5. Voz sobre IP (VoIP) y Telefonía sobre IP (ToIP)

River Quispe Tacas[1] y Germán Suárez Gómez[2]

5.1. Fundamentos de los servicios de VoIP y ToIP

Voz sobre IP (VoIP, *Voice over IP*) es un grupo de recursos que hacen posible que la señal de voz viaje a través de redes TCP/IP. El tráfico de VoIP puede circular por cualquier red TCP/IP, incluyendo aquellas conectadas a Internet. Esto significa que se envía la señal de voz (digitalizada) en paquetes, en lugar de enviarla (en forma digital o analógica) a través de circuitos utilizables sólo para telefonía como en la RTPC/PSTN (Red Telefónica Pública Conmutada/*Public Switched Telephone Network*).

Telefonía sobre IP (ToIP, *Telephony over IP*) es el conjunto de nuevas funcionalidades de telefonía que se pueden ofrecer gracias al envío de la voz sobre el protocolo IP en redes de datos TCP/IP.

La voz ha de digitalizarse para ser transmitida por la red IP. Para ello se hace uso de códecs que realizan la codificación y compresión del audio antes de su transmisión, y luego su decodificación y descompresión en recepción, para entregar una señal audible. Según el códec empleado en la transmisión, se utilizará más o menos ancho de banda y recursos del sistema de cómputo. La cantidad de ancho de banda utilizado suele ser directamente proporcional a la calidad de los datos transmitidos. Entre los códecs más comunes se encuentran los siguientes:

- G.711: Estándar de la UIT-T para la digitalización de audio en telefonía fija. Representa las señales de audio mediante muestras codificadas en una señal digital con tasa de muestreo de 8.000 muestras por segundo con un flujo de datos de 64 kbps. Existen dos tipos:
    - Ley μ: Usado sobre todo en Norte América y Japón. Se basa en un algoritmo de compresión logarítmico de 16 segmentos para representar cada muestra en palabras de 8 bits.

[1]Pontificia Universidad Católica del Perú (PUCP), Perú
[2]Vodafone, España

  - Ley A: Usado en Europa y en el resto del mundo. Se basa en un algoritmo de compresión logarítmico de 14 segmentos para representar cada muestra en palabras de 8 bits.

- G.723.1: Estándar de la UIT-T que comprime la voz en tramas de 30 ms y opera a 5,3 y 6,3 kbps.
- G.726: Estándar de la UIT-T basado en ADPCM (*Adaptative Differential Pulse Code Modulation*). Permite trabajar con velocidades de 16, 24, 32 y 40 kbps. Este códec proporciona una disminución considerable del ancho de banda sin aumentar en gran medida la carga computacional.
- G.729: Estándar de la UIT-T usado sobre todo en aplicaciones de VoIP por los bajos requerimientos en ancho de banda. Opera con tasas de 8 kbps pero existen extensiones para tasas de 6,4 y 11,8 kbps para peor o mejor calidad de voz respectivamente.
- GSM (*Global System for Mobile Communications*): Familia de códecs para telefonía móvil estandarizados por el ETSI. En VoIP se ha venido usando el GSM FR (*Full Rate*), estandarizado como GSM 06.10, que tiene una implementación libre y opera a 13 kbps con una carga de CPU aceptable. En telefonía móvil se están imponiendo versiones mejoradas como el GSM AMR (*Adaptive Multi-Rate*), que ofrece 8 tasas de operación entre 4,75 y 12,2 kbps.
- iLBC (*Internet Low Bit rate Codec*): Es un códec libre que implementa un algoritmo complejo desarrollado por Global IP Sound (GIPS), el cual ofrece una buena relación ancho de banda/calidad de voz a cambio de una mayor carga computacional. Opera a 13,3 y 15,2 kbps.
- Speex: Es un códec libre para voz[3] que implementa un algoritmo capaz de variar la velocidad de transmisión dependiendo de las condiciones actuales de la red (VBR: *Variable Bit Rate*). El ancho de banda puede variar desde 2,15 a 22,4 kbps.

En la actualidad no es posible garantizar la calidad de servicio de VoIP sobre Internet porque se presentan diversos problemas de retardos; pero en redes LAN sí es posible controlar en cierto grado estos problemas. La máxima latencia (tiempo transcurrido desde el instante en que se genera un paquete hasta que se recibe) aceptable en VoIP es de 300 ms ida y vuelta (150 ms en cada dirección). Para lograr una mejor calidad de servicio se emplean los siguientes criterios:

- Supresión de silencios, que otorga más eficiencia a la hora de realizar una transmisión de voz, ya que se aprovecha mejor el ancho de banda al transmitir menos información.
- Compresión de cabeceras aplicando los estándares RTP/RTCP (*Real-time Transport Protocol/Real-time Transport Control Protocol*).
- Priorización de los paquetes que requieran menor latencia.

