

# El Protocolo SIP:

## PROPUESTA DEL IETF PARA LA TRANSMISIÓN DE VOIP

Por MsC. Nelson Poyeaux Cabrera

Especialista en Sistemas de Computación, Administrador de red, Presidencia de ETECSA

[nelsonp@etecsa.cu](mailto:nelsonp@etecsa.cu)

En el ámbito de las tecnologías relacionadas con las redes y los servicios IP, han ganado ventajas la sencillez, la flexibilidad y la robustez del protocolo SIP que se perfila como una apuesta clara del futuro.

El SIP —*Session Initiation Protocol* / Protocolo de Inicio de Sesión— es un protocolo de aplicación desarrollado por el IETF dentro del grupo MMUSIC —*Multiparty Multimedia Session Control*— y especificado en la RFC 2543. Permite a los usuarios participar en sesiones de intercambio de información multimedia a través de los mecanismos de establecimiento, modificación y finalización de llamadas. Existen dos modos básicos de identificar y participar en sesiones multimedia :

- ♦ Mecanismo de anuncio en el que las sesiones son anunciadas mediante correo electrónico, páginas web, grupos de noticias o mediante el Protocolo de Anuncio de Sesiones (SAP).

- ♦ Mecanismo de invitación donde los usuarios son invitados a participar mediante el Protocolo de Establecimiento de Sesiones (SIP).

Entre ambos, SIP ha sido propuesto como un mecanismo genérico para el soporte de mecanismos de señalización del servicio de telefonía IP. Soporta cinco elementos funcionales para el establecimiento y la terminación de comunicaciones multimedia: localización de usuarios, inter-

cambio / negociación de capacidades de los terminales, disponibilidad de usuarios, establecimiento de llamada, mantenimiento de llamada.

El protocolo SIP se utiliza para el establecimiento de las comunicaciones en tiempo real sobre redes IP. El contenido de una sesión puede ser de cualquier tipo —voz, video, datos— y puede modificarse en cualquier momento, añadiendo capacidades de los usuarios o variando el número de los mismos que participan en la comunicación. Las similitudes con HTTP y SMTP, hacen que SIP sea ideal para integrar la voz como un servicio más en Internet; además utiliza el protocolo MIME —*Multipurpose Internet Mail Extension* / Extensión de Correo de Internet Multipropósito— para la descripción de contenidos, el sistema de direccionamiento propuesto es del tipo de las URL de Internet y la localización de los servidores se hace vía DNS, características que convierten a SIP en el mecanismo ideal para las comunicaciones en un entorno de convergencia como el actual.

SIP tiene dos tipos de mensajes: peticiones (*request*) y respuesta (*response*). El mensaje de petición es emitido desde el cliente terminal al servidor terminal. El encabezado del mensaje de petición y de respuesta contiene campos similares:

- ♦ *Start Line*: se usa para indicar el tipo de paquete, la dirección y la versión de SIP, por ejemplo, <INVITE SP sip:roberto@teleinfo.com.ar SP SIP/2.0 CRLF>.

- ♦ *General Header*: el encabezamiento general contiene las informaciones: *Call-ID* se genera en cada llamada para identificarla. Contiene la dirección del dominio del *host*, por ejemplo, <CallID: 1876@foo.bar.com>.

*Cseq* se inicia en un número aleatorio e identifica en forma secuencial a cada petición, por ejemplo, <Cseq: 1234 Invite>

*From* es la dirección del origen de la llamada y se encuentra presente en toda petición y respuesta, por ejemplo, <From: "MyName" <sip:myaccount@company.com>>.

*To* es la dirección del destino de la llamada y se encuentra presente en todo *request* y *response*, por ejemplo, <To: "Helpdesk" <sip:helpdesk@company.com;tag=287447>>. El *tag* es usado cuando el mismo destino designa a varios puntos finales.

*Via* sirve para recordar la ruta del *request*; por eso cada proxy en la ruta añade una línea de vía, por ejemplo, <Via:SIP/2.0/UDP PXY1.provider.com;received 10.0.03>.

*Encryption* identifica un mensaje que ha sido encriptado para seguridad —proceso descrito en SDP, es del tipo MD5—, por ejemplo, <Encryption: PGP versión=2.6.2, encoding=ascii>.

- ♦ *Additional*s: además del encabezado general, pueden transportarse campos adicionales, por ejemplo, *Expire* que indica el tiempo de validez de registro, *Priority* que indica la prioridad del mensaje, etc.

