Servicio de Mensajería Unificada
y
Comunicaciones Multimedia en Tiempo Real sobre Redes IP

por

Macia Nicolás

Rodríguez Christian Adrián

Objetivo

Realizar la integración de los servicios propios de las redes de comunicación convencionales sobre una infraestructura IP utilizando herramientas open source. Para ello se implementará un servicio de mensajería que contemple las comunicaciones multimedia entre máquinas conectadas a la red de datos como también llamadas telefónicas hacia y desde la PSTN.

INDICE

Prefacio

Antes de empezar a desarrollar la tesis, teníamos algunos conocimientos sobre telefonía IP, también denominado VoIP. Analizamos el protocolo de señalización H.323, que por aquel entonces marcaba la tendencia, y el protocolo RTP que permite el intercambio de datos multimedia. De este modo, realizamos algunas pruebas no solo de telefonía IP sino también de integración con la red telefónica.

Al momento de empezar la tesis de grado apuntamos a determinar la mejor manera de implantar telefonía IP y cómo integrarla con los servicios de comunicación provistos por Internet como son el correo electrónico y la mensajería instantánea.

A lo largo de la etapa de recopilación de información notamos la gran cantidad de extensiones que tenía el trabajo encarado, entre las que podemos mencionar:

- Desempeño y calidad de servicio en VoIP.
- Protocolos de telefonía IP y su integración con centrales telefónicas convencionales.
- Portal de Mensajería unificada integrando en una única interfaz de usuario todos los servicios propios de Internet incluidas las comunicaciones de VoIP.

Organización del Material

El material está compuesto por 10 capítulos agrupados en tres partes: la primera introduce los conceptos de la PSTN, la segunda analiza las tecnologías VoIP y la tercera muestra las pruebas realizadas.

La Parte I, compuesta de los Capítulo 1 y 2, introduce al lector en los orígenes y funcionamiento de la PSTN, como así también se muestran algunos protocolos propios de esta red.

La segunda parte, compuesta por los capítulos 3 al 8 dan una visión detallada de los diferentes protocolos VoIP. Para ello el Capítulo 3 introduce al lector con el protocolo de señalización H.323. Como alternativa, el Capítulo 2, presenta el protocolo de señalización SIP. El siguiente capítulo analiza las similitudes y diferencias entre SIP y H.323. La codificación y decodificación de los datos multimedia, así como las alternativas de compresión, son analizadas en el Capítulo 6. Luego el transporte de los datos codificados, es responsabilidad del protocolo RTP cuyas características se nombran en el Capítulo 7. Finalmente el Capítulo 8 cierra la Parte II, con el análisis de los Gateways cuya tarea es la de integrar dos redes de comunicaciones diferentes.

La Parte III, describe las pruebas efectuadas sobre los diferentes ambientes VoIP. Esta parte comprende el Capítulo 9, donde se evalúan productos que utilizan SIP, el Capítulo 10 aquellos que usan H.323. Por último en el Capítulo 11 se da una solución a la integración de las tecnologías presentadas.

Finalmente, en una conclusión del trabajo, se consideran los puntos abordados a lo largo de la tesis, explicando por qué hoy por hoy la integración entre Internet y la PSTN no es completa.

Siglas, acrónimos y referencias

Las palabras clave, siglas y acrónimos que se sucedan a lo largo del texto serán referenciadas en el apartado llamado glosario. En este apartado dará una breve descripción y en algunos casos indicará referencias externas.

Introducción

El nacimiento de la telefonía se remonta a fines del siglo XIX. Esta fue concebida para proveer comunicaciones de voz a las personas. Desde entonces ha sufrido diversos cambios tecnológicos pero su filosofía se ha mantenido intacta. Actualmente es la red con mayor cantidad de usuarios.

Por otro lado, las redes digitales de datos fueron concebidas para proporcionar comunicaciones de datos, audio y/o video a las personas. Desde entonces, las distintas redes digitales que surgieron se fueron integrando a la hoy conocida red de redes: Internet; la cual se caracteriza por ser una red de mejor esfuerzo que utiliza TCP/IP como pilar de su arquitectura.

Entre los servicios de comunicaciones propios de las redes telefónicas, podemos mencionar: llamadas telefónicas, contestador automático, llamada en espera, etc. Estos servicios se restringen a los provistos por las empresas a sus usuarios y se caracterizan por su alta disponibilidad y robustez.

Entre los servicios propios de Internet, podemos mencionar: correo electrónico, mensajería instantánea, chat, etc. Estos se crearon a fines del siglo pasado y marcaron un cambio radical en la forma de comunicación de las personas.

Cabe mencionar que estos servicios están íntimamente ligados a la red sobre la que funcionan, dado que cada una de estas redes fue concebida con distintos propósitos e incluso con 100 años de diferencia. Por esto no es de esperar la misma calidad, disponibilidad y/o confiabilidad en sus servicios. De esta forma, ambas redes dieron lugar a dos universos diferentes sobre los que surgieron distintos servicios de comunicaciones.

Antes del apogeo de lo que hoy en día se conoce como la “red de redes”, o Internet, los usuarios utilizaban las redes de datos como medio de comunicación para compartir archivos en redes de corto alcance llamadas redes de área local. Sin embargo, estas primeras redes eran propietarias por lo que los problemas de compatibilidad no permitían interconectar redes de diferentes proveedores.

Con el correr del tiempo y ante la necesidad de integrar redes propietarias, organizaciones de estandarización como son ISO, EIA, IEEE, IETF, W3C, entre otras, se encargaron de definir estándares que permitan la interoperabilidad entre los distintos fabricantes de diferentes tecnologías.

ARPANET, el predecesor de Internet, era por aquel entonces una red creada con fines militares por el Departamento de Defensa de Estados Unidos. Finalmente se

conectó a cientos de Universidades y dependencias de Gobierno a través de líneas telefónicas rentadas, pero cuando se añadieron redes satelitales y de radio se generaron incompatibilidades entre los sistemas, de modo que surgió un nuevo modelo de referencia, el modelo TCP/IP, brindando a sistemas de diferentes arquitecturas la capacidad de conectarse entre sí.

El Modelo TCP/IP es el utilizado hoy en día en Internet y dio lugar a servicios como correo electrónico, web, ftp, irc, etc. A partir de este momento el término comunicación tomó un nuevo rumbo ya que esta nueva red llamada Internet, creció hasta convertirse en un éxito comercial con miles de millones de dólares anuales en inversiones, siendo un servicio al alcance de cualquier persona, cambiando hábitos cotidianos como por ejemplo leer un periódico de cualquier país, intercambiar opiniones, publicar trabajos, exponer datos personales, realizar compras, etc. Con estos nuevos mecanismos de comunicación, las distancias y los tiempos se redujeron al máximo.

El crecimiento de Internet fue exponencial. La integración con otras redes propietarias que operaban a su par y que obviamente carecían de la infraestructura de Internet, no tardó en llegar. Este fue y sigue siendo el caso de redes como SNA, X.25 e IPX, las cuales arrendaban un enlace dedicado para conectar dos puntos distantes y luego tuvieron la posibilidad de integrarse a Internet[◊] y utilizar el servicio de transporte que esta provee. La historia muestra que la tendencia es la integración de redes diferentes, y la telefonía no será una excepción.

Por todo esto, hoy por hoy es inconcebible considerar la implementación de un nuevo servicio de comunicaciones que no funcione en Internet. La gran cantidad de usuarios potenciales y los costos asociados justifica en gran medida tales decisiones de implementación.

La razón que motivó este trabajo fue unificar dos redes naturalmente diferentes, como son la red telefónica e Internet, considerando además la convivencia de los servicios que estas proveen.

Dentro de las posibilidades que supone semejante integración aparece un concepto muy importante que es el de Mensajería Unificada. La Mensajería Unificada es la consolidación de los distintos tipos y fuentes de comunicación. Se trata de tener todas las fuentes posibles de los mensajes en un sólo lugar, independientemente del tipo de mensajes: mensajes cortos, correo electrónico, llamadas telefónicas, correo de voz, fax, etc. Además considera una unifica interfaz de usuario accesible a través de Single Sign On.

Desde el punto de vista de la implementación, es fundamental la integración de los servicios mencionados anteriormente para el desarrollo de una arquitectura cliente/servidor de mensajería unificada. Dicha integración aun no está resuelta.

Las prestaciones del servicio de comunicaciones de voz provistas por la PSTN son excelentes, sin embargo, las comunicaciones VoIP en Internet no están a la misma altura. Esto se debe a que la PSTN fue diseñada para tal fin, alocando los recursos necesarios para garantizar una buena calidad de servicio, mientras que por su parte, Internet permite el uso de sus recursos sin discriminar el servicio que lo utiliza.

Los servicios propios de Internet nacieron en forma natural de acuerdo a la filosofía de la red. Los servicios de comunicaciones que tienen la particularidad de ser en tiempo real son los más difíciles de ser provistos, por ser servicios que necesitan ciertos parámetros de QoS como ser ancho de banda, delay y jitter. Estos servicios son propios de redes conmutadas de circuitos, como son por ejemplo ISDN. Por el

[◊] Esto es posible por medio de tuneles ipsec, ip-ip, gre, entre otros.

contrario, Internet es una red de paquetes en la que no existe ningún tipo de QoS a menos que se la implemente.

Por lo tanto, el único servicio que no es propio de Internet, el cual es uno de los más importantes a integrar en un entorno de mensajería unificada, es el de las llamadas telefónicas.

Este trabajo esta orientado a abordar la mensajería unificada en Internet, desde la problemática que supone la integración del servicio de comunicaciones de voz de la PSTN. Para ello se analizará como proveer servicios de voz en Internet con las mismas prestaciones que en la PSTN y como permitir interactuar estas dos redes en forma transparente, de modo de, por ejemplo, poder llamar desde una PC a un teléfono o viceversa.

Introducción a las comunicaciones multimedia

En la PSTN, las comunicaciones telefónicas pueden describirse fácilmente desde la perspectiva del usuario, debido a la masiva adopción de esta tecnología por parte de la sociedad contemporánea. Una gran parte de la población mundial conoce el teléfono desde su niñez, y en consecuencia está familiarizada con esta red en forma natural.

Partiendo de la premisa de que cada extremo a comunicar posee un número telefónico y un teléfono, básicamente el proceso para iniciar una llamada telefónica puede describirse de la siguiente manera: el extremo que desee iniciar la llamada debe obtener tono de marcado. Para ello simplemente debe levantar el auricular. Luego se procede con el discado del número correspondiente al destinatario, e inmediatamente se disparan dos señales con el fin de alertar tanto al destinatario de una llamada entrante, como al llamador del estado de su comunicación. Finalmente si el teléfono en el destinatario es descolgado mientras está sonando, se establece un canal de voz entre ambos extremos hasta que alguno de los dos participantes finalice la comunicación.

A pesar de ser explicativo, el párrafo anterior oculta un sin fin de operaciones que se producen internamente y en forma transparente a los usuarios. Estos procesos internos son tan complejos como necesarios para poder llevar a cabo una simple llamada telefónica, principalmente por los servicios actualmente disponibles en la mayoría de las compañías de telefonía. Los proveedores de servicios de telecomunicaciones ofrecen una variada gama de servicios entre los que se pueden nombrar movilidad, llamada en espera, retransmisión de llamadas, buzón de mensajes, selección entre diferentes compañías locales o de larga distancia, etc.

Tratar de implementar estos mismos servicios en redes de datos, particularmente en Internet, es un desafío que parece ser inminente a la vez de complejo. Al comenzar un análisis de objetivos, los primeros impedimentos que surgen son:

- Facturación: ¿cuál es el costo de una llamada? La respuesta sería por tiempo y distancia, pero ¿cómo estimo la distancia entre dos nodos de Internet? ¿Cómo calculo el tiempo de una llamada en una red de mejor esfuerzo?
- Teléfono: ¿cuál es el dispositivo que suplantaría al teléfono? Una computadora parece ser la respuesta pero no es intuitivo usar un micrófono y parlantes. Además sería menos drástico tener un dispositivo de características similares al teléfono.
- Direccionamiento: ¿Cómo funcionaría un direccionamiento telefónico sobre el modelo de direcciones IP? Un número de teléfono no es directamente asignable a una dirección IP ya que muchas redes utilizan DHCP como mecanismo de asignación de direcciones. El problema de un ambiente DHCP es que un “teléfono IP” no tendría la misma dirección en dos momentos diferentes, convirtiéndose en inaccesible. Problemas similares ocurren cuando se utiliza NAT.
- Calidad de Servicio: como Internet se basa en una filosofía de mejor esfuerzo sin garantizar calidad de servicio, no parecería ser una red candidata para el tráfico sensible al tiempo como es el de voz.

- Duración de una llamada: este tema se relaciona con la facturación. Es necesario establecer el inicio y la finalización de las comunicaciones en Internet. Para controlar la duración debe existir alguna entidad diferente de los participantes en la comunicación que lleve un registro del tiempo transcurrido.

Los esfuerzos por proveer estos servicios en redes de paquetes, han dado lugar a estándares de señalización que en algunos casos copian características de los servicios existentes en la PSTN y en otros los extienden o introducen nuevas capacidades.

Conceptualmente, el proceso de señalización debe ser similar tanto en Internet como en la PSTN, proveyendo mecanismos para iniciar, mantener y terminar comunicaciones. Los protocolos estándar de señalización para redes de paquetes son H.323 y SIP. La forma de operar de los mismos responde a las concepciones propias de sus creadores a la hora de escribir recomendaciones, es por eso que H.323 se parece mas en su funcionamiento al de la PSTN mientras que el funcionamiento de SIP es mas parecido al de protocolos de Internet como HTTP, SMTP, etc.

Para el caso de H.323, su desarrollo está a cargo de la ITU-T, la cual es el sector de estandarización de telecomunicaciones de la ITU. El objetivo de la ITU-T es el de desarrollar recomendaciones de alta calidad que cubran todos los aspectos de las telecomunicaciones, en particular las comunicaciones en redes de telefonía. Sus publicaciones se pueden encontrar en http://www.itu.int/ITU-T/publications/recs.html, las cuales no son gratuitas.

Para el caso de SIP, el desarrollo del mismo está a cargo del grupo de trabajo MMUSIC de la IETF, la cual a diferencia de la ITU, es una comunidad abierta. Sus publicaciones se pueden encontrar en http://www.ietf.org/rfc.html, las cuales son gratuitas. Cabe destacar que la mayoría de los protocolos de comunicación en Internet son estándares abiertos y evolutivos mantenidos por esta entidad.

El funcionamiento de estos protocolos es diferente dado que la lógica empleada por cada uno es distinta, siendo incluso el direccionamiento empleado por estos también diferente. Sin embargo, aunque la competencia por ser el protocolo de facto para la señalización de VoIP seguirá existiendo, la tendencia marca una confluencia entre SIP y H.323, por medio de productos que integren ambas tecnologías.

Mencionados ya los estándares de señalización, continuamos con la siguiente fase correspondiente al intercambio de datos multimedia. Este intercambio de datos es independiente del protocolo de señalización que se utilice, puesto que es existe un estándar llamado RTP utilizado para tal fin.

A continuación se introducirá al lector en la evolución, funcionamiento y protocolos propios de la PSTN. Luego, se procederá al estudio en profundidad de las tecnologías VoIP mencionadas en párrafos anteriores y su integración con otras tecnologías.

Parte I – Red Telefónica

Capítulo 1: Red Pública de Telefonía o PSTN

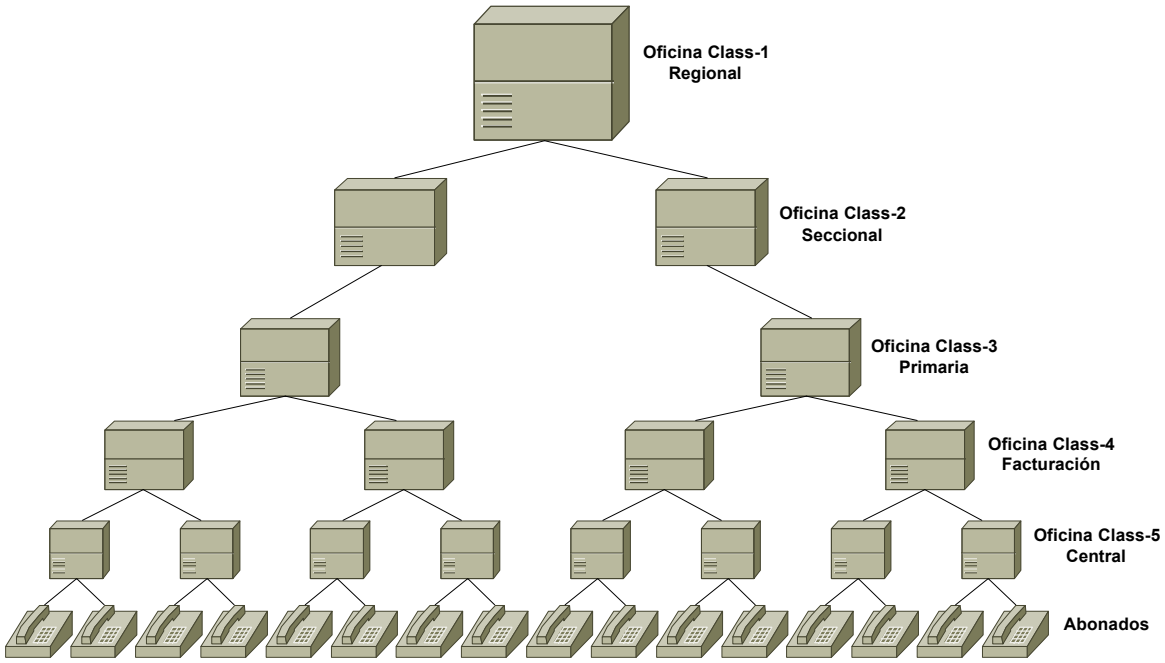
Evolución

Suponiendo la existencia de tres o cuatro teléfonos en una localidad, parece sensato conectar cada teléfono con los restantes y encontrar un método simple para comunicar un teléfono con otro. Sin embargo, si existieran tres o cuatro mil teléfonos en una localidad, el método mencionado queda obsoleto. Para este caso parecería apropiado conectar cada teléfono con alguna oficina central y realizar en ella las operaciones de conmutación. Esta tarea podría hacerse manualmente utilizando conectores y enchufes, o con dispositivos electromecánicos o electrónicos. Cualquiera sea el caso, esta solución introduce una oficina central o CO y es la que ha sido elegida por la industria de las telecomunicaciones. Una analogía que simplificaría el entendimiento del funcionamiento de una CO es una PBX, cuya función es idéntica a la de una oficina central pero en escala reducida, interconectando teléfonos de una pequeña oficina.

Al conectar estos miles de teléfonos a una CO se crea una topología en estrella en la que todas las líneas van desde un único teléfono hasta el núcleo de la estrella, es decir la CO. Estas líneas que conectan abonados con una CO, reciben el nombre de bucle local.

Del esquema de abonados conectados a una CO, surge la incógnita de cómo es posible efectuar llamadas que no terminen dentro de la cobertura correspondiente a una CO particular, es decir, cómo es posible llamar a otras ciudades, provincias o incluso a otros países. La respuesta es conectar varias COs entre sí en forma jerárquica, en donde cada una se conecta con otra de nivel superior. La oficina local, también llamada oficina final, recibe el nombre de oficina Class-5. Ésta se conecta con una oficina Class-4 y así siguiendo. Cada país posee unas pocas oficinas de nivel superior, denominadas oficinas Class-1.

Como se mencionó en el párrafo anterior, las únicas oficinas que conectan abonados son las oficinas Class-5. El resto de las oficinas en la jerarquía conectan COs de niveles inferiores como se muestra en la siguiente figura.



Las líneas que conectan oficinas de conmutación entre sí, son conocidas como trunks.

El número asociado al término Class**,** identifica el tipo de servicios de señalización que se provee como ser identificación de llamada, redirección de llamada, numeración 800, etc.

A esta red en su totalidad, es decir desde un simple abonado que posee un teléfono convencional hasta una oficina Class-1, se la denomina red pública de conmutación telefónica o PSTN.

Entonces, para permitir llamadas entre abonados, las COs completan las conexiones desde un abonado a otro, o desde un abonado a un trunk que termina en otra CO que a su vez se conecta con otro abonado. Para completar las conexiones las COs llevan a cabo tareas de conmutación. Estas tareas fueron las más afectadas durante las distintas eras que subsistió la telefonía, donde se produjeron cambios tecnológicos que van desde un operador que conmutaba conexiones en forma manual a través de enchufes, hasta la utilización de ESS de tercera generación donde las tareas de conmutación son íntegramente eléctricas.

La evolución de las tecnologías de conmutación fue acompañada de una migración masiva de las redes telefónicas, transformando las antiguas redes analógicas en líneas digitales.

Funcionamiento

Como es sabido, cuando un abonado realiza llamadas telefónicas, utiliza un dispositivo telefónico que interactúa con la PSTN. Esta interacción es necesaria y hace posible la comunicación ya que comprende todas las funciones de señalización entre el usuario y la red telefónica.

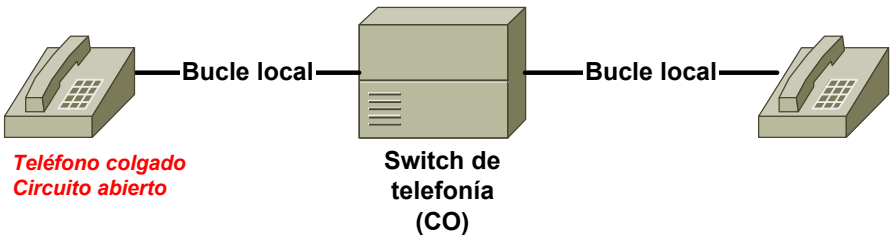
Esta señalización, que se da entre el abonado y la red, depende tanto de las características del dispositivo del abonado como de sus necesidades. Para el caso

más común, en el que el abonado posee un teléfono analógico, la señalización ocurre en la misma línea telefónica, es decir, tanto el tráfico de voz como la señalización comparten el mismo canal. Esta técnica, en la cuál las señales de control se transportan sobre el mismo canal que los datos, en este caso la voz, se denomina señalización intracanal e involucra tonos de marcado, tonos de ocupado, tonos de indicación de llamada y tonos de identificación de dígitos de marcado.

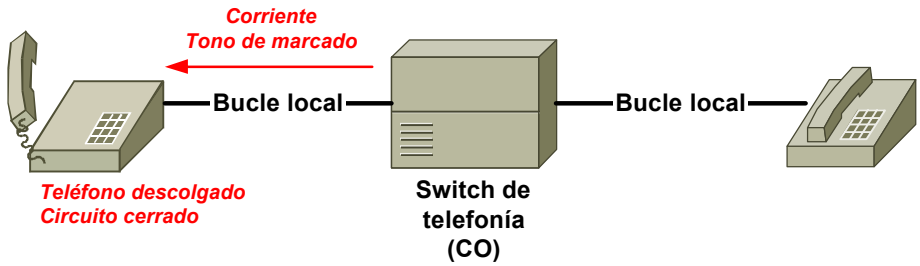
Es posible describir el proceso de señalización en la PSTN, a partir de tres procedimientos básicos, los cuales se combinan de modo de hacer posible la comunicación. Dichos procedimientos son:

- Supervisión: involucra la detección de cambios en el estado del bucle local o de un trunk. Una vez detectados estos cambios, se genera una respuesta predeterminada como por ejemplo, cerrar el circuito para conectar una llamada.
- Direccionamiento: involucra el pasaje de los dígitos marcados a una PBX o CO mediante pulsos o tonos. Los dígitos marcados proveen al conmutador el camino de conexiones al teléfono llamado.
- Indicación: permite la generación de tonos audibles que indican ciertas condiciones como ser llamadas entrantes o teléfono ocupado.

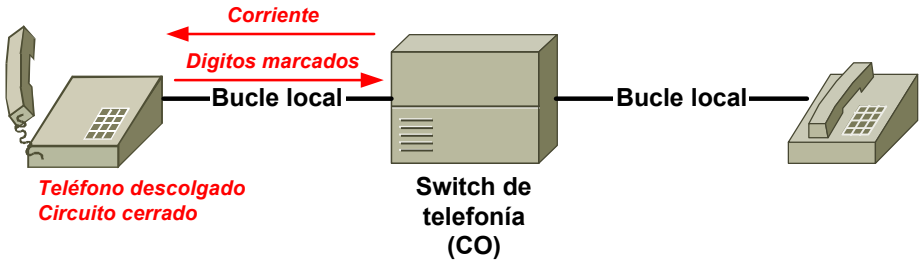
El progreso de una llamada telefónica puede encontrarse en uno de cinco estados: colgado, descolgado, marcando, conmutando, sonando y en curso.



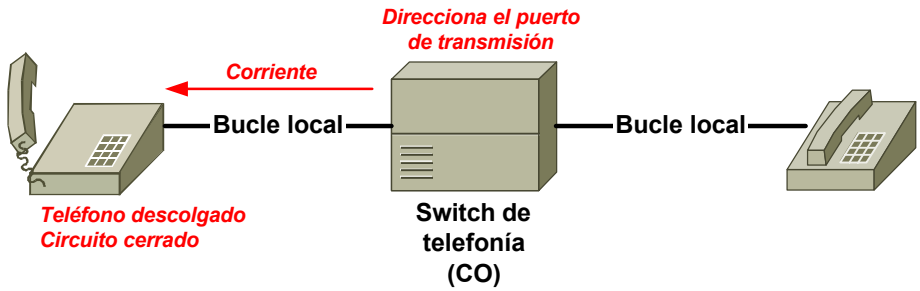
Antes de iniciar una llamada, el teléfono se encuentra en condición de listo esperando que alguien levante el auricular, es decir se encuentra en el estado colgado. Durante este momento el circuito desde el teléfono hasta el conmutador de telefonía en la CO se encuentra abierto. El conmutador mantiene la fuente de alimentación para este circuito previniéndose la pérdida del servicio telefónico por cortes eléctricos en la ubicación del teléfono.



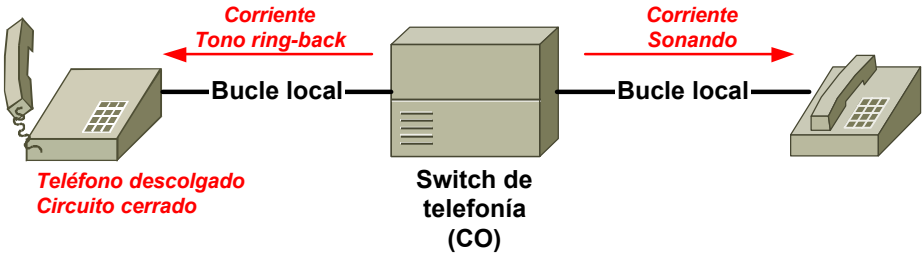
Cuando alguien desea iniciar una llamada, levanta el auricular del teléfono. Este hecho, cierra el circuito entre el conmutador en la CO y el teléfono permitiendo que la corriente circule. Cuando el conmutador detecta este flujo de corriente, transmite un tono de marcado al teléfono. Dicho tono indica al cliente que puede iniciar el marcado de los dígitos. No es posible garantizar que el cliente siempre obtendrá tono de marcado ya que puede suceder que todos los circuitos estén siendo utilizados. La CO sólo entregará tono después de haber reservado espacio para almacenar la dirección entrante.



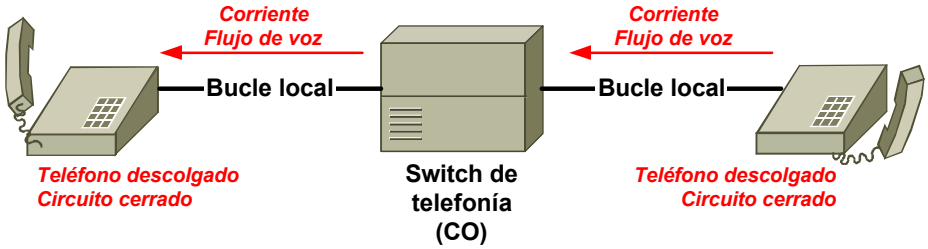
La fase de marcado permite al cliente ingresar el número telefónico o dirección perteneciente a un teléfono en otra ubicación. El marcado puede realizarse tanto por pulsos o tonos que son transmitidos al conmutador.



En la fase de conmutación, el conmutador traduce los pulsos o tonos recibidos en una dirección de puerto que se conecte con el teléfono de la entidad llamada. Esta conexión puede ir directamente al bucle local del teléfono requerido, para llamadas locales, o atravesar otros conmutadores antes de alcanzar el destino final si es el caso de una llamada de larga distancia.



Una vez que el conmutador se conecta con la línea llamada, envía una señal a la misma haciendo sonar el teléfono destino. Mientras el teléfono esté sonando, un tono de respuesta vuelve al teléfono que inició la llamada indicando que el teléfono destino está sonando. En caso que el destino esté ocupado, vuelve la señal de ocupado.

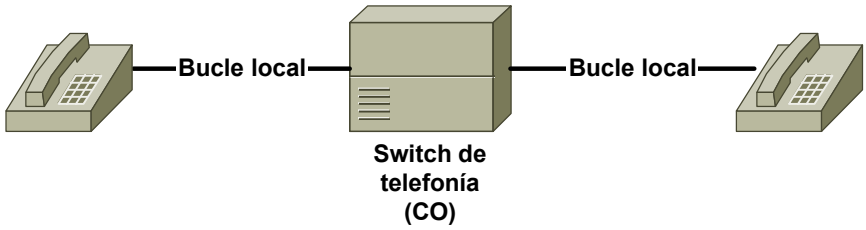


La última fase comienza cuando el cliente que es llamado levanta el auricular del teléfono que se encuentra sonando. Tan pronto como el destino levanta el auricular, nuevamente se comienza la fase de descuelgue. De esta forma se cierra circuito del lado del destino comenzando a fluir corriente hacia el conmutador. Cuando el mismo detecta este flujo de corriente, completa la conexión de voz hacia el teléfono que inició la llamada. A partir de este momento ambos extremos pueden hablar.

