Capítulo III
Red VoIP

3.1 Introducción

La industria y su deseo de combinar voz y datos han llevado al desarrollo de nuevos y diversos conceptos y tecnologías, como la voz en paquetes. Los paquetes de voz comprenden varios estándares y protocolos. Las aplicaciones usan estos estándares y protocolos para proveer servicios rentables y de valor agregado para los usuarios.

La voz en paquetes permite a un dispositivo enviar tráfico de voz sobre una red IP/Frame Relay/ATM. En el caso de VoIP, el DSP que se encuentra en los segmentos del Gateway de voz segmenta la señal de voz en tramas. El Gateway de voz combina estas tramas para formar un paquete IP y enviarlo por la red IP. En el punto de destino, sucede la acción contraria que consiste en convertir la información de voz que está en el paquete IP en la señal original de voz.

En una red con capacidad de comunicación VoIP también el Gateway es el encargado de convertir las señales analógicas que entren por sus interfaces en paquetes de voz comprimidos y así poder ser transportados por la red.

Dentro de la comunicación VoIP existen elementos y conceptos importantes e imprescindibles como:

- **Direccionamiento**: Utilizado para identificar origen y destino de las llamadas.
- **Enrutamiento**: Es el proceso mediante el cual se encuentra el mejor camino a seguir por un paquete desde la fuente al destino basado en métricas.
- **Señalización**: Avisa a las terminales y elementos de red de su estado.
- **Terminales**: Existen dos tipos: las terminales de hardware y las terminales de software. Las primeras son los teléfonos, mientras que las terminales de software se ejecutan desde una computadora.
- **Gatekeeper**: Es el elemento encargado de sustituir a la central telefónica. Su función principal es el control de llamadas y gestión del sistema de direccionamiento. Cada terminal antes de realizar una llamada debe consultar al gatekeeper si es posible. Este elemento también es capaz de redireccionar llamadas al Gateway más indicado o un nuevo destino en caso de que el original no esté disponible.
- **Gateway**: Permiten que toda llamada dirigida a la red telefónica pueda establecerse sin intervención del usuario.

Existen diversos requerimientos y recomendaciones relacionadas a la infraestructura de una red LAN y WAN útiles para garantizar en medida de lo posible, una óptima calidad en la implementación de soluciones de VoIP en dichas redes. También es importante crear un buen diseño de red y para ello es primordial conocer todas las advertencias y las entrañas de la tecnología de red que se desplegará y que pretenderá contar con alta disponibilidad. Conociendo el tipo de interfaces y protocolos de señalización soportados por el PBX o los sistemas claves es posible elegir los componentes correctos de hardware y software en la solución VoIP que se quiere implementar. De esta manera es posible que un único estándar permita interoperabilidad de aplicaciones con diferente hardware y software distintos sobre IP, tomando como base éste estándar se presentan los requerimientos de hardware, software y servicios de comunicación necesarios para el diseño.

Por estas razones este capítulo aborda los conceptos más importantes involucrados en el desempeño de una red VoIP.

3.2 Protocolos

En cualquier ámbito de las telecomunicaciones existen protocolos que se refieren al conjunto de reglas estandarizadas que son útiles para que se asegure un intercambio de datos fiable a través de diversos canales de comunicación. Es por estas razones que también se crean protocolos para VoIP, cuyo mecanismo de conexión abarca una serie de transacciones de señalización entre terminales.

Al hablar de VoIP se hace referencia a un conjunto de protocolos que conforman las redes IP, existe una serie de ellos que proporcionan servicios en tiempo real y definen la manera en que por ejemplo los códecs se conectan entre sí y hacia otras redes usando VoIP.

En VoIP existen dos tipos de protocolos:

- Los protocolos que proveen el control de llamada y señalización y
- Protocolos que transportan la carga útil (RTP, RTCP, UDP e IP)

3.2.1 Protocolos de señalización

En VoIP se usa IP para las decisiones de ruteo, UDP para la entrega de paquetes y RTP/RTCP para transportar en tiempo real.

Los protocolos de señalización son los responsables de localizar una terminal, negociar varias funciones, de iniciar y finalizar las llamadas de voz en una red VoIP. Comúnmente existen diferentes protocolos usados en las redes VoIP entre los cuales se encuentran: H.323, MGCP, SCCP y SIP. Dichos protocolos difieren en arquitectura, control de llamada y otros servicios.

3.2.1.1 H.323

H.323 es la especificación de la ITU-T [11] para la transmisión de audio, video y datos a través de una red IP, incluyendo internet. Los productos y aplicaciones deben ser compatibles con H.323 pueden comunicarse y ser interoperables el uno con el otro. La especificación del protocolo H.323 describe como se crea y mantiene una sesión entre dos terminales.

Los componentes de H.323 son los siguientes:

Protocolo	Característica
H.225	Señalización de llamada
RAS	Registro, admisión y estado de funciones
Q.931	Señalización de inicio de llamada
H.235	Protocolo de seguridad
H.245	Capacidad de negociación
H.450	Servicios suplementarios
H.246	Interoperabilidad con redes de circuitos conmutados
H.26x	Códecs de video
G.7xx	Códecs de voz

Tabla 3.1 Componentes de H.323

H.323 está basado en varios protocolos como se ilustra en la figura 1. Estos protocolos son provistos tanto por mecanismos de entrega de paquetes confiable y no confiable sobre la misma red.

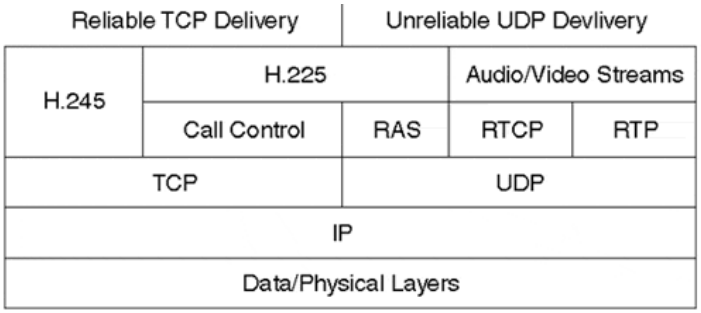

Figura 3.1 Paquete de voz

Los componentes principales para H.323 son: H.225, H.245 y RAS[14].

[14] Las terminales H.323 usan este protocolo para comunicarse con los gatekeepers H.323 para manejar registro/administración/estado.

