Lima, Diciembre 2001

Estándares relacionados a la tecnología Voz sobre IP (VoIP): Su Clasificación

Autor: Enrique Landaure
elandaure@yahoo.com
http://www.geocities.com/elandaure/voip.html

En este artículo se presentan los dos principales estándares de la tecnología Voz sobre IP, ITU-T H.323 y IETF SIP. Se revisan sus modos de funcionamiento y se hace una comparación entre ambos. Mas adelante se mencionan cuáles son los protocolos de soporte de estos estándares explicando sus modos de operación. Finalmente, y a manera de resumen se presentan una tabla de los principales protocolos involucrados y sus principales funciones.

1. Estándar H.323

Este es el estándar de la Unión Internacional de Telecomunicaciones (ITU-T) que los fabricantes deben cumplir para proveer servicios de Voz sobre IP. Esta recomendación provee los requerimientos técnicos para comunicación de Voz sobre LAN en las que se asume no hay control de la calidad de servicio (QoS). Fue desarrollado originalmente, en el año 1,996, sólo para soportar conferencias multimedia en LANs, pero luego fue extendida para soportar Voz sobre IP. En relativo poco tiempo ha pasado por varias actualizaciones, siendo la última la versión 4 del mes de Febrero del 2001.

1.1 Componentes del H.323

El sistema de Voz Sobre IP está compuesto por los siguientes elementos o entidades:

1.1.1 **Terminales**. Pueden ser teléfonos tradicionales (analógicos, RDSI, GSM, etc.), computadoras personales con tarjeta de sonido, parlantes y micrófono (o *handset)*, o teléfonos IP. Estos elementos proveen comunicaciones en tiempo real en dos vías. Todos los terminales deben soportar H.245, Q.931, RAS (Registration Admission Status) y RTP (Real Time Transport Protocol). H.245 es usado para permitir el uso de los canales, Q.931 es requerido para señalización y establecimiento de la llamada, RTP es el protocolo de transporte en tiempo real que lleva los paquetes de voz mientras que RAS es usado para interactuar con el *gatekeeper*. Estos elementos pueden también incluir protocolos para conferencia de datos, codificadores de voz y soporte para MCU. Un terminal H.323 puede comunicarse con otro terminal H.323, gateway o un MCU.

1.1.2 **Gateways**. Un gateway es la entidad que provee comunicaciones en tiempo real en dos vías entre terminales H.323 en la red IP y otros terminales ITU en la red conmutada, o con otro gateway H.323. Realizan la función de traducción entre diferentes formatos de transmisión, por ejemplo de H.225 a H.221. También son capaces de traducir entre codificadores de audio y video.

El gateway es la interfaz entre la PSTN y la Internet, toman la voz de la PSTN y la colocan en la red IP y viceversa. Los gateways son opcionales cuando los terminales en una simple LAN pueden comunicarse entre sí directamente. Cuando los terminales en la red necesitan comunicarse con otra entidad en alguna otra red, pueden hacerlo vía gateways usando los protocolos H.245 y Q.931.

1.1.3 Gatekeepers. Es el componente mas importante de un sistema H.323 ya que hace las funciones de un manager. Actúa como el punto central para todas las llamadas dentro de su zona (una zona es el conjunto del gatekeeper y las entidades registradas con él) y provee servicios a las entidades registradas. Algunas de las funcionalidades aparecen a continuación:

- *Traducción de direcciones*. Traducción de una dirección alias a la dirección de transporte usando la tabla de traducción, la cual es actualizada usando los mensajes Registration.

- *Control de Admisiones*. Gatekeepers pueden otorgar o denegar accesos basados en autorización de la llamada, direcciones de origen y destino o algún otro criterio.

- *Señalización de llamadas*. El Gatekeeper puede escoger completar la señalización de la llamada con las entidades y puede procesar la llamada misma. Alternativamente, el Gatekeeper puede guiar a las entidades a conectarse al Canal de Señalización de la Llamada directamente entre sí.

- *Autorización de la Llamada*. El Gatekeeper puede rechazar llamadas de un terminal debido a una falla de autorización a través del uso de la señalización H.225. Las razones para el rechazo podrían ser el acceso restringido durante algunos períodos de tiempo o acceso restringido hacia o desde terminales o gateways particulares.

- *Administración del Ancho de Banda*. Control del número de terminales H.323 permitidos simultáneamente accediendo a la red. A través de la señalización H.225, el Gatekeeper puede rechazar llamadas desde un terminal debido a limitaciones de ancho de banda.

- *Administración de la Llamada*. El Gatekeeper puede mantener una lista de llamadas H.323 en curso. Esta información puede ser útil para indicar que el terminal llamado se encuentra ocupado, y para proveer información para la función de Administración de Ancho de Banda.

1.1.4 Unidades de Conferencia Multipunto (MCU). Administran conferencias multipartitas. El MCU es el elemento de la red que provee capacidad para 3 ó mas terminales y gateways para participar en una conferencia multipartita. El MCU consiste de un obligatorio Controlador Multipartita (MC) y opcionales

Procesadores Multipartita (MP). El MC determina las características comunes de los terminales usando H.245 pero no realiza la multiplexación de audio, video y datos. La multiplexación de los flujos de medios es administrada por el MP bajo el control del MC.

La figura 1 muestra la interacción de todos los componentes H.323.