Capítulo 2: Protocolos propios de la PSTN

En las redes de conmutación de circuitos, las señales de control se utilizan tanto para gestionar la red como las llamadas telefónicas. Estas señales de control corresponden al intercambio de información entre el abonado y los conmutadores, entre los conmutadores entre sí, y entre los conmutadores y los puntos de control de servicio, los cuales proveen acceso a bases de datos propias de las compañías telefónicas. Ejemplos de estos flujos de control son los requerimientos de abonados, la conmutación de dichos requerimientos, y consultas a los puntos de control, ya sea para conmutar llamadas a servicios especiales como un 0-800, validar tarjetas telefónicas prepagas, etc.

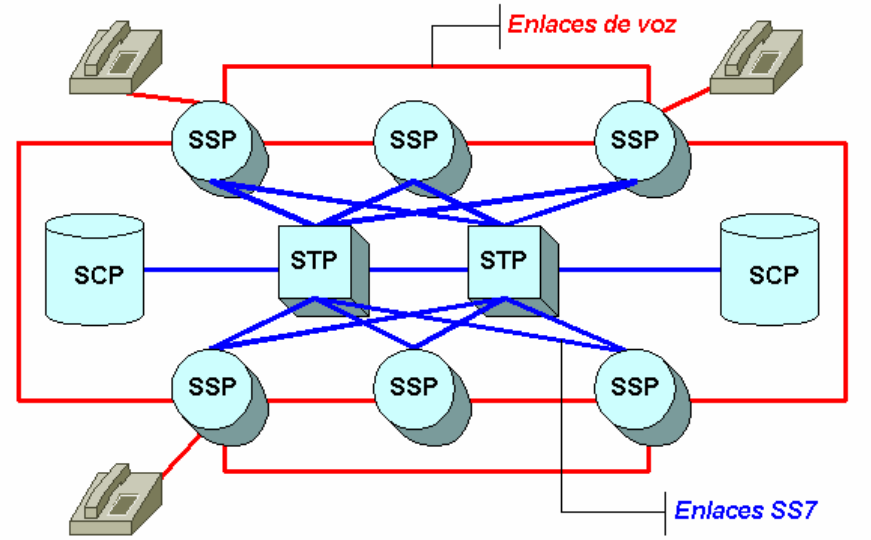
Como vimos anteriormente, la señalización que se da entre el abonado y la red se denomin*a señalización intracanal,* dado que la voz y la señalización comparten el mismo canal.



En contraposición, la señalización que ocurre dentro de la red, es decir, entre los conmutadores, se la conoce como señalización por canal común. En esta técnica, las señales de control se transmiten por rutas completamente independientes de los canales de voz, pudiendo una misma ruta transportar señales de control para varios canales de voz. SS7 es el esquema de señalización por canal común más usado. Este estándar, es una recomendación desarrollada por la ITU-T, diseñada para uso especifico en redes ISDN. Sin embargo, es una norma abierta de señalización que puede utilizarse en cualquier red de conmutación de circuitos digitales.

Los mensajes de control SS7, son paquetes pequeños que se encaminan a través de la red, de modo que aunque la red que está siendo controlada sea una red de conmutación de circuitos, la señalización de control se basa en una tecnología de conmutación de paquetes.

Es importante distinguir la función de señalización de la función de transferencia de información dado que la primera esta basada en la conmutación de paquetes mientras que la segunda está basada en la conmutación de circuitos. Esta distinción permite ver la red desde dos puntos de vista diferentes, los cuales distinguen dos planos, cada uno con su propio ámbito y función, a los que llamaremos plano de control y plano de información.



En el dibujo anterior, las líneas rojas identifican el plano de información, mientras que las azules el plano de control SS7. La entidades presentadas en dicho dibujo se describen mas adelante.

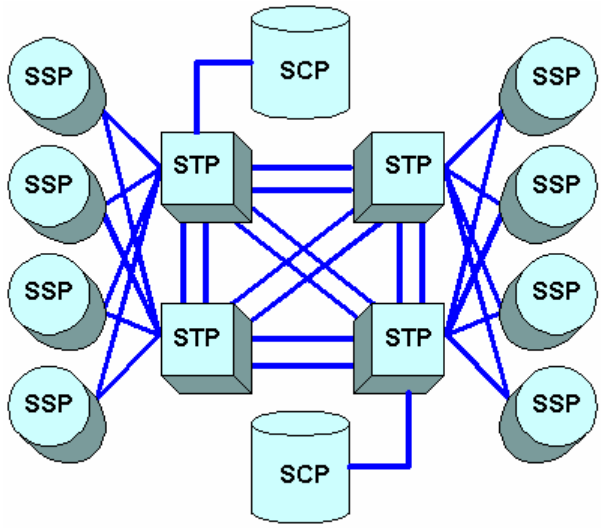
SS7 ofrece las siguientes capacidades de señalización:

- Establecimiento y terminación de llamadas entre abonados del servicio telefónico
- Acceso a sistemas back-office para obtener información de los abonados directa o indirectamente relacionada con una llamada telefónica.
- Acceso a sistemas back-office para registrar información de facturación de los clientes.
- Acceso a sistemas back-office para la traducción de números de tarifa gratis, como los 0-800, tarifa especial, como por ejemplo 0-810 o 0-610, y otras capacidades avanzadas propias de las redes telefónicas.
- Servicios de movilidad, como autenticación y roaming.
- Servicio de portabilidad numérica local, permitiendo a los abonados seleccionar entre diferentes compañías de telefonía manteniendo el mismo número telefónico.
- Acceso a servicios avanzados como redirección de llamadas, conferencias tripartitas e identificación de llamada por número y/o nombre.

Para proveer los servicios mencionados, las redes SS7 se componen de nodos llamados puntos de señalización, los cuales son identificados en forma única por un valor numérico llamado *código del punto*. Estos identificadores se utilizan en los mensajes transmitidos entre los puntos de señalización para identificar al origen y al destino de los mismos.

Los puntos de señalización se clasifican en:

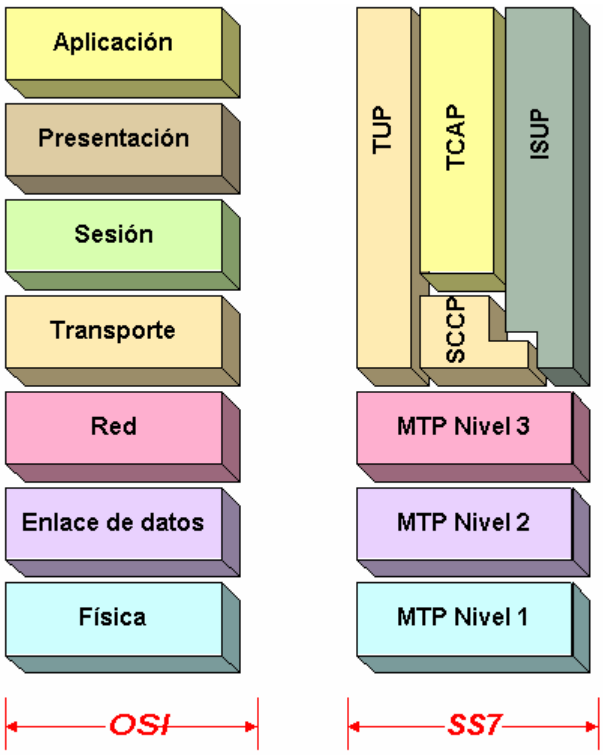
- SSP: puntos de conmutación de servicio
- STP: puntos de transferencia de señal
- SCP: puntos de control de servicio



Un SSP es un conmutador local que se conecta con los abonados de la forma descripta en la señalización intracanal y con otros conmutadores enviando y recibiendo mensajes de control SS7. Además, un SSP puede consultar una base de datos centralizada o SCP para determinar cómo llevar a cabo una llamada, como es el caso de los números 0-800 o 0-610. El SCP envía una respuesta al SSP que efectuó la consulta indicando los números asociados al número recibido. Por su parte los STP ofrecen servicios de enrutamiento y transferencia de los mensajes originados en los SSP.

Una vez que se ha establecido una conexión entre dos abonados, la información se transfiere desde un usuario al otro, extremo a extremo. En el plano de información, se establece un circuito desde el conmutador local de un usuario hasta el del destinatario, habiéndose realizado el enrutamiento quizás, a través de uno o más nodos de conmutación de circuitos.

El protocolo SS7 se divide en abstracciones funcionales llamadas niveles. Estos niveles pueden asociarse con las capas del modelo de referencia OSI.



Los niveles que se muestran en la figura se encargan de las siguientes tareas:

Message Transfer Part o MTP

Se divide en tres subniveles. El nivel más bajo, nivel 1, se corresponde con la capa física del modelo OSI y define las características eléctricas, físicas y funcionales del enlace digital de señalización. Entre las interfaces físicas definidas se encuentran:

- E1: 2048 Kbps, 32 canales de 64 kbps
- DS1: 1544 kbps, 24 canales de 64 kbps
- V.35: sólo un canal de 64 kbps
- DS0: sólo un canal de 64 kbps
- DS0A: sólo un canal de 56 kbps

El nivel 2 asegura la transmisión de mensajes punto a punto a través de un enlace de señalización. Para ello implementa control de flujo, numeración de mensajes, y control de errores. Al encontrar errores, los mensajes son retransmitidos.

Finalmente el nivel 3 provee enrutamiento de mensajes entre puntos de señalización de una red SS7. Su tarea principal es modificar las rutas de los mensajes en caso de fallas en enlaces o congestión de la red conmutada.

ISDN User Part o ISUP

Define el protocolo utilizado para establecer, mantener y liberar trunks que transportan voz y datos entre puntos finales, es decir abonados. ISUP es utilizado tanto en llamadas ISDN como en llamadas que no lo son. Las llamadas que se originan y terminan en el mismo conmutador no utilizan señalización ISUP debido a que no necesitan trunks para completar sus llamadas.

Telephone User Part o TUP

En algunos países como Argentina, Brasil y China se utiliza TUP para el inicio y terminación de llamadas. TUP sólo maneja circuitos analógicos razón por la que muchos países lo han reemplazado por ISUP.

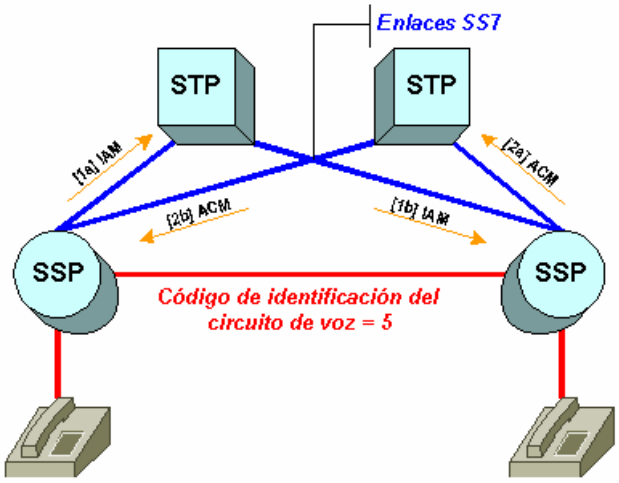
Signaling Connection Control Part o SCCP

Provee varios servicios pero principalmente se encarga del transporte de datos orientados y no orientados a la conexión. Es utilizado como capa de transporte para los servicios basados en TCAP.

Transaction Capabilities Applications Part o TCAP

Permite el intercambio de datos entre aplicaciones utilizando el servicio no orientada a la conexión provisto por SCCP. Las consultas y respuestas enviadas entre los SSP y SCP se encapsulan en mensajes TCAP.

Para comprender cómo interactúan las partes en una comunicación SS7 se presenta un ejemplo simplificado donde un abonado realiza una llamada a otro. En el ejemplo se considera que los abonados se conectan a diferentes SSP.



1. Cuando un abonado llama a otro ubicado en un conmutador diferente, el SSP origen transmite un mensaje ISUP initial address message o IAM con el fin de reservar un trunk libre desde el conmutador local al remoto. Este mensaje contiene el código del punto origen y destino, el código de identificación del circuito (circuito 5 en la figura), números marcados y opcionalmente el número y nombre de la entidad que a la que llama. [1a]
2. El IAM es enrutado a través del STP correspondiente al SSP local. [1b]
3. El conmutador en el destino examina el número marcado, presente en el mensaje IAM, determina que él sirve al destinatario final y que está disponible. Entonces el conmutador destino hace sonar la línea del abonado y transmite un mensaje ISUP address complete message o ACM al SSP origen a través del STP local. [2a] y [2b]
4. Cuando el conmutador origen recibe el ACM conecta al abonado que inició la llamada en forma directa al trunk para que la conexión se establezca punto a punto entre los abonados.

En el ejemplo presentado los SSP están conectados en forma directa a través de un trunk. Si los conmutadores origen y destino no se encontraran directamente conectados, el conmutador origen debe transmitir un IAM para reservar un circuito en el trunk hasta un conmutador intermedio. Luego el conmutador intermedio envía una aceptación ACM de la solicitud de reserva del circuito en el trunk y transmite un IAM para reservar otro circuito en el trunk hasta llegar al conmutador destino.

Conclusión

Como se mencionó en la introducción de esta sección, las comunicaciones en la PSTN pueden describirse de forma simplificada, pero por detrás existe un mundo difícil de comprender que facilita las operaciones y provee servicios enriquecedores a los abonados.

SS7 es un estándar complejo de analizar desde una perspectiva computacional como es la de nuestro alcance, es por ello que se hizo una breve introducción del protocolo dejando en claro cuales son sus funciones básicas. Para más información ver [Ref. SS7].

Parte II – Tecnologías VoIP

Introducción

La utilización de la red telefónica como medio para enviar y recibir paquetes provenientes de comunicaciones digitales no trajo complicaciones salvo por la escasez de ancho de banda. El inverso de la situación anterior, es decir transportar voz sobre una red de paquetes, siempre pareció imposible por la baja calidad de los enlaces, el excesivo tamaño de los datos multimedia, la escasez de hardware, etc.

Sin embargo, actualmente tener comunicaciones de voz, e incluso video, sobre redes IP es una realidad siempre que se respeten los límites de latencia y jitter. A pesar que estos parámetros pueden ser satisfechos por una red, sobre todo si es el caso de una LAN, existen pequeñas precauciones que pueden adoptar los extremos de la comunicación para mejorar el desempeño de la red.

La clave para enviar audio en cualquier red de paquetes es la compresión, y VoIP no es una excepción. La compresión ofrece diversas ventajas, y en particular preserva el ancho de banda requerido para la transferencia de datos.

La compresión es efectuada por agentes llamados codecs, diferenciándose esencialmente por sus capacidades de compresión y calidad. Estos agentes intervienen en ambos extremos de forma opuesta. Previo al envío de una señal analógica, el codec entra en acción transformando dicha señal en datos digitales que se encapsulan dentro de un paquete IP. Luego, el receptor procesa este paquete aplicando el codec en sentido inverso decodificando los datos digitales para así recuperar la señal analógica.

Como Internet está expuesta a la perdida de paquetes por posibles congestionamientos, los codecs están preparados para tratar estos problemas a través de algoritmos extremadamente complejos de recuperación de datos. Esto es posible, ya que a diferencia de otra información sensible a la pérdida de paquetes como son los mensajes de texto, el audio y video delegan la interpretación en la percepción humana capaz de reconstruir una palabra o una secuencia de imágenes a partir de un contexto.

En las próximas secciones se presentarán los dos estándares más usados en Internet para realizar comunicaciones de audio y video: H.323, protocolo que utiliza conceptos de su predecesor H.320 utilizado en las redes ISDN, y SIP, un protocolo moderno cuyos pilotes son otros protocolos usados en Internet. En cada uno de ellos se realizará un análisis de los mensajes involucrados en las comunicaciones, el flujo de los mismos, las entidades definidas por cada estándar, y finalmente se compararán en forma objetiva. Luego se continuará con el estudio de un tercer protocolo, utilizado por SIP y H.323 para el transporte de tráfico multimedia en tiempo real, RTP. Como se verá en las subsiguientes secciones, este estándar fue definido intencionalmente en forma aislada con el fin de ser independiente de las alternativas de señalización e incluso de su uso específico, ya que no sólo se aprovecha en comunicaciones sino también en sesiones de streaming de emisoras de radio, conferencias, etc.

Capítulo 3: Protocolo de señalización H.323

El estándar H.323 provee los fundamentos para las comunicaciones de audio, video y datos sobre redes de datos, especialmente Internet. Este protocolo es uno más de la familia H.32x definidos por la ITU-T, siendo H.323 el estándar para las comunicaciones multimedia sobre redes LAN sin proveer una calidad de servicio asegurada. Pertenecer a esta familia de protocolos convierte a H.323 en una excelente alternativa debido a la gran confiabilidad mostrada en otros protocolos de la familia, como por ejemplo es el caso de H.320, utilizado en redes ISDN. Para julio de 2003, se definió la versión 5 del estándar después de haber sufrido modificaciones desde su primer versión cuya aprobación data de octubre de 1996. Las redes de área local que hoy día dominan el mercado son TCP/IP e IPX sobre Ethernet, FastEthernet y TokenRing. Es por ello que H.323 es una opción flexible para la implementación de comunicaciones multimedia sobre las diferentes tecnologías LAN.

A continuación se enumeran las definiciones que especifica el estándar:

- Comunicaciones multimedia en tiempo real punto a punto y multipunto
- Interoperabilidad de redes
- Capacidades de clientes heterogéneos
- Codecs de audio y video
- Administración y accounting

Además de los puntos anteriores, H.323 define las entidades que intervienen en las comunicaciones de forma genérica, hecho que puede prestarse a confusión dado que en la práctica uno tiende a materializar dichas entidades en dispositivos y en realidad sucede que muchas veces son conceptos abstractos, o simplemente se convierten en un dispositivo que unifica una serie de funciones. Las componentes especificadas por el estándar corresponden a una de las siguientes clases:

- Terminales
- Gateways
- Gatekeepers
- Unidades de control multipunto o MCUs

Terminales

Una terminal puede ser una PC o un dispositivo de propósito específico, como por ejemplo un teléfono con capacidades IP que ejecute aplicaciones multimedia. Básicamente una terminal debe soportar comunicaciones de audio, pero opcionalmente puede soportar video o datos. El objetivo principal de estas entidades es interactuar con otras.

Gateway

Sin entrar en detalles, un gateway conecta dos redes diferentes. H.323 define gateways para permitir la conectividad entre una red H.323 y una red que no lo sea, como por ejemplo H.320, donde se podría conectar una terminal H.323 y un teléfono ISDN. Para conectar diferentes redes, un gateway debe traducir señales de control de llamada y realizar una conversión del formato de datos multimedia entre las diferentes redes. Además de las conversiones necesarias, los gateways conmutan los datos entre ambas redes.

Gatekeepers

El gatekeeper puede ser considerado el cerebro de una red H.323. Es el punto central para todas las llamadas dentro de la red H.323. A pesar de que no es necesario que exista un gatekeeper en toda red H.323, ellos proveen importantes servicios como por ejemplo:

- Direccionamiento
- Autorización y autenticación
- Administración del ancho de banda
- Facturación

Los gatekeepers también pueden brindar servicios de enrutamiento de llamadas. Una red H.323 que no cuente con un gatekeeper no tendrá dichas características. En cambio si este está presente las terminales deben usar sus servicios.

El estándar define servicios obligatorios que el gatekeeper debe proveer y otros opcionales. Los servicios obligatorios son:

- Traducción de direcciones: las llamadas originadas dentro de la red H.323 pueden usar alias para especificar la dirección destino. Para las llamadas originadas fuera de la red H.323, generalmente recibidas por un gateway, el gatekeeper debe traducir el número telefónico en una dirección de la terminal destino.
- Control de admisión: el gatekeeper controla la admisión a la red H.323.
- Control de ancho de banda: el gatekeeper provee soporte para el control del ancho de banda. Por ejemplo, permite establecer y controlar la cantidad de llamadas simultaneas.

Los servicios opcionales que provee un gatekeeper son:

- Enrutamiento de señalización de llamadas: las terminales y gateways envían mensajes de señalización al gatekeeper, quien se encarga de su entrega al destinatario. Alternativamente, es posible enviar dichos mensajes directamente al destino, es decir evitando al gatekeeper perdiéndose la capacidad del monitoreo de llamadas. El enrutamiento a través del gatekeeper ofrece un mejor desempeño de la red, dado que es posible tomar

decisiones de enrutamiento teniendo en cuenta congestiones de la red y ancho de banda disponible.

- Autorización de llamada: cuando un punto final envía un mensaje de señalización al gatekeeper, es posible aceptar o rechazarlo. Las razones de rechazo pueden ser restricciones basadas en el acceso, en el tiempo, por origen o destino, etc. Esto se realiza de acuerdo a la especificación H.225[Ref. H.225].
- Administración de llamadas: el gatekeeper puede mantener información de todas las llamadas activas, de modo de poder controlar su zona, proveyendo información a la función de administración de ancho de banda o enrutamiento. De esta forma, se aprovechan rutas alternativas que balancean la carga en los enlaces.

Unidad de control multipunto o MCU

Las MCUs administran las conferencias de tres o más terminales H.323, de forma tal que un interesado en unirse en una conferencia debe conectarse con la MCU en vez de administrar una comunicación con cada participante. La MCU administra los recursos de una conferencia, negocia con las terminales el tipo de codec de audio y video a usar, y manipula el flujo de datos.

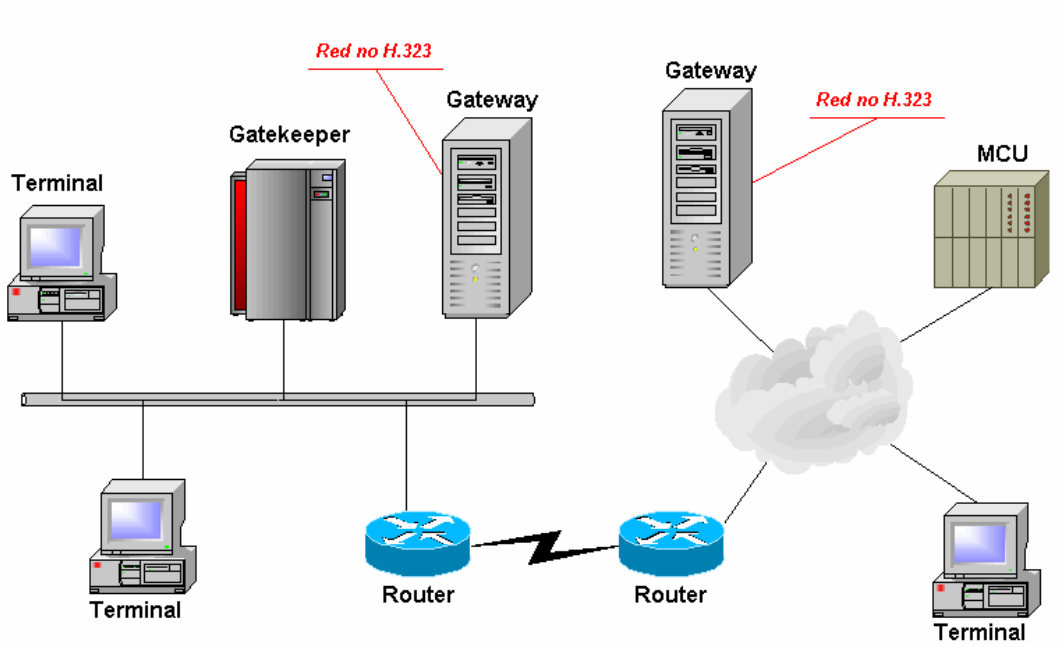
Una MCU puede descomponerse en dos partes fundamentales:

- *Controlador multipunto o MC:* provee funciones de control
- *Procesador multipunto o MP:* recibe y procesa flujos de audio, video y/o datos. Su tarea es la de multiplexar los datos propios de la conferencia minimizando la carga de procesamiento en los puntos finales.

Es importante volver a remarcar que dentro del estándar H.323, los gatekeepers, gateways y MCUs son componentes diferentes; sin embargo, varios fabricantes deciden agruparlos en un único dispositivo físico.

Zonas H.323

Una zona H.323 es un conjunto de terminales, gateways y MCU manejados por un único gatekeeper. Una zona incluye al menos una terminal y puede o no incluir gateway o MCU. Es de destacar que una zona H.323 es independiente de la topología de red y puede estar compuesta por múltiples segmentos de redes conectados por routers u otros dispositivos.



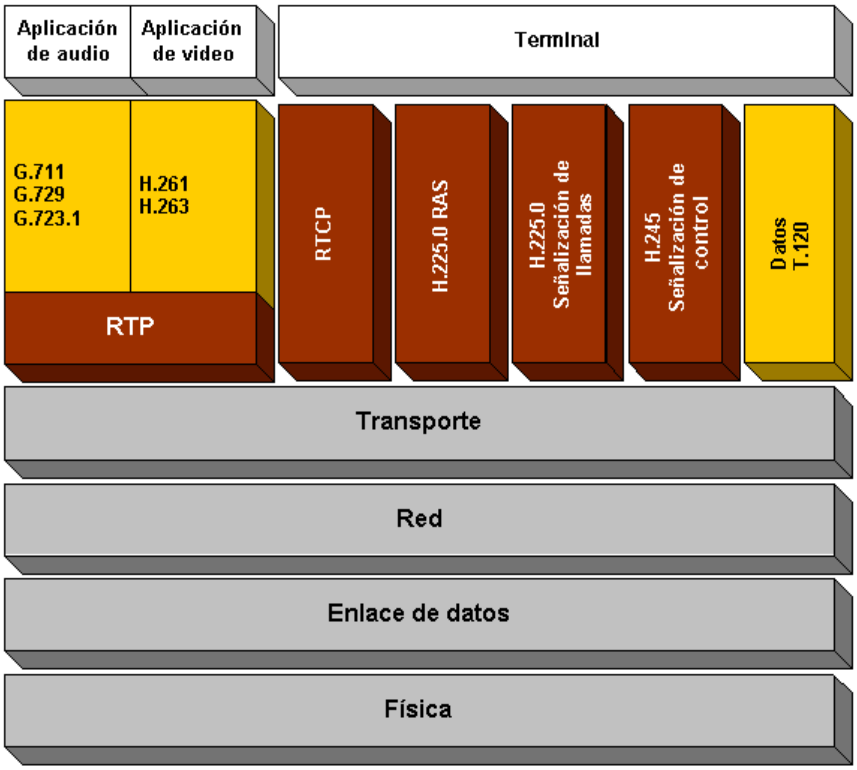
En el dibujo puede observarse una zona que abarca una topología que comprende más de un segmento delimitado por dos routers, e incluso atraviesan una WAN.

Dejando de lado la topología, podemos observar las entidades que participan. Las terminales intercambian datos de audio/video pudiendo ser simples PCs de escritorio o teléfonos H.323. Estas terminales mantienen comunicaciones que son administradas por el gatekeeper. El gateway permite a cualquier terminal de la zona H.323 establecer comunicaciones con entidades pertenecientes a otro tipo de red diferente, como por ejemplo teléfonos de la PSTN. Por su parte, la MCU es la encargada de llevar a cabo conferencias de más de dos participantes.

Protocolos definidos en el estándar

La suite de protocolos H.323 fue diseñada para operar de forma independiente sobre la capa de transporte de la red subyacente. Se vale de un conjunto de protocolos entre los que figuran:

- Codecs de audio y video.
- RAS: registración, admisión y estado.
- H.225.0: señalización de llamadas.
- H.245: señalización de control.
- T.120: conferencia de datos.
- RTP: protocolo de transporte en tiempo real.
- RTCP: protocolo de control en tiempo real.



Distribución de los protocolos en el modelo OSI

Codecs

Como se mencionó en la introducción de VoIP, los codecs cumplen una función importante en las comunicaciones multimedia. H.323 define qué codecs son obligatorios para establecer una comunicación y cuáles son opcionales según el tipo de datos a intercambiar.

Básicamente, el estándar define cuál es la configuración mínima que debe soportar cualquier entidad H.323. Esta configuración mínima se refiere a una comunicación de audio donde el codec propuesto es G.711 y cuya propiedad es la de codificar los datos utilizando 64kbps. Además, se proponen otros codecs opcionales como G.722 de 64, 56, y 48 kbps, G.723.1 de 5.3 y 6.3 kbps, G.728 de 16 kbps, y G.729 de 8 kbps.

Respecto a las comunicaciones de video, en el estándar están contempladas como una caso opcional. Sin embargo, cualquier terminal que provea comunicaciones de video, debe soportar al menos el codec de video H.261.

RAS

El protocolo RAS permite a las terminales y gateways comunicarse con un gatekeeper, para realizar tareas de registración, control de admisión y estado. A continuación se listan los mensajes que intervienen en RAS:

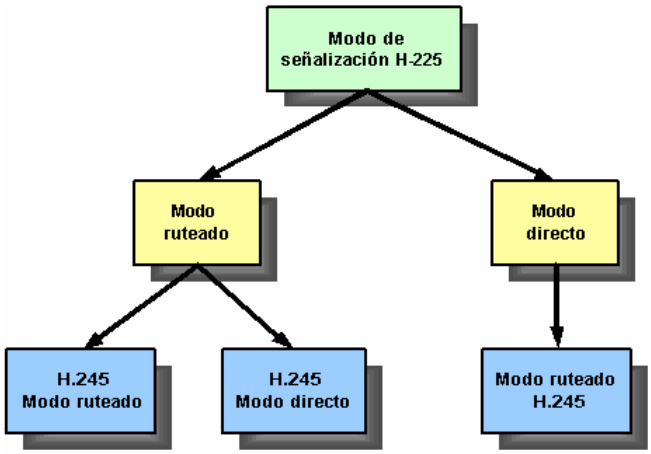
- ARQ: admision request
- ACF: admision confirm

- ARJ: admision reject
- DRQ: Pedido de desconexión
- DCF: confirmacion de desconexión
- BRQ: bandwidth request
- BCF: bandwidth confirm
- BRJ: bandwidth reject

Más adelante, se completará el entendimiento de cómo trabaja este protocolo, por medio de ejemplos plasmados en los flujos de mensajes que intercambian las entidades con el gatekeeper.

H.225.0

H.225.0 es un subconjunto del protocolo de control de comunicaciones Q.931 el cual provee señalización de llamadas, es decir permite el establecimiento y finalización de las mismas. La señalización H.225 para el establecimiento de llamadas puede realizarse en forma directa entre los puntos finales de la comunicación, modo directo, o a través del gatekeeper de la zona, modo ruteado. Estas formas de encaminamiento impactan sobre la forma de encaminamiento que tendrá el protocolo H.245.



