



**CLIENTE MÓVIL SIP PARA UN SISTEMA IVR BASADO EN LA
ARQUITECTURA DE SERVICIOS IMS**

**JOSÉ FERNANDO MUÑOZ BERMEO
XIMENA VELASCO MELO**

ANEXO A

ARQUITECTURA DE SERVICIOS IMS, PROTOCOLO SIP Y SERVICIOS IVR SOBRE IMS

**Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Telemática
Popayán
2006**



TABLA DE CONTENIDO

	Página
A1.ARQUITECTURA DE SERVICIOS IMS Y SERVICIOS IVR SOBRE IMS	5
A1.1.ARQUITECTURA DE SERVICIOS IMS	5
A1.1.1. Origen y definición	5
A1.1.2. Características	6
A1.1.3. Arquitectura	8
A1.1.3.1. Capa de Acceso y Transporte	9
A1.1.3.2. Capa de Control de Sesión	9
A1.1.3.3. Capa de Aplicaciones	13
A1.1.4. Estado actual y evolución	16
A1.1.5. Principales tecnologías y protocolos	17
A1.1.5.1. SIP	17
A1.1.5.2. IPv6	17
A1.1.5.3. RTP y RTCP	17
A1.1.5.4. COPS	17
A1.1.5.5. DIAMETER	18
A1.1.5.6. RSVP Y DIFFSERV	18
A1.1.5.7. MEGACO	18
A1.1.5.8. SIGTRAN	18
A1.2. PROTOCOLO DE INICIO DE SESIÓN SIP	18
A1.2.1. Definición y origen de SIP	18
A1.2.2. Funciones	19
A1.2.3. Pila de protocolos	20
A1.2.4. Dirección	21
A1.2.5. Arquitectura	21
A1.2.5.1. Entidades	21
A1.2.5.1.1. Principales componentes	21
A1.2.5.1.2. Servidores SIP: Servidor de Medios	23
A1.2.6. Peticiones y Respuestas	24
A1.2.6.1. Peticiones	24
A1.2.6.2. Códigos de respuesta	25
A1.2.6.3. Formato de las peticiones y respuestas	26
A1.2.7. Flujo de mensajes SIP: ejemplos	29
A1.2.7.1. Registro	29
A1.2.7.2. Establecimiento de sesión	29
A1.2.7.3. Establecimiento de sesión involucrando un servidor Proxy SIP	30
A1.2.7.4. Mensajes instantáneos	30
A1.2.7.5. Terminación y cancelación de sesión	31
A1.2.7.6. Control de llamada REFER	32
A1.2.8. Creación de servicios	33



A2. OPERACIÓN DE IMS EN CONJUNCIÓN CON SIP	34
A2.1. IMS Y SIP	34
A2.2. CABECERAS DEL PROTOCOLO SIP PARA IMS	34
A2.2.1. P-Associated-URI	34
A2.2.2. P-Called-Party-ID	35
A2.2.3. P-Visited-Network-ID	35
A2.2.4. P-Access-Network-Info	35
A2.2.5. P-Charging-Function-Addresses	36
A2.2.6. P-Charging-Vector	36
A2.3. PROCEDIMIENTO DE REGISTRO	37
A2.4. ESTABLECIMIENTO DE SESIÓN	38
A2.5. FLUJOS DE SESIÓN	40
A3. SERVICIOS IVR SOBRE IMS	42
A3.1. SISTEMAS IVR	42
A3.1.1. IVR (Interactive Voice Response)	42
A3.1.2. Conversión de Texto a Habla	43
A3.1.2.1. El proceso de la síntesis del habla	44
A3.1.3. Reconocimiento Automático del Habla (RAH)	45
A3.1.3.1. Tipos de motores de reconocimiento	46
A3.1.3.2. Análisis de la Señal Acústica	46
A3.1.3.2.1. Reconocimiento basado en Patrones	47
A3.1.3.2.2. Modelos Ocultos de Markov	47
A3.2. SERVICIOS IVR EN IMS	48
A3.2.1. La Voz Sobre IP (VoIP) en las Redes Móviles	48
A3.2.1.1. Establecimiento del camino de habla	49
A3.2.1.2. Señalización	50
A3.2.1.3. Transporte de los paquetes de habla	51
A3.2.2. Servicios Multimedia en IMS	53
A3.2.2.1. Arquitectura de la MRF	53
A3.2.2.2. Interacciones que involucran al MRFC y al MRFP	54
A3.2.2.3. Requerimientos funcionales para la provisión de capacidades multimedia	54
A3.2.2.4. Procedimientos del MRFC y MRFP en la provisión de ASR, TTS, y detección DTMF	56
A3.2.2.5. Requerimientos de la arquitectura IMS para los servicios de reconocimiento del habla	56
A3.2.2.5.1. Técnicas para mejorar el desempeño del reconocimiento de habla	56
A3.2.2.5.2. Impactos en la arquitectura IMS por las mejoras en el reconocimiento de habla	58
A3.2.2.5.3. Tarifación en IMS por la funcionalidad ASR	59
A3.2.2.5.4. Seguridad en IMS para la funcionalidad ASR	59
A3.2.3. Servicios IVR en IMS que se adapten a las redes 2.5G actuales	60
A4. GLOSARIO	61
A5. BIBLIOGRAFÍA	66



LISTA DE FIGURAS

	Página
Figura A1. Arquitectura de Servicios IMS	8
Figura A2. Operación entre el HSS y el I-CSCF	12
Figura A3. Pila de protocolos que actúan junto con SIP	20
Figura A4. Formato de una petición SIP	26
Figura A5. Formato del protocolo SDP	29
Figura A6. Registro mediante mensajes SIP	29
Figura A7. Establecimiento de sesión mediante mensajes SIP	30
Figura A8. Establecimiento de sesión involucrando un servidor Proxy SIP	30
Figura A9. Envío de mensajes instantáneos mediante mensajes SIP	31
Figura A10. Terminación de sesión mediante mensajes SIP	31
Figura A11. Cancelación de sesión mediante mensajes SIP	32
Figura A12. Transferencia de llamada mediante mensajes SIP	32
Figura A13. Procedimiento de registro en IMS	37
Figura A14. Procedimiento de establecimiento de sesión en IMS	39
Figura A15. Proceso de Síntesis del Habla	44
Figura A16. Proceso de Reconocimiento del Habla basado en patrones	47
Figura A17. Pila de protocolos utilizada en VoIP para redes móviles	50
Figura A18. Transporte de SIP y SDP en IP	51
Figura A19. Transporte del habla en IP	52
Figura A20. Arquitectura de la MRF	53

LISTA DE TABLAS

	Página
Tabla A1. Códigos de respuesta SIP	26
Tabla A2. Cabeceras de un mensaje SIP	27



A1. ARQUITECTURA DE SERVICIOS IMS Y SERVICIOS IVR SOBRE IMS

A1.1. ARQUITECTURA DE SERVICIOS IMS

A1.1.1. Origen y definición [1][2][3][4]

Las redes de primera generación (1G), año 1980, soportaban los servicios de voz tradicionales. La segunda generación de redes (2G), año 1990, tenía voz y una tasa de transferencia de datos en el rango de 14.4-28.8Kbps, en los servicios de datos. Una de las redes dominantes en la 2G es el Sistema Global para Comunicaciones Móviles (Global System for Mobile Communications, GSM). Esta tecnología fue diseñada para circuitos conmutados de voz, transferencia de datos a tasas pequeñas, y no tiene un buen diseño para los paquetes IP. Para continuar con la evolución, el Servicio General de Paquetes de Radio (General Packet Radio Service, GPRS), se sumó a las redes GSM y estas redes a su vez hacen referencia a una generación 2.5G. No obstante las tasas de datos en GPRS, el Sistema de Telecomunicaciones Móviles Universales (Universal Mobile Telecommunications System, UMTS), la tercera generación (3G) de los sistemas inalámbricos, y que es estandarizado por el 3GPP, puede ofrecer tasas de datos en el rango de 2Mbps y dar soporte a video, audio y servicios de localización. La estandarización de UMTS ha progresado por fases y en su desarrollo y evolución se observa la tendencia de ir migrando hacia una infraestructura de transporte IP.

La primera fase la constituye el Release 99 de UMTS, el cual se caracteriza por proponer como elemento nuevo, una red de acceso radio UTRAN basada en ATM. El Núcleo de Red (Core Network, CN) reutiliza la infraestructura de los sistemas GSM y GPRS: Multiplexación por División de Tiempo (Time Division Multiplexing, TDM) para el dominio de la Conmutación de Circuitos (Circuit Switching, CS) e IP para el dominio de la Conmutación de Paquetes (Packet Switching, PS).

El Release 4 de UMTS añade como novedad la posibilidad de una infraestructura IP para el dominio CS, del núcleo de la red, por lo que todo el CN se plantea ahora sobre una infraestructura IP, preparando el camino para "all-IP" o "Todo IP".

Posteriormente, el Release 5 de UMTS es donde se introduce el Subsistema Multimedia IP (IP Multimedia Subsystem, IMS) en el dominio PS, para dar soporte a terminales con aplicaciones multimedia, a la vez que se plantea también una infraestructura de transporte IP alternativa para la red de acceso radio. Con este subsistema se propone una red convergente basada en paquetes, sobre una sola red, para ofrecer tanto servicios tradicionales de telefonía, como de paquetes conmutados.

El objetivo de IMS es llevar la evolución de las redes de operadores fijos y móviles hacia un escenario de convergencia. Este estándar, desarrollado por el 3GPP y el 3GPP2, se basa en el Protocolo de Inicio de Sesión (Session Initiation Protocol, SIP), estandarizado por el IETF. IMS permite la creación de Redes de Próxima Generación (Next Generation Network, NGN), además de nuevas aplicaciones multimedia capaces de suministrar al usuario final una experiencia de comunicaciones integrada, independientemente del tipo de aplicación, método de acceso a la red,



e infraestructura de la misma, de modo que los usuarios finales tengan acceso a servicios desde cualquier sitio, en cualquier momento, y sobre múltiples tecnologías de acceso [5].

Por otro lado, IMS posibilita a la industria de las telecomunicaciones implementar la capa de control de sesiones y servicios de una manera estructurada y estandarizada, lo cual permitirá a las empresas de comunicaciones poder desplegar servicios de una manera más eficiente y rápida, orientada a una integración horizontal de las plataformas, donde los facilitadores de servicios y las funciones comunes se pueden reutilizar para múltiples aplicaciones. Los operadores podrán proteger sus inversiones en tecnología, reduciendo los costos y el tiempo de introducción de nuevos servicios, al tomar ventaja de una infraestructura común de comunicaciones.

IMS es un sistema totalmente independiente del acceso, lo que permite un escenario de desarrollo de servicios sobre IP y servicios multimedia para usuarios conectados a través de redes fijas, Líneas de Abonado Digitales Asimétricas (Asymmetric Digital Subscriber Line, ADSL), WLAN, WIMAX, GPRS, WCDMA, CDMA2000, UMTS, o cualquier otro acceso, lo cual posibilita la convergencia, con el consecuente beneficio para operadores (creación de nuevos servicios y sinergias en las inversiones) y usuarios (facilidad de uso y acceso a servicios desde cualquier equipo terminal).

IMS permite y optimiza servicios móviles multimedia en tiempo real y no real. Entre los servicios IMS se pueden incluir el de Presionar Para Hablar para celulares (Push To Talk Over Cellular, PoC), compartir video en tiempo real, juegos interactivos, servicios de mensajería instantánea, mensajería de voz, telefonía de voz y video, y videoconferencia. [6]

Algunas designaciones de IMS por parte de organismos de estandarización son las siguientes:

- 3GPP: IMS en las fases 3GPP R5 y 3GPP R6
- 3GPP2: MMD
- TISPAN: Multimedia NGN
- PacketCable: Multimedia PacketCable

A1.1.2. Características

Como IMS fue diseñado para una red evolucionada de GSM, hereda ciertas cualidades intrínsecas del mundo móvil, como lo son [4][7]:

- La itinerancia. IMS soporta la itinerancia o “roaming” de tipo nativo, lo que se define como la capacidad del sistema de admitir y dar servicio a abonados de otros operadores que emplean la misma tecnología, y con los cuales se tiene el acuerdo de negocio pertinente.
- La interconexión con redes y servicios heredados. Existen elementos IMS para el interfuncionamiento entre las sesiones multimedia con los componentes de audio y las redes Telefónicas Públicas Conmutadas (Public Switched Telephone Network, PSTN), GSM o el dominio de conmutación de circuitos de la propia red 3G.
- La seguridad integrada. Uno de los factores clave del éxito de GSM fue la introducción de la tarjeta SIM para brindar mecanismos de seguridad. IMS requiere autenticación de abonado y especifica sus propios mecanismos y arquitectura de seguridad, los cuales son independientes de los propios de UMTS. De este modo, la suscripción IMS se soporta por una aplicación lógica llamada ISIM (IMS SIM), que ejecuta funciones de autenticación de abonado durante su registro en IMS. La ISIM contiene datos de la suscripción de abonado, de la misma forma que



la SIM en GSM y la USIM en 3G, y reside, junto con la aplicación USIM, en la tarjeta inteligente física. En este sentido, un abonado que desee acceder a IMS, en primer lugar deberá autenticarse y registrarse con el núcleo de red UMTS empleando la USIM, y posteriormente autenticarse y registrarse con IMS utilizando la ISIM.

- La Calidad de Servicio (QoS). El subsistema GPRS de 3G hace uso de la arquitectura de QoS de UMTS, que define una jerarquía de portadoras adecuadas para servicios con diferentes requisitos de QoS. Por lo anterior, se han definido funciones nuevas e interfaces opcionales en GPRS 3G, en el marco del Release 5, que permiten que IMS controle y autorice el uso de recursos del subsistema de transporte GPRS. Estas funciones son opcionales, y, en caso de no estar presentes en GPRS, la asignación de recursos a las sesiones IMS se controla mediante los mecanismos ordinarios propios y la suscripción de abonado de GPRS.
- La provisión de servicios. IMS posibilita un desarrollo de servicios rápido y simplificado al seguir el modelo Internet. Las aplicaciones creadas pueden modificar el transcurso de una sesión multimedia de una forma muy similar a cómo lo hacen las aplicaciones de red inteligente, que pueden iniciar y modificar una llamada de voz, con la ventaja de la simplicidad y facilidad del desarrollo de las aplicaciones Web. De este modo, se pueden establecer una serie de criterios en la suscripción de usuario IMS, de forma que el control de una sesión se traspase a un servidor de aplicaciones, posibilitando la implementación de todos los servicios suplementarios tradicionales, así como de nuevos servicios avanzados. En este sentido, IMS define componentes de pasarela que interactúan con las plataformas del plano de servicios tradicional de la red 3G, como la Arquitectura de Servicios Abiertos (Open Services Architecture, OSA) y las Aplicaciones Personalizadas para Redes Móviles de Lógica Mejorada (Customized Applications for Mobile Networks Enhanced Logic, CAMEL). De esta forma, las aplicaciones heredadas OSA y CAMEL pueden también actuar sobre las sesiones multimedia IP, permitiendo la provisión de servicios IMS desarrollados por terceras partes.
- La tarificación y facturación. En la tarificación de servicios multimedia IP intervienen: el sistema de facturación de GPRS y el sistema de facturación de IMS. Este último registra los datos relacionados con la sesión IMS, tales como los usuarios implicados, la duración, los componentes multimedia empleados y la QoS autorizada, y los asocia a los registros correspondientes de tarificación de GPRS que se originaron como consecuencia del transporte de los flujos multimedia y la señalización de IMS, en el subsistema de transporte GPRS. De esta forma, es posible facturar los servicios según su duración, contenidos, volumen de datos, destino de la sesión o las diferentes combinaciones de los anteriores. Por otro lado, el sistema soporta tanto tarificación pospago y prepago, necesaria para atender a todo el mercado de clientes potenciales.

Otras características de IMS son las siguientes [8][9]:

- Soporte a servicios de comunicaciones no tiempo real, pseudo tiempo real y tiempo real.
- Conexiones orientadas por la sesión. Varias sesiones y servicios simultáneamente sobre la misma conexión de red
- Movilidad sin restricciones (movilidad de servicios, movilidad del usuario, y movilidad de la sesión), también permitiendo el control de servicios de la red origen
- Convergencia fija-móvil de servicios y de la operación de la red
- Independencia de la red de acceso. Servicios distintos debido al tipo de acceso
- Interfaces abiertas y estandarizadas para todos los elementos de red
- Base de datos centralizada para simplificación de la operación
- Interoperabilidad terminal – red. Variedad de terminales, terminales libres
- Permite al operador mantener el control y el modelo de negocio actual, ya que las dependencias con la red le ayudan a asegurar su posición. De esta forma, IMS asegura que el operador no sea relegado a un mero transportador de datos.



- Al ser un estándar aceptado por la industria, IMS permite un despliegue de servicios multimedia a un mercado masivo.
- La eficiencia en costos ayuda a disminuir el riesgo en el lanzamiento de nuevos servicios.

A1.1.3. Arquitectura [4][7][8][10]

La arquitectura de servicios IMS es una colección de funciones lógicas que se dividen en 3 capas:

- la Capa de Acceso y Transporte,
- la Capa de Control,
- y la Capa de Aplicaciones.

Estas capas y sus componentes básicos se muestran en la Figura A1.

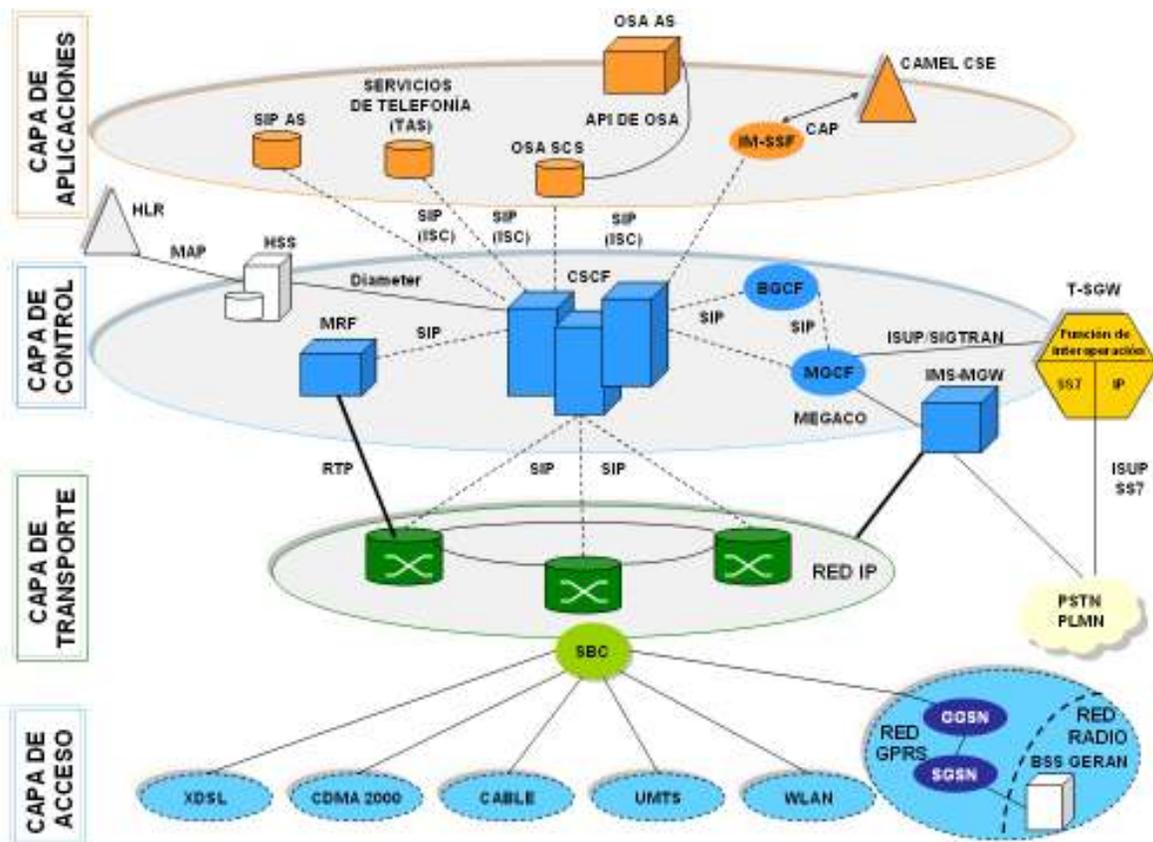


Figura A1: Arquitectura de Servicios IMS

En las siguientes secciones se describen las características más importantes de las entidades que conforman la arquitectura IMS.



A1.1.3.1. Capa de Acceso y Transporte

Conecta las diferentes redes de acceso con el corazón de la arquitectura IMS, mediante el empleo de pasarelas y servidores de control, los cuales se encargan, por ejemplo, de la conversión de los flujos VoIP al formato TDM de la PSTN.

Inicia y termina la señalización SIP mediante la cual se pueden configurar sesiones y proveer los servicios que se soportan utilizando el protocolo de Transporte en Tiempo Real (Realtime Transport Protocol, RTP).

A1.1.3.2. Capa de Control de Sesión

Abarca los servidores de control de red para la gestión de las llamadas (establecimiento, modificación y liberación); contiene la Función de Control de la Sesión de Llamada (Call Session Control Function, CSCF), que constituye la entidad funcional clave encargada de proveer el registro de los extremos finales y el enrutamiento de los mensajes de señalización SIP hacia el servidor de aplicación apropiado, además realiza funciones de soporte como aprovisionamiento, tarificación, operación y mantenimiento. La CSCF interopera con la capa de transporte y de acceso para garantizar la calidad de servicio a través de todos los servicios.

Esta capa contiene también la base de datos local de suscriptores (Home Subscriber Server, HSS), la cual mantiene un perfil único de servicio por cada usuario final. Dicho perfil almacena toda la información de servicio y preferencias de un usuario en una localización central, lo cual incluye la información actual de registro del usuario final (como por ejemplo la dirección IP), información de itinerancia, servicios de telefonía (como por ejemplo información de reenvío de llamada), información del servicio de mensajería instantánea (por ejemplo lista de contactos), opciones del correo de voz (por ejemplo saludos), etc. Como esta información se encuentra centralizada, las aplicaciones pueden compartirla para crear directorios personales unificados, múltiples tipos de usuarios con la información de presencia, y servicios combinados. También, este orden centralizado simplifica enormemente la administración de los datos de usuario.

Por otro lado, la capa de control de sesión también incluye la Función de Recursos Multimedia (Media Resource Function, MRF), la cual se encarga de los recursos multimedia, y la Función de Control de la Pasarela de Medios (Media Gateway Control Function, MGCF), la cual se encarga de hacer interoperar la señalización SIP con la señalización utilizada por la pasarela de medios (por ejemplo H.248).

- **CSCF (Call State Control Function)**

La entidad funcional clave de la capa de control es el nodo CSCF, este es análogo a un servidor SIP en la arquitectura IETF.