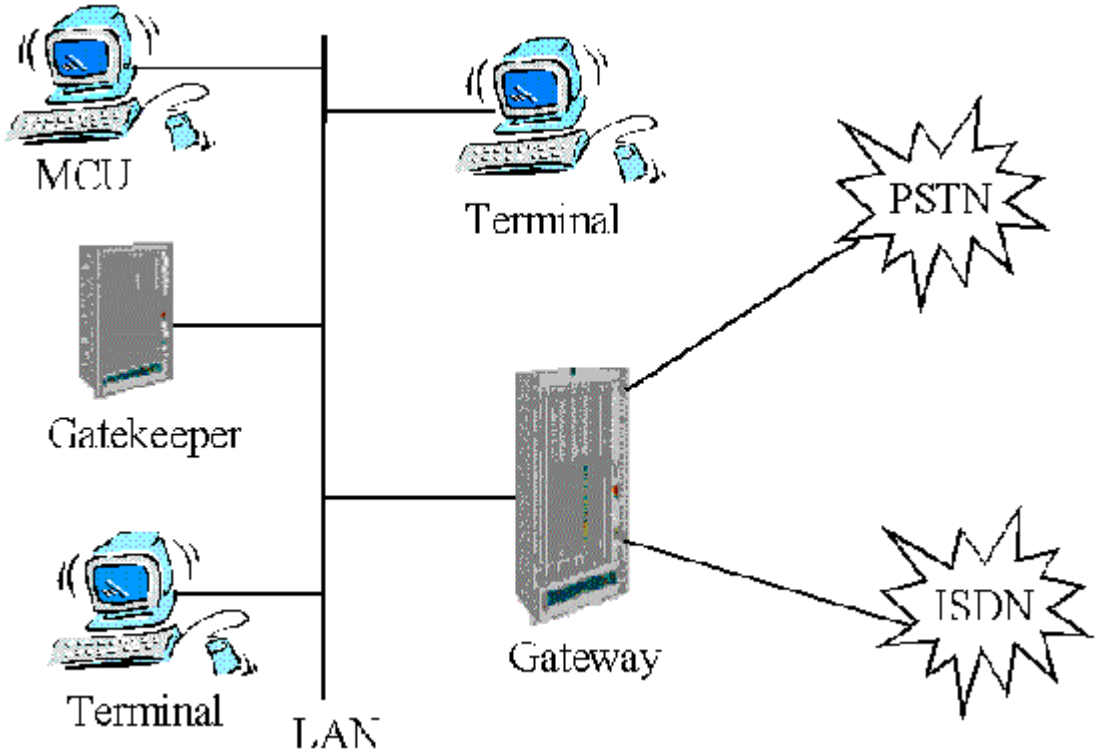

Figura 1. Interacción de los componentes H.323

Estos componentes pueden ser implementados en hardware o en software, integrados o separadamente. Se comunican entre sí gracias a los protocolos de señalización y transporte

El establecimiento y el mantenimiento de conexiones H.323 se realiza sobre los protocolos TCP o UDP:

- Q.931 sobre TCP que se realiza a través del conocido puerto 1720 para negociar el puerto de conexión del protocolo H.245.
- H.245 sobre TCP para realizar las negociaciones de los parámetros (codificadores entre otros) y realiza las conexiones UDP para RTP y RTCP.
- RTP y RTCP sobre UDP en que se usan conexiones UDP para mantener los flujos asociados con el tráfico H.323.

1.2 Pila de Protocolos del H.323

La siguiente figura muestra la pila de protocolos del estándar H.323. Los paquetes de audio, video y registro usan como protocolo de transporte al protocolo no confiable UDP (User

Datagram Protocol). Excepto por el protocolo T.120, que es usado para la definición de conferencias de datos, los demás protocolos se estudian mas adelante.

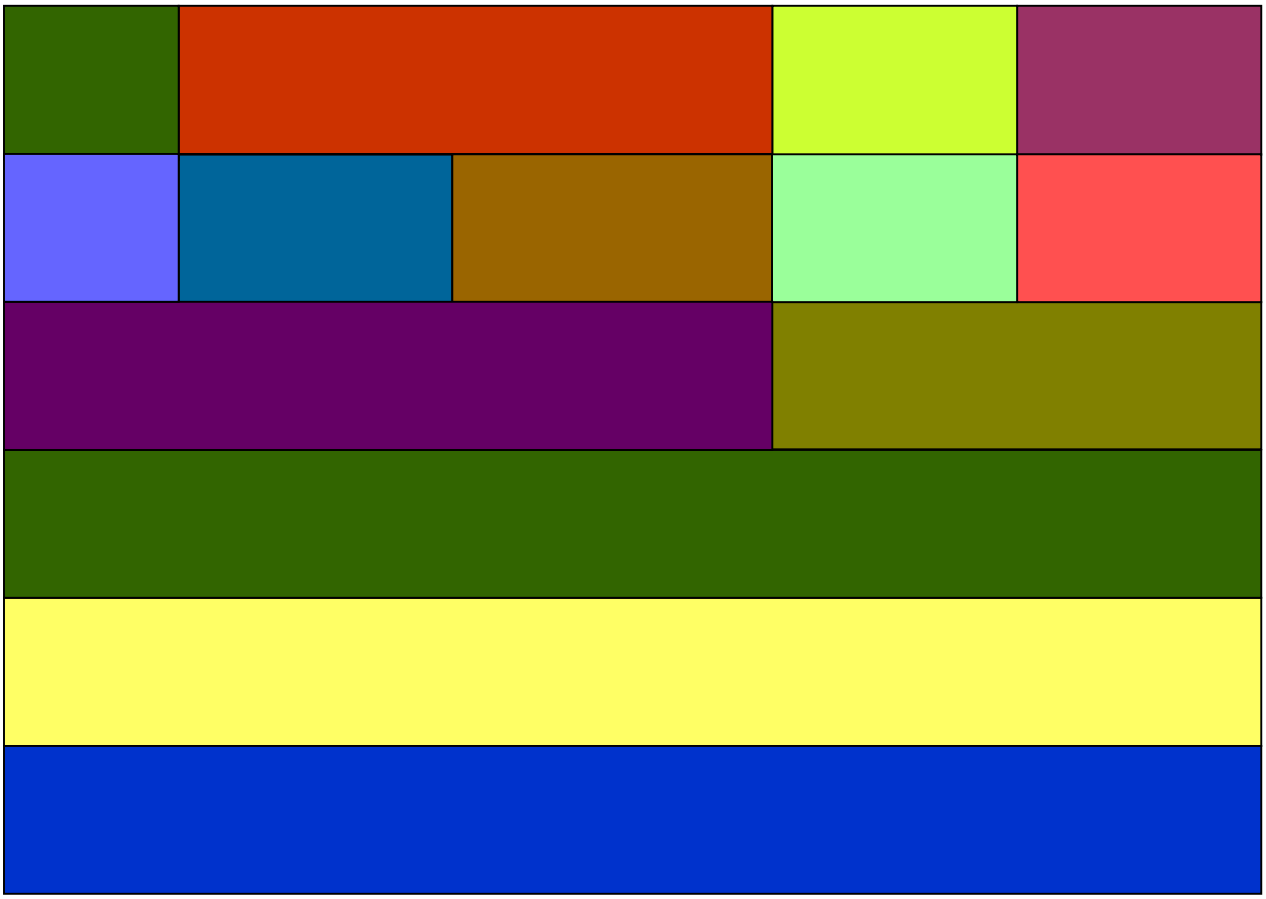

Figura 2. Pila de Protocolos de H.323

1.3 Control y Señalización en H.323

H.323 provee tres protocolos de control: señalización de llamada H.225.0/Q.931, RAS H.225.0 y control de medios H.245.