Con el propósito de administrar el inicio y fin de las llamadas, H.225.0 se ayuda de los siguientes mensajes:

- SETUP
- CALL PROCEEDING
- ALERTING
- CONNECT
- RELEASE COMPLETE

Al igual que con RAS, se deja para más adelante la forma en que estos mensajes son intercambiados debido a que es mucho más didáctico utilizar flujos de mensajes.

H.245

La función de H.245 es intercambiar mensajes de control entre los extremos de una conferencia con el propósito de controlar la llamada entre los puntos finales H.323. Estos mensajes de control llevan información relacionada con las siguientes características:

- Intercambio de capacidades: codecs de audio, codecs de video, etc
- Apertura y cierre de canales lógicos usados para la transmisión de datos.
- Mensajes de control de flujo.

Este protocolo es de vital importancia al momento de llevar a cabo las llamadas en redes de datos puesto que, a diferencia de la PSTN donde todos los usuarios disponen del mismo dispositivo de comunicación, en esta red los usuarios no solo tienen distintos dispositivos sino que los mismos pueden tener dispositivos con diferentes periféricos e incluso diferentes codecs. Por ello es necesaria la presencia de este protocolo que permite la negociación de capacidades entre los extremos de la comunicación.

Los distintos tipos de mensajes H.245 son:

- TerminalCapabilitySet
- TerminalCapabilitySetAck
- OpenLogicalChannel
- OpenLogicalChannelAck
- EndSessionCommand

Datos T.120

El estándar T.120 define las conferencias documentales y aplicaciones compartidas también llamadas conferencias de datos. La recomendación especifica cómo distribuir archivos e información gráfica en tiempo real de forma eficiente y confiable durante una conferencia. Los ejemplos de uso más evidentes incluyen pizarras compartidas, intercambio de imágenes y ventanas de aplicaciones compartidas.

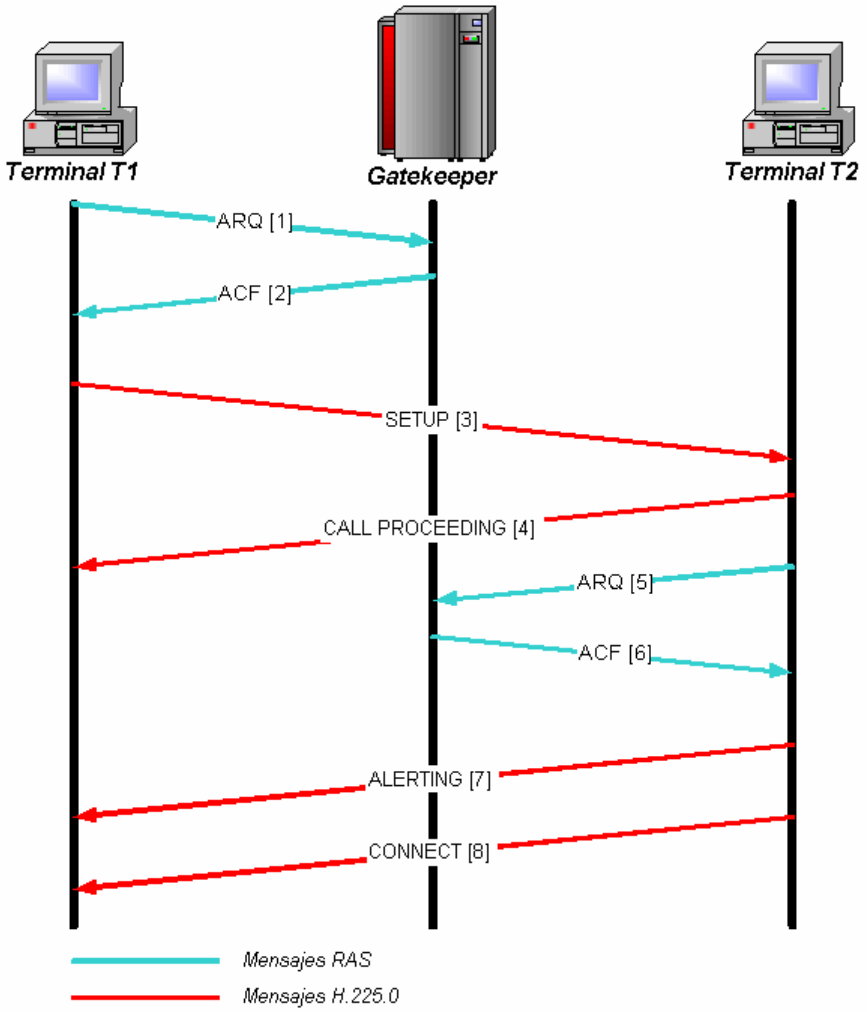
RTP y RTCP

H.323 se vale de estos protocolos para el transporte en tiempo real de la información involucrada en una conferencia. Estos protocolos son descriptos en el capítulo 7 llamado protocolos de transporte de datos multimedia.

Flujos de mensajes

A modo de ejemplo se muestra como es la secuencia de mensajes involucrados en una llamada entre dos terminales H.323 T1 y T2, en la que se supone que la llamada se efectúa a través del gatekeeper, pero el encaminamiento de la misma es directo, es decir, no se realiza a través del gatekeeper.

Inicialmente se muestra el flujo de mensajes involucrados en el establecimiento de la llamada, para ello se supone que tanto T1 como T2 ya están registrados con el gatekeeper por medio de los mensajes RRQ “Registration Request” y RCF “Registrtion Confirm”:



- Luego de estar registrado T1 puede iniciar o aceptar una llamada solamente luego de haber pedido admisión al gatekeeper. Para ello envía un mensaje *ARQ [1]* al gatekeeper para pedir admisión. Además en este momento indica al gatekeeper que utilizará señalización de llamada directa.

- El gatekeeper confirma la admisión de T1 enviando un *ACF [2]* a esta terminal. Además en este mismo mensaje se indica que T1 puede utilizar la señalización de llamada directa y se provee además de información para la localización de T2.

- T1 envía un mensaje *SETUP [3]* o establecimiento de llamada a T2 requiriendo una conexión.

- T2 responde con un mensaje *CALL PROCEEDING [4]* a T1, donde no está confirmando la conexión sino que está indicando al otro extremo que está procesando la solicitud de conexión.

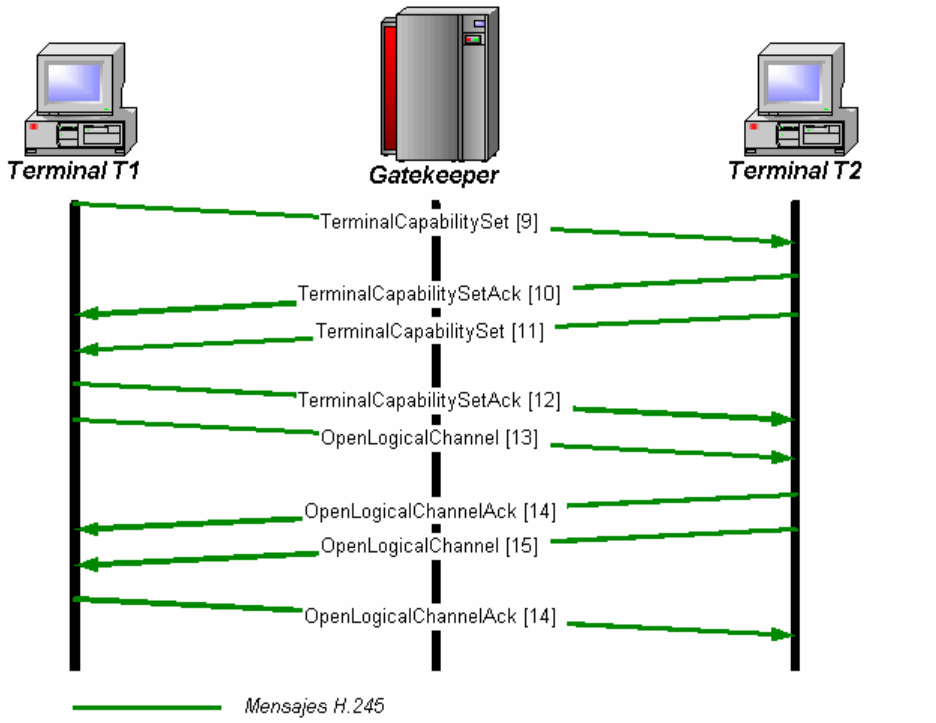
- T2 pide admisión a su gatekeeper enviando un mensaje *ARQ [5]*.

- El gatekeeper confirma a T2 que puede aceptar la llamada. *ACF [6]*.

- T2 alerta a T1 del establecimiento de la conexión enviándole un mensaje *ALERTING [7]*, el cual indica que en el otro extremo el teléfono está sonando.

- Finalmente T2 confirma el establecimiento de la conexión enviando un mensaje *CONNECT [8]* a T1, y la llamada se establece.

Una vez conectados ambos extremos, entra en juego H.245 acordando entre las partes determinados parámetros de modo de hacer posible el intercambio de datos multimedia. A continuación se muestra el flujo de mensajes de control H.245 en la señalización H.323.



- T1, quien inició la llamada, envía un mensaje *TerminalCapabilitySet [9]* a T2 para intercambiar capacidades.

- Si T2 confirma la recepción de las capacidades de T1 mediante un mensaje *TerminalCapabilitySetAck [10]* a T1, reconociendo sus capacidades.

- Luego T2 envía sus capacidades a T1 enviándole un mensaje *TerminalCapabilitySet [11]*

- T1 le envía un mensaje *TerminalCapabilitySetAck [12]* a T2, reconociendo sus las capacidades.

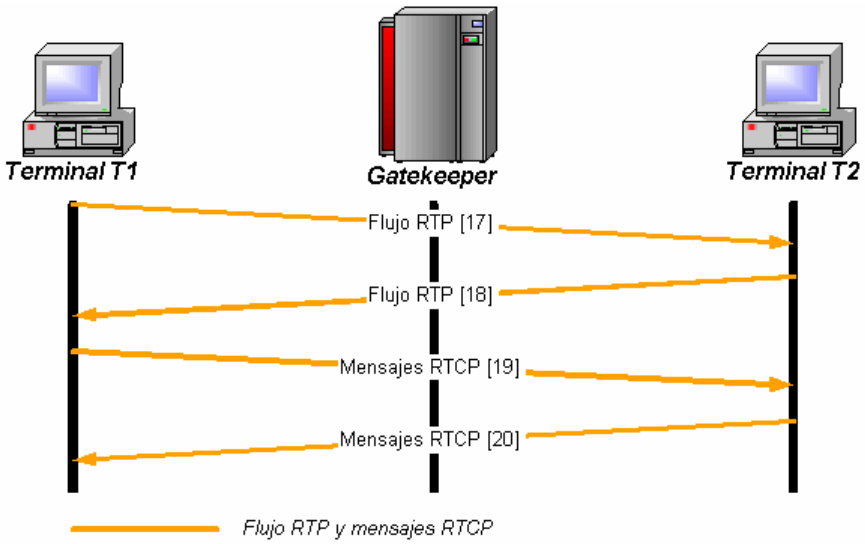
- T1 envía un mensaje *OpenLogicalChannel [13]* a T2, solicitando la apertura de un canal lógico. En este mensaje se incluye entre otros el codec a utilizar y el puerto UDP para la transmisión de los datos propiamente dichos.

- T2 asiente el establecimiento del canal lógico unidireccional de T1 a T2 por medio de un mensaje *OpenLogicalChannelAck [14]*. Incluido en el mensaje está el puerto UDP para la recepción de los datos RTP que deberá ser usado por T1 cuando le envíe información a T2.

- Luego T2 abre un canal lógico con T1 enviándole un mensaje *OpenLogicalChannel [15]*.

- T1 asiente el establecimiento del otro canal lógico unidireccional de T2 a T1 por medio de un mensaje *OpenLogicalChannelAck [16]*. Ahora, la comunicación entre T1 y T2 es bidireccional para la transmisión de paquetes RTP.

En este punto está establecido un canal RTP/RTCP bidireccional por el cual T1 y T2 intercambian datos multimedia:



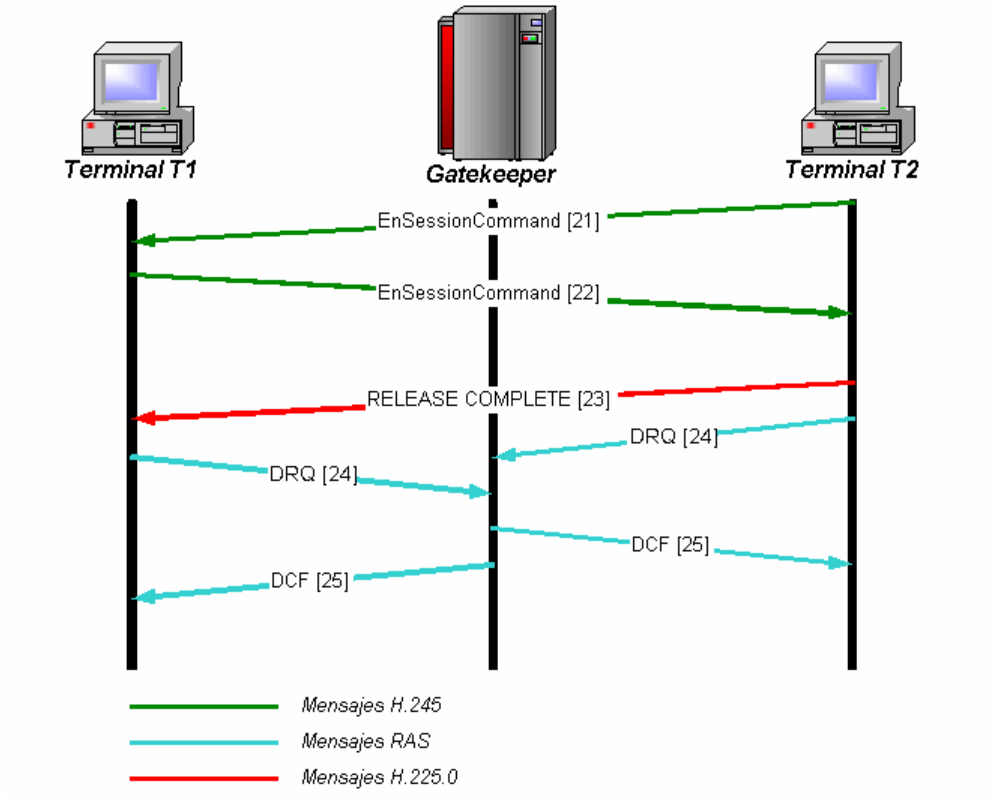
- T1 envía el stream de datos encapsulado en RTP a T2.

- T2 envía el stream de datos encapsulado en RTP a T1.

- T1 envía mensajes RTCP a T2.

- T2 envía mensajes RTCP a T1.

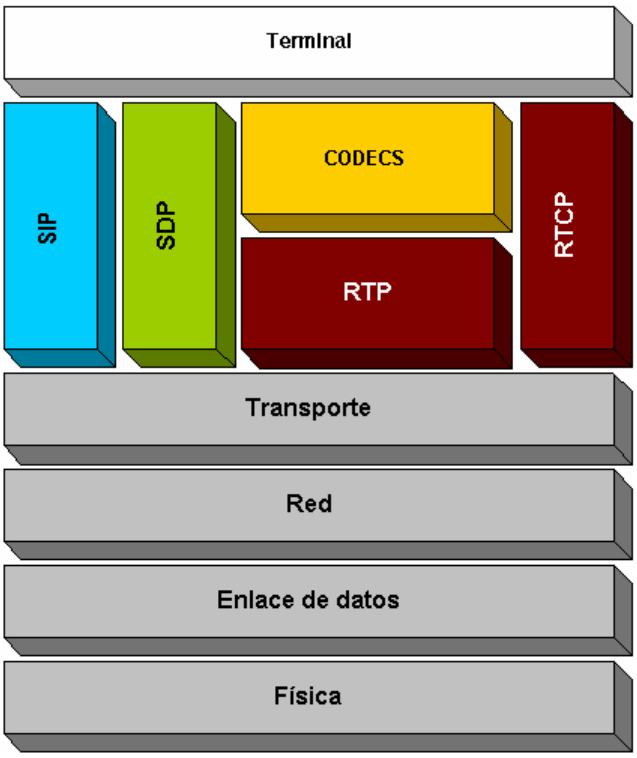
Finalmente cuando una de las partes desea terminar la comunicación, se produce la siguiente secuencia de mensajes:



- T2 inicia la liberación de la llamada. Para ello envía un mensaje H.245 *EndSessionCommand [21]* a T1.
- T1 confirma la liberación enviando un mensaje H.245 *EndSessionCommand [22]* a T2.
- T2 completa la liberación enviando un mensaje H.225 *RELEASE COMPLETE [23]* a T1.
- T1 y T2 se desconectan del gatekeeper enviándole un mensaje RAS *DRQ [24]*.
- El gatekeeper se desconecta de T1 y T2 por medio del envío de un mensaje RAS *DCF [25]*.

Capítulo 4: Protocolo de señalización SIP

El otro estándar existente para VoIP es SIP, siendo su principal característica la de definir el proceso de señalización entre entidades. Este protocolo está ampliamente desarrollado en la RFC 2543 correspondiente al grupo de trabajo MMUSIC perteneciente a la IETF. A diferencia de H.323, SIP fue diseñado específicamente para Internet aprovechando la flexibilidad del modelo TCP/IP. Su primer Draft aparece a principios de 1996, y luego de varias extensiones del mismo, se lanza la RFC 2543 en marzo de 1999. Su arquitectura fue cuidadosamente definida de forma tal de simplificar el desarrollo de aplicaciones multimedia, ubicando el conjunto de protocolos que describe el estándar por sobre la capa de transporte del modelo OSI.



La operación del protocolo se basa en la creación, modificación y terminación de *sesiones* o llamadas entre uno o más participantes. Estas sesiones incluyen conferencias multimedia de audio y video a través de Internet. A diferencia de H.323, no se consideran conferencias de datos multimedia como mensajes de texto o pizarras compartidas. Sin embargo, es posible encontrar aplicaciones que proveen esta funcionalidad, no respetando el estándar sino extendiéndolo.

SIP utiliza URLs para identificar al origen y el destino de las sesiones. Para ello utiliza una codificación como la del correo electrónico, por ejemplo bill@info.unlp.edu.ar. Esta codificación puede ser dependiente o independiente del host y además puede incluir números de puertos en la URL. El puerto por defecto es el port UDP 5060.

Como característica central y a diferencia de la telefonía convencional, SIP ofrece mecanismos de autenticación y control de acceso, permitiendo a los clientes

rechazar llamadas no autorizadas. Además de los servicios presentes en la PSTN se ofrecen servicios de redirección, lo que facilita la movilidad de los usuarios.

A continuación se describen las entidades especificadas en el estándar las cuales, como ya se ha mencionado en el caso de H.323, es difícil materializarlas en un dispositivo de hardware.

Las distintas entidades que pueden existir en un esquema de comunicaciones SIP pueden clasificarse como clientes, servidores o ambos. Los clientes, inician *requerimientos* SIP y pueden o no interactuar en forma directa con los usuarios, mientras que los servidores aceptan y envían *respuestas* a requerimientos.

- User agent client (UAC): es un cliente que inicia requerimientos SIP.
- User agent server (UAS): es un servidor que contacta al usuario cuando se recibe un requerimiento SIP y retorna una respuesta en nombre de este. Dicha respuesta acepta, rechaza o redirige el requerimiento.
- User agent (UA): es una aplicación o dispositivo que contiene tanto un UAC como UAS. Ejemplos de esta entidad son los teléfonos SIP o aplicaciones que utilizan SIP para poder intercambiar audio y/o video.
- Proxy o proxy server: actúa como cliente y servidor, cumpliendo la función de agente intermediario. Su propósito es llevar a cabo requerimientos en nombre de otros clientes. Para ello, al recibir un requerimiento, puede que lo complete en forma interna o lo deba retransmitir a otros proxies.
- Redirect server: su tarea es aceptar requerimientos SIP desde clientes y traducirlos en cero o más direcciones; luego devuelve este resultado al cliente. A diferencia de los proxy servers, no inicia nuevos requerimientos SIP. Tampoco acepta llamadas, este sólo responde pedidos.
- Registrar Server: solo acepta requerimientos REGISTER de tal forma de registrar la o las ubicaciones de cada usuario.
- Location Server: provee un servicio utilizado por el proxy server o por el redirect server para obtener información sobre la ubicación de la entidad llamada.

Protocolos definidos en el estándar

Al ser SIP un protocolo simple, no engloba tantos protocolos como lo hace H.323. SIP puede considerarse un protocolo que define una serie de mensajes para la señalización, y solamente se complementa de otros tres protocolos extras: SDP para la negociación de capacidades, codecs para la compresión de datos y RTP/RTCP como mecanismo de transporte de datos multimedia en tiempo real.

Como SIP fue concebido teniendo en cuenta que sería un protocolo más de Internet, el diseño de los mensajes respetó la línea de dos protocolos existentes: SMTP y HTTP. Es por ello que SIP es un protocolo textual, fácil de debuguear y comprender por aquellos que están familiarizados con el estándar HTTP. Incluso los mecanismos de autenticación provistos por el protocolo son similares a los empleados entre los servidores web y browsers.

Mensajes SIP

SIP es el encargado de iniciar, mantener y terminar las comunicaciones desde la perspectiva de señalización. Para ello define dos categorías de mensajes:

- Requerimientos: estos mensajes son enviados por los clientes con el fin de iniciar o cambiar el estado de una comunicación

- Respuestas: estos mensajes son emitidos por los servidores en respuesta a solicitudes previas realizadas por los clientes.

Requerimientos:

A continuación se nombran los requerimientos definidos por el estándar dejando evidencia de la simpleza del protocolo:

- INVITE: corresponde al primer mensaje enviado por el llamador. Generalmente, el mensaje INVITE contiene una descripción SDP con las capacidades del que llama.

- ACK: este mensaje es enviado por el llamador en respuesta a una confirmación exitosa con código 200 aceptando un INVITE previamente enviado. Este requerimiento indica que el llamador ha recibido la confirmación a un requerimiento INVITE. El contenido de un ACK puede contener una descripción SDP. Si la confirmación recibida fue exitosa, y no contenía una descripción SDP, se asume que la propuesta del INVITE será usada.

- OPTIONS: este mensaje se utiliza para consultar las capacidades de un UA.

- BYE: un cliente envía este mensaje a un UA para terminar la llamada. El que envíe este mensaje cesa la transmisión de datos multimedia sin importar la respuesta del otro extremo.

- CANCEL: cancela un requerimiento en progreso, pero no tiene efecto sobre una comunicación establecida.

- REGISTER: este mensaje es utilizado para registrar la dirección listada en el encabezado del mensaje SIP ante un servidor de registración.

Los mensajes anteriores completan la mitad funcional de SIP restando por describir las respuestas a estos mensajes. Las respuestas forman una extensa lista debido al buen manejo de errores provisto por el protocolo. Existen diferentes clases de respuestas clasificadas basándose en códigos numéricos de tres cifras, dónde el primer dígito del identificador es el que permite distinguir la categoría de una respuesta. A continuación se explican por categoría las posibles respuestas que ofrece SIP.

Respuestas informativas:

Códigos: 100 al 199.

- Código 100 – Trying: retornado por un proxy, redirect o UAS a un UAC indicando que algún requerimiento está siendo procesado.

- Código 180 – Ringing: indica que el teléfono virtual o real está sonando en el otro extremo.

- Código 181 – Call forwarding: retornado por un proxy server indicando que la llamada fue redirigida e informando en el cuerpo del mensaje a dónde fue redirigida.

- Código 182 – Queued for service: este mensaje es útil en aplicaciones que permiten encolar llamadas entrantes hasta terminar con las pendientes.

Respuestas Satisfactorias:

Código 200. El término utilizado es *"OK"* y como su nombre lo indica, significa que el requerimiento enviado previamente fue ejecutado en forma satisfactoria.

Respuestas de Redirección:

Códigos: 300 al 399. Indican que la llamada necesita más procesamiento antes de poder determinar si es posible completarla.

- Código 300: la dirección en el requerimiento resuelve en más de una opción. Se retornan las distintas alternativas para que el llamador seleccione una y redirija la llamada.

- Código 301: el usuario llamado se ha retirado de su ubicación indicándose en el encabezado del mensaje la posible ubicación del mismo. Es posible recibir una lista de ubicaciones o un código de no encontrado.

- Código 302: el usuario llamado se ha movido temporalmente y puede encontrarse en la dirección retornada.

- Código 305: el cliente llamado no puede ser accedido en forma directa y debe contactarse a través de un proxy.

Respuestas de fallas en requerimientos de los clientes:

Códigos: 400 al 499. El servidor no puede satisfacer el requerimiento.

Respuestas de fallas en servidores:

Códigos: 500 al 599. El requerimiento pudo haber sido válido, pero el servidor no ha podido ejecutarlo.

Respuestas de fallas globales:

Códigos 600 al 699. El requerimiento de un cliente no pudo ser atendido por ningún servidor.

SDP

Es posible encontrar una analogía entre este protocolo y el estándar H.245 utilizado por H.323. Como ya se mencionó en esa sección, las redes de datos presentan un problema considerable, ausente en la PSTN, respecto de las capacidades de los extremos de una comunicación ya sea en cuanto a hardware o codecs a utilizar. La solución es contar con algún mecanismo de intercambio de capacidades y negociación de parámetros mínimos indispensables para el inicio de una llamada. Esta es la tarea de SDP.

Este protocolo definido en la RFC 2327 provee los mecanismos y la sintaxis para describir sesiones multimedia, siendo el protocolo elegido en comunicaciones IP basadas en texto como son SIP y MGCP.

RTP y RTCP

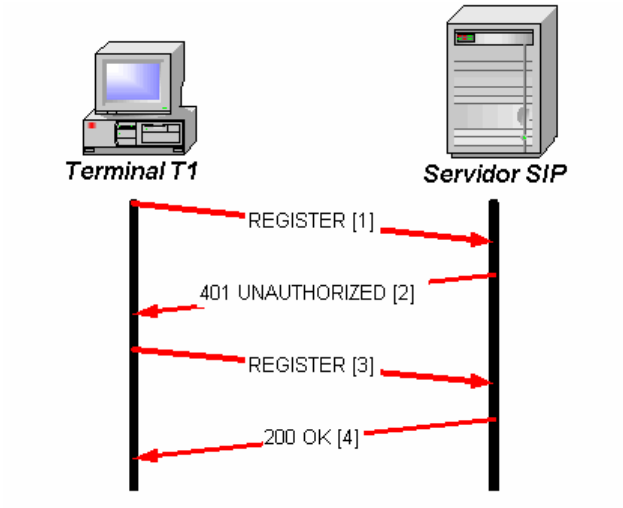
SIP también utiliza RTP/RTCP para el transporte en tiempo real de información multimedia. Estos protocolos se describen en las secciones siguientes.

Flujo de Mensajes

Como primer ejemplo, se considera el caso de llamar a una entidad que podría ubicarse en diferentes dispositivos físicos; como máquinas o teléfonos SIP. Por esta razón se provee un mecanismo para registrar la o las ubicaciones de un usuario en forma dinámica ante un servidor SIP. Este servidor, llamado registrar server, mantiene información de ubicación, mientras que el redirect server responde solicitudes de ubicación de un usuario retornando para cada pregunta una o varias ubicaciones, según indique el location server.

La acción a tomar cuando se recibe una lista de ubicaciones, depende del tipo de servidor SIP. Por ejemplo, un proxy server podría llamar a cada dirección en la lista secuencial o paralelamente hasta obtener una respuesta exitosa o un rechazo de la llamada por parte de la entidad llamada.

El proceso de registración de usuarios se realiza a través de requerimientos *REGISTER*. A continuación se muestra el intercambio de dichos paquetes, donde un UAC se registra ante un servidor SIP:



- El UAC envía un requerimiento *REGISTER [1]* al servidor.
- El servidor desafía al UAC para autenticarlo enviando una respuesta con código 401, *UNAUTHORIZED [2]* indicando la desautorización.
- El UAC debe ingresar su nombre de usuario y contraseña, encriptar la información correspondiente al desafío impuesto por el servidor SIP y enviar nuevamente un requerimiento *REGISTER [3]* con esta información encriptada.
- El servidor valida la credencial del usuario y lo registra en su base de datos de contactos. Luego el servidor retorna una respuesta *200 OK [4]* al UAC.

Una vez registrados los usuarios, se procede con las llamadas. Cuando se realiza una llamada, la misma depende de la configuración de los clientes dependiendo del uso o no de un proxy, de la misma forma en que un browser puede configurarse para que utilice o no un proxy HTTP. Si se considera la existencia de un proxy, cuando el cliente envía un requerimiento INVITE, éste se envía al proxy para que se encargue de su entrega al destinatario indicado en la URL. Si por el contrario, no se utiliza un proxy, entonces dicho requerimiento es enviado directamente al destinatario que indica la URL.

El requerimiento INVITE típicamente contiene una descripción de sesión, escrita en formato SDP, que provee a la entidad llamada con suficiente información para unirse a la sesión. La descripción de la sesión enumera los tipos de medios y formatos que el llamador pretende utilizar y, si el destinatario desea aceptar la llamada, responde a la invitación retornando una descripción similar. Una vez establecida la llamada, los datos multimedia se intercambian a través de RTP. Finalmente se concluye la llamada con el envío de un requerimiento BYE, el cual a su vez es confirmado.

El intercambio de requerimientos y respuestas para el caso básico de una llamada, sin intervención de proxy alguno, se muestra en la figura 1. En la figura 2 se realiza el mismo procedimiento, pero con la presencia de un proxy server, mientras que en la figura 3 se lo realiza con un redirect server.