Su función es procesar los mensajes de señalización para controlar la sesión multimedia de los usuarios. El núcleo de la red de conmutación de paquetes existente se utiliza para proporcionar el camino que siguen los datos de sesión multimedia, y los CSCFs se utilizan para establecer las sesiones y para controlar otras características de la llamada.

Los CSCFs realizan una serie de funciones, entre las cuales se tienen: la función de control de la sesión multimedia, la función de traducción de la dirección, y el manejo del perfil del suscriptor.



El CSCF puede desempeñar tres papeles: el papel de Proxy-CSCF (P-CSCF), el papel de Interrogating-CSCF (I-CSCF) y el papel de Serving-CSCF (S-CSCF). El P-CSCF es el primer punto de contacto de un móvil en la red IMS. El S-CSCF es responsable de gestionar la sesión del móvil. La función del I-CSCF es determinar el S-CSCF basado en carga o capacidad de la red. Estos tres papeles pueden soportar capacidades de cortafuegos. Se presenta a continuación cada rol.

▪ Proxy CSCF (P-CSCF)

El Proxy Controlador de la Función de Sesión de Llamada (P-CSCF) es el punto de entrada al subsistema IMS (ya estén en su red origen o visitada) y recibe directamente la señalización IMS desde el terminal, vía GPRS, en una red 2.5G. Implementa las funciones de protección de señalización, control de recursos del subsistema de transporte y encaminamiento de mensajes SIP. Además, es el punto donde se ejerce la política de calidad de servicio (control de recursos y calidad de las sesiones en coordinación con la red de acceso), proporciona el control local para los servicios de emergencia, y también realiza asistencia telefónica de los planes locales de numeración bajo dirección del Serving-CSCF.

En itinerancia, es el nodo en la red visitada que se encarga de enrutar la señalización de registro y sesión desde los terminales que se encuentran en situación de itinerancia hasta la red IMS nativa. Además, ejecuta las funciones comunes a los demás CSCF: el procesado y enrutado de señalización, la consulta del perfil de usuario en el HSS y colecta información para la tarificación.

El Proxy-CSCF es análogo al Servidor Proxy en la arquitectura SIP.

▪ Serving CSCF (S-CSCF)

A cada usuario registrado en IMS se le asigna un S-CSCF, el cual gestiona las sesiones SIP coordinando con los otros los elementos de red el control de las sesiones, y se encarga de enrutar las sesiones destinadas o iniciadas por el usuario. También realiza el registro y autenticación del abonado IMS, proporciona la seguridad de la sesión y la provisión de los servicios IMS. Por esto último, el S-CSCF puede transferir la señalización de un registro o sesión hacia un servidor de aplicaciones SIP, o transferirla hacia una pasarela OSA o CAMEL, que traduce SIP en Parte de la Aplicación CAMEL (CAMEL Application Part, CAP). Además, aplica las políticas o normas del operador de red y genera los registros de tarificación.

Puede haber varios S-CSCFs en la red con varias funcionalidades, estos se pueden añadir según las necesidades, basándose en las capacidades de los nodos o en los requisitos de capacidad de la red. El S-CSCF en la red origen es el responsable de todo el control de la sesión, pero podría remitir una petición específica a un P-CSCF en la red visitada, basándose en los requisitos de la esta. El S-CSCF se puede elegir de diferentes maneras, ya sea basándose en los servicios solicitados o en las capacidades del móvil.

Una ventaja importante de esta arquitectura es que la red origen proporciona las características de servicio. Esto significa que el móvil no está restringido a las capacidades de la red visitada como se ve en la red inalámbrica actual (si un MSC no apoya una característica a la que el usuario se ha suscrito, este no podrá utilizar esa característica). Esta capacidad de permitir que el usuario pueda conseguir siempre el acceso a sus características suscritas se conoce como el Entorno de Origen Virtual (Virtual Home Environment, VHE).



▪ Interrogating CSCF (I-CSCF)

Es un nodo intermedio que da soporte a la operación IMS. El I-CSCF ayuda a otros nodos a determinar el siguiente salto de los mensajes SIP, y a establecer un camino para la señalización encaminando las peticiones o respuestas al S-CSCF. Durante el registro, el I-CSCF obtiene del HSS la dirección del S-CSCF, basándose en parámetros como capacidad y carga, y de esta forma el P-CSCF se ayuda del I-CSCF para determinar el S-CSCF que ha de servir a cada usuario.

En situaciones de itinerancia y en sesiones al interior de una red, el I-CSCF es el punto de entrada, conocido por la red IMS externa e indica el siguiente salto a realizar para la señalización. De manera opcional, el I-CSCF efectúa funciones de ocultación de la configuración y la topología de la red IMS ante redes externas, de forma que los elementos ajenos a IMS no puedan averiguar cómo se gestiona la señalización internamente (por ejemplo, el número, el nombre y la capacidad de los CSCF). Esta última función se denomina Pasarela de Interoperación para el Ocultamiento de la Topología (Topology Hiding Inter-network Gateway, THIG).

El I-CSCF puede realizar también algunas formas de tarificación. Si el I-CSCF es la entrada en la red origen, debe soportar funciones de cortafuegos.

Este es un nodo opcional en la arquitectura de IMS. El sistema puede ser configurado de modo que el P-CSCF entre en contacto con el S-CSCF directamente.

• HSS (Home Subscriber Server)

El Servidor Local de Suscriptor (Home Subscriber Server, HSS), es una evolución introducida en el Registro Local de Ubicación (Home Location Register, HLR) para adaptarlo a los servicios IP, y hereda las funciones de: almacenar y gestionar el perfil del servicio IMS del abonado, almacenar las claves de seguridad y generar vectores de autenticación, almacenar el nodo S-CSCF con el que el abonado se registró, registrar el estado de los abonados y esto se utiliza a su vez para el tratamiento de la información de presencia de los mismos.

La arquitectura de servicios asigna un perfil de servicio en la HSS a cada usuario IMS. Este perfil es un conjunto de criterios de filtrado, una información estática que corresponde a la suscripción de un usuario a un servicio del dominio IMS. El HSS se comunica con el I-CSCF y el S-CSCF para proporcionar la información sobre la localización del cliente y su información de suscripción, a través de la interfaz Cx (protocolo DIAMETER), un protocolo que no es de la IETF pero está basado en IP.

En la Figura A2 se puede ver lo que ocurre cuando existen varios HSS dentro de la arquitectura IMS, y en este caso, es la Función de Ubicación del Suscriptor (Subscription Locator Function, SLF), otra base de datos de esta arquitectura, la que a través de una Respuesta de Información de Localización (Location Information Answer, LIA) especifica cual HSS maneja la información de un determinado usuario, previa Petición de la Información de Ubicación (Location Information Request, LIR).

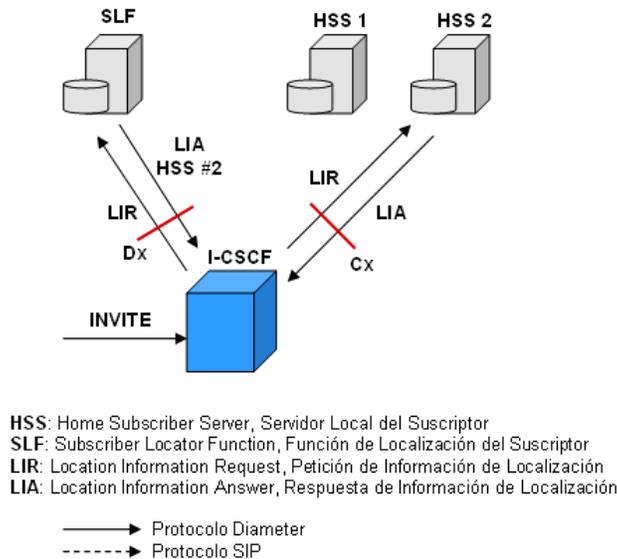


Figura A2: Operación entre el HSS y el I-CSCF

- **MRF (Media Resource Function)**

La MRF permite multiconferencias mezclando los flujos multimedia de varios participantes. Está dividida en el Controlador de la Función de Recursos Multimedia (Media Resource Function Controller, MRFC) y el Procesador de la Función de Recursos Multimedia (Media Resource Function Processor, MRFP), los cuales son similares en función a la Función de Control de la Pasarela de Medios (Media Gateway Control Function, MGCF) y a la Pasarela de Medios (Media Gateway, MGW). El MRFC controla los flujos y recursos multimedia del MRFP, a través de MEGACO, basándose en información suministrada por el S-CSCF y el servidor de aplicaciones (si este existe). Más adelante se tratará más a fondo esta entidad.

En un entorno donde todas las sesiones fueran entre dispositivos de usuario con capacidad IP, no se necesitaría otra cosa que los CSCFs y el HSS. Pero habrá un periodo de transición muy largo para eliminar totalmente las redes antiguas PSTN y móviles. Por esto se necesitan dentro del subsistema multimedia IP los siguientes nodos:

- Función de Control de la Pasarela de Medios (Media Gateway Control Function, MGCF)
- Pasarela de Medios Multimedia IP (IP Multi-Media Gateway, IM-MGW)
- Función de control de la Pasarela de Fuga (Breakout Gateway Control Function, BGCF)
- Pasarela de Señalización Troncal (Trunking Signalling Gateway, T-SGW)

- **MGCF (Media Gateway Control Function)**

La MGCF hace parte de la arquitectura de interoperación de IMS con las redes de circuitos, por ello maneja la conexión entre la portadora PSTN y el flujo IP. La MGCF controla una o más IM-MGWs, permitiendo más escalabilidad en la red. Por simplicidad, la MGCF se podría colocar integrada con la IM-MGW.



La MGCF tiene como funciones: convertir mensajes SIP/SDP en mensajes de MEGACO o de la Parte de Usuario ISDN (ISDN User Part, ISUP), recibir un mensaje SIP del CSCF y determinar qué conexión realizar con la MGW, crear el mensaje apropiado de ISUP y enviarlo, vía IP, a la SGW.

- **IM-MGW (IP Multi-Media Gateway)**

La MGCF es quien controla la interconexión de redes, pero la Pasarela de Medios es la que se encarga del procesamiento de la información multimedia entre los usuarios finales. Su función principal es convertir medios de un formato a otro. En UMTS esto sucede sobre todo entre la Modulación por Impulsos Codificados (PCM) en la PSTN y un formato de codificación de voz basado en IP.

- **BGCF (Breakout Gateway Control Function)**

La BGCF elige la red en la que debe ocurrir el acceso a la PSTN o a la red de conmutación de circuitos. Si la BGCF determina que ésta salida ocurre en la misma red, en la que ella misma se encuentra, entonces elegirá la MGCF encargada de la interacción con la red PSTN. Si la salida ocurre en otra red, la BGCF encaminará la señalización a la BGCF de la red seleccionada.

- **T-SGW (Signaling Gateway)**

Una característica esencial de IMS es que la mayoría de la comunicación entre los componentes está basada en IP. Solamente hay dos interfaces que no están basadas en IP y ambas se utilizan para interactuar con una red tradicional (la PSTN o una red móvil anterior). Estas dos interfaces son el camino portador y el de señalización hacia la red antigua.

La PSTN actualmente sólo entiende SS7 y el problema es que SS7 tiene limitaciones y no es tan flexible como IP. Para evitar la necesidad de que la MGCF soporte SS7 se creó la Pasarela de Señalización Troncal. Su función es convertir SS7 a IP, y esto lo hace convirtiendo las capas más bajas de SS7 en IP, ya que los protocolos de la capa de aplicación no se verán afectados.

- **IMS Application Level Gateway**

La IMS-ALG proporciona a la pila de protocolos SIP/SDP la funcionalidad de aplicación necesaria para que interactúen aplicaciones IPv4 e IPv6. Cuando la IMS-ALG recibe un mensaje SIP de los CSCFs o de una red externa SIP IPv4, cambia los parámetros apropiados de SIP/SDP, traduciendo las direcciones IPv6 a IPv4 y viceversa.

A1.1.3.3. Capa de Aplicaciones

Contiene los servidores de aplicaciones y de contenido, que son los encargados de proveer la lógica de servicio del usuario final. Tanto IMS como la señalización SIP son lo suficientemente flexibles como para soportar una variedad de servidores de aplicaciones, tanto telefónicas como no telefónicas. Por ejemplo, los estándares de SIP se desarrollaron para servicios de telefonía y para servicios Multimedia IP.

IMS define tres formas diferentes de ofrecer servicios. Se enuncian a continuación:



- **Servicios Nativos SIP:** En los últimos años, se han desarrollado una amplia variedad de tecnologías por parte de varias organizaciones para desarrollar servicios SIP, entre las cuales de pueden incluir: SIP Servlets, CPL, SIP CGI y JAIN. Uno o más servidores de aplicaciones SIP se pueden utilizar para desplegar servicios usando estas tecnologías.
- **Servicios tradicionales:** Mientras que se requieren nuevos e innovadores servicios, no se puede ignorar los servicios de la telefonía tradicional.
- **Servicios de terceras partes:** UMTS ha definido la arquitectura de Arquitectura de Servicios Abiertos para permitir que terceros ofrezcan servicios a través de la red UMTS. OSA ofrece una API segura para que puedan acceder a la red UMTS. Por lo tanto, los suscriptores no están restringidos a los servicios ofrecidos por los operadores.

El S-CSCF interactúa con diversas plataformas de servicios a través de la Interfaz de Control de Servicios Multimedia IP (IP multimedia Interface Service Control, IM-ISC) que está basada en SIP y sus extensiones. Sin embargo, los entornos de OSA y CAMEL no soportan la interfaz ISC, por lo que es el servidor de Capacidad de Servicio de OSA (Service Capability Server, SCS) el que realiza la mediación entre la ISC y el API de OSA. La Función de Conmutación de Servicios - Multimedia IP (IP Multimedia - Services Switching Function, IM-SSF) realiza la mediación entre la ISC y el Protocolo de Aplicaciones CAMEL (CAMEL Application Protocol, CAP).

Los tres anteriores se pueden desglosar con más detalle en las siguientes entidades:

- **Servidor de aplicaciones de telefonía**

La arquitectura IMS soporta servidores de aplicaciones múltiples para servicios de telefonía. El Servidor de Aplicaciones de Telefonía (Telephony Application Server, TAS) es un agente de usuario SIP “back-to-back” que mantiene el estado de la llamada. El servidor TAS contiene la lógica de servicio que provee los servicios básicos de procesamiento de llamada, incluyendo análisis de dígitos, enrutamiento, configuración de llamada, llamada en espera, reenvío de llamada, conferencia, etc. Además, este servidor provee la lógica de servicio para la invocación de los servidores de medios cuando se trata de soportar un progreso apropiado de los tonos y anuncios durante la llamada. Si las llamadas se originan o terminan en la PSTN, el servidor TAS provee la señalización SIP a la MGCF, para dar instrucciones a las pasarelas de medios de convertir el flujo de bits que contienen la voz TDM/PSTN al flujo RTP/IP y dirigirlo a la dirección IP del teléfono correspondiente.

- **Función de conmutación de servicios - Multimedia IP (IM-SSF)**

Esta provee la interoperación de los mensajes SIP con los mensajes correspondientes de: CAMEL, ANSI-41, Protocolo de Aplicación de Red Inteligente (Intelligent Network Application Protocol, INAP) o la Parte de Aplicación de las Capacidades de Transacción (Transaction Capabilities Application Part, TCAP). Esta interoperación permite a los teléfonos IP, soportados por IMS, acceder a servicios tales como los servicios requeridos por nombre, servicios 800, servicios de LNP, etc.



- **Servidores de aplicaciones de telefonía suplementarias**

La capa del servidor de aplicaciones puede contener también servidores independientes o autónomos para proveer servicios de telefonía suplementarios al inicio, final o en medio de la llamada. Incluyen estos servicios: hacer clic o presionar para marcar, clic para transferir, clic para conferencia, servicios de correo de voz, servicios IVR, servicios VoIP en Redes Privadas Virtuales (Virtual Private Network, VPN), servicios de facturación prepagada, y servicios de bloqueo de llamadas entrantes o salientes.

- **Servidores de aplicaciones no telefónicas**

La capa del servidor de aplicaciones puede también contener servidores de aplicaciones basados en SIP para operar fuera del modelo de telefonía. Estos servidores de aplicaciones pueden interoperar con clientes finales para proveer servicios tales como: mensajería instantánea (IM), PTT, o servicios basados en presencia.

Por medio de la implementación de estos servicios no telefónicos, basados en SIP, y en una arquitectura IMS común, es posible interoperar los servicios telefónicos con aquellos que no lo son, para de esta forma crear servicios de comunicaciones nuevos y combinados. Un ejemplo de combinación de servicios, es el caso de una lista de contactos, en la cual se debe hacer clic para conectarse con algún amigo, se muestra la información de presencia y de disponibilidad del usuario final, y además provee una interfaz a través de la cual se puede escoger y hacer clic en servicios de comunicaciones múltiples. Otro ejemplo, es la utilización de una cuenta única de servicios prepagados para servicios de telefonía y de Video Bajo Demanda (Video On Demand, VoD).

- **Pasarela de Acceso a Servicios Abiertos (OSA-GW)**

La arquitectura IMS permite a los proveedores de servicios la flexibilidad de adicionar servicios en sus redes VoIP, mediante la interacción con las aplicaciones heredadas o a través de la integración con servidores de aplicaciones basados en SIP, que ellos mismos pueden comprar o desarrollar. Sin embargo, los desarrolladores de aplicaciones de las empresas trabajan frecuentemente bajo un contexto de Tecnologías de Información (Information Technology, IT) y no están familiarizados con la variedad de protocolos complejos de señalización telefónica (tales como SS7, ANSI41, CAMEL, SIP, ISND, etc.). La interoperación entre SIP y el API de Parlay para las redes de telefonía, definida en conjunto con el 3GPP y la ETSI, se provee a través de la Pasarela de la Arquitectura de Servicios Abiertos (Open Services Architecture - Gateway, OSA-GW) que hace parte de la capa del servidor de aplicaciones de la arquitectura IMS.

Como protocolo para conectar la gran mayoría de los nodos de IMS entre sí y con el resto de los elementos que componen la red móvil, se ha determinado en el 3GPP la utilización del Protocolo de Inicio de Sesión SIP, el cual se tratará más adelante. Por el momento, se puede agregar que IMS utiliza también a SIP para establecer llamadas de multimedia y voz en Internet y se lo eligió por su simplicidad, extensibilidad y su amplia disponibilidad.



A1.1.4. Estado actual y evolución [2][7][11][12]

IMS está definido en las versiones del sistema 3GPP del Release 5 y 6. En el Release 5, ya finalizado, se definen perfectamente la arquitectura y las entidades funcionales, aunque en las especificaciones de esta Release solamente se describe de forma detallada el funcionamiento básico de IMS, dejándole al Release 6 bastantes funcionalidades de este último. Así pues el Release 5 y Release 6 están muy relacionados.

Dentro del marco del Release 5 se plantea la utilización de SIP para el control de sesiones IP multimedia en cuanto a la señalización entre el Equipo de Usuario (User Equipment, UE) e IMS, y también entre entidades funcionales IMS. Igualmente, la especificación del subsistema IMS utiliza SIP para completar llamadas de voz y llamadas multimedia con otras redes (redes telefónicas e Internet).

En 3GPP Release 6 (IMS Fase 2) se completa y amplía el concepto relativo a los servicios IP multimedia. IMS Release 6 incorpora, respecto a su versión de Release 5, las siguientes funcionalidades o mejoras:

- La descripción completa de la operación para el interfuncionamiento con las redes de circuitos (conversión de SIP a ISUP).
- La descripción completa de la operación para el interfuncionamiento con las redes multimedia IP (interoperación del "3GPP SIP" con "IETF SIP"), incluyendo los aspectos para la interconexión y escenarios de migración entre implementaciones IPv4 y IPv6.
- La descripción completa del servicio de conferencias multiparticipante.
- La especificación del servicio de mensajería IMS.
- La especificación del servicio de presencia IMS.
- La optimización del subsistema de transporte de paquetes y red de acceso radio para el soporte eficiente y flexible de multimedia IP.
- La introducción del concepto de independencia de la red de acceso: para permitir el acceso a IMS desde cualquier red de conectividad IP.
- Se mejora y completa los Servicios de Llamada de Voz Grupales (Voice Group Call Services, VGCS) con mecanismos de cifrado y gestión de claves de cifrado.
- Mejoras en políticas de QoS. Si en el Release 5 apareció la interfaz Go como punto de acceso desde el Nodo de Soporte de la Pasarela GPRS (Gateway GPRS Support Node, GGSN) a IMS, para garantizar una calidad de servicio a nivel de portadora (contexto PDP en acceso GPRS), esta interfaz es superada en el Release 6 con una granularidad a nivel de flujo IP.
- Mejoras en los servicios de localización, mejoras en la seguridad, reconocimiento de voz, servicios activados por voz, extensión de AMR para Banda Ancha (Wide Band - Adaptive Multi Rate, WB-AMR) para alta calidad de audio, etc.

El Release 7 se encuentra en desarrollo, y en este se trata la habilitación de velocidades mayores y mejoras en la capacidad, así como el soporte a servicios mejorados de tiempo real tales como la VoIP, los juegos interactivos, y el PTT en celulares. Las mejoras incluyen características como Múltiples Entradas Múltiples Salidas (Multiple Input Multiple Output, MIMO), alcanzar tasas teóricas mayores de las actuales (14Mbps), además de la mejora en el rendimiento promedio de la celda. Otras de las características del Release 7 incluyen: mejoras en la Red de Acceso Radio (Radio Access Network, RAN) tales como la conectividad continua, mejoras en la configuración de la latencia, mejoras en cuanto al núcleo de red y a IMS en lo que respecta a la telefonía multimedia, y la Convergencia de las Políticas y la Facturación (Policy and Charging Convergence, PCC).



A1.1.5. Principales tecnologías y protocolos [4][7][13]

A1.1.5.1. SIP

La señalización en IMS se efectúa mediante el protocolo SIP, que la Fuerza de Trabajo en Ingeniería de Internet (Internet Engineering Task Force, IETF) diseñó para la gestión de sesiones multimedia en Internet. SIP aporta las funciones para el registro, establecimiento, liberación y mantenimiento de las sesiones IMS, lo que incluye funciones de enrutado de sesiones e identificación de usuarios y nodos, y también habilita todo tipo de servicios suplementarios.

Más información de SIP se puede encontrar en la RFC3261 [14].

A1.1.5.2. IPv6

IMS se ha definido desde el inicio como una red y un servicio que fundamenta completamente su transporte de red sobre IPv6. El subsistema GPRS 3G que proporciona acceso a dicha red IPv6 modifica sus especificaciones para soportar el transporte de datagramas IPv6 desde el terminal de usuario hasta IMS, así como la configuración y la asignación de direcciones de red. Por otro lado, el terminal IMS ha de soportar la pila de protocolos de IPv6, y posiblemente de IPv4, por ello se han realizado trabajos en la generación de mecanismos para el interfuncionamiento IPv4/IPv6.