## Métodos de mensajes

Se han definido seis métodos para los mensajes de solicitud respuesta —*request-response*—:

1. *Invite*: para invitar al usuario a realizar una conexión. Localiza e identifica al usuario.
2. *Bye*: para la terminación de una llamada entre usuarios.
3. *Options*: información de capacidades que pueden ser configuradas entre agentes o mediante un servidor SIP.
4. *ACK*: usado para reconocer que el mensaje *Invite* puede ser aceptado.
5. *Cancel*: termina una búsqueda de un usuario.
6. *Register*: emitido en un mensaje multicast para localizar al servidor SIP.

Los clientes SIP envían peticiones —*Request Messages*— a un servidor que, una vez procesadas, contesta con una respuesta —*Response Messages*—. Las respuestas son del tipo HTTP:

- 1xx Informacional —100 probando, 180 timbrando, 181 llamadas siendo reenviadas—.
- 2xx Éxito —200 OK, 202 aceptado—.
- 3xx Redirección —300 selección múltiple, 301 movido permanentemente, 302 movido temporalmente—.
- 4xx Error de cliente —400 petición incorrecta, 404 no encontrado, 482 lazo detectado, 486 ocupado aquí—.
- 5xx Falla de servidor —500 error interno del servidor, 501 no implementado—.
- 6xx Falla global —600 en cualquier lugar ocupado, 603 rechazo—.

Los terminales SIP pueden generar tanto peticiones como

respuestas, al estar formados por el denominado Cliente del Agente de Usuario (UAC) y Servidor del Agente de Usuario (UAS).

Los terminales SIP pueden establecer llamadas de voz directamente sin la intervención de elementos intermedios. La figura 1 muestra un ejemplo de conexión entre los usuarios *user1* con dirección IP 172.16.10.1 y *user2* con dirección IP 172.16.1.2 mediante el envío de una petición *Invite Request*, en la cual el *user1* indica al *user2* las capacidades de recepción de audio y el puerto donde espera recibir dicho audio —*port 12345*—. Al recibir la petición, el *user2* puede establecer inmediatamente el canal de voz y enviar la aceptación de conexión mediante el envío de *OK Response*, en la cual incluye la información complementaria para el establecimiento del canal opuesto —codificación GSM, puerto 54321 en nuestro ejemplo—. Tras el intercambio de señal de audio, cualquiera de los participantes puede finalizar la llamada a través del envío de mensaje *Bye Request* que, a su vez, debe recibir mensaje de confirmación (OK).

Para codificar los mensajes SIP se utiliza la sintaxis de mensajes definidos en HTTP/1.1 y el contenido de cada mensaje sigue las recomendaciones del Protocolo de Descripción de Sesiones (SDP).

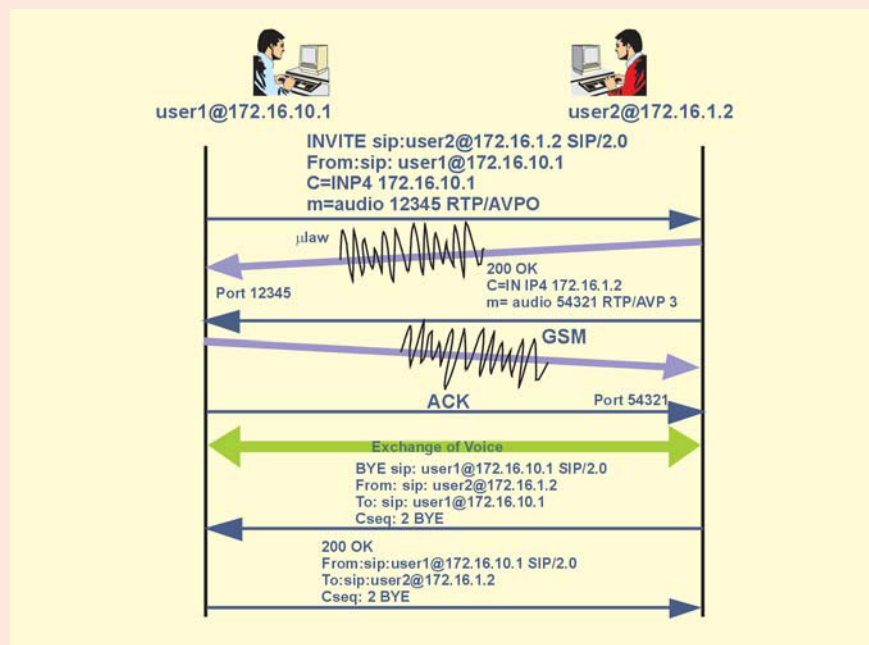


Figura 1 Establecimiento de llamada SIP

La arquitectura SIP define cuatro tipos de servidores:

**Servidor proxy**: se encarga de encaminar peticiones / respuestas hacia el destino final. El encaminamiento se realiza salto a salto de un servidor a otro hasta alcanzar el destino final. Para estos casos existe un parámetro incluido en las peticiones / respuestas denominado *vía* que incluye los sistemas intermedios que han participado en el proceso de encaminamiento

lo cual evita bucles y permite forzar que las respuestas sigan el mismo camino de las peticiones. Esto afecta únicamente a la información de control pues el transporte de medios, salvo en el caso de requerir transcodificación intermedia, se realiza directamente entre origen y destino.

**Servidor de redirección:** realiza una función equivalente al servidor proxy, pero a diferencia de este no progresa la llamada, sino que contesta a un *Invite* con un mensaje de redirección indicándole cómo contactar con el destino.

**Servidor de registro:** mantiene la localización actual de un usuario, se utiliza para que los terminales registren la localización en la que se encuentran. Este servidor facilita la movilidad de usuarios, pues la actualiza dinámicamente.

**Agente de llamada —Call Agent—:** realiza las funciones de los tres servidores anteriores, además de poder realizar acciones como localizar a un usuario mediante la redirección de la llamada a una o varias localizaciones, implementar servicios de redirección como reenvío si está ocupado, reenvío si no contesta, etc., implementar filtrado de llamada en función del origen o del instante de la llamada, almacenar información de administración de llamadas, y realizar cualquier otra función de gestión.

Las direcciones SIP son identificadas mediante los denominados URI —*Uniform Resource Identifiers* / Identificadores de Recursos Uniformes—, que siguen la estructura *user@host*, donde *user* corresponde a un nombre, identificador o número telefónico, y *host* es el dominio al que pertenece el usuario o dirección de red.

En la figura 2 se muestra un ejemplo de interacción entre servi-

dores SIP. En este ejemplo david, desde su oficina (company.es), desea llamar al usuario jmoreno del dominio upm.es. Para realizar la llamada envía una petición —*Notify Request*— al servidor SIP de su organización, que actúa como un servidor proxy y, tras consultar el DNS, localiza el servidor SIP del dominio upm.es, reenviándole la petición. Este servidor, que actúa como servidor de redirección, contesta a la petición indicándole que el usuario jmoreno se encuentra localizado en otro dominio (uc3m.es). El servidor sip.company.es progresa de nuevo la llamada hacia el servidor SIP del dominio uc3m.es, que ahora actúa como un agente —*Call Agent*— y después de consultar la base de datos DNS intenta localizar al usuario en el sistema *host1.uc3m.es*. Tras un período de espera y al no contestar el usuario en dicho terminal, cancela la llamada e intenta localizar al usuario en *host2.uc3m.es*, quien ahora contesta. La aceptación de llamada progresa hasta el origen, pasa por

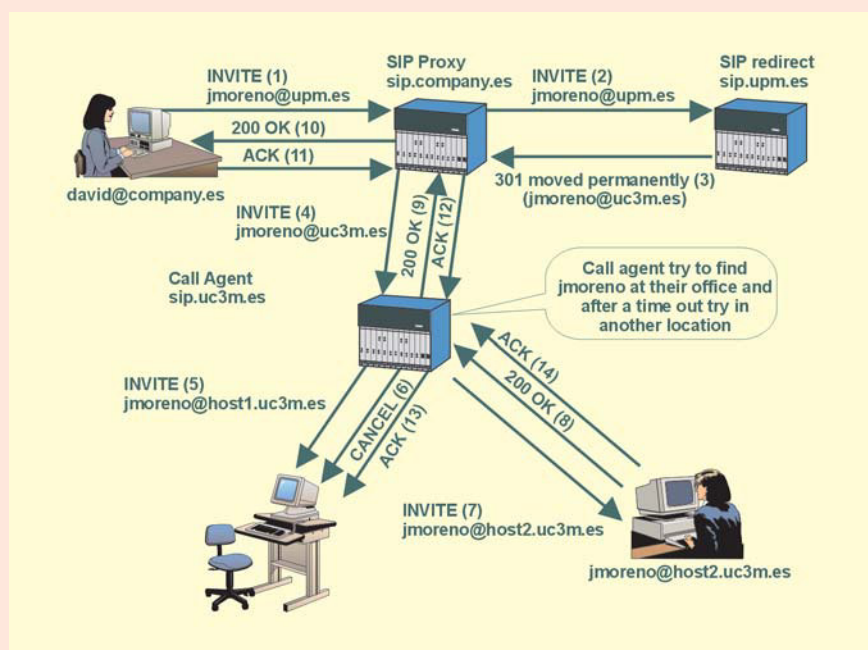


Figura 2 Ejemplo de servidores SIP

los servidores de redirección involucrados, en este momento origen y destino pueden establecer los canales de voz.