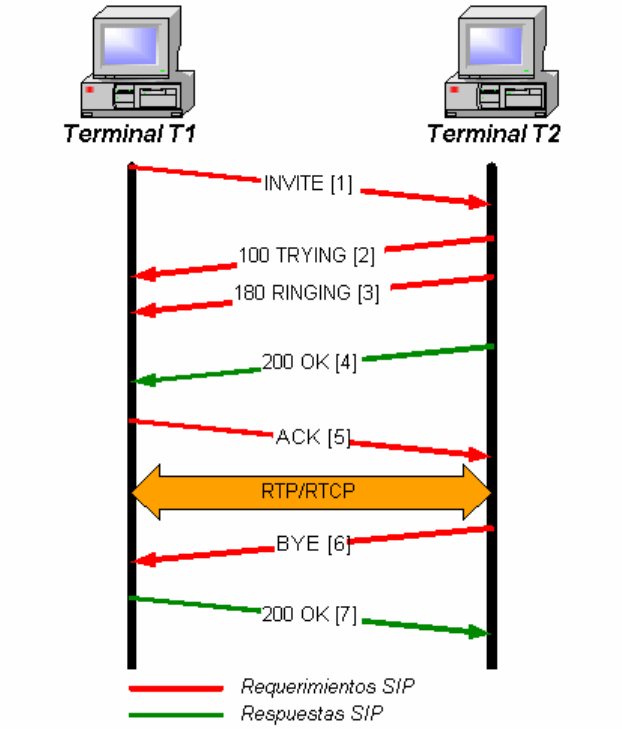


Figura 1

De la figura 1 se observa el pedido de conexión mediante un mensaje SIP INVITE. La correspondiente respuesta, el mensaje SIP TRYING indica al extremo que origino la llamado, que la misma esta siendo procesada, mientras que la siguiente respuesta informativa RINGING indica que el dispositivo llamado le esta advirtiendo al usuario respecto de la llamada entrante. Por ultimo, la respuesta satisfactoria OK indica que la llamada quedó establecida. Dicha respuesta se confirma mediante un ACK. Finalmente, luego del intercambio de datos la llamada se cierra por medio de unn mensaje BYE el cual es correspondido con un OK.

Tanto en la figura 2 como el la figura 3 se omitieron las respuestas TRYING y RINGING con el fin de ayudar a la claridad del flujo presentado.

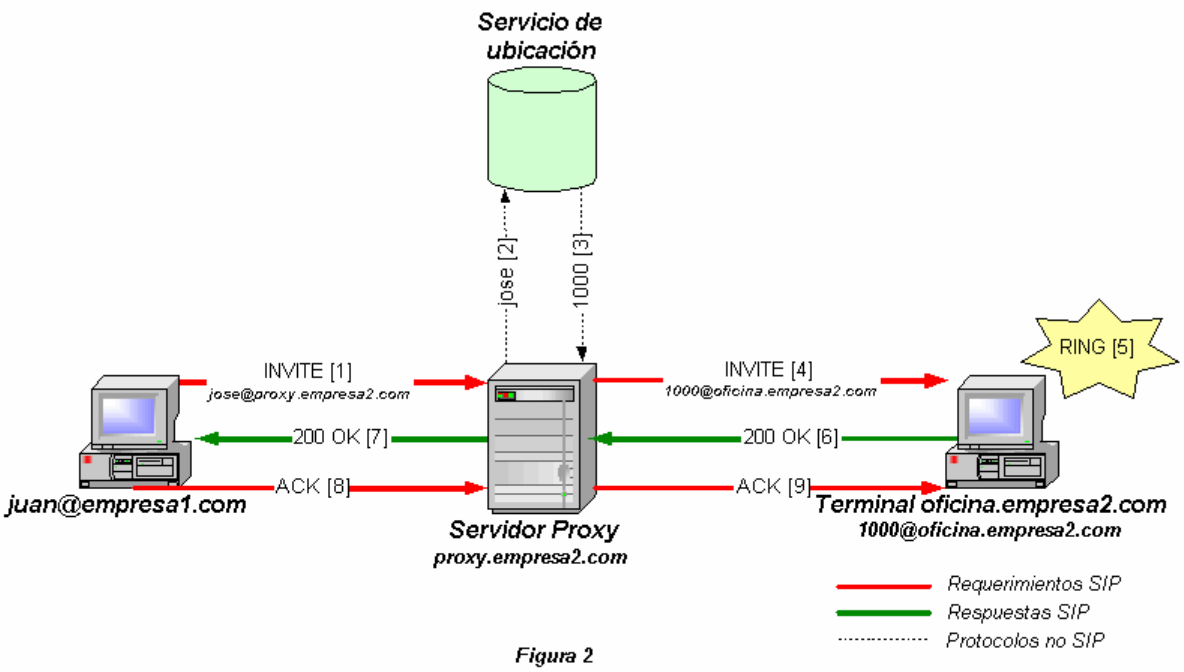


Figura 2

En la figura 2, el proxy acepta el requerimiento *INVITE [1]*, dirigido a la dirección *"jose@proxy.empresa2.com".* Luego, el proxy utiliza el servicio de ubicación con toda o parte de la dirección, *jose [2]*, y obtiene una ubicación más precisa, *1000 [3]*. Mas tarde, el proxy envía un requerimiento SIP *INVITE [4]* a la, o las direcciones retornadas por el servicio de ubicación en nombre del usuario que genero la llamada. El extremo llamado advierte la invitación y emite una indicación sonora o visual de *RING [5]* independientemente de aceptación o no de la llamada por el usuario. En caso de aceptar la llamada, el usuario llamado retorna una indicación exitosa *200 OK [6]* al proxy server. Del mismo modo, el proxy retorna una respuesta *200 OK [7]* a la entidad que originalmente inició la llamada. La recepción de esta respuesta es confirmada utilizando un requerimiento *ACK [8]*, que es restransmitido por el proxy hacia la entidad llamada *[9]*.

En la figura 3, el redirect server acepta un requerimiento *INVITE [1]*, contacta el servicio de ubicación tal cual sucedía en el ejemplo anterior pero, en vez de contactar la nueva dirección encontrada, retorna la dirección a la entidad que llama utilizando una respuesta *302 MOVED TEMPORARILY [4]*, hecho que es confirmado por medio de un requerimiento *ACK [5]*. Luego, la entidad que llama realiza un nuevo requerimiento a la dirección retornada por el redirect server, *INVITE [6]*. En el ejemplo, la llamada es exitosa por la respuesta *200 OK [7]*. La llamada se completa con el requerimiento *ACK [8]*.

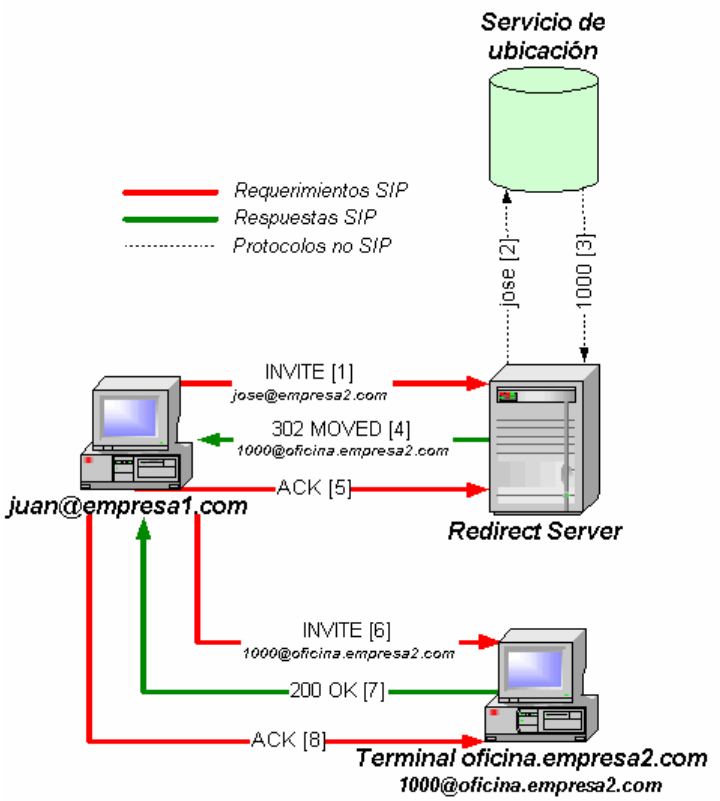

Servicio de ubicación
Requerimientos SIP
Respuestas SIP
Protocolos no SIP
jose [2]
1000 [3]
INVITE [1]
jose@empresa2.com
302 MOVED [4]
1000@oficina.empresa2.com
ACK [5]
juan@empresa1.com
Redirect Server
INVITE [6]
1000@oficina.empresa2.com
200 OK [7]
ACK [8]
Terminal oficina.empresa2.com
1000@oficina.empresa2.com

Figura 3

Capítulo 5: Comparación: H.323 - SIP

Durante la introducción de ambos protocolos, se marcaron diferencias funcionales. Estas diferencias se presentan en diversas comparaciones que contrastan los protocolos de una forma competitiva, siendo muchas veces poco objetivas ya que por cuestiones comerciales se favorece a uno u otro según convenga. Sin embargo, tanto SIP como H.323 se encuentran en una etapa de crecimiento, y por lo tanto hay varios puntos indefinidos. Por esta razón la perspectiva de comparación se achica, puesto que es imposible contrastar características que no son comparables, y se simplifica si en el análisis sólo se estudian aquellas características actualmente definidas.

Es probable que en un futuro H.323 y SIP converjan en un único estándar, pero por lo pronto ambos siguen caminos separados. Una comparación justa podría hacerse cuando SIP alcance un estado de implementación más avanzado, ya que actualmente muchas capacidades están disponibles como drafts. Esto termina favoreciendo a H.323, mostrándose más robusto y completo. Sin embargo, puede anticiparse que SIP sigue la misma línea que atravesó H.323.

Creación

H.323 fue diseñado teniendo en cuenta los requerimientos presentes en las comunicaciones multimedia sobre redes IP, incluyendo conferencias de audio, video y datos. Como resultado, es razonable que los usuarios esperen el mismo nivel de robustez e interoperabilidad que pueden encontrar en la PSTN actual.

SIP, en cambio, fue diseñado para establecer sesiones entre usuarios, siendo una componente modular y flexible propia de la arquitectura de Internet. En SIP, el concepto de llamada es un tanto pobre, siendo en sí una sesión con intercambio de flujos multimedia. Además no existe una definición estandarizada para conferencias multimedia, y su integración con otros estándares queda libre a la implementación de los vendedores. En consecuencia, la expectativa de los usuarios para este protocolo es del mismo nivel de robustez e interoperabilidad que puede hallarse en otros servicios presentes en Internet. Esto significa que los usuarios pueden sentir la misma sensación de incompatibilidad presente en un browser que intenta acceder a un servidor web, donde pueden presentarse problemas de versión con el browser o la necesidad de plug-ins.

Además es importante considerar las fechas en que cada estándar fue aprobado: SIP fue considerado estándar a principios de 1999, mientras que H.323 lo fue tres años antes, a fines de 1996.

Confiabilidad

H.323 define un número de funciones para manejar la falla de entidades intermedias. Por ejemplo, si un gatekeeper falla, el protocolo está diseñado para utilizar un gatekeeper alternativo. Por otra parte, si una llamada que está siendo

enrutada a través de entidades intermedias de señalización falla, la misma se reenvía a través de una ruta alternativa de tal forma de no interrumpir la comunicación.

En contraposición, SIP no define procedimientos para manejar las fallas de dispositivos, es decir que si por ejemplo un cliente SIP falla, no existe forma alguna de que el proxy detecte la falla a menos que envíe mensajes de pedido de conexión INVITE al dispositivo y espere a que expiren. Es más, si un proxy SIP falla, los clientes no tienen forma de detectar este hecho. Además de tener problemas para manejar errores previos al inicio de sesiones, SIP no posee mecanismos de recuperación de llamadas en progreso. Algunos procedimientos SIP de confiabilidad y balanceo de carga se definieron en documentos drafts pero aún no son parte del estándar.

Definición y codificación de mensajes

La definición de los mensajes en SIP y H.323 es diferente. Por su parte H.323 utiliza ASN.1, una notación estructurada, fácil de entender, estandarizada y extremadamente precisa. SIP en cambio utiliza ABNF, una notación sintáctica mayormente utilizada para la especificación formal de lenguajes a través de gramáticas.

La codificación de mensajes está estrechamente relacionada con la definición de los mismos. Por su parte H.323 utiliza un formato binario y compacto, ideal para conexiones de escaso ancho de banda, mientras que SIP codifica sus mensajes en texto ASCII formateado, legible para los humanos.

Tanto la definición como la codificación no implican una gran mejora en el aprovechamiento del ancho de banda ni acelera el proceso de codificación. Sí es cierto que por ejemplo SIP, es mucho más simple de codificar, depurar e incluso entender que H.323.

Extensibilidad

Comercialmente, la extensibilidad es de gran importancia para los vendedores de servicios. Esta facilidad es más directa en SIP ya que la extensión del estándar es tan simple como agregar nuevas líneas al encabezado de los mensajes. Sin embargo, lo que es simple termina convirtiéndose en un arma de doble filo, ya que el protocolo se presta a posibles incompatibilidades entre capacidades propias de diferentes vendedores.

Direccionamiento

Tanto SIP como H.323 utilizan para el formato de sus direcciones, el ya tradicional estilo de una URL.

Movilidad

En referencia a la funcionalidad provista por ambos protocolos, la resolución de direcciones juega un factor primordial ya que es el pilote de la movilidad en VoIP.

Esta característica es factible en H.323 a través de los gatekeepres mediante el uso de otros protocolos como ser H.225, TRIP y/o DNS. SIP, define entidades diferentes que resuelven la movilidad de los usuarios. Las entidades en cuestión corresponden a los servidores proxy y redirect, que también colaboran con otros protocolos como TRIP y/o DNS.

Negociación de capacidades

Otro factor esencial para ambos protocolos es la negociación de capacidades. Este proceso permite a los puntos finales H.323 intercambiar información para negociar qué canales abrir, incluyendo audio, video y canales de datos. Cada canal individual puede abrirse y cerrarse durante una llamada sin interrumpir otros canales. Las entidades SIP no pueden intercambiar facilidades sino que sólo pueden proponer canales multimedia, y la otra entidad está limitada al conjunto de canales propuestos. Se ha propuesto que las entidades SIP puedan utilizar mensajes RE-INVITE para renegociar capacidades. Sin embargo, las pruebas no fueron exitosas, y entre los problemas encontrados se incluye la posibilidad de perder la comunicación.

Interoperabilidad con la PSTN

En la interacción con la PSTN, H.323 parece ser superior a SIP. Como H.323 se basó en protocolos tradicionales de la PSTN, como por ejemplo Q.931, pareciera que H.323 es ideal para interactuar con la PSTN. La forma en que un gateway se integra en una arquitectura puramente H.323 está bien definida por el estándar. Por su parte SIP no se basó en la PSTN para su definición por lo que su integración es más indirecta que la integración de H.323. Tampoco, a pesar que existen implementaciones, existe una arquitectura SIP que describa la interacción con un gateway.

Transporte de datos multimedia

Un factor que comparten ambos protocolos es el medio de transporte de datos multimedia utilizado, *RTP/RTCP*..

Conferencias

Las conferencias de más de dos participantes son factibles en ambos protocolos haciendo que cada nodo final maneje la conferencia por su cuenta, es decir mantenga más de una comunicación. Una alternativa, sólo disponible en H.323, es el uso de una entidad externa que se encarga de la multiplexación de los diferentes flujos multimedia involucrados. Esta entidad es la MCU.

Capítulo 6: Codecs

Como ya se mencionó en secciones anteriores, antes de transmitir audio y video a través de una red de paquetes, es necesario digitalizar y comprimir estos datos. Básicamente este mecanismo, también llamado codificación, es similar al utilizado en la digitalización de audio analógico para su grabación en discos compactos o CDs. La necesidad de digitalizar las señales de audio y video se basa en el hecho de que la unidad de información es el bit. La compresión es importante debido al gran volumen de bytes requeridos en la digitalización de audio y video, puesto que es deseable consumir el menor ancho de banda posible. Para el caso de una imagen de 1024 x 1024 pixeles donde cada pixel se codifica con 24 bits, se necesitan 3 Mbytes. Luego, transmitir dicha imagen a través de un enlace de 64 Kbps, tomaría 7 minutos. Para mejorar este tiempo, puede utilizarse algún mecanismo de compresión, por ejemplo con una relación de 10 a 1, reduciendo el tamaño del archivo a 300 Kbytes, el tiempo de transmisión sería de 6 segundos.

El proceso inverso a la codificación, llamado decodificación, transforma una señal digital en una analógica con el fin de poder reproducirla como audio o video. En consecuencia, las aplicaciones multimedia de audio y video necesitan de herramientas para la codificación y decodificación de señales digitales, denominadas codecs.

Codecs de Audio

Los algoritmos empleados por los codecs que proveen capacidades de compresión, en adición a la digitalización, son extremadamente complejos involucrando formulas matemáticas que van más allá de este estudio. Sin embargo, el mecanismo universal de digitalización, llamado PCM donde sólo se digitalizan datos analógicos sin compresión, es fácil de comprender y consiste de los siguientes pasos:

1. La señal de audio analógica se muestrea a una velocidad fija, digamos 8000 muestras por segundo. El valor de cada muestra es un valor real.

2. Luego, cada muestra se redondea a un valor entero. A esta operación se la denomina cuantización. El total de valores corresponde a una potencia de 2, como por ejemplo 256 valores de cuantización.

3. Cada valor de cuantización se representa con una cantidad fija de bits. Por ejemplo, si se tienen 256 valores posibles de cuantización, entonces cada uno se representa por 1 byte. Cada cuantización se representa en binario y todas las representaciones en binario se concatenan formando una señal de audio digital.

Como un ejemplo del proceso antes descripto, se supone una señal muestreada a una frecuencia de 8000 muestras por segundo. Cada muestra es cuantizada y representada con 8 bits, dando como resultado una señal digital de 64000 bits por segundo. Esta señal puede decodificarse o convertirse en una señal analógica para su reproducción. Sin embargo, la señal analógica decodificada es diferente de la señal originalmente codificada. Incrementando la velocidad de muestreo y el total de valores de cuantización, la calidad de la señal decodificada se aproxima a la original. Aquí

puede notarse que la calidad de la señal decodificada y los requerimientos de almacenamiento son directamente proporcionales.

Generalmente se utiliza PCM para la codificación de voz, con una frecuencia de muestreo de 8000 muestras por segundo y 8 bits por muestra, es decir 64 Kbps. Los discos compactos también utilizan PCM pero la frecuencia de muestreo es de 44100 muestras por segundo y 16 bits por muestra; dando una frecuencia de 705,6 Kbps para mono y 1,411 Mbps para stereo.

Si se analizan los requerimientos de ancho de banda para los valores mencionados en los ejemplos anteriores, puede notarse que la transmisión de música en stereo excede el ancho de banda existente al momento. Incluso los 64 Kbps necesarios para la transmisión de voz exceden las capacidades de un usuario dial-up. Por estas razones PCM no es una opción adecuada para la codificación de audio en Internet. Alternativamente existen técnicas de compresión que reducen la cantidad de bits necesarios para una señal de audio digital.

Entre las técnicas de compresión de audio más populares se encuentran:

Codec	Requerimientos
GSM	13 Kbps
G.711 (PCM)	64 Kbps
G.728	16 Kbps
G.729A	8,5 Kbps
G.723.1	6,4 y 5,3 Kbps

Además de los codecs mencionados existen otros propietarios como el provisto por RealNetworks y utilizado en su aplicación por excelencia: realplayer. Otra técnica de compresión de audio revolucionaria y popular que provee calidad de música en stereo cercana a la provista en CDs es MPEG Layer 3, o también denominada MP3. MP3 comprime música utilizando 128 o 112 Kbps sin degradación distinguible para el oído humano. Una de las ventajas provistas por este codec es la capacidad de particionar un archivo en pequeños fragmentos, cada uno de ellos reproducible en forma independiente. Este formato de archivo lo convierte en una excelente opción para las transmisiones de streaming a través de Internet. Por otro lado, el estándar de compresión utilizado en MP3 es extremadamente complejo, lo que lo hace caro en términos computacionales e inadecuado para transmisiones en tiempo real por el tiempo necesario para su codificación.

Es importante remarcar que los codecs mencionados antes salvo MP3 son buenos candidatos para la compresión de voz ya que no requieren grandes capacidades de cómputo ni una calidad alta para estar a la par de una conversación telefónica.

Codecs de video

Un video es una secuencia de imágenes, reproducidas a una velocidad constante, como por ejemplo 24 o 30 imágenes por segundo. Una imagen digital sin compresión puede verse como una matriz de pixeles, donde cada pixel es codificado con una cierta cantidad de bits representando un color. En la digitalización de videos existen dos tipos de redundancias que se explotan para lograr compresión de datos. Una corresponde a la redundancia espacial que se refiere a la redundancia en una imagen. Por ejemplo, una imagen compuesta de zonas similares puede comprimirse

fácilmente. El otro tipo de redundancia llamada redundancia temporal se refiere a una sucesión de imágenes. Por ejemplo, es factible que una imagen y las subsecuentes coincidan debido al escaso movimiento en el video.

MPEG es el estándar de compresión de video más popular al momento y se clasifica de la siguiente forma:

Codec	Descripción	Requerimientos
MPEG 1	Calidad CD-ROM	1,5 Mbps
MPEG 2	Calidad DVD	3-6 Mbps
MPEG 4	Calidad DVD	910-3000 Kbps

Al igual que sucede con MP3 para la codificación de audio, MPEG no es una buena alternativa de codificación para transmitir video en tiempo real debido al excesivo tiempo requerido para su conversión a este formato. Sin embargo, por sus cualidades en cuanto al aprovechamiento del ancho de banda, MPEG se utiliza en aplicaciones de streaming.

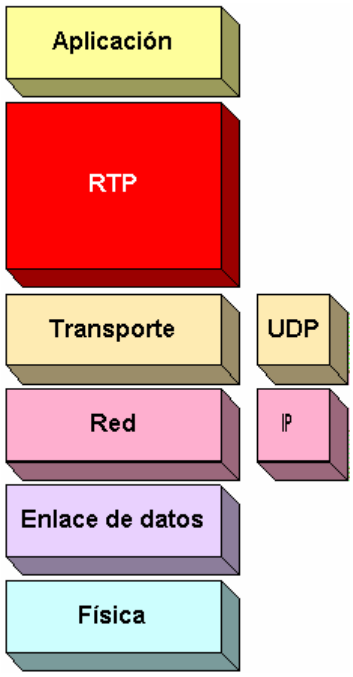
Existen numerosas alternativas a MPEG siendo la mayoría propietarias. Por ejemplo, RealNetworks provee el formato Real Movie, Apple Quicktime, Microsoft Windows Media, entre otros. Pero ninguno de estos codecs sirve para la codificación en tiempo real.

La opción estándar propuesta por H.323 para intercambio de video en tiempo real es el codec H.261, cuyos requerimientos de ancho de banda y tiempo de codificación lo convierten en la opción adecuada.

Capítulo 7: Protocolos de transporte de datos multimedia

RTP

El transporte de datos multimedia en tiempo real, ya sea para aplicaciones de streaming o conferencias, está cubierto por el estándar RTP. Este protocolo está definido en la RFC 1889 de la IETF ubicándose por sobre la capa de transporte del modelo OSI.



Generalmente RTP corre sobre UDP debido a sus ventajas en cuanto a velocidad respecto a TCP, ya que este último realiza control de flujo y errores.

Como el principal problema de las aplicaciones con requerimientos de tiempo real que corren en red es el retardo de paquetes, RTP provee datos útiles a las aplicaciones que lo utilizan para combatir retardos y pérdida de paquetes. Sin embargo, es importante destacar que RTP no garantiza calidad de servicio y no realiza reserva de recursos a lo largo del camino de su conexión.

Ejemplos de aplicaciones que utilizan RTP son:

- Aplicaciones de audio y video interactivas
- Streaming de audio y video
- Streaming de audio y video en broadcast

El protocolo no está limitado sólo a aplicaciones unicast, sino que además se permiten conexiones multicast.

RTP asigna a cada fuente multimedia, como por ejemplo un micrófono o cámara, un flujo independiente de paquetes RTP. Entonces para el caso de una videoconferencia entre dos personas, se necesitan cuatro flujos RTP: dos para la transmisión del audio, uno para cada dirección, y dos para la transmisión de video,

también uno en cada dirección. A pesar de ello, no siempre se utilizan flujos independientes de audio y video ya que algunas técnicas de codificación, como MPEG-1 y MPEG-2, utilizan un único flujo para audio y video.

La estructura del paquete RTP es la siguiente:

0			7
V	**P**	**X**	**CSRC count**
M	**Tipo de dato**		
Número de secuencia (2 bytes)			
Timestamp (4 bytes)			
SSRC (4 bytes)			
CSRC (0-60 bytes)			

El campo más importante de los paquetes RTP es el tipo de dato, el cual indica qué codec se va a utilizar para enviar datos. El valor de este campo se define en la etapa de señalización. Sin embargo, durante la conexión alguno de los participantes puede cambiar la codificación notificando al otro extremo a través de este campo. La siguiente tabla, muestra algunos valores válidos para este campo:

Valor del campo	Formato de Audio
0	PCM μ-law
1	1016
3	GSM
7	LPC
9	G.722
14	MPEG Audio
15	G.728
26	Motion JPEG
31	H.261
32	MPEG-1 video
33	MPEG-2 video

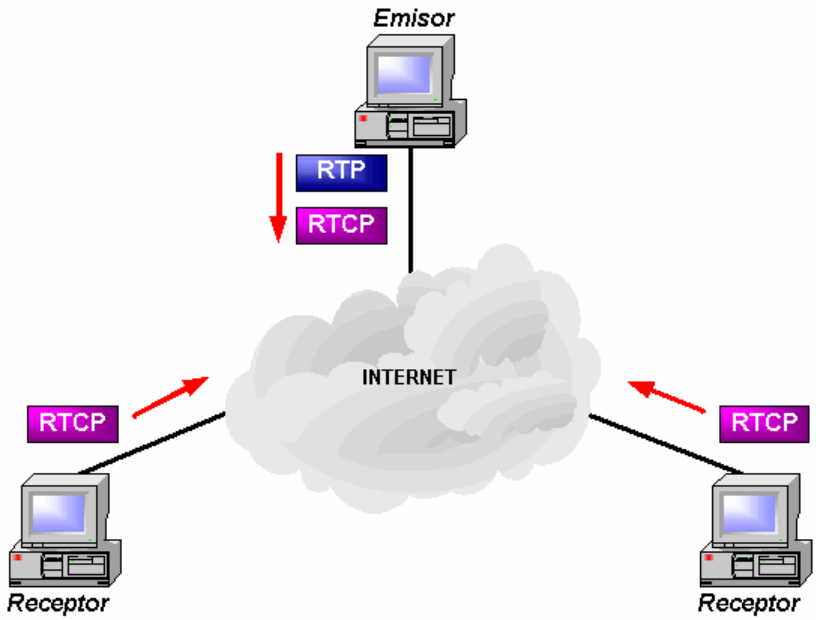
Otro campo importante es el número de secuencia. Este valor de 16 bits se incrementa para cada paquete enviado, y debe ser utilizado por el receptor para detectar pérdida de paquetes y restaurar la secuencia de los mismos. La razón de estos controles radica en el uso de UDP como protocolo subyacente, no contando con controles de flujo o errores. Ante la detección de errores o pérdida de paquetes, existen diversas técnicas de corrección como por ejemplo, para el caso de video, agregar una imagen de color pleno en sustitución del paquete perdido, o interpolación de imágenes.

Existen otros dos campos de importancia dentro de los paquetes RTP; el timestamp que refleja el instante en el que se generó el primer byte del dato multimedia y el SSRC el cual se elige aleatoriamente cuando se crea una conexión para poder identificar un flujo de datos RTP de otro.

Por ultimo, el CSRC describe una lista de orígenes que contribuyeron a un stream combinado de datos generado por un mixer RTP. Un ejemplo de uso es en una conferencia de audio donde un mixer identifica a todos los participantes cuyos discursos fueron combinados para producir el paquete generado, permitiendo al que recibe dicho paquete identificar a todos los participantes por medio de esta lista, mientras que todos los paquetes contienen el mismo SSRC.

RTCP

El protocolo de control que acompaña a RTP en la ***RFC 1889*** *es* RTCP. RTCP no es un protocolo de señalización. RTCP está basado en la transmisión periódica de paquetes de control de manera de proveer i nformación estadística, en tiempo real, a todos los participantes de una conferencia.



Como se muestra en la figura, los paquetes RTCP son transmitidos por cada participante al resto.

Los paquetes RTCP no encapsulan datos multimedia de audio o video como es el caso de RTP, sino que son paquetes que se envían periódicamente con reportes de emisores y/o receptores anunciando estadísticas que pueden utilizarse por las aplicaciones. Estos datos estadísticos, incluyen delay, jitter, número de paquetes enviados, paquetes recibidos y paquetes perdidos. El estándar no especifica qué es lo que la aplicación debe realizar con esta información. Las aplicaciones podrían cambiar la codificación sacrificando calidad y así aumentar la tasa de transferencia.

RTCP permite a los receptores de alguna conexión RTP generar reportes indicando la siguiente información:

- SSRC: reporte correspondiente a la sesión RTP para la cuál se está generando un reporte.

- Fracción de paquetes RTP perdidos: cada receptor calcula el número de paquetes RTP perdidos dividido por el total de paquetes enviados.

- Último número de secuencia recibido: corresponde al último número de secuencia recibido en el flujo de paquetes RTP.

- Jitter: retardo entre arribos, calculado como el promedio de tiempos de arribo entre los sucesivos paquetes RTP.

Estos reportes son enviados a todos los participantes utilizando multicast.

Por su parte, cada emisor de paquetes RTP también genera reportes de envío RTCP. Estos paquetes contienen información relacionada al flujo RTP, incluyendo:

- SSRC correspondiente a la sesión RTP.

- El timestamp del paquete RTP generado más recientemente.

- El número de paquetes enviados hasta el momento.