[3] `http://www.speex.org/`

5.2. Protocolos de VoIP y ToIP

Los protocolos usados para llevar las señales de voz sobre la red IP son comúnmente llamados protocolos de voz sobre IP. El objetivo de VoIP es dividir en paquetes los flujos de audio para transportarlos sobre redes basadas en IP. Los protocolos de las redes IP no fueron diseñados originalmente para el transporte en tiempo real de audio o cualquier otro tipo de flujo de audio/video, por lo que se han creado diversos protocolos para VoIP (Figura 5.1) cuyo mecanismo de conexión incluye una serie de transacciones de señalización entre terminales, que establecen flujos de audio para cada dirección de la conversación. En los siguientes apartados se describen los más utilizados.

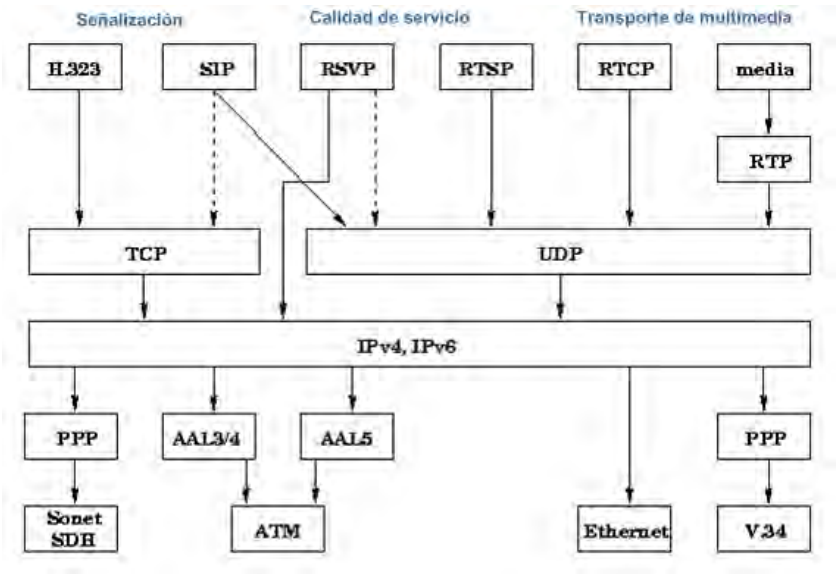


Figura 5.1.: Protocolos VoIP.

5.2.1. SIP (*Session Initiation Protocol*)

Es un protocolo desarrollado por el IETF (*Internet Engineering Task Force*) como el estándar RFC 3261, para la iniciación, moderación y finalización de sesiones multimedia entre dos pares (*unicast*) o multipares (*multicast*). SIP ofrece flexibilidad para controlar sesiones multimedia, como llamadas de voz y video, videoconferencia, mensajería instantánea, juegos en línea y telefonía IP. Una sesión puede ser una simple llamada telefónica de doble vía o una conferencia multimedia con muchos participantes.

Es un protocolo de señalización orientado a conexiones terminal a terminal (*end-to-end*). Esto quiere decir que toda la lógica se encuentra almacenada en los dispositivos terminales (salvo el enrutamiento de mensajes SIP). La ventaja es la estabilidad que se obtiene porque los servidores no son saturados con mensajes SIP, y la desventaja es que los encabezados son mucho mayores.

Es un protocolo de la capa de aplicaciones de la familia TCP/IP; está relacionado estrechamente con el protocolo SDP (*Session Description Protocol*) y coexiste junto con otros protocolos del mismo nivel y funciones, como el H.323 (apartado 5.2.2). Está basado en una arquitectura cliente-servidor similar a HTTP y SMTP; esta similitud es natural ya que SIP fue diseñado para incorporar la telefonía como un servicio más de Internet.

SIP no es un protocolo de propósito general; su objetivo es ayudar a establecer y finalizar la comunicación. Se apoya en otros protocolos para lograr una llamada telefónica, o una sesión de video-conferencia o de mensajería instantánea, etc. Los protocolos que comúnmente colaboran con SIP son: RTSP (*Real-Time Streaming Protocol*) para el control de flujos y sesión, SDP para describir los flujos, RTP/RTCP para el transporte de datos en tiempo real, y RSVP (*Resource Reservation Protocol*) junto a DiffServ (*Differentiated Services*) para gestionar la calidad de servicio y la reserva de recursos.

En las redes TCP/IP, las conversaciones que utilizan señalización del tipo SIP hacen uso de RTP para llevar las conversaciones (flujos de audio/video) de un terminal a otro (Figura 5.2). De la misma forma que en una conversación existen dos flujos de voz, en una conversación en una red TCP/IP se tiene dos flujos de paquetes RTP.