H.225 (Señalización de control de llamada)

En las redes H.323, los procedimientos de control de llamada están basados en la recomendación H.225 de la UIT, la cual especifica el uso y soporte de los mensajes de señalización Q.931. Un canal confiable de control de llamada es creado a través de la red IP en el puerto TCP 1720. Este puerto es el que inicia los mensajes de control entre dos terminales con el propósito de conectar, mantener y desconectar llamadas.

Los mensajes actuales de control y mensajes "keepalive" se mueven a puertos efímeros después del setup inicial de la llamada. Pero 1720 es el puerto bien conocido para las llamadas H.323. H.225 también especifica el uso de mensajes Q.932 para servicios suplementarios.

Los mensajes Q.931 y Q.932 que son más comúnmente usados en las redes H.323 son los siguientes:

- Inicio – Es un mensaje enviado por la entidad H.323 que llama como intento para establecer una conexión a la entidad H.323 llamada.
- Seguimiento de llamada – Es un mensaje de regreso enviado por la entidad llamada a la entidad que llama para avisar que el seguimiento de llamada ya inició.
- Alerta – Es un mensaje de vuelta desde la entidad llamada avisando que el ring de la parte llamada inició.
- Conectar – Mensaje de vuelta del usuario llamado hacia el que llama indicando que la parte llamada ha contestado.
- Liberación completa – Enviada por la terminal iniciando la desconexión, lo que indica que la llamada ha sido liberada.
- Instalación – Mensaje Q.932 usado para pedir o acusar de recibo servicios suplementarios.

H.245 (Control)

H.245 maneja los mensajes de control de principio a fin entre entidades H.323. Los procedimientos H.245 establecen canales para la transmisión de audio, video, datos e información del canal de control. Una terminal establece un canal H.245 por cada llamada con la terminal participante. Un canal de control confiable es creado en IP usando una asignación dinámica de puerto TCP en el mensaje final de señalización de llamada. El intercambio de capacidades, el abrir y cerrar de los canales lógicos, modos preferentes y mensajes de control se llevan a cabo sobre este canal de control.

3.2.1.2 SCCP (Skinny Call Control Protocol)

Por sus siglas en inglés SCCP, es un protocolo del cual Cisco es propietario y está basado en una arquitectura cliente-servidor. Los clientes pueden tratarse de cualquier teléfono Cisco o un softphone IP Cisco. El servidor se trata del CUCM[15].

El CUCM maneja el control de inicio de llamada y el teléfono es el responsable del procesamiento de los paquetes RTP/RTCP. Los mensajes SCCP se transportan por el puerto TCP 2000. La ventaja de este protocolo es que como usa TCP como protocolo de capa cuatro, los mensajes pueden aprovechar la funcionalidad de corrección de errores y garantizar la entrega de paquetes.

Se envían mensajes constantes entre el teléfono cliente y el CUCM para cualquier cosa que el usuario haga en el teléfono. Es importante tener en cuenta que este modelo de cliente-servidor entre la terminal y el CUCM es sólo para señalización; los paquetes de voz encapsulados en RTP y RTCP son transportados directamente de una terminal a otra. Para mayor información ver la referencia [12].

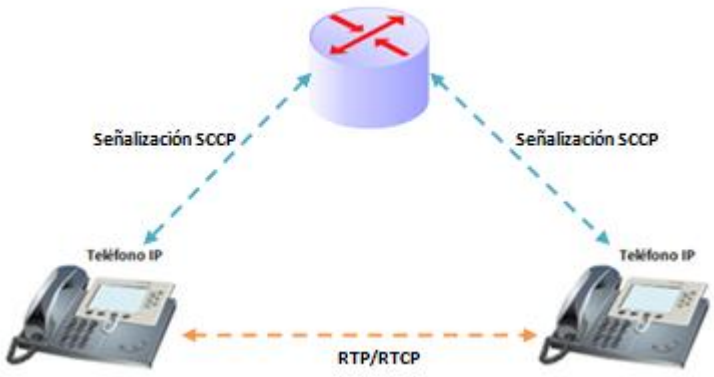

Figura 3.2 Operación de SCCP

[15] Cisco Unified Communications Manager

3.2.1.3 MGCP (Media Gateway Control Protocol)

MGCP controla VoIP a través de elementos de control de llamada externos. Este protocolo está basado en una arquitectura cliente-servidor lo que significa que la inteligencia recae en el CUCM facilitando la función de ruteo de voz. Es un estándar de la IETF[16] y es uno de los más recientes y más sencillos de configurar.

MGCP puede ser usado tanto en TCP como UDP y la información correrá en los puertos 2428 y 2427 por default respectivamente. Para un estudio más detallado analizar la referencia [12].

3.2.1.4 SIP (Session Initiation Protocol)

En el año de 1996 se presentó ante la IETF un prototipo de SIP conocido como SIPv1 pero no fue hasta 2002 que se publicó la RFC 3261 [8] en la cual se introducían todas las características y modificaciones realizadas por el grupo SIP creado en 1999.

Al igual que los otros protocolos SIP es usado para iniciar, mantener y finalizar las sesiones multimedia incluyendo telefonía por internet, conferencias y otras aplicaciones similares las cuales involucran datos, voz y video.

SIP soporta tanto sesiones multicast como unicast así como también llamadas punto a punto o multipunto. Para establecer y terminar dichas llamadas se transita por estas cinco facetas SIP:

- Localización de usuario
- Capacidad de usuario
- Disponibilidad de usuario
- Inicio de llamada
- Manejo de llamada

Los componentes principales en un sistema SIP son los agentes de usuario y los servidores de red. Las partes que llaman y las llamadas se identifican por medio de direcciones SIP ya que las partes necesitan localizarse entre sí.

[16] Internet Engineering Task Force: Tiene como objetivos el contribuir a la ingeniería de Internet, actuando en diversas áreas, como transporte, encaminamiento, seguridad.

Agente de usuario

El agente de usuario es una aplicación que contiene el UAC[17] y el UAS[18] o también llamados cliente y servidor respectivamente. El cliente manda las solicitudes SIP y actúa como agente de llamadas de usuario, mientras que el servidor recibe las solicitudes y regresa una respuesta en nombre del usuario, actúa como el agente de usuario llamado.

Servidores de red

Existen dos tipos de servidores SIP: los servidores proxy y servidores de redirección.

- **Servidores Proxy** – Actúa como otros clientes y contiene funciones tanto de cliente como servidor. Un servidor de este tipo es capaz de interpretar y reescribir los encabezados de solicitud antes de enviarlos a otros servidores.
- **Servidores de redirección** – Acepta las solicitudes SIP y envía una respuesta redirigida hacia el cliente con la dirección del siguiente servidor. Estos servidores no aceptan llamadas ni procesan ni envían solicitudes SIP.