- El número de bytes enviados hasta el momento

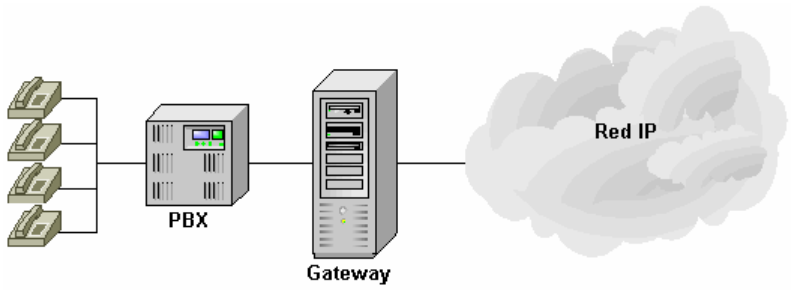
Estos reportes pueden utilizarse para asociar diferentes flujos multimedia de un conjunto de sesiones RTP. Por ejemplo, si en una videoconferencia con más de dos participantes uno de ellos se desconecta en forma abrupta, éste puede reestablecerse utilizando los reportes RTCP. Estos reportes identifican quiénes participan de la videoconferencia y qué sesiones RTP se mantienen con el resto de los conferencistas, de modo de poder sincronizar audio y video con cada uno de ellos.

Por último quedan mencionar los paquetes de descripción del origen RTCP cuya funcionalidad es informar datos sobre el origen como dirección de correo electrónico, el nombre del emisor y la aplicación que genera el flujo RTP. Además se incluye el SSRC asociado a la conexión RTP. Estos paquetes proveen un mecanismo de mapeo entre el identificador de origen o SSRC y el nombre de la máquina y/o usuario.

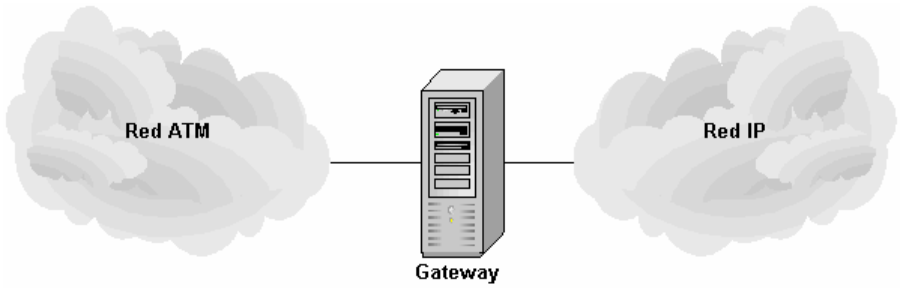
Finalmente cabe destacar que los paquetes RTCP pueden apilarse, es decir que ante la necesidad de enviar los tres tipos de reportes, éstos pueden concatenarse en un único paquete UDP.

Capítulo 8: Gateways

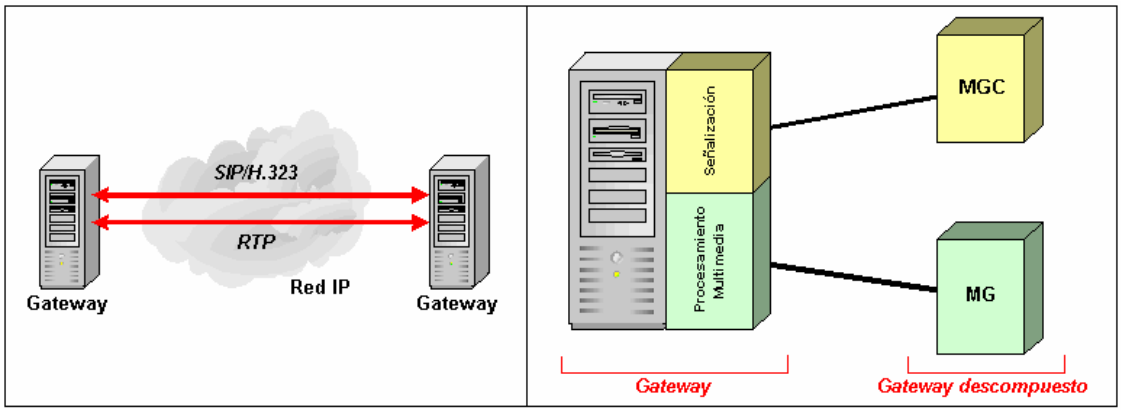
Estas entidades son las encargadas de conectar dos tipos de redes diferentes. Para el caso particular de las redes telefónicas, un gateway se ubicaría entre el dominio conmutado de la PSTN y el dominio de paquetes de una red IP por ejemplo. Un gateway de esta clase podría conectar un teléfono analógico, una PBX o un conmutador de telefonía a una pc en una red IP.



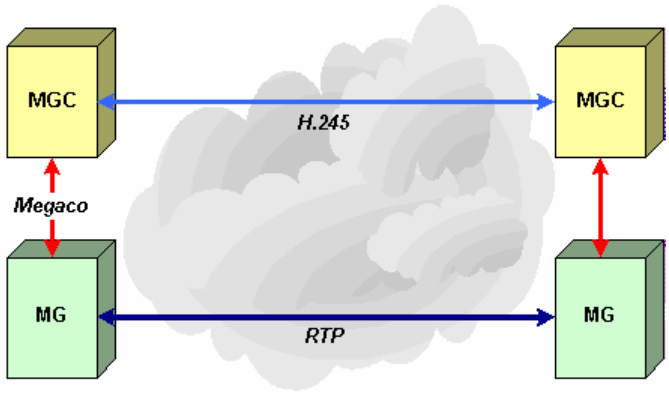
También existen gateways que conectan dos dominios de paquetes diferentes; por ejemplo, un dispositivo podría trabajar con voz sobre ATM en uno de sus dominios y VoIP sobre el otro. Incluso un dispositivo que efectúe la transcodificación entre dos algoritmos de compresión pero mantenga el mismo mecanismo de transporte, es considerado un gateway.



En algunas implementaciones se descompone la tarea del gateway en funciones distribuidas entre diferentes entidades. Este no es el caso de un gateway convencional H.323 o SIP, donde el mismo dispositivo provee conversión de voz analógica a paquetes, además de traducción de mensajes de señalización entre los diferentes dominios.



Los gateways distribuidos, dividen las funciones de señalización de las funciones de manipulación de datos multimedia. En este tipo de gateways, las funciones que trabajan sobre los datos multimedia se asigna a dispositivos llamados media gateway o MG y el manejo de señalización a otra clase de dispositivos denominados media gateways controllers o MGC. Estos últimos son también vulgarmente conocidos como call agent, softswitch. Entre estas nuevas entidades, MGC y MG, existe una relación maestro / esclavo respectivamente. Sin embargo, tanto entre MGs – intercambio de datos multimedia – como entre MGCs – intercambio de datos de señalización – existe una relación par a par.



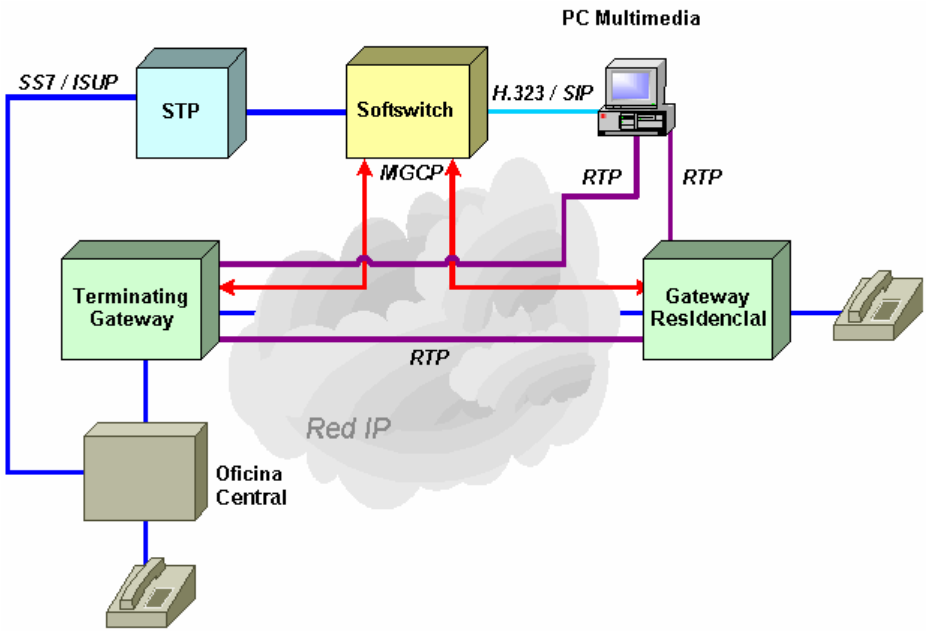
Como puede observarse en la figura anterior, los protocolos entre MGs puede ser cualquiera de los formatos de paquetes para el envío de datos de audio y video, que comúnmente son:

- RTP para redes IP
- AAL1/AAL2/AAL5 para redes ATM
- FRF.11 para redes Frame Relay

Por su parte, entre MGCs puede utilizarse cualquiera de los siguientes protocolos de señalización:

- H.323
- SIP

Existen diversas razones por las cuales se dividieron las funciones de los gateways. Un ejemplo de ello son los grandes gateways de telefonía con una inmensa cantidad de puertos TDM[1] de entrada y salida. Dichos gateways poseen requerimientos de cómputo importantes para efectuar el proceso de señalización, que es intensivo en procesamiento pero escaso en tareas de entrada / salida. Por esta razón es más eficiente mover el código de procesamiento intensivo de señalización a máquinas de bajo costo, mientras que las tareas de traducción de datos multimedia se delega a hardware de propósito especial como son los MGs. Otra de las razones de esta división, es la escalabilidad, confiabilidad y bajo costo que otorga esta solución.



Como puede apreciarse en el gráfico, el gateway residencial podría estar ubicado en una pequeña empresa. Aquí surge la disyuntiva respecto de la confiabilidad en las tareas de señalización cuando queda en manos de un usuario final. Como alternativa, el usuario recibe un MG como dispositivo hogareño mientras que el MGC reside en manos de un operador confiable.

A continuación se describen brevemente los protocolo de comunicación utilizados entre el MG y el MGC como son MGCP y Megaco.

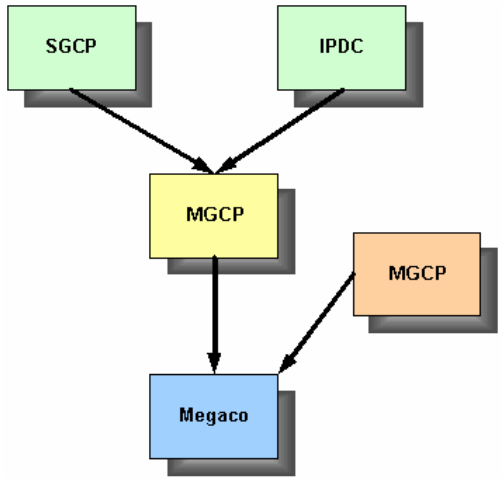
MGCP

MGCP o Media Gateway Control Protocol, se define en la RFC 2705 [Ref. MGCP] correspondiente a la IETF. Este protocolo surge de la combinación de otros dos protocolos anteriores:

- IPDC o IP Device Control, desarrollado por Level 3 Communications
- SGCP o Simple Gateway Control Protocol, desarrollado por Bellcore y Cisco

[1] TDM: Multiplexación por división del tiempo

Tanto MGCP como IPDC y SGCP describen la interfaz entre MG y MGC.



Como muestra la figura, IPDC y SGCP surgen en forma independiente y con el mismo propósito. Sin embargo, IPDC tuvo una influencia comercial mucho más fuerte que SGCP, mientras que SGCP se mostró más robusto técnicamente. Entonces el tiempo llevó a la combinación de ambas perspectivas rescatando las mejores características de ambos. Es de esta forma como surge MGCP.

Durante el desarrollo de MGCP, se enfatizó en aspectos de simplicidad y confiabilidad, y actualmente, existen varias implementaciones completas de MGCP.

Megaco

Independientemente de MGCP, la ITU-T comenzó a trabajar en la definición de un protocolo similar. Mientras tanto, Lucent Technologies, contribuyó con otro protocolo de características similares que dio a conocerse como MDCP o Media Device Control Protocol.

A raíz de este hecho y a través de un acuerdo histórico entre la ITU y la IETF, se decidió colaborar en la definición de un único protocolo para el control de media devices. Utilizando MGCP y MDCP como punto de partida, surgen el grupo de trabajo MeGaCo, perteneciente a la IETF, y el grupo de estudio 16 de la ITU-T para comenzar a trabajar en la definición del nuevo protocolo. Los resultados de este estudio pueden hallarse en la RFC 3015 de la IETF[Ref. MeGaCo], como en la recomendación H.248 de la ITU-T[Ref. H.248]. Anecdóticamente ambos documentos son idénticos salvo por su organización.

Parte III – Casos de prueba

Al analizar los productos VoIP disponibles, se encontraron una infinidad de opciones con distintas características. Sin embargo, a groso modo es posible categorizar los productos según la implementación adoptada en cada caso. Una gran cantidad de productos ofrecen servicios VoIP en forma propietaria, es decir no utilizan ninguno de los estándares antes descriptos. Por otra parte existen productos que se apegan a alguno de los estándares o incluso a más de uno.

Como la finalidad de este trabajo es el análisis de estándares y la integración entre los mismos, las pruebas sólo consideraron productos que respetan los estándares más importantes, es decir, SIP y H.323.

Tal vez el éxito de estos protocolos, tiene una estrecha relación con el apoyo dado por empresas cuyo alcance es masivo, imponiendo restricciones en el mercado y finalmente marcando tendencias. Tal es el caso de productos como Netmeeting o Messenger de Microsoft que utilizan H.323 y SIP, respectivamente.

La lista de productos para VoIP es gigantesca, existiendo componentes de hardware y software con capacidades similares. Analizando la variedad de productos, existen diferencias marcadas que permiten una clasificación de los mismos. Por un lado están las terminales utilizadas por los usuarios finales, como es el caso de los teléfonos IP o cualquier software que permita comunicaciones multimedia. Por otro lado está la gama de servidores VoIP, ofreciendo servicios de comunicación entre las partes.

La gran mayoría de productos VoIP, comparten un factor de gran importancia como es el costo. Todos los productos comerciales de este tipo son excesivamente costosos ya sean en versiones de hardware o software, y la razón de este hecho es asegurar a sus clientes una reducción máxima de los gastos en telefonía. También existe una reducida línea de productos open source que ofrecen similares servicios a los productos comerciales, pero la puesta a punto es mucho más costosa.

En general, las soluciones comerciales VoIP prestan excelentes servicios a los usuarios finales, incluso al administrador ya que facilita sus tareas. Sin embargo, esta alternativa se vuelve inalcanzable para una pequeña o mediana empresa debido a los elevados costos que deben afrontar en licencias y dispositivos. Es esencialmente por esta razón, que para este trabajo consideramos la evaluación de alternativas open source para poder hacer un deployment SIP y H323 tratando de maximizar la calidad de los servicios ofrecidos a los usuarios finales, siempre teniendo en cuenta que las tareas de administración se volverían más complejas, sobre todo a la hora de integrar productos de diferentes empresas.

Para poder definir un ámbito de pruebas que satisfaga nuestras expectativas, se siguieron los siguientes pasos:

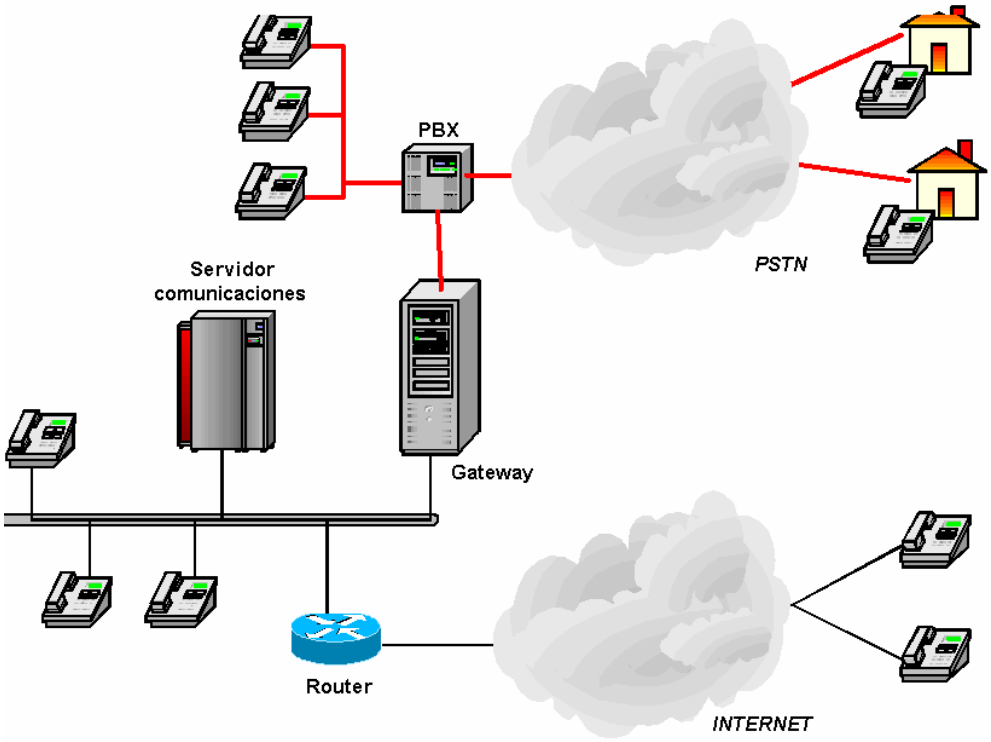
1. Evaluación de alternativas open source
2. Deployment de una zona SIP
3. Evaluación de servicios SIP
4. Deployment de una zona H323
5. Evaluación de servicios H.323

6. Integración de ambas zonas

El deployment de las zonas SIP y H.323 será progresivo, es decir, se partirá de zonas extremadamente simples, como es el caso de dos clientes conectados directamente, y a medida que se vaya avanzando en las pruebas se irán agregando nuevas entidades. Los proveedores de servicios de cada una de las zonas, se corresponden con alguna alternativa open source que implemente el conjunto de servidores SIP o H.323, mientras que los clientes pueden ser en algunos casos comerciales, ya que varias empresas publican demos de este tipo de productos.

Los puntos más ricos en contenido práctico, y que nos permitieron llegar a conclusiones precisas, corresponden a las evaluaciones de los servicios ofrecidos por SIP contrastándolos con los propios de H.323. Básicamente, en cada caso se evaluaron capacidades definidas en el estándar, interacción con el usuario y proveedores de servicios de diferentes vendedores, ya sean open source como demos comerciales.

Finalmente, al completar el análisis de cada protocolo, veremos cómo es posible integrarlos entre sí e incluso con la PSTN.



Hardware utilizado

Muchos de los servicios otorgados por las aplicaciones VoIP, sean clientes o servidores, se pueden presentar en dos formatos:

- ✓ Hardware VoIP
- ✓ Software VoIP

Ambas alternativas ofrecen iguales servicios, siempre teniendo en cuenta las limitaciones de un hardware, en cuanto a actualizaciones, mientras que la segunda opción es más versátil.

Claramente un producto implementado en hardware es comercial, mientras que una implementación por software sobre una PC o workstation puede ser una alternativa comercial o gratuita.

En nuestro caso contamos con varias PCs Windows y Linux, restringiendo las pruebas sobre aplicaciones de software VoIP. En particular contamos con una placa especial que permite enriquecer las pruebas de interacción con la PSTN. Esta placa ISA es provista por la empresa ***Quicknet's Technologies***, y el nombre del producto es ***Internet Linejack***. La característica de esta placa es que funciona como gateway telefónico / fax entre la PSTN y una red IP.

Además de las PCs, contamos con un dispositivo de Cisco Systems correspondiente a su línea más económica: ATA 186. Este producto, posee tres interfaces: dos RJ-11 y un RJ-45 permitiendo conectar dos teléfonos analógicos a una red IP. La función provista por el ATA, es la de hacer que un par de teléfonos analógicos interactúen como teléfonos IP, utilizando tanto SIP como H.323.

Capítulo 9: Pruebas de productos SIP

Servidores

SIP define una serie de servidores que son necesarios para un desempeño comparable al ofrecido por una central telefónica. Estos servidores, muchas veces se implementan dentro el mismo dispositivo o aparecen distribuidos en la LAN como servicios diferentes.

Durante la búsqueda de productos libres, se encontraron varias librerías disponibles en diferentes lenguajes que permiten la implementación del stack SIP. Sin embargo, no existen muchas alternativas ya armadas que implementen estos servicios. Sólo se encontraron dos productos:

- ✓ Vocal de Vovida
- ✓ SER de Iptel

Ambos productos implementan el protocolo SIP utilizando el lenguaje C y C++ para su desarrollo. En ambos casos, los servicios brindados se corresponden con las siguientes entidades del estándar:

- ✓ Proxy Server
- ✓ Redirect Server
- ✓ Location Server

Además de proveer estos servicios, anexan con la distribución una interfaz amigable de configuración que básicamente permite administrar los usuarios.

Vocal

Vocal es una suite de servidores distribuidos que componen un framework de voz sobre IP utilizando principalmente SIP para comunicar entidades, además de gateways MGCP y H.323. Vocal también admite teléfonos analógicos a través de gateways residenciales, como por ejemplo el ATA 186 de Cisco Systems.

Desde una perspectiva generalizada, puede verse al sistema Vocal como un conjunto de componentes básicas, a decir:

- ✓ Sistema Vocal: inteligencia de la aplicación de telefonía. Se compone de un conjunto de servidores.
- ✓ Interfaz de usuario: facilita la administración de usuarios y servicios así como también el monitoreo del sistema.

- ✓ Teléfonos IP: Vocal soporta la interacción con diversos clientes SIP ya sean implementados por software o hardware.

- ✓ Traductores: las aplicaciones MGCP o H.323, requieren de una traducción de sus mensajes en mensajes SIP para así poder interactuar con Vocal.

- ✓ Gateways: Vocal no sólo provee traducción entre protocolos, sino que además permite la interoperabilidad con otras redes. Vocal trabaja con dos tipos de gateways:

  - ❑ Residential Gateways: traducen señales analógicas en paquetes IP, permitiendo a teléfonos analógicos realizar y recibir llamadas SIP.

  - ❑ Trunking Gateways: permiten intercambiar llamadas entre una red SIP y abonados de la PSTN, traduciendo mensajes SIP en alguna de las siguientes señales:
    - Analógicas
    - Señalización de canal asociado (CAS)
    - PRI

Entrando un poco más en detalle, el sistema Vocal se compone de una variedad de servidores que pueden trabajar en forma distribuida:

- ✓ Marshal Server: es una implementación del servidor Proxy definido por SIP y actúa como el punto de contacto inicial de todo tipo de señalización. Entre sus capacidades provee autenticación, reenvío y facturación.

- ✓ Redirect Server: es una implementación combinada de los servidores de redirección, ubicación y registración definidos por SIP. Este servidor almacena contactos y capacidades para todos los usuarios registrados, así como planes de discado.

- ✓ Administrador de Red: permite al administrador monitorear el sistema a través de SNMP.

- ✓ Servidor VoiceMail: servicio que permite almacenar mensajes de voz y enviarlos por mail cuando el usuario a ser contactado no está disponible.

- ✓ Servidor de capacidades: es otra implementación del servidor Proxy descripto por SIP. Estos servidores proveen capacidades básicas como retransmisión y bloqueo de llamadas.

- ✓ Servidor de administración: almacena registros para cada usuario del sistema y servidores, y distribuye esta información a través de todo el sistema. Para su administración provee una interfaz gráfica basada en web.

Estos servidores pueden instalarse en una misma máquina o en forma distribuida.

Entre el abanico de servicios provistos por el sistema podemos hacer una división de dos categorías según:

- ***Servicios de la entidad llamadora***

  - Bloqueo de llamadas: el administrador o el usuario puede utilizar este servicio para bloquear llamadas de larga distancia o a números de tarifas especiales como ser 0600.

  - Bloqueo de la identificación de llamadas: este servicio puede ser habilitado por el usuario para evitar su identificación en el equipo remoto.

- ***Servicios de la entidad receptora***

  - Retransmitir todas las llamadas: permite redirigir todas las llamadas entrantes a un número específico o a un servicio de mensajería de voz como ser el voice mail.

  - Retransmitir llamadas no atendidas u ocupadas: permite redirigir llamadas entrantes en caso de:
    - El receptor está ocupado, es decir actualmente en una conferencia
    - El receptor no contesta la llamada después de un determinado número de rings.

  - Retorno de llamada: permite al usuario volver a marcar el último número marcado. El usuario marca *69 y automáticamente se llama al último número marcado.

  - Buzón de voz: la entidad llamada puede utilizar el servicio de voice mail en combinación con la retransmisión de llamadas no atendidas u ocupadas para recibir por mail mensajes de voz.

Instalación

Vocal es compatible con las plataformas Linux/i386, Windows/i386 y Solaris/Sparc. Los requerimientos mínimos de instalación en cuanto a hardware son:

- ✓ *Pentium II 480 Mhz*
- ✓ *128 Mb RAM*
- ✓ *Espacio en disco de 1 GB*

Inicialmente se optó por una instalación de un sistema distribuido, dividiendo el Marshal Server del Redirect Server. En este caso de prueba, la captura de paquetes simplificó el análisis del flujo de mensajes involucrados en las llamadas, pero los requerimientos de hardware necesarios para otras pruebas nos llevó a una reinstalación esta vez centralizada.

SER

SIP Express Router o SER, es un servidor open source basado en SIP que puede actuar como servidor registrar, proxy o redirect. Su gran distinción es el alto grado de configuración provisto, permitiendo la creación de políticas de ruteo, admisión de las llamadas, definición de planes de discado y traducción de números. Esta capacidad le otorga flexibilidad de configuración a costa de alta complejidad.

Otra característica importante, es su arquitectura. SER permite agregar nuevas aplicaciones en forma de módulos, los cuales se acoplan extendiendo su funcionalidad. Los módulos que acompañan la distribución brindan:

- ✓ Soporte SNMP
- ✓ Soporte RADIUS
- ✓ SMS gateway
- ✓ SIMPLE/Jabber gateway

Para el manejo de los usuarios del sistema, existe una interfaz web que se distribuye como otra aplicación llamada ***Serweb***, sobre la cual se pueden mantener perfiles de usuario, listas de control de acceso, envío de mensajes instantáneos y registrar las llamadas realizadas y perdidas.

Instalación

SER es compatible con las plataformas Linux/i386 y Solaris/Sparc. En nuestro caso, se montó sobre la plataforma Linux. A diferencia de Vocal, este producto no impone tantas restricciones, demostrando ser un producto más liviano y rápido.

Clientes

Clientes hardware

Producto	Firmware	Descripción
ATA 186	2.15	Es un adaptador que permite conectar dos teléfonos analógicos a la red de datos, de modo de poder integrar dichos teléfonos a una solución de VoIP. Las alternativas son: SIP, H.323 y MGCP

Clientes software

Producto	Plataforma	URL	Características
Kphone 3.0	Linux	http://www.wirlab.net/kphone	✓ Configuración de proxy ✓ Autenticación Digest y Basic ✓ Mensajería instantánea ✓ Llamadas de voz ✓ Llamadas de video ✓ Soporte de NAT (STUN) ✓ Presencia: online/offline
Linphone 0.10.0	Linux	http://www.linphone.org	✓ Configuración de proxy ✓ Autenticación Digest ✓ Llamadas de voz ✓ Soporte de NAT (STUN)
Sipset 0.8.16	Linux / Windows	http://www.vovida.org	✓ Configuración de proxy/Registrar ✓ Autenticación Digest y Basic ✓ Llamadas de voz ✓ Llamadas de video
Windows Messenger 4.6	Windows	http://messenger.microsoft.com	✓ Autenticación Digest y Basic ✓ Mensajería instantánea ✓ Llamadas de voz ✓ Presencia: online/offline

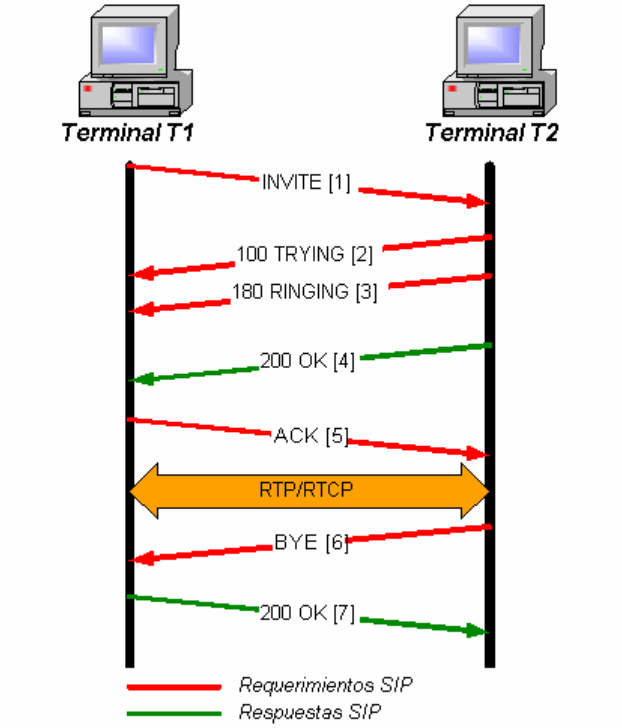
Descripción de las pruebas realizadas sobre una zona SIP

Como se mencionó en la introducción de esta sección, el deployment de esta zona irá creciendo en complejidad a medida que se vayan presentando nuevas entidades.

UA a UA sin registración

El primer caso de prueba, toma dos clientes SIP, los cuales para establecer la llamada telefónica, solo deben discar la dirección IP del otro.

A continuación se detalla el flujo de mensajes intervinientes:



El significado preciso de estos mensajes se detalla en las siguientes líneas:

1. INVITE: el usuario en la terminal 1, levanta el tubo de su teléfono IP y marca el teléfono, en este caso la dirección IP, de la terminal 2. Esto provoca el envío del requerimiento INVITE hacia terminal 2. Este mensaje contiene una oferta SDP describiendo cómo la terminal 1 recibiría RTP.

2. 100 Trying: cuando la terminal 2 recibe el requerimiento INVITE, inmediatamente envía otro requerimiento 100 Trying indicando que se inicia el procesamiento del mensaje anterior.

3. Al recibir un INVITE, Terminal 2 comienza a sonar emitiendo una señal de alerta por llamada entrante. Este hecho es informado a Terminal 1 por medio de el requerimiento 180 Ringing donde además se encapsula una respuesta SDP determinando qué codec se utilizará y cómo recibirá RTP.