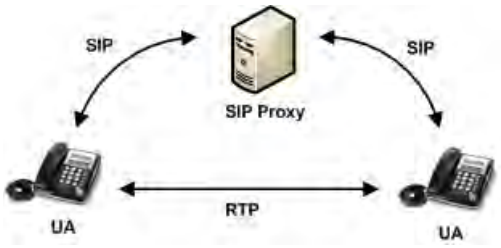


Figura 5.2.: La señalización SIP y las conversaciones de voz (RTP) viajan por caminos distintos.

El principal problema que afecta el funcionamiento de RTP son los NAT (*Network Address Translator*)[4]. El efecto de un NAT en VoIP es que no se pueden recibir conexiones iniciadas desde el exterior; en consecuencia, el que inicia la llamada detrás de un NAT no puede escuchar a la otra parte. Si los dos comunicantes se encuentran detrás de sus respectivos NAT, ningún flujo de audio originado llegará a su destino final. Para este problema ya existen soluciones implementadas en Asterisk (apartado 5.2.3).

5.2.1.1. Elementos de SIP

Los elementos básicos de un sistema SIP son los agentes de usuario (UA, *User Agent*) y los servidores. Estos últimos pueden ser de diferentes tipos: *Proxy*, de Registro y de Redirección. La configuración más simple para establecer una sesión SIP utiliza sólo

[4] Los NAT son traductores de direcciones IP, usados principalmente para permitir a máquinas conectadas a LAN con direcciones IP privadas, acceder a servidores en Internet (que usan direcciones IP públicas).

dos UA conectados uno a otro. El protocolo SIP permite el establecimiento de sesiones multimedia entre dos o más usuarios mediante el intercambio de mensajes entre las partes.

Agentes de Usuario (UA). Son los puntos extremos del protocolo SIP, es decir, los que emiten y procesan los mensajes del protocolo. Un videoteléfono, un teléfono, una aplicación cliente y cualquier otro dispositivo similar es un agente de usuario para SIP. El protocolo SIP no se ocupa de la interfaz de estos dispositivos con el usuario final; sólo se interesa por los mensajes que estos generan y cómo se comportan al recibir determinados mensajes.

Los agentes de usuario se comportan como clientes (UAC: *User Agent Clients*) y como servidores (UAS: *User Agent Servers*). Un agente de usuario se comporta como UAC cuando realiza una petición, y como UAS cuando la recibe y responde a la misma. Por esto los agentes de usuario deben implementar un UAC y un UAS.

Servidores de Registro. SIP permite establecer la ubicación física de un usuario determinado, esto es, en qué punto de la red está conectado. Para ello se vale del mecanismo de registro. Cada usuario tiene una dirección lógica que es invariable respecto de su ubicación física; una dirección lógica del protocolo SIP tiene la forma usuario@dominio. La dirección física, en cambio, es dependiente del lugar en donde el usuario está conectado (su dirección IP). Cuando un usuario inicializa su terminal (e.g. conectando su teléfono o abriendo su aplicación de telefonía SIP) el agente de usuario SIP que reside en dicho terminal envía una petición con el método REGISTER a un Servidor de Registro, informando a qué dirección física debe asociarse la dirección lógica del usuario. El Servidor de Registro realiza entonces la asociación, la cual tiene un período de vigencia que termina si no es renovada, y también se puede deshacer mediante un desregistro.

Un Servidor de Registro es comúnmente sólo una entidad lógica, y la mayoría de las veces se localiza junto con el Servidor *Proxy*.

Servidores *Proxy* y de Redirección. Para encaminar un mensaje entre un agente de usuario cliente y un agente de usuario servidor normalmente se recurre a los servidores.

El *Proxy* se encarga de encaminar las invitaciones de la sesión para llevarlas hasta el UA llamado. El servidor de Redirección genera una respuesta que indica al que origina la comunicación, la dirección del destino o la de otro servidor que lo acerque al destino; este tipo de servidor sólo escucha peticiones y retorna respuestas que contienen la localización actual de un usuario en particular o de otro servidor.

La principal diferencia entre un servidor *Proxy* y un servidor de Redirección es que el primero se queda formando parte de la comunicación entre el UAC y el (o los) UAS, mientras que el servidor de Redirección, una vez que indica al UAC cómo encaminar el mensaje, ya no interviene más. Un mismo servidor puede actuar como Redirección o como *Proxy* dependiendo de la situación.

Un conjunto de usuarios que pertenecen a una compañía o proveedor de servicios de comunicaciones, conforman un dominio. Este dominio, que se indica en una dirección SIP después del caracter "@", es atendido por al menos un servidor. Un agente de usuario normalmente encamina todas sus peticiones hacia un servidor de su propio dominio, el cual determina (por sus propios medios o valiéndose de otros servidores) la ubicación de los usuarios que son llamados por el agente de usuario en cuestión. El servidor que recibe las peticiones originadas por los usuarios de un dominio hacia otros dominios recibe el nombre de Servidor Saliente (*Outbound Server*). Por su parte, un servidor que recibe las peticiones destinadas a un dominio específico es denominado Servidor Entrante (*Inbound Server*).