Direcciones SIP

Las direcciones SIP son también llamadas URLs[19] y existen de la forma: *usuarios@dominio* similar al e-mail. La porción del usuario puede ser un nombre o un número telefónico, y la porción de dominio puede ser un nombre de dominio o dirección de red.

Localización de un servidor y de un usuario

Para la localización de un servidor un cliente puede enviar una solicitud SIP de dos maneras, ya sea directamente o por medio de la IP y el puerto correspondiente del URL SIP. La primera manera de hacerlo es sencilla pues la aplicación conoce el servidor proxy, mientras que la segunda manera mencionada es más complicada pues presenta las siguientes complicaciones:

- El cliente debe determinar la dirección IP y el número de puerto del servidor para el cual la solicitud es destinada.

[17] User-agent client
[18] User-agent server
[19] Universal Resource Locators

- Si el número de puerto no está enlistado en el URL SIP solicitado, que por default es el 5060.
- Si el número de puerto no está enlistado en el URL SIP solicitado, el cliente primero deberá intentar conectarse usando UDP y después TCP.
- El cliente consulta el servidor DNS para la IP de dominio, por lo que si no encuentra registros, el cliente es incapaz de localizar el servidor y continuar con su solicitud.

La localización de un usuario depende mucho del tipo de servidor que se utilice pues el usuario podría moverse hacia diferentes sistemas finales. La localización de estos sistemas debe estar registrada en el servidor SIP o en otros servidores que no son SIP.

Cuando se usa un servidor SIP proxy, este puede intentar direcciones en paralelo hasta que la llamada sea satisfactoria, sin embargo, al usar un servidor SIP redirigido se regresará el listado completo de locaciones y se habilitará el usuario directamente.

Mensajes SIP

Existen dos tipos de mensajes SIP, las solicitudes iniciadas por los clientes y las respuestas enviadas por los servidores. Cada mensaje contiene un encabezado el cual especifica los detalles de la comunicación. Los mensajes SIP se envían sobre TCP o UDP.

Los encabezados de los mensajes SIP especifican la parte que llama, la parte llamada, ruta y tipo de mensaje de la llamada. Existen cuatro grupos de encabezados:

- Encabezados generales – Aplicados para solicitudes y respuestas
- Encabezados de entidad – Dan información acerca del tipo de mensaje y longitud
- Encabezados de solicitud – Permite incluir al cliente información adicional de solicitud
- Encabezados de respuesta – Permite al servidor incluir información adicional de respuesta

La siguiente tabla muestra algunos encabezados:

Generales	Entidad	Solicitud	Respuesta
Accept	Content-Encoding	Authorization	Allow
Accept-Encoding	Content-Length	Contact	Proxy-Authenticate
Accept-Language	Content-Type	Hide	Retry-After
Call-ID		Max-Forwards	Server

Date		Proxy-Authorization	WWW-Authenticate
Encryption		Proxy-Require	
Expires		Route	
From		Require	
Record-Route		Response-Key	
Timestamp		Subject	
To		User-Agent	
Via			

Tabla 3.2 Grupos de encabezados principales

Mensajes de Solicitud

Existen seis solicitudes SIP:

- INVITE – Indica que el usuario o servicio está invitado a participar en una sesión.
- ACK – Representa la confirmación final para concluir la transacción iniciada con INVITE.
- OPTIONS – Permite preguntar y recolectar capacidades de agentes de usuario y servidores
- BYE – Usado por las dos partes para liberar una llamada.
- CANCEL – Sirve para cancelar cualquier solicitud en progreso
- REGISTER – Registra la locación de clientes con los servidores SIP.

Mensajes de respuesta

Son los mensajes enviados en respuesta a una solicitud e indican el éxito o fallo de la llamada, incluyendo el estado del servidor.

Clase de respuesta	Código de estado	Explanation
Informational	100	Trying
	180	Ringing
	181	Call is being forwarded
	182	Queued
Success	200	OK
	300	Multiple choices

	301	Moved permanently
	302	Moved temporarily
	303	See other
	305	Use proxy
	380	Alternative service
Client-Error	400	Bad request
	401	Unauthorized
	402	Payment required
	403	Forbidden
	404	Not found
	405	Method not allowed
	406	Not acceptable
	407	Proxy authentication required
Client-Error	408	Request timeout
	409	Conflict
	410	Gone
	411	Length required
	413	Request entity too large
	414	Requested URL too large
	415	Unsupported media type
	420	Bad extension
	480	Temporarily not available
	481	Call leg or transaction doesn't exist
	482	Loop detected
	483	Too many hops
	484	Address incomplete
	485	Ambiguous
	486	Busy here
Server-Error	500	Internal server error
	501	Not implemented

	502	Bad gateway
	503	Service unavailable
	504	Gateway timeout
	505	SIP version not supported
Global Failure	600	Busy everywhere
	603	Decline
	604	Does not exist anywhere
	606	Not acceptable

Tabla 3.3 Respuestas SIP

Operación básica de SIP

Los servidores SIP manejan las solicitudes de dos maneras y la operación de estas se basa en invitar a un participante a la llamada. Los dos modos de operación del servidor SIP son: los modos de servidor proxy y el de servidor de redirección.

Los pasos para llevar a cabo una llamada de dos vías en el modo proxy son los siguientes:

- El servidor proxy acepta la solicitud INVITE del cliente.
- El servidor proxy identifica la localización usando las direcciones suministradas y los servicios de locación.
- Una solicitud INVITE es emitida hacia la locación obtenida.
- El agente de usuario de la parte llamada alarma al usuario y regresa una indicación de éxito al servidor proxy involucrado.
- Una respuesta de OK se envía del servidor proxy a la parte que llama.
- La parte que llama confirma mediante una petición ACK, la cual se envía por el servidor proxy hacia la parte llamada.

