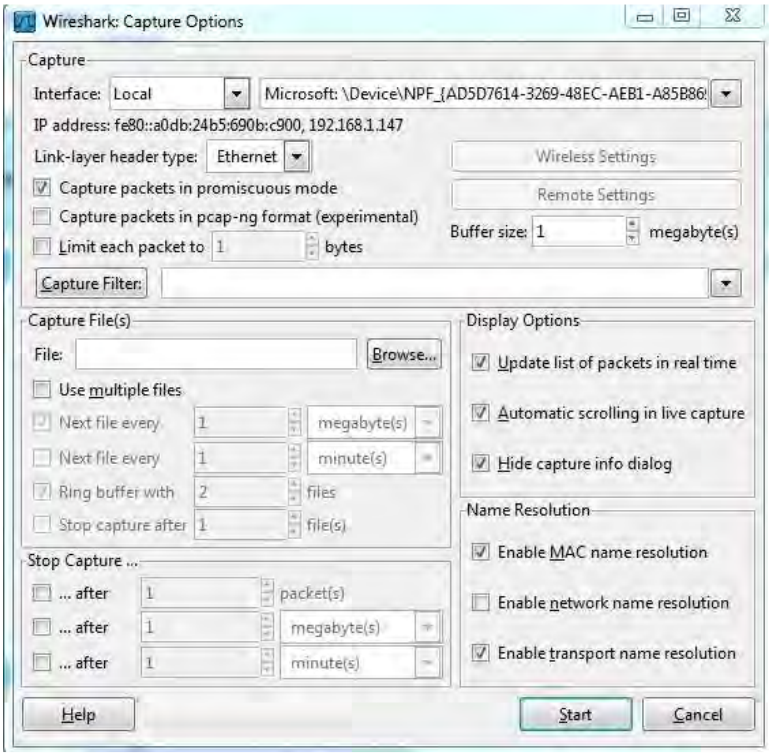


Figura 6 - 2. Opciones de Captura del Wireshark

Posteriormente, es necesario verificar que la dirección de la PC donde se encuentra instalado el Wireshark sea seleccionada en la sección IP address y procedemos a configurar los siguientes campos como sigue:

Capture filter: Aquí especificaremos los filtros a utilizar durante la medición, tales como, tipo de paquetes que se capturaran (TCP, UDP, etc) y entre que dispositivos. Especificaremos las direcciones IP origen y destino en el siguiente formato:

Host [IP origen] and host [IP destino] and UDP

Tenemos que considerar que debido a que estos dos gatekeepers se comunican a través de internet, las direcciones IP deben ser direcciones públicas. En este caso el filtro quedó de la siguiente manera:

Host 192.100.164.19 and host 148.247.21.3 and UDP

De esta manera ya estamos limitando al Wireshark a sólo capturar los paquetes UDP que se intercambian los gatekeepers GK-UQROO y GK-MARTE.

En el apartado de File designamos el nombre del archivo con extensión ".pcap" junto con la ruta en la cual se guardará. Posteriormente debemos activar las casillas de verificación "Use multiple files", después la segunda casilla "next file every" la configuramos a una hora, y por último "stop capture after" lo configuramos con el valor de nueve.

De esta manera ya se configuró el Wireshark para que capture paquetes UDP entre GK-UQROO y GK-MARTE en intervalos de a una hora, durante nueve horas, como se ilustra en la Tabla 6 - 4.
Al final de la captura se generaran nueve archivos con extensión ".pcap". Para obtener las trazas mostradas en la Tabla 6 - 4, cada uno de estos archivos debe pasar por un proceso que se describirá a continuación:

Paso 1: Abrir uno de los archivos ".pcap". Una vez abierto, decodificamos los flujos UDP como RTP mediante la opción "decode as" del menú desplegable "Analyze". Ver Figura 6 – 3.

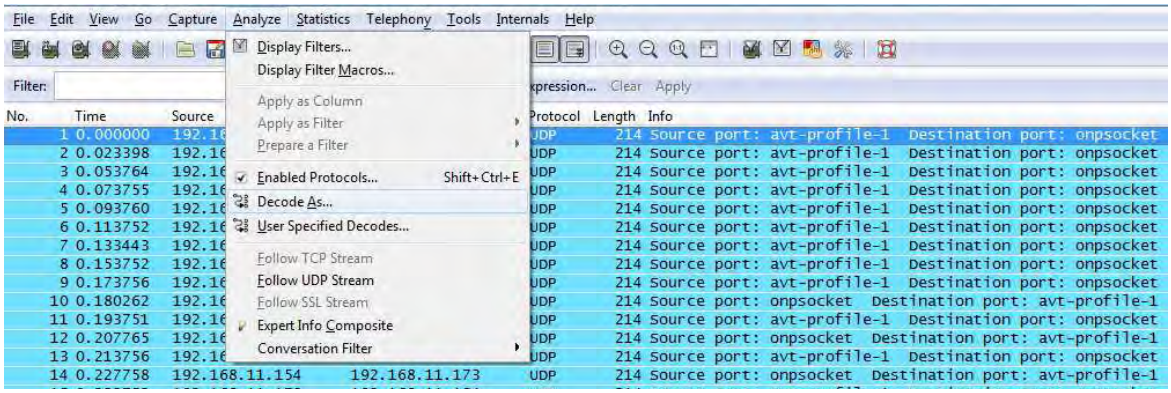


Figura 6 - 3. Decodificación Paquetes UDP

Paso 2: Elegir RTP en el apartado derecho y clic en aceptar (Figura 6 – 4)

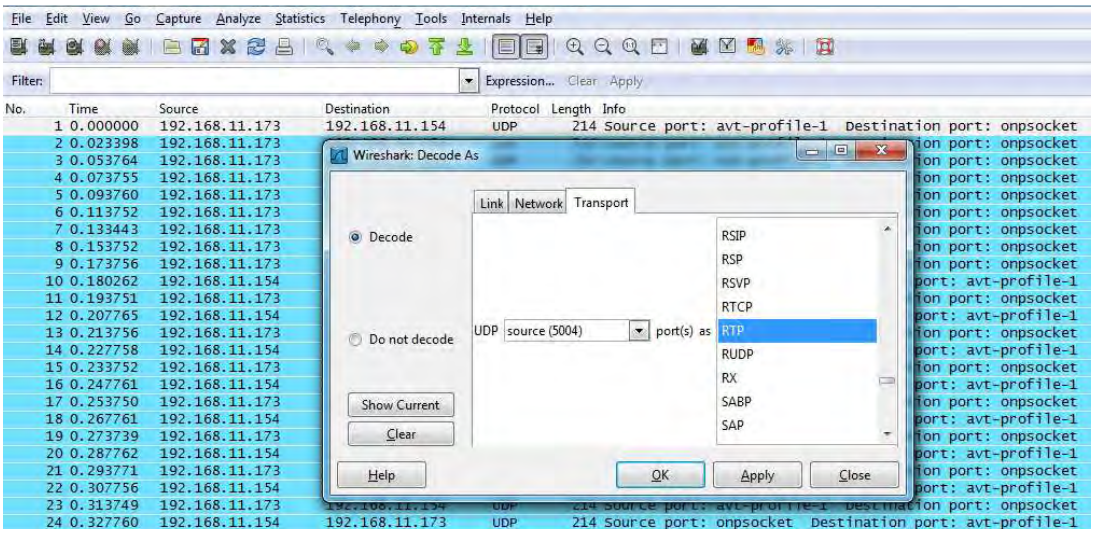


Figura 6 - 4. Seleccionar RTP

Se empezará a decodificar el paquete UDP a RTP, dependiendo de las características del equipo es lo que tardará este proceso.

Paso 3: Este proceso hay que repetirlo hasta que no quede ningún flujo UDP visible en la pantalla principal de Wireshark y sólo sea vean flujos RTP, ver figura 6 - 5.

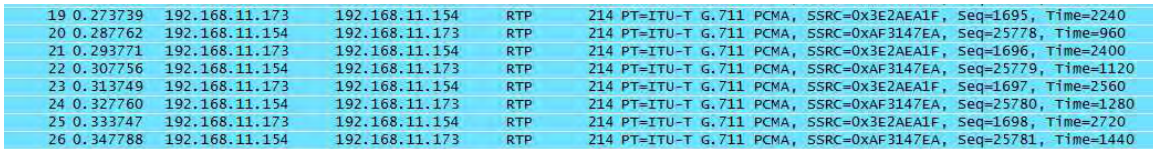


Figura 6 - 5. Flujos RTP

Paso 4: Ahora podemos ver los flujos de datos de las llamadas, para ello nos vamos al menú Telephony y en el submenú RTP seleccionamos “Show all streams”, como se muestra en la Figura 6 – 6.

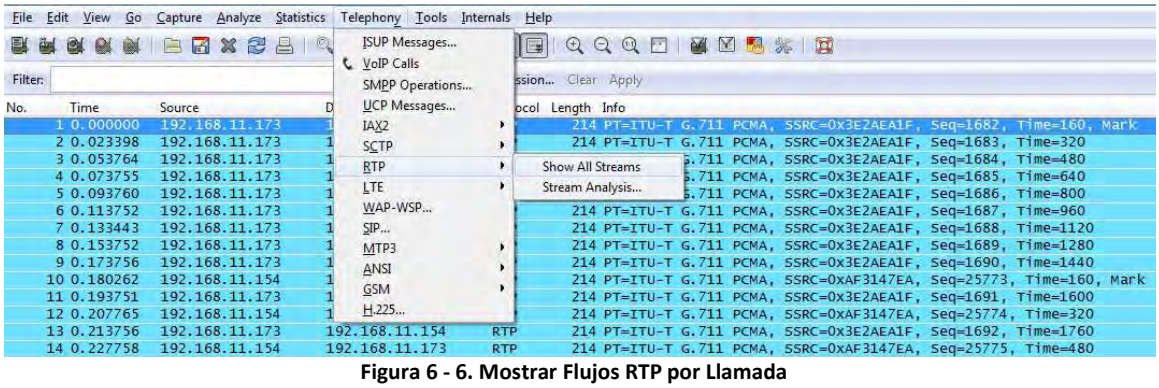


Figura 6 - 6. Mostrar Flujos RTP por Llamada

Nos aparecerá la siguiente pantalla con los flujos de llamada; en este caso uno de entrada y uno de salida, ver Figura 6 - 7.

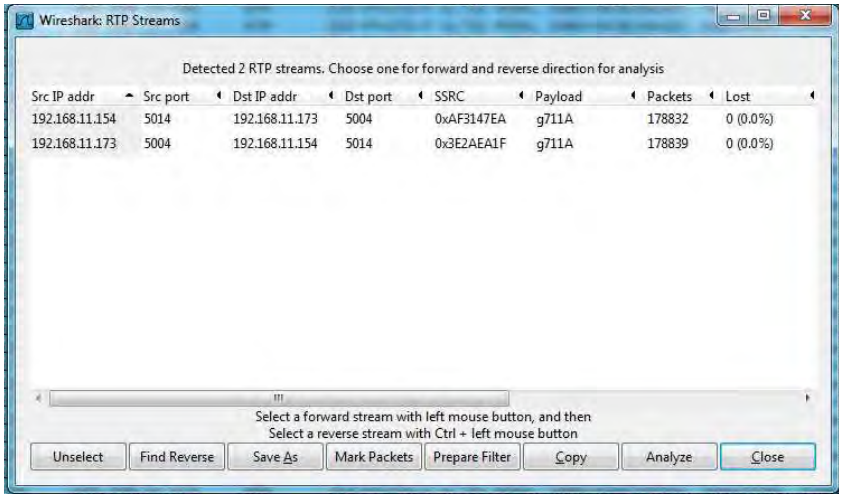


Figura 6 - 7. Flujos RTP por Llamada

Dependiendo del número de llamadas realizadas es el número de flujos RTP por llamada que nos aparecerá, son dos flujos por llamada, esto quiere decir que por 3 llamadas tendríamos 6 flujos; 3 de ida y 3 de vuelta.

Paso 5: Wireshark es capaz de obtener una estadística a cerca de algunos parámetros de red como jitter de arribo, pérdidas de paquetes, etc. Para esto, en la ventana “streams RTP” (figura 6-7) seleccionamos uno de los flujos y damos clic en “Analize”. En el cuadro de diálogo que aparecerá a continuación podremos ver los valores de jitter de arribo entre paquetes consecutivos, los números de secuencia de los paquetes, etc., ver Figura 6 - 8.