5.2.1.2. Mensajes SIP

Existen dos tipos básicos de mensajes SIP: Peticiones y Respuestas. Ambos tipos emplean un formato de mensaje genérico, que consiste en una línea inicial (*Start Line*) seguida de uno o más campos de cabecera (*Message Header*), una línea vacía que indica el final de las cabeceras, y por último el cuerpo del mensaje (*Message Body*), que es opcional.

La línea inicial contiene la versión del protocolo, y el método y direcciones involucradas en la sesión, en el caso de las Peticiones, o el estado de la sesión, en el caso de las Respuestas. La cabecera contiene información relacionada con la llamada en formato de texto; por ejemplo, el origen y destino de la petición, el identificador de la llamada, etc. El cuerpo del mensaje o carga útil lleva la información, comúnmente mensajes SDP o ISUP (*ISDN User Part*) en caso de interfuncionamiento con la RTPC.

Las Peticiones se emplean para iniciar alguna acción o para solicitar información. La línea inicial de un mensaje de Petición (llamada también *Request Line*) incluye el nombre del método al que invoca, que puede ser uno de los siguientes:

- INVITE: Utilizado para invitar un usuario a participar en una sesión o para modificar parámetros.
- ACK: Confirma el establecimiento de una sesión.
- OPTION: Solicita información sobre las capacidades de un servidor.
- BYE: Indica la finalización de una sesión.
- CANCEL: Cancela una petición pendiente.
- REGISTER: Registra un UA.
- PRACK: Confirmación de respuesta provisional.

Las Peticiones no contienen por lo general un cuerpo de mensaje, porque no lo requieren.

Las Respuestas se generan como retorno de una petición, devolviendo un código numérico de estado. La línea inicial de un mensaje de Respuesta (llamada también *Status*

Line) incluye el código de respuesta y una pequeña descripción de ese código. Hay seis clases de códigos de respuesta, a saber:

- **1xx**: Mensaje provisional. La petición fue recibida pero se desconoce aún el resultado del procesamiento. El emisor se abstiene de enviar retransmisiones después de recibir una respuesta de este tipo. Son ejemplos el código 180 (Ringing) y el 100 (Trying).
- **2xx**: Éxito. Son respuestas finales positivas. La petición fue recibida y procesada exitosamente. Por ejemplo, 200 (OK) significa que el extremo llamado aceptó la invitación a la sesión.
- **3xx**: Redirección: Son usados para redireccionar las llamadas. Dan información acerca de la nueva localización de un usuario o sobre un *Proxy* alterno que puede resolver satisfactoriamente alguna petición. El emisor del mensaje de petición debe reenviar su petición a otro para que su petición sea atendida.
- **4xx**: Fallo de método. Son respuestas finales negativas. Falla del lado del emisor, mala sintaxis del mensaje, etc.
- **5xx**: Fallos de servidor. Falla del lado del servidor. Aparentemente la petición es válida pero el *Proxy* es incapaz de procesarla. El emisor debe reintentar después.
- **6xx**: Fallos globales. La petición no puede ser atendida en ningún *Proxy*.

5.2.1.3. Transacciones y Diálogos SIP

Una transacción SIP es una secuencia de mensajes entre dos elementos de red. Una transacción corresponde a una petición y todas las respuestas a esa petición. Esto quiere decir que una transacción incluirá cero o más respuestas provisionales y una o más respuestas finales. En el caso de un mensaje INVITE, puede ser dividido por un *Proxy* y por lo tanto tendrá múltiples respuestas finales. Las entidades SIP que almacenan el estado de las transacciones se denominan *Stateful* y llevan un registro de cada transacción.

Un diálogo SIP es una conversación par a par (*peer-to-peer*) entre dos UA. Los diálogos son identificados usando los campos Call-ID, From y To. Los mensajes que tienen estos campos iguales pertenecen al mismo diálogo. El campo CSEQ es utilizado para ordenar los mensajes en un diálogo. De hecho, CSEQ representa el número de transacción. De forma simple se puede decir que un diálogo es una secuencia de transacción.

5.2.1.4. Flujo de establecimiento de una sesión SIP

En una sesión SIP común se encuentran la siguientes etapas:

- **Registro** (Figura 5.3): Para que un usuario pueda ser llamado por otro, debe registrarse primero ante el *Proxy*. El registro consiste en el envío de un mensaje REGISTER seguido de su correspondiente respuesta 200 (OK). En caso de que el usuario no haya dado credenciales válidas, recibirá por respuesta un mensaje

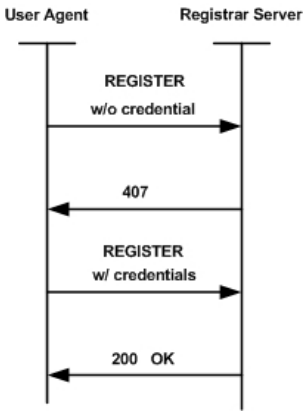


Figura 5.3.: Registro SIP.