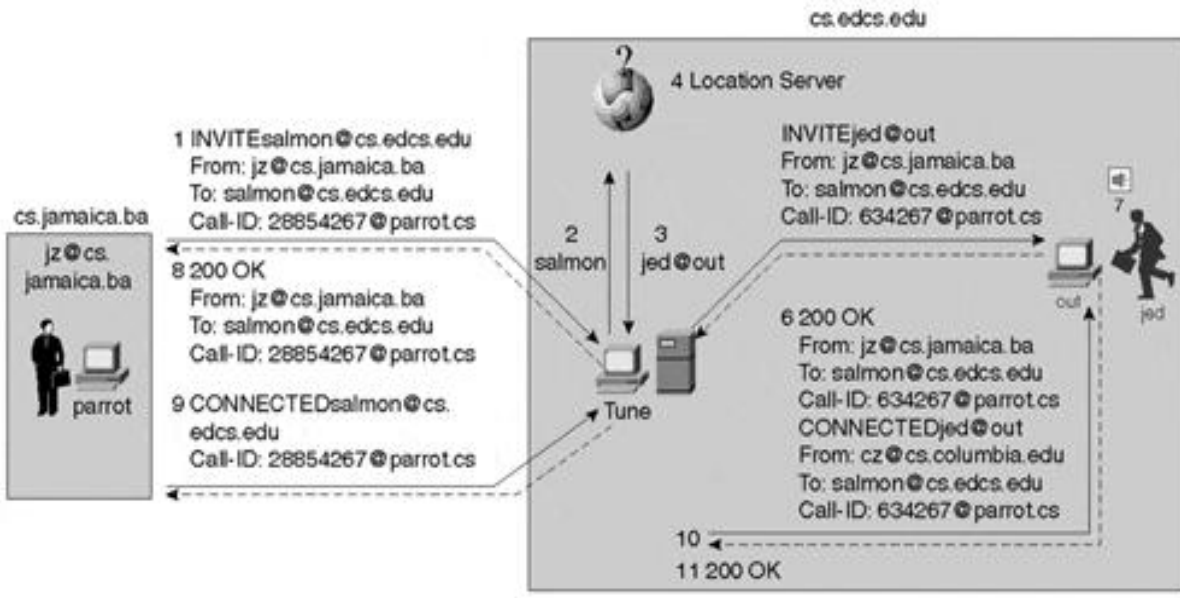

Figura 3.3 Ejemplo de operación SIP (modo proxy) [4]

Los pasos para llevar a cabo una llamada en el modo de redirección son los siguientes:

- El servidor acepta la petición INVITE de la parte que llama y contacta los servicios de locación con la información suministrada.
- Luego de localizar al usuario, el servidor regresa la dirección directamente a la parte que llama.
- El agente de usuario envía un ACK al servidor como comprobante de una transacción completada.
- El agente de usuario envía una petición INVITE directamente a la dirección regresada por el servidor.
- La parte llamada da una indicación de éxito OK y la parte que llama regresa un ACK.

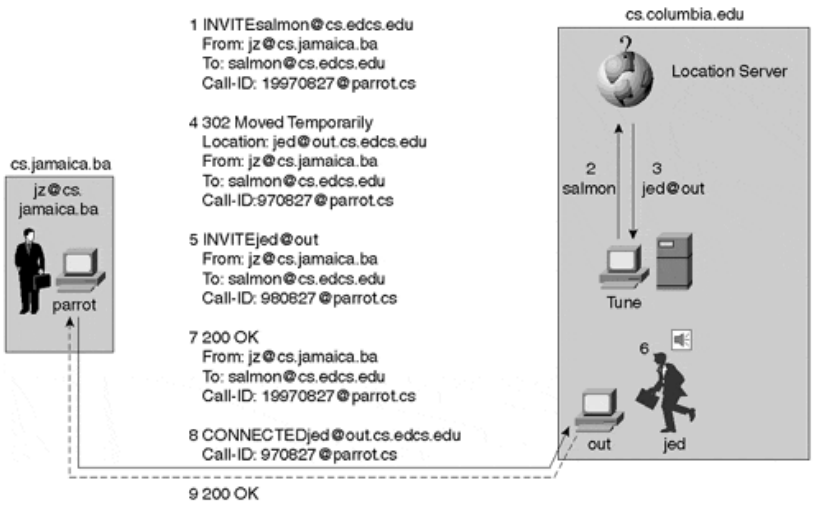

Figura 3.4. Ejemplo de operación SIP (modo redireccionado)[4]

En esta tesis se usa un dispositivo que funge como servidor SIP trabajando en modo proxy y se lleva a cabo un proceso similar al de la figura 3.3.

3.2.1.5 Comparación de protocolos de señalización de voz

Es importante poder comparar y diferenciar los beneficios y características de los protocolos de señalización abordados anteriormente. La siguiente tabla muestra una comparación de dichos protocolos:

Protocolo	Estándar	Arquitectura	Control de llamada	Usos del CUCM
SCCP	Propiedad de Cisco	Cliente-servidor	Centralizado	Gw de voz/trunk y punto final a CUCM
MGCP	IETF	Cliente-servidor	Centralizado	Gw de voz/trunk
H.323	ITU	P2P[20]	Distribuido	Gw de voz/trunk
SIP	IETF	P2P	Distribuido	Gw de voz/trunk y punto final a CUCM

Tabla 3.4 Comparación de protocolos de señalización de voz [3]

[20] Peer to peer: Se refiere a una red en la que todos o algunos aspectos funcionan sin clientes ni servidores fijos, actúan simultáneamente como clientes y servidores respecto a los demás modos de la red. Las redes P2P permiten el intercambio directo de información, en cualquier formato, entre los ordenadores interconectados.

3.2.2 Protocolos de Transporte

Como es bien sabido sobre IP recaen dos tipos de tráfico: los de UDP y los de TCP. Se sabe que al usar TCP se tendrá una conexión confiable en comparación con UDP.

Debido a que el tráfico de voz es muy sensible a los retrasos de tiempo, la solución más lógica es usar UDP/IP para transportar la voz. La IETF adoptó RTP para tiempo real o sensibilidad al retardo. VoIP viaja en la parte superior de RTP, el cual viaja en la parte superior de UDP. Por lo tanto VoIP es transportado con un encabezado de paquete RTP/UDP/IP como se muestra en la figura 3.5.