H.225/Q.931 es usado en conjunto con H.323 y provee la señalización para el control de la llamada. Para establecer una llamada desde el terminal origen hasta el del destino, el RAS (registro, admisión y señalización) del H.225 es usado. Después que la llamada ha sido establecida, el H.245 es usado para negociar la transmisión de medios.

1.3.1 RAS H.225. El canal RAS es usado para la comunicación entre los terminales y el gatekeeper. Los procedimientos definidos por el canal RAS son:

- *Descubrimiento del Gatekeeper*. Este es el proceso que un terminal usa para determinar con cuál de los gatekeeper se irá a registrar. El terminal normalmente difunde un mensaje GRQ (Gatekeeper Request) preguntando por su gatekeeper. Uno o mas gatekeepers puede responder con el mensaje de confirmación del Gatekeeper GCF (Gatekeeper Confirmation) con el que se muestran como disponibles para atender al terminal. La respuesta incluye la dirección de transporte del canal RAS del gatekeeper. Los gatekeepers que no estén disponibles para registrar al terminalse lo harán saber enviando un mensaje de rechazo GRJ (Gatekeeper Reject). Si más de un gatekeeper responde con GCF, entonces el terminal puede escoger el gatekeeper y registrarse con él. Si ningún gatekeeper responde dentro de un intervalo de tiempo determinado, el terminal puede retransmitir el mensaje GRQ.

- *Registro de los terminales*. Este es el proceso por el que un terminal se asocia a una zona e informa al gatekeeper sus direcciones de transporte y alias. Todos los terminales usualmente se registran con el gatekeeper que fue identificado a través del proceso de descubrimiento. Un terminal debe enviar un mensaje Solicitud de Registro RRQ (Registration Request) a un gatekeeper. Este es enviado a la dirección de transporte del canal RAS del gatekeeper. El terminal tiene la dirección de red del gatekeeper desde el procedimiento de descubrimiento del gatekeeper y usa el muy conocido identificador TSAP del canal RAS. El gatekeeper debe responder ya sea con un mensaje de Confirmación de Registro RCF (Registration Confirmation) o con un mensaje de Rechazo de Registro RRJ (Registration Reject). El gatekeeper debe asegurar que cada dirección alias se traduzca únicamente en una simple dirección de transporte. Un terminal puede cancelar su registro enviando un URQ (Unregister Request) al gatekeeper. El gatekeeper debe responder con mensaje de confirmación de cancelación de registro UCF (Unregister Confirmation).

- *Localización de los terminales.* Un terminal o gatekeeper que tiene la dirección alias de un terminal y que desea obtener la información de contacto envía el mensaje de petición de localización LRQ (Location Request). El gatekeeper con el que el terminal requerido es registrado debe responder con el mensaje de confirmación de localización LCF (Location Confirmation) conteniendo la información de contacto del terminal o de su gatekeeper. Todos los gatekeepers con los que el terminal requerido no está registrado deben retornar el mensaje de rechazo de localización LRJ (Location Request) si es que ellos recibieron el LRQ en el canal RAS.

- *Admisiones, cambio de ancho de banda, estado y liberación*. El canal RAS es también usado para la transmisión de mensajes de Admisiones, cambio de ancho de banda, estado y liberación. Estos mensajes son intercambiados entre un terminal y un gatekeeper y son usados para proveer funciones de control de admisiones y administración del ancho de banda. El mensaje de requerimiento de admisiones ARQ (Admissions Request) especifica el

ancho de banda de la llamada requerida. El gatekeeper puede reducir el ancho de banda de la llamda requerida con el mensaje Confirmación de Admsiones ACF (Admissions Confirm). Un terminal o gatekeeper puede intentar modificar el ancho de banda de la llamada durante la llamada usando el mensaje de requerimiento de cambio de ancho de banda BRQ (Bandwidth Change Request).

1.3.2 Señalización de la llamada con H.225.0

El canal de señalización de la llamada es usado para llevar los mensajes de control del protocolo H.225. En redes que no existe ningún gatekeeper, los mensajes de señalización de llamadas son pasados directamente entre el terminal que llama y el llamado usando Direcciones de Transporte de Señalización de Llamada. Se asume que el terminal que llama conoce la Dirección de Transporte de Señalización de Llamada del terminal llamado y así se puede comunicar directamente. En redes que sí tienen el gatekeeper, el intercambio de mensajes de admisión inicial toma lugar entre el terminal que llama y el gatekeeper usando la dirección de transporte del canal RAS del gatekeeper. La señalización de la llamada es hecha sobre el canal confiable TCP.

- ***Enrutamiento del canal de Señalización de Llamada***. Los mensajes de Señalización de Llamadas pueden ser pasados de dos maneras. La primera manera es Señalización de la llamada enrutada del gatekeeper donde los mensajes de señalización de la llamada son enrutados a través del gatekeeper entre los terminales. La otra alternativa es Señalización Directa de la Llamada de terminales donde los mensajes de señalización son pasados directamente entre los terminales. Los mensajes de admisión son intercambiados con el gatekeeper sobre el canal RAS, seguidos por un intercambio de mensajes de señalización en el Canal de Señalización de la Llamada el que a su vez es seguido por el establecimiento del canal de control H.245.
- *Enrutamiento del Canal de Control*. Cuando la señalización de la llamada enrutada del gatekeeper es usada, hay dos métodos para enrutar el canal de control H.245. La primera alternativa es establecer el canal de control H.245 directamente entre los terminales mientras que en el segundo caso, el establecimiento del canal de control H.245 es hecho a través del gatekeeper.

1.3.3 Control de medios y de conferencia H.245

H.245 es el protocolo de control de medios que el sistema H.323 usa después que la fase de establecimiento de la llamada ha sido concluida. H.245 es usado para negociar y establecer todo los canales de medios llevados por RTP/RTCP. Las funcionalidades ofrecidas por H.245 son:

- *Determinación del maestro y esclavo*. H.245 apunta al Controlador Multipunto (MC) el cual se mantiene como responsable para el control central en casos donde una llamada es extendida a una conferencia.

- *Capacidad de Intercambio*. H.245 es usado para negociar las características cuando una llamada ha sido establecida. La capacidad de intercambio puede ocurrir en cualquier momento durante la llamada, por lo que se permiten renegociaciones en cualquier momento.

- *Control del Canal de Medios*. Después que los terminales de una conferencia han intercambiado sus características, ellos pueden abrir y cerrar canales lógicos de los medios. Dentro de H.245 los canales de medios son abstraídos como canales lógicos (que sólo son identificadores).