407, con lo cual tendrá que reenviar el mensaje de Registro hasta que tenga éxito.

- **Invitación a una sesión** (Figura 5.4): Una invitación inicia con el mensaje INVITE dirigido comúnmente al *Proxy*. Este responde con 100 (Trying) para detener las retransmisiones y reenvía las peticiones hacia el usuario llamado. Todas las respuestas provisionales generadas por el usuario llamado son entregadas al usuario origen. Por ejemplo, 180 (Ringing) que es un mensaje que se envía cuando el usuario es contactado y comienza a timbrar. La respuesta 200 (OK) se genera en cuanto el usuario llamado descuelga el auricular.
- **Terminación de sesión** (Figura 5.5): Una sesión es finalizada cuando uno de los usuarios envía el mensaje BYE al otro extremo. El otro usuario confirma el final de la conversación enviando por respuesta un mensaje 200 (OK). La transacción que finaliza la sesión se realiza de un extremo a otro sin pasar por el *Proxy,* a menos que en el mismo se haya establecido un proceso de Registro de ruta. Existen situaciones en las que el *Proxy* requiere permanecer en la ruta de todos los mensajes con fines de control del tráfico o, por ejemplo, cuando existe un NAT. El *Proxy* logra esto insertando el campo RECORD ROUTE en las cabeceras de los mensajes SIP.

5.2.1.5. Protocolo de Descripción de Sesión (SDP)

SDP es un formato para describir parámetros de inicialización de flujo audiovisual. Está diseñado para transportar información de la sesión hacia los destinatarios, así como información de los flujos audiovisuales referentes a la misma. Permite además asociar más de un flujo audiovisual a una misma sesión; por ejemplo, en una misma sesión puede existir un flujo para audio y uno más para video o transferencia de documentos.

SDP es usado exclusivamente para la descripción y negociación de los parámetros de sesión; no transporta el flujo audiovisual en sí. Fue pensado para trabajar en conjunto

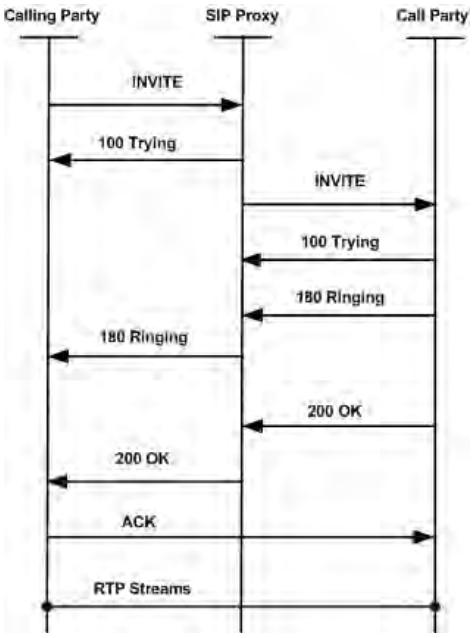
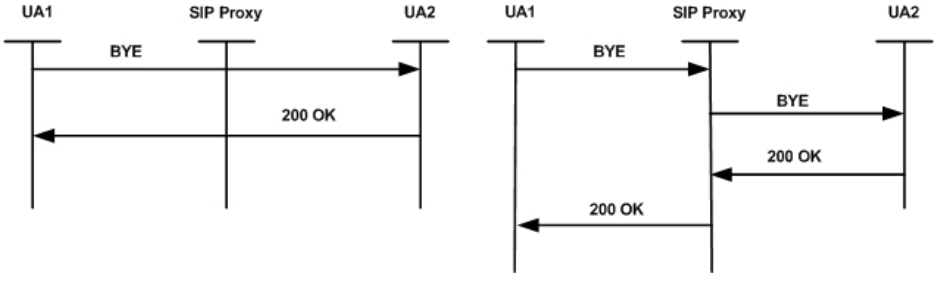


Figura 5.4.: Inicio de una sesión SIP.



a) Sin registro de ruta b) Con registro de ruta

Figura 5.5.: Fin de una sesión SIP.

con otros protocolos como SIP, Megaco o HTTP. El transporte de información acerca de los flujos audiovisuales permite a los destinatarios participar en la sesión si ellos soportan dichos flujos. Además, SDP permite la negociación de los parámetros de flujo tales como la tasa de muestreo de la señal, el tamaño de los paquetes, etc.