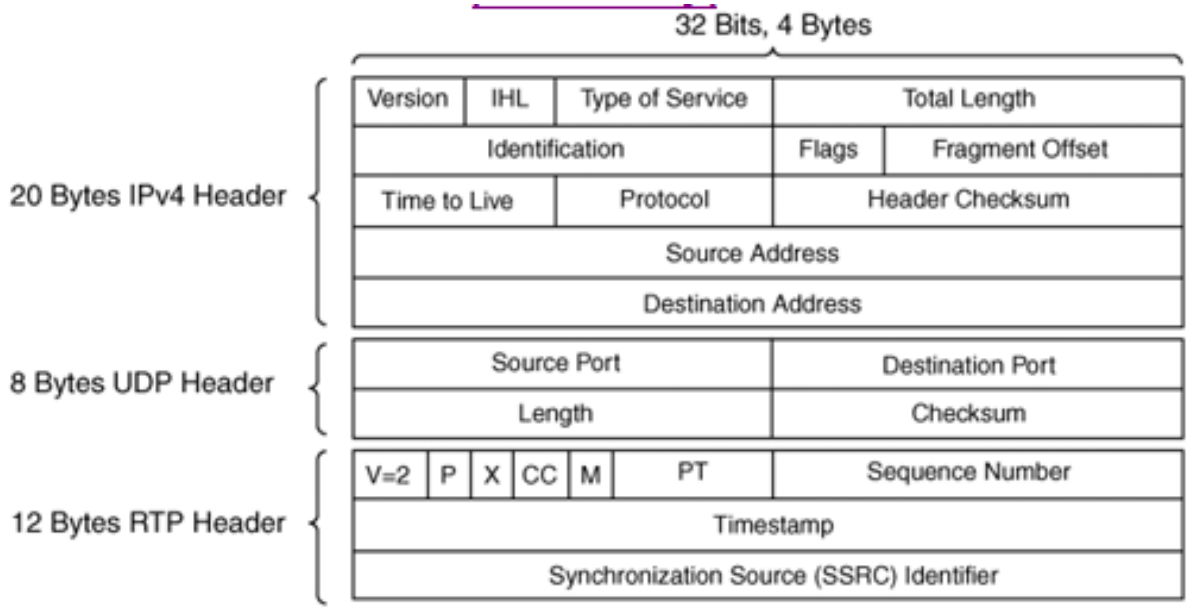

Figura 3.5 Encabezados de protocolos RTP, UDP e IP

3.2.2.1 RTP (Real-Time Transport Protocol)

RTP es un estándar de la IETF RFC 1889 [9] y 3050 [13] para la entrega unicast y multicast de voz y video. El protocolo de transporte que usa RTP es casi siempre UDP pero es un servicio no confiable basado en el mejor esfuerzo y aunque puede llegar a sonar como algo perjudicial en realidad es el mejor método para transportar este tipo de datos.

UDP al ser un servicio basado en el mejor esfuerzo no intenta retransmitir ni reordenar paquetes como lo haría TCP. La explicación de por qué UDP es la mejor opción para transporte es simple: si tratáramos de retransmitir un paquete de voz perdido, al hacerlo y

que el paquete alcanzara su destino, el sonido contenido no tendría sentido pues estaría siendo entregado fuera de tiempo.

RTP por medio de su encabezado (ver figura 3.5) proporciona un campo llamado "timestap" el cual se pone en cada paquete de voz digitalizada y ayuda a corregir el problema de retardo de llegada.

3.2.2.2 cRTP (Compress RTP)

cRTP es una opción que surgió para mitigar un poco el problema que aún se tenía al utilizar RTP, debido a que la voz es muy sensible al retardo. cRTP toma los 40 bytes del conjunto de encabezados y los corta entre dos y cinco bytes.

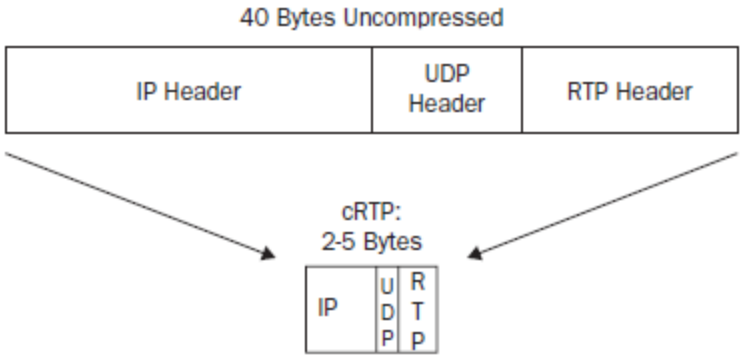

Figura 3.6 De RTP a cRTP

Lo que hace cRTP es que una vez que la información es conocida en los dos extremos del cable y ya que mucha de la información contenida en los encabezados UDP/IP/RTP es estática, cRTP quita esa información y al no enviar esta información se conserva ancho de banda. Este protocolo es más eficiente en enlaces WAN con velocidades T1 y menores, enlaces con mayores velocidades no obtienen beneficio alguno. Para un entendimiento a profundidad revisar la referencia [14].

3.2.2.3 RTCP (Real-Time Control Protocol)

RTCP trabaja directamente con RTP para proveer un monitoreo de la transmisión de los datos RTP encapsulados. Los paquetes RTCP se envían a los participantes de una determinada transmisión RTP. La función principal de RTCP es proporcionar un respaldo acerca de la calidad de las transmisiones RTP. La aplicación de tiempo real puede usar esta información para adaptar las características de la codificación si el protocolo detecta

congestión y si se encuentra congestionado, el receptor puede informar al receptor que use otro códec de menor calidad y por lo tanto colaborar para evitar cuellos de botella.

Alguna de la información que RTCP rastrea de RTP es la siguiente:

- Cuenta total de paquetes de la transmisión
- Paquetes perdidos
- Retraso
- Cantidad de Jitter

RTP usa puertos pares de UDP mientras que RTCP usa el siguiente número impar más alto.

3.2.3 Protocolos de enrutamiento

Un router es aquel dispositivo que conecta múltiples redes, es decir, tiene varias interfaces y cada una de ellas pertenece a una red IP diferente. Cuando un router recibe un paquete IP en una interfaz, determina la interfaz que va usar para enviarlo a su destino, usando su tabla de enrutamiento para determinar la mejor ruta.

Generalmente cada red a la que se conecta un router requiere una interfaz separada. En dichas interfaces se pueden conectar tanto LAN como WAN. Es probable que un router reciba un paquete encapsulado en un tipo de trama de enlace de datos, como una trama Ethernet, y al enviar el paquete, el router lo encapsula en otro tipo de trama, como PPP. La encapsulación de enlace de datos dependerá del tipo de interfaz del router y del tipo de medio al que se conecta. Las tecnologías de enlace de datos a las que se conecta un router pueden ser LAN, como Ethernet y conexiones WAN como una conexión T1 que usa PPP, Frame relay y ATM.

Los routers usan protocolos de rutas estáticas y de enrutamiento dinámico para aprender sobre redes remotas y construir sus tablas de enrutamiento.