- *Control de Conferencia*. En conferencias, H.245 provee a los terminales información de ellos mismos y establece el modelo de flujo de medios entre todos los terminales.

1.4 Establecimiento de la llamada en H.323

El procedimiento de establecimiento de una llamada incluye:

- Descubrimiento del gatekeeper el cual podría tomar el manejo del terminal.
- Registro del terminal con su gatekeeper.
- Terminal entra a la fase de establecimiento de la llamada.
- El intercambio de características toma lugar entre los terminales y el gatekeeper.
- La llamada es establecida.
- Cuando el terminal está conectado, éste puede terminar la llamada. La terminación puede también ser hecha por el gatekeeper.

2. Estándar SIP

Este es el estándar de la IETF para establecimiento de conexiones VoIP. Es un protocolo de control de la capa de aplicación para creación, modificación y terminación de sesiones con uno o más participantes. La arquitectura de SIP es similar a HTTP (protocolo cliente / servidor). Las solicitudes son generadas por el cliente y enviadas al servidor. El servidor procesa las solicitudes y envía una respuesta al cliente. Una solicitud y su respuesta conforman una transacción. SIP tiene los mensajes INVITE y ACK que definen el proceso de abrir un canal confiable sobre el que los mensajes de control de la llamada pueden pasar. SIP hace las mínimas suposiciones acerca del protocolo de transporte subyacente. Este protocolo provee por sí mismo confiabilidad y no depende de esta característica del TCP. SIP depende del protocolo de Descripción de Sesión (SDP) para transmitir la negociación para identificación de codificadores. SIP soporta descripciones de sesión que permite a los participantes estar de acuerdo en un juego de tipos de medios compatibles. También soporta movilidad del usuario a través de solicitudes de redirección o vía proxy a la localización actual del usuario. Los servicios que SIP provee incluyen:

- *Localización de Usuarios*. Determinación del sistema final a ser usado para la comunicación.

- *Establecimiento de la llamada*. Timbrado y establecimiento de los parámetros de la llamada en ambos terminales de la llamada.

- *Disponibilidad del usuario*. Determinación del deseo del usuario llamado para aceptar llamadas entrantes.

- *Capacidades del usuario*. Determinación de los medios y sus parámetros a ser usados.

- *Manejo de la llamada*. Transferencia y terminación de las llamadas.

2.1 Componentes de SIP

El estándar SIP consiste de dos componentes: agentes del usuario y servicios de red.

2.1.1 Agentes del usuario. Un agente de usuario es un sistema final actuando en favor del usuario. Hay dos partes: un cliente y un servidor. La porción del cliente es llamada el agente de usuario cliente (UAC) mientras que la porción del servidor es llamada agente de usuario servidor (UAS). El UAC es usado para iniciar una solicitud SIP mientras que el UAS es usado para recibir solicitudes y retornar respuestas en favor del usuario.

2.1.2 Servicios de red. Hay tres tipos de servidores dentro de una red. Un servidor de registro recibe actualizaciones concernientes a las localizaciones actualizadas de los usuarios. Un servidor proxy de solicitudes recibidas las reenvía al servidor siguiente (next-hop server) el cual tiene más información sobre la localización del usuario llamado. Un servidor de redirección de solicitudes recibidas, determina el servidor siguiente (next-hop) y retorna la dirección del siguiente servidor (next-hop) al cliente.

2.2 Mensajes SIP

SIP define un gran número de mensajes. Estos mensajes son usados para comunicación entre el cliente y el servidor SIP. Estos mensajes son:

- *INVITE*. Para invitar a un usuario a una llamada.
- *BYE*. Para terminar una conexión entre dos terminales.
- *ACK*. Para intercambio confiable de mensajes de invitación.
- *OPTIONS*. Para conseguir información sobre las capacidades o características de la llamada.
- *REGISTER*. Da información sobre la localización del usuario al servidor de registro SIP.
- *CANCEL*. Para terminar la búsqueda de un usuario.

2.3 Modo de Operación SIP

Llamadores y llamados son identificados por direcciones SIP. Cuando se hace una llamada SIP, el que llama necesita primero localizar el servidor apropiado y enviarle una solicitud. El que llama puede alcanzar directamente al llamado o indirectamente a través de servidores de redirección. El campo Call ID en la cabecera del mensaje SIP identifica

unívocamente las llamadas. A continuación se explica el modo de operación del estándar SIP.

2.3.1 Identificación de Direcciones en SIP. Los servidores SIP son identificados por un URL SIP el que es de la forma sip:username@host. Una dirección SIP puede designar a un individuo o a un grupo entero.

2.3.2 Localización de un servidor IP. El cliente puede enviar la solicitud a un servidor proxy SIP o puede enviarla directamente a la dirección IP y puerto correspondientes al identificador de solicitudes uniforme URI (Uniform Request Identifier).

2.3.3 Transacción SIP. Una vez que el host de la solicitud URI identificó a un servidor SIP, el cliente puede enviar solicitudes a ese servidor. Una solicitud junto con las respuestas conforman una transacción SIP. Las solicitudes pueden ser enviadas a través del protocolo confiable TCP o del protocolo no confiable UDP.

2.3.4 Invitación SIP. Una satisfactoria invitación consiste de dos solicitudes: un INVITE seguido de un ACK. La solicitud INVITE pregunta al llamado unirse a una conferencia particular o establecer una conversación de dos participantes. Después que el llamado haya aceptado participar en la llamada, el que llama confirma esta aceptación recibiendo el mensaje ACK. La solicitud INVITE contiene una descripción de sesión que provee a la parte llamada con suficiente información para unirse a la sesión. Si el llamado desea aceptar la llamada, este responde a la invitación retornando una descripción de sesión similar.

2.3.5 Localización de un Usuario. La persona llamada puede cambiar su localización con el tiempo. Estas localizaciones pueden ser dinámicamente registradas con el servidor SIP. Cuando el servidor SIP es consultado sobre la localización del que se desea llamar, retorna una lista de las posibles localizaciones. Un servidor de localizaciones en un sistema SIP realmente es el que genera la lista y se la pasa al servidor SIP.

2.3.6 Cambio de una Sesión Existente. Algunas veces se necesita cambiar los parámetros de una sesión existente. Esto es hecho reenviando el mensaje INVITE con el mismo *Call ID* pero con un nuevo cuerpo que contendrá la nueva información.