La información que SDP incluye en sus paquetes de forma general es la siguiente:

- La versión del protocolo.
- El nombre de la sesión y su propósito.
- El tiempo que la sesión está activa.
- Los medios relacionados con la sesión (video, audio, formatos para video y audio, etc.)
- Las direcciones IP y los puertos pertinentes para el establecimiento de la sesión.
- Los atributos específicos de la sesión o de los medios dentro de ella.

5.2.1.6. Protocolos RTP/RTCP

Son los protocolos usados para transportar flujos de audio/video en Telefonía IP. RTP es utilizado para transportar flujos en tiempo real (*real-time streaming*) y RTCP para monitorear la calidad del servicio, así como para transportar información acerca de los participantes en la sesión. Sus funciones generales son:

- Identificación del tipo de carga útil transportada (códecs de audio/video).
- Verificación de la entrega de los paquetes en orden (usando marcas de tiempo) y, si resulta necesario, reordenamiento de los bloques fuera de orden.
- Transporte de información de sincronización para la codificación y decodificación.
- Monitoreo de la entrega de la información.

RTP utiliza UDP para el transporte de la información y aprovecha la suma de verificación (*checksum*) del mismo para verificar la integridad de los datos. RTCP también utiliza UDP para enviar paquetes de control hacia todos los participantes de una sesión.

5.2.2. H.323

Forma parte del grupo de recomendaciones H.300 de la UIT-T que define el funcionamiento de sistemas y equipos terminales para servicios audiovisuales. Particularmente, H.323 es una recomendación que agrupa diferentes estándares para especificar un sistema de comunicaciones multimedia a través de redes de paquetes IP. Su primera versión fue definida en el año 1996, tiempo en el cual no había disponible ningún estándar que permitiera establecer mecanismos de interoperabilidad entre fabricantes y desarrolladores de sistemas de VoIP; por este motivo se convirtió en el protocolo más utilizado y de mayor aceptación en el mercado. Actualmente sigue siendo utilizado en gran medida por los grandes operadores de VoIP, y a la par del protocolo SIP es uno

de los estándares más utilizados por los desarrolladores de soluciones IP. La versión actual de la recomendación es la H.323v7, que fue publicada en el 2009.

Los protocolos más relevantes involucrados en H.323 son:

- H.225: Es el encargado de definir los procesos de señalización de las llamadas, así como de la gestión del registro y las características de los usuarios del sistema.
- H.245. Su labor es controlar las llamadas, definiendo los parámetros para el establecimiento, mantenimiento y cierre de los canales lógicos utilizados.
- H.450.x: Establece los servicios suplementarios de H.323, como desvío y llamada en espera.
- H.235: Define los mecanismos de seguridad y autenticación para las comunicaciones multimedia.

Es importante destacar que los protocolos anteriores se encargan de la señalización de las comunicaciones; una vez establecido el canal H.323, se utiliza el protocolo RTP para el transporte de los paquetes audiovisuales involucrados en la llamada.

Componentes y topología: Un sistema de VoIP basado en H.323 consta de 4 elementos fundamentales: termínales, pasarelas (*gateways*), MCU (Unidades de Control Multipunto) y controladores de acceso (*gatekeepers*). Estos elementos se agrupan en zonas, constituidas por diversos nodos H.323 gestionados por un solo controlador de acceso.

- **Terminales**: Son componentes en los que terminan las comunicaciones de voz y opcionalmente video y datos. Es obligatorio que los terminales soporten comunicaciones con el códec G.711 y los protocolos H.245, H.225 y RAS (Registro, Admisión y Estado). Otros protocolos y códecs son opcionales según los tipos de servicios que se estén prestando.
- **Controladores de acceso**: Son los nodos centrales de un sistema H.323. Se encargan de controlar las comunicaciones y la conexión entre los terminales. Su presencia no es necesaria para la realización de comunicaciones entre terminales de un mismo segmento, aunque sí es recomendable. Tienen las siguientes tareas fundamentales:
  - Conversión de direcciones de terminales H.323 a direcciones IP o E.164, para que sea posible la comunicación con terminales de otros segmentos o de una RTPC.
  - Administración del ancho de banda, asignando un ancho de banda a cada conferencia entre terminales y estableciendo comunicaciones hasta que se alcanza el ancho de banda máximo permitido, momento en el cual empieza a rechazar las solicitudes desde los terminales.
  - Control de admisión, a través del protocolo RAS, aceptando o negando solicitudes dependiendo del terminal o pasarela que las esté realizando.

  En caso de que una conferencia incluya a más de dos terminales, el controlador de acceso redirecciona la señalización al MCU que presta soporte a la multiconferencia.

- **Pasarelas**: Es un nodo opcional dentro de una zona H.323, encargado de garantizar la compatibilidad con otro tipo de redes distintas a H.323, como redes SIP o RTPC. Se encarga de la conversión de los protocolos de señalización de las llamadas y también de los formatos de audio y video entre las redes.
- **MCU**: Es un elemento también opcional, encargado de brindar el soporte para las conferencias que constan de tres o más terminales H.323. Está constituido por dos componentes: el MC (Controlador Multipunto), que controla la conexión con los diferentes terminales, definiendo el códec y el ancho de banda entre otros, y el MP (Procesador Multipunto), que lleva a cabo la multidifusión de los datos de audio y video entre los distintos terminales.