3.2.3.1 Enrutamiento estático

Las rutas estáticas se utilizan generalmente cuando se enruta desde una red a una red de conexión única, es decir, una red a la que se accede por una sola ruta. Si en una red de conexión única se ejecuta un protocolo de enrutamiento se considera un desperdicio de recursos pues sólo existe una manera de enviar tráfico que no sea local.

En este tipo de enrutamiento las redes remotas se agregan a la tabla de enrutamiento configurando las rutas estáticas. Una ruta estática debe incluir la dirección de red, la máscara de subred de la red remota, junto a la dirección IP del router del siguiente salto o la interfaz de salida. En la tabla de enrutamiento se indican con la letra S. Agregando rutas estáticas se hace posible la existencia de tablas de enrutamiento más pequeñas y resulta en un proceso de búsqueda en dicha tabla más eficiente pues existen menos rutas para buscar.

3.2.3.2 Protocolos de enrutamiento dinámico

Estos protocolos se utilizan para facilitar y mejorar el intercambio de información de enrutamiento entre los routers con la selección de las mejores rutas. Estos protocolos permiten a los routers compartir información en forma dinámica sobre redes remotas y así agregar automáticamente en sus propias tablas de enrutamiento.

El método que se usa un protocolo de enrutamiento para lograr su propósito depende de las características de cada protocolo, pero en general las operaciones que lleva a cabo son:

1. El router envía y recibe mensajes de enrutamiento en sus interfaces.
2. El router comparte mensajes e información de enrutamiento con otros routers que usan el mismo protocolo.
3. Los routers intercambian información de enrutamiento y aprenden rutas.
4. Cuando un roter detecta algún cambio, el protocolo de enrutamiento puede anunciar el cambio a otros routers.

Dichos protocolos de enrutamiento se clasifican en protocolos de Gateway interior (IGP) y exterior (EGP). Los primeros usados en redes bajo control de una única organización mientras que los últimos son usados en redes controladas por diferentes administraciones como por ejemplo el internet.

Protocolos de enrutamiento por vector de distancia

Los protocolos de vector de distancia es una de las clasificaciones que pueden tener los protocolos de Gateway interior. Este tipo de protocolos se refiere a que las rutas son publicadas como vectores de distancia y dirección, es decir, se define la distancia en términos de una métrica como el conteo de saltos y la dirección es el siguiente router o la

interfaz de salida. A continuación se mencionan algunos protocolos y sus métricas de funcionamiento:

- **RIP** (Routing Information Protocol): Utiliza conteo de saltos, si el conteo de saltos en una red es mayor a quince, no se podrá suministrar ruta para dicha red. Envía actualizaciones cada treinta segundos.
- **IGRP** (Interior Gateway Routing Protocol): Es desarrollado por Cisco. Considera el ancho de banda, el retardo, la carga y la confiabilidad. Actualmente se considera obsoleto.
- **EIGRP** (Enhanced IGRP): Utiliza DUAL[21] para calcular la ruta más corta. No existen actualizaciones periódicas, sólo si existe un cambio de topología

También existen los protocolos de estado de enlace. Un router configurado con un protocolo de enrutamiento de estado de enlace crea una vista completa o topología de la red al reunir información proveniente de los demás routers. Estos protocolos son conocidos por usar el algoritmo SPF[22] y son OSPF[23] y IS-IS[24].

Toda esta información se encuentra más detallada en la referencia [15].

3.3 Códecs

VoIP funciona digitalizando la voz en paquetes de datos, enviándola a través de la red, se realiza la reconversión de digital a analógica en la terminal destino de la comunicación. La señal análoga del teléfono es digitalizada en señales PCM[25] por medio del codificador/decodificador de voz.

Las muestras PCM pasan por el algoritmo de compresión, el cuál comprime la voz y la fracciona en paquetes que pueden ser transmitidos en la red WAN. Al otro extremo del canal de comunicación se realiza el proceso inverso.

Los códecs son usados dentro del mundo VoIP para codificar y decodificar los datos de voz. Estos códecs nos pueden ayudar a usar menor número de bits por conversación de

[21] Algoritmo por difusión dual:
[22] Short Path First
[23] Open Short Path First
[24] Intermediate System-to-Intermediate System
[25] PCM: Pulse code modulation

voz, por lo que se traduce en mayor número de llamadas simultáneamente en un ancho de banda finito. La compresión tiene como objetivo eliminar la redundancia de los datos que son enviados. Usualmente entre más comprimida sea la señal de voz más recursos usará el DSP, por lo que estos códecs se clasifican por su complejidad.

3.3.1 ITU G.711

Este estándar [17] también se conoce como PCM. Este códec muestrea la señal de voz a una frecuencia de 8 000 muestras por segundo. Esto proporciona una mejor calidad a comparación de la mayoría de los códecs empleados.

Existen dos técnicas comunes de compresión binaria G.711 en la mayoría de los servicios de voz: una es llamada la ley μ que es la más usada en los Estados Unidos, Canadá y Japón; mientras que también existe la ley A que es mayormente usada en el resto del mundo. Para lograr interoperabilidad entre estas técnicas PCM debe haber una traducción de un códec a otro.

Una llamada telefónica requiere 64 Kbps en el cable. De acuerdo al teorema de muestreo de Nyquist tendremos 8 000 muestras de voz cada segundo. Cada muestra es de 8 bits; por lo que al multiplicar 8 000 x 8, obtendremos 64 Kbps, lo que significa que G.711 no usa compresión y es la alternativa cuando existe suficiente ancho de banda.

3.3.2 ITU G.729

El muestreo que provee este códec es el mismo que el de G.711. La diferencia de G.711 radica en la compresión, pues G.729 usa una técnica llamada CS-ACELP[26] la cual se basa en métodos alternos de muestreo y expresiones algebraicas como libro de códigos para predecir la representación numérica real. Estas expresiones algebraicas se envían al sitio remoto, donde estas son decodificadas y el audio es sintetizado para imitar el audio original; la predicción y sintetización de forma de onda de audio degrada la calidad de la señal de voz haciendo que la voz del que habla suene robótica.

La ventaja de este códec es que permite una compresión de voz que sólo requiere de 8 Kbps por llamada en vez de los 64 Kbps requeridos por el G.711. Esto significa que se

[26] Conjugative-structure algebraic-code-excided linear prediction.

podrían hacer ocho llamadas en el espacio de una que estuviera usando G.711, lo que sería bueno para compensar el despliegue de VoIP en un enlace WAN de poca rapidez.