Por otro lado, además de las ventajas inherentes a IPv6 como son la QoS y la seguridad integradas, así como la autoconfiguración y el mayor espacio de direccionamiento, el tráfico del plano de usuario se transfiere directamente entre terminales siguiendo el esquema “peer-to-peer” o “de igual a igual”. Por tanto, IPv6 simplifica el transporte de este modelo de tráfico en las redes IPv4 privadas, como son la mayor parte de las redes GPRS existentes en el mundo.

Más información de IPv6 se puede encontrar en la RFC791 [15].

A1.1.5.3. RTP y RTCP

La IETF desarrolló los protocolos RTP/RTCP de envío de flujos multimedia de audio y vídeo basados en el protocolo UDP. Estos protocolos se utilizan en IMS para el transporte y secuenciación de flujos IP multimedia del plano de usuario, RTP para la transmisión de la voz y vídeo sobre IP, y el protocolo RTCP para el control de la transmisión de los paquetes RTP.

Más información de RTP y RTCP se puede encontrar en las RFC1889 [16] y RFC3605 [17], respectivamente.

A1.1.5.4. COPS

El protocolo de Servicio de Políticas Común y Abierto (Common Open Policy Service, COPS) se utiliza para el control de los recursos de GPRS, de señalización de sesión y señalización de transporte, mediante el uso de políticas o normas de asignación de los mismos en función de los objetivos de calidad establecidos.

Más información de COPS se puede encontrar en la RFC2748 [18].



A1.1.5.5. DIAMETER

Es un protocolo que proporciona un marco de Autenticación, Autorización y Tasación (Authentication, Authorization and Accounting, AAA) para aplicaciones de acceso a redes, movilidad en redes IP e itinerancia. Diameter es compatible con Radius. Se compone de una especificación base más una serie de especificaciones dependientes de la aplicación. Principalmente se emplea como heredero de la Parte de Aplicación Móvil (Mobile Application Part, MAP) para el diálogo con el nodo HSS de IMS.

A1.1.5.6. RSVP Y DIFFSERV

Los protocolos de Reservación de Recursos (Resource Reservation Protocol, RSVP) y DiffServ los utiliza IMS para asegurar la QoS extremo a extremo, especialmente cuando la conectividad IP requerida se extiende más allá de la red móvil GPRS.

Más información de RSVP y DiffServ se puede encontrar en las RFC2205 [19] y RFC2475 [20], respectivamente.

A1.1.5.7. MEGACO

Este protocolo se utiliza en arquitecturas SIP y H.323 para el control de las Pasarelas de Medios que permiten el interfuncionamiento con las redes TDM.

Más información de MEGACO o H.248 se puede encontrar en la RFC 3525 [21].

A1.1.5.8. SIGTRAN

El grupo de trabajo del IETF denominado SIGTRAN, ha definido un conjunto de protocolos y adaptadores de protocolos estándares que permiten el transporte de información del Sistema de Señalización número 7 (Signaling System 7, SS7) sobre redes basadas en el protocolo IP.

El transporte de señalización SS7 sobre redes IP es la clave para la integración de los nuevos componentes que aparecen en IMS (MGW, MGC y Sistemas VoIP) y en las redes telefónicas actuales (RTC y RDSI).

En SIGTRAN se define un protocolo de transporte nuevo, orientado a conexión, y denominado SCTP como alternativa al tradicional TCP. El diseño de SCTP incluye procedimientos para evitar comportamientos del protocolo que provoquen congestión y resistencia ante ataques externos basados en falsificación de direcciones y avalanchas de peticiones de servicio.

A1.2. PROTOCOLO DE INICIO DE SESIÓN SIP

A1.2.1. Definición y origen de SIP [2][7]

SIP son las siglas en inglés del Protocolo para el Inicio de Sesión el cual ha existido desde 1995, y fue ratificado en 1999 bajo el RFC 2543 por la Fuerza de Trabajo de Ingeniería para Internet (Internet Engineering Task Force, IETF). El protocolo actual está definido en el RFC 3261.



Este protocolo fue desarrollado inicialmente por el grupo de trabajo MMUSIC (Multiparty Multimedia Session Control, Control de Sesión Multimedia Multipartita) de la IETF, para proporcionar, esencialmente, 'Presencia' y 'Movilidad' dentro de una red IP. El grupo de trabajo de SIP, dentro de la IETF es hoy unos de los más activos, entre los cuales existen 6 dedicados a mensajería instantánea y presencia. Aparte del grupo principal, existe uno dedicado exclusivamente a mensajería [22] (SIMPLE) y otro de apoyo, denominado SIPPING. El grupo IPTEL de encarga CPL y MMUSIC del desarrollo del Protocolo de Descripción de la Sesión (Session Describing Protocol, SDP). Como otros estándares, SIP está contemplado dentro de la base de datos RFC, siendo un protocolo libre y abierto a nuevas modificaciones y no ligado a ninguna empresa ni entidad privada.

Este es un protocolo de señalización del nivel de aplicación diseñado para el establecimiento, modificación, mantenimiento y terminación de sesiones interactivas sobre redes IP para diferentes tipos de aplicaciones. El inicio de sesión, cambio o término de la misma, son independientes del tipo de medio o aplicación que se use en la llamada; dichas sesiones incluyen elementos multimedia como voz, datos, video y otros formatos, y pueden ser desde una llamada de teléfono IP, una conversación breve a través de un servicio de mensajes instantáneos donde se pueden compartir páginas Web y documentos, Chat o juegos interactivos, hasta una videoconferencia.

SIP aporta las funciones para el registro, enrutado de sesiones, identificación de usuarios y nodos, localización y disponibilidad de usuarios, negociación de las capacidades de comunicación y también habilita todo tipo de servicios suplementarios.

La especificación del subsistema IMS utiliza SIP para completar llamadas de voz y llamadas multimedia con otras redes (redes telefónicas e Internet). La elección de SIP por el Foro 3GPP se debe a lo simple y lo relativamente fácil de extender de dicho protocolo. Es por eso que la IETF ha ido añadiendo al protocolo básico extensiones y cabeceras privadas para adaptar su uso a las necesidades del entorno móvil, y a las particularidades de una red IMS. En este sentido, se habla del perfil 3GPP del protocolo SIP, una variante personalizada para la red 3G IMS.

A1.2.2. Funciones [4][23]

- Resolución de Direcciones
 - Hecha por agentes o servidores, no siempre implica traslación a IP, generalmente se hace una sola vez.
 - Determina la localización del punto extremo objetivo; SIP soporta la resolución de dirección, mapeo de nombre, y redirección de llamada.
- Funciones de Sesión
 - Establecimiento: SIP establece una sesión entre los puntos extremos si la llamada puede completarse.
 - Determinación de la disponibilidad del punto extremo: si una llamada no puede completarse porque la otra parte no esta disponible, SIP determina si la otra parte se encuentra en ese momento utilizando el teléfono o si no respondió la llamada, después de haberle timbrado según la cantidad de timbres definidos. SIP luego retornará un mensaje indicando el porque la otra parte no estaba disponible.
 - Negociación de medios: determinación de las capacidades multimedia del punto extremo objetivo, a través de SDP, SIP determina el nivel mínimo de servicios comunes entre los dos puntos. Las sesiones solo se establecen utilizando las capacidades multimedia que pueden ser soportadas por ambos extremos.



- Modificación: SIP soporta cambios durante la llamada, tales como adición a la conferencia de otro punto extremo, o el cambio de una característica multimedia o de un dispositivo codificador-decodificador o codec.
- Transferencia y terminación: SIP soporta la transferencia de llamadas de un punto extremo a otro. Durante una transferencia de llamada, SIP simplemente establece una sesión entre el transferido y el nuevo punto extremo (el cual especifica la parte que transfiere) y termina la sesión entre el transferido y la parte que transfiere. Al final de la llamada SIP termina la sesión entre todas las partes.
- Cancelación
- Señalización en llamada
- Control de llamada
- Configuración de QoS

Los tipos de Sesiones que puede establecer SIP son los siguientes:

- ✓ Entre dos partes (llamadas telefónicas)
 - ✓ Multiparte (todos pueden participar)
 - ✓ Multicast (uno envía, muchos reciben)
- Funciones no relacionadas con la sesión
- Movilidad: SIP permite implementar dos tipos de movilidad diferentes: 1. La movilidad personal, que permite que el usuario pueda ser alcanzado en un dispositivo cualquiera, mediante los servicios de proxy y redirección y 2. La movilidad intrínseca provista por la ubicuidad del protocolo IP.
 - Transporte de Mensajes
 - Suscripción a eventos
 - Autenticación y contabilidad
 - Otras funciones (SIP es extensible)

A1.2.3. Pila de protocolos [23][24]

El protocolo SIP no es una solución completa en sí mismo, sino uno de los protocolos necesarios para permitir el intercambio de flujos multimedia a través de la red. Por ello, SIP trabaja en conjunto con otros protocolos. Actúa como transportador del protocolo SDP, el cual describe el contenido de los medios de la sesión (puertos IP usados, codec, etc.). Para el transporte de contenido de voz y video sobre IP se utiliza el protocolo RTP, y el protocolo RTCP para el control de la transmisión de paquetes RTP. Esto se muestra en la Figura A3.

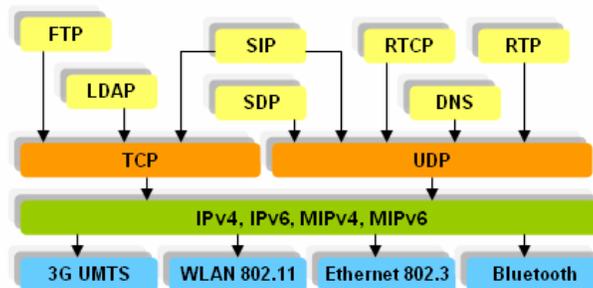


Figura A3: Pila de protocolos que actúan junto con SIP



A1.2.4. Dirección [25]

Una dirección SIP hace referencia a una URI que indica un usuario o servicio. Una URI es una cadena de caracteres compacta que se utiliza para identificar un recurso ya sea abstracto o físico. Las URIs proporcionan un significado simple y extensible para la identificación de recursos, y un direccionamiento global.

La sintaxis de una URI consiste en una secuencia de caracteres que se encuentran separados por caracteres reservados tales como: ";", ":", "@", "?", y "/".

```
URI = scheme ":" hier-part [ "?" query ] [ "#" fragment ]
```

En el caso de una dirección SIP, `scheme` puede ser `sip` o `sips`, y la cadena de caracteres siguiente corresponde al cliente SIP el cual se identifica por una dirección que tiene un formato similar al formato del correo electrónico.

```
[ userinfo "@" ] host [ ":" port ]
```

La parte `userinfo` puede ser un nombre del usuario o un número de teléfono. La parte `host` puede ser un nombre de dominio, un nombre de un servidor, o una dirección IP. La parte `port` es un dígito que indica el puerto a través del cual se establecerá la comunicación SIP. Debe incluir `host`, puede incluir puerto y nombre de usuario.

Una dirección SIP puede ser una combinación de Nombres de Dominio Totalmente Cualificados (Fully-Qualified Domain Name, FQDN), o números E.164, o ambos.

A continuación se muestran ejemplos típicos de direcciones SIP:

```
sip: usernickname@unicauca.edu.co: 5060  
sip: usernickname @172.16.255.96:5061  
sip: 556254@127.0.0.1:5062  
sip: +33-143592334@gw.tn.com; user = phone  
sip: "John Doe" <jdoe@cisco.com> (dirección FQDN)  
sip:14085551234@gateway.com; user=phone (dirección E.164)
```

A1.2.5. Arquitectura

SIP es un protocolo basado en texto y en la arquitectura cliente-servidor, por lo que todos los procesos se ejecutan mediante el intercambio de mensajes en forma de peticiones y respuestas entre una entidad cliente y otra que actúa como servidor.

A1.2.5.1. Entidades

A1.2.5.1.1. Principales componentes [26][27][28]



- **Agente de Usuario (User Agent)**

Corresponde a una aplicación software ubicada en la estación terminal. Puede operar como cliente, realizando peticiones SIP, y en este caso se tratará de un Agente de Usuario Cliente (User Agent Client, UAC), o como servidor, generando respuestas, y tratándose en este caso de un Agente de Usuario Servidor (User Agent Server, UAS). En ambos procesos, el agente de usuario es capaz de soportar una comunicación básica, en la cual administra el estado de la llamada.

Se puede encontrar en los terminales de los usuarios y en otros equipos de control y pasarelas.

- **Servidor Proxy (SIP Proxy)**

Actúa como cliente y servidor, y es el responsable primario del encaminamiento de mensajes entre equipos finales. Se encarga de interpretar, reescribir o traducir un mensaje, en caso de ser necesario, antes de enrutarlo hacia su destino final, y retiene la información por cuestiones de contabilidad o facturación. Los Servidores Proxy SIP pueden usar varios métodos para intentar resolver la dirección destino solicitada, incluyendo búsquedas en el DNS, en bases de datos o relevando la labor hacia el siguiente Servidor Proxy.

- **Servidor de Redirección (Redirect Server)**

Se encarga de aceptar peticiones, y a partir de la dirección del destinatario final obtiene la dirección del siguiente elemento de la cadena. Dicha dirección es devuelta al cliente para que éste sea capaz, por sí mismo, de contactar al destinatario final. Al contrario de un servidor Proxy, el Redirect no inicia sus propios mensajes SIP, sólo responde y al contrario de un agente de usuario, no acepta o termina llamadas.

- **Servidor de Registro (Register Server)**

Acepta peticiones de registro por parte de los UA y almacena dicha información para proporcionarla a otros UA. Un usuario puede estar registrado con múltiples dispositivos y un dispositivo puede tener registrado múltiples usuarios, pero cada usuario es responsable de registrar y mantener el registro en sus diferentes dispositivos.

- **Servidor de Localización (Location Server)**

Suministra la información sobre la posible localización del destinatario de la llamada a los servidores de Redirección o Proxy. Es alimentado por el servidor de Registro o por interfaces de aprovisionamiento de usuarios. Consiste en una base de datos y no utiliza SIP para comunicarse con los otros servidores.

El servidor Location es un sistema concebido para ser alimentado desde diversas fuentes, además del servidor Register, por ejemplo, de la información que pueden aportar sistemas Itinerancia de telefonía móvil, coordenadas GPS o atributos sobre un usuario exportados por un servidor LDAP (hora de recepción de llamadas, idioma, disponibilidad, etc.).

La división entre servidores Proxy, Register, Redirect y Location no es física, sino conceptual, con lo cual dichas funciones pueden coexistir en un único servidor, o estar separadas por motivos de escalabilidad, redundancia o rendimiento. Estos son los elementos clave para ofrecer "Presencia" y



"Movilidad" a un usuario SIP. Al combinarse dichos servidores se hace posible localizar a los usuarios en distintos medios, convirtiendo a SIP en una arquitectura muy flexible e ideal para usuarios móviles ya que estos pueden emplear varios esquemas, simultáneamente, para localizar a otros usuarios, aún cuando se trate de usuarios remotos, para lo cual, los servidores Proxy y de Redirect pueden reenviar la petición de conexión al lugar en donde se encuentra el usuario. Las sesiones pueden incluir a varios participantes.

SIP utiliza direcciones de tipo correo electrónico para identificar a los usuarios, en lugar de los dispositivos que utilizan, por lo cual es un protocolo que no depende del dispositivo. Durante el proceso de registro, un usuario emplea un mensaje SIP para indicar su dirección IP (dirección física) y correo electrónico (dirección lógica) al servidor de Registro que prestará el servicio. A su vez, este servidor informa de la localización del usuario (dirección IP física) a un servidor jerárquicamente superior conocido como Location.

Cuando un usuario o terminal SIP desea establecer una comunicación con otro, envía un mensaje a su servidor Proxy para conocer la dirección física que tiene en ese momento el destinatario de la comunicación. En esta fase, el llamante indicará al servidor Proxy la dirección lógica del usuario llamado, por ejemplo, la dirección de correo electrónico de este último. Con esta información, el servidor Proxy escalará la consulta al servidor Location quien resolverá la localización física del usuario llamado.

A1.2.5.1.2. Servidores SIP: Servidor de Medios [4]

Existen varios tipos de servidores SIP:

- **Servidores de aplicación (Application servers):** permiten el control de la ejecución de la aplicación.
- **Servidores de medios (Media servers):** permitiendo el control de medios.
- **Servidores de llamada (Call servers, CSCF):** responsables del control y enrutamiento de llamadas, así como de la invocación de los servicios internos (ejecutados por el CSCF) y externos (ejecutados por servidores de aplicación).
- **Servidores de mensajería (Messaging servers):** responsables del almacenamiento de mensajes cortos, del envío de los mensajes y de la notificación de nuevos mensajes. Este es un tipo particular de servidor de aplicación.

Funcionalidades del servidor de medios

La funcionalidades del servidor de medios incluyen el control de medios y de los recursos en sí: Anuncios (audio y vídeo), reconocimiento automático de la voz, generación de información de tasación, IVR, grabación de mensajes vocales, conversión de texto a voz, traducción entre diferentes dispositivos codificadores-decodificadores o "codecs", gestión de multipartes y de conferencia, replicación del contenido audio y vídeo RTP, interfaces normalizadas y abiertas (SIP, VoiceXML).

Interfaces del servidor de medios

- Control: Protocolo SIP.
- Multimedia: Protocolo RTP para el transporte de voz y vídeo y RTCP para el control del transporte.



- Servidor de archivo externo: Los archivos multimedia (audio, vídeo) son almacenados en el servidor de medios o en un servidor de archivo externo. El protocolo de Transferencia de Hipertexto (Hypertext Transfer Protocol, HTTP) permite leer y almacenar los ficheros en un servidor de archivo externo.
- ASR y TTS: Las funcionalidades ASR y TTS pueden ser provistas por el servidor de medios o por un servidor externo. El protocolo de control del servidor ASR/TTS es el Protocolo de Control de los Recursos Multimedia (Media Resource Control Protocol, MRCP).
- Tasación: Protocolo DIAMETER para recuperar los Registros Detallados de Llamada (Call Detailed Record, CDR) del servidor de medios.
- Gestión de sesión interactiva: Protocolo de Transmisión Continua en Tiempo Real (Real-Time Streaming Protocol, RTSP).
- Gestión: La gestión de los servidores de medios es hecha por interfaces del Protocolo Simple de Administración de Red (Simple Network Management Protocol, SNMP), Interfaces por Línea de Comandos (Command Line Interface, CLI) y HTTP.

A1.2.6. Peticiones y Respuestas

A1.2.6.1. Peticiones [23][29]

- **INVITE**
Permite el establecimiento de una sesión. Este método lo deben soportar los servidores Proxy, Redirect, agentes de usuario servidores y agentes de usuario clientes.
- **ACK**
Confirma que el cliente recibió la respuesta final a un mensaje INVITE. Este método sólo tiene sentido dentro de un flujo que se inició con un INVITE. Puede contener en el cuerpo del mensaje la descripción final de la sesión que va a utilizar el usuario llamado.
- **BYE**
Permite la liberación de una sesión previamente establecida. Esta petición puede solicitarse tanto por el usuario que inicio la sesión como por el usuario llamado. Este método lo deben soportar los servidores Proxy y debería ser soportado por los servidores Redirect y agentes de usuario que actúen como servidores.
- **OPTIONS**
Se utiliza para solicitar opciones, capacidades y el estado de un UA o de un servidor.
- **CANCEL**
Se utiliza para anular la petición de un INVITE en curso. Esta petición no afecta una petición que se halla procesado y ejecutado completamente, o a una respuesta que ya haya sido enviada de regreso.
- **REGISTER**
Permite el registro de un agente de usuario. De esta forma, el servidor Register conoce la localización del usuario.

Dentro de las extensiones de las peticiones SIP se encuentran los métodos:



- **SUBSCRIBE** : Solicitud de notificación de eventos que ocurran en otro agente de usuario (RFC 3265)
- **UNSUBSCRIBE** : Cancelar notificación de Evento
- **NOTIFY** : Notifica un evento suscrito a través de SUBSCRIBE (RFC 3265)
- **PUBLISH** : Publica un estado
- **MESSAGE** : Transporta un mensaje de datos fuera de una sesión (RFC 3428)
- **REFER** : Transfiere el usuario a una URL (RFC 3515)
- **UPDATE** : Modifica la sesión antes de su aceptación (RFC 3311)
- **INFO** : Permite el intercambio de señalización durante la llamada (RFC 2976)
- **PRACK** : acusa las respuestas provisionales (RFC 3262)

A1.2.6.2. Códigos de respuesta [29]

Tras haber recibido e interpretado una petición SIP, el llamado devuelve una respuesta SIP. Existen seis clases de respuestas:

CLASE	DESCRIPCIÓN
1xx	Información provisional: Requerimiento recibido y en progreso pero no terminado 100 "Trying" 180 "Ringing" 181 "Call Is Being Forwarded" 182 "Queued" 183 "Session Progress"
2xx	Completo: Requerimiento recibido, comprendido y completado satisfactoriamente 200 "OK" 202 "Accepted"
3xx	Redirección: La petición requiere otros tratamientos antes de poder determinar si puede ser realizada y por ello debería redireccionarse 300 "Multiples Choices" 301 "Moved Permanently" 302 "Moved Temporarily" 303 "See Other" 305 "Use Proxy" 380 "Alternative Service"
4xx	Error del cliente: Error en la petición, esta no puede ser interpretada y deberá modificarse antes de intentar volver a enviarla 400 "Bad request" 401 "Unauthorized" 403 "Forbidden" 404 "Not found" 407 "Proxy authentication required" 482 "Loop Detected" 483 "Too Many Hops" 486 "Busy here" 487 "Request terminated"
5xx	Error del servidor: El servidor no puede procesar una petición válida aparentemente 500 "Sever internal error"



501 "Not Implemented"
502 "Bad gateway"
503 "Service Unavailable"
504 "Gateway Timeout"
505 "SIP Version Not Supported"

6xx Falla global: La petición no puede procesarse por ningun servidor

600 "Busy everywhere"
603 "Decline2"
604 "Does Not Exist Anywhere"
606 "Not Acceptable"

Tabla A1: Códigos de respuesta SIP

A1.2.6.3. Formato de las peticiones y respuestas [29]

Los mensajes SIP se componen de tres partes, como se puede ver en la Figura A4:

Línea de Inicio	{	INVITE sip:usuariob@uniedu.co SIP/2.0
Cabeceras	{	Via: SIP/2.0/UDP station1.uniedu.co:5060 Max-Forwards: 70 To: Usuario B<sip:usuariob@uniedu.co> From: Usuario A <sip:usuarioa@uniedu.co> Call-Id: 256321@station1.uniedu.co Cseq: 1 INVITE Contact: usuarioa@172.16.255.193 Content-Type: application/sdp Content-Length: 162
Línea en blanco	{	
Cuerpo del mensaje	{	v = 0 c = IN IP4 172.16.255.193 m = audio 45450 RTP/AVP 0

Figura A4: Formato de una petición SIP

- Línea de inicio

Cada mensaje SIP empieza con esta línea. En esta línea se transporta el tipo de mensaje, que es el método de una petición o el código de respuesta, y la versión del protocolo. Por lo anterior, la línea de inicio puede ser o una Línea de Petición, para las peticiones, o una Línea de Estado, para las respuestas, con las características que se describen a continuación:

- La línea de petición incluye el `request-URI` el cual indica el usuario o el servicio al cual se le dirige la petición.
- La línea de estado mantiene un valor numérico que corresponde al código de estado y un valor textual de la frase asociada.