5.2.3. IAX (*Inter Asterisk eXchange*)

El protocolo IAX (ahora referido generalmente como IAX2 por su segunda versión) es uno de los protocolos utilizados por la centralita Asterisk (Sección 5.3) para manejar conexiones VoIP entre sus servidores, y entre servidores y clientes VoIP que lo utilizan.

IAX es robusto y muy simple en comparación con otros protocolos. Permite manejar una gran cantidad de códecs y un gran número de flujos de audio/video, lo que significa que puede ser utilizado para transportar virtualmente cualquier tipo de datos. Esta capacidad lo hace muy útil para realizar videoconferencias o presentaciones remotas.

IAX utiliza un único puerto UDP, generalmente el 4569, para comunicaciones de señalización y datos entre puntos terminales. El tráfico de voz es transmitido en banda (*in-band*)[5], lo que hace a IAX2 un protocolo casi transparente a los cortafuegos y realmente eficaz para trabajar dentro de redes internas. En esto se diferencia de SIP, que utiliza una conexión RTP fuera de banda (*out-of-band*)[6] para entregar la información.

IAX soporta entroncamiento (*trunking*), mediante el cual un sólo enlace permite enviar datos y señalización por múltiples canales. Cuando se realiza entroncamiento, los datos de múltiples llamadas son manejados en un único conjunto de paquetes, lo que significa que un datagrama IP puede entregar información para más llamadas sin crear latencia adicional. Esto es una gran ventaja para los usuarios de VoIP, pues las cabeceras IP ocupan un gran porcentaje del ancho de banda utilizado; en contraparte se consumen mayores recursos de equipo de cómputo.

El principal objetivo de IAX ha sido minimizar el ancho de banda utilizado en la transmisión de voz y vídeo a través de la red IP, con particular atención al control y a las llamadas de voz, y proveyendo un soporte nativo para ser transparente a los NAT. La estructura básica de IAX se fundamenta en la multiplexación de la señalización y el flujo de datos sobre un mismo puerto UDP entre dos sistemas.

[5] Comunicaciones que tienen lugar dentro de un método de comunicación previamente establecido.
[6] Se refiere a las comunicaciones que tienen lugar fuera de un método de comunicación previamente establecido.

5.3. La centralita telefónica Asterisk

Asterisk es un programa bajo licencia GPL creado por Digium Inc, que implementa una centralita (PBX) completa utilizando un equipamiento relativamente económico. Trabaja sobre Linux y otras plataformas, pero en Linux cuenta con el mayor soporte.

Puede trabajar con la mayoría de los equipos estándares de telefonía y operar con otras redes de telefonía global tradicional.

Ha sido adoptado en algunos entornos corporativos como una solución de bajo coste junto con otras aplicaciones para mejorar sus prestaciones (como el servidor SIP Express Router[7]). Puede interoperar con terminales IP actuando como un registrador y como pasarela entre ambos.

Incluye muchas características que anteriormente sólo estaban disponibles en costosos sistemas propietarios PBX, tales como:

- Buzón de voz.
- Conferencias.
- Respuesta Interactiva de Voz (IVR, *Interactive Voice Response*).
- Compatibilidad con SIP, H.323, IAX y MGCP.
- Creación de nuevas funcionalidades.
- Llamadas de conferencia.
- Llamada en espera.
- Transferencia de llamadas, internas y externas.
- Soporte para llamadas tripartitas.
- Identificación de llamadas.
- Música en espera y en transferencia (ficheros MP3 actualizables por el usuario).
- Soporte para fax.
- Grabación de llamadas entrantes y salientes.
- Monitorización de llamadas en curso.

La versión actual de Asterisk es la 1.8. Según los desarrolladores, sus características más fuertes son la estabilidad (como en la versión 1.4) y la seguridad, pero sobre todo la variedad de características que trae:

- Soporte para cifrado AES de 128 bits.
- Soporte para IPv6.
- Códec G.722 incluido, con capacidad de transcodificación con los demás códecs.
- CEL, un nuevo CDR (*Call Detail Record*) mucho más completo.

[7] `http://www.iptel.org/ser/`.

5.4. Los terminales de telefonía IP

Un terminal telefónico IP es un dispositivo completamente digital y programable que permite realizar una comunicación de voz o vídeo utilizando el protocolo IP, en una red LAN o a través de Internet.

Suelen tener más opciones y ventajas que un teléfono convencional; algunos pueden tener múltiples líneas, incluir cámara de vídeo para realizar videoconferencias, y dan la posibilidad de configurar la calidad del servicio (QoS) o una LAN virtual (VLAN). La configuración se realiza mediante un sistema de administración que puede ser accedido vía Web en una dirección IP asignada para tal fin.