La principal característica de SIP frente a H.323 es su simplicidad, dentro de H.323 se incluye un conjunto de protocolos perfectamente integrados que participan en el establecimiento y mantenimiento de conferencias multimedia: Q.931 para el establecimiento de llamada, H.225 para la señalización, H.245 para la negociación de capacidades y el establecimiento de canales, H.450.x para la definición de servicios suplementarios —*Call Park, Call Pickup, Call Hold, Call Transfer, Call Diversion, MWI*—, RAS para el registro de terminales y el control de

admisión, RTP/RTCP para el transporte y sucesión de los flujos multimedia, G.711/G.712 para la especificación de los *codecs*, T.120 para colaboración y conferencia de datos. Esto identifica una de las características menos agradables de este protocolo, y que siempre han argumentado sus detractores: la excesiva complejidad frente a la sencillez del modelo Internet en que se basa SIP. De hecho SIP se podría comparar, *grosso modo*, con las partes de Q.931 y H.225 de H.323.

### Líneas de trabajo y esfuerzos de estandarización relacionados con el protocolo SIP

#### Interoperabilidad del protocolo SIP con redes tradicionales

Teniendo en cuenta el entorno de convergencia donde deben integrarse Internet, la Red Telefónica Pública Conmutada (RTPC) tradicional y toda la pléyade de tecnologías y servicios móviles que están proliferando, SIP debe ser capaz de establecer sesiones sobre una red multiservicio IP y contemplar un escenario de convivencia entre los nuevos servicios y una gran cantidad de servicios tradicionales que se prestarán sobre la RTPC; de manera que se han originado iniciativas para garantizar la portabilidad de esos servicios entre las que se encuentran:

*Pint*: el objetivo de esta iniciativa es permitir el acceso vía Web a contenidos de voz y habilitar el despliegue de servicios de *dial/fax-on-click*.

*Spirits*: trata el camino recíproco, por ejemplo, peticiones que se

originan en la RTPC y que se sirven desde la red IP.

*Trip* —ruteo telefónico sobre IP—: está pensado para que los proveedores de servicio intercambien información de encaminamiento, así evitan el sobredimensionamiento de la infraestructura de pasarelas. Utilizan protocolos como BGP, OSPF y SCSP.

SIP-T —SIP para teléfonos—: se trata de buscar un mecanismo para interconectar la red IP con la RTPC con la utilización de SIP. El objetivo es permitir la gestión de servicios tipo red inteligente en un entorno tipo Internet de forma transparente y homogénea. SIP-T realiza la correspondencia entre las funciones de SIP y los requisitos de ISUP, por ejemplo, se encapsula la información de los mensajes ISUP en el cuerpo de los mensajes SIP, las dependencias de los de ISUP se mantienen al traducir la información a las cabeceras de los mensajes SIP, mantienen así su integridad hasta entregarla en los puntos de interconexión. La señalización en el transcurso de la llamada se pasa con un método INFO.

WTA: pretende ser un entorno de trabajo para el desarrollo de servicios tipo red inteligente, que siempre trata de buscar la interoperabilidad entre RTPC/IN y los servicios IP. No está claro cómo se realizaría la correspondencia entre los mensajes SIP y WAP. Se investiga con diferentes escenarios SIP  $\leftrightarrow$  WAP con aplicaciones como la Mensajería Instantánea (IM) aprovechando la similitud con el servicio de mensajes cortos (SMS) de la telefonía móvil celular.

SIP y el 3GPP —*Third Generation Partnership Project*—: el grupo de trabajo del 3GPP se dedica a la elaboración de especificaciones técnicas para la tercera generación de sistemas móviles. La propuesta de este grupo, desde el punto de vista conceptual, se basa en la utilización de las tecnologías IP para la comunicación extremo a extremo de terminales móviles heterogéneos; luego deben contemplar los desarrollos del IETF y, de hecho, han elegido SIP para el control de llamada y señalización.