2.4 Ejemplo de Operación SIP

A continuación se presenta como ejemplo de la operación SIP el caso en que un participante invita a un cliente a una llamada. Un cliente SIP crea un mensaje INVITE para luisenrique@power.com, el que es normalmente enviado a un servidor proxy. Este servidor proxy intenta obtener la dirección IP del servidor SIP que administra las solicitudes del dominio requerido. El servidor proxy consulta al Servidor de Localización para determinar el servidor siguiente (next-hop). El servidor de localización es un no-SIP

que guarda información sobre los servidores siguientes (next-hop) para diferentes usuarios. Al obtener la dirección IP del servidor siguiente (next-hop), el servidor proxy reenvía el mensaje INVITE al servidor siguiente (next-hop). Después que el servidor agente del usuario (UAS) ha sido alcanzado, se envía una respuesta de regreso al servidor proxy. El servidor proxy a su vez envía de regreso una respuesta al cliente. El cliente entonces confirma que ha recibido la respuesta enviando un ACK. El intercambio de mensajes es mostrado en la Figura 3. En este caso, se ha asumido que la solicitud INVITE del cliente fue derivada al servidor proxy. Sin embargo, si hubiese sido derivada al servidor de redirección, éste retornaría la dirección IP del servidor siguiente (next-hop) al cliente. El cliente entonces se comunica directamente con el UAS.

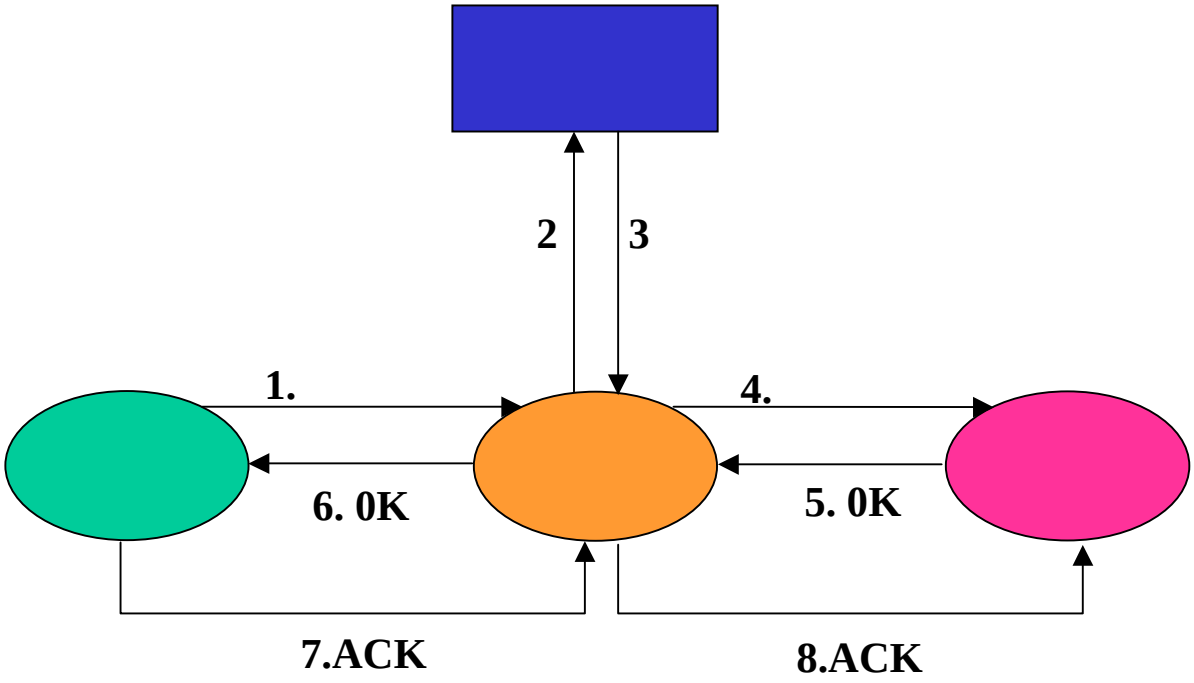

Figura 3. Ejemplo de Operación SIP

3. Comparación entre los estándares H.323 y SIP

Los que proponen SIP claman que dado que H.323 fue diseñado pensando en señalización ATM y RDSI, H.323 no está bien diseñado para controlar sistemas de voz sobre IP. Ellos dicen que H.323 es inherentemente complejo, tiene *overheads* y por tanto es ineficiente para VoIP. También mencionan que H.323 carece de la extensibilidad requerida del protocolo de señalización para VoIP. Como SIP ha sido diseñado manteniendo a Internet en mente, se evitan la complejidad y problemas de extensibilidad. SIP reutiliza la mayoría de los campos cabecera, reglas de codificación, códigos de errores y mecanismos de autenticación de HTTP. H.323 define cientos de elementos mientras que SIP tiene sólo 37 cabeceras, cada una con un pequeño número de valores y parámetros. H.323 usa una

representación binaria para sus mensajes, la cual está basada en ASN.1 mientras que SIP codifica sus mensajes como texto similar a HTTP. H.323 no es muy escalable ya que fue diseñado para usarse en una LAN y así aparecen problemas al escalar aunque en versiones nuevas se sugieran técnicas para resolverlos. H.323 es también limitada en la detección del loop en búsquedas complejas en múltiples dominios. Esto puede hacerse en forma alternativa grabando los mensajes pero esta técnica no es muy escalable. De otro lado, SIP usa un método de detección de loop revisando la historia de los mensajes en los campos cabeceras. La ventaja de SIP es que está respaldada por IETF, uno de las organizaciones de estándares más importantes mientras que H.323 tiene una gran parte del mercado copado.

La Tabla 1 presenta las diferencias entre ambos estándares:

H.323	SIP
Protocolo complejo	Comparativamente mas simple
Representación binaria de sus mensajes	Representación textual
Requiere compatibilidad hacia atrás	No requiere compatibilidad hacia atrás
No es muy modular	Es muy modular
No es muy escalable	Altamente escalable
Señalización compleja	Señalización simple
Gran porción del mercado	Respaldado por el IETF
Cientos de elementos	Sólo 37 cabeceras
Detección de loop es difícil	Detección de loop es más fácil

Tabla 1. Comparación entre H.323 y SIP

4. Protocolos de Soporte

SIP trabaja en conjunto con RSVP (Resource Reservation Protocol), RTP/RTCP (Real time Transport Protocol), RTSP (Real time Streaming Protocol), SAP (Session Announcement Protocol) y SDP (Session Description Protocol). RTP/RTCP es usado para transporte de datos en tiempo real, RSVP para reservación de recursos, RTSP para entregas controladas de flujos, SAP para sesiones de anuncio multimedia y SDP para describir sesiones multimedia. H.323 también trabaja en conjunto con RTP y RTCP (Real-time Control Protocol). Los actuales gateways de voz usualmente están compuestos de dos partes: el gateway de señalización y el gateway de medios. El gateway de señalización se comunica con el gateway de medios usando MGCP (Media Gateway Access Protocol). MGCP puede interoperar tanto con SIP como con H.323. La Figura 4 muestra los protocolos de señalización y transporte requeridos para entrega de voz sobre IP:

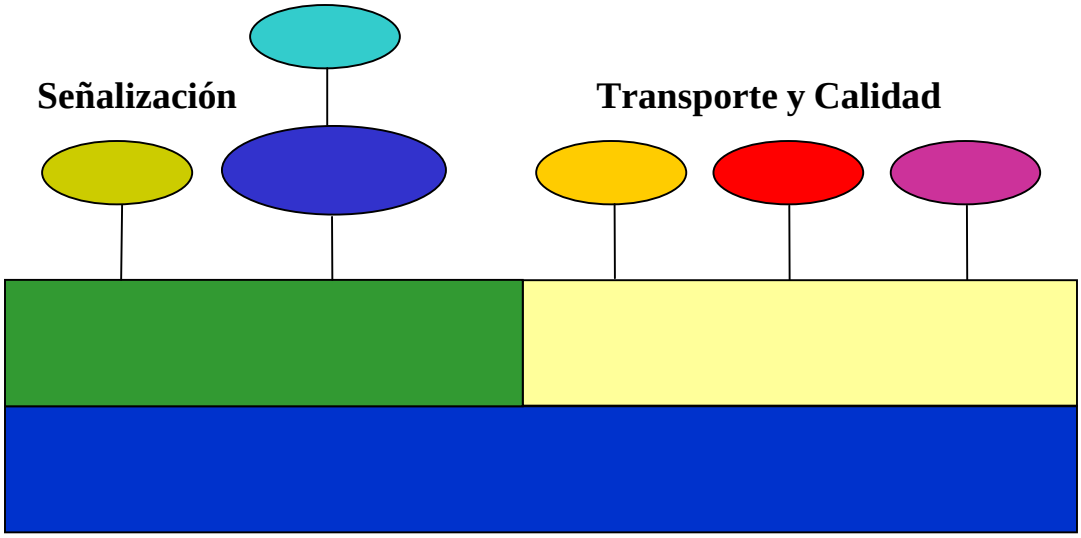

Figura 4. Protocolos de Señalización SIP y H.323 con algunos de sus protocolos de soporte

4.1 Media Gateway Control Protocol (MGCP)

Este protocolo define la comunicación entre elementos de control de llamada (Agentes de Control) y gateways de telefonía. Los agentes de llamadas son también conocidos como Controladores de Gateway de Medios. Este es un protocolo de control que permite a un coordinador central monitorear los eventos en teléfonos y gateways IP e instruirlos para enviar los medios a específicas direcciones. Este es el resultado de la unión de Simple Gateway Control Protocol y Internet Protocol Device Control. La inteligencia del control de llamada está localizada fuera de los gateways y es manejada por elementos externos de control de llamada, el Agente de llamada. MGCP asume que esos elementos de control de llamada o Agentes de Llamada se sincronizarán entre sí para enviar comandos coherentes a los gateways que estén bajo su control. Este es un protocolo maestro / esclavo, donde se espera que los gateways ejecuten comandos enviados por los Agentes de Llamadas. Este ha introducido los conceptos de conexiones y terminales para establecer caminos de voz entre dos participantes, y los conceptos de eventos y señales para establecer y concluir llamadas. Dado que el principal énfasis de MGCP es la simplicidad y confiabilidad, permite que las dificultades de programación sean concentradas en Agentes de Llamadas, así se habilitará a los proveedores de servicios a desarrollar sistemas de acceso local confiables y baratos.

4.1.1 Terminales y Conexiones. Los terminales son las fuentes de los datos. Un ejemplo puede ser una interfaz en un gateway que termina una conexión principal conectada a un switch PSTN. Las conexiones pueden ser punto a punto o multipunto. Una conexión puede ser una asociación entre dos terminales

(punto a punto) o una asociación entre múltiples terminales (multipunto). Una vez que la asociación está establecida, la transferencia de datos puede tomar lugar. Las conexiones pueden ser establecidas sobre un número de redes base como TCP/IP, ATM etc.

4.1.2 Eventos y Señales. Un agente de llamada puede requerir ser notificado sobre ciertos eventos que ocurren en un terminal, como *descolgado, colgado o dígitos marcándose* y pueden solicitar que una cierta señal sea aplicada a un terminal como tono de marcado, tono de ocupado o timbrado. Eventos y señales son agrupados en paquetes que son soportados por un tipo particular de terminal, por ejemplo un paquete puede soportar a cierto grupo de eventos y señales para líneas de acceso analógicas.

4.1.3 Creación de Conexiones. Las conexiones son creadas en el agente de llamada en cada terminal que estará involucrado en la llamada. Cuando dos terminales son localizados en gateways que son administrados por el mismo agente de llamada, la creación es hecha por medio de los tres siguientes pasos:

- El Agente de llamada pide al primer gateway crear una conexión en el primer terminal. La respuesta enviada por el gateway incluye una descripción de sesión que contiene información pertinente requerida por terceras partes para poder enviar paquetes a la nueva conexión que ha sido creada.

- El Agente de llamada entonces envía la descripción de sesión del primer gateway al segundo gateway y le pide crear una conexión en el segundo terminal. El segundo gateway responde enviando su propia descripción de sesión.

- El Agente de llamada usa un comando de modificación de conexión para proveer esta segunda descripción de sesión al primer terminal. Ahora la comunicación puede ocurrir en ambas direcciones.

4.1.4 Comandos. El MGCP implementa la interfaz de control del gateway de medios como un juego de transacciones. Las transacciones son compuestas de un comando y una respuesta obligatoria. Hay ocho tipos de comandos:

- *CreateConnection*. Este comando es usado para atachar un terminal a una dirección y puerto específicos IP. Para crear una conexión, la solicitud CreateConnection es requerida también por el terminal remoto. Si la solicitud es satisfactoriamente reconocida por el gateway, un ConnectionId es retornado el que será usado para identificar unívocamente a la conexión.

- *ModifyConnection*. Este comando es usado por el agente de llamada para modificar los parámetros de una conexión activa. El ConnectionId es pasado para identificar la conexión.

- *DeleteConnection*. Este comando es usado ya sea por un agente de llamada o por el gateway para eliminar una conexión existente. La respuesta incluye una lista de parámetros sobre el estado de la conexión.