Los principales tipos de terminales de telefonía IP son (Figura 5.6):

Teléfonos IP: Un teléfono IP suele ser un equipo con forma de teléfono, aunque con la particularidad de que utiliza una conexión de red de datos en lugar de una conexión de red telefónica.

ATA (Adaptador de Teléfono Analógico): Son dispositivos que permiten conectar un teléfono analógico o RDSI a una red de VoIP. Disponen de un sistema de administración y gestión similar a los teléfonos IP, por lo que poseen también dirección IP, y las mismas ventajas que cualquier terminal IP.

Teléfonos IP inalámbricos: Son similares a los teléfonos móviles (o celulares) y permiten utilizar redes inalámbricas para conectarse al servidor de VoIP. Existen teléfonos móviles con soporte de Wi-Fi y DECT (*Digital Enhanced Cordless Telecommunications*) para ser utilizados dentro de una LAN.

***Softphone*:** Es un programa que simula un teléfono convencional, y se instala en una computadora donde interactúa con micrófonos y auriculares/altavoces. Hace posible usar la computadora para realizar llamadas a otros *softphones* o a otros teléfonos convencionales, como cualquier otro teléfono IP, usando VoIP. Permite hacer parte de una red de telefonía IP, pero también conectarse a un proveedor de servicios de telefonía por Internet gratuito o de pago.

Figura 5.6.: Terminales de telefonía IP.

5.5. Interconexión con la red telefónica fija y la celular

Los sistemas de Telefonía IP, como Asterisk, permiten integrar una red de telefonía IP con redes telefónicas tradicionales por medio de interfaces analógicas y digitales. La conexión con líneas analógicas se hace a través de interfaces FXO (*Foreign eXchange Office*) y FXS (*Foreign eXchange Subscriber*); la conexión con líneas digitales RDSI se logra por medio de interfaces del tipo BRI (*Basic Rate Interface*) y PRI (*Primary Rate Interface*); para acceder a la RTPC se puede utilizar una interfaz o pasarela FXO (puede ser de múltiples líneas); para la interconexión con la red móvil se puede conseguir un conversor celular. En la Figura 5.7 se muestra un esquema de la interconexión de una centralita Asterisk con otras redes de telefonía tradicional.

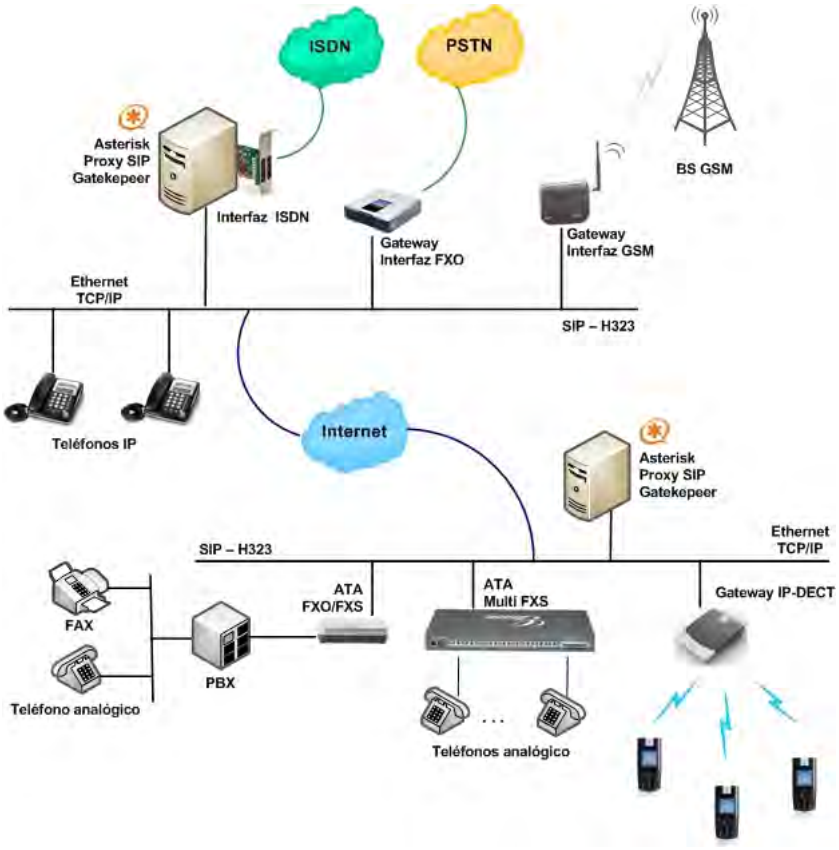


Figura 5.7.: Interconexión de Asterisk con la red telefónica fija y la celular.