- Cabeceras

Los campos de las cabeceras SIP transportan los atributos que proporcionan información adicional acerca del mensaje. Son similares a los campos de las cabeceras HTTP semántica y sintácticamente, y por esta razón manejan el siguiente formato:



<nombre>:<valor>

Las cabeceras pueden abarcar líneas múltiples. Algunas cabeceras como *Via*, *Contact*, *Route* y *Record-Route* pueden aparecer varias veces en un mensaje, o pueden tener varios valores separados por comas.

En la siguiente tabla se muestra las cabeceras que puede tener un mensaje SIP:

Cabeceras que aplican tanto para las peticiones como para las respuestas	Cabeceras que definen información acerca del cuerpo del mensaje o acerca del recurso que se identifica en la petición	Cabeceras que le permiten al cliente pasar información que no se puede poner en la línea de estado, acerca de la respuesta	Cabeceras que le permiten al servidor pasar información que no se puede poner en la línea de estado, acerca de la respuesta
Accept	Content-Encoding	Authorization	Allow
Accept-Encoding	Content-Length	Contact	Proxy-Authenticate
Accept-Language	Content-Type	Hide	Retry-After
Call-ID		Max-Forwards	Server
Contact		Organization	Unsupported
Cseq		Priority	Warning
Date		Proxy-Authorization	WWW-Authenticate
Encryption		Proxy-Require	
Expires		Route	
From		Require	
Record-Route		Response-Key	
Timestamp		Subject	
To		User-Agent	
Via			

Tabla A2: Cabeceras de un mensaje SIP

- **Cuerpo del mensaje**

El cuerpo del mensaje se utiliza para describir la sesión que se va a iniciar (por ejemplo en una sesión multimedia, los tipos de codecs de audio y de video, y las tasas de muestreo) o por otra parte se puede utilizar para que contenga texto o datos binarios de cualquier tipo, que tengan relación de alguna manera con la sesión. El cuerpo del mensaje puede aparecer tanto en las peticiones como en las respuestas.

SIP hace una distinción clara entre la información de señalización, la cual se transporta en la línea de inicio y cabeceras, y la información que describe la sesión, la cual se encuentra fuera del alcance de SIP.

Entre algunos tipos de los cuerpos de los mensajes SIP se tienen:

- SDP
- MIME

Otros se encuentran por definir en la IETF y en implementaciones específicas.



SDP (Session Description Protocol) [30]

Es un lenguaje para describir sesiones de audio, video y multimedia (RFC 2327). SIP utiliza SDP para definir el tipo de medios de la comunicación, para ello utiliza un modelo de oferta/respuesta: en la oferta se ofrecen los distintos tipos de protocolos, codecs, carga útil, direcciones y puertos para establecer la comunicación; en la respuesta se define cuales serán utilizados.

Las cabeceras de SDP se transportan utilizando el cuerpo del mensaje SIP, en los mensajes INVITE/200/ACK. Si se realiza la oferta en el INVITE, se debe responder en el 200. Si el INVITE no contiene un oferta SDP, se debe proveer una oferta en la respuesta 200 y ser contestada en el cuerpo del ACK.

SDP tiene campos opcionales y requeridos.

La descripción de la sesión por medio de SDP consiste en un número de líneas de texto, de la forma <tipo>=<valor>, donde <tipo> debe ser exactamente un carácter y <valor> corresponde a un texto estructurado cuyo formato depende del <tipo>.

La regla que sigue SDP es la siguiente:

```
session-description = proto-version      ; v=  
                    origin-field        ; o=  
                    session-name-field  ; s=  
                    information-field    ; i=*  
                    uri-field           ; u=*  
                    email-fields        ; e=*  
                    phone-fields        ; p=*  
                    connection-field    ; c=*  
                    bandwidth-fields    ; b=* (cero o más)  
                    time-fields         ; t= r=* (uno o más, con cero o más 'r=')  
                    key-field           ; k=*  
                    attribute-fields    ; a=* (cero o más)  
                    media-descriptions  ; m= i=* c=* b=* k=* a=* (cero o más)  
  
                    *_ opcionales
```

A continuación, en la Figura A5, se muestra un ejemplo:

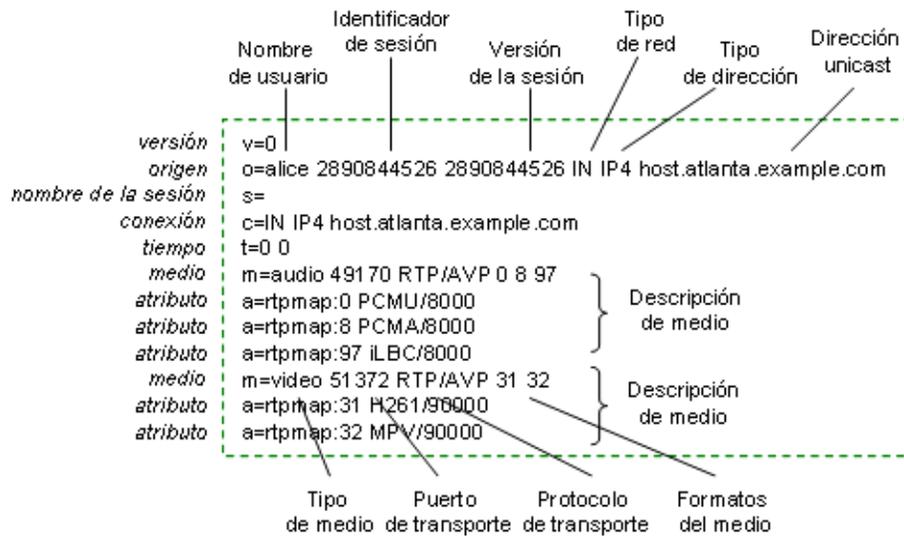


Figura A5: Formato del protocolo SDP

A1.2.7. Flujo de mensajes SIP: ejemplos

A1.2.7.1. Registro

El registro mediante mensajes SIP se puede ver en la Figura A6.



Figura A6: Registro mediante mensajes SIP

A1.2.7.2. Establecimiento de sesión

El establecimiento de una sesión mediante mensajes SIP se puede ver en la Figura A7.



Figura A7: Establecimiento de sesión mediante mensajes SIP

A1.2.7.3. Establecimiento de sesión involucrando un servidor Proxy SIP

El establecimiento de una sesión involucrando un servidor Proxy SIP se puede ver en la Figura A8.

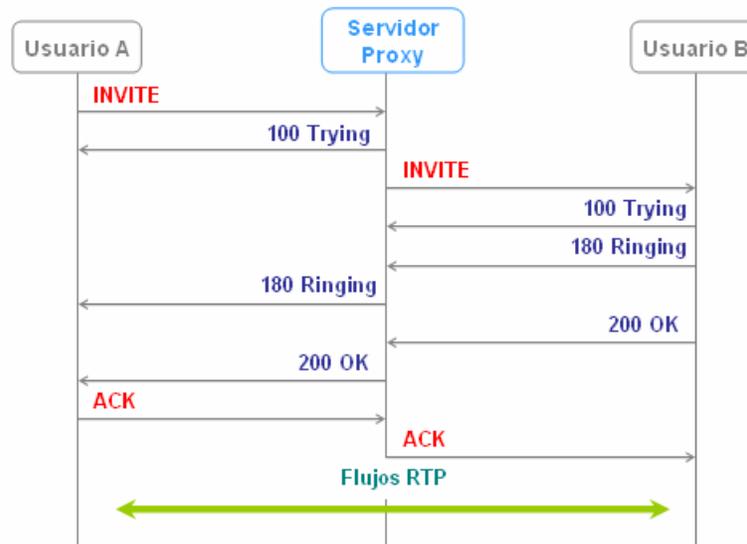


Figura A8: Establecimiento de sesión involucrando un servidor Proxy SIP

A1.2.7.4. Mensajes instantáneos

El envío de mensajes instantáneos mediante mensajes SIP se puede ver en la Figura A9.

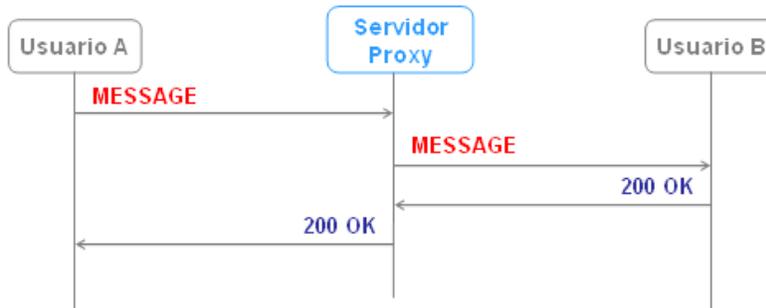


Figura A9: Envío de mensajes instantáneos mediante mensajes SIP

A1.2.7.5. Terminación y cancelación de sesión

La terminación y la cancelación de sesión mediante mensajes SIP se pueden ver en las Figuras A10 y A11 respectivamente.

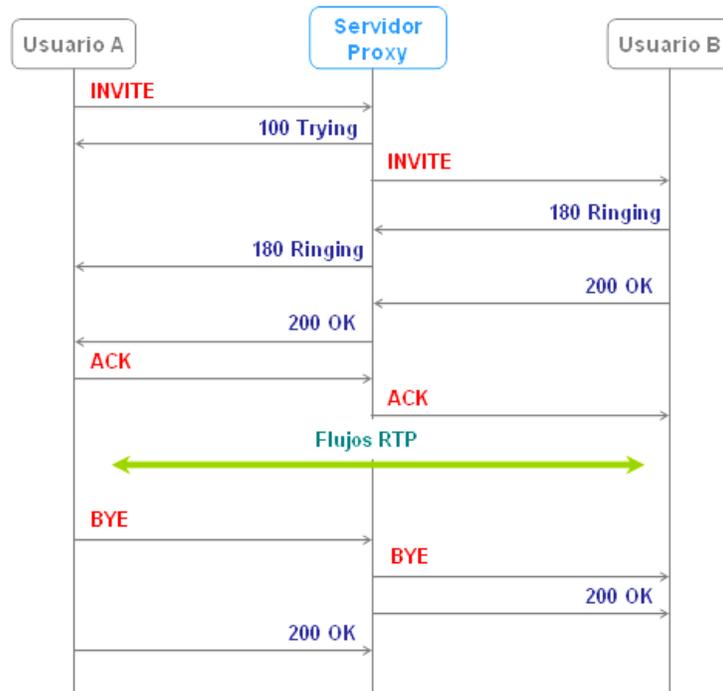


Figura A10: Terminación de sesión mediante mensajes SIP

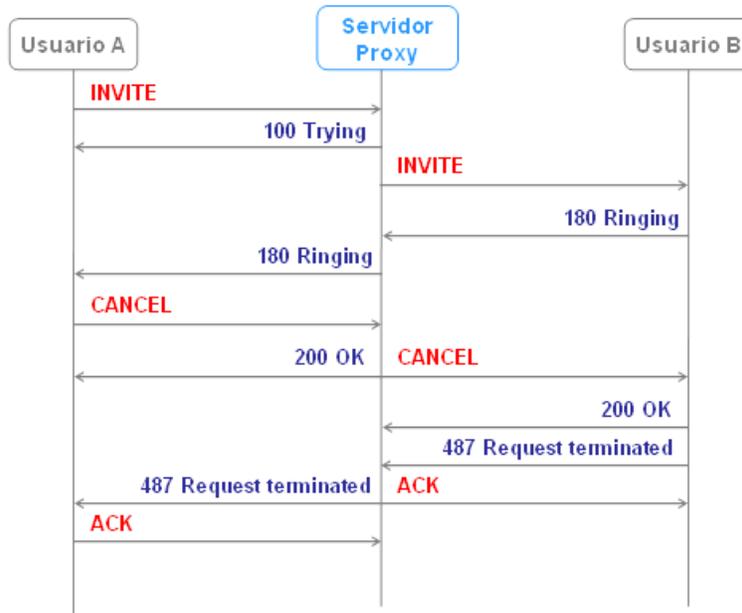


Figura A11: Cancelación de sesión mediante mensajes SIP

A1.2.7.6. Control de llamada REFER

La transferencia de llamada mediante mensajes SIP se puede ver en la Figura A12.

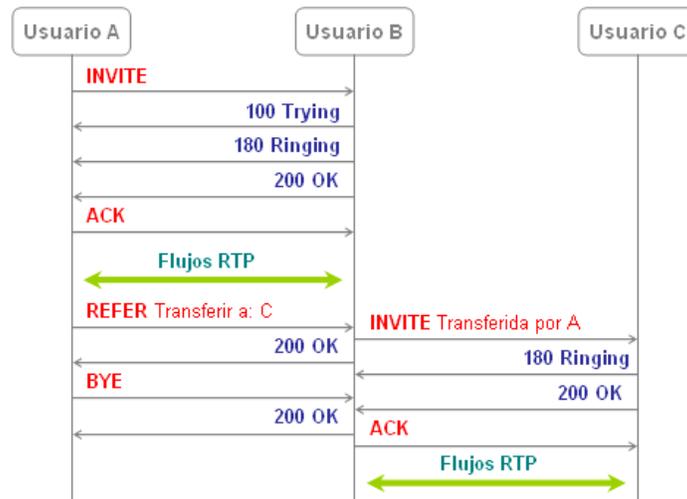


Figura A12: Transferencia de llamada mediante mensajes SIP



A1.2.8. Creación de servicios [23][31]

Unas de las características de SIP son la flexibilidad y facilidad para la de creación de servicios. Estos, pueden residir en los agentes de usuario, servidores Proxy y Redirect.

Por ejemplo, los servicios pueden ser implementados utilizando:

- **SIP CGI**

Es 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 archivo estándar de salida. Los scripts de CGI se pueden escribir en Perl, C, C++, TCL, Java, etc.

- **SIP CPL**

Fue la primera API que se desarrolló para SIP, y basada en XML. Está pensado 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 Java 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 propio servidor Web o de aplicaciones.

- **Java-enhanced SIP**

Los mensajes SIP se extienden para poder llevar applets de java.

- **JAIN SIP**

JAIN SIP es una API de bajo nivel, que hace una implementación del RFC 3261.

Estas tecnologías son extensiones del protocolo básico SIP.



A2. OPERACIÓN DE IMS EN CONJUNCIÓN CON SIP

A2.1. IMS Y SIP [32]

El 3GPP adopta el protocolo SIP, el cual fue estandarizado inicialmente por la IETF y le hizo los ajustes necesarios para proveer soporte completo a los requerimientos de una red IMS, los cuales son básicamente la definición de varias extensiones. Estas extensiones adaptan el protocolo SIP para su uso dentro de la arquitectura IMS, y son definidas en el estándar TS.24.229 del 3GPP.

SIP ha sido extendido con el fin de soportar numerosos servicios, tales como servicios de presencia, PTT, mensajería instantánea, transferencia de llamada, conferencia, servicios complementarios de telefonía, etc.

Por definición SIP no es un protocolo diseñado para una red o una aplicación específica. Para utilizar SIP, se debe definir un Perfil de uso. Los perfiles de uso trabajan como plantillas, y proporcionan un ambiente variado y flexible según los requisitos particulares.

El perfil de SIP que se creó para IMS es uno de los más importantes en el ámbito de las telecomunicaciones, pues no solo involucra las redes móviles sino toda la industria de las telecomunicaciones en conjunto y según los expertos actualmente es el más apropiado para las NGN.

Para una mejor comprensión de las funciones de las entidades IMS, en la secciones A2.3 y A2.4 se muestran dos de los procedimientos más comunes en IMS: el proceso de registro (necesario para que el abonado pueda acceder a los servicios IP multimedia) y el establecimiento de sesión (que permite iniciar las comunicaciones con otros abonados y con los servicios multimedia).

A2.2. CABECERAS DEL PROTOCOLO SIP PARA IMS [33][34]

A2.2.1. P-Associated-URI [35]

La cabecera `P-Associated-URI` contiene la lista de identidades públicas que el usuario autoriza, es decir, contiene un conjunto de URI's relacionadas con una dirección de registro.

La primera URI en la lista de las identidades públicas del usuario, provista por el HSS al S-CSCF, indica la identidad pública de usuario, que el S-CSCF usará por defecto. Un servidor de registro (Register) contiene información que permite a una URI (dirección de registro) asociarse a cero o más URI's, generalmente todas éstas se asignan a un usuario específico.

Esta extensión de SIP le permite al UAC saber, sobre un registro realizado, cuales son las URI's asociadas, si las hay. Esta cabecera es aplicable en redes SIP, donde el proveedor asigna a un usuario un conjunto de identidades que podrá utilizar, en este caso el proveedor tendrá



conocimiento de todas las identidades de cada usuario, tanto de las que tiene en uso, como las que tiene restringidas.

El servidor de registro inserta la cabecera `P-Associated-URI` en una respuesta `200 OK` de una petición `REGISTER`, y el valor del campo de cabecera se llena con las URI's asociadas a la URI de la dirección de registro. El servidor realiza este procedimiento para un proceso de registro, tanto de "re-registro", como de "de-registro".

A2.2.2. P-Called-Party-ID

Esta cabecera es insertada por un servidor Proxy, con el valor de la dirección lógica que registró un usuario, por lo general en una petición `INVITE`, antes de sustituirla con la dirección que va a utilizar para encaminar la petición al usuario. De este modo se asegura que el destino reciba la dirección lógica correspondiente a la petición.

A2.2.3. P-Visited-Network-ID

Esta cabecera está diseñada para realizar las funciones relacionadas con la itinerancia de los usuarios, es decir, cuando un usuario se traslada de su red local a una red visitada.

Esta cabecera se utiliza para transmitir el identificador de red visitada desde el Servidor de Registro al Proxy de la red local. Dicho identificador es entendido por el servidor de Registro o Proxy tanto de la red local como de la red visitada. Generalmente, la red local autoriza a un usuario para entrar en una red visitada, siempre y cuando existan acuerdos de itinerancia entre dichas redes.

A2.2.4. P-Access-Network-Info

Esta cabecera contiene información sobre la red de acceso que el UA está utilizando, esta información es conocida por el UA y requerida por el Proxy a la hora de proveer los servicios.

El protocolo SIP fue pensado para trabajar independiente de la tecnología de acceso, por lo tanto esta información no es de uso general del protocolo SIP. Además, la información que se transporta en esta cabecera es alterable fácilmente, la protección de esta información depende de la existencia de acuerdos y relaciones de seguridad entre los servidores Proxy que utilizarán el contenido de esta cabecera, y depende también del conocimiento que el UA tenga de esas relaciones. Es por eso que este mecanismo es apropiado solo para ambientes donde existan los elementos de seguridad apropiados.

Cuando un UA envía una solicitud o respuesta SIP al servidor proveedor de servicios, inserta esta cabecera en el mensaje SIP. La cabecera contiene información de la red de acceso que el UA está usando para tener conectividad IP. Generalmente, esta cabecera la ignoran los servidores intermedios que se encuentran entre el UA y el servidor proveedor del servicio. Este último, puede "leer" y usar la información que se encuentra en la cabecera, y luego quitarla del mensaje. Para



borrar la cabecera también deben existir transacciones seguras entre el UA y el servidor que provee los servicios, estas transacciones generalmente soportan mecanismos de seguridad como AKA.

El servidor que provee los servicios al usuario esta ubicado generalmente en su red local y por lo tanto es confiable, entonces se debe borrar la cabecera cuando la señalización SIP es enviada a un servidor ubicado en un dominio poco confiable.

A2.2.5. P-Charging-Function-Addresses [36]

Esta cabecera contiene los nombres de los servidores o las direcciones IP de los nodos que reciben la información de facturación. En esta cabecera se introduce una lista de direcciones de una o más entidades de facturación a las cuales un Proxy puede enviar información relacionada con la tarificación. La solución que provee el 3GPP define dos tipos de entidades funcionales de tarificación una es la CCF (Charging Collection Function) y la otra es ECF (Event Charging Function).

- CCF: se utiliza para la facturación de usuarios post-pago. [37]
- ECF: se utiliza para la facturación de usuarios prepago. [37]

La cabecera P-Charging-Function-Addresses no se incluye en los mensajes SIP enviados fuera del dominio de la red, tampoco si el dominio no provee una función de facturación.

A2.2.6. P-Charging-Vector [36]

Cabecera que proporciona información para poder correlacionar los registros de tarificación generados, por cada una de las entidades de red involucradas en una misma sesión.

Existen tres tipos de información de correlación que se puede transmitir: El valor de la identidad de tarificación de IMS, llamada ICID (IMS Charging Identity), la dirección del servidor Proxy SIP que crea el valor del ID, y el IOI (Inter operator identifiers).

- ICID: se utiliza para correlacionar archivos de tarificación, y debe ser un valor global único.
- IOI: se utiliza para identificar la red que origina y donde finaliza una transacción SIP. Existe un parámetro IOI para cada lado de la transacción, es decir uno para la red origen y otro para la red destino.

El vector de tarificación es definido como un recolector de información de tarificación, el cual se utiliza para transportar información de facturación, como por ejemplo el valor del ICID.

Más información de estas cabeceras se podrá encontrar en el RFC 3455 del IETF. [33]



A2.3. PROCEDIMIENTO DE REGISTRO [4]

Antes de utilizar los servicios del dominio IMS, tales como establecer una sesión multimedia o recibir una sesión, el usuario debe registrarse en IMS. Este proceso involucra un P-CSCF de la red visitada si el usuario está en una red visitada o un Proxy Controlador de la Función de Sesión de Llamada (Proxy-CSCF, P-CSCF) de la red nominal si el usuario está en su red.