Información más detallada se puede encontrar en [18].

3.3.3 ITU G.729a

Es un códec muy parecido al G.729 pues usan el mismo ancho de banda de 8Kbps por llamada pero difieren en el tipo de algoritmo usado, por lo que el códec G.729 es considerado de complejidad alta mientras que el G.729a es considerado de complejidad media.

3.3.4 ITU G.728

Este estándar describe el códec G.728, el cual opera a 16 Kbps. El algoritmo que usa es llamado LD-CELP[27] el cual calcula su predicción mediante un filtro codificador lineal predictivo de orden cincuenta, la excitación es generada por medio de un vector de cuantización.

3.3.5 Otros códecs

El tratamiento de la voz, incluyendo la codificación, decodificación y compresión es un tema que debe analizarse profundamente a la hora de implementar VoIP en una red de determinadas características. Por esta razón existen diversos estándares de la ITU que explican los algoritmos y esquemas de codificación utilizados en cada uno de ellos, unos más complejos que otros. Algunos otros estándares que son de importantes conocer, son los siguientes:

- **G.726** – Describe la codificación mediante el algoritmo ADPCM[28] codificando a 40, 32, 24 y 16 Kbps.
- **G.722 –** Usa la tecnología SB-ADPCM[29] y permite operar a 48, 56 y 64 Kbps.

También existe el iLBC[30] que usa ya sea 20 ó 30 ms de muestras de voz y terminan por consumir 15.2 ó 13.3 Kbps respectivamente. Uno de los beneficios de este códec es que

[27] Low-delay code excited linear prediction.
[28] Adaptative Differential Pulse Code Modulation
[29] Sub-band Adaptative Differential Pulse Code Modulation

tiene la capacidad de manejar la pérdida de paquetes, pues las técnicas usadas por este códec permiten que la pérdida de paquetes sea prácticamente no perceptible para el usuario. Este estándar no es definido por la ITU, sino que fue propuesto por una colaboración de líderes del mundo VoIP y espera ser aceptado universalmente.

Con la gran diversidad de códecs existentes, se debe tomar en cuenta las ventajas y desventajas que cada uno representaría para nuestra red. A continuación se presenta una tabla comparativa de algunos de los códecs:

Códec	Bit Rate [Kbps]	Método	Retraso del algoritmo [ms]	Calidad
G.711	64	PCM (μ ó A)	0.125	4
G.722/6/7	16 – 40	ADPCM	0.125	2.4 – 4
G.728	16	LD-CELP	0.625	3.61
G.729a	8	CS-ACELP	10	3.7
G.729	8	CS-ACELP	15	3.9
G.723.1	6.3	MP-MLQ	30	3.9
G.723.1	5.3	ACELP	30	3.65

Tabla 3.5 Comparación de algunos códecs.

3.4 Problemas de diseño

Para crear un diseño apropiado de red, es importante conocer todas las debilidades y comportamiento interno de las tecnologías relacionadas con el funcionamiento de la red. Los problemas más comunes e importantes a enfrentar en una red VoIP deben de ser tratados uno por uno y con sumo cuidado para obtener el desempeño deseado.

[30] Internet Low Bandwidth Codec

3.4.1 Retraso/Latencia

El retraso o latencia en VoIP es caracterizado por la cantidad de tiempo que toma al diálogo salir de la boca del hablante hasta que alcanza el oído del escucha.

Los tres tipos de retardo que son sustanciales en las redes telefónicas hoy en día son:

- **Retardo de propagación** – Es debido al medio de transporte de la red (fibra óptica, cobre, etc.)
- **Retardo de serialización** – Es la cantidad de tiempo que lleva colocar un bit o byte en una interfaz.
- **Manejo del retardo** – Define diferentes causas de retraso como la paquetización actual, compresión y conmutación de paquetes y son causadas por dispositivos que envían las tramas a través de la red.

La ITU-T en su recomendación G.114 [19] especifica que para una buena calidad de voz, no debe existir un retardo mayor a 150 ms de un punto a otro, por lo cual es un parámetro que se debe tener en consideración y no debemos rebasar.

También existe el retardo PDD[31] el cual consiste en la cantidad de tiempo que pasa entre marcar un número y que el teléfono al que se está llamando suene. Este retraso aumenta cuando hay retrasos en la señalización o pérdida de paquetes.

3.4.2 Jitter

El jitter hace referencia a la variación de tiempo entre llegada de paquetes y es un factor que tiene un impacto significativo en la calidad de voz. Es un parámetro que sólo se presenta en redes de conmutación de paquetes. En un ambiente de paquetes de voz se espera transmitir confiablemente paquetes a intervalos regulares, pero estos paquetes pueden no llegar a estos intervalos regulares de tiempo en la terminal receptora. La diferencia entre el tiempo que es esperado el paquete y el tiempo real en que es recibido es el jitter.

[31] Post Dial Delay

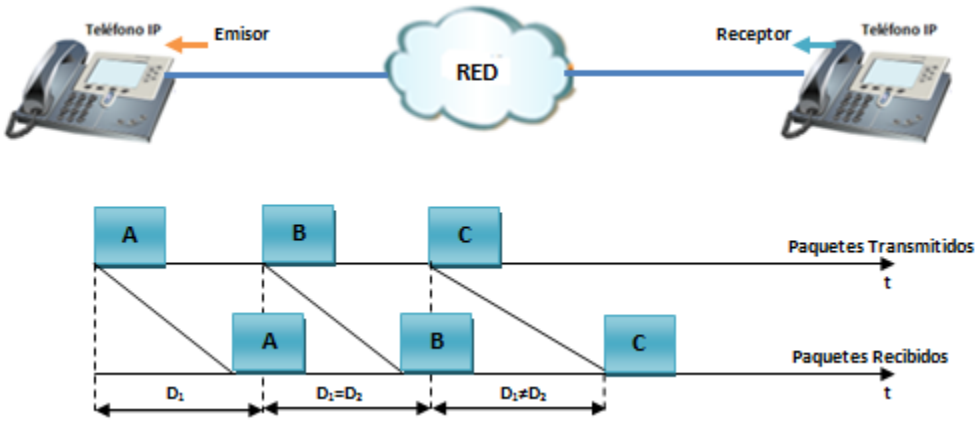

Figura 3.7 Jitter

Para mitigar el problema del jitter existen los buffers de jitter pero estos contribuyen directamente en el retraso total de la red. En los dispositivos de interconexión cisco su IOS[32] permite por medio de las estampas de tiempo de RTP determinar el nivel de jitter en la red si es que existe.