4. Cuando el usuario en Terminal 2 contesta la llamada, se envía una respuesta 200 OK a Terminal 1.

5. Al arribar a Terminal 1 la respuesta 200 OK, se envía un requerimiento ACK aceptando la conexión. Luego comienza el flujo de mensajes RTP / RTCP, es decir la conferencia de audio y/o video.

6. Cuando el usuario en la Terminal 2 cuelga el teléfono se envía un requerimiento BYE al otro extremo y ya no se envían nuevos paquetes RTP.

7. Finalmente, la Terminal 1 responde con una respuesta 200 OK aceptando la desconexión.

Si bien es posible realizar llamadas de esta forma, este caso carece de sentido ya que, por ejemplo en un ambiente DHCP sería complicado determinar la IP del destino para poder iniciar una llamada.

Los sistemas basados en SIP usan una red de servidores que permiten realizar el control de las llamadas, localización de usuarios, etc. El uso de un Proxy Server y un Servidor de Registración solucionan el problema anterior, ya que permiten definir usuarios que representen a las personas u oficinas del sistema. Esto permite la movilidad de los usuarios, puesto que se pueden redireccionar las llamadas según el lugar en el cual se registraron.

UA a UA registrados

Este caso supone la comunicación entre dos usuarios registrados ante un servidor Registrar. Como este factor fue conflictivo en las pruebas realizadas, haremos un comentario de cuáles fueron los impedimentos que se nos plantearon. Se distinguen los acontecimientos en ambos productos evaluados:

- ✓ Vocal: la registración de un cliente contra Vocal no presentó problemas salvo en los casos que se utilizó autenticación. Los mecanismos de autenticación provistos por Vocal son básico y MD5. El primer problema se presentó con dos clientes: Linphone y ATA 186, que soportan autenticación digest MD5 como única alternativa y por lo tanto no es posible utilizarlos con autenticación básica. El resto de los clientes probados soportan ambos mecanismos. Sin embargo al utilizar autenticación digest tanto en Kphone como Linphone no se tuvo éxito en una primera instancia.

  Luego de una minuciosa investigación sobre las posibles causas de este problema se determinó que Vocal no utilizaba comillas en el valor del campo proxy-authenticate en los encabezados de mensajes de solicitud de autenticación donde se indican los valores de realm y nonce. Este bug se documentó con el número 544 y fue motivo del patch aplicado al sistema Vocal instalado. Una vez patcheado, Vocal modificó el envío de estos mensajes en forma correcta y Kphone logró autenticarse, mientras que Linphone continuó fallando debido a problemas de implementación en su stack SIP, más conocido como libosip, en lo que hace a la autenticación.

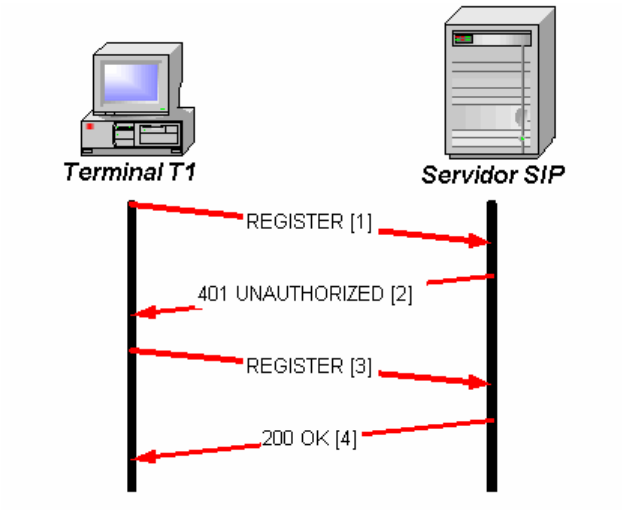
- ✓ SER: el único mecanismo de autenticación en SER es digest MD5, y no se presentaron problemas mayores.

Una vez solucionado el problema de la registración, se probaron capacidades de movilidad de un usuario, es decir: un usuario se registra desde un lugar y recibe llamadas entrantes. Luego cambia su ubicación, se registra y recibe llamadas entrantes en su nueva ubicación.

Las pruebas de movilidad dieron como resultado una diferencia importante de funcionalidad entre Vocal y SER. Por su parte SER, al recibir llamadas dirigidas a usuarios registrados en múltiples ubicaciones, dispara la misma llamada hacia todas las ubicaciones registradas. Si bien este comportamiento por parte del servidor, se apega más a la especificación del estándar, puede no ser un efecto deseable, puesto que muchas veces un usuario olvida desregistrarse y no sería correcto que todas las entidades registradas puedan contestar la llamada. Sin embargo SER, permite modificar su funcionamiento dando como resultado el mismo comportamiento que Vocal.

El caso de prueba planteado en este punto, puede subdividirse en dos casos: comunicaciones entre dos puntos con la presencia de un servidor proxy, o con la ausencia de este servidor.

A continuación se detallan los pasos involucrados en el proceso de registración de una terminal frente a un servidor registrar o Redirect sin la intervención de un proxy:



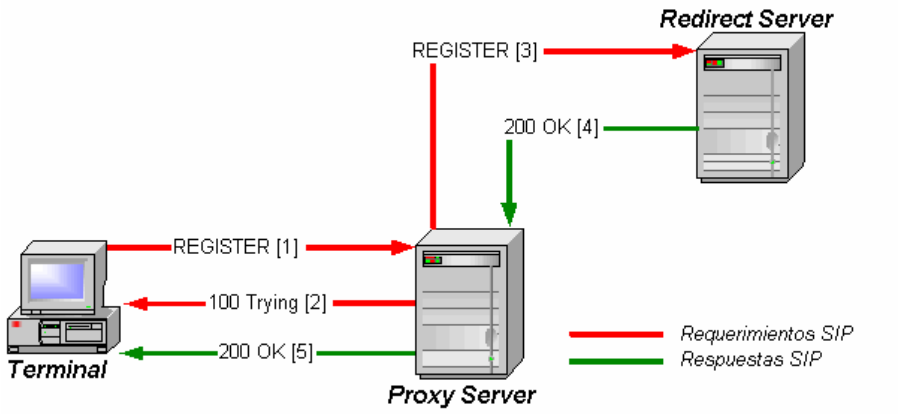
1. La terminal envía un requerimiento REGISTER al servidor SIP.

2. El servidor rechaza la solicitud enviando un mensaje 401 UNAUTHORIZED. Esto no significa que el usuario que está intentando registrarse no lo pueda hacer, sino en realidad se está previendo el tipo de autenticación que se espera, así como otros parámetros necesarios para este proceso.

3. Una vez que se tienen todos los datos necesarios se reenvía la solicitud REGISTER.

4. Una vez que el servidor SIP, autentica al usuario satisfactoriamente, envía una respuesta 200 OK, indicando el resultado exitoso de la registración. En caso contrario volvería a retornar el mismo mensaje que en [2].

El segundo caso corresponde al mismo ejemplo antes descrito pero esta vez con la intervención de un servidor proxy:



1. Inmediatamente que la Terminal es conectada a la red, se envía un requerimiento REGISTER al proxy.

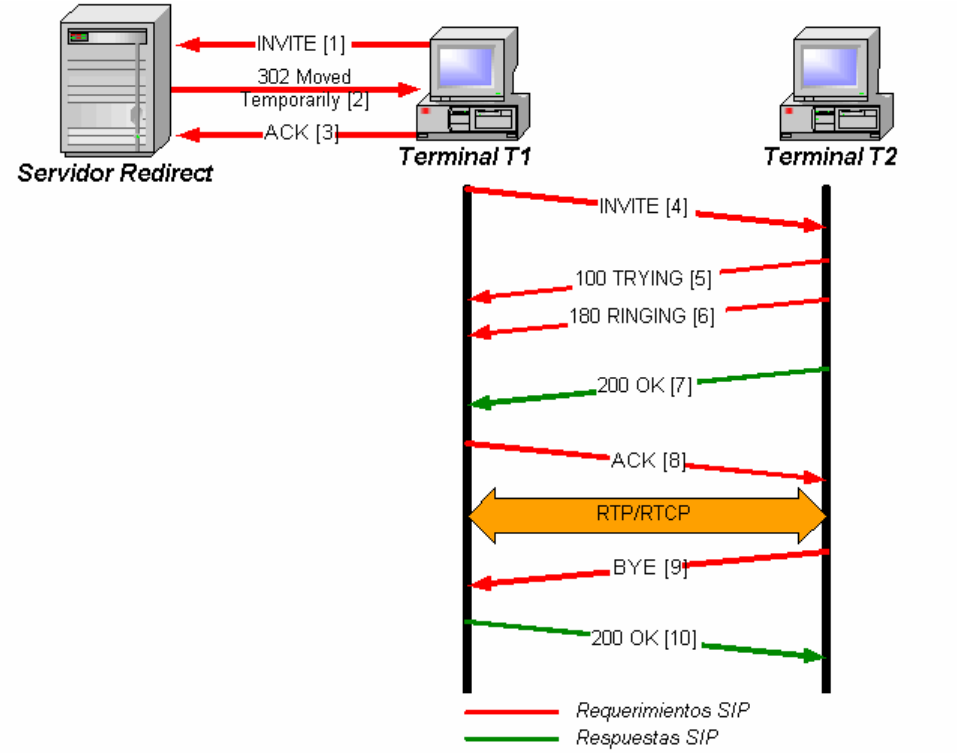
2. Ni bien el proxy recibe este mensaje responde a la terminal con 100 Trying. Esto indica que el proxy recibió el mensaje y lo está procesando.

3. Simplificando el procedimiento de registración, el requerimiento REGISTER se reenvía al servidor Redirect.

4. El servidor Redirect acepta el requerimiento REGISTER enviando una respuesta 200 OK.

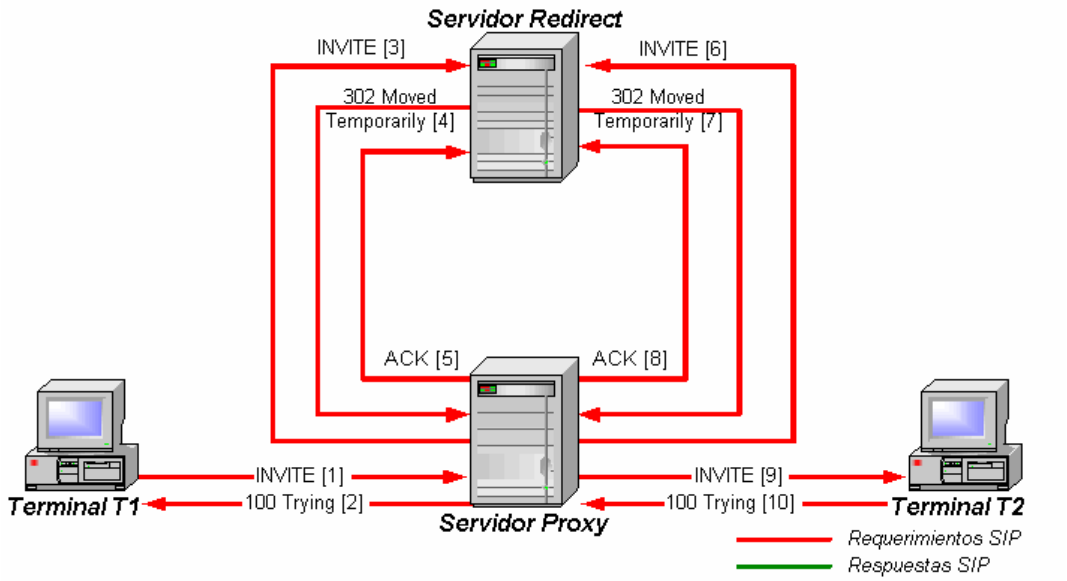
5. El proxy forwardea esta respuesta a la terminal, finalizando el proceso de registración.

Suponiendo registradas dos entidades, se procede al caso de prueba en que se conectan dos entidades registradas. Una vez más, el ejemplo se divide en dos casos, comenzando por el caso en que los participantes no utilizan un servidor proxy:

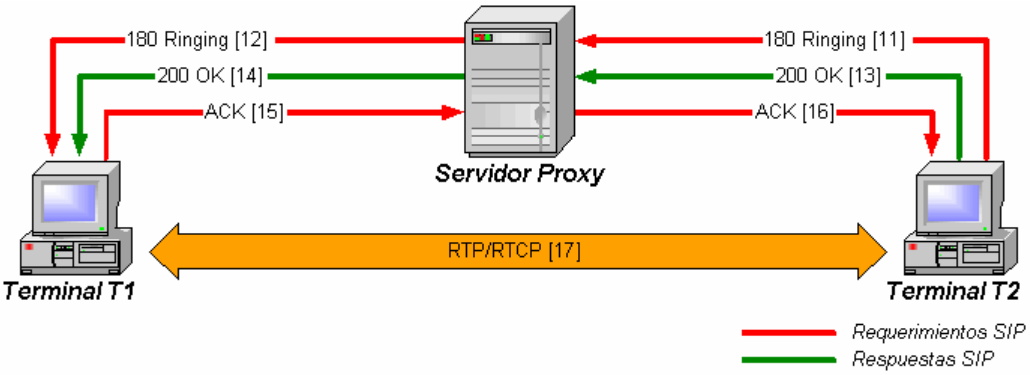


La diferencia entre esta secuencia de mensajes y el caso de comunicación entre dos UAs no registrados, radica en los primeros tres mensajes, donde la Terminal 1 solicita al servidor Redirect la ubicación del usuario llamado. Luego el servidor devuelve la ubicación a la Terminal 1 pudiendo efectivizar la llamada.

Ahora continuamos con las pruebas entre dos entidades que se conectan a través de un servidor proxy. Para simplificar los gráficos dividimos el transcurso de una llamada en tres etapas:

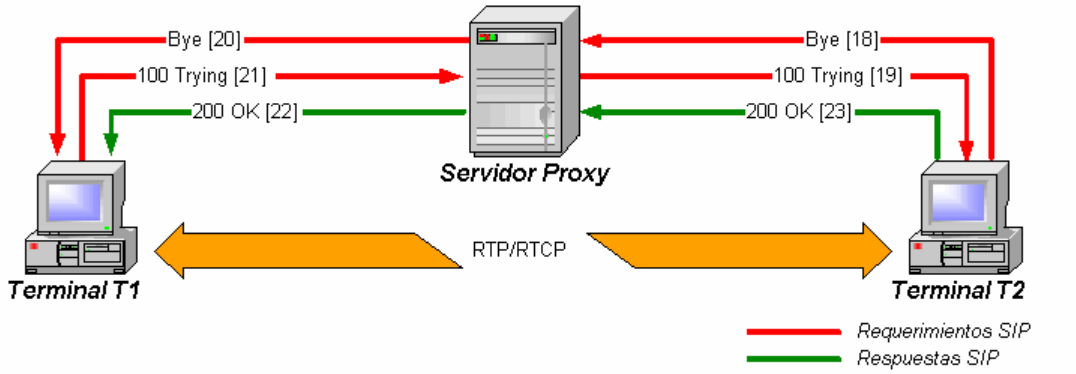


La primer etapa corresponde al inicio de la llamada. Inicialmente, la Terminal 1 envía un requerimiento INVITE al servidor Proxy, generando una consulta al servidor Redirect para ubicar el destino, a pesar de estar conectado directamente a él. Una vez ubicado el destino, se reenvía el INVITE desde el servidor Proxy al destinatario en la Terminal 2.



La segunda etapa se sucede a la recepción del requerimiento INVITE por la Terminal 2. Aquí, el destinatario hace sonar el teléfono alertando al usuario por la llamada entrante correspondiente a la Terminal 1. Además de sonar, Terminal 2 informa al otro extremo de este hecho con un requerimiento 180 Ringing.

Cuando el usuario atiende la llamada en Terminal 2, se envía una respuesta 200 OK hacia Terminal 1 indicando que la llamada fue aceptada. Finalmente se inicia el flujo RTP / RTCP cuando Terminal 1 envía una aceptación ACK devuelta a Terminal 2.

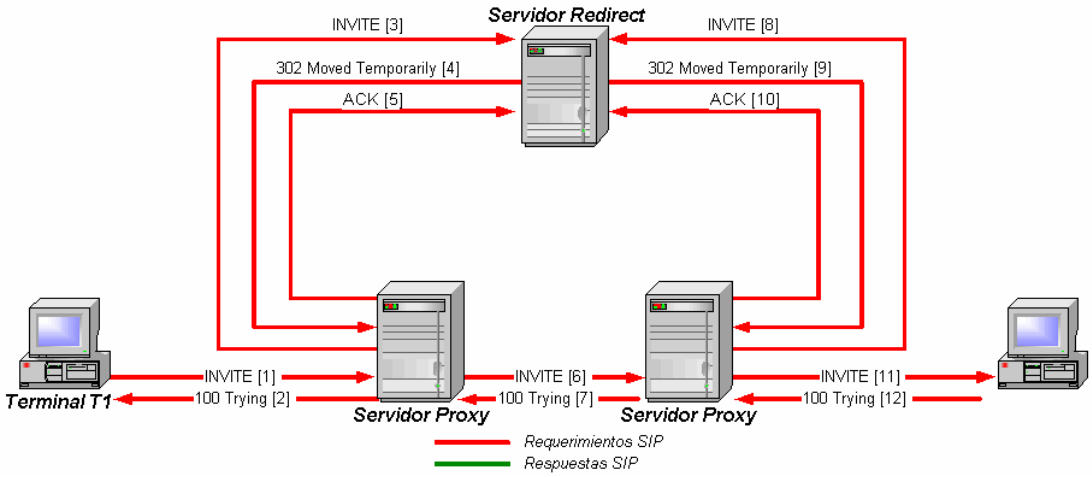


La última etapa corresponde a la descjonexión de la llamada. En el ejemplo suponemos que el usuario en Terminal 2 inicia la desconexión.

UA a UA conectados a servidores proxy diferentes

Esta prueba surgió principalmente como un test de integración de productos. Básicamente se desea conectar dos teléfonos SIP a dos servidores Proxy diferentes. Para ello, se utilizaron dos Proxys correspondientes a diferentes vendedores: en un extremo se utilizó Vocal, mientras que en el otro extremo, SER.

La diferencia más importante se da en la fase de solicitud de conexión en una llamada. El siguiente gráfico muestra la secuencia de mensajes que se suceden



Servicios adicionales

Particularmente vocal, hace hincapié en este punto. Son capacidades adicionales provistas a las entidades que inician llamadas como a las que reciben. A continuación se enumeran estos servicios adicionales, clasificándolos según el tipo de servicio ofrecido:

Servicios ofrecidos al llamador

- Bloqueo de llamadas
- Bloqueo de identificación de llamadas
- Retorno de llamada

Servicios ofrecidos al receptor

- Redirigir todas las llamadas
- Redirigir llamadas cuando esté ocupado
- Redirigir llamadas cuando no conteste
- Retorno de llamada
- Bloqueo de llamadas recibidas
- Voice Mail

Observaciones

Durante la implementación de los casos de prueba en una zona SIP, se presentaron una gran variedad de dificultades.

Descartando las dificultades de instalación y configuración, la mayor parte de los problemas se dieron por el anarquismo existente en las diferentes implementaciones de los protocolos involucrados en una sesión SIP, donde no todos los vendedores respetan la RFC, otorgando servicios vistosos a los usuarios a costa de perder compatibilidad.

Esto suena un tanto duro, pero comercialmente sería una ventaja puesto que obligaría a los clientes de estas empresas a mantener una línea de productos del mismo vendedor si se desea utilizar un grupo reducido de servicios.

Capítulo 10: Pruebas de productos H.323

Servidores

Al igual que para SIP, se analizaron alternativas de código abierto que implementan el conjunto de protocolos propios de H.323. El resultado de esta búsqueda, centralizó nuestros estudios sobre un proyecto emprendido por la empresa australiana Equivalence Pty. Ltd., el proyecto Open H.323. Este proyecto implementa una serie de librerías correspondientes al conjunto de protocolos descriptos por la ITU.

A diferencia de otros paquetes, el proyecto Open H.323 no instala ninguna aplicación H.323 más que un par de ejemplos simples con fines de testeo. En realidad esta instalación registra una serie de librerías en el sistema operativo que luego son utilizadas por otros productos, sean estos libres o comerciales.

Es por esta razón que a continuación se introducirá brevemente qué es el proyecto Open H.323, y luego se procede a la descripción de las aplicaciones que lo utilizan.

El proyecto Open H.323

Como ya se mencionó, este proyecto provee librerías que implementan el conjunto de protocolos H.323. A pesar de ser un producto coordinado por una empresa comercial, Equivalence Pty. Ltd., es una distribución open source que puede utilizarse tanto para aplicaciones libres como comerciales.

Instalación

Las librerías provistas por Open H.323 son compatibles con los sistemas operativos más frecuentes: Linux, Windows, y cualquier tipo de UNIX. En nuestro caso, se escogió Linux como candidato para realizar las pruebas. En cuanto a los requerimientos de hardware, la capacidad de procesamiento no es fundamental sino la cantidad de RAM disponible, necesitándose al menos 256 Mb de RAM y 128 de swap.

La particularidad de este producto es que se ayuda de una librería llamada PWLIB. Esta librería se creó hace varios años con el objetivo de producir aplicaciones que corrieran tanto en Windows como en sistemas X-Windows. También existió la intención de ampliar la portabilidad a Macintosh pero no prosperó. Dado al gran crecimiento que tuvo PWLIB, la portabilidad de ventanas gráficas se convirtió en una componente más del conjunto, ofreciéndose portabilidad en otros aspectos tales como:

- ✓ Entrada / Salida
- ✓ Multithreading, para la implementación de demonios en UNIX y servicios en Windows
- ✓ Varios protocolos de Internet

Volviendo a la instalación de Open H.323, es necesario seguir los siguientes pasos en orden:

- ✓ Instalación de PWLIB
- ✓ Instalación de Open H.323

Las opciones de compilación de ambos productos son simples, siendo la opción más importante la que permite seleccionar el tipo de librerías a generar, es decir, librerías dinámicas o estáticas.

A continuación se describen los productos que utilizan estas librerías para implementar diferentes entidades H.323

Gatekeepers

Como se ha mencionado en el análisis de las entidades H.323, el gatekeeper representa la entidad más importante en una zona H.323 y por esta razón la selección de una implementación debería considerar completitud en cuanto a servicios contemplados.

En la búsqueda de productos que implementen la funcionalidad de esta entidad, se hallaron tres opciones:

- ✓ OpenGatekeeper de Egoboo
- ✓ OpenGk de Equivalence
- ✓ OpenH323 Gatekeeper o GNU Gatekeeper

El producto OpenGatekeeper de Egooboo se encuentra actualmente discontinuado, razón por la cual se descartó su utilización en las pruebas realizadas. Por su parte el OpenGk provisto como parte del proyecto Open H.323 es una implementación precaria del gatekeeper por ser éste simplemente una herramienta con propósitos de testeo.

Finalmente, se llegó al análisis del GNU Gatekeeper, cuyas cualidades son excelentes por hacer un buen aprovechamiento de las funciones provistas por las librerías Open H.323, ajustarse al estándar y ser estable en cuanto a su funcionalidad. Hoy día, GNU Gatekeeper o gnugk provee los siguientes servicios:

- ✓ Traducción de direcciones
- ✓ Control de admisión
- ✓ Control y manejo de ancho de banda
- ✓ Control sobre una zona
- ✓ Control de señalización
- ✓ Control y autorización de llamadas

Desde un punto de vista más técnico, gnugk soporta dos backends opcionales para su funcionamiento: MySQL y LDAP. Además en la última versión se optimizó la admisión de miles de usuarios funcionando en modo ruteo.

Instalación

La instalación de gnugk supone que se disponen los siguientes paquetes en el sistema:

- ✓ PWLIB
- ✓ Open H.323
- ✓ MySQL (opcional)
- ✓ LDAP (opcional)

En particular, este software no necesita alguna técnica especial para su compilación. Simplemente es necesario compilar los fuentes, copiar los binarios en algún directorio según conveniencia y generar un archivo de configuración para su funcionamiento.

Además de proveer la funcionalidad de gatekeeper, el software provee una interfaz de administración remota que puede ser accedida a través de telnet y se basa en línea de comandos.

Existen otras dos interfaces de configuración un poco más amigables que la tradicional línea de comandos. Estas interfaces son gráficas y sugieren el reemplazo de la interfaz por línea de comandos provista por defecto.

Configuración

Las capacidades configurables de este software son sumamente flexibles, permitiendo cambiar su modo de funcionamiento. A continuación se describen las más importantes.

- ✓ Modo ruteo: de esta forma todos los mensajes de señalización involucrados en las llamadas pasan a través del gatekeeper. Esto le da mayor control sobre las llamadas. En este modo, todo el tráfico multimedia perteneciente a una comunicación no pasaría por el gatekeeper, sólo la señalización. Los posibles modos de ruteo son los siguientes:
  - ❑ Clase I: el gatekeeper no rutea mensajes. El canal de control H.245 y los canales lógicos se establecen entre los puntos finales.
  - ❑ Clase II: el canal de control H.245 se rutea entre los puntos finales a través del gatekeeper, mientras que los canales lógicos se establecen directamente entre los puntos finales.
  - ❑ Clase III: todo el tráfico pasa por el gatekeeper, incluyendo RTP/RTCP para audio y video, o T.120 para datos. Este sería el caso de un proxy.

- ✓ Autenticación: define una serie de reglas que identifican qué usuarios y máquinas pueden conectarse para la administración del gatekeeper.

- ✓ Numeración E.164: indica los números que deben rutearse a gateways específicos. Además se definen las reglas de reescritura de números E.164.

- ✓ Backend MySQL o LDAP: permite calcular por usuario el tiempo de las llamadas para su facturación. Además provee mecanismos de autenticación H.235 como SHA-1 y MD5.

MCU

Esta entidad es la más codiciada comercialmente, y por esta razón se encontró solo una implementación open source llamada openmcu. Dicha implementación es parte del proyecto Open H.323 y provee las siguientes capacidades:

- ✓ No requiere ningún codec por hardware

- ✓ Soporta los siguientes codecs de audio:
  - ❑ G.711
  - ❑ GSM
  - ❑ MS-GSM de Microsoft
  - ❑ LPC-10

- ✓ Soporta el codec de video H.261

- ✓ Maneja múltiples conferencias al mismo tiempo utilizando el concepto de habitaciones

- ✓ Permite iniciar llamadas desde la MCU

Su funcionamiento básicamente consta en esperar solicitudes de conexión por parte de los usuarios. Cuando una solicitud es recibida, se determina la conferencia requerida por medio de la habitación. La forma de unirse a una conferencia en una habitación es llamando a la siguiente URL: nombre_habitacion@hostname_mcu.

Además de su funcionalidad como MCU, provee una interfaz basada en menúes que permite su administración y visualización de estadísticas.

Instalación

Openmcu se compila por sobre las librerías provistas por Open H.323, imponiendo entonces los siguientes requerimientos:

- ✓ PWLIB

- ✓ Open H.323

Su compilación no requiere consideraciones extra, siendo extremadamente simple, dando como resultado un binario que puede copiarse a cualquier directorio para su posterior uso.

Clientes

Clientes hardware

Producto	Firmware	Descripción
ATA 186	2.15	Es un adaptador que permite conectar dos teléfonos analógicos a la red de datos, de modo de poder integrar dichos teléfonos a una solución de VoIP. Las alternativas son: SIP, H.323 y MGCP

Clientes software

Producto	Plataforma	URL
Cphone	Linux / Windows	http://cphone.sourceforge.net/
Gnomemeeting	Linux	http://www.gnomemeeting.org
Windows Netmeeting 3.01	Windows	http://www.microsoft.com/windows/netmeeting/

Otras aplicaciones

El proyecto Open H.323 provee unas serie de herramientas que extienden la especificación de la ITU, enriqueciendo el conjunto de servicios ofrecidos a los clientes. Entre estos servicios están:

- ✓ Sistema IVR
- ✓ Entidad con soporte T.38, para la recepción de fax

Descripción de las pruebas realizadas sobre una zona SIP

Durante la realización de las pruebas, se determinó que los distintos ejemplos planteados mostraban diferencias en la primer etapa correspondiente a una llamada. Es por esta razón que analizaremos por partes las distintas fases involucradas en una comunicación, a decir:

- Fase A: configuración de llamada
- Fase B: intercambio de capacidades
- Fase C: establecimiento de la comunicación audiovisual
- Fase D: finalización

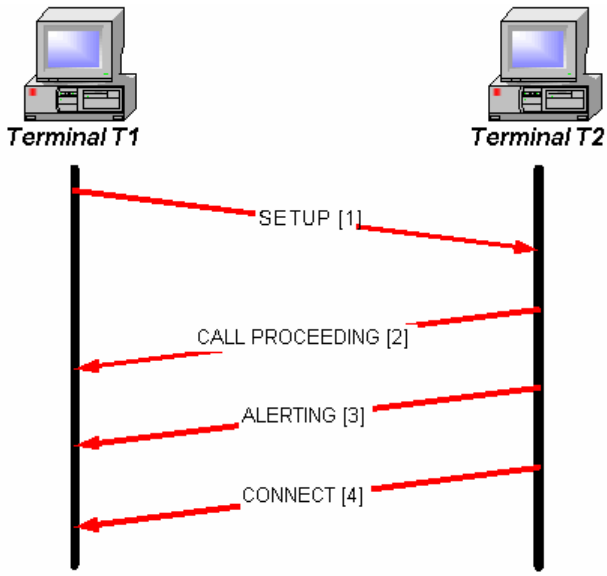
Por lo tanto mostraremos el comportamiento del protocolo H.323 en la fase A según las distintas pruebas que contemplan entidades diferentes en cada caso.

Fase A: configuración de llamada

La configuración de llamada hace uso de los mensajes de control definidos en el estándar H.225.0. Sin embargo, cualquier requerimiento de reserva de ancho de banda debe efectuarse en una etapa previa.

Esta fase, permite una variedad de combinaciones, agregando complejidad a medida que se introducen nuevas entidades H.323.