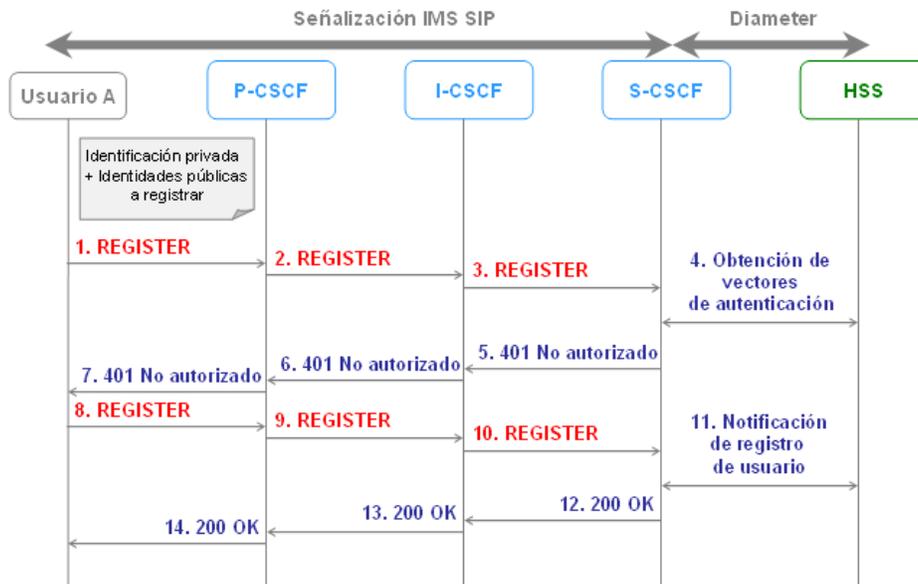


Figura A13: Procedimiento de registro en IMS

El procedimiento de registro en IMS se representa en la Figura A13. Este consta de las siguientes fases:

- En primer lugar, como paso previo para acceder a IMS, el usuario debe registrarse en el sistema. Mediante este proceso se activan las identidades públicas que el usuario desea emplear en sus sesiones multimedia y se establece el S-CSCF que le aportará el servicio. Para llevarlo a cabo, se emplea la señalización SIP y un algoritmo de autorización/autenticación por desafío de usuario a red, y viceversa, que recibe el nombre de Clave de Acuerdo y Autenticación IMS (IMS Authentication and Key Agreement, IMS-AKA).

Más información acerca de IMS-AKA se puede encontrar en 3GPP TSG SA WG3 Security [38] y 3GPP TS 33.102 V7.0.0 [39].

- A continuación, el usuario inicia el proceso enviando un mensaje SIP REGISTER hacia el P-CSCF, que detecta que se trata de un mensaje que no se encuentra protegido por ninguna asociación de seguridad previa; es decir, se trata de un mensaje de registro inicial. En ese mensaje se encuentran la identidad privada del usuario, almacenada en la ISIM, y las identidades públicas que desea registrar para su posterior uso.

En esta fase, el P-CSCF envía el mensaje hacia un I-CSCF, que se encarga de seleccionar un S-CSCF hacia el que reenvía la petición de registro. Cuando el S-CSCF recibe el mensaje,



comprueba que no se trata de un usuario que ya este registrado y contacta con el HSS para obtener los vectores de autenticación, necesarios para el algoritmo IMS-AKA.

Posteriormente, para solicitar la autenticación, devuelve hacia el terminal móvil un mensaje SIP 401 "No autorizado", en el que se incluyen ciertos números generados aleatoriamente, así como las claves para el cifrado y protección de la integridad de la señalización IMS.

- Por último, el usuario, en base al mensaje de desafío especificado en la anterior fase, comprueba la identidad de la red IMS y genera un mensaje nuevo SIP del tipo REGISTER. Este segundo mensaje contiene una respuesta formada a partir del algoritmo de autenticación IMS-AKA.

Cuando el mensaje llega al S-CSCF, el usuario es registrado finalmente después de comprobar la veracidad de su identidad. Posteriormente, el S-CSCF indica al HSS que aquel ha registrado al abonado satisfactoriamente y descarga desde allí la suscripción IMS del usuario. El proceso finaliza con el asentimiento SIP 200 OK enviado hacia el terminal móvil.

Gracias al registro IMS (utilizando el protocolo SIP):

- La red IMS notifica al HSS de la localización actual del terminal del usuario y actualiza su perfil.
- El usuario es autenticado antes de poder acceder a los servicios IMS.
- El dominio IMS nominal del usuario selecciona un S-CSCF apropiado responsable de la invocación de sus servicios. Estos últimos se encuentran en servidores de aplicación. El S-CSCF tiene el perfil del usuario (como la Información de Suscripción en el Servidor de Aplicaciones o ASSI) que le permite conocer la lista de servicios a invocar.

A2.4. ESTABLECIMIENTO DE SESIÓN [4]

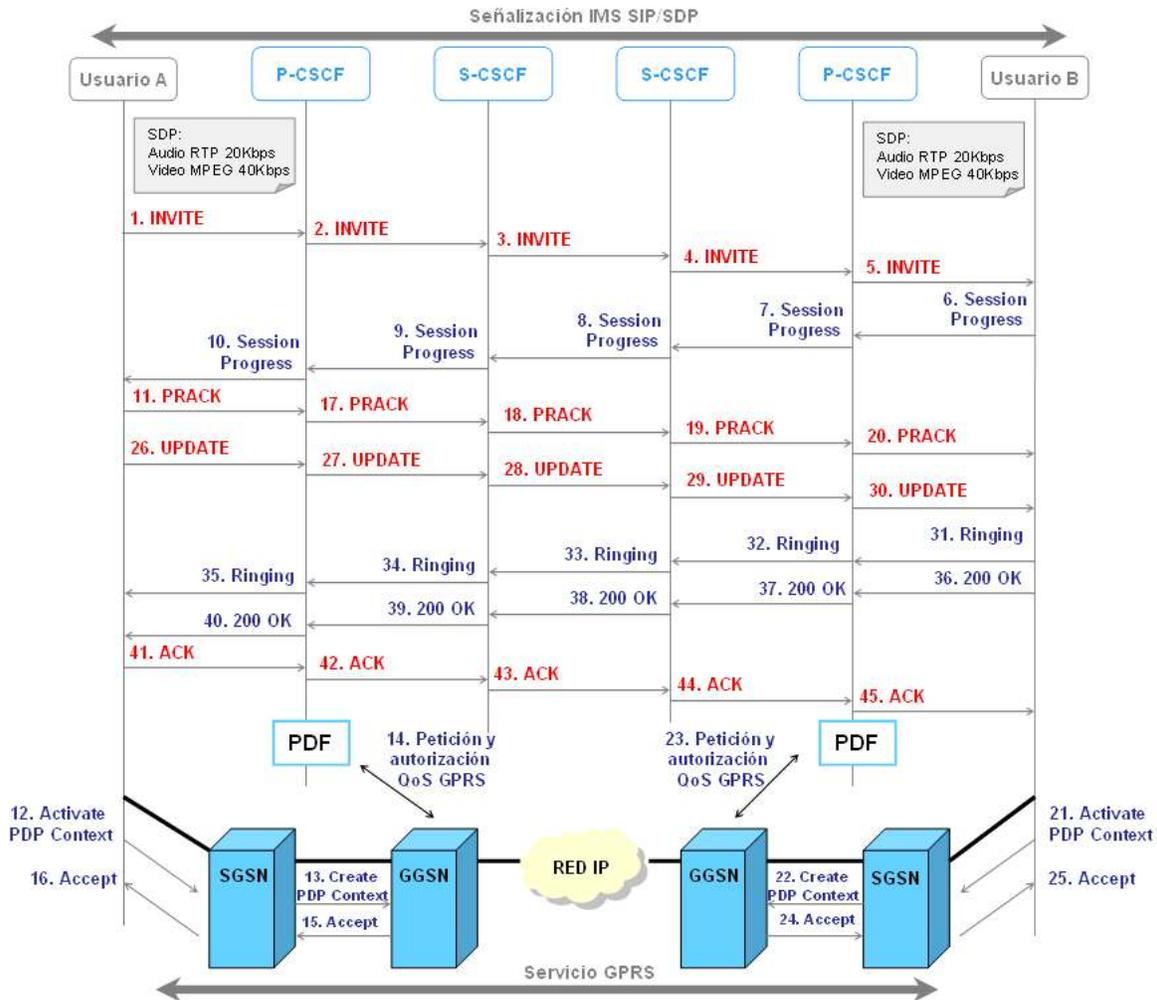


Figura A14: Procedimiento de establecimiento de sesión en IMS

El procedimiento de inicio de sesión se muestra en la Figura A14. Las fases de este proceso son las siguientes:

- En primer lugar, una vez que el usuario ha sido registrado en el subsistema IMS, aquel puede acceder a los servicios IP multimedia que proporciona IMS. De esta manera el usuario, por ejemplo, podría establecer una sesión de videoconferencia con otro de una red diferente. Para ello, utilizará el subsistema IMS para intercambiar información de señalización mediante los protocolos SIP y el Protocolo de Descripción de Sesión (Session Describing Protocol, SDP) con el usuario con el que se quiere comunicar. El objetivo de este intercambio de señalización es el establecimiento de una sesión, mediante la cual se contactará con el nodo destino, se negociarán los parámetros de sesión y se activarán los recursos GPRS necesarios para soportar la sesión multimedia.



Para realizar lo anterior, el usuario origen debe enviar a través de IMS un mensaje SIP `INVITE`, en el que añadirá también el mensaje SDP que describe las capacidades de la sesión que pretende establecer.

En ese mensaje SDP estarán incluidos los medios que quiere transmitir, la tasa binaria a la que se transmitirá cada medio, los protocolos utilizados para la transmisión de los medios, los codecs que se utilizarán, etc. La señalización SIP y SDP llegará al usuario remoto pasando por los nodos IMS de la red origen y destino.

- A continuación, el terminal destino retorna al nodo origen un mensaje SIP de progreso de sesión (`183`), en el que se añade un mensaje SDP con la respuesta del nodo destino al ofrecimiento de los parámetros SDP del nodo origen. Estos parámetros pueden haber sido modificados en función de las capacidades del terminal o las preferencias del usuario.
- Seguidamente el terminal origen envía un mensaje `PRACK` como respuesta al de progreso de sesión, en el que está incluida la oferta SDP final. En este momento es cuando se activan los recursos necesarios a nivel GPRS en la red origen, para soportar los medios que se han negociado. Si se ha habilitado el control de QoS mediante COPS, IMS puede interactuar con el nivel GPRS para autorizar los recursos y la QoS para cada medio.

Por su parte, en la red destino se activan los recursos GPRS, una vez que llega el `PRACK` con el mensaje SDP final. Para simplificar, en la Figura A14 no se ha mostrado el mensaje SIP `200 OK` de respuesta del `PRACK` desde el terminal destino hacia el origen. Cuando éste recibe el mensaje `200 OK`, envía un mensaje SIP `UPDATE` para indicar al destino que ha tenido éxito la operación de activación GPRS.

Es entonces cuando el terminal destino avisa acerca de la llamada que está entrando, a la vez que envía la señalización SIP para indicar al terminal origen que el usuario destino está siendo alertado.

- Por último, cuando el usuario destino descuelga para recibir la sesión de videoconferencia, su terminal envía otro mensaje `200 OK`, con el que se confirma el establecimiento definitivo de la sesión desde el usuario remoto. Si se implementa el control QoS desde IMS, este mensaje activa a su paso, a través de IMS, el plano de transporte GPRS para autorizar la transferencia de paquetes, y empieza el intercambio de tráfico de usuario, compuesto de audio y vídeo, entre los terminales que soportan el servicio GPRS.

A2.5. FLUJOS DE SESIÓN [6]

Existen dos tipos de flujo, uno es el flujo de mensajes de señalización y otro es el flujo de datos del usuario.

Los mensajes de señalización irán del móvil a través de la Red Terrestre de Acceso Radio UMTS (UMTS Terrestrial Radio Access Network, UTRAN) o el Subsistema de la Estación Base (Base Station Subsystem, BSS) de Acceso Radio GPRS EDGE (GPRS EDGE Radio Access, GERAN), al Nodo de Soporte del Servidor GPRS (Server GPRS Support Node, SGSN) y al Nodo de Soporte de la Pasarela GPRS (Gateway GPRS Support Node, GGSN), para luego ir al CSCF y finalmente a la red destino (a otra red de IMS Release 5, a una red de MGCF con la Pasarela de Medios, o simplemente hacia la red IP). Es importante distinguir entre qué componentes procesan el mensaje



y cuáles lo enrutan. Cuando el móvil envía una petición para establecer un servicio, la petición se envía al S-CSCF (vía el P-CSCF y el I-CSCF). El SGSN y el GGSN realizarán solamente la función de enrutadores. No miran el contenido del mensaje, sólo miran la dirección IP del destino y encaminan el mensaje de acuerdo a ésta.

La información del usuario se transmitirá del móvil, a través del SGSN y el GGSN hacia la red destino (al GGSN de otra red de Release 5, o a la red PSTN pasando por la Pasarela de Medios MGW, o hacia Internet), saltándose la red de CSCFs. Es importante notar que en IMS por tanto se sigue la filosofía de que señalización y datos toman caminos distintos en la red.



A3. SERVICIOS IVR SOBRE IMS

A3.1. SISTEMAS IVR

A3.1.1. IVR (Interactive Voice Response)

IVR es una tecnología de automatización de contacto con el usuario, es decir pretende que un operador humano se reemplace por un programa interactivo de alta calidad que provea respuestas a las peticiones de los usuarios, realizadas a través del teclado telefónico o comandos de voz. [40]

Esto difiere de CATI (Entrevista telefónica asistida por computador) donde un entrevistador lee las preguntas e ingresa la respuestas en un computador [41]. Muchos atributos de CATI son compartidos por IVR incluyendo salto lógico automático o ramificación basada en respuestas, entrevistas que pueden ser pausadas y reanudadas en un tiempo posterior, validación de respuestas, control de grabación, etc. Además, un sistema IVR puede ofrecer ventajas considerables tales como: Un sistema IVR puede correr automática y continuamente, los llamantes pueden responder a su propia velocidad; Mejoramiento en la exactitud de la entrevista (cuando se utiliza para ese fin); Reducción de costos y un alto grado de confidencialidad [42].

IVR ha sido utilizado inicialmente en la construcción de “Calls Center” o Centros de Llamada y en compañías grandes para manipular un amplio número de llamadas recibidas en los departamentos de atención al usuario. Estos servicios han sido utilizados en el sistema bancario y en sistemas de reservación. [40]

Los sistemas modernos IVR tienen las siguientes características [43]:

- Facilidad de reconocimiento de voz de llamante.
- Soporte para múltiples lenguajes.
- Capacidad de integración con otras plataformas IVR.

La mayoría de los sistemas utilizan interfaces DTMF donde un usuario marca un número y un computador responde la llamada e inicia el proceso de dialogo. Una vez la sesión es terminada, el sistema IVR desconecta la llamada.

Para 1980 DTMF era el estándar de FACTO para IVR. Sin embargo esta interfaz está siendo reemplazada por una técnica capaz de reconocer la voz con más precisión. En 1980 el ARPA fundó una organización de trabajo en el desarrollo de sistemas de reconocimiento de voz basado en computador. [44]

Aunque las tecnologías de reconocimiento de voz han avanzado notablemente los sistemas basados en DTMF todavía sobreviven debido a que son menos costosos y de desarrollo rápido comparado con los sistemas de voz natural. [45]

La tecnología IVR moderna se soporta en dos tecnologías afines:

- TTS (Síntesis de Voz)
- ASR (Reconocimiento de Voz)



A3.1.2. Conversión de Texto a Habla [43][44][45]

Permite a los computadores y otros sistemas electrónicos simular respuestas de voz, es decir, convertir textos escritos en locuciones de voz natural.

Las dos características utilizadas para describir la calidad de la síntesis de voz son la *naturalidad* e *inteligibilidad*. La *naturalidad* se refiere a cuánto el resultado se aproxima a la voz de una persona real. La *inteligibilidad* a la facilidad de entender el habla sintetizada. La síntesis ideal debe de ser a la vez natural e inteligible, y cada tecnología intenta conseguir el máximo de ambas.

A continuación se explican de manera general las tecnologías más utilizadas en la conversión de texto a voz.

- **Síntesis por selección de unidades**

La síntesis por selección de unidades utiliza una base de datos de voz grabada (más de una hora de habla grabada). Durante la creación de la base de datos, el habla se segmenta en algunas o todas de las siguientes unidades: fonemas, sílabas, palabras, frases y oraciones. Típicamente, la división en segmentos se realiza usando un reconocedor de voz modificado para forzar su alineamiento con un texto conocido. Después se corrige manualmente, utilizando representaciones como la forma de onda y el espectrograma. Se crea un índice de las unidades en la base de datos basada en parámetros acústicos de la segmentación como la duración, la posición en la sílaba y los fonemas vecinos. En tiempo de ejecución, el objetivo deseado se crea determinando la mejor cadena de candidatos de la base de datos (selección de unidades). Este proceso se logra utilizando fundamentalmente un árbol de decisión ponderado adecuadamente.

La selección de unidades da la máxima naturalidad debido a que no aplica en gran medida el procesamiento digital de la señal al habla grabada, lo que a menudo hace que el sonido grabado suene menos natural, aunque algunos sistemas usan un poco de procesamiento de señal en la concatenación para suavizar las formas de onda. En este sentido, la salida de la mejor selección de unidades es a menudo indistinguible de la voz humana real, especialmente en contextos en los que el sistema ha sido adaptado. Por ejemplo, un sistema de síntesis de voz que se encargue de dar informaciones de vuelos puede ganar en naturalidad si la base de datos fue construida a base de grabaciones de informaciones de vuelos, ya que será más probable que aparezcan unidades apropiadas e incluso cadenas enteras en la base de datos. Sin embargo, la máxima naturalidad a menudo requiere que la base de datos sea muy amplia, llegando en algunos sistemas a los Gigabytes de datos grabados.

- **Síntesis de difonos**

La síntesis de difonos utiliza una base de datos mínima conteniendo todos los difonos que pueden aparecer en un lenguaje dado. El número de difonos depende de la fonotáctica del lenguaje: el español tiene unos 800 difonos, el alemán unos 2500. En la síntesis de difonos, la base de datos contiene un sólo ejemplo de cada difono. En tiempo de ejecución, la prosodia de una oración se impone a estas unidades mínimas mediante el procesamiento digital de la señal, como codificación lineal predictiva, PSOLA o MBROLA.



- **Síntesis específica para un dominio**

La síntesis específica para un dominio concatena palabras y frases grabadas para crear salidas completas. Se utiliza en aplicaciones donde la variedad de textos que el sistema puede producir está limitada a un dominio particular, como anuncios de salidas de trenes o información meteorológica.

Esta tecnología es muy sencilla de implementar, y se ha utilizado comercialmente durante largo tiempo en aparatos como relojes y calculadoras parlantes. La naturalidad de estos sistemas puede ser muy alta porque la variedad de oraciones está limitada y corresponde a la entonación y la prosodia de las grabaciones originales. Sin embargo, al estar limitados a unas ciertas frases y palabras de la base de datos, no son de propósito general y sólo pueden sintetizar la combinación de palabras y frases para las que fueron diseñados.

A3.1.2.1. El proceso de la síntesis del habla [46]

En idiomas como el español y el portugués los sintetizadores permiten localizar las vocales sobre las que recae el acento para evitar ambigüedades.

Otra mejora que ofrecen los sintetizadores de voz es la localización de pausas no marcadas ortográficamente, estas pausas se insertan cuando se detecta que no es posible leer un fragmento demasiado largo sin hacer ninguna pausa intermedia.

Los programas de conversión de texto a voz toman la mayoría de entradas desde un archivo de texto. Mientras se recibe el texto, es necesario convertirlo a Unicode o a otro formato nativo.

Como se observa en la Figura A15, los sintetizadores actuales realizan procesos complementarios para mejorar la calidad de la salida de voz, esta tarea consiste en asignar palabras convencionales a números, abreviaturas o símbolos para facilitar su correspondencia con el lenguaje hablado, a este proceso se le denomina: normalización.

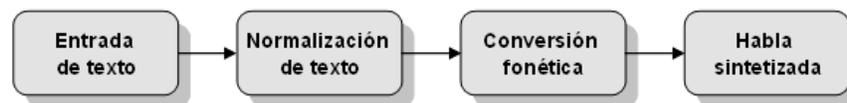


Figura A15: Proceso de Síntesis del Habla

El proceso de normalizar texto es pocas veces simple. Los textos están llenos de homógrafos, números y abreviaturas que tienen que ser transformados en una representación fonética. Sin embargo, en lenguas donde la correspondencia entre el texto escrito y su equivalente fonético es poca (por ejemplo el inglés) o ninguna (por ejemplo el mandarín), la creación de estos sistemas se complica.

Muchos sistemas de texto a voz no generan representaciones semánticas de los textos de entradas, ya que los sistemas para hacerlo no son fiables o efectivos. Como resultado, se utilizan varias técnicas heurísticas para estimar la manera correcta de eliminar la ambigüedad de los homógrafos, y entre estas se tienen buscar palabras vecinas y usar estadísticas sobre la frecuencia de aparición de las palabras.



Decidir como convertir números en palabras es otro problema que tienen que solucionar los sintetizadores de voz. Es un desafío en gran parte simple el programar un sistema que convierta números en palabras, como por ejemplo transformar 1325 en "mil trescientos veinticinco". Sin embargo, los números aparecen en diferentes contextos, y mientras 1325 puede ser un número ordinal, "uno tres dos cinco" pueden ser los últimos dígitos de una identificación o "trece veinticinco" un número de teléfono. A menudo un sistema de síntesis de voz puede inferir como expandir un número con base en las palabras o números vecinos y la puntuación, y algunos sistemas proporcionan un sistema para especificar el tipo de contexto si este resulta ambiguo.

El paso siguiente a la normalización del texto es la conversión fonética que se describió anteriormente (síntesis por unidades, dítonos, etc.), hasta llegar obtener habla sintetizada representada por muestras o archivos de audio.

A3.1.3. Reconocimiento Automático del Habla (RAH) [47]

Se define como la conversión de una señal acústica (recogida por un micrófono o por un teléfono) en su equivalente escrito. Los sistemas ASR se basan en un modelo de generación de habla estudiando el proceso de comunicación humano, donde el flujo de aire es modulado para producir potencia acústica en un determinado rango de frecuencias.

En primer lugar, se debe distinguir los sistemas de reconocimiento del habla de los sistemas de comprensión del habla.

Se suele considerar que la comprensión del habla es un concepto más amplio, que si bien incluye entre otras partes a un sistema de reconocimiento su objetivo es el de capturar la semántica del mensaje, y de esta forma no solamente transcribirlo en texto sino entenderlo correctamente.

La interfaz de voz es muy prometedora puesto que actualmente la mayoría de la gente sigue tecleando unas 60 palabras por minuto (en el mejor de los casos) cuando podría llegar a pronunciar unas 200 en el mismo tiempo. Sin embargo, las dificultades asociadas con estos sistemas son:

- Cantidad de información en una señal de voz.
- Flujo continuo de voz (concatenación de palabras).
- Variabilidad entre las palabras dichas del mismo locutor (debido a la velocidad, estado de ánimo, etc).
- Variabilidad entre locutores.
- Ruido, calidad del micrófono, etc.