- *NotificationRequest*. Si un agente de llamada desea ser informado sobre la ocurrencia de eventos especificados en un terminal, entonces éste puede enviar su requerimiento al gateway. Los eventos pueden ser: transición de descolgado, flash-hook, detección de tono de continuidad, etc. Una notificación puede ser requerida por un evento de detección de tono de continuidad en el gateway.

- *Notify*. La respuesta a NotificationRequest es enviada por el gateway por medio del comando Notify. El comando de notificación incluye una lista de eventos que el gateway observará.

- *AuditEndpoint*. Este comando es usado por el agente de llamada para conseguir los detalles sobre el estado de un terminal o varios terminales y la respuesta del gateway que contiene la información requerida.

- *AuditConnection*. Para obtener información para una conexión específica de un terminal, el agente de llamada usa este comando. La conexión es identificada por el ConnectionId y la respuesta del gateway que contiene la información requerida.

- *RestartInProgress*. Este comando es usado por el gateway para indicar que un terminal o un conjunto de terminales han entrado o salido de servicio. Este comando también incluye un parámetro que indica el tipo de reinicio (natural, forzado o retardado).

4.2 RTP and RTCP (Real-time Transport Protocol and Real-time Control Protocol)

RTP soporta la transferencia de medios en tiempo real (audio y video) sobre redes de conmutación de paquetes. Este protocolo es usado por SIP y H.323. El protocolo de transporte debe permitir al receptor detectar pérdidas en paquetes y también proveer información de tiempos tal que el receptor puede correctamente compensar el retardo de jitter. La cabecera RTP contiene información que usa el receptor para reconstruir los medios y también contiene información que especifica como flujos de bits codificados son divididos en paquetes. RTP no reserva recursos en la red sino que provee información de tal modo que el receptor pueda recuperar en presencia de pérdidas y de retardo por jitter.

4.2.1 Funciones de RTP. Las funciones provistas por RTP incluyen:

- *Numeración Secuencial*. El número de secuencia en el paquete RTP es usado para detección de paquetes perdidos.

- *Identificación de carga*. En Internet, es frecuentemente requerido para cambiar la codificación de los medios en forma dinámica para ajustar la

disponibilidad del ancho de banda. Para proveer esta funcionalidad, un identificador de carga es incluido en cada paquete RTP para describir la codificación del medio transmitido.

- *Identificación de marco*. Video y audio son ennviados en unidades lógicas llamadas marcos. Para indicar el principio y final del marco, un bit indicador del marco ha sido provisto.

- *Identificación de fuente*. En una conferencia se tienen varios participantes. Así un identificador es requerido para determinar el originador del marco. Para esto el identificador de sincronización de fuente (SSRC) ha sido provisto.

- Sincronización intramedios: Para compensar los diferentes retardos jitter de los paquetes dentro del mismo flujo, RTP provee marcas de tiempo o *timestamps* las que son necesitadas para ejecutar los paquetes en los buffers.

4.2.2 Servicios adicionales de RTCP. RTCP es un protocolo de control y trabaja en conjunto con RTP. En una sesión RTP, los participantes periódicamente envían paquetes RTCP para obtener información útil sobre QoS, etc. Los servicios adicionales que RTCP provee a los participantes son:

- *Información de retroalimentación de QoS*. RTCP es usado para reportar la calidad de servicio. La información provista incluye un número de paquetes perdidos, tiempo de ida y vuelta, y jitter. Esta información es utilizada por las fuentes para ajustar sus tasas de datos.

- *Control de Sesión*. Para el uso de paquetes BYE, RTCP permite a los participantes indicar que ellos están dejando la sesión.

- *Identificación*. Información como dirección de correo electrónico, nombre y número de teléfono es incluida en los paquetes RTCP de tal modo que los usuarios pueden conocer la identidad de los otros usuarios para esa sesión.

- *Sincronización entre medios*. Aún así el video y el audio son normalmente enviados sobre diferentes flujos, se necesita sincronizarlos en el receptor de tal modo que se puedan ejecutar juntos con coherencia. RTCP provee la información que es requerida para sincronización de flujos.

4.3 Real-Time Streaming Protocol (RTSP)

RTSP, protocolo de flujos en tiempo real, es un protocolo cliente/servidor que provee control sobre la entrega de flujos de medios en tiempo real. Provee funcionalidades para flujos de audio y video como pausa, adelanto, retroceso y posicionamiento deseado. Provee los medios para escoger los canales de distribución (como UDP y TCP), y mecanismos de distribución basados en RTP. RTSP establece y controla flujos continuos de audio y video entre los servidores de medios y los clientes. Un servidor de medios provee servicios de

ejecución y grabación de los flujos de medios mientras que un cliente requiere datos continuos de audio o video desde el servidor de medios. RTSP actúa como el *'control remoto de red'* entre el servidor y el cliente.

4.3.1 Funciones de RTSP. Soporta las siguientes operaciones:

- *Obtención de medios desde el servidor de medios*. El cliente puede requerir una descripción de presentación, y pedir al servidor que establezca una sesión para enviar los datos requeridos. El servidor puede enviar la presentación a una conferencia o solamente al cliente que lo solicita.

- *Invitación de un servidor de medios a una conferencia*. El servidor de medios puede ser invitado a la conferencia para ejecutar medios o grabar una presentación.

- *Adición de medios de una presentación existente*. El servidor o el cliente pueden notificarse entre sí sobre los medios adicionales que estén disponibles.

4.3.2 Características de RTSP. Entre las características de RTSP se mencionan las siguientes:

- RTSP es un protocolo de nivel de aplicación con sintaxis y operaciones similares a HTTP, pero trabaja con audio y video. Usa URLs como los existentes en HTTP.

- Un servidor RTSP necesita mantener estados, usando SETUP, TEARDOWN y otros métodos.

- A diferencia de HTTP, en RTSP tanto servidores como clientes pueden ser requeridos.

- RTSP es implementado en múltiples plataformas de sistemas operativos y permite interoperar entre clientes y servidores de diferentes fabricantes.

4.4 Resource Reservation Protocol (RSVP)

El retardo de red y la Calidad de Servicio QoS son los factores más críticos en la convergencia de voz y datos. La solución más prometedora a este problema ha sido desarrollada por el IETF RSVP. RSVP puede priorizar y garantizar latencia para flujos de tráfico IP específicos. RSVP habilita una red conmutada de paquetes para emular a una red mas determinística como la red de circuitos conmutados.