Dos terminales no registradas



El caso básico corresponde a la conexión entre dos extremos no registrados. En este escenario, ambos extremos se conectan directamente de la siguiente forma:

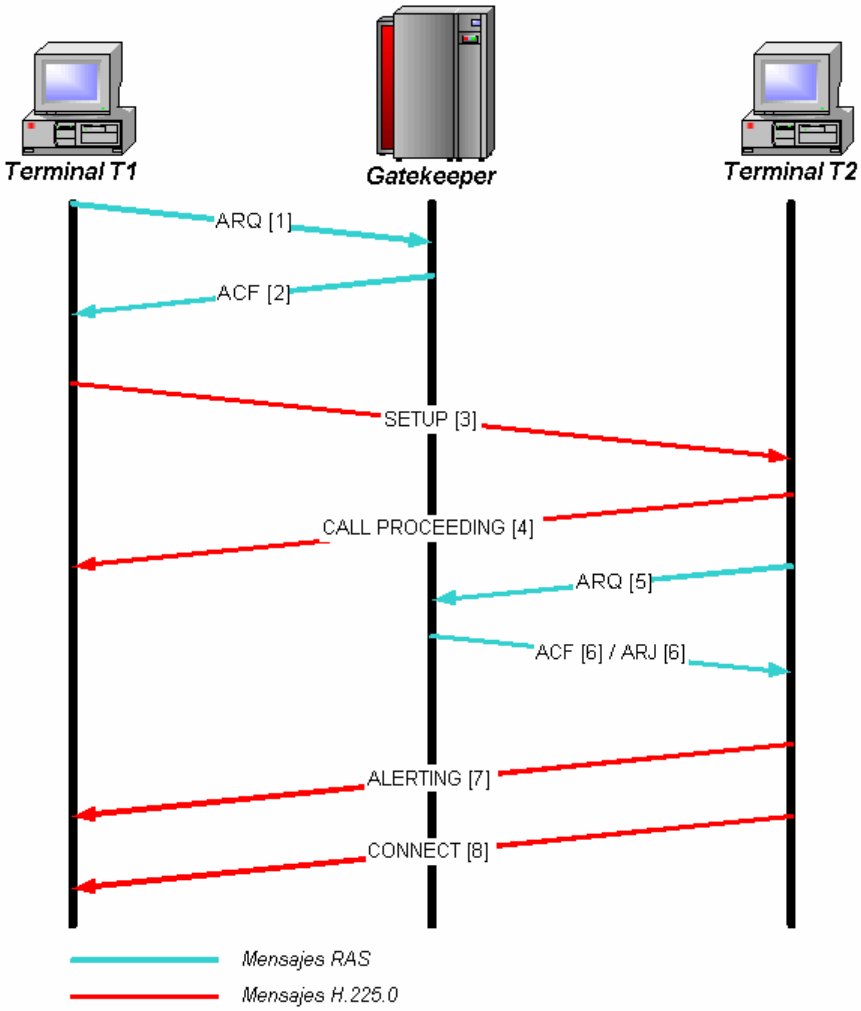
1. La terminal 1 inicia la llamada enviando un mensaje Setup[1] hacia la terminal 2, asumiendo que conoce su nombre de host.

2. Inmediatamente que la terminal 2 recibe el mensaje anterior, envía dos señales de recepción: Call proceeding[2] y alerting[3]

3. Luego la terminal 2 responde con un mensaje Connect[4] incluyendo en él, una dirección H.245 de control del canal de transporte para su posterior utilización en la señalización H.245.

Dos terminales registradas a un gatekeeper. Modo directo

En esta prueba, suponemos dos entidades registradas a un mismo Gatekeeper:



1. La terminal 1 es quien inicia la llamada, comenzando el intercambio de mensajes ARQ[1] / ACF[2] con el gatekeeper, a fines de concretar su admisión.

2. El gatekeeper, que utiliza señalización de llamada directa, retorna a la terminal 1 la dirección correspondiente a la terminal 2, en la confirmación ACF[2] de admisión.

3. Luego, la terminal 1 envía un mensaje Setup[3] directamente a la terminal 2 utilizando la dirección enviada por el gatekeeper en el paso anterior.

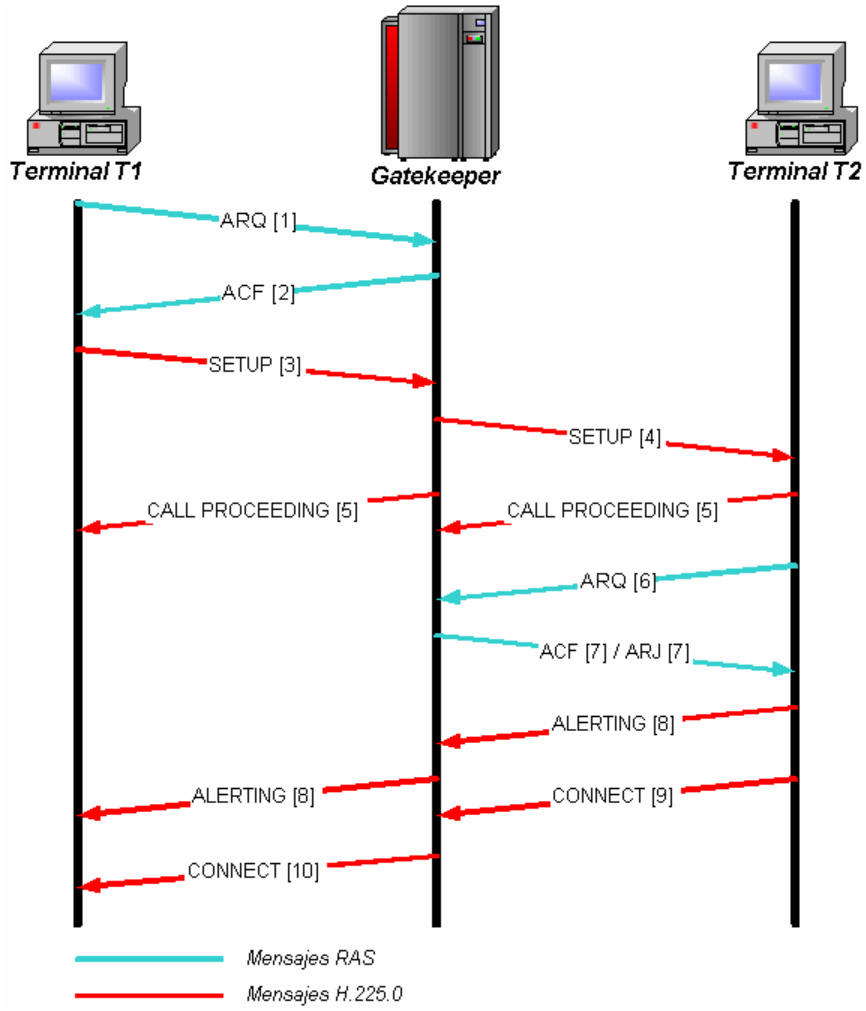
4. Si la terminal 2 desea aceptar la llamada entonces inicia un diálogo de admisión ARQ[5] / ACF[6] con el gatekeeper.

5. Es posible que dicha terminal reciba un ARJ[6], en cuyo caso envía un mensaje Release Complete a la terminal 1.

6. En otro caso, un ACF es recibido por la terminal 2. Entonces responde con un mensaje Connect[8] conteniendo una dirección H.245 de control del canal de transporte para su posterior utilización en la señalización H.245.

Dos terminales registradas a un gatekeeper. Modo ruteo

En esta prueba, suponemos dos entidades registradas a un mismo Gatekeeper:



1. La terminal 1 es quien inicia la llamada, comenzando el intercambio de mensajes ARQ[1] / ACF[2] con el gatekeeper, a fines de concretar su admisión.

2. El gatekeeper, que utiliza señalización de llamada ruteada, retorna a la terminal 1 la dirección correspondiente a sí mismo, en la confirmación ACF[2] de admisión.

3. Luego, la terminal 1 envía un mensaje Setup[3] al gatekeeper utilizando la dirección recibida en el paso anterior.

4. Al recibir el mensaje Setup, el gatekeeper luego envía otro mensaje Setup[4] a la terminal 2.

5. Si la terminal 2 desea aceptar la llamada entonces inicia un diálogo de admisión ARQ[6] / ACF[7] con el gatekeeper.

6. Es posible que dicha terminal reciba un ARJ[7], en cuyo caso envía un mensaje Release Complete al gatekeeper.

7. En otro caso, un ACF es recibido por la terminal 2. Entonces responde al gatekeeper un mensaje Connect[9] conteniendo una dirección H.245 de control del canal de transporte para su posterior utilización en la señalización H.245.

8. El gatekeeper reenvía este mensaje a la terminal 1, Connect[10], modificando o no la dirección H.245 indicada por la terminal 2, dependiendo de si se utiliza ruteo del canal de control H.245 o no.

Conferencias con MCU

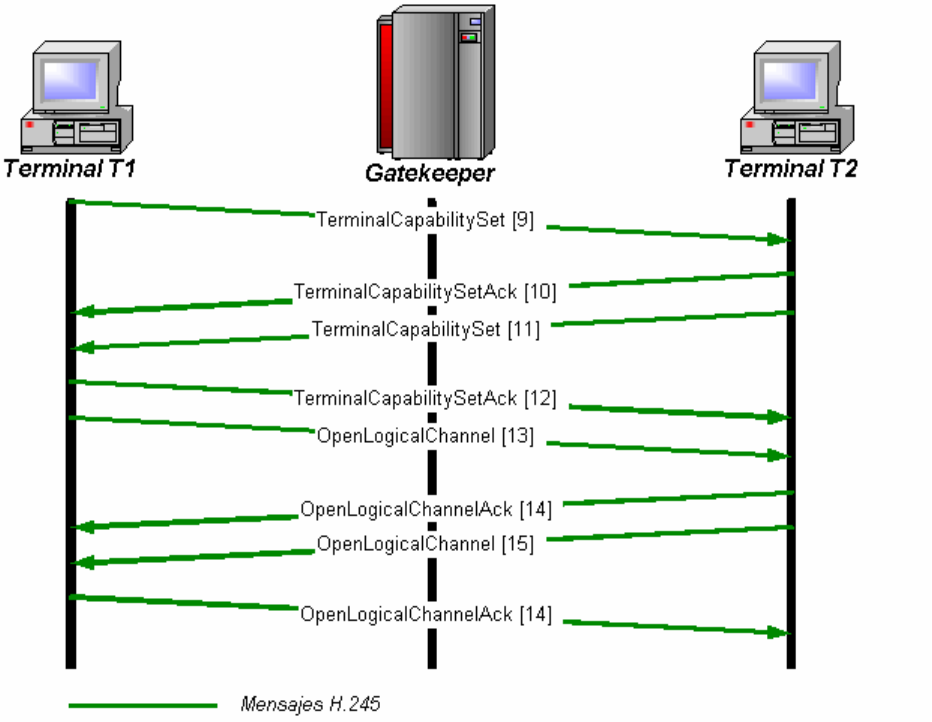
Para el caso de conferencias utilizando una MCU, todas las terminales intercambian mensajes de señalización con la MCU. La fase A entre una terminal y una MCU procede de igual forma que sucede entre dos terminales, tal cual se describió en los casos anteriores, donde el rol de la MCU es el de una terminal que recibe llamadas o las inicia.

Cuando se efectúan conferencias utilizando MCUs, los canales de control H.245 se abren entre las terminales y el MC dentro de la MCU, mientras que los canales de audio, video y datos son abiertos entre las terminales y el MP dentro de la MCU. En contraste con las conferencias descentralizadas, donde se forma una malla de canales de comunicación, notamos que la MCU forma un punto central donde todos los canales en cuestión forman una estrella.

Fase B: intercambio de capacidades

Una vez que ambos extremos han intercambiado mensajes de configuración de llamada propios de la fase A, las terminales deben establecer un canal de control H.245. El propósito de este canal de control es el intercambio de capacidades y la apertura de los canales de datos multimedia.

A continuación se muestra el flujo de mensajes de control en la señalización H.323.



1. T1, quien inició la llamada, envía un mensaje TerminalCapabilitySet [9] a T2 para intercambiar capacidades.

2. Si T2 está de acuerdo con las capacidades especificadas por T1, envía un mensaje TerminalCapabilitySetAck [10] a T1, reconociendo sus las capacidades.

3. Luego T2 intercambia sus capacidades con T1 enviándole un mensaje TerminalCapabilitySet [11].

4. T1 le envía un mensaje TerminalCapabilitySetAck [12] a T2, reconociendo sus las capacidades.

5. T1 envía un mensaje OpenLogicalChannel [13] a T2, solicitando la apertura de un canal lógico. En este mensaje se incluye el número de puerto para el canal RTCP.

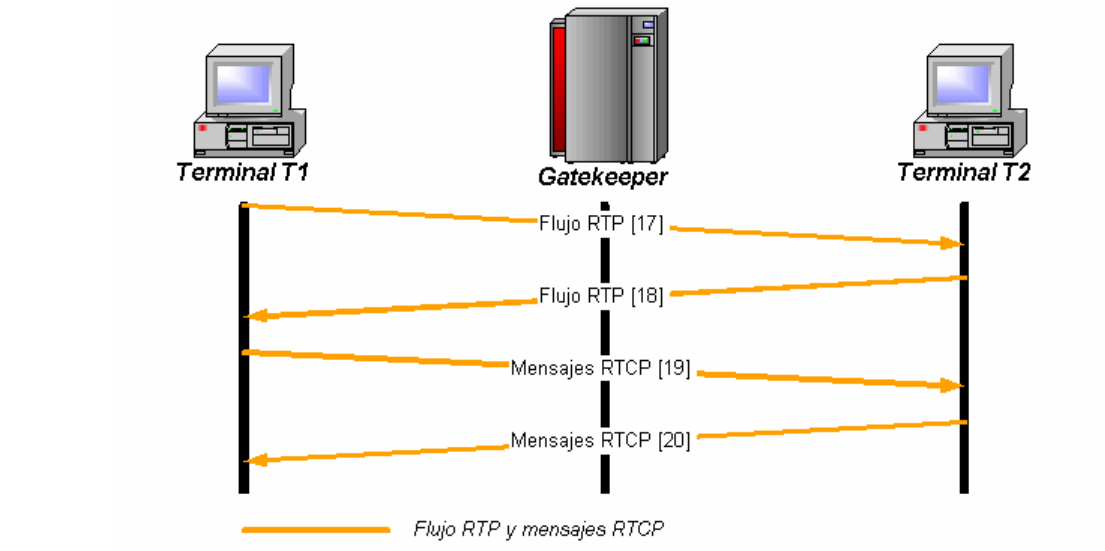
6. T2 asiente el establecimiento del canal lógico unidireccional de T1 a T2 por medio de un mensaje OpenLogicalChannelAck [14]. Incluido en el mensaje están el puerto que utilizará RTP definido por T2 para ser usado por T1 para el envío de paquetes RTP y el puerto RTCP recibido de T1 previamente.

7. Luego T2 abre un canal lógico con T1 enviándole un mensaje OpenLogicalChannel [15].

8. T1 asiente el establecimiento del otro canal lógico unidireccional de T2 a T1 por medio de un mensaje OpenLogicalChannelAck [16]. Ahora, la comunicación entre T1 y T2 es bidireccional para la transmisión de paquetes RTP.

Fase C: establecimiento de la comunicación audiovisual

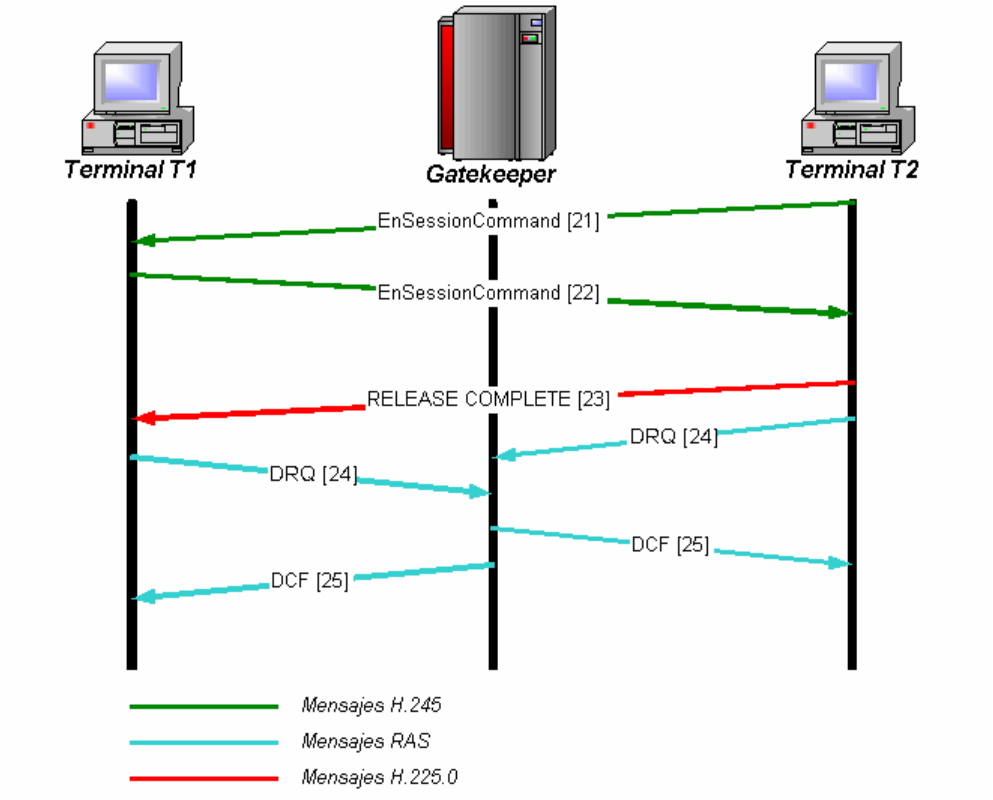
En este punto ya está establecido un canal RTP/RTCP bidireccional por el cual T1 y T2 intercambian datos multimedia. Para el caso de datos de audio o video se utiliza un canal no confiable como es el caso de UDP, mientras que para intercambio de datos se utiliza un canal confiable como TCP. A continuación se muestran los flujos presentes entre dos terminales que por ejemplo mantienen una conversación:



1. T1 envía el stream de datos encapsulado en RTP a T2.
2. T2 envía el stream de datos encapsulado en RTP a T1.
3. T1 envía mensajes RTCP a T2.
4. T2 envía mensajes RTCP a T1.

Fase D: finalización

Cuando una de las partes desea terminar la comunicación, se produce la siguiente secuencia de mensajes:



1. T2 inicia la liberación de la llamada. Para ello envía un mensaje H.245 EndSessionCommand [21] a T1.

2. T1 confirma la liberación enviando un mensaje H.245 EndSessionCommand [22] a T2.

3. T2 completa la liberación enviando un mensaje H.225 RELEASE COMPLETE [23] a T1.

4. T1 y T2 se desconectan del gatekeeper enviándole un mensaje RAS DRQ [24].

5. El gatekeeper se desconecta de T1 y T2 por medio del envío de un mensaje RAS DCF [25].

Observaciones

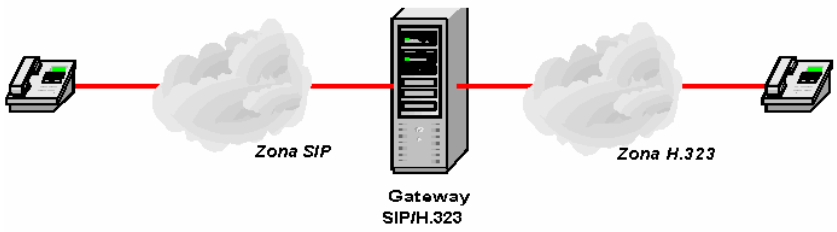
A diferencia de SIP, con H.323 no se presentaron dificultades. Atribuimos este hecho a que H.323 se muestra más íntegro, tal vez más refinado que SIP. Otro factor en favor de H.323 es que las herramientas disponibles para las pruebas, salvo Netmeeting de Microsoft y el dispositivo ATA 186 de Cisco, utilizaban la misma librería minimizando las incompatibilidades.

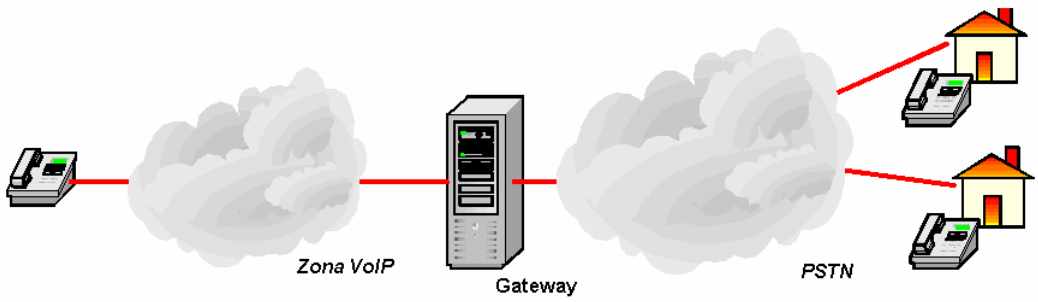
Por otra parte la instalación y configuración de los productos no introdujo complicaciones.

Capítulo 11: Pruebas de integración

Simplificamos las pruebas de integración en dos puntos básicos:

- Establecer comunicaciones entre Microsoft Netmeeting y Windows Messenger
- Establecer comunicaciones entre cualquier cliente VoIP y un teléfono de la PSTN o celular.





La explicación a esta simplificación, es la integración detrás de cada caso. El primer caso muestra cómo integrar los dos productos VoIP pivote de Microsoft, de uso masivo e incompatibles entre sí, a través de la interacción de los protocolos H.323 y SIP. El segundo caso plasma el propósito de este trabajo en una prueba donde un usuario de la PSTN llama a un cliente VoIP y viceversa.

Para llevar a cabo estas pruebas se aprovechó la infraestructura utilizada en los casos anteriores. Luego fue necesario instalar un gateway, provisto por Vocal, que permitió la interacción entre ambos protocolos. Además se necesitó de un dispositivo de hardware, la placa Internet Linejack, para permitir las llamadas entre VoIP y teléfonos de la PSTN.

Conclusiones

La aparición de redes de datos tuvo como principal causa la necesidad de intercambiar datos entre dispositivos digitales, facilitando las formas de compartir y optimizar recursos. Es de esta forma como diferentes redes de datos fueron surgiendo pero sólo TCP/IP se impuso sobre el resto constituyendo el núcleo de la hoy conocida Internet.

Internet se ha convertido en una red de acceso masivo, de bajo costo y que además provee a sus usuarios de una variada cantidad de servicios de comunicación. Entre estos servicios vale destacar la importancia del correo electrónico y la mensajería instantánea, ya que han sido adoptados como formas cotidianas de comunicación.

Paralelamente a las tecnologías basadas en Internet, el servicio telefónico siguió siendo la alternativa de comunicación más frecuente. Este servicio se apoya sobre una red independiente de Internet, cuya filosofía e infraestructura no coinciden. Incluso la telefonía mostró un gran crecimiento hacia la actualidad por el auge de la telefonía celular, que provee movilidad a sus usuarios.

Estas dos redes diferentes, que han evolucionado paralelamente, han tenido su encuentro en la última década. A continuación se nombran las principales razones por las cuales, inicialmente las comunicaciones telefónicas no se integraron con una red de datos como es Internet:

- El servicio telefónico brinda al usuario la seguridad de disponer del mismo cuando lo necesite, mientras que en Internet, la disponibilidad de sus servicios no está asegurada por ser ésta una red de mejor esfuerzo.
- El tipo de red subyacente en la PSTN asegura una buena calidad de servicio para el transporte de voz, ya que es garantizada desde el momento en que la comunicación fue establecida. Para el caso de comunicaciones de voz en Internet, brindar la misma calidad que la PSTN no es trivial, y muchas veces se vuelve imposible por diversos factores relacionados con la performance de la red.
- El teléfono es un dispositivo de fácil utilización incluso para el menos idóneo. Sin embargo, en VoIP existen numerosos dispositivos y aplicaciones de comunicación que son de un nivel de complejidad mucho mayor al de un teléfono tradicional.
- Los servicios propios de Internet generalmente transportan datos estáticos como ser texto e imágenes, por lo que son sensibles a la perdida de paquetes, es decir no importa la demora sino la integridad de los datos. Si bien es deseable que los datos arriben al destino lo antes posible, pueden tolerarse perdida de paquetes y grandes retardos.
- Totalmente en contraposición al caso anterior, transmitir audio y video en Internet es un tanto más complejo puesto que los servicios que transportan este tipo de datos son altamente sensibles a los retardos y son pocos tolerantes a la perdida de paquetes. Esto se debe a que datos que alcanzan el destino con un gran retardo resultan obsoletos si el momento de reproducción de los mismos ya ocurrió.

Compatibilizar la red telefónica convencional con la red de datos IP, impuso desafíos complejos, donde los protocolos de señalización y el buen desempeño de la red han demostrado ser puntos claves para el éxito.

Los protocolos de VoIP son varios, siendo los de señalización los más importantes. SIP y H.323 son los estándares que hoy día se disputan el mercado de esta tecnología. Mas allá de cuál sea la alternativa escogida, la integración de éstos protocolos con la PSTN está resuelta.

Por otro lado, el desempeño de la red es vital para que las comunicaciones VoIP sean satisfactorias. Las alternativas para atacar este punto son varias, si embargo la clave es mantener bajos los valores de jitter y perdida de paquetes.

La relación entre la calidad de las comunicaciones y los protocolos de señalización VoIP, está dada por la negociación de los codecs a utilizar. Este proceso, llevado a acabo por SDP en el caso de SIP y por H.245 para H.323, permite negociar, entre otras cosas, el codec a utilizar. Este punto es de vital importancia ya que un excelente codec será aquel que demuestre una buena calidad de los datos multimedia, a la vez de lograr una alta compresión en el menor tiempo posible.

A pesar de estas restricciones para que VoIP tenga una calidad aceptable, es necesario que sea aceptada por la gente. No sirve de mucho brindar una solución de VoIP utilizando PCs como dispositivos de comunicación si los que van a comunicarse no tienen un manejo mínimo de las posibilidades básicas ofrecidas por dichos dispositivos.

El siguiente análisis refleja en distintos escenarios, los factores a tener en cuenta en la implantación de una solución de VoIP.

Caso I: integración de una central telefónica convencional con VoIP.

Este caso contempla aquellas empresas que mantienen una central telefónica sin capacidades VoIP y desean adoptar esta nueva tecnología, para por ejemplo minimizar costos de comunicación entre sucursales distantes o incluso agregar internos donde no hay cobertura del tendido telefónico.

En este escenario, ya poseemos una infraestructura de comunicación donde cada sucursal posee su propia central de telefonía. La principal ventaja de esto es que disponemos del telefono convencional como elemento de comunicación para los usuarios.

Para el caso de centrales telefónicas tradicionales, VoIP es posible asignándole un interno de la misma a un gateway VoIP, el cual permita enrutar las llamadas entre sucursales a través de la red de datos. Cuando queramos realizar llamadas VoIP tendríamos que marcar el interno correspondiente a dicho gateway y este nos proveerá de una nueva línea de marcado. Los destinos llamados a través de este gateway se limitan a direcciones numéricas las cuales se encargara de resolver un gatekeeper o un location server, dependiendo si la tecnología de VoIP utilizada es H.323 o SIP respectivamente.

Para el caso de centrales telefónicas de ultima generacion, es probable que ya vengan capacitadas con un puerto de conexión LAN, la cual nos permitirá conectar nuestra red de internos con la red de paquetes. En este caso, la misma central telefónica podria oficiar de gateway, para lo cual tendría que utilizar, al igual que el caso anterior, los servicios provistos por un gatekeeper o un location server.

En este escenario, la calidad de las llamadas de voz para el caso de las llamadas internas es la misma que la de la PSTN, puesto que la central telefónica se encarga de ellas. Por otro lado, para el caso de las llamadas VoIP, deberían analizarse los parámetros de jitter y perdida de paquetes.

Para el caso de sucursales remotamente conectadas via Internet, si los parametros de delay y jitter no son buenos, debería analizarse la posibilidad de tener un enlace dedicado entre las mismas para el transporte de voz y datos, estableciendo las reservas necesarias para los canales de voz necesarios.

Caso II: solución íntegramente VoIP

En este escenario, los dispositivos de comunicación que podemos utilizar son varios. A continuación se listan las posibilidades de las que disponemos:

- Teléfono analógico: es una excelente solución para usuarios que no están familiarizados con dispositivos más sofisticados. Su costo es bajo, pero se necesita de alguna especie de gateway que los adapte a una red de datos.
- Teléfono digital: este dispositivo es en realidad una computadora de propósito especifico. Su costo es alto y su manipulación es relativamente difícil, ya que maneja todas las capacidades provistas por VoIP.
- Computadora de propósito general: su costo relativamente es bajo teniendo en cuenta que sus funciones son varias y que ya disponemos de las mismas. Su utilización como medio de comunicación es poco natural. Sin embargo, su manipulación provee mayor flexibilidad ya que integra un interno telefónico, correo electrónico y mensajería instantánea en un mismo dispositivo.

La elección del dispositivo a utilizar no está restringida a sólo uno de los anteriores, sino que pueden convivir entre las diferentes alternativas.

Respecto a la calidad del servicio necesaria para VoIP, hay que tener en cuenta la performance de la red ya que las comunicaciones pueden verse degradadas ante la aparición de factores externos como podría llegar a ser una descarga masiva producida por herramientas peer-to-peer. Sin embargo, las redes de área local actuales alcanzan grandes velocidades por lo que a este nivel no sería notoria una gran degradación.

En este caso, queda considerar llamadas a teléfonos de la PSTN, por lo que sería necesario instalar un gateway.

Finalmente, al igual que en el caso anterior, si existiera alguna sucursal remota conectada a través de la red de datos, debería analizarse la posibilidad de establecer las reservas necesarias sobre el canal para los canales de voz que se quieran.

En definitiva, una buena implementación VoIP será aquella en la cuál se considere:

- Los clientes deben utilizar codecs adecuados a la red subyacente
- La implementación de QoS es de gran ayuda, evitando efectos indeseables en las comunicaciones

- La selección del protocolo de señalización teniendo en cuenta qué servicios de comunicación se desean ofrecer
    - Conferencias
    - Audio
    - Video
    - Fax
    - Imágenes
    - Aplicaciones colaborativas
- Cálculo de la máxima cantidad de usuarios que pueden compartir el enlace simultáneamente en comunicaciones VoIP

A lo largo del trabajo, se fueron obteniendo diversas conclusiones acerca de las distintas componentes del universo VoIP. Sin embargo, las conclusiones más impactantes fueron obtenidas cuando se pisó el plano práctico. Ya se han expuesto los casos de prueba realizados y plantearon los inconvenientes que se presentaron a la hora de testear cada protocolo de señalización. Fue precisamente después de este momento donde adquirimos una visión más personal de cada alternativa VoIP.

Muchos de los detalles propios de SIP y H.323 fueron observados en los correspondientes capítulos de pruebas, permitiéndonos en ese punto hallar conclusiones propias. Las mismas se encuentran al finalizar cada capítulo.

Una vez adquiridos los conocimientos teóricos de cada alternativa y evaluado productos VoIP actuales, podemos asegurar que actualmente H.323 se muestra más robusto que SIP, siempre que su uso considere conferencias, comunicaciones de audio, video, etc. Esto no quiere decir, que descartemos SIP, ya que sería una buena opción para el reemplazo de una central telefónica convencional, pero no sería adecuado su uso en conferencias, videoconferencias o comunicaciones que consideren datos diferentes de audio.

En cuanto a los dispositivos e implementaciones de VoIP, podemos decir que muchas compañías agregan capacidades y se desprenden de los estándares. En algunos casos prometen consumos de ancho de banda excesivamente bajos, pero sólo entre equipos del mismo fabricante, logrando compresiones a través de mecanismos físicos, en vez del uso de codecs. Estas observaciones son las que, a nuestro parecer, a largo plazo tal vez terminen siendo parte de un estándar, pero durante el lapso que es una capacidad extra de un fabricante, no es más que un impedimento para la interoperabilidad entre diversos productos.

Glosario

Accounting: mide el consumo de recursos utilizados por un usuario durante su acceso. Por ejemplo la cantidad de tiempo de sistema utilizado o la cantidad de datos enviados y/o recibidos durante una sesión. El accounting se obtiene a partir del logueo de información estadística de la sesión, y generalmente se utiliza para el control de autorización, facturación, utilización de recursos, y actividades de planeamiento.

ACF: Admision ConFirm, mensaje H.323.

ALERTING: mensaje H.323.

Alias: nombre alternativo para algo o alguien.

Ancho de banda: cuánta información puede ser transportada en un período de tiempo dado, generalmente segundos, sobre un enlace de cable o inalámbrico.

Aplicaciones multimedia: aplicaciones que utilizan recursos de audio, video o gráficos.

ARJ: Admision ReJect, mensaje H.323

Arpanet: fue la red pilote de Internet. Fue creada por fuentes militares de USA y consistía de un número de computadoras individuales conectadas a través de líneas arrendadas. Arpanet, fue reemplazada en 1980 por una nueva red militar, científica y universitaria creada por la National Science Funation.

ARQ: Admision ReQuest, mensaje H.323.

ASN.1: Abstract Syntax Notation One. Lenguaje formal para describir en forma abstracta mensajes a ser intercambiados entre una gran cantidad de aplicaciones. La especificación de su notación básica está descripta en ITU-T Rec. X.680 (2002).

ATA-186: equipo de Cisco Systems, que funciona como un adaptador de teléfonos analógicos en teléfonos VoIP.

Autenticación: es el proceso de determinar si alguien o algo es quien o lo que dice ser.

Back-office: área o departamento de una empresa encargada del procesamiento de los aspectos operativos vinculados a una operación ya sea desde el registro del cliente hasta la liquidación de las transacciones.

BCF: Bandwidth ConFirm, mensaje H.323.

BRJ: Bandwidth ReJect, mensaje H.323.

BRQ: Bandwidth ReQuest, mensaje H.323.

Bucle local: conexión cableada desde la oficina central de telefonía hasta el aparato del abonado.

Bug: error de codificación en un programa de computación.

C: lenguaje de programación.

C++: lenguaje de programación.

Calidad de servicio, QoS: permite calcular tiempos de transmisión, frecuencia de errores y otras características y, en cierta medida, garantizarlos. Existen dos alternativas para lograr QoS: RSVP y COPS.

CALL PROCEEDING: mensaje H.323.

Cisco Systems: empresa dedicada al networking. http://www.cisco.com

Class: Custom Local Area Signaling Services. Tipo de servicios ofrecidos por un oficina de telefonía.

CO: Central Office. Oficina central de telefonía a la cuál se conectan los abonados.

Codec: Acrónimo que se refiere a un algoritmo que realiza compresión y descompresión de los datos.

Codificación: proceso por el cual se toma una señal analógica y se la convierte en digital.

CONNECT: mensaje H.323.

Correo electrónico: el correo electrónico, también llamado E-MAIL (Electronic Mail), es una forma de enviar correo, mensajes o cartas electrónicas de una computadora a otra. Tanto la persona que envía el correo electrónico, como la persona que lo recibe, debe tener una cuenta en una maquina de Internet.

DCF: Disangage ConFirm, mensaje H.323

Decodificacion: proceso inverso a la codificación, por medio del cual una señal digital se convierte en una señal analógica.

Delay: retardo que sufren los paquetes al ser transportados por la red

DHCP: Dynamic Host Configuration Protocol, es un protocolo que permite la configuración automática de los host que utilizan TCP/IP. Definido en la RFC 2131 http://www.ietf.org/rfcs/rfc2131.html

MD5: descripto en la RFC 1321, http:// www.ietf.org/rfcs/rfc1321.html, es un algoritmo que toma un mensaje de entrada de cualquier longitud y produce una salida fija de 128 bit llamada digest. Dado que es casi imposible encontrar dos mensajes que produzcan el mismo digest, este algoritmo es utilizado para firma digital y autenticación.

Digitalización: proceso por el cual se toma una señal analógica y se la convierte en digital.

Digest MD5: ver MD5.

DNS: el Domain Name System es un sistema por medio del cual son ubicados los nombres de dominio propios de Internet y traducidos en direcciones IP. Un nombre de dominio es una forma amigable de referenciar direcciones de Internet. Dado a que el mantenimiento centralizado de la lista de nombres de dominio sería poco práctica, esta lista es administrada en forma jerárquica por medio de bases de datos distribuidas.

Draft: es un tipo de reporte técnico de un trabajo en progreso, una forma preliminar de un documento futuro. Luego de una serie de revisiones, este puede pasar a ser una RFC o quedar obsoleto.

DRQ: Disengage ReQuest, mensaje H.323

DS: Digital Signal X es el término utilizado para las transmisiones digitales basadas en DS0. Tanto el sistema de cableado norteamericano como el europeo, sistemas T y E respectivamente, operan utilizando series DS. DS0 es la base de las digital signal X.

DS0: Disgital Signal 0 es un enlace con una velocidad de transmisión de 64 kbps, el cual es generalmente utilizado para un canal telefónico de voz.

DS1: Digital Signal 1 es un enlace compuesto de 24 DS0, transmitidas utilizando PCM y TDM. Es la señal utilizada en T1.

E1: Conexión por medio de la línea telefónica que puede transportar datos con una velocidad de hasta 1,920 Mbps. Según el estándar europeo (ITU), un E1 está formado por 30 canales de datos de 64 kbps más 2 canales de señalización. E1 es la versión europea de T1 (DS-1). Velocidades disponibles:

E1: 30 canales, 2,048 Mbps
E2: 120 canales, 8,448 Mbps
E3: 480 canales, 34,368 Mbps
E4: 1920 canales, 139,264 Mbps
E5: 7680 canales, 565,148 Mbps

EIA: Electronic Industries Association abarca organizaciones individuales que en conjunto han acordado sobre ciertos estándares de transmisión.

EndSessionCommand: mensaje H.323

Equivalence Pty. Ltd: comenzaron el proyecto Open H.323, y son los principales mantenedores del código. http://www.equival.com.au

ESS: Electronic Switching System

Ethernet: tecnología LAN especificada en el estándar IEEE 802.3. Esta tecnología fue desarrollada por Xerox. El sistema Ethernet más utilizado es el 10Base-T, alcanzando velocidades de hasta 10 Mbps. En estas redes, los dispositivos se conectan al cable y compiten por el acceso utilizando el protocolo CSMA/CD.

FastEthernet: o 100Base-T, es una extensión de Ethernet alcanzando velocidades de hasta 100 Mbps.

FTP: File Transfer Protocol es un estándar que permite el intercambio de archivos a través de Internet. Este protocolo se encuentra definido en la RFC959. http://www.ietf.org/rfc/rfc0959.txt

G.711: es un estándar internacional definido por la ITU-T para la codificación de audio sobre canales telefónicos de 64 kbps. Es una codificación PCM que opera a 8 kHz, utilizando 8 bits por muestra. G.711 puede codificar frecuencias entre 0 y 4 kHz. Las compañías telefónicas pueden seleccionar dos variantes de G.711: A-law y mu-law. A-law es el estándar utilizado en circuitos internacionales.

G.722: es un estándar definido por la ITU-T para la codificación de señales de audio muestreadas a 16000 muestras por segundo. Este codec se basa en SB-ADPCM, donde la señal se divide en dos sub bandas y las muestras de ambas bandas se codifican utilizando ADPCM.

G.723.1: es un estándar definido por la ITU-T para la compresión de audio. Su aplicación típica es en telefonía sobre redes de paquetes como VoIP.

G.728: es un estándar definido por la ITU-T para la compresión de audio. Su aplicación típica es en telefonía sobre redes de paquetes, especialmente IP. Es un codec robusto y de muy buena calidad, comparable con ADPCM de 32 Kbps.

G.729: sofisticado estándar definido por la ITU-T para la compresión de audio. Codifica señales a 8 Kbps. Este codec incluye la detección de actividad de voz y reducción de ruido.

Gatekeeper: punto central para todas las llamadas dentro de una red H.323. Entre sus servicios se pueden mencionar: direccionamiento, autorización y autenticación, administración de ancho de banda, facturación.

Gateway: En general se trata de una pasarela entre dos redes. Técnicamente se trata de un dispositivo repetidor electrónico que intercepta y adecua señales eléctricas de una red a otra. En Telefonía IP se entiende que estamos hablando de un dispositivo que actúa de pasarela entre la red telefónica y una red IP. Es capaz de convertir las llamadas de voz y fax, en tiempo real, en paquetes IP con destino a una red IP, por ejemplo Internet. Originalmente sólo trataban llamadas de voz, realizando la compresión/descompresión, paquetización, enrutado de la llamada y el control de la señalización. Hoy en día muchos son capaces de manejar fax e incluir interfaces con controladores externos, como gatekeepers, soft-switches o sistemas de facturación.

H.225.0: protocolo de señalización de llamadas y RAS utilizado en H.323.

H.245: protocolo de señalización de control utilizado para el intercambio de capacidades entre terminales H.323.

H.261: estándar para la codificación de video publicado por la ITU en 1990. Fue diseñado para líneas ISDN.

H.263: estándar para la codificación de video que extiende y mejora a H.261. Inicialmente fue diseñado para enlaces de baja calidad.

H.320: estándar internacional para las comunicaciones multimedia sobre redes ISDN, incluyendo audio, video y conferencia de datos definido por la ITU-T.

H.323: estándar internacional para las comunicaciones multimedia sobre redes de paquetes, incluyendo LAN, WAN e Internet, definido por la ITU-T.

HTTP: Hypertext Transfer Protocol es un protocolo de capa de aplicación que puede utilizarse para muchas tareas además de su uso para hipertexto. El protocolo se encuentra definido en la RFC 2616. http://www.ietf.org/rfc/rfc2616.txt

IEEE: Institute of Electrical and Electronics Engineers, Inc. Es una asociación de profesionales técnicos sin fines de lucro que cuenta con más de 360000 miembros en aproximadamente 175 piases. http://www.ieee.org

IETF: Internet Engineering Task Force. Es una comunidad abierta e internacional de diseñadores, operadores, vendedores e investigadores con un mismo interés, la evolución de la arquitectura de Internet. http://www.ietf.org

Internet: red de redes en donde los usuarios de cualquier máquina pueden, mientras tengan permiso, obtener información de otras máquinas. Fue concebida por el gobierno de los Estados Unidos con el fin de interconectar Universidades. Hoy día Internet es pública y accesible por cientos de millones de personas en el mundo.

Internet Linejack: placa ISA cuya función es la de gateway VoIP de bajo costo. Su función está limitada a una única línea. http://www.quicknet.net/products/ilj.htm

IP: Internet Protocol es un protocolo de capa de red que define el direccionamiento y la entrega de paquetes en Internet. Este protocolo se encuentra definido en la RFC 791. http://www.ietf.org/rfc/rfc0791.txt

Iptel: grupo de investigadores dedicados a soluciones SIP desde 1995. http://www.iptel.org

IPX: Internetwork Packet eXchange es un protocolo de red utilizado por el sistema operativo Novell NetWare.

IRC: Internet Relay Chat protocol es un protocolo que permite a personas conectadas a Internet unirse a discusiones en tiempo real. A diferencia de los sistemas de chat antiguos, IRC no está limitado a dos personas. Este protocolo esta definido en la RFC 1459.

ISDN: Integrated Services Digital Network es un estándar de comunicaciones internacional para el envío de voz, video y datos sobre líneas telefónicas digitales.

ISO: International Organization of Standarization es una organización internacional fundada en 1946 que ha definido una gran cantidad de estándares en el universo de la Informática.

ISUP: ISDN User Part define el protocolo y procedimientos utilizados para iniciar, mantener y finalizar trunks que transportan voz y datos de llamadas sobre la PSTN. Las llamadas que se originan y terminan en el mismo conmutador no utilizan señalización ISUP.

ITU: International Telecommunication Union es una organización responsable de aceptar tratados internacionales, regulaciones y estándares que abarquen telecomunicaciones. En un principio las funciones de estandarización eran tarea de un grupo dentro de la ITU llamado CCITT, pero después de una reorganización en 1992, la CCITT dejó de existir como entidad separada.

ITU-T: ITU Telecommunication Standarization Sector es uno de los tres sectores de la ITU. La misión de la ITU-T es asegurar la producción de estándares de alta calidad en forma eficiente.

IVR: Interactive Voice Response permite utilizar los tonos telefónicos para interactuar con una base de datos y adquirir información o ingresar datos en la misma.

J1: La versión japonesa del sistema E en Europa o T en Norteamérica.

J1: 24 canales, 1.544 Mbps
J2: 96 canales, 6.312 Mbps
J3: 480 canales, 32.064 Mbps
J4: 1440 canales, 97.728 Mbps
J5: 5760 canales, 400.352 Mbps

Jabber: protocolo XML para el intercambio de mensajes y presencia en tiempo real entre dos puntos de Internet. Jabber mantiene una plataforma similar a los sistemas como AIM, ICQ y MSN pero es open source, extensible a través de XML, descentralizado, y cualquier servidor Jabber puede instalarse aislado de forma tal de poder utilizarse en Intranets aumentando los niveles de seguridad.

Java: lenguaje de programación.

Jitter: variación del delay en comunicaciones multimedia causando la degradación de la señal transmitida

LAN: Local Area Network es una red que abarca un área pequeña. Existen diferentes tipos de LAN siendo Ethernet la más utilizada.

Latencia: Retardo que sufren los paquetes al ser transportados por la red, es decir, el tiempo que toma a un paquete llegar desde el origen al destino.

Lip synchronization: Es la sincronización de audio y la correspondiente señal de video de manera que no se note la falta de simultaneidad entre ellos.

Matriz: tabla de valores bidimensional

MCU: Multipoint Control Unit. En un elemento de una red H.323. Esta provee soporte para conferencias multipunto entre tres o más terminales. Esta también provee administración de recursos y negociación de codecs entre los participantes.

Mensajeria instantánea: también llamado IM o IMing, es un servicio que permite ver si un amigo, compañero o contacto está conectado a Internet y, si está conectado, intercambiar mensajes instantáneos con él.

Mensajeria unificada: Servicio que unifica todas las formas de comunicación.

Mixer: es un dispositivo que recolecta múltiples paquetes multimedia provenientes de múltiples orígenes, los combina en un solo paquete, y lo reenvía a su destino.

MMUSIC: Multiparty MUltimedia SessIon Control; grupo de trabajo de la IETF, creado para desarrollar protocolos de tele conferencia en Internet y comunicaciones multimedia.

MTP: Message Transfer Part. Mensaje SS7.

Muestreo: Proceso en el cual se realizan reiteradas observaciones a una frecuencia determinada, sobre una variable.

NAT: Network Address Traslation, es una función propia de un router, la cual consiste en la tralación de una dirección IP, usada en una red, a una dirección IP diferente perteneciente a otra red. NAT es descrito en la RFC 1631: http://www.ietf.org/rfc/rfc1631.txt

Oficina central: también conocida como CO, es una oficina de telecomunicaciones centralizada en una localidad específica que maneja el servicio telefónico en esa localidad. La CO maneja switchea llamadas locales y de larga distancia.

OpenH323: es un proyecto cuya meta es crear una implementación open source, interoperable y completamente funcional del protocolo de tele conferencia H.323 de la ITU-T.

OpenLogicalChannel: Mensaje H.323

OpenLogicalChannelAck: Mensaje H.323

OSI: Open System Interconnection, es un modelo de referencia sobre como los mensajes deben ser transmitidos entre los distintos integrantes de una red de telecomunicaciones.

Patch: Un arreglo temporal para un bug de un programa.

PBX: Private Branch eXchange, es una red telefónica privada usada dentro de una empresa. Los usuarios de la PBX comparten un cierto numero de líneas directas para realizar llamadas externas.

PCM: Pulse Code Modulation, es una técnica de muestreo para digitalizar señales analógicas, especialmente señales de audio. PCM hace el muestreo de la señal 8000 veces por segundo; Cada muestra es representada con 8 bits por lo que consume un total de 64 Kbps.

Pixel: Picture element. Es un punto en una imagen. Los monitores gráficos muestran las imágenes dividiendo la pantalla en millones de pixels, organizados en filas y columnas. El numero de bits usados para representar cada píxel determina la cantidad de colores o niveles de gris que se pueden mostrar.

Port: En redes TCP/IP y UDP, un port es un lugar de conexión lógica. Los números de puertos van del 0 al 65536. Los puertos del 0 al 1024 son utilizados por ciertos servidores mientras que el resto son utilizados en forma dinámica por las aplicaciones con necesidades de comunicación.

PSTN: Public Switched Telephone Network. Este término se refiere al sistema de teléfonos internacional que basa en cables de cobre el transporte de datos de voz analógicos.

Puerto: ver port.

Punto final: En H.323, se refiere de esta manera a cualquiera de los extremos de una comunicación.

PWLIB: Es una librería de clases diseñada para dar soporte al proyecto OpenH323.

Q.931: es un protocolo de control de conexiones propio de ISDN definido por la ITU-T, comparable con TCP en el stack de protocolos de Internet. Q.931 no provee control de flujo y retransmisiones, dado que las capas subyacentes se asumen confiables.

QoS: ver calidad de servicio.

Quicknet's Technologies: empresa dedicada a la creación de productos de telefonía sobre Internet. http://www.quicknet.net

Radius: Remote Authentication Dial-In User Service es un sistema de accounting y autenticación, muy utilizado por ISPs.

RAS: Registration, Admission and Status.

Red de conmutación de circuitos: Red de datos en la cual el intercambio de información esta precedido por una etapa de establecimiento de un circuito lógico por el cual circulará dicha información. La PSTN es un ejemplo de este tipo de redes.

Red de conmutación de paquetes: Red de datos en la cual la información se encapsula en paquetes que son enviados en forma independiente del resto, pudiendo éstos tomar caminos diferentes o incluso llegar desordenados.

Red de mejor esfuerzo: Red de paquetes donde la pérdida de los mismos es aceptable.

RELEASE COMPLETE: mensaje H.323.

RJ-11: Registered Jack 11 es un conector de cuatro o seis hilos utilizado principalmente en las conexiones de equipos de telefonía.

RJ-45: Registered Jack 45 es un conector de 8 hilos utilizado en las conexiones LAN, especialmente Ethernet.

Roaming: es la capacidad de movilidad de un usuario cualquiera sea el servicio. El término puede utilizarse en diversos ámbitos como ser wireless, telefonía, IP, etc.

Router: dispositivo que retransmite paquetes entre redes diferentes. Estos dispositivos utilizan los encabezados de los paquetes y tablas de retransmisión para determinar el mejor camino que un paquete debe tomar.

RTCP: Real Time Control Protocol es un protocolo definido conjuntamente con RTP en la RFC 1889. Su motivo es mantener datos estadísticos sobre las conexiones RTP que pueden ser utilizados por las aplicaciones en tiempo real. http://www.ietf.org/rfc/rfc1889.txt

RTP: Real Time Protocol es un protocolo de Internet definido en la RFC 1889 para la transmisión de datos en tiempo real como es el caso del audio y el video. El protocolo no garantiza la entrega de datos en tiempo real, sino que provee mecanismos para que las aplicaciones que envían y reciben datos soporten streaming. http://www.ietf.org/rfc/rfc1889.txt

SDP: Session Description Protocol es un estándar definido en la RFC 2327 para la descripción de sesiones de streaming. SDP no es un protocolo de transporte sino un método por el cual se

detallan las características de las transmisiones multimedia como por ejemplo protocolos, codecs, formatos, etc. http://www.ietf.org/rfc/rfc2327.txt

Ser: SIP Express Router es un servidor SIP opensource cuyo mantenimiento es efectuado por Iptel. http://www.iptel.org

Serweb: interfaz web de administración para Ser. http://www.iptel.org

SETUP: mensaje SIP.

Single Sign On: proceso de autenticación cliente / servidor donde un usuario o cliente, ingresando su nombre y contraseña tiene acceso a más de una aplicación o un número de recursos dentro de la empresa. Single Sign On evita que un usuario ingrese información de autenticación cada vez que cambia de una aplicación a otra.

SIP: Session Initiation Protocol es un protocolo de señalización para conferencias sobre Internet definido en la RFC 2543. Además de conferencias, SIP es utilizado para notificación de eventos, presencia y mensajería instantánea. El protocolo permite iniciar llamadas, setear sus parámetros, rutearlas, autenticar usuarios y otras operaciones. http://www.ietf.org/rfc/rfc2543.txt

SMTP: Simple Mail Transfer Protocol es un protocolo utilizado para el envío de correos electrónicos entre servidores. La definición detallada del protocolo puede encontrarse en la RFC 821. Un cliente de correo electrónico se conecta a un servidor SMTP para entregar a éste último el mensaje a enviar. Luego el servidor se comunica con el servidor SMTP remoto quien, luego de recibir el mensaje, lo coloca en la casilla de correo correspondiente al dueño de la cuenta. La recuperación del correo, es decir la lectura de los buzones de correo electrónico se efectúa con protocolos como IMAP o POP. http://www.ietf.org/rfc/rfc0821.txt

SMS: Short Message Service es un servicio similar al utilizado por los pagers permitiendo el envío de mensajes de texto cortos a teléfonos celulares.

SNA: Systems Network Architecture es un conjunto de protocolos desarrollados por IBM. Originalmente fue diseñado para los mainframes IBM, sin embargo, hoy día soporta la interconexión de workstations.

SNMP: Simple Network Management Protocol es un conjunto de protocolos que permiten el manejo de redes complejas. SNMP trabaja enviando mensajes, llamados PDUs, a diferentes partes de la red. Los dispositivos que soportan SNMP, llamados agentes, almacenan datos sobre ellos mismos en MIBS y retornan estos datos a los servidores SNMP. http://www.ietf.org/rfc/rfc1157.txt

SS7: Signaling System Nº 7. SS7 es un estándar global para telecomunicaciones definido por la ITU-T. Define los procedimientos y protocolos mediante los cuales los elementos de la PSTN, intercambian información sobre una red de señalización digital para establecer, enrutar, facturar y controlar llamadas, tanto a terminales fijos como móviles. Más información en http://www.pt.com/tutorials/ss7/

Streaming: Técnica para la transferencia de datos de tal forma que puedan ser procesados como un flujo permanente y contínuo.

T1: Circuito digital punto a punto dedicado a 1,544 Mbps proporcionado por las compañías telefónicas Norteamérica. Ver E1 y J1 para los equivalentes europeos y japonés, respectivamente. Permite la transmisión de voz y datos y en muchos casos se utilizan para proporcionar conexiones a Internet.
T1 (DS1): 24 canales, 1,544 Mbps
T2 (DS2): 96 canalels, 6,312 Mbps
T3 (DS3): 672 canales, 44,736 Mbps
T4 (DS4): 4032 canales, 274,176 Mbps

T.120: Estándar definido por la ITU-T para las conferencias de datos, como por ejemplo, compartir programas o ventanas, pizarras o transferencia de archivos.

T.38: estándar definido por la ITU-T para el envío de fax sobre redes IP en tiempo real. La recomendación permite el uso de TCP o UDP.

TCAP: Transaction Capabilities Application Part provee servicios inteligentes en redes SS7. Entre estos servicios, TCAP es utilizado por un SCP con el fin de determinar los números asociados a un número 0800, 0810, o 0610. Otro ejemplo son las tarjetas telefónicas que utilizan TCAP para su validación.

TCP: Transmission Control Protocol es uno de los principales protocolos del stack TCP/IP definido en la RFC 793. TCP permite que dos máquinas establezcan una conexión e intercambien flujos de información. TCP garantiza la entrega de datos y que el orden de entrega de los datos es el mismo en que fueron enviados.

TCP/IP: es una serie de protocolos desarrollados para compartir recursos a través de una red, independientemente de los diferentes tipos de tecnología, software y modalidades que puedan utilizar los distintos servidores y clientes. Originalmente fue desarrollado por el Departamento de Defensa de los Estados Unidos. Entre los miembros de la familia de protocolos que hacen posible la interconexión y el tráfico de red en Internet están: FTP, SMTP, NNTP, entre otros. Los dos protocolos más importantes son los que dan nombre a la familia: "IP" y "TCP".

TerminalCapabilitySet: mensaje H.323

TerminalCapabilitySetAck: mensaje H.323

Token Ring: arquitectura de red LAN desarrollada por IBM. La especificación de token ring ha sido estandarizada por la IEEE como el estándar IEEE 802.5

TRIP: Telephony Routing over IP, definido en la RFC 3219. trata los casos en que las llamadas telefónicas deben ser ruteadas entre dominios.

Trunks: canal de comunicación entre dos puntos. Generalmente hace referencia a canales telefónicos de gran ancho de banda entre centros de conmutación que manipulan simultáneamente varias señales de voz y datos.

TUP: Telephony User Part es utilizado en algunas partes del mundo, como por ejemplo Argentina, Brasil y China, donde es necesario utilizar esta tecnología para soportar llamadas telefónicas básicas. TUP sólo maneja circuitos analógicos, y en muchos países fue reemplazado por ISUP.

UDP: User Datagram Protocol es un protocolo estándar de TCP/IP definido en la RFC 768. UDP es un protocolo de transporte, al igual que TCP, pero no garantiza entrega segura de los datos, ni control de flujo ni secuencia.

V.35: estándar de la ITU para el intercambio de datos a alta velocidad en forma sincrónica. Es una interfaz generalmente utilizada para conectar router a DSUs.

Vocal: servidor SIP opensource desarrollado por Vovida. http://www.vovida.org

VoIP: Voice Over IP es una categoría de hardware y software que permite a los usuarios de Internet utilizar este medio para efectuar llamadas telefónicas.

Vovida: sitio dedicado a la comunidad de las comunicaciones motivando las discusiones en el uso de software open source en ambientes de telecomunicaciones y datos. http://www.vovida.org

W3C: World Wild Web Consortium desarrolla tecnologías de interoperabilidad (especificaciones, guías, software, y herramientas) para llevar la WEB a su máximo potencial. W3C es un foro para el comercio, la información, las comunicaciones, y el entendimiento colectivo. http://www.w3c.org/

WAN: Wide Area Network es una red geográficamente dispersa. Generalmente una WAN conecta dos o más LANs.

X.25: X.25 es una recomendación del CCITT para el interfaz entre un DTE y un DCE sobre la PSTN. Generalmente, X.25 cubre las capas 1 a 3 del modelo de comunicaciones OSI, aunque muchas veces se utiliza este término para referirse específicamente a la capa de paquetes 3. X.25 se transporta dentro del campo Información de las tramas LAPB.

Bibliografía y Referencias

El libro “IP Telephony – The Integration of Robust VoIP Services” de Bill Douskalis, al igual que el sitio http://www.protocols.com fue material de consulta en todos los aspectos de esta tesis.

Respecto a cada uno de los temas particulares abordados durante este trabajo, hay bibliografía y referencias particulares para cada uno de ellos. Estas son:

PSTN

- Señalización en la PSTN: White paper de cisco Voice Network Signaling and Control: http://www.cisco.com/warp/public/788/signalling/net_signal_control.html
- http://www.iec.org/online/tutorials/fund_telecom/
- Capitulo de “Redes de Conmutación de Circuitos” del libro de Stallings Data and Computer Communications.
- http://www.iec.org/online/tutorials/ss7/
- http://www.tekelec.com/ss7/members/ssp.asp
- http://www.pt.com/tutorials/ss7/

H.323

- http://www.iec.org/online/tutorials/h323/index.html
- http://www.vovida.org/document/Training/1_VoIP_Overview/index_files/frame.htm
- http://www.openh323.org/
- http://www.packetizer.com/iptel/h323/

SIP

- http://www.vovida.org/document/Training/1_VoIP_Overview/index_files/frame.htm
- http://www.vovida.org/document/pdf/sip.pdf
- http://www.cs.columbia.edu/sip/
- http://www.packetizer.com/iptel/h323/

Comparación e/ SIP y H.323

- http://www.packetizer.com/iptel/h323_vs_sip/
- Service Architectures in H.323 and SIP – A Comparison, Josef Glasmann, Wolfgang Kellerer, Munich University of Technology (TUM), Harald Müller, Siemens AG, Germany: http://www.h323forum.org/papers/Service_Architecures_SIP-H323.pdf
- Lip sinchronization: http://www.its.bldrdoc.gov/fs-1037/dir-021/_3069.htm

Gateways

- http://www.tmcnet.com/articles/itmag/1299/1299it.htm

RTP

- http://www.vovida.org/
- http://www.cs.columbia.edu/~hgs/rtp/

Codecs

- Codecs de audio: http://www-obile.ecs.soton.ac.uk/speech_codecs/standards/
- MPEG 4: http://www.showshifter.com/support/codecinfo.htm
- General de codecs: http://www.seas.upenn.edu/~ross/book/emerge/multimedia.htm