Por otro lado, el respaldo tecnológico que contribuye al avance de esta tecnología está relacionado con los siguientes campos:

- Microprocesadores más poderosos
- Integración de computadoras y telefonía (CTI)
- Crecimiento de aplicaciones multimedia



A3.1.3.1. Tipos de motores de reconocimiento [48][49]

- 1) *En función del hablante*: según el grupo de locutores que se utilizan en el entrenamiento y evaluación del sistema se pueden distinguir 3 tipos de reconocedores:
 - Reconocedores *monolocutores* y se evalúa para un único hablante.
 - Reconocedores *multilocutor* el entrenamiento y evaluación se realiza sobre un grupo de locutores específico.
 - Motores *independientes del hablante* en los cuales la evaluación del sistema se efectúa sobre un conjunto de locutores diferentes para entrenarlo.

- 2) *En función de la manera de hablar y la aplicación*: existen los siguientes 3 tipos:
 - Reconocedores de palabras completas emitidas de forma aislada entre sí.
 - Reconocedores de palabras conectadas, se utilizan las palabras como unidades de reconocimiento, pero éstas pueden ser emitidas secuencialmente y con pausas entre ellas.
 - Reconocedores de voz continuos, los cuales realizan las tareas atendiendo a unidades más pequeñas que la palabra (fonemas, difonos, etc) y sobre frases completas sin necesidad de establecer silencios entre las palabras que las constituyen.

- 3) *En función de las distorsiones de la señal*: se pueden tener en cuenta las distorsiones producidas en la señal de voz por distintos factores. En este sentido, se pueden separar los reconocedores en dos tipos:
 - *Reconocedores de habla limpia*, son aquellos entrenados y probados en condiciones de laboratorio.
 - *Reconocedores robustos*, los cuales permiten su uso en ambientes reales donde varios factores complican la tarea (ruido, reverberación, etc.).

A3.1.3.2. Análisis de la Señal Acústica [48][49]

El proceso de resonancia acústica es el más importante para determinar las propiedades del sonido. La principal estructura resonante, particularmente para vocales, es el tracto vocal, que inicia desde la laringe y se extiende hacia arriba en la faringe la boca y los labios. En algunos sonidos también se incluye la cavidad nasal que hace más complicado el sistema resonante.

Los modos resonantes del tracto vocal son conocidos como formantes y por lo general son numerados desde la frecuencia más baja. Por consiguiente son referidos como F_1 , F_2 , F_3 . En general F_1 y F_2 usualmente en el rango de 250hz a 3Khz son las más significativas en la determinación de las propiedades fonéticas de los sonidos del habla. Pero algunos formantes de frecuencia más alta pueden ser importantes para ciertos fonemas

La frecuencia fundamental de la señal de habla se encuentra típicamente en el rango de 50-200Hz, para hombres adultos y cerca de una octava más alta para mujeres.

Uno de los métodos empleados con el fin de determinar las propiedades del habla es el espectro de señal de corto tiempo o función de ventanas de tiempo, equivalente a la magnitud de la



transformada de Fourier multiplicada por una franja de tiempo (unos pocos milisegundos), y el reconocimiento basado en *patrones*, que se verá a continuación.

A3.1.3.2.1. Reconocimiento basado en Patrones [49]

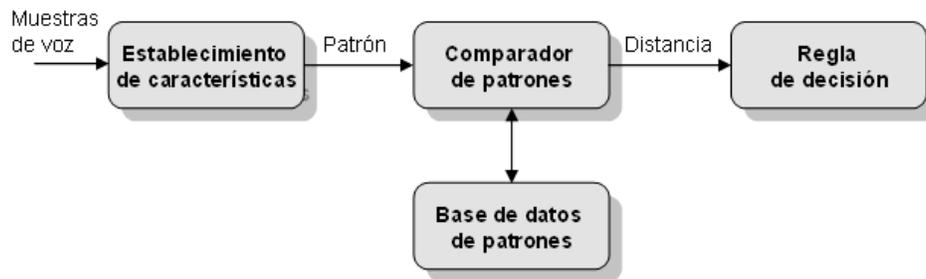


Figura A16: Proceso de Reconocimiento del Habla basado en patrones

Esta teoría ha dado lugar a toda una familia de técnicas de reconocimiento de voz, causando un considerable avance durante las décadas de los 70's y 80's.

El principio, que se puede ver en la Figura A16, consiste en lo siguiente: se almacenan patrones de voz típicos, como modelos de referencia, en un diccionario de palabras candidato. El reconocimiento se lleva a cabo comparando la expresión desconocida con los patrones almacenados y seleccionando el que más se le parece.

Por lo general, se construyen patrones para las palabras completas. La ventaja es que se evitan errores debidos a la segmentación o clasificación de unidades pequeñas que pueden variar acústicamente, como los fonemas. Y la desventaja es que cada palabra requiere de su patrón y el tiempo para preparar y compararlos se vuelve demasiado grande al incrementar el tamaño del vocabulario.

De este método, nació uno que ha sido ampliamente utilizado en el desarrollo de reconocedores de voz, el cual corresponde al de Modelos Ocultos de Markov (MOM).

A3.1.3.2.2. Modelos Ocultos de Markov [50]

Los modelos ocultos de Markov (MOM) constituyen una de las técnicas que se ha utilizado con más éxito en el reconocimiento automático del habla (RAH).

Principalmente, esta técnica ha permitido modelar adecuadamente la gran variabilidad en el tiempo de la señal de voz. En la terminología del ASR con MOM suele hacerse referencia no sólo a la técnica de los modelos ocultos de Markov, propiamente dicha, sino también a una larga lista de adaptaciones y técnicas asociadas que se fueron incorporando para solucionar el problema de RAH.

Si se considera el proceso de la comunicación oral se podría pensar que para cada texto el locutor activa un sistema y da como salida una determinada emisión sonora. Para comenzar a entender como se aplican los MOM al RAH se debe empezar por imaginar que para cada una de las posibles emisiones se puede encontrar un modelo capaz de imitar al sistema y activado por el locutor. Es decir, un modelo que sea capaz de generar la misma emisión que generó el locutor a



partir del texto que había en su mente. De esta forma, se puede suponer que se cuenta con tantos modelos como posibles emisiones pueda hacer el locutor y, para cada modelo un texto asociado. En caso de que se conozcan perfectamente estos modelos, se podría utilizar el camino inverso para resolver el problema de RAH. Teniendo una determinada emisión del locutor, lo que se haría es buscar cuál de todos los modelos generaría el sonido más parecido al que generó el locutor. Al encontrar el modelo que genera el sonido más parecido a la emisión del locutor entonces también se habrá encontrado el texto, ya que se había dicho que todos los modelos estaban asociados a un determinado texto.

A3.2. SERVICIOS IVR EN IMS

Uno de los objetivos del presente Trabajo de Grado es la implementación de un prototipo de cliente SIP para dispositivos móviles con capacidades de acceso a servicios IVR en base a la arquitectura de servicios IMS. Por lo tanto, antes de exponer las características, requerimientos y provisión de servicios IVR soportados por IMS, es necesario conocer y definir el escenario a través del cual los usuarios van a tener acceso a estos. El primer tema que se debe tratar en esta sección es entonces el acceso a los servicios VoIP para usuarios móviles. Una vez identificado y caracterizado el escenario de acceso, se procederá a describir la forma como la arquitectura IMS especifica actualmente la prestación de servicios de VoIP, más específicos, como lo son los servicios de IVR (ASR y TTS), enfocando el estudio en aquellos que se ajustan a las capacidades tecnológicas actuales de las redes de 2.5G.

A3.2.1. La Voz Sobre IP (VoIP) en las Redes Móviles [7][13][51]

VoIP es la tecnología que permite el soporte de comunicaciones de voz en redes IP. Apareció a mediados de los años 90, y probablemente fue en el año 1998 cuando disfrutó de su primer despegue importante. Pero solo hasta el año 2003 se puede considerar el ascenso definitivo de esta tecnología.

La red telefónica basada en la conmutación de circuitos se ha utilizado para la transmisión de la voz durante más de 100 años. Sin embargo, en la actualidad, el crecimiento del tráfico y de las redes de datos es muy superior al de voz. Las redes IP son las redes de paquetes más comúnmente extendidas en todo el mundo, por lo que VoIP se convierte en la solución más atractiva, dado los bajos costos de Internet; enmarcándose además en la convergencia hacia una red multiservicio única.

Los operadores móviles han hecho una clara apuesta de futuro por esta tecnología, como queda reflejado en los últimos Releases 5 y 6 del 3GPP con la introducción del subsistema IMS basado en SIP. Sin embargo, las características de la interfaz aire en comparación con los sistemas de transmisión cableados, por un lado, y la movilidad de usuario que ofrece el sistema, por otro, dificultan la adopción de la tecnología IP y requieren una mayor adecuación técnica para cumplir los requisitos estrictos de temporización y QoS que el servicio clásico de voz impone, además de otros requerimientos adicionales tales como la codificación de voz, interconexión de redes, etc.

A continuación, se describen de manera resumida algunos temas que se deben tener en cuenta al hablar de la VoIP.



- **La QoS en VoIP**

Las redes de paquetes degradan por su naturaleza la calidad de servicio de las comunicaciones de voz, debido a varios factores:

El retardo. Los participantes en una conversación telefónica consideran retardos inaceptables, extremo a extremo, aquellos superiores a 200-250ms. Este retardo depende de la carga del sistema.

El jitter. Se asocia a la variación del tiempo entre las llegadas de los paquetes pertenecientes a un mismo flujo. La solución natural al jitter es el empleo de un buffer en recepción que amortigüe la varianza del retardo (en el buffer entrarán paquetes con jitter y se sacarán a una tasa constante). Sin embargo, el buffer repercute negativamente en el retardo.

La pérdida de paquetes. Debido a la característica temporal de las comunicaciones de voz, la retransmisión de los paquetes no es factible. Las dos técnicas empleadas para recuperación ante pérdidas son:

- Interpolación de la información perdida a partir de la conocida. Es una solución acertada cuando las pérdidas son pocas.
- Enviar información redundante. Es menos sensible a las pérdidas de ráfagas, pero aumenta el retardo y el consumo de ancho de banda.

- **Codificación de la voz**

Las premisas de la codificación de la voz son obtener representaciones que sean compactas, que degraden poco la señal original y que requieran poca capacidad de procesamiento. En VoIP se emplean codecs con tasas más de 10 veces inferiores a los 64Kbps de la G.711 (MIC), que es el estándar en las redes de telefonía, todo ello con una calidad de voz bastante aceptable.

La ITU-T normaliza los esquemas de codificación de voz CELP, MP-MLQ, PCM y la Modulación por Impulsos Codificados Diferencial y Adaptativa (Adaptive Differential Pulse Code Modulation, ADPCM) en sus recomendaciones de la serie G. Entre los estándares de codificación más populares para telefonía y voz por paquetes, además del G711, se incluyen: el G.726 (codificación ADPCM a 40, 32, 24, 16 Kbps), el G.729 (describe la compresión CELP que permite que la voz sea codificada en corrientes de 8Kbps), y el G.723.1 (describe una técnica de compresión que se puede utilizar para comprimir voz u otros componentes de la señal de audio de los servicios multimedia a una baja velocidad de bit).

- **Interconexión de redes**

Entre las capacidades que deberá tener una red IP para poder ofrecer servicios de VoIP (Computador a Computador, Computador a Teléfono, Teléfono a Teléfono, etc.) tiene gran importancia la interoperación con la PSTN.

A3.2.1.1. Establecimiento del camino de habla

Así como en las redes de VoIP fijas, en las móviles la voz digitalizada de cada usuario se transporta en paquetes IP a través de la red y entre los terminales de usuario. El camino que toman estos paquetes a través de la red es el camino del habla. A diferencia del entorno de la conmutación de circuitos, los paquetes pueden tomar rutas diferentes a través del núcleo de la red



IP para llegar a un punto común, en lugar de estar forzados a seguir un circuito específico. Sin embargo, si la red no está congestionada es probable que los paquetes sigan una misma ruta.

Para el establecimiento de un camino para el habla y para sincronizar a los usuarios y sus equipos, se programan funcionalidades de control de llamada tanto en los equipos de usuario como en la red. Dichas funciones de control de llamada se comunican utilizando mensajes de señalización. El control de la llamada, por ejemplo, permite poder pasar al terminal del usuario las direcciones del punto final para poder establecer el camino del habla, la negociación de los recursos, que se necesitan para la llamada, tanto de la red como del equipo de usuario, entre los cuales se tienen los codecs y la calidad de servicio que se requiere.

A3.2.1.2. Señalización

Los dos protocolos principales de señalización para la VoIP son SIP de la IETF y H.323 de la UIT. Es probable que ambos protocolos convivan a mediano plazo, pero dependiendo del escenario dentro del cual un determinado tipo de servicio se despliegue, se debe decidir qué protocolo es más válido. El presente trabajo se enmarca dentro de un escenario basado en la arquitectura de servicios IMS, por lo cual es SIP el protocolo utilizado para la señalización de los servicios VoIP que se van a estudiar en las secciones posteriores.

En la Figura A17 se muestra la pila de protocolos utilizada para VoIP al emplear SIP como protocolo de señalización y RTP como el protocolo utilizado para la transmisión de los datos de voz en los terminales móviles.

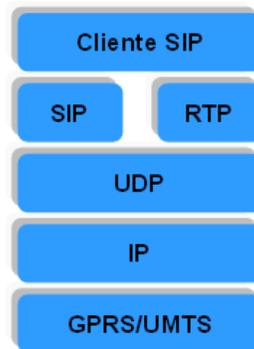


Figura A17: Pila de protocolos utilizada en VoIP para redes móviles

El protocolo de señalización SIP se utiliza para el registro y el control de la llamada. En términos generales, el control de la llamada hace referencia a la invitación y sincronización de varios participantes en una llamada. Esto permite a los participantes describir y compartir la información de las características de sus terminales y del camino de habla a utilizar entre ellos. La anterior se conoce como la información de descripción de la sesión, y puede incluir, por ejemplo, el codec a utilizar para el habla y el ancho de banda necesario para los caminos de esta.

SIP incluye mecanismos de confiabilidad que pueden utilizarse cuando el transporte no es seguro, pero estos se pueden omitir si se trata de protocolos de transporte como el Protocolo de Control de la Transmisión (Transmission Control Protocol, TCP) o el Protocolo de Control del Flujo de la Transmisión (Stream Control Transmission Protocol, SCTP). Estos protocolos son transportados sobre IP, la red GPRS, y el dominio IM del núcleo de la red IP como se observa en la Figura A18.

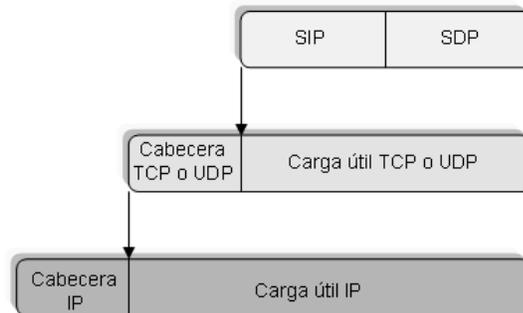


Figura A18: Transporte de SIP y SDP en IP

A3.2.1.3. Transporte de los paquetes de habla

Con el fin de que el habla se pueda enviar y recibir en paquetes IP, el habla del usuario se muestrea por medio del equipo de él mismo y se codifica para la transmisión (utilizando por ejemplo el codec de voz de multi-tasa adaptativa AMR). Una vez se halla tomado un número de muestras, que usualmente se encuentran entre los 10 y los 40 milisegundos, las muestras codificadas se empaquetan y se envían a la red. El tiempo que toma realizar el empaquetamiento de las muestras de habla, le adiciona un retardo considerable al camino de datos por lo cual se puede necesitar dispositivos de cancelación de eco, ya sea dentro del terminal de usuario o dentro de la red. Sin embargo, sería ineficiente enviar simplemente paquetes de muestras de habla más pequeños, ya que esto incrementaría el ancho de banda necesario, además de los requerimientos de enrutadores IP adicionales, para encaminar más paquetes de habla. Adicional al retardo en la transmisión extremo a extremo, también se encuentra uno al final del camino del habla, donde los paquetes tienen que pasar a través de un buffer para que el habla digitalizada se sincronice antes de que se reproduzca, y este retardo es debido al jitter de los paquetes.

Los paquetes de habla son transportados en paquetes del Protocolo de Datagramas de Usuario (User Datagram Protocol, UDP) e IP por la red GPRS y el núcleo de la red IP, entre los equipos de usuario. Para sincronizar las muestras y controlar la tasa de muestreo, se requiere de un protocolo de entramado. Los protocolos de la IETF para el entramado de la voz y de la multimedia son RTP y RTCP, los cuales son llevados en paquetes UDP/IP. Esto se muestra en la Figura A19.

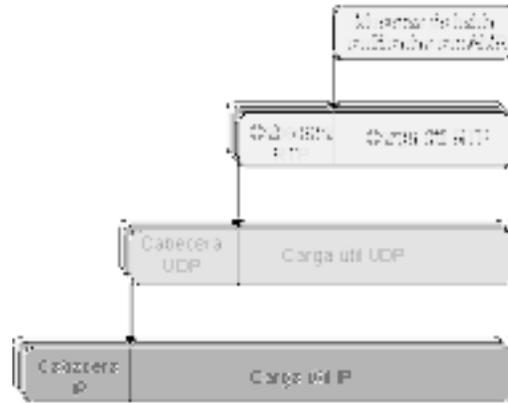


Figura A19: Transporte del habla en IP

Para que el equipo de usuario sea capaz de enviar y recibir paquetes de habla hacia y desde el dominio IMS, este debe activar de manera bidireccional el contexto del Protocolo de Paquetes de Datos (Packet Data Protocol, PDP), entre él mismo y el dominio IMS. De esta forma se podrá asignar el ancho de banda y la calidad de servicio requerida, para el transporte de los paquetes de habla, sobre la UTRAN y la red GPRS. El punto de entrada al dominio IMS va a tener cortafuegos para la seguridad y la prevención de ataques a servicios, los cuales a su vez también pueden controlarse dinámicamente por medio del control de la sesión, esto último con el fin de prevenir el envío o recepción de paquetes de habla antes de que se establezca una sesión. Mediante la desactivación del contexto PDP GPRS adecuado, se puede desconectar el camino de habla entre el equipo de usuario y el dominio IMS.

La red GPRS especifica una clase de servicio conversacional para proporcionar una calidad de servicio mayor que la del "mejor esfuerzo". En esta se priorizan los paquetes de habla por medio de los parámetros: retardo bajo y jitter bajo. De forma semejante, para proporcionar una calidad de servicio alta dentro del dominio IMS se pueden utilizar mecanismos tales como Diffserv o el Protocolo de Reservación de Recursos (Resource Reservation Protocol, RSVP).

El codec por defecto que todos los terminales de la Release 5 deben soportar es el AMR, aunque también se podrían soportar otros codecs. Debido a que la señalización del control de la sesión soporta una operación de traducción libre, la codificación de habla AMR se puede utilizar extremo a extremo entre los terminales de usuario, sin la necesidad de tener que traducir a otro estándar. Sin embargo, con una codificación del habla AMR a 12.2kbps, y un muestreo de habla de 20ms, da como resultado una carga útil de habla RTP de valor igual a la mitad del tamaño de la combinación de los encabezados de los paquetes IP, UDP y RTP. Lo anterior produce un uso del ancho de banda ineficiente, especialmente en el acceso radio que es bastante costoso. Al incrementar el tamaño de las muestras se reduce el problema, pero se incrementa el retardo de los paquetes de habla extremo a extremo, así como se incrementa la probabilidad de la pérdida de paquetes en la interfaz radio. Una solución a lo anterior es llevar a cabo la compresión de encabezados entre el equipo de usuario y la UTRAN, que es donde el ancho de banda es más costoso.

Un beneficio de la utilización de GPRS para llevar los caminos del habla, así como con los de señalización, es que GPRS controla el "handover", o la transferencia de los caminos del habla de una estación base a otra, mientras los usuarios se mueven entre las celdas de radio. Sin embargo, se requiere que se mejoren los procedimientos de "handover" de GPRS, para asegurar la calidad de servicio requerida por la voz cuando ocurre este proceso.



A3.2.2. Servicios Multimedia en IMS [52][53][54]

A medida que las arquitecturas de redes de telecomunicaciones han evolucionado, ha habido una tendencia fuerte hacia la centralización de las capacidades de procesamiento multimedia en un nodo lógico y especializado. Este se conoce genéricamente como el Servidor de Medios, y de manera equivalente IMS lo ha llamado la MRF, la cual se descompuso posteriormente en dos subnodos llamados el Controlador de la Función de Recursos Multimedia (Media Resource Function Controller, MRFC) y el Procesador de Funciones de Recursos Multimedia (Media Resource Function Processor, MRFP). El MRFC muestra una interfaz SIP a los otros componentes de la red, y proporciona las funciones de señalización y control, mientras que el MRFP proporciona las capacidades de procesamiento de medios.

Típicamente la MRF es invocada por las aplicaciones que se encuentran corriendo en el Servidor de Aplicaciones (Application Server, AS) y a través del CSCF. La interfaz que el componente MRFC, de la MRF, muestra a la red está basada en SIP. La multimedia termina y se origina desde el MRFP, a través de la interfaz Mb, además el MRFP utiliza RTP para el transporte de la multimedia a través de las redes IP.

En los despliegues modernos, los servicios a gran escala del reconocimiento del habla y de la síntesis de la misma se proporcionan por servidores de habla especializados, los cuales son controlados por un protocolo de la IETF que tiene el nombre de Protocolo de Control de los Recursos Multimedia (Media Resource Control Protocol, MRCP). El MRFC incorporará un cliente MRCP el cual se encarga del control de los servidores de habla distribuidos y el MRFP suministrará y recibirá los medios de los servidores de habla.

A continuación, se describe con más detalle los requerimientos funcionales y el flujo de información que genera los procedimientos entre el MRFC y el MRFP, además de su interacción con las entidades de la arquitectura IMS que necesitan de dichos procesos. El desarrollo de esta temática se basa en los Reportes y Especificaciones Técnicas del 3GPP que abordan los Servicios de Habla dentro de la arquitectura IMS.

A3.2.2.1. Arquitectura de la MRF

La arquitectura básica de la MRF se puede ver en la Figura A20.

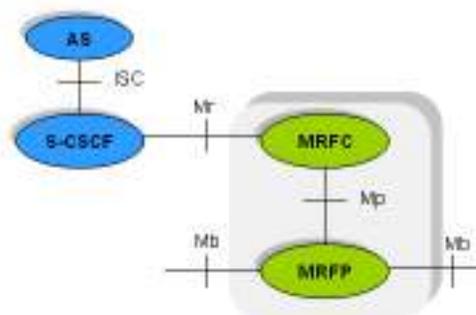


Figura A20: Arquitectura de la MRF



Las tareas del MRFC son las siguientes:

- Controlar los recursos de los flujos multimedia del MRFP.
- Interpretar la información que proviene del AS y del S-CSCF (por ejemplo el identificador de la sesión) y en consecuencia controla al MRFP.
- Generar los CDRs

El punto de referencia Mp le permite al MRFC controlar los recursos de los flujos multimedia que provee el MRFP. El punto de referencia Mp tiene las siguientes propiedades:

- Obedece completamente al estándar H.248 [21]
- Se puede llevar a cabo un trabajo de definición de extensiones sobre la interfaz para una arquitectura abierta.