En un escenario de 3G, los usuarios se identificarán mediante una URL SIP o un número E.164; el sistema portador —GPRS o Mobile IP— se ocupará de la micromovilidad, mientras que SIP soportará la macromovilidad —entre dominios—. SIP se encargará de encaminar la señalización de manera que los servicios estén disponibles desde la red que origina la llamada hasta la que la termina.

#### Interfaces de Software

La similitud del protocolo SIP con los protocolos HTTP y el SMTP y el hecho de que se base en el formato de texto plano, hacen que sea del agrado de los desarrolladores Web, por lo tanto, se realiza un esfuerzo considerable en la definición de API para el desarrollo de servicios:

CPL —*Call Processing Language / Lenguaje de Procesamiento de Llamada*—: fue la primera API que se desarrolló para SIP; aunque es más un lenguaje de *scripts* basado en XML. Está concebido para la creación de servicios para el usuario final. No maneja variables, ni bucles, ni puede ejecutar programas externos.



Incorpora primitivas para tomar decisiones y acometer acciones atendiendo a los parámetros de una llamada.

SIP-CGI: es prácticamente idéntica a la CGI de HTTP. Se trata de *scripts* que residen en el servidor y se dedican a pasar parámetros de los mensajes a diferentes procesos a través de variables de entorno; esperan la respuesta de esos procesos en forma de instrucciones de vuelta que pasarán mediante el correspondiente descriptor de fichero estándar de salida. Los *scripts* de CGI se pueden escribir en Perl, C, C++, TCL, Java, etc., prácticamente cualquier variante es válida.

SIP-Servlets: de la misma manera que existe CGI para SIP, también está la alternativa de utilizar servlets; el concepto es el mismo que con HTTP, no se utilizan procesos separados para pasar los mensajes, sino que se realiza mediante una clase de Java que se ejecuta sobre una máquina virtual (JVM) en el servidor Web o de aplicaciones —es la alternativa Java a los *scripts* de CGI, las ASP o los *plug-in* propios escritos en C—. Se supone que al estar escritos en Java son independientes de la plataforma.

JAIN API: es una extensión de Java para la creación de servicios sobre redes tipo RTPC/IN. El objetivo de esta iniciativa es atacar los aspectos de portabilidad de los servicios *Write-Once-Run-Anywhere*, convergencia de redes —dar servicio sobre cualquier infraestructura de red: IP, ATM, TDM, WLAN— y acceso a servicios —la abstracción del nivel de servicio debe permitir el acceso a cualquier recur-

so de red—. Se trabaja en tres líneas dentro de esta iniciativa:

1. JAIN SIP es una API de bajo nivel, que hace una implementación de la RFC 2543.
2. JAIN SIP Lite es una API de alto nivel; permite una mayor abstracción del SIP por parte de los desarrolladores.
3. JAIN SIP Servlets es la misma alternativa a la CGI que se comentó anteriormente.

Parlay: tiene como objetivo que las aplicaciones puedan acceder a las funcionalidades que ofrecen las redes de telecomunicaciones de forma segura. Es otro esfuerzo en la línea de acercar los mundos de las tecnologías de la informática y las telecomunicaciones tradicionales y conseguir que funcionen en un escenario integrado de convergencia tecnológica. Se definen dos categorías: servicio y entorno.

### Aplicaciones del protocolo SIP

La filosofía y arquitectura que propone SIP, y que se ha perfilado en los apartados anteriores, ofrece muchas posibilidades para el desarrollo de servicios en un entorno convergente. A partir de los ingredientes básicos —voz, video y datos— pueden componerse infinidad de servicios nuevos de valor añadido para el entorno corporativo, e incluso, doméstico. No debe olvidarse que el desarrollo depende tanto de los proveedores de servicio como de los departamentos de tecnologías de la

informática, y hasta de los usuarios finales.

La incorporación de la voz como un servicio vía Web ofrece muchas posibilidades para el comercio electrónico: puede incluirse un servicio de click para marcar en un sitio Web y poner en contacto al usuario —que en Internet puede ser el cliente final del banco— con un experto de producto, con una línea de atención al cliente. Puede pensarse en un verdadero centro de contactos Web en el que el cliente contacte con un agente por cualquier medio: voz, correo electrónico, mensajería instantánea, videoconferencia.

