

VOZ SOBRE IP. TÉCNICAS DE PAQUETIZACIÓN Y CALIDAD DE SERVICIO

TOLL P. YUNIET DEL C. ⁽¹⁾

GUZMÁN O. OMAR A. ⁽²⁾

RIL G. YOHANDRI ⁽¹⁾

VIGOA M. LILIAN ⁽⁴⁾

Recibido: 02-02-11 / Aceptado: 13-07-11

RESUMEN

En la presente investigación se hace una breve panorámica acerca de la tecnología VoIP o Voz sobre IP (Protocolo de Internet) y del proceso de paquetización de voz en una red de paquetes IP. Se identifican una serie de técnicas y mecanismos para realizar las llamadas telefónicas a través del muestreo y digitalización de la voz, procediendo posteriormente a su encapsulado en paquetes. Durante el trabajo se relacionan algunos de los códecs, el algoritmo de compresión que utilizan, la calidad intrínseca y la demora añadida, utilizados en la telefonía tradicional. También se hace alusión a la principal ventaja de todo el proceso de VoIP que no es más que el aprovechamiento de una red para brindar múltiples servicios y poder obtener una gestión segura y confiable en la red y garantizar el ahorro de ancho de banda.

Palabras Clave: Calidad de Servicio; Paquetización; Paquetes IP; Protocolo de Internet; Voz sobre IP.

⁽¹⁾ **Ing. Yuniét Del Carmen Toll Palma:** Profesor de teleinformática I (Universidad de las Ciencias Informáticas, Cuba), especialista de Información del Centro Tecnologías para la Formación de la Universidad de las Ciencias Informáticas, Ingeniera en Ciencias Informáticas y Cursando Maestría de Calidad de Software (Universidad de las Ciencias Informáticas), ytoll@uci.cu

⁽²⁾ **Ing Omar Alexander Guzmán Obregón,** Profesor de teleinformática I (Universidad de las Ciencias Informáticas), Ingeniero en Telecomunicaciones y Electrónica (Universidad Central de Las Villas, Cuba), Cursando Maestría de Sistemas Digitales (Instituto Superior Politécnico José Antonio Echeverría, Cuba), omaralexander@uci.cu

⁽³⁾ **Ing Yohandri Ril Gil,** Asesor Técnico Docente del Departamento Docente Central de Sistemas Digitales (Universidad de las Ciencias Informática, Cuba), Ingeniero en Telecomunicaciones y Electrónica (Instituto Superior Politécnico José Antonio Echeverría, Cuba), Cursando Maestría de Telemática (Instituto Superior Politécnico José Antonio Echeverría, Cuba), rilltt@uci.cu

⁽⁴⁾ **Ing. Lilian Vigoa Machin:** Profesor de Práctica Profesional IIII (Universidad de las Ciencias Informáticas), especialista de Información del Centro Tecnologías para la Formación de la Universidad de las Ciencias Informáticas, Ingeniera en Ciencias Informáticas, Imachin@uci.cu.

Voice about IP. techniques package and quality of service

ABSTRACT

In this research there is a brief overview about VoIP or Voice about IP (Internet Protocol) and the process of voice package in a IP packet network. In this study a number of techniques and mechanisms to make phone calls through sampling and digitalization of the voice, coming after its encapsulation in packets is identified. During the work some of the codecs are related, the compression algorithm that is used, the intrinsic quality and the added delay, used in traditional telephony. It also refers to the main advantage of the whole process of VoIP is simply the use of a network to provide multiple services and management to obtain a safe, reliable and secure network saving bandwidth.

Keywords: Quality of Service; Package; IP Packages; Internet Protocol; Voice about IP.

INTRODUCCIÓN

El desarrollo de técnicas avanzadas de digitalización de voz, el empleo de mecanismos para el control y priorización de tráfico y los protocolos de transmisión en tiempo real, favorecen la aparición de nuevos estándares para el establecimiento de la calidad de servicios en redes IP. De este modo se cuenta con un entorno donde es posible transmitir voz, datos y video