Las tareas del MRFP son las siguientes:

- Controlar la entidad portadora en el punto de referencia Mb.
- Proporcionar los recursos que va a controlar el MRFC.
- Mezclar los flujos multimedia de entrada (cuando por ejemplo se manejan múltiples partes).
- Originar flujo multimedia (por ejemplo para anuncios multimedia).
- Procesar flujo multimedia (por ejemplo análisis de medios).
- Control de piso o “floor” (por ejemplo la gestión de los derechos de acceso para compartir recursos durante una conferencia).

Las tareas de un Servidor de Aplicaciones con relación a la MRF son: facturación de conferencia y gestión de la información de facturación (momento en el que empezó, duración, lista de participantes).

El protocolo que se utiliza en el punto de referencia Mr es SIP.

A3.2.2.2. Interacciones que involucran al MRFC y al MRFP

- **Interacciones entre el UE y el MRFC**

En algunos casos un operador puede querer que el MRFC este disponible directamente para el UE, por ejemplo para soportar sesiones multipartita “de igual a igual” iniciadas por el UE.

- **Interacciones entre el MRFC y el AS basadas en el control del servicio**

Basado en el control del servicio, los recursos del MRFC y el MRFP también se pueden utilizar dentro de una red IMS. En este caso es el AS el que interactúa con el MRFC, y los mensajes de control de la sesión se pasan desde este al MRFC a través del S-CSCF.

A3.2.2.3. Requerimientos funcionales para la provisión de capacidades multimedia

IMS no define las aplicaciones o servicios que pueden ofertarse al usuario final, sino la infraestructura y capacidades del servicio que los operadores o proveedores de servicio pueden emplear para construir sus propias aplicaciones y producir su oferta de servicios. En este sentido, IMS no impone límites, son la capacidad de la red de acceso y las características de los terminales las que las fijan.



Las capacidades que la MRF le proporciona a la arquitectura IMS son: reproducción de tonos, reproducción de anuncios, conversión de texto a habla, grabación de audio, detección de Tono Dual Multifrecuencia (Dual Tone Multi Frequency, DTMF), reconocimiento automático de habla, reproducción de multimedia, grabación de multimedia, conferencia de audio, conferencia multimedia, entre otras.

A continuación, se describen los requerimientos funcionales de las capacidades de conversión de texto a habla, reconocimiento automático de habla, y detección DTMF.

- **Conversión de Texto a Habla**

- El MRFC puede estar en capacidad de pedirle al MRFP que reproduzca el texto a una de varias, a múltiples o todas las partes que se encuentren conectadas en una llamada o sesión, ya sea por valor (indicando directamente el texto) o por referencia (a través de una URI).
- El MRFC puede estar en capacidad de pedirle al MRFP que reproduzca un texto en un bucle hasta que le ordene detenerse.
- El MRFC puede estar en capacidad de pedirle al MRFP que reproduzca un texto durante un número fijo de veces.
- El MRFC puede estar en capacidad de pedir detección DTMF durante la reproducción de un texto.
- El MRFC puede estar en capacidad de pedirle al MRFP que termine la reproducción de un texto cuando se detecte un dígito DTMF.
- El MRFC puede estar en capacidad de indicarle al MRFP qué tipo de lenguaje se va a utilizar para reproducir el texto (el tipo de lenguaje hace referencia al lenguaje fuente del texto, por ejemplo puede ser no necesaria una función de traducción en el MRFP).
- El MRFC puede estar en capacidad de pedirle al MRFP que indique cuando un texto se ha reproducido de manera exitosa.
- El MRFP puede estar en capacidad de indicar los casos en que ocurren errores, entre los cuales, se tiene cuando el texto no se reprodujo exitosamente.

En este apartado se realiza una anotación relacionada con la poca claridad que se tiene acerca de cómo el MRFC obtiene el texto y en qué formato se debe encontrar (tamaño, si esta segmentado para que se concatene, etc.), ya que esto depende de la definición de la interfaz Mr.

- **Reconocimiento Automático de Habla**

La función del Reconocimiento Automático de Habla consiste en que un reconocedor procese la voz de entrada del usuario y la compare con unos "datos específicos", para luego producir un resultado que represente la detección de la misma. En IMS, el MRFP actúa como el reconocedor, y este se encuentra bajo el control del MRFC.

- El MRFC está en capacidad de pedirle al MRFP que inicie el reconocimiento automático del habla.
- El MRFC puede estar en capacidad de pedir la detección DTMF cuando se está ejecutando el ASR.
- El MRFC puede estar en capacidad de pedirle al MRFP que detenga el ASR en el caso en que se detecte un dígito DTMF predefinido.
- El MRFC puede estar en capacidad de indicarle al MRFP el tipo del lenguaje del suscriptor.
- El MRFC puede estar en capacidad de indicarle al MRFP los datos específicos con los cuales se va a comparar la voz de entrada.



- El MRFC puede estar en capacidad de pedirle al MRFP que le indique cuando un proceso de ASR específico se halla culminado exitosamente.
- Cuando un proceso de ASR se culmina exitosamente el MRFP puede notificarle al MRFC el nivel de confidencialidad.
- El MRFP estará en capacidad de indicar los casos de error en los cuales el proceso de ASR no se ejecutó exitosamente.

En este apartado se realizan las siguientes anotaciones:

- Se encuentra para estudio futuro la configuración del dígito DTMF predefinido, ya que puede ser dentro del AS, MRFC o del MRFP.
- La definición de los “datos específicos” es para estudios futuros.
- Se necesita definir el nivel de confidencialidad.

- **Detección DTMF**

- El MRFC debe estar en capacidad de pedirle al MRFP que detecte y reporte los dígitos DTMF.

En este apartado se realiza la anotación de la necesidad de revisar si se requiere detección de tonos DTMF dentro de banda, adicionalmente al informe de los Eventos de Telefonía DTMF RTP. Más adelante, cuando se reporten los eventos de telefonía DTMF, se necesita revisar si se debe reportar su recepción (para qué propósito, donde se valida el evento de telefonía DTMF, duración de un tono DTMF válido).

A3.2.2.4. Procedimientos del MRFC y MRFP en la provisión de ASR, TTS, y detección DTMF

En la especificación técnica donde se trata la descripción de los procedimientos del MRFC y el MRFP para la provisión de capacidades multimedia, aún están por definirse los procedimientos para la conversión de texto a habla, el reconocimiento automático de habla, y la detección DTMF.

A3.2.2.5. Requerimientos de la arquitectura IMS para los servicios de reconocimiento del habla

Los servicios de habla generan requerimientos adicionales a la arquitectura IMS, ya que se deben aplicar una serie de técnicas para mejorar el desempeño en lo que respecta al reconocimiento del habla. Con base en un Reporte Técnico del 3GPP, en esta sección se describe la forma como debe mejorarse el sistema 3GPP existente para proporcionar la señalización que controle estas técnicas.

A3.2.2.5.1. Técnicas para mejorar el desempeño del reconocimiento de habla

Las plataformas de ASR tienden a realizar la siguiente secuencia de operaciones: cancelación de eco, extracción e interpretación. El proceso de cancelación de eco se utiliza para permitir el envío de comandos a la plataforma ASR mientras ésta se encuentra reproduciendo anuncios de voz. (Sin la cancelación de eco la plataforma ASR no podría distinguir entre su voz y la voz del usuario). En el proceso de extracción, se extraen los fonemas de la señal de habla de entrada y se transforman en palabras mediante los modelos acústicos. Luego, el motor de reconocimiento de habla realiza una búsqueda de las palabras pronunciadas en la gramática creada por el Lenguaje Gramatical



Específico (Grammar Specific Language, GSL). El proceso de interpretación se utiliza para extraer la interpretación semántica de la secuencia de palabras.

- **Supresión adicional de ruido en el terminal**

Entre los factores que afectan el desempeño del ASR se tienen: calidad de la señal, tamaño de la gramática, confusión acústica, etc. Los desafíos que tiene el reconocimiento del habla incluyen: el ruido de fondo, el equipo de habla, los equipos manos libres, y la pérdida de la calidad de la señal debido a la compresión. Una gran cantidad de aplicaciones inalámbricas son capaces de ejecutarse en ambientes ruidosos y parece ser que al reducir el impacto del ruido de fondo se puede mejorar el desempeño del ASR.

El rendimiento del ASR se puede mejorar si este le pide al Equipo de Usuario (User Equipment, UE) que adicione supresión extra de ruido.

- **Deshabilitar la transmisión no continua para mejorar el desempeño del ASR**

En un caso normal, el terminal utiliza la transmisión no continua para reducir las transmisiones en el enlace de subida. Las tramas descriptoras de silencio se envían a una velocidad baja y la Unidad de Adaptación de Tasa y Traducción (Transcoder & Rate Adaptation Unit, TRAU) las utiliza para enviar, en el enlace de subida, ruido aceptable para mejorar la calidad normal del habla.

Una consecuencia de esto es que durante una parte pequeña de la conversación el codec de habla codifica el ruido de fondo durante los periodos de silencio. Esto generará problemas en la plataforma ASR debido a que el codec de habla utilizará los periodos de silencio para estimar el ruido de fondo.

Si la unidad ASR puede ordenarle al dispositivo móvil deshabilitar la Transmisión No Continua (Discontinuous Transmission, DTX), entonces, el codec de voz siempre se utilizaría para codificar el ruido de fondo, y de esta forma podría permitirse que la unidad ASR realice una estimación más exacta del ruido.

El grado de mejoría en el desempeño del reconocimiento de la voz tendrá que justificar el flujo de potencia adicional y el incremento de la interferencia debido a que no se habilite el DTX.

- **Indicar la utilización o no del DTX**

Es posible que la unidad ASR encuentre como beneficioso el hecho de conocer cuando el DTX esta siendo o no utilizado. Esta información podría ser llevada desde la TRAU hasta la unidad ASR por medio de extensiones de la señalización Operación Libre de Tandem (Tandem Free Operation, TFO).

- **Indicar a la plataforma ASR el modo del codec**



El desempeño del reconocimiento de habla puede ser influenciado por el tipo de codec que se utilice para la compresión en entornos inalámbricos ruidosos. Por lo tanto, resultaría útil proveer a la plataforma ASR una indicación del codec utilizado.

- **Ordenar que se proporcione una codificación de alta velocidad para el modo de reconocimiento de habla**

El desempeño en el reconocimiento del ASR puede disminuir significativamente por tasas de codificación por debajo de 8-10Kbps. Por lo tanto, es necesario resaltar que durante las sesiones ASR no se considere la optimización de la capacidad.

En realidad, la velocidad actual que se utiliza depende de las condiciones del canal de radio, y del grado de error y de protección que se requieren. Lo anterior maximizará la calidad actual del habla que se entrega al reconocedor.

- **Para el modo de banda ancha AMR se necesita la señalización TFO en la plataforma ASR**

Para mejorar la calidad del habla en llamadas de voz normales el AMR de banda ancha se puede implementar en el terminal. Para llevar la señal de banda ancha a la plataforma ASR es necesario, por ejemplo, algo como el TFO.

- **Acentos regionales y las plataformas ASR**

Las personas de diferentes regiones hablan con diferentes acentos por lo cual pueden sufrir un peor desempeño del ASR. Potencialmente, el HLR (o plataformas CAMEL) podría proporcionar al Centro Móvil de Conmutación Visitado (Visited Mobile Switching Centre, VMSC) la "información del acento regional", la cual a su vez se pasa a la plataforma ASR para permitir que sus unidades utilicen una gramática específica de acuerdo al acento de una región.

A3.2.2.5.2. Impactos en la arquitectura IMS por las mejoras en el reconocimiento de habla

La utilización de un codec de Reconocimiento de voz distribuido (Distributed Speech Recognition, DSR) para el reconocimiento del habla puede ocasionar algunos impactos en IMS, los cuales incluyen:

a) Imaginarse el caso en que la plataforma ASR se encuentra ejecutando una aplicación de marcado activado por voz. El móvil necesita utilizar el codec DSR cuando hable con la plataforma ASR, y tendrá que cambiar al codec AMR cuando quiera conectarse con un usuario. El tema importante en este escenario, es la forma como la señalización IMS va a hacer que el codec se cambie entre dichos codecs.

b) Cuando el móvil está utilizando el codec DSR con la plataforma ASR, en el enlace de bajada se tiene que utilizar el codec AMR. Por lo tanto, los parámetros SDP que se incluyen por el móvil, tienen que reflejar estos flujos unidireccionales de medios. Además, es muy probable que esto tenga impacto en muchas de las implementaciones de SIP que esperaban utilizar el mismo codec en ambas direcciones.

El primero de estos temas también aplica si se utiliza el codec AMR para Banda Ancha (Wide Band - Adaptive Multi Rate, WB-AMR) para el reconocimiento de habla y el otro usuario soporta AMR



para Banda Estrecha (Narrow Band – Adaptive Multi Rate, NB-AMR) solamente. La utilización del codec NB-AMR para el reconocimiento del habla parece evitar este asunto.

Un servidor ASR puede actuar como un agente de usuario SIP “back-to-back” para la conexión de un usuario con otro, o puede actuar como un agente que transfiere la llamada de un usuario. Así sea que la plataforma ASR este en modo de conexión a un usuario o en el modo de reenvío, el SDP tiene que soportar la situación de que el enlace de subida utilice un tipo de codec y el enlace de bajada utilice otro tipo de codec.

La manera de modelar un canal asimétrico en SDP es mediante la utilización de los atributos *recvonly* y *sendonly*, lo cual le permite a SIP soportar tanto la simetría como la asimetría en los canales.

Un ejemplo que indique en el SDP de un mensaje INVITE un enlace de subida con DSR y un enlace de bajada con AMR es el siguiente:

```
v=0
o=SESPhone 0 1 IN IP4 10.132.30.33
s=asymmetric ses dsr session
c=IN IP4 10.132.30.33
t=0 0
m=audio 9002 RTP/AVP 97
a=rtpmap:97 AMR/8000
a=recvonly
m=audio 9002 RTP/AVP 101
a=rtpmap:101 dsr-es202212/8000
a=sendonly
```

El anterior ejemplo utiliza IPv4 pero se puede extender fácilmente a IPv6.

A3.2.2.5.3. Tarificación en IMS por la funcionalidad ASR

En los casos de llamadas persona a máquina, o persona a persona a través del servidor ASR, el S-CSCF y/o el servidor ASR deberían generar los CDRs, ya sea que se realice la tarificación por eventos, duración o volumen, etc.

Pero si se trata al servidor ASR como un Servidor de Aplicaciones IMS, como el MRF, se pueden reutilizar los mecanismos de tarificación de IMS existentes.

A3.2.2.5.4. Seguridad en IMS para la funcionalidad ASR

En el dominio IMS, el ASR utiliza las identidades pública y/o privada, que se encuentran en los mensajes SIP, para verificar con una base de datos de administración si el usuario tiene acceso permitido al servicio. Cuando se habla de la seguridad a nivel de aplicación, ya se entra a tratar con temas como la verificación de los credenciales del usuario, reconocimiento del hablante, o la utilización de la tecnología de autenticación biométrica.

A3.2.3. Servicios IVR en IMS que se adapten a las redes 2.5G actuales



Actualmente, la tecnología de VoIP tiene un grado de desarrollo mucho mayor en las operadoras fijas que en las móviles. Sin embargo, en el camino de evolución hacia lo que se denomina “todo IP” las operadoras móviles han dirigido parte de sus esfuerzos hacia esta tecnología, y un reflejo de ello ha sido la introducción de IMS a los sistemas UMTS, utilizando SIP como protocolo de señalización para el control de las sesiones. El camino de especificación e implementación de las redes 3G de manera completa es aún largo, sin embargo el proceso de estandarización contribuye a la superación de barreras tecnológicas que se presentan debido a la utilización de tecnologías como la VoIP en entornos móviles, donde las principales limitantes son la capacidad y costo actuales del acceso radio, y la falta de garantías para la prestación de servicios con calidad.

El servicio clásico de voz impone requisitos estrictos de temporización, y por lo dicho anteriormente, las características y capacidades actuales de la interfaz aire requieren de una mayor adecuación técnica, que se encuentra en proceso de especificación y estandarización, para llegar a ser un medio de transmisión de datos con buenas prestaciones. Sin embargo, existen aplicaciones de voz empaquetada con requisitos de tiempo menos restrictivos, y cuya puesta en servicio se ha considerado factible en un menor plazo que las llamadas tradicionales con tecnología VoIP. Estos son los casos de los servicios persona a persona donde la voz se transmite a ráfagas a través de un canal semidúplex, o servicios PoC, y los servicios persona a contenido, donde los requerimientos técnicos de acceso son inclusive menores que los primeros ya que la transmisión en tiempo real de la información no es un aspecto obligatorio, y la gestión de la sesión es menos compleja.

Ambos casos permiten un mayor ahorro de recursos de red y admiten una mayor tolerancia al retardo para proporcionar una percepción aceptable por parte del usuario. Esto permitiría su despliegue sobre la red de datos móvil de forma efectiva, y les permitirá a los operadores introducir y experimentar de forma segura con la voz sobre IP, en un entorno real, a través de un nuevo servicio, sin sustituir ni experimentar con el servicio de llamadas.

Los sistemas IVR proporcionan diversas capacidades de interacción con los usuarios, y dicha interacción está regida por el tipo de servicio que se ofrezca. Los dos tipos de servicios nombrados anteriormente y las diversas aplicaciones que de ellos se desprenden pueden ser proveídas y soportadas por sistemas IVR, sin embargo, por los alcances y objetivos del presente Trabajo de Grado se van a seleccionar los servicios persona a contenido, los cuales permitirán demostrar de una manera básica las funcionalidades ofrecidas por los sistemas ASR y TTS al enmarcarse dentro de la arquitectura IMS.

Entre los servicios persona a contenido, a través de IVR, se pueden encontrar: reservaciones en el área del turismo, servicios de información geográfica y meteorológica, servicios de ayuda a usuarios por parte de empresas, consulta de información personal de usuarios como por ejemplo su correo electrónico o su disponibilidad bancaria, envío de mensajes cortos, etc.



A4. GLOSARIO

1G	Primera Generación de sistemas de comunicaciones móviles
2G	Segunda Generación de sistemas de comunicaciones móviles
2G	2.5 Generación de sistemas de comunicaciones móviles
3G	Tercera generación de sistemas de comunicaciones móviles
3GPP	Third Generation Partnership Project, Proyecto Conjunto de Tercera Generación
3GPP2	Third Generation Partnership Project 2, Proyecto Conjunto de Tercera Generación 2
AAA	Authentication, Authorization and Accounting, Autenticación, Autorización y Tasación
ADSL	Asymmetric Digital Subscriber Line, Línea de Abonado Digital Asimétrica
ADPCM	Adaptive Differential Pulse Code Modulation, Modulación por Impulsos Codificados Diferencial y Adaptativa
AMR	Adaptive Multi Rate, Codificador de voz de multi-tasa adaptativa. Codifica y decodifica las muestras de habla digitalmente.
ANSI	American National Standard Institute, Instituto Nacional de Estándares Americanos.
Applet	Componente de software que corre en el contexto de otro programa, por ejemplo un navegador Web
API	Application Programming Interface, Interfaz de Programación de Aplicaciones
AS	Application Server, Servidor de Aplicaciones
ASR	Automatic Speech Recognition, Reconocimiento Automático de Habla
ATM	Asynchronous Transfer Mode, Modo de Transferencia Asíncrono
ASSI	Application Server Subscription Information, Información de Suscripción en el Servidor de Aplicaciones
Back-to-back	Una entidad que puede tener dos roles como el de cliente y el de servidor
BGCF	Breakout Gateway Control Function, Función de control de la Pasarela de Fuga
BSS	Base Station Subsystem, Subsistema de la Estación Base
CAMEL	Customized Applications for Mobile Networks Enhanced Logic, Aplicaciones Personalizadas para Redes Móviles de Lógica Mejorada
CAP	CAMEL Application Part, Parte de la Aplicación CAMEL
CDMA2000	Code Division Multiple Access 2000, una tecnología 3G
CDR	Call Detailed Record, Registro Detallado de la Llamada
CELP	Codificación que se basa en procedimientos de búsqueda de análisis-por-síntesis, cuantización de vectores con pesos y predicción lineal.
CGI	Common Gateway Interface, Interfaz Común de Pasarela
CLI	Command Line Interface, Interfaz por Línea de Comandos
CN	Core Network, Núcleo de Red
Codec	Dispositivo codificador-decodificador
COPS	Common Open Policy Service, Servicio de Políticas Común y Abierto
CPL	Call Processing Language, Lenguaje de Procesamiento de Llamada
CS	Circuit Switching, Conmutación de Circuitos
CSCF	Call Session Control Function, Función de Control de la Sesión de Llamada
Cx	Punto de referencia entre HSS y CSCF
DiffServ	Differentiated Services Internet QoS model, Modelo de Calidad de Servicio en Internet basado en Servicios Diferenciados
Dominio IM	Dominio Multimedia IP
DSR	Distributed Speech Recognition, Reconocimiento de voz distribuido
DTMF	Dual Tone Multi Frequency, Tono Dual Multifrecuencia
DTX	Discontinuous Transmisión, Transmisión no Continúa