Con el advenimiento de RSVP, VoIP se hace realidad hoy en día. Con RSVP, habilitado, podemos realizar comunicaciones de voz con retardo tolerable en una red de datos. Las solicitudes RSVP generalmente resultarán en recursos siendo reservados en cada nodo a lo largo de la ruta de datos. RSVP solicita recursos en, solamente, una dirección, por lo tanto trata a un emisor como distinto lógicamente que al receptor, aunque el mismo proceso de aplicación puede actuar tanto como emisor como receptor al mismo tiempo. RSVP no es en sí un protocolo de enrutamiento, está diseñado para operar con protocolos actuales y futuros

unicast y *multicast*. Para acomodar eficientemente requerimientos de grandes grupos, membresías grupales dinámicas y requerimientos diversos de los receptores, RSVP hace a los receptores responsables de solicitar una QoS específica. La QoS solicitada por una aplicación del host receptor es pasada al proceso RSVP local. El protocolo RSVP entonces lleva la solicitud a todos los nodos a lo largo del camino inverso de los datos hasta la fuente de datos. RSVP tiene los siguientes atributos:

- Está orientado a los receptores
- Soporta comunicaciones unicast y multicast
- Mantiene el estado en routers y hosts, proporcionando soporte natural para cambios de membresía dinámicos.
- Provee una operación transparente a través de los routers que no lo soportan.

4.5 Session Description Protocol (SDP)

SDP está destinado a describir las sesiones multimedios para el propósito de anuncio de la sesión, invitación de la sesión, etc. El propósito de SDP es transportar información sobre los flujos de medios en sesiones multimedios para permitir que los recipientes de una descripción de sesión participen en la sesión. SDP incluye la siguiente información:

- Nombre y propósito de la sesión
- Dirección y número de puerto
- Tiempos de inicio y parada
- Información para recibir dichos medios
- Información sobre el ancho de banda a ser usado por la conferencia.
- Información de Contacto de la persona responsable de la sesión.

Esta información es transportada en formato de texto simple. Cuando una llamada es establecida usando SIP, el mensaje INVITE contiene un cuerpo SDP describiendo los parámetros de la sesión aceptables para el que llama. La respuesta desde el llamado incluye un cuerpo SDP que describe las propias capacidades del llamado. En general, SDP debe transportar suficiente información para habilitar una sesión y anunciar los recursos a ser usados que los que no participan deben conocer. La información de medios que SDP envía son: tipo de medio (audio o video), protocolo de transporte (RTP, UDP, etc) y formato del medio (video MPEG, video H.263, etc).

4.6 Session Announcement Protocol (SAP)

Este protocolo es usado para hacer conocidas las conferencias y otras sesiones multicast. El anunciador SAP periódicamente envía un paquete de anuncio a una dirección y puerto multicast conocidos (número de puerto 9875).

Un oyente SAP se entera del alcance multicast usando el protocolo Multicast Scope Zone Announcement Protocol ya que oye en la dirección y puerto conocidos el alcance determinado. No hay mecanismo de reunión - El anunciador SAP no está al tanto de la presencia o ausencia de alguno de los oyentes SAP. Un anuncio SAP es difundido con el mismo alcance que la sesión que se está anunciando, asegurando que los recipientes del anuncio pueden ser también recipientes potenciales de la sesión siendo difundida.

Si una sesión usa direcciones en rangos de alcance administrativos múltiples, es necesario para el anunciador enviar idénticas copias del anuncio a cada rango del alcance administrativo.
Múltiples anunciadores pueden anunciar una sesión simple, como una ayuda a la robustez frente a la pérdida de paquetes y falla de uno o mas anunciadores. El período de tiempo entre repeticiones de un anuncio es escogido tal que el ancho de banda total usado por todo los anuncios en un simple grupo SAP permanece bajo un límite previamente configurado. Cada anunciador debe escuchar los otros anuncios para determinar el número total de sesiones siendo anunciadas en un grupo particular. SAP está dirigido a anunciar la existencia de sesiones de invitación de área amplia (*wide-area-multicast*) de larga vida e involucra un gran retardo de inicio antes que un juego completo de anuncios sea escuchado por un oyente. Para reducir los retardos inherentes en SAP, se recomienda implementar el proceso proxy-caches. Un proxy SAP debe escuchar a todos los grupos SAP en su alcance y mantener una lista actualizada de todas las sesiones anunciadas junto con el tiempo en que cada anuncio fue recibido. SAP también contiene mecanismos para asegurar la integridad de anuncio de sesiones, para autenticación del emisor de un anuncio y para encriptación de tales anuncios.

5. Resumen de los Estándares y Protocolos de VoIP

En este artículo, se han presentado los protocolos de señalización H.323 (estándar ITU-T) y SIP (estándar IETF). Se compararon ambos protocolos haciéndose notar que aunque H.323 tiene mayor porción del mercado actualmente, SIP es un mejor protocolo debido a su simplicidad y escalabilidad. También se presentó el protocolo MGCP, el que es un protocolo gateway por el que el Agente de Llamada controla al gateway de señalización. Tanto H.323 como SIP necesitan algunos protocolos de tiempo real que llevan a cabo el verdadero transporte de la voz y video. RTP y RTCP se usan para el transporte y control en tiempo real. RTSP es usado para proveer entrega controlada de flujos de medios. También se revisan algunos protocolos que son requeridos en conjunto con SIP así como dar a conocer la sesión (SAP) y dar una descripción de la sesión (SDP). RSVP es usado para reservar recursos en la red y por lo tanto proveer alguna Calidad de Servicio QoS. En la Tabla 2 se resumen los protocolos y estándares revisados:

Protocolo	Descripción
H.323 (estándar ITU-T)	Protocolo principal que provee interoperabilidad
H.225	Provee señalización y registro de llamada.
H.245	Negocia el uso de los canales de medios.
SIP (estándar IETF)	Protocolo que provee Voz sobre IP
MGCP	Protocolo gateway que define la comunicación entre el agente de llamada y el gateway de señalización
RTP	Provee transporte en tiempo real sobre redes de paquetes conmutados
RTCP	Protocolo de control que provee retroalimentación a la aplicación
RSVP	Responsable de proveer QoS reservando recursos
RTSP	Provee control en la distribución de flujos de medios en tiempo real
SDP	Describe la sesion multimedios
SAP	Publica las conferencias/sesiones multicast

Tabla 2. Funciones de los principales protocolos y estándares

6. Referencias

- The Internet Telephony – Jiri Kuthan.
- Voice Over IP: Protocols and Standards – Rakesh Arora.
- Información diversa e Internet.