En el escenario de convergencia del que se habla han surgido todo tipo de terminales móviles con capacidades muy heterogéneas que exigen la disponibilidad de medios que permitan el acceso a servicios asociados al usuario que los demanda, y no a los requisitos propios de cada sistema o plataforma. Una vez más parece que SIP se perfila como el candidato perfecto, pues permite a los terminales negociar el medio de comunicación y las capacidades y recursos de cada uno. Además, los clientes SIP son lo suficientemente ligeros como para residir en prácticamente cualquier dispositivo, ya sea un teléfono celular, una Agenda Portátil Digital (PDA), o una computadora portátil; en una arquitectura basada en SIP pueden ofrecerse servicios de red inteligente a una computadora de escritorio de la misma forma que se hace con un teléfono fijo. Nadie desconoce que habrá limitaciones relacionadas con el ancho de banda y las dimensiones de las pantallas de los diferentes dispositivos.

Se ha planteado que SIP era un candidato; pero es mucho más que eso, porque el 3GPP lo ha seleccionado como protocolo de señalización para los sistemas móviles de tercera generación.

La arquitectura que propone SIP permite también tener los perfiles de los usuarios en una base de datos centralizada a la que accedan los dispositivos. De esta forma, cualquier actualización que se haga en esos perfiles será recogida por el dispositivo que utilicen aquellos en cada caso para el acceso a los servicios. Precisamente esta capacidad para tener la información de usuario asociada al usuario y no al dispositivo resulta crucial para las aplicaciones de mensajería unificada; además las similitudes con SMTP y la utilización del DNS hacen que la integración de la voz con el correo electrónico sea prácticamente inmediata. También puede valorarse el SIP para el lanzamiento de servicios asociados a la localización de usuarios móviles.


## Conclusiones

El escenario actual seguirá evolucionando hacia la convergencia tecnológica efectiva: la tendencia es ir a un escenario final donde esté disponible una red multiservicio que integre todo tipo de contenidos —voz, video y datos— y que permita entregarlos de forma perso-

nalizada a cualquier tipo de usuario, en cualquier terminal, con la calidad requerida e independientemente de la ubicación de aquel.

Hoy esa evolución, ante un escenario basado en redes y servicios IP, ha ganado muchos puntos la sencillez, flexibilidad y robustez de SIP; y se perfila como propuesta futura que, por otra parte, se veía venir como contrapartida desde el mundo de Internet ante las iniciativas del mundo institucional. El esfuerzo de convergencia se centra actualmente en los puntos de interconexión entre la RTPC y la red IP y, por lo tanto, interesará que los servidores SIP estén junto a los *softswitches*; pero, según se avance hacia un escenario más integrado, la atención se centrará en la infraestructura IP, con lo cual la función de los *softswitches* se alejará de los puntos de interconexión. Finalmente, en un entorno IP puro la función de creación de servicios se distribuirá por toda la red, donde puede extenderse el modelo ASP para dar servicios de voz. Tanto los ASP como los ISP, incluso los usuarios finales, pueden crear sus servicios.

Ante la creciente demanda de servicios de multiconferencia y multimedia en tiempo real, con funcionalidades de colaboración, parece que son los desarrollos li-

gados a SIP los que brindan una respuesta más satisfactoria a los retos que presenta la implementación de tales servicios. Además, en un mundo IP con la Web como canal de comunicación por defecto y una oferta cada vez mayor de tecnologías de acceso y soluciones de movilidad, tiene sentido apostar decididamente por SIP. 

## Bibliografía

Díaz-Guerra, Carlos. "Visión de las Telecomunicaciones 2003-2006". Revista AHCIET, no. 96 (noviembre 2003): 18-24.

Lima, Armando F. "IP Telephony y NGN" (2002). Disponible en: [web.frm.utn.edu.ar/codarec/Cursos/NGN\\_IPTel.pdf](http://web.frm.utn.edu.ar/codarec/Cursos/NGN_IPTel.pdf) (Consultado: 15-10-05).

Moreno, José Ignacio; Soto, Ignacio. "Protocolos de señalización para el transporte de voz sobre Redes IP". Disponible en: <http://www.ati.es> (Consultado: 15-10-05).

RFC 2327 - SDP: Session Description Protocol (04/1998). Handley, M; Jacobson, V. Disponible en: <http://www.faqs.org/rfcs/rfc2327> (Consultado: 5-11-05).

Ríos, Javier; García Moraima. "Softswitch" (Monografías). Disponible en: <http://www.monografias.com/trabajos14/softswitch/softswitch.shtml> (Consultado: 26-10-05).

Simon Znaty, Effort. *Voz sobre IP*. SIP. OT: Observatorio Tecnológico, The 30 Cents Consultancy.