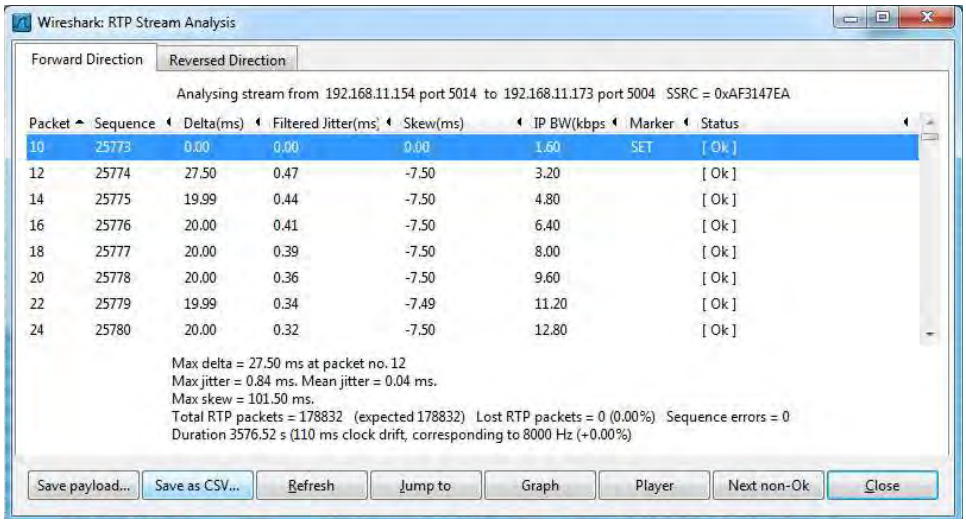


Figura 6 - 8. Estadísticas de Paquetes de un Flujo RTP

Para guardar estos datos podemos hacerlo con el botón “Save as CSV” el cual guardará un archivo en formato CSV con los parámetros de QoS de un flujo RTP. Hay que repetir el proceso desde el paso 5 con los demás flujos RTP.

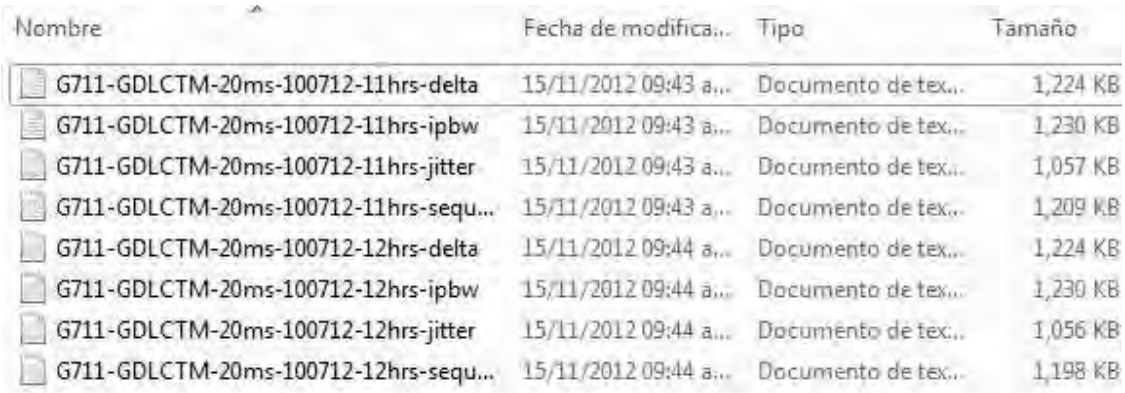
Cabe mencionar que el proceso de filtrado realizado anteriormente, corresponde a una hora de medición, por lo tanto hay que repetirlo para las demás horas de cada día de medición.

Para obtener finalmente las trazas, se abrieron los archivos CSV mediante Excel (ver Figura 6 - 9), se extrajeron las columnas contenidas en el mismo y se guardaron en archivos TXT (ver Figura 6 - 10).

	A	B	C	D	E	F	G	H	I	J	K
1	Forward										
2	Packet	Sequence	Time stamp	Delta (ms)	Jitter (ms)	Skew(ms)	IP BW (kbps)	Marker	Status		
3	74	53140	1120	0	0	0	1.6		[Ok]	08/27/2012 1	214
4	76	53141	1280	22.79	0.17	-2.79	3.2		[Ok]	08/27/2012 1	214
5	78	53142	1440	18.75	0.24	-1.54	4.8		[Ok]	08/27/2012 1	214
6	80	53143	1600	16.77	0.43	1.69	6.4		[Ok]	08/27/2012 1	214
7	83	53144	1760	35.24	1.35	-13.55	8		[Ok]	08/27/2012 1	214
8	85	53145	1920	19.95	1.27	-13.5	9.6		[Ok]	08/27/2012 1	214
9	88	53146	2080	39	2.38	-32.5	11.2		[Ok]	08/27/2012 1	214
10	90	53147	2240	21.01	2.29	-33.51	12.8		[Ok]	08/27/2012 1	214

Figura 6 - 9. Ejemplos de un Archivo .csv

Por cada archivo CSV, este programa creará cuatro archivos TXT, correspondientes a un flujo RTP capturado en una hora determinada.



Nombre	Fecha de modifica...	Tipo	Tamaño
G711-GDLCTM-20ms-100712-11hrs-delta	15/11/2012 09:43 a...	Documento de tex...	1,224 KB
G711-GDLCTM-20ms-100712-11hrs-ipbw	15/11/2012 09:43 a...	Documento de tex...	1,230 KB
G711-GDLCTM-20ms-100712-11hrs-jitter	15/11/2012 09:43 a...	Documento de tex...	1,057 KB
G711-GDLCTM-20ms-100712-11hrs-sequ...	15/11/2012 09:43 a...	Documento de tex...	1,209 KB
G711-GDLCTM-20ms-100712-12hrs-delta	15/11/2012 09:44 a...	Documento de tex...	1,224 KB
G711-GDLCTM-20ms-100712-12hrs-ipbw	15/11/2012 09:44 a...	Documento de tex...	1,230 KB
G711-GDLCTM-20ms-100712-12hrs-jitter	15/11/2012 09:44 a...	Documento de tex...	1,056 KB
G711-GDLCTM-20ms-100712-12hrs-sequ...	15/11/2012 09:44 a...	Documento de tex...	1,198 KB

Figura 6 - 10. Archivos extraídos en formato .txt

Capítulo 7
Análisis de Desempeño y Caracterización de Parámetros de QoS

En la presente sección se presenta el análisis y caracterización del conjunto de mediciones que se muestran en la Tabla 6.4.

Como se menciona en la sección 5.2, dos importantes métodos para evaluar el desempeño en una comunicación de voz son el Modelo E y el MOS. En esta tesis se realizó el análisis de desempeño del conjunto de llamadas de prueba, mediante el cálculo del factor R (ver ecuación 5.7) y el correspondiente mapeo a valores de MOS (ver ecuación 5.10).La Figura 7-1 muestra los valores promedio de MOS correspondiente a los Conjuntos de Trazas: CT1, CT5, CT6, CT7 y CT8 por cada hora de medición.

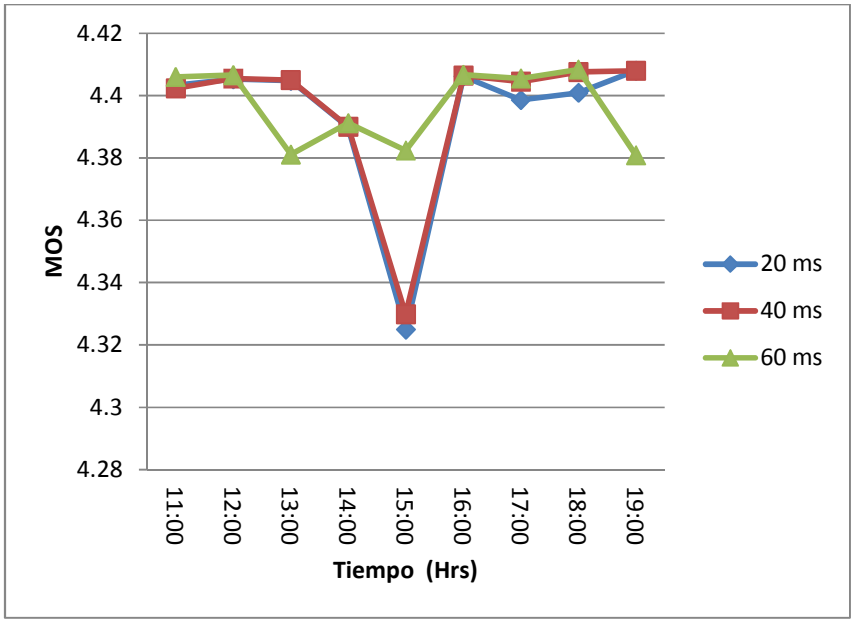


Figura 7 - 1. Valores Promedio de MOS: G.711

De acuerdo a la Tabla 6 - 3, en la Figura 7 - 1 se presentan los valores promedio de MOS a lo largo de todos los días de medición, haciendo uso del CODEC G.711 y tamaños de paquete de 20ms, 40ms y 60ms. Se observa que mientras más pequeño sea el tamaño de paquete, menor es la calidad de servicio, es decir, flujos de paquetes de voz de tamaño mayor son menos sensibles a degradaciones en la calidad de la voz. Por otro lado, se puede observar que en estos flujos de voz, los valores críticos de calidad de servicio se dieron a las 13:00, 15:00 y 19:00 horas, por tal motivo, más adelante se hace un estudio más detallado sobre el comportamiento del OWD, PLR y jitter de arribo en estos instantes de tiempo.

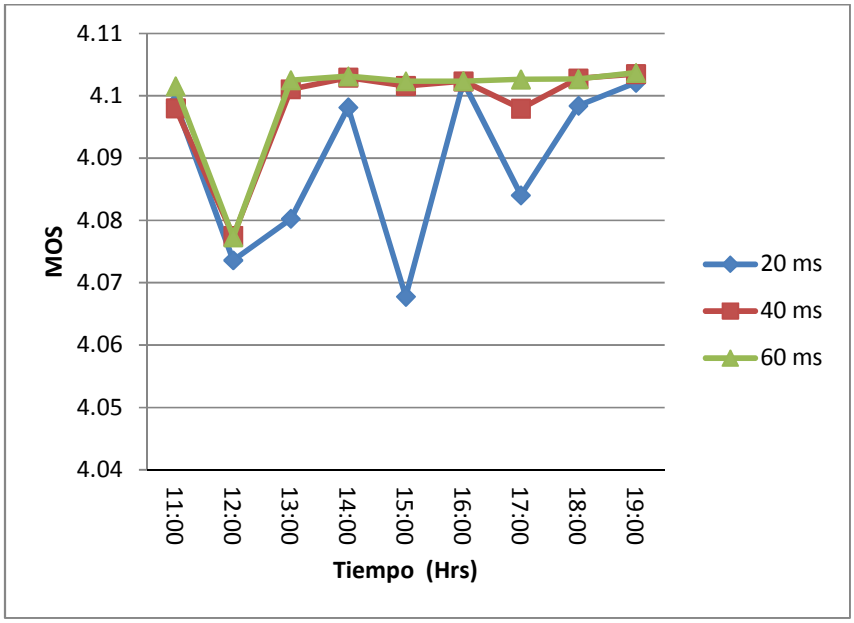


Figura 7 - 2. Valores Promedios de MOS por Hora: G.729

La Figura 7 - 2 muestra los valores promedio de MOS por hora, correspondiente a CT2, CT3 y CT4; en los cuales se utilizó el CODEC G.729 y tamaños de paquete de 20ms, 40ms y 60ms. En esta figura se puede observar el mismo comportamiento al de la Figura 7 - 2, es decir, paquetes de voz de tamaño mayor son menos sensibles a degradaciones en la calidad de la voz. Sin embargo, se puede ver que en estos flujos de voz, los valores críticos de calidad de servicio se dieron a las 12:00 y 15:00 horas, por tal motivo, más adelante se hace un estudio detallado sobre el comportamiento del OWD, PLR y jitter de arribo en estas horas de medición.

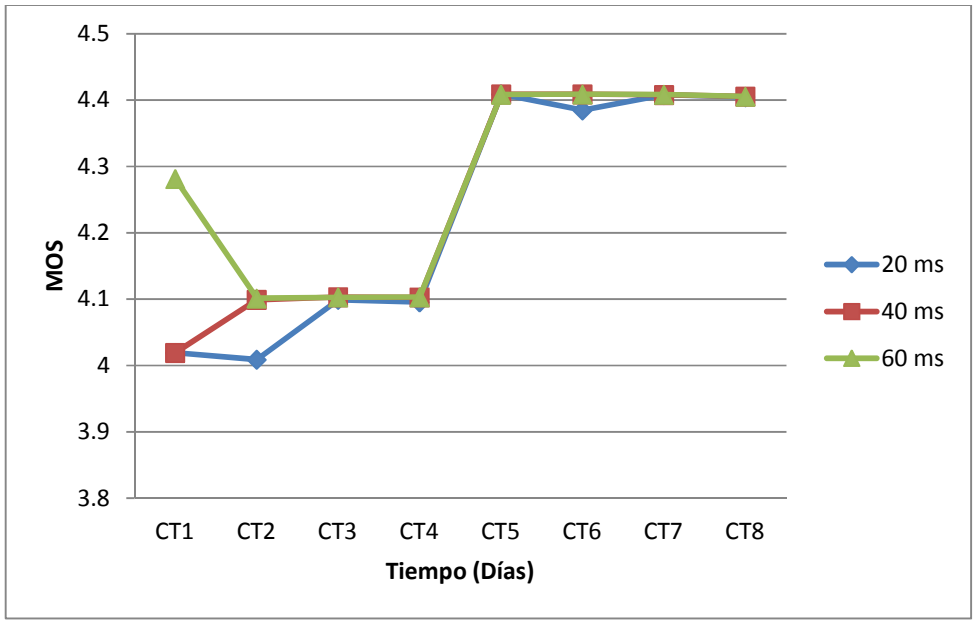


Figura 7 - 3. Valores Promedio de MOS por Día: G.711 y G.729

Para identificar los días que presentaron mayor degradación en la calidad de la voz, se calcularon los valores promedio de MOS por día, como se muestra en la Figura 7 - 3. Cómo se puede observar en la Figura 7 – 3, los valores más bajos de MOS se dieron en el conjunto de trazas CT1 y CT2. El conjunto de trazas CT1 corresponden a una medición con G. 711, mientras que el conjunto de trazas CT2 pertenecen a G.729.

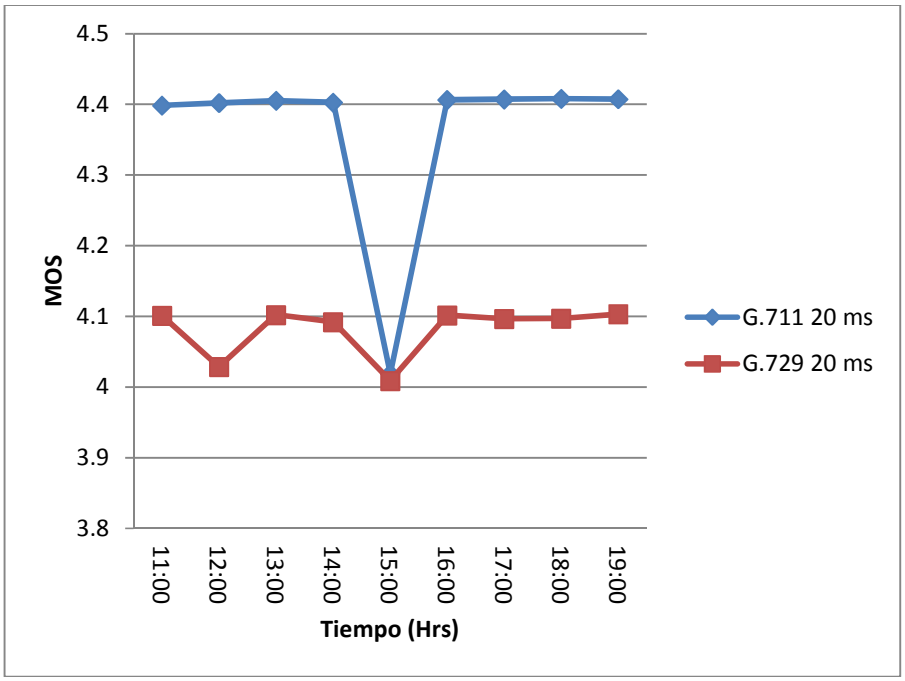


Figura 7 - 4. Valores de MOS para Flujos de 20ms: G.711 y G.729

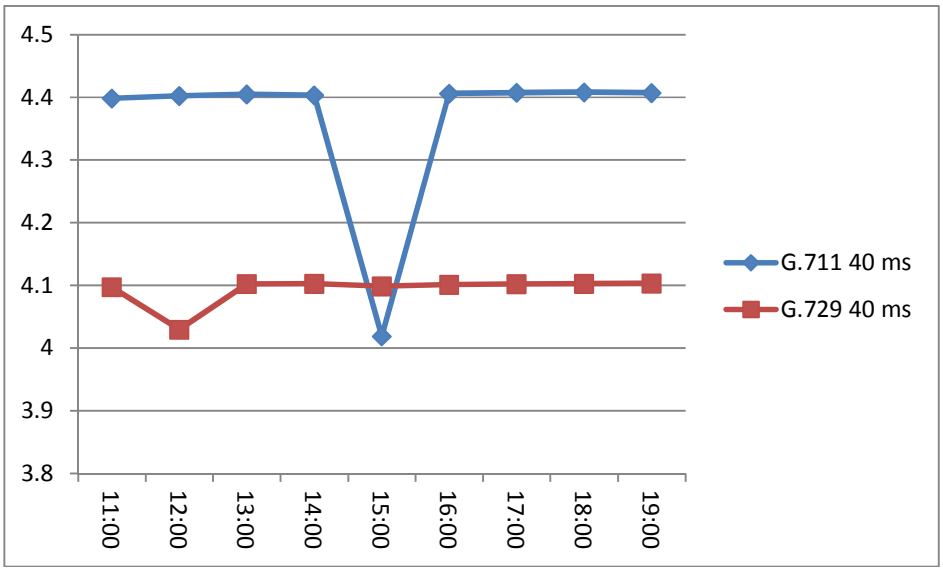


Figura 7 - 5. Valores de MOS para flujos de 40ms: G.711 y G.729

Con el objetivo de evaluar el desempeño a nivel de tipo de codificador, se realizó la comparativa de los valores de MOS entre CT1 y CT2 con cada uno de los distintos tamaños de paquete utilizados (ver Figuras 7 - 4, 7 - 5 y 7 - 6).

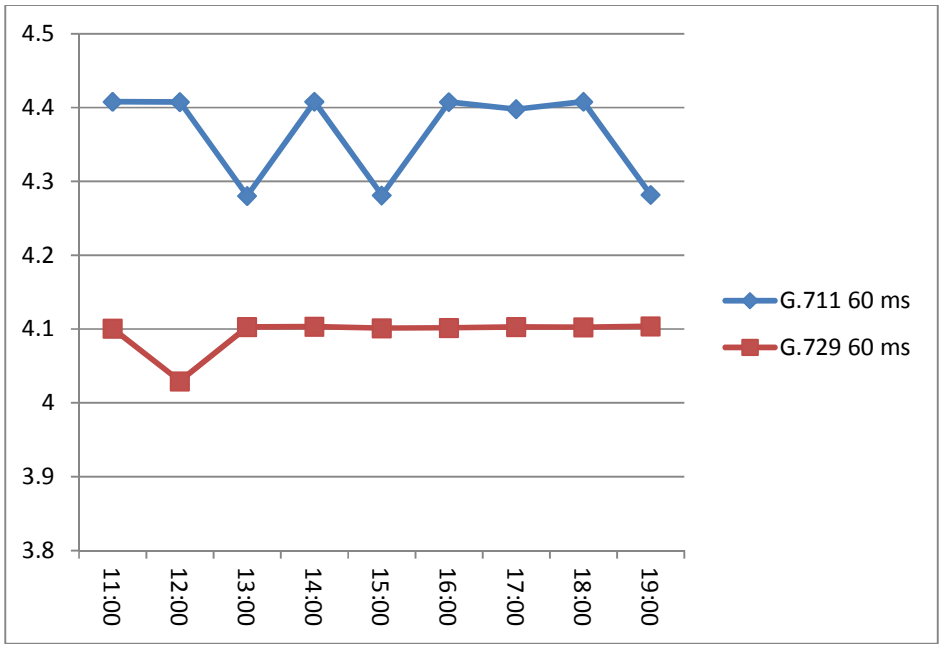


Figura 7 - 6. Valores de MOS para 60ms: G.711 y G.729

Las Figuras 7 – 4, 7 – 5 y 7 – 6 hacen una comparación del nivel de calidad de servicio en función de valores de MOS entre flujos de voz codificados con G.711 y G.729 a diferentes tamaños de paquete. Claramente se nota que G.711 siempre está por encima de G.729. Esto es debido a que G.711 no comprime la voz y por tal motivo garantiza mayor calidad de servicio en la comunicación que G.729.

Con el objetivo de estudiar la relación entre OWD, PLR y MOS en flujos de voz codificados mediante G.711, se realizó la comparación entre estos parámetros en CT1 (ver Figuras 7 - 7, 7 - 8 y 7 - 9).

La comparación entre el OWD y MOS se muestra en la Figura 7-7 y se puede observar que no existe relación alguna entre estos dos parámetros, debido a que los valores de OWD presentados son muy bajos y por tanto no tienen impacto en la calidad de servicio.

La Figura 7-8 ilustra la comparación entre PLR y MOS, de la cual podemos ver que existe una alta correlación entre los valores de PLR y MOS, es decir, a valores grandes de PLR (1.3%-3.7%) corresponden valores bajos de MOS (4.28-4.02).

La comparación entre OWD y PLR se presenta en la Figura 7-9, de esta figura podemos ver que no existe correlación alguna entre estos dos parámetros.

De este análisis, se puede concluir que las pérdidas de paquetes tuvieron mayor impacto en la calidad de servicio sobre los flujos de voz bajo el esquema de codificación G.711.

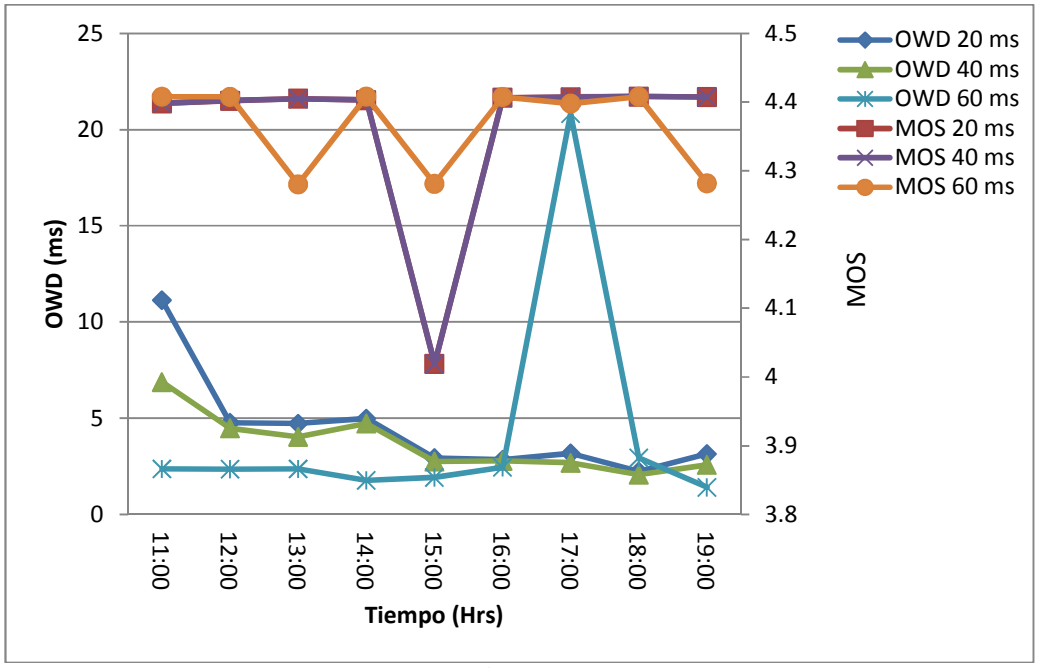


Figura 7 - 7. Comparación entre OWD y MOS para CT1

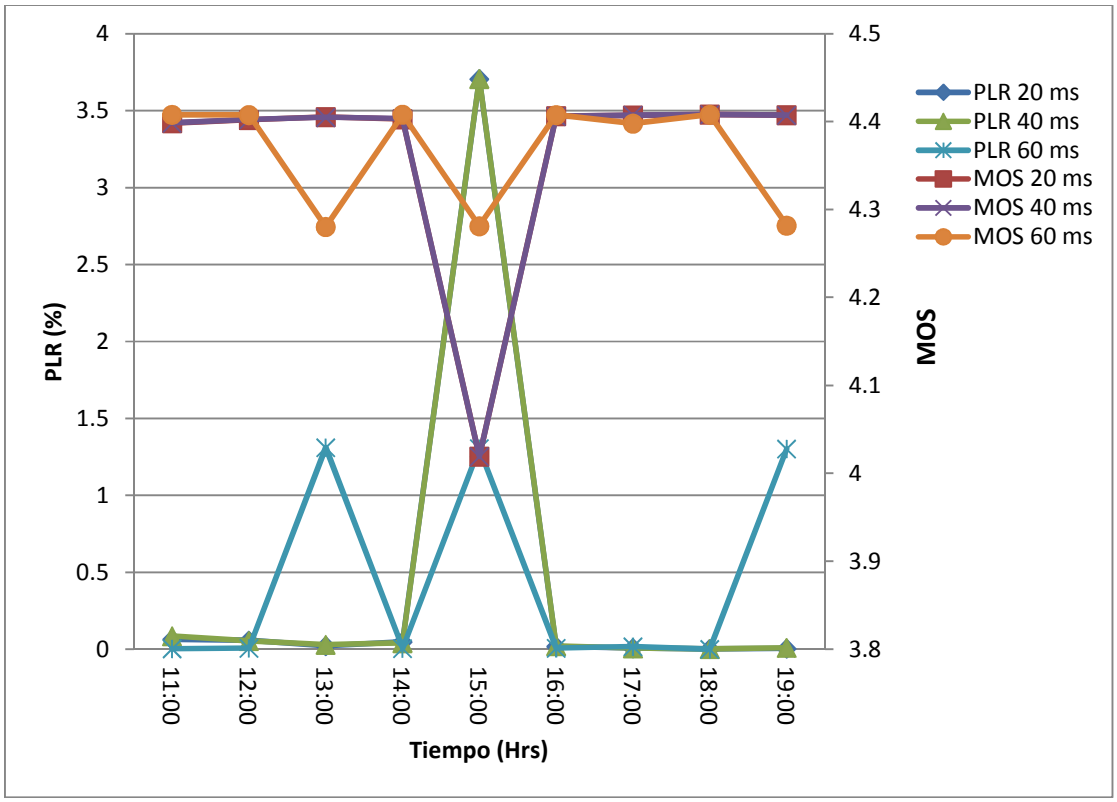


Figura 7 - 8. Comparación entre PLR y MOS para CT1

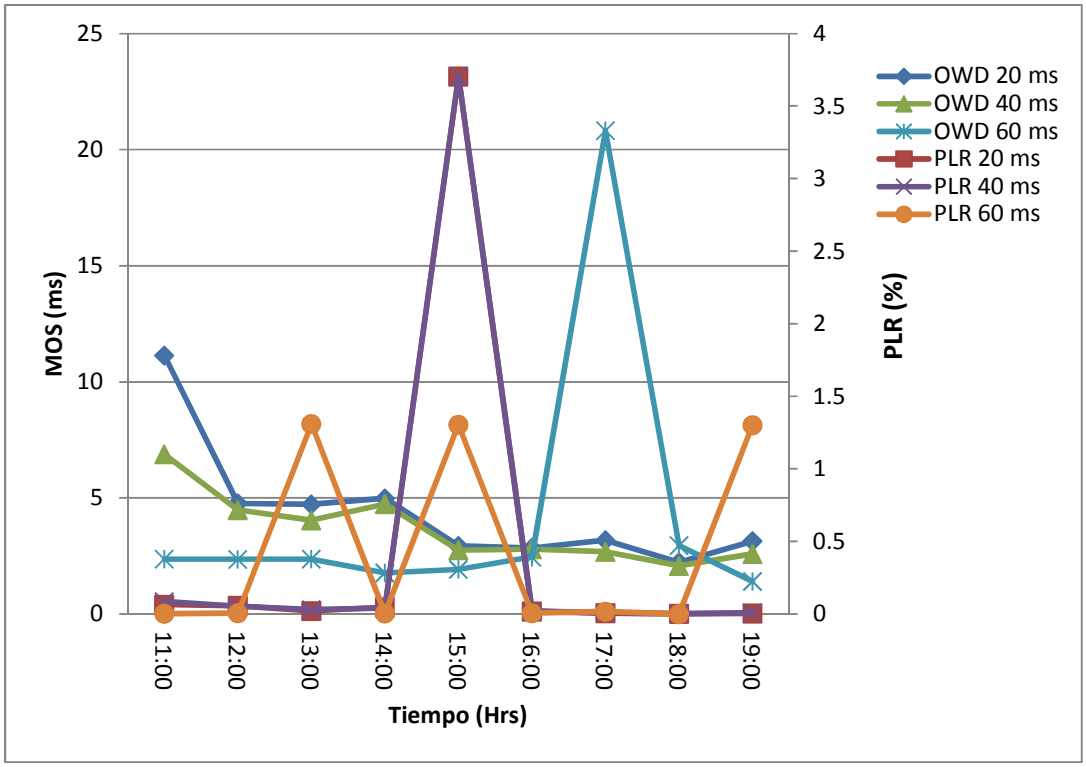


Figura 7 - 9. Comparación entre OWD y PLR para CT1

Debido a que el PLR tuvo mayor impacto en la QoS, se analizó el comportamiento de dicho parámetro en las horas donde se detectaron los niveles más bajos de QoS (ver Figura 7-8).

En [11] se presenta que existe una relación entre el PLR y el jitter de arribo, donde esta relación puede ser deducida a partir de la ecuación (5.5) como sigue:

Si el paquete $K-1$ se pierde:

$$IAT(K,K-2)=J^{K}(L)+(2)IDT \qquad 7.1$$

Entonces, si n paquetes consecutivos se pierden,

$$IAT(K,K-n-1)=J^{K}(L)+(n+1)(IDT) \qquad 7.2$$

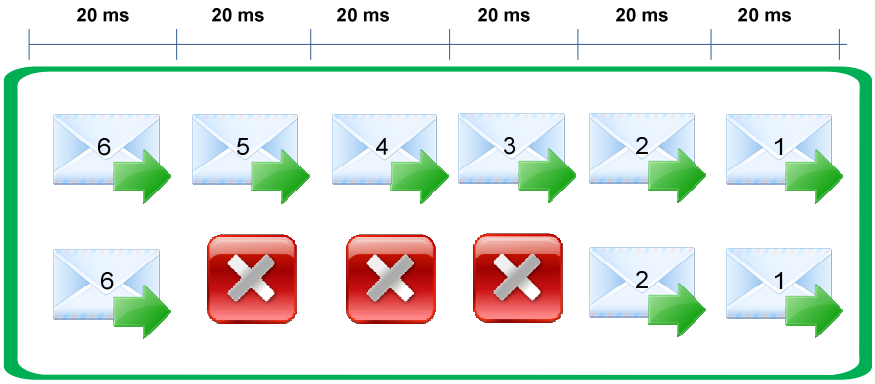


Figura 7 - 10. Relación entre Jitter de Arribo y PLR

Tomando como referencia la ecuación (7.2) y la Figura 7 – 10 se puede ver cómo están relacionados el jitter de arribo y el PLR. Por ejemplo, en esta tesis usamos tamaños de paquete de 20ms, 40ms y 60ms, los cuales son transmitidos a tasa constante, es decir 1 paquete/20ms, 1 paquete/40ms y 1 paquete/60ms, respectivamente. Sin embargo, cuando los paquetes de voz son transportados sobre una red IP y arriban al receptor, estos pueden experimentar variaciones de retardo y pérdida de paquetes.

En ausencia de pérdidas, los paquetes arriban a su destino a retardos variables, pero aproximados al tamaño de paquete. Sin embargo cuando hay pérdidas, estas variaciones de retardos (jitter de arribo) aumentan en base al múltiplo $(n+1)$, donde n es el número de paquetes perdidos de manera consecutiva.

Suponiendo que el flujo de paquetes de la Figura 7-11 son transportados sobre una red ideal donde los paquetes arriban a la misma tasa de salida ($J^K(L)=0$), entonces, si se pierden tres paquetes consecutivos (2, 3 y 4), el valor de jitter de arribo entre los últimos dos paquetes consecutivos (2 y 6) será igual a 80ms, $IAT(K,K-3-1)=0+(3+1)(20ms)=80ms$.

Realizando el proceso inverso se podría también calcular el número de paquetes perdidos; tomamos el valor de jitter de arribo (80 ms) lo dividimos entre 20 ms que es el tamaño de paquete y le restamos 1 y obtenemos que se perdieron 3 paquetes. De esta manera, el jitter de arribo y el PLR están relacionados, por lo tanto, teniendo uno es posible calcular el otro.

En base a la relación mencionada anteriormente se analizó el comportamiento del PLR y jitter de arribo en las horas donde se detectaron los niveles más bajos de QoS:

Trazas: (CT1, 60 ms, 13:00hrs), (CT1, 60 ms, 15:00hrs), y (CT1, 60 ms, 19:00hrs)
Traza: (CT1, 40 ms, 15:00hrs)
Traza: (CT1, 20 ms, 15:00hrs)

Las Figuras 7 - 11 y 7 - 12 muestran el vector de pérdida (como se define en la sección 5.1.1) y la traza de jitter de arribo (CT1, 60 ms, 13:00hrs), respectivamente.

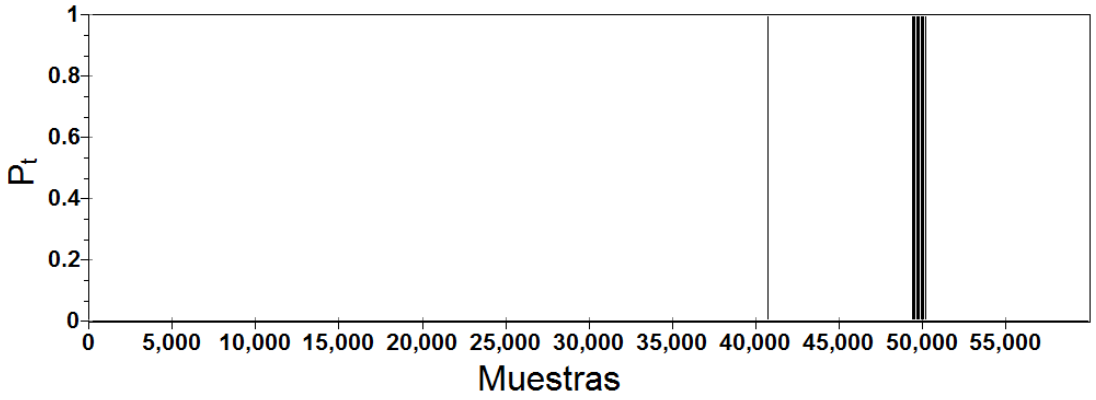


Figura 7 - 11. Vector de Pérdida: (CT1, 60 ms, 13:00hrs)

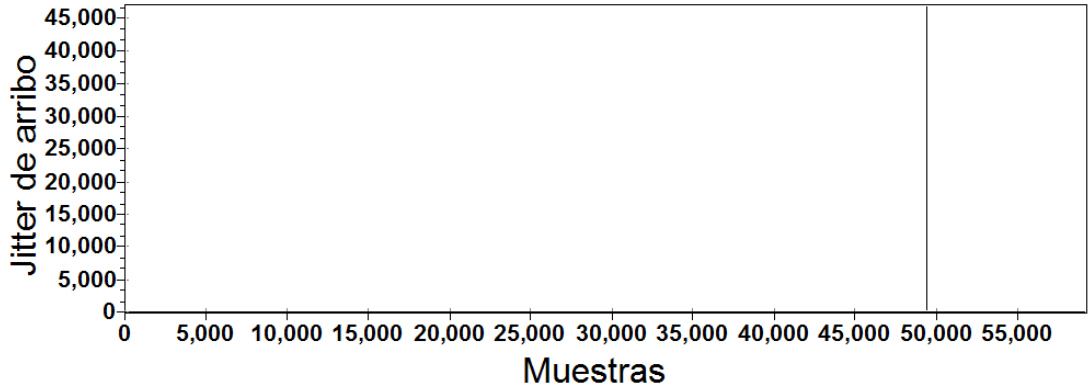


Figura 7 - 12. Jitter de arribo: (CT1, 60 ms, 13:00hrs)

Como se puede ver en la Figura 7-13, se presenta un valor atípico de jitter de arribo en el orden de 47039.92ms, por tanto esto equivale a una ráfaga de 783 paquetes perdidos de manera consecutiva, como se muestra en la Figura 7-11. Mediante este valor de jitter la longitud de la ráfaga, se puede determina que se perdieron 47.22 segundos de la comunicación. Ya que a 60 ms se transmiten 16.6 paquetes por segundo, entonces: (47,039.92 ms / 60 ms) / 16.6 = 47.22 segundos. Las Figuras 7-13 y 7-14 muestran el vector de pérdida y la traza de jitter de arribo (CT1, 60 ms, 15:00hrs), respectivamente.

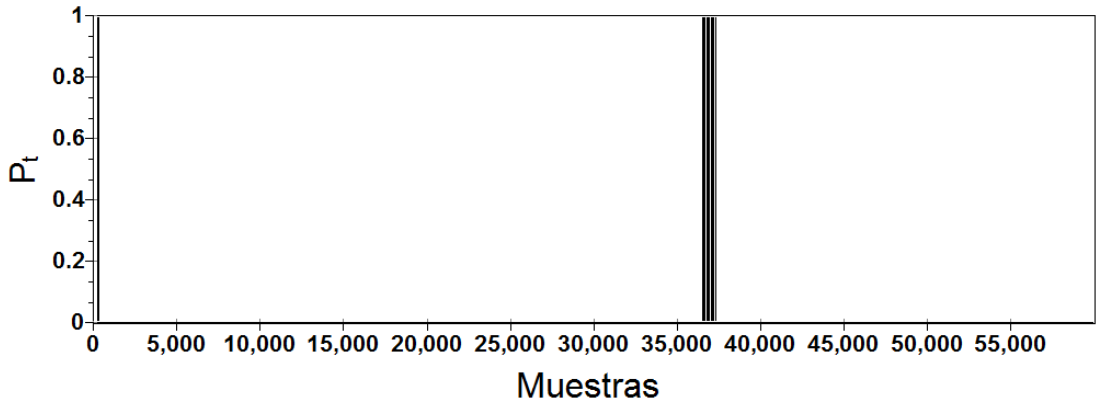


Figura 7 - 13. Vector de Pérdida: (CT1, 60 ms, 15:00hrs)

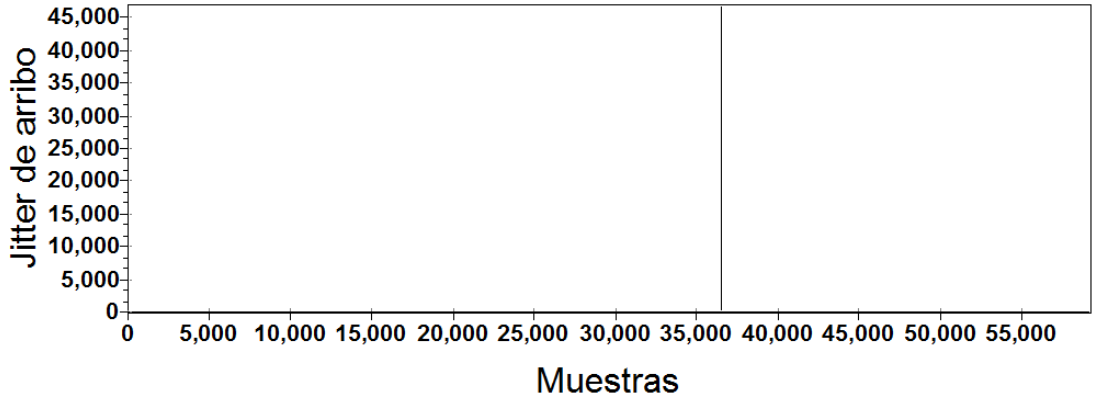


Figura 7 - 14. Jitter de arribo: (CT1, 60 ms, 15:00hrs)

Como se puede apreciar en la Figura 7 - 14, se presenta un valor poco común de jitter de arribo en el orden de 46860.04ms, esto quiere decir que hubo una ráfaga de 780 paquetes perdidos consecutivamente, como se ilustra en la Figura 7 - 13. Por lo tanto, se perdieron 47.04 segundos de voz.

Las Figuras 7 - 15 y 7 - 16 ilustran el vector de pérdida y la traza de jitter de arribo, correspondientes a CT1, de las 19:00hrs pero ahora a 60 ms.

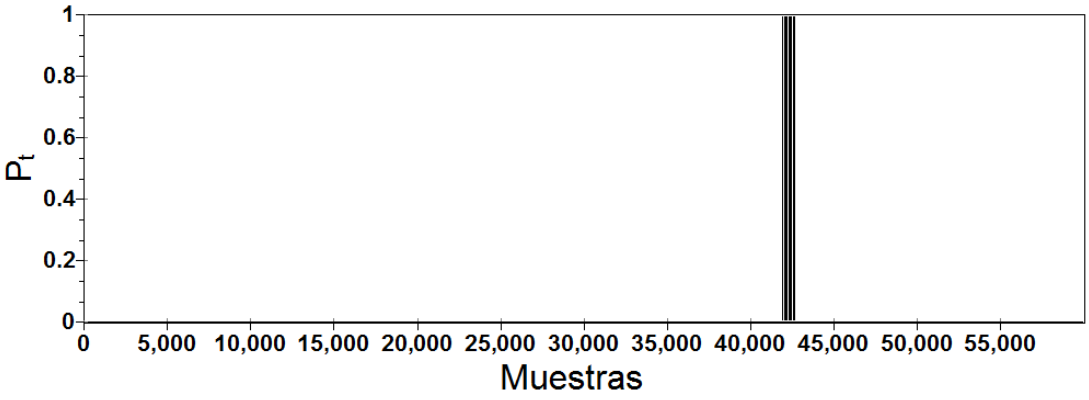


Figura 7 - 15. Vector de Pérdida: (CT1, 60 ms, 19:00hrs)

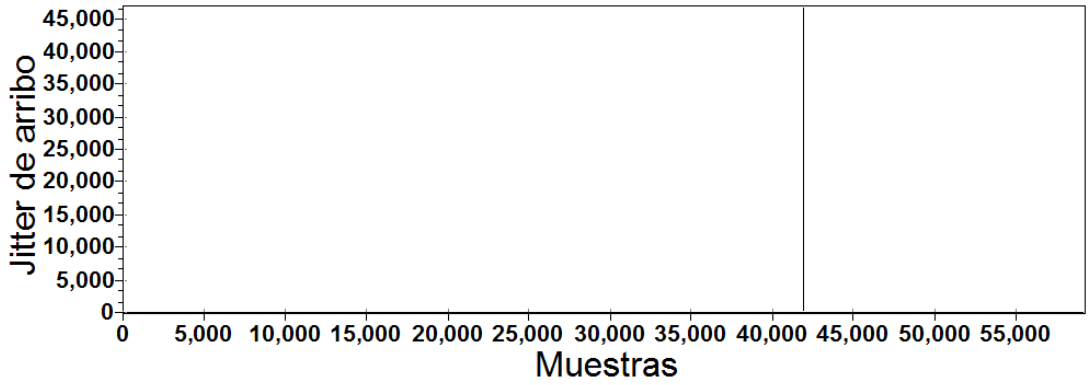


Figura 7 - 16. Jitter de arribo: (CT1, 60 ms, 19:00hrs)

Como se puede ver en la Figura 7-16, se presenta nuevamente un valor de jitter de arribo atípico en el orden de 46920.25ms, lo cual equivale a una ráfaga de 781 paquetes perdidos de manera consecutiva, como se muestra en la Figura 7-15. El valor de jitter y la longitud de ráfaga reportada en esta ocasión determinan que se perdieron 47.10 segundos de comunicación.

Como se puede observar en las Figuras 7-17 y 7-18, se muestran el vector de pérdida y la traza de jitter de arribo, correspondiente al CT1, a las 15:00hrs a una tasa de 40ms.

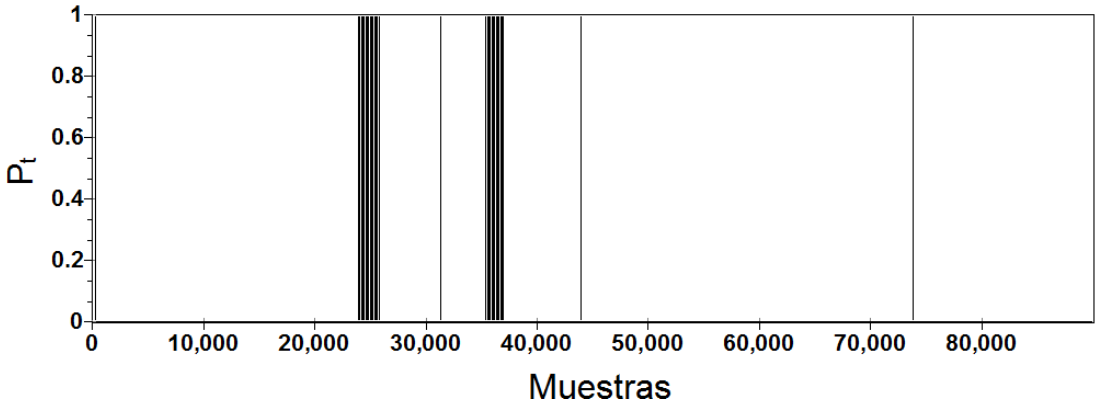


Figura 7 - 17. Vector de Pérdida: (CT1, 40 ms, 15:00hrs)

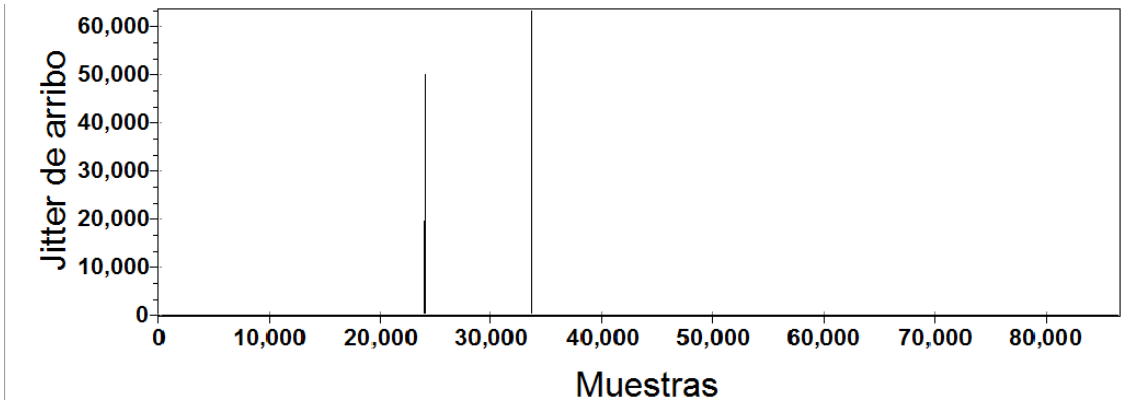


Figura 7 - 18. Jitter de arribo: (CT1, 40 ms, 15:00hrs)

La Figura 7-18 muestra la presencia de dos valores atípicos de jitter de arribo en el orden de 49959.96ms y 63754.92ms, por lo tanto, esto equivale a una ráfaga de 1248 y 1593 paquetes perdidos de manera consecutiva, respectivamente, como se muestra en la Figura 7 - 18. Los valores atípicos de jitter y las longitudes de las ráfagas de pérdida, determinan que se perdieron 49.95 y 63.75 segundos de voz.
En las Figuras 7 - 19 y 7 - 20 se muestra el vector de pérdida y la traza de jitter de arribo del CT1, a una tasa de 20ms, correspondiente a las 15:00hrs.

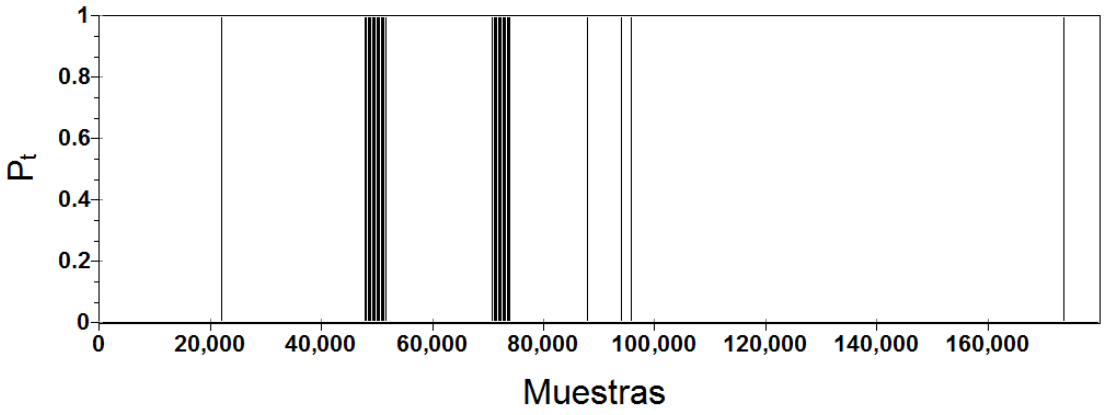


Figura 7 - 19. Vector de Pérdida: (CT1, 20 ms, 15:00hrs)

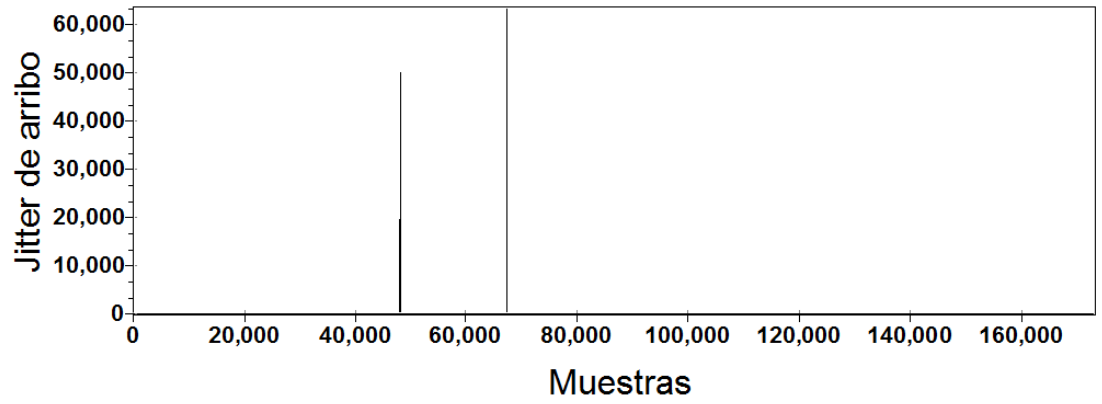


Figura 7 - 20. Jitter de arribo: (CT1, 20 ms, 15:00hrs)

En la Figura 7 - 20, se presentan dos valores atípicos de jitter de arribo en el orden de 49960.24ms y 63699.89ms, por tanto esto equivale a una ráfaga de 2497 y 3184 paquetes perdidos de manera consecutiva, respectivamente, como se muestra en la Figura 7-16. Por consiguiente, se perdieron 49.96 y 63.69 segundos de voz, en cada evento atípico.

Algo muy importante a tomar en cuenta del análisis anterior es que a un tamaño de paquete de 20 ms hay alrededor 113.65 segundos de audio perdidos, es decir, casi dos minutos de conversación. Cualquier persona que estuviera realizando una llamada en ese instante habría experimentado una calidad de servicio

deficiente debido a que no hubiera podido escuchar casi dos minutos de la conversación.
De lo anterior, se puede concluir que la calidad de servicio se está viendo afectada severamente por las pérdidas en ráfagas en los flujos de voz codificados mediante G.711.

Con el objetivo de estudiar la relación entre OWD, PLR y MOS en flujos de voz codificados mediante G.729, se realizó la comparación entre estos parámetros en CT2 (ver Figuras 7 - 21, 7 - 22 y 7 - 23).
La comparación entre el OWD y MOS se muestra en la Figura 7 - 23 y se puede observar que cuando se presentan valores grandes de OWD (mayor a 100ms), existe correlación entre este parámetro y el MOS, es decir, valores grandes de OWD tienen significativo impacto en la calidad de servicio.
La Figura 7 - 22 ilustra la comparación entre PLR y MOS, de la cual podemos ver que el PLR tiene gran influencia en la QoS.

La comparación entre OWD y PLR se presenta en la Figura 7 - 23, de esta figura podemos ver que no existe correlación alguna entre estos dos parámetros.
De este análisis, se puede concluir que para G.729, una menor cantidad de pérdidas de paquetes tiene significativo impacto en la calidad de servicio a comparación de G.711 y cuando existen valores de grandes de OWD (mayor a 100ms), este parámetro impacta negativamente también la QoS.

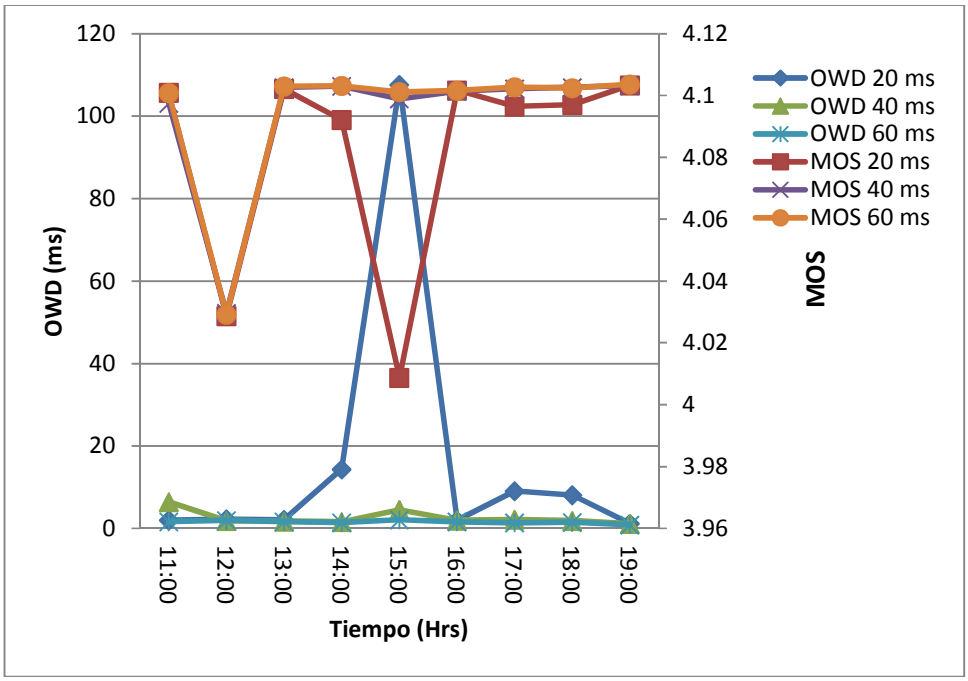


Figura 7 - 21. Comparación OWD - MOS en G.729

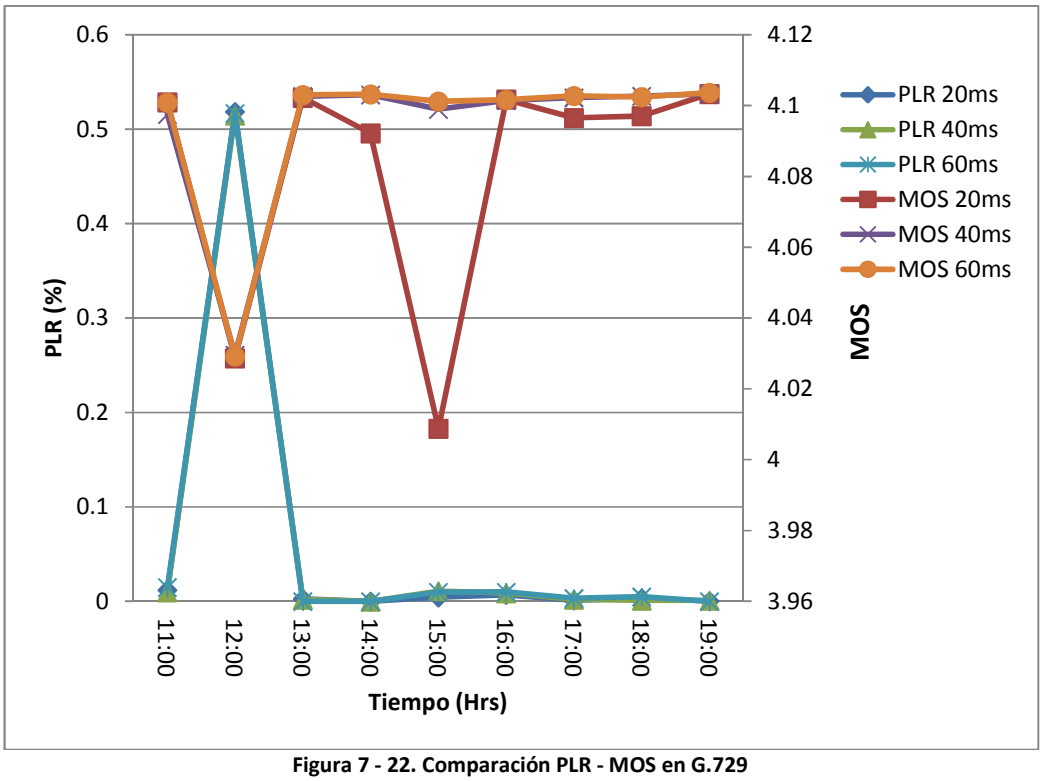


Figura 7 - 22. Comparación PLR - MOS en G.729

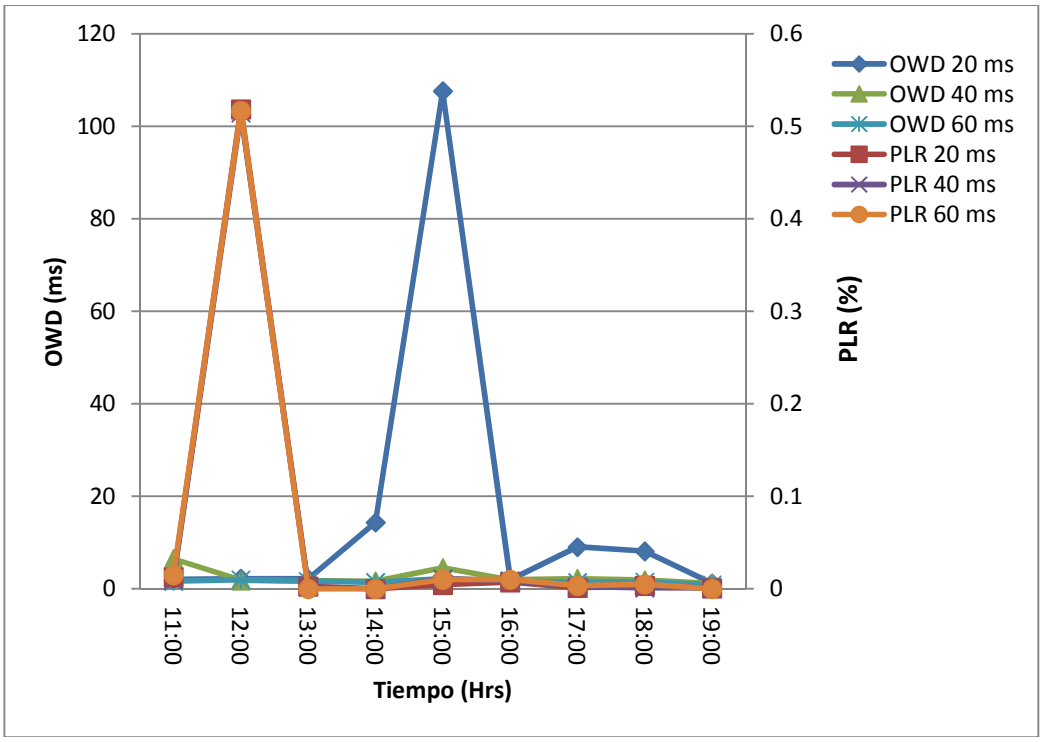


Figura 7 - 23. Comparación OWD - PLR en G.729

En base a la relación mencionada en la ecuación (5.7), se analizó el comportamiento del PLR y jitter de arribo en las horas donde se detectaron los niveles más bajos de QoS:

- Traza: (CT2, 60 ms, 12:00hrs)
- Traza: (CT2, 40 ms, 12:00hrs)
- Traza: (CT2, 20 ms, 12:00hrs)

Las Figuras 7-25 y 7-26 muestran el vector de pérdida y la traza de jitter de arribo (CT2, 60 ms, 12:00hrs), respectivamente.

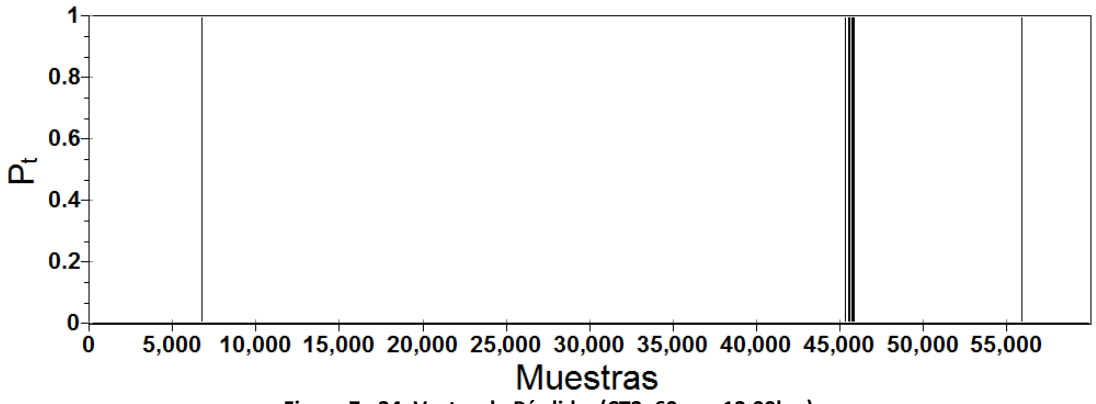


Figura 7 - 24. Vector de Pérdida: (CT2, 60 ms, 12:00hrs)

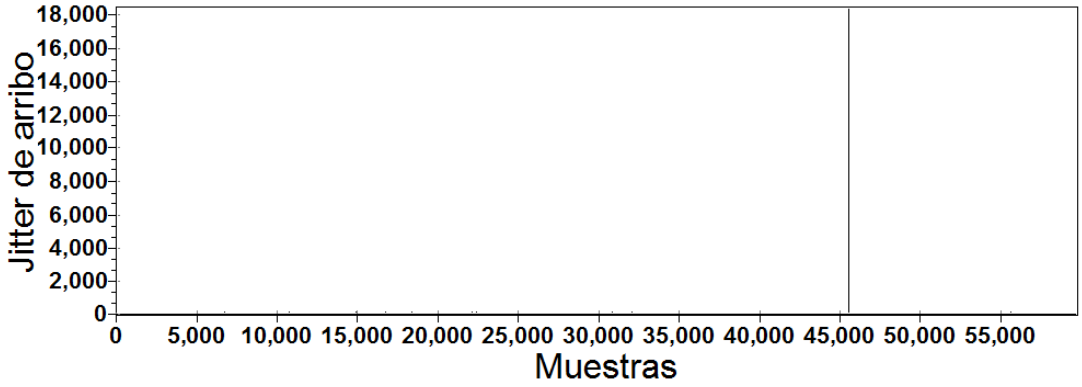


Figura 7 - 25. Jitter de arribo: (CT2, 60 ms, 12:00hrs)

Como se puede ver en la Figura 7 - 25, se presenta un valor atípico de jitter de arribo en el orden de 18480ms, lo cual nos indica que se produjo una ráfaga de 307 paquetes perdidos consecutivamente, como se observa en la Figura 7 - 25. Por tanto, podemos determinar que se perdieron 18.55 segundos de voz.

Las Figuras 7 - 26 y 7 - 27 ilustran el comportamiento del vector de pérdida y la traza de jitter de arribo (CT2, 60 ms, 12:00hrs), respectivamente.

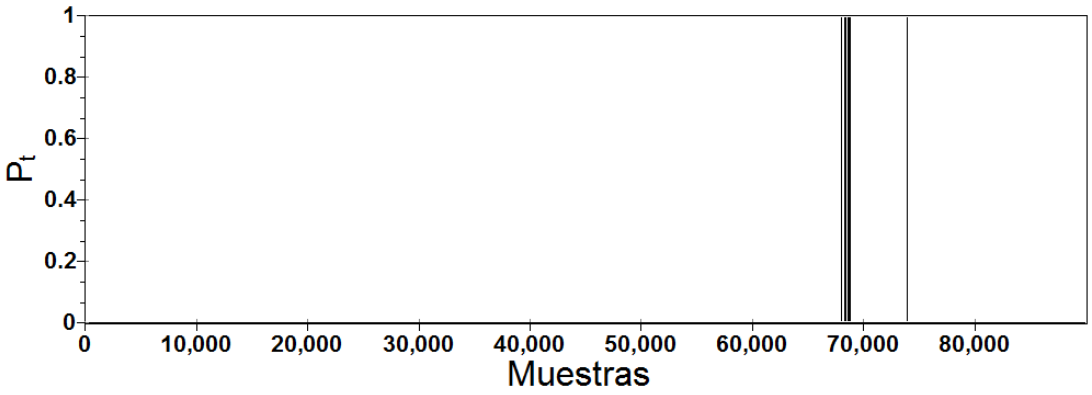


Figura 7 - 26. Vector de Pérdida: (CT2, 40 ms, 12:00hrs)

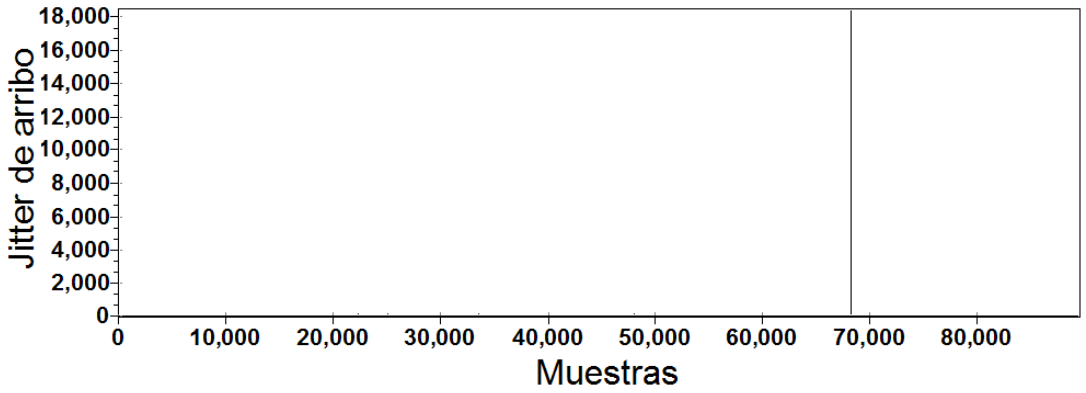


Figura 7 - 27. Jitter de arribo: (CT2, 40 ms, 12:00hrs)

La Figura 7-27 presenta el comportamiento del jitter de arribo y se puede apreciar un valor atípico de éste parámetro en el orden de los 18480.05ms, lo que equivale a un conjunto de 461 paquetes perdidos consecutivamente, como se muestra en la Figura 7 - 26. Este valor de jitter, en conjunto con la longitud de ráfaga, nos ayuda a determinar que se perdieron 18.48 segundos de voz.

Las Figuras 7 - 28 y 7 - 29 muestran el vector de pérdida y la traza de jitter de arribo (CT2, 60 ms, 12:00hrs), respectivamente.

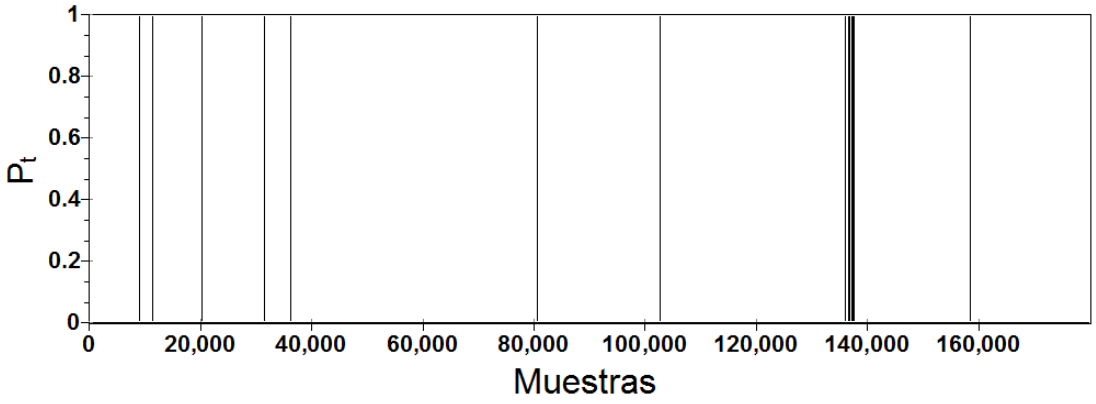


Figura 7 - 28. Vector de Pérdida: (CT2, 20 ms, 12:00hrs)

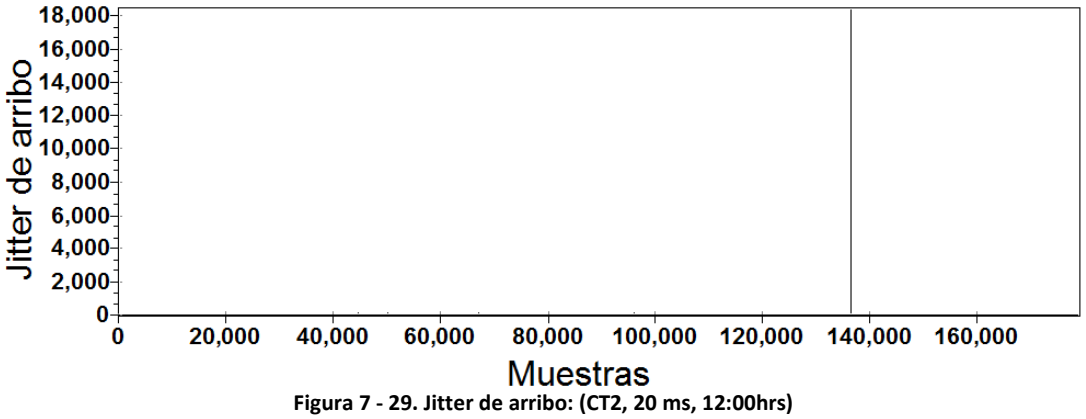


Figura 7 - 29. Jitter de arribo: (CT2, 20 ms, 12:00hrs)

Como se puede ver en la Figura 7-29, se presenta de nuevo un valor atípico de jitter de arribo, esta vez de 18459.54ms, por lo tanto, hubo una ráfaga de 922 paquetes perdidos de manera consecutivos, como se puede observar en la Figura 7-28. Por tanto, se perdieron 18.45 segundos de voz.

De lo anterior, se puede concluir que la calidad de servicio se esta viendo afectada severamente por las pérdidas en rafagas en los flujos de voz codificados mediante G.729.

En base a la conclusión encontrada, respecto al impacto significativo que provocan las pérdidas de paquetes en la calidad de servicio de las llamadas de prueba, se presenta un análisis complementario. Dicho análisis complementario consiste en hacer la comparativa entre el MOS y PLR para los dos esquemas de codificación utilizados (G.711 y G.729) por tamaño de paquete (20ms, 40ms, 60ms), como se muestra en la Figuras 7 - 30, 7 - 31 y 7 - 32.

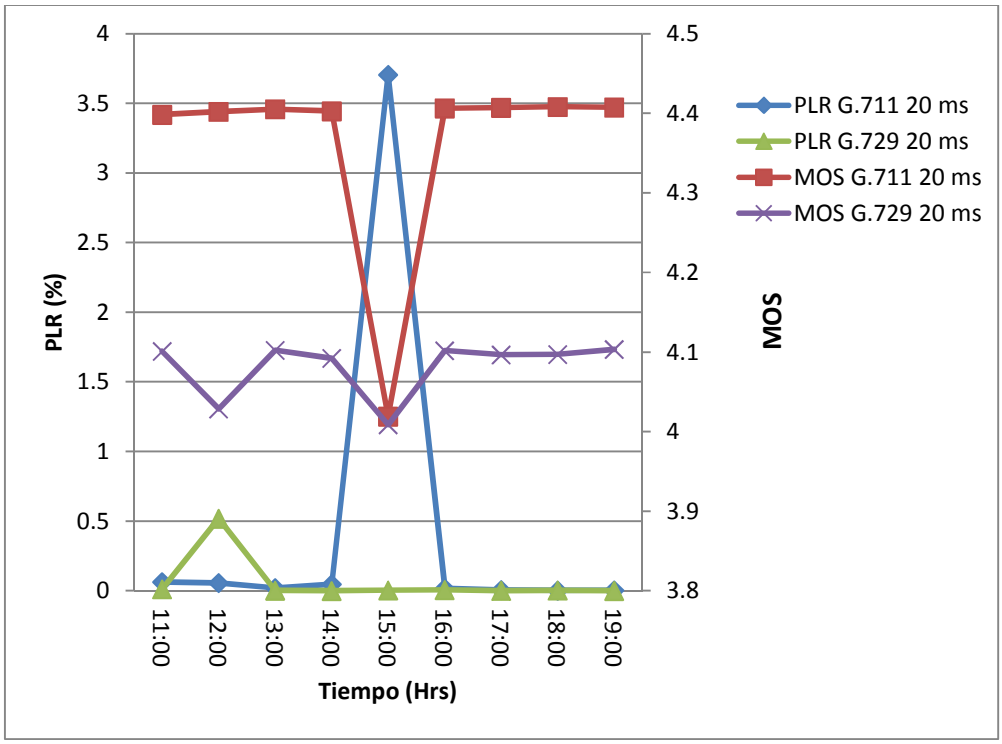


Figura 7 - 300. PLR vs MOS en G.711 y G729: 20 ms

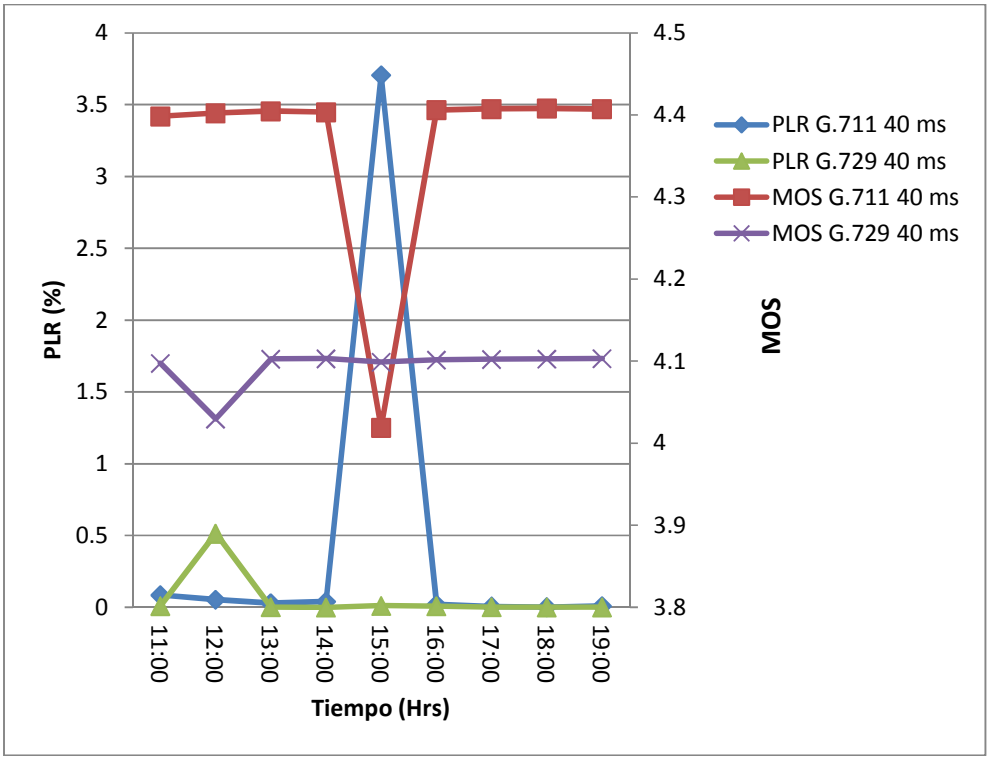


Figura 7 - 311. PLR vs MOS en G.711 y G729: 40 ms

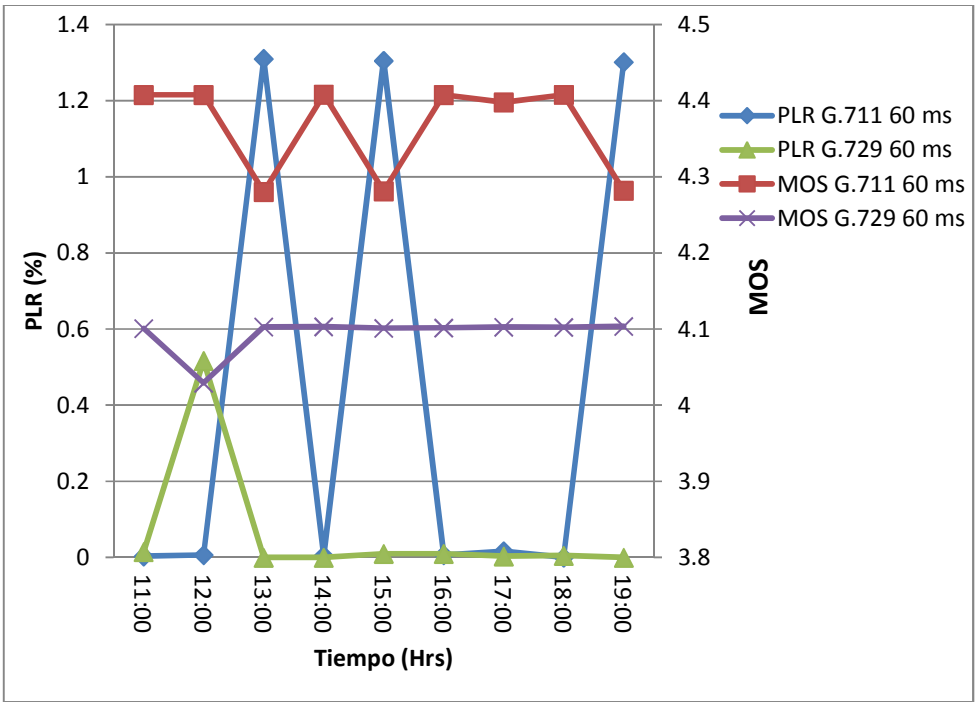


Figura 7 - 32. PLR vs MOS en G.711 y G729: 60 ms

La Tabla 7-1 y las Figuras 7 - 30, 7 - 31 y 7 - 32, resumen el comportamiento del parámetro pérdida de paquetes. Se puede observar que tanto para G.711 como para G.729 con diversos tamaños de paquete, el nivel de pérdida de paquetes es determinante en el nivel de QoS alcanzado en cada llamada de prueba.

Por otro lado, también se puede apreciar que aún cuando el porcentaje de pérdidas de paquetes para G.711 era mayor que para G.729, la calificación de MOS estuvo dentro del mismo rango en ambos CODECs. Esto deja en evidencia que el CODEC G.729, es mucho más sensible al PLR que G.711.

Tabla 7 - 1. Valores críticos de MOS con ambos CODECs

CODEC	Tamaño de paquete (ms)	Valor de MOS obtenido	Nivel más alto de PLR (%)	Nivel de satisfacción
G.711	20	4.019085	3.705185	Algunos insatisfechos
	40	4.019066	3.706873	Algunos insatisfechos
	60	4.281167	1.305174	Satisfechos
G.729	20	4.028704	0.518351	Algunos insatisfechos
	40	4.029599	0.514462	Algunos insatisfechos
	60	4.029184	0.516684	Algunos insatisfechos

Capítulo 8
Conclusiones

La tecnología VoIP utiliza las redes IP para transportar voz, por tanto, el primer paso será, la digitalización de la señal de voz mediante CODECs (G.711, G.729, etc.), seguido de la paquetización de las muestras de voz o adición de los encabezados IP (IP-UDP-RTP) y finalmente el envío de estos sobre la red IP. Si, además, se desea prestar el servicio de telefonía, será necesario ofrecer todas las funciones propias de una red telefónica, tales como la señalización de llamada, entre otras funciones avanzadas.

Dos de los protocolos o estándares más utilizados para la señalización de llamadas de voz sobre redes IP son: H.323 y SIP. H.323 fue ratificado por la ITU-T, y está conformado por un conjunto de protocolos para soportar voz, video y datos sobre una red IP. SIP pertenece a la IETF, y es un protocolo de control de señalización de la capa de aplicación para crear, modificar y terminar sesiones con uno o más participantes.

Ambos protocolos consisten de tres componentes lógicos principales; una terminal, servidor de señalización y un gateway. Las actuales implementaciones de H.323 y SIP no proveen QoS, sin embargo, VoIP es uno de los servicios más sensibles a la QoS y demanda estrictos niveles del mismo. El nivel de calidad de servicio en VoIP depende de varios parámetros, sin embargo, los que tienen mayor impacto son: OWD, jitter y PLR.

Algunos de estos parámetros están estrechamente relacionados, por ejemplo el PLR con el jitter de arribo guardan una relación de acuerdo a la siguiente ecuación:

$$IAT(K, K-n-1) = J^{K}(L) + (n+1)(IDT)$$

donde, $IAT(K, K-n-1)$ es el jitter de arribo entre dos paquetes consecutivos, $J^K(L)$ es el jitter de OWD, IDT es el tiempo de interpartida y n determina la longitud de la ráfaga o el numero de paquetes perdidos de manera consecutiva en una comunicación. Por tanto los valores de jitter de arribo pueden ser usados para determinar las longitudes de las ráfagas que se presenten en una llamada y en consecuencia la cantidad de información perdida en unidad de tiempo (ms).

Derivado de los puntos anteriores, en esta tesis se evaluó el desempeño y se caracterizaron los principales parámetros de QoS (jitter de arribo, OWD y PLR) de un conjunto de llamadas de prueba bajo diferentes configuraciones (tipo de códec: G.711 y G.729 / tamaño de paquete: 20ms, 40ms y 60ms).

Para poder realizar el presente estudio, se implementó un escenario de red VoIP-H.323 representativo, en términos de llamadas larga distancia (ver Figura 6 - 1).

Sobre el escenario implementado se generó tráfico VoIP mediante el establecimiento de un conjunto de llamadas de prueba, bajo las configuraciones mostradas en la Tabla 6 - 3.

Posteriormente se capturaron los patrones de tráfico correspondientes a las llamadas de prueba y se colectó un conjunto de trazas (ver Tabla 6 -3) de la siguiente manera:

- Tres llamadas de prueba fueron establecidas de manera simultáneas entre las terminales PC1/TELE1, PC2/TELE2 y PC3/TELE3, ver Figura 6 - 1.
- Las configuraciones usadas en las llamadas de prueba están basadas en dos parámetros: tipo de códec (G.711 y G.729) y tamaño de paquete de voz (20ms, 40ms y 60ms).
- Los periodos de medición fueron de 60 minutos por cada llamada de prueba (duración de llamada).
- Por cada periodo de medición (una hora), tres trazas de jitter de arribo, OWD y PLR fueron obtenidas.

- Los ocho conjuntos de trazas seleccionados contienen 23.76 millones de paquetes RTP, correspondiente a 216 trazas de jitter de arribo, 216 trazas de OWD y 216 trazas de PLR, medidos en horas típicas de trabajo.
-

Finalmente, como resultado de nuestro estudio sobre el comportamiento de los principales parámetros de QoS, se puede concluir lo siguiente:

- Mientras más pequeño sea el tamaño de paquete usado en la llamada de prueba, menor es la calidad de servicio, es decir, flujos de paquetes de voz de tamaño mayor son menos sensibles a degradaciones en la calidad de la voz.
- Los valores de OWD encontrados no tuvieron impacto significativo en la QoS de las llamadas de prueba.
- Existe una gran correlación entre los valores de MOS y PLR, es decir, valores grandes de PLR (1.3%-3.7%) corresponden valores bajos de MOS (4.28-4.02).
- Un fenómeno que afecta fuertemente a la calidad de servicio en las llamadas de prueba son las pérdidas en ráfagas, es decir pérdidas consecutivas entre paquetes.
- El CODEC G.729, resultó ser mucho más sensible al PLR que G.711, puesto que, aún cuando el porcentaje de pérdidas de paquetes para G.711 era mayor que para G.729, la calificación de MOS estuvo dentro del mismo rango en ambos CODECs.
- La red IP de la Universidad de Quintana Roo está apta para poder soportar llamadas telefónicas mediante los esquemas de codificación G.777 y G.729, sin embargo, es recomendable usar en horas pico, el CODEC G.729 a bajo tamaño de paquete de voz (20ms), para que la calidad de servicio se vea afectada lo menos posible.

Referencias

- [1] Shigeru Kashihara, ***VoIp Technologies***, Janeza Trdine 9, 51000 Rijeka, Croatia, 2011.
- [2] Julio Gómez López, Francisco Gil Montoya, ***VoIP y Asterisk redescubriendo la telefonía***, Editorial RA-MA Madrid, España, 2008.
- [3] Jonathan Davidson, James Peters, ***Voice over IP Fundamentals***, Cisco Press, 201 West 103rd Street, Indianapolis, IN 46290 USA, 2000.
- [4] Bates, Gallon, Bocci, Walker, Taylor, ***Converged Multimedia Networks***, John Wiley & Sons, Ltd, 2006.
- [5] Recomendación I.120, ***Redes Digitales de Servicios Integrados***, 1993.
- [6] John Fox, Karen Loutsch, and Michelle O'Brien, ***ISDN: Linking the Information Highway to the Classroom,*** 1993***.***
- [7] José Manuel Huidobro Moya y David Roldán Martínez, ***Tecnología VoIP y telefonía IP***, Creaciones Copyright, S. L., España, 2006.
- [8] Michael A. Gallo y William M. Hancock, ***Comunicación entre computadoras y tecnologías de redes,*** 2002.
- [9] Scott Keagy, ***Integración de redes de voz y datos***, Cisco Press, 2001
- [10] Allan Sulkin, ***PBX Systems for IP telephony***, Migrating Enterprise Communication, McGraw-Hill TELECOM, 2003.
- [11] HomeroToral Cruz, ***QoS Parameters Modeling of Self-similar VoIP Traffic and an to the E-Model***, Tesis de doctorado, 2010
- [12] J. Willamowius, **"OpenH323 Gatekeeper: The GNU Gatekeeper,"** http://www.gnugk.org/, 2009.
- [13] ITU-T, **"H.323: Packet-based multimedia communication systems"**, Telecommunication Standardization Sector, Geneva, Switzerland, 2009.
- [14] Colin Perkins, **RTP: Audio and video for the Internet**, Publisher Addison Wesley, 2003.

- [15] G. Combs, Wireshark: A Network Protocol Analyzer, http://www.wireshark.org/, 2010.
- [16] ITU-T, **“G.107: The E-Model, a computational model for use in transmission planning,”** Telecommunication Standardization Sector, Geneva, Switzerland, 2009.