Dx	Punto de referencia entre CSCF y SLF
E.164	Recomendación que describe el plan internacional de numeración telefónica
EDGE	Enhanced Data Rates for Global Evolution, Tasas de Datos Mejorados para la Evolución Global
FQDN	Fully-Qualified Domain Name, Nombre de Dominio Totalmente Cualificado
G.711	Técnica de codificación de voz PCM de 64Kbps
GERAN	GPRS EDGE Radio Access, Acceso Radio GPRS EDGE
GGSN	Gateway GPRS Support Node, Nodo de Soporte de la Pasarela GPRS
Go	Punto de referencia entre GGSN y PDF
GPRS	General Packet Radio Service, Servicio General de Paquetes de Radio
GPS	Global Positioning System, Sistema de Posicionamiento Global
GSL	Grammar Specific Language, Lenguaje Gramatical Específico
GSM	Global System for Mobile Communications, Sistema Global para Comunicaciones Móviles
H.248	También conocido como protocolo Megaco
H.323	Estándar de la ITU-T para voz y videoconferencia interactiva en tiempo real, para redes de Área Local e Internet
HLR	Home Location Register, Registro Local de Ubicación
HSS	Home Subscriber Server, Servidor Local del Suscriptor
HTTP	Hypertext Transfer Protocol, Protocolo de Transferencia de Hipertexto
I-CSCF	CSCF de Interrogación o “Interrogating”
IETF	Internet Engineering Task Force, Grupo de Trabajo en Ingeniería de Internet
IM	Instant Messaging, Mensajería Instantánea
IM-SSF	IP Multimedia - Services Switching Function, Función de Conmutación de Servicios - Multimedia IP
IM-SSP	IP Multimedia – Service Switching Point, Punto de Conmutación de Servicios - Multimedia IP
IMS	IP Multimedia Subsystem, Subsistema Multimedia IP
IMS-AKA	IMS Authentication and Key Agreement, Clave de Acuerdo y Autenticación IMS
IMS-ALG	IMS Application Level Gateway, Pasarela a Nivel de Aplicación IMS
IM-MGW	IP Multi-Media Gateway, Pasarela de Medios Multimedia IP
INAP	Intelligent Network Application Protocol, Protocolo de Aplicación de Red Inteligente
IP	Internet Protocol, Protocolo Internet
IPv4	Versión 4 del Protocolo Internet
IPv6	Versión 6 del Protocolo Internet
ISC	IP multimedia Service Control, Control del Servicio Multimedia IP
ISDN	Integrated Services Data Network, Red Digital de Servicios Integrados
ISIM	Aplicación lógica que reside en una tarjeta de circuitos integrados
ISUP	ISDN User Part, Parte de Usuario ISDN
IT	Information Technology, Tecnología de la Información
IVR	Interactive Voice Response, Respuesta de Voz Interactiva
JAIN	Java APIs for Integrated Networks, APIs de Java para Redes Integradas
JVM	Java Virtual Machine, Máquina Virtual de Java
LDAP	Lightweight Directory Access Protocol, Protocolo Ligero de Acceso a Directorio
LIA	Location Information Answer, Respuesta de Información de Localización
LIR	Location Information Request, Petición de Información de Ubicación
LNP	Local Number Portability, Portabilidad de Número Local
MAP	Mobile Application Part, Parte de Aplicación Móvil
Mb	Punto de referencia para servicios en redes IPv6
MEGACO	Media Gateway Control, Control de Interfaces de Medios
MG	Media Gateway, Pasarela de Medios
MGC	Media Gateway Control, Control de la Pasarela de Medios



MGCF	Media Gateway Control Function, Función de Control de la Pasarela de Medios
MGW	Media Gateway, Pasarela de Medios
MIC	Modulación por Impulsos Codificados
MIME	Multipurpose Internet Mail Extensions, Extensiones Multipropósito del Correo Electrónico de Internet
MIMO	Multiple Input Multiple Output, Múltiples Entradas Múltiples Salidas
MMUSIC	Multiparty Multimedia Session Control, Control de Sesión Multimedia Multipartita
MMD	Multimedia Domain, Dominio Multimedia
Mp	Punto de referencia entre MRFC y MRFP
MP-MLQ	Modo de compresión de VoIP
Mr	Punto de referencia entre CSCF y MRFC
MRCP	Media Resource Control Protocol, Protocolo de Control de los Recursos de Multimedia
MRF	Media Resource Function, Función de Recursos Multimedia
MRFC	Media Resource Function Controller, Controlador de la Función de Recursos Multimedia
MRFP	Media Resource Function Processor, Procesador de la Función de Recursos Multimedia
MSC	Mobile Switching Center, Centro de Conmutación Móvil
MSFC	Media Server Function Control, Función de Control del Servidor de Medios
Multicast	Multidifusión
NB-AMR	Narrow Band – Adaptive Multi Rate, AMR para Banda Estrecha
NGN	Next Generation Network, Red de Próxima Generación
OSA	Open Services Architecture, Arquitectura de Servicios Abiertos
OSA-GW	Open Services Architecture - Gateway, Pasarela de la Arquitectura de Servicios Abiertos
PCC	Policy and Charging Convergente, Convergencia de las Políticas y la Facturación
PCM	Pulse Code Modulation, Modulación por Impulsos Codificados
PDF	Policy Decision Function, Función de Decisión de Política
PDP	Packet Data Protocol, Protocolo de Paquetes de Datos
Perl	Practical Extraction and Reporting Language, Lenguaje de Reporte y Extracción Práctica
P-CSCF	Proxy-CSCF, Proxy Controlador de la Función de Sesión de Llamada
Plug-in	Es un módulo opcional que puede ser agregado a un programa
PTT	Push To Talk, Presionar Para Hablar
PS	Packet Switching, Conmutación de Paquetes
PoC	PTT para celulares
PSTN	Public Switched Telephone Network, Red Telefónica Pública Conmutada
QoS	Quality of Service, Calidad de Servicio
Radius	Remote Authentication Dialing User Service, Protocolo de Autenticación y autorización para aplicaciones de acceso a la red
RAN	Radio Access Network, Red de Acceso Radio
RDSI	Red Digital de Servicios Integrados
Release	Versión o Edición
RFC	Request For Comments, Petición de Comentarios. Conjunto de notas técnicas y organizativas donde se describen los estándares o recomendaciones de Internet
RNC	Radio Network Controller, Controlador de la Red Radio
RSVP	Resource Reservation Protocol, Protocolo de Reservación de Recursos
RTC	Red Telefónica Conmutada
RTP	Realtime Transport Protocol, Protocolo de Transporte en Tiempo Real
RTCP	Real Time Control Protocol, Protocolo de Control en Tiempo Real



RTSP	Real-Time Streaming Protocol, Protocolo de Transmisión Continua en Tiempo Real
Script	Tipo de programa que consiste de una serie de instrucciones que serán utilizadas por otra aplicación
SCS	Service Capability Server, Servidor de Capacidad de Servicio
S-CSCF	CSCF de Servicio o "Serving"
SCTP	Stream Control Transmission Protocol, Protocolo de Control del Flujo de la Transmisión
SDP	Session Describing Protocol, Protocolo de Descripción de Sesión
SGSN	Server GPRS Support Node, Nodo de Soporte del Servidor GPRS
SGW	Signaling Gateway, Pasarela de Señalización
SIGTRAN	Signalling Transport, Protocolo de Transporte de Señalización
SIM	Aplicación lógica que reside en una tarjeta de circuitos integrados
SIMPLE	SIP for Instant Messaging and Presence Leveraging Extensions, SIP para Mensajería Instantánea y Extensiones para el servicio de Presencia
SIP	Session Initiation Protocol, Protocolo de Inicio de Sesión
SIPPING	Session Initiation Proposal Investigation
SLF	Subscription Locator Function, Función de Ubicación del Suscriptor
SNMP	Simple Network Management Protocol, Protocolo Simple de Administración de Red
SS7	Signaling System 7, Sistema de Señalización No.7
SSF	Service Switching Function, Función de Conmutación de Servicio
TAS	Telephony Application Server, Servidor de Aplicaciones de Telefonía
TCAP	Transaction Capabilities Application Part, Parte de Aplicación de las Capacidades de Transacción
TCP	Transmission Control Protocol, Protocolo de Control de la Transmisión
TDM	Time Division Multiplexing, Multiplexación por División de Tiempo
TFO	Tandem Free Operation, Operación Libre de Tandem
THIG	Topology Hiding Inter-network Gateway, Pasarela de Interoperación para el Ocultamiento de la Topología
TISPAN	TISPAN es el grupo de ETSI dedicado a la estandarización de servicios y redes fijas
TRAU	Transcoder & Rate Adaptation Unit, Unidad de Adaptación de Tasa y Traducción
T-SGW	Trunking Signalling Gateway, Pasarela de Señalización Troncal
TTS	Text To Speech, Conversión de Texto a Habla
UAC	User Agent Client, Agente de Usuario Cliente
UAS	User Agent Server, Agente de Usuario Servidor
UDP	User Datagram Protocol, Protocolo de Datagramas de Usuario
UE	User Equipment, Equipo de Usuario
UIT	Unión Internacional de Telecomunicaciones
UMTS	Universal Mobile Telecommunications System, Sistema de Telecomunicaciones Móviles Universales
Unicast	Unidifusión
URI	Uniform Resource Identifier, Identificador Uniforme de Recursos
USIM	Aplicación lógica que reside en una tarjeta de circuitos integrados
UTRAN	UMTS Terrestrial Radio Access Network, Red Terrestre de Acceso Radio UMTS
VGCS	Voice Group Call Services, Servicios de Llamada de Voz Grupales
VHE	Virtual Home Environment, Entorno de Origen Virtual
VLR	Visitor Location Register, Base de datos temporal para los usuarios Visitantes
VMSC	Visited Mobile Switching Centre, Centro Móvil de Conmutación Visitado
VoD	Video On Demand, Video Bajo Demanda
VoIP	Voice over IP, Voz sobre IP
VPN	Virtual Private Network, Red Privada Virtual
WB-AMR	Wide Band - Adaptive Multi Rate, AMR para Banda Ancha



WCDMA	Wideband CDMA, CDMA de Banda Ancha
WIMAX	Worldwide Interoperability for Microwave Access, Interoperabilidad Mundial para Acceso por Microondas
WLAN	Wireless Local Area Network, Red Inalámbrica de Área Local
XML	eXtensible Markup Language, Lenguaje de Mercado Extensible



A5. BIBLIOGRAFÍA

- [1] **Luis Miralles Pechuán.** “*El nuevo sistema multimedia conocido como IMS que adoptarán las redes UMTS*”. [En línea]. Trabajo presentado en la asignatura Redes de Ordenadores, Curso 2004-2005, Universidad de Valencia. Disponible en web:
<<http://www.uv.es/~montanan/redes/trabajos/IMS.doc>> [Consulta: Octubre de 2005]
- [2] **Manuel Moreno Martín, Manuel Álvarez-Campana Fernández-Corredor, Joan Vinyes Sanz.** “*Propuesta de utilización de SIP como protocolo de señalización en la red de acceso radio de sistemas UMTS*”. [En línea]. Disponible en web:
<<http://www.ahciet.net/comun/portales/1000/10002/10007/10299/docs/011-72-80.pdf>> [Consulta: Abril de 2006]
- [3] **Anselmo Carrascosa.** “*Nuevos Servicios Multimedia, Convergencia y Evolución con IMS*”. [En línea]. Agosto-Septiembre 2005. Disponible en web:
<<http://www.coit.es/publicaciones/bit/bit152/84-87.pdf>> [Consulta: Abril de 2006]
- [4] **Simon ZNATY.** “*NGN: Red de Nueva Generación*”. Agosto 2006. Disponible en: <<http://www.efort.com>> [Consulta: Agosto de 2006]
- [5] **Expocomm Argentina 2005.** *Seminario de Tecnología y Mercado*. [En línea]. Septiembre 2005. Disponible en web: <http://www.ejkreed.com/extras/img/web_727_brochurestm.pdf> [Consulta: Abril 14 de 2005]
- [6] **3G Americas.** “*Gateway IP Multimedia Subsystem (IMS) para nuevos e innovadores servicios*”. [En línea]. Julio 30 de 2004. Disponible en web:
<http://www.3gamericas.org/Spanish/News_Room/DisplayPressRelease.cfm?id=886&s=SPN> [Consulta: Enero de 2006]
- [7] **Teléfono I+D.** “*Las Telecomunicaciones y la Movilidad en la Sociedad de la Información*”. [En línea]. Febrero 2005. Disponible en web:
<<http://www.telefonica.es/sociedaddelainformacion/pdf/publicaciones/movilidad/telecoymovilidad.pdf>> [Consulta: Abril de 2006]
- [8] **Flávio Tonioli Mariotto.** “*Tendencias en Convergencia Móvil*”. [En línea]. Disponible en web:
<<http://www.comunidadmovil.com.co/media/Flavio%20Mariotto2.pdf>> [Consulta: Abril de 2006]
- [9] **Elena Romero.** “*Evolución de las infraestructuras hacia un nuevo entorno. El estándar IMS como elemento clave*”. Customer Solutions & System Integration, Ericsson. [En línea]. Mayo 2005. Disponible en web:
<http://www.ericsson.com/es/documentos/estandar_ims.pdf> [Consulta: Septiembre de 2005]
- [10] **Lucent Technologies.** “*IP Multimedia Subsystem (IMS) Service Architecture*”. [En línea]. Febrero 2005. Disponible en web:
<http://www.lucent.com/livelink/090094038005df2f_White_paper.pdf> [Consulta: Septiembre de 2005]
- [11] **Javier González Gallego.** “*El Release 6 de UMTS*”. [En línea]. Febrero-Marzo 2005. Disponible en web: <<http://www.coit.es/publicaciones/bit/bit149/56-58.pdf>> [Consulta: Abril de 2006]



- [12] **3GNewsroom**. “*UMTS ready to meet demands for high speed wireless data services*”. [En línea]. Julio 2006. Disponible en web: <http://www.3gnewsroom.com/3g_news/jul_06/news_7095.shtml> [Consulta: Agosto de 2006]
- [13] **María Victoria de Diego Bartolomé, Diego Gallego Pérez, José Antonio López Mora, Alberto Gómez Vicente**. “*UMTS: hacia una red todo IP*”. [En línea]. Enero 2002. Disponible en web: <<http://bibliotecavirtual.clacso.org.ar/ar/libros/peru/grade/arregui/art4.pdf>> [Consulta: Abril de 2006]
- [14] **SIP**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc3261.txt>>
- [15] **IPv6**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc791.txt>>
- [16] **RTP**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc1889.txt>>
- [17] **RTCP**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc3605.txt>>
- [18] **COPS**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc2748.txt>>
- [19] **RSVP**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc2205.txt>>
- [20] **DiffServ**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc2475.txt>>
- [21] **MEGACO**. [En línea]. Disponible en web: <<http://tools.ietf.org/html/rfc3525>>
- [22] **SIMPLE**. [En línea]. Disponible en web: <<http://www.rfc-editor.org/rfc/rfc3428.txt>>
- [23] **Diego Andrés Acosta**. “*SIP Session Initiation Protocol*”. [En línea]. Julio 2002. Disponible en web: <<http://greco.dit.upm.es/~david/TAR/trabajos2002/01-SIP-%20Diego-Acosta.pdf>> [Consulta: Abril de 2006]
- [24] **Douglas Howie, Mika Ylianttila, Erkki Harjula, Jaakko Sauvola**. “*STATE-OF-THE-ART SIP FOR MOBILE APPLICATION SUPERNETWORKING*”. [En línea]. Disponible en web: <<http://www.mediateam oulu.fi/publications/pdf/548.pdf#search=%22%22SIP%20plugin%22%20%2B%20j2me%22>> [Consulta: Septiembre de 2006]
- [25] **Tech-invite**. “*ABNF for URI Generic Syntax – RFC 3986*”. [En línea]. Octubre 2005. Disponible en web: <<http://www.tech-invite.com/Ti-uri-abnf.html>> [Consulta: Octubre de 2006]
- [26] **CIENTEC**. “*SIP: hacia la era de las comunicaciones interactivas*”. [En línea]. Disponible en web: <<http://www.cientec.com/tendencias/tendencias28.asp>> [Consulta: Septiembre de 2005]
- [27] **Mariano Stokle, NORTEL**. “*This is the way (SIP Tutorial V3)*”. Disponible en pdf: <Session Initiation Protocol (SIP) Tutorial - Slides.pdf> [Consulta: Septiembre de 2005]
- [28] **Unitronics comunicaciones**. “*Arquitecturas en Telefonía IP y Factores de convergencia Voz/Datos*”. [En línea]. Disponible en web: <<http://www.comunicaciones.unitronics.es/tecnologia/TelefoniaIP.htm>> [Consulta: Octubre 19 de 2005]
- [29] **Inmaculada Espigares del Pozo**. “*An Implementation of the Internet Call Waiting Service using SIP*”. [En línea]. Helsinki University Of Technology, Laboratory of Telecommunications Technology. Diciembre 1999. Disponible en web: <http://www.cs.columbia.edu/sip/articles/Espi9912_Implementation.pdf> [Consulta: Octubre de 2006]



[30] **Tech-invite**. “*ABNF Grammar for SDP – Session Description Protocol (RFC 4566)*”. [En línea]. Septiembre 2006. Disponible en web: <<http://www.tech-invite.com/Ti-sdp-abnf.html>> [Consulta: Octubre de 2006]

[31] **M D Cookson, D G Smith**. “*3G service control*”. [En línea]. Enero 2001. Disponible en web: <[http://www.sipcenter.com/sip.nsf/html/WEBB5YH724/\\$FILE/3G_Service_Ctl.pdf](http://www.sipcenter.com/sip.nsf/html/WEBB5YH724/$FILE/3G_Service_Ctl.pdf)> [Consulta: Abril de 2006]

[32] **RADVISION**. “*IMS SIP and Signaling. The RADVISION Perspective*” A Technology Overview. White Paper. [En línea]. 2006. Disponible en web: <<http://www.radvision.com/NR/rdonlyres/FC60D840-1FE5-4F82-A6A2-088D2D4AADCB/0/IMSSIPWhitePaper.pdf>> [Consulta: Agosto de 2006]

[33] **RFC 3455**. “*Private Header (P-Header) Extensions to the Session Initiation Protocol (SIP) for the 3rd-Generation Partnership Project (3GPP)*”. [En línea]. Enero 2003. Disponible en web: <<http://www.rfc-archive.org/getrfc.php?rfc=3455>> [Consulta: Agosto de 2006]

[34] **Telefónica**. “*Características técnicas de las interfaces de Telefónica España. Interfaz para la conexión de terminales a los servicios de voz sobre IP*”. [En línea]. Noviembre 2005. Disponible en web: <http://www.recursosvoip.com/docs/spanish/ITE_BA_009_v1.pdf> [Consulta: Agosto de 2006]

[35] **3GPP, TSG-CN**. “*Missing statements regarding P-Charging-Function-Addresses*”. [En línea]. Marzo 2004. Disponible en web: <http://www.3gpp.org/ftp/tsg_cn/tsg_cn/TSGN_23/Docs/PDF/NP-040039.pdf> [Consulta: Agosto de 2006]

[36] **3GPP, TR 45**. “*All-IP Core Network Multimedia Domain. IP Multimedia Subsystem – Charging Architecture*”. [En línea]. Disponible en web: <http://ftp.tiaonline.org/tr-45/tr452/Public/PN-3-4935.007_X.P0013.007_IMS_Charging_Architecture/PN-3-4935.07%20v033%20IMS%20Charg%20Arch%20Post%20Ballot.doc> [Consulta: Agosto de 2006]

[37] **3GPP, TS 32.200**. “*3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Telecommunication management; Charging management; Charging principles. Release 5*”. [En línea]. Septiembre 2005. Disponible en web: <http://www.3gpp.org/ftp/Specs/archive/32_series/32.200/32200-590.zip> [Consulta: Agosto de 2006]

[38] **Günther Horn, Dirk Kröselberg, Klaus Müller**. “*IMS authentication and integrity/confidentiality protection*”. Siemens AG, Corporate Technology, Competence Centre Security. [En línea]. Noviembre 2000. Disponible en web: <http://www.3gpp.org/ftp/tsg_sa/WG3_Security/TSGS3_15bis_Munich/Docs/PDF/S3z000035.pdf> [Consulta: Abril de 2006]

[39] **3GPP**. “*3GPP TS 33.102 V7.0.0 (2005-12) 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security architecture (Release 7)*”. [En línea]. Diciembre 2005. Disponible en web: <http://www.arib.or.jp/IMT-2000/V600Dec06/2_T63/ARIB-STD-T63/Rel7/33/A33102-700.pdf> [Consulta: Abril de 2006]

[40] **Dulude Louise**. “*Automated telephone answering systems and aging*”. Conference Papers – American Association for Public Opinion Research. Mayo 2002. Vol. 21 Issue 3, p171-184. [Consulta: Abril de 2006]

[41] **Singer Eleanor**. “*Does Voice Matter? An Interactive Voice Response (IVR) Experiment*”. Behaviour & Information Technology. 2003 Annual Meeting, Nashville, TN. [Consulta: Abril de 2006]



- [42] **Corkrey Ross**. “*Interactive Voice Response: Review of studies 1989–2000*”. Behavior Research Methods, Instruments, & Computers. Agosto 2002. Vol. 34 Issue 3. [Consulta: Abril de 2006]
- [43] **Kuperstein Michael**. “*Computer Speech Rec at Every Call Center*”. Communications Convergence. Diciembre 2003. Vol. 11 Issue 12. [Consulta: Abril de 2006]
- [44] **Caton, Michael**. “*Banking on speech services*”. Agosto 2004. Vol. 21 Issue 35. [Consulta: Abril de 2006]
- [45] **Lenke Nils**. “*Supercharging man/machine dialog*”. Electronic Engineering Times. Marzo 1999. Issue 1052. [Consulta: Abril de 2006]
- [46] **B.H. Juang**. “*Automatic Speech Recognition – A Brief History of the Technology Development*”. [En línea]. Disponible en web: <http://www.caip.rutgers.edu/~lrr/lrr%20papers/353_LALI-ASRHistory-final-10-8.pdf> [Consulta: Abril de 2006]
- [47] **Richard V**. “*A segmental speech coder based on a concatenative TTS*”. Speech Communication. Septiembre 2002. Vol. 38. [Consulta: Abril de 2006]
- [48] **Baum, David**. “*Will your PC talk back by 2001?*”. InfoWorld. Mayo 1994. Vol. 16 Issue 20. [Consulta: Abril de 2006]
- [49] **Van Santen**. “*Synthesis of prosody using multi-level unit sequences*”. Speech Communication. Julio 2005. Vol. 46 Issue 3/4. [Consulta: Abril de 2006]
- [50] **Scharenborg, Odette**. “*Early recognition of polysyllabic words in continuous speech*”. Computer Speech & Language. Enero 2006. Vol. 21. [Consulta: Abril de 2006]
- [51] **M C Bale**. “*Voice and Internet multimedia in UMTS networks*”. [En línea]. Enero 2001. Disponible en web: <[http://www.sipcenter.com/sip.nsf/html/WEBB5YH724/\\$FILE/Voice_and_MM_in_3G.pdf](http://www.sipcenter.com/sip.nsf/html/WEBB5YH724/$FILE/Voice_and_MM_in_3G.pdf)> [Consulta: Abril de 2006]
- [52] **Voxpilot**. “*Interactive Services for 3G and IMS Networks: The Next Wave*”. [En línea]. Mayo 2006. Disponible en web: <<http://www.voxpilot.com/pdf/IVVR3G.pdf>> [Consulta: Agosto de 2006]
- [53] **3GPP**. “*3GPP TR 23.877 V6.0.0 (2004-03) 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects Architectural aspects of Speech Enabled Services; (Release 6)*”. [En línea]. Marzo 2004. Disponible en web: <http://www.3gpp.org/ftp/specs/archive/23_series/23.877/23877-600.zip> [Consulta: Abril de 2004]
- [54] **3GPP**. “*3GPP TS 23.333 V0.3.0 (2006-09) 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Multimedia Resource Function Controller (MRFC) – Multimedia Resource Function Processor (MRFP) Mp interface: Procedures Descriptions (Release 7)*”. [En línea]. Septiembre 2006. Disponible en web: <<http://www.3gpp.org/ftp/specs/latest-drafts/23333-030.zip>> [Consulta: Octubre de 2006]