3.4.3 Eco

El eco es el efecto de escuchar tu propia voz mientras hablas o aún después de cierto tiempo de haber producido un sonido. Este efecto causa en la mayoría de las veces interrupciones y rompe la cadencia en una conversación.

En la telefonía tradicional el eco es causado por un desajuste de impedancia en la conversión de los cuatro a los dos cables del bucle local, problema que es resuelto con canceladores de eco.

En las redes basadas en paquetes de hoy en día los canceladores de eco son funciones llevadas a cabo por códecs de bajo bit rate, los cuales son operados por un DSP.

3.4.4 Pérdida de paquetes

La calidad de VoIP puede ser impactada dramáticamente por la pérdida de paquetes. El diseño para una red VoIP debería no perder ni un solo paquete de voz, incluyendo los de señalización y paquetes RTP, pero en la realidad esto no es posible.

[32] Internetwork Operating System

La pérdida de paquetes es causada por una calidad pobre de la red, como muy altos BERs en diversos enlaces o congestión en la red.

La pérdida de paquetes sucede por ejemplo cuando los buffers ya sean de un switch o de un router que tienen conectados cuatro teléfonos y que están llamando simultáneamente, llegan a sufrir un desbordamiento debido a que a la salida se toma un paquete que no concuerda con la secuencia de paquetes y se manda uno incorrecto.

En redes VoIP es importante poder transportar la voz en un tiempo y de manera confiable, así como también es importante contar con mecanismos para hacer que la voz de alguna manera sea resistente a la pérdida de paquetes.

3.5 QoS (Calidad de Servicio)

Cuando se habla de calidad de servicio ¿quién determina que es bueno y qué es malo? En el mercado de las telecomunicaciones y específicamente hablando de voz, por medio de una encuesta subjetiva se puede obtener una idea de la calidad de llamada:

Puntaje	Escala de opinión	Esfuerzo para escuchar
5	Excelente	Relajación posible, sin esfuerzo
4	Buena	Atención necesaria; ningún esfuerzo apreciable
3	Justa	Esfuerzo moderado
2	Pobre	Esfuerzo considerable
1	Mala	No se entiende con un esfuerzo razonable

Tabla 3.6 Puntaje subjetivo de calidad de llamada

Ya sabemos los problemas que se tiene en la implementación de una red VoIP, QoS puede ayudar a resolver estos problemas. La calidad de servicio (QoS) se refiere a la habilidad de identificar tráfico sensible al tiempo y darle prioridad por encima de otro tipo de tráfico. Desafortunadamente QoS no puede resolver todos los problemas y en específico el retraso de propagación, el retraso que introducen los códecs, el retraso del muestreo ni tampoco el retraso de la digitalización.

En una red siempre existen limitaciones en cuanto al ancho de banda y latencia. La primera limitación puede ocasionar cuellos de botella. Un cuello de botella se refiere a enlaces de red que interconectan dos nodos donde la cantidad de tráfico enviada por una interfaz excede la capacidad de la misma. La segunda limitación está relacionada con el retardo y jitter que se puede presentar en la red.

Al hablar de retardo se pueden presentar el caso del retardo fijo, el cuál prácticamente no altera la red y es el que está presente en todas las redes sin excepción; también se puede presentar un retardo variable, el cual es el que se busca eliminar a través de la implementación de QoS. El retardo variable se presenta cuando en un cuello de botella el tráfico sensible al tiempo tiene que esperar en una cola de paquetes y esperar a que los que están adelante sean enviados. Implementando QoS podemos dar prioridad a la voz por encima de cualquier otro tipo de tráfico que no es sensible al tiempo y los paquetes.

Otra ventaja de implementar QoS es que cuando en un cuello de botella la cola o fila de paquetes se empieza a llenar y llega la pérdida de paquetes, los paquetes perdidos serán los de datos menos importantes. Esto se logra usando clasificación de QoS.

Los requerimientos más importantes para poder implementar QoS y que la red no experimente ningún tipo de problema son los siguientes:

- Retraso de inicio a fin menor o igual a 150 ms (ITU G.114)
- Jitter igual o menor a 30 ms
- 1 % o menos pérdida de paquetes

3.5.1 Mecanismos de QoS

3.5.1.1 Clasificación de tráfico

La clasificación de tráfico es el proceso de identificar los paquetes que son sensibles al tiempo, tarea que se debe realizar para que el equipo sea capaz de identificar claramente cierto tipo de tráfico. En este caso el crear VLANS de voz hace más fácil identificar el tráfico de voz ya que se puede asumir que cualquier paquete proveniente de una VLAN de voz debe ser clasificado como tal.

3.5.1.2 Marcado de paquetes

Este proceso consiste de marcar paquetes críticos para que el resto de la red pueda identificarlos y darles prioridad sobre el demás tráfico. En este concepto también se introduce la definición de CoS (Class of service) refiriéndose a un campo en una trama Ethernet el cual es marcado con un número entre cero a siete, entre mayor sea el valor de CoS mayor será la prioridad que se dará a esta información. La voz está marcada por default con un valor de cinco. Los datos que no están marcados con CoS tienen un valor de cero. La CoS es usada por los switches para que se ordenen los datos en fila de forma apropiada.

Para dispositivos de capa tres el marcado de los paquetes se hace con un identificador llamado ToS (Type of service).

3.5.2 Enfilamiento de datos

El enfilamiento del tráfico es ordenar cierto tipo de tráfico para ser transportado a través de interfaces WAN o LAN. Existen muchas técnicas para enfilar el tráfico, pero la que es considerada óptima para el tráfico de voz es la de LLQ (Low latency queuing) pues ayuda a eliminar el retado variable, jitter y pérdida de paquetes que se presentan en la red.

En un switch LLQ crea una estricta prioridad para enfilar el tráfico de voz.

3.5.3 Límites de confianza de QoS

En una red los procesos de clasificación y de marcado deben iniciarse lo más cerca de un punto final, pero dependiendo de la red y de la confiabilidad de los equipos ese límite puede ser modificado. A este tipo de criterio se le llama límite de confianza.

Si se tiene completo control de los puntos finales, entonces se tiene control sobre el CoS y ToS generados y el límite de confianza puede llegar al teléfono IP e incluso a las PC, pero si no se tiene tanto control sobre la red se podría empezar a marcar los valores de CoS y ToS desde el Switch y así sucesivamente.

Figura 3.8 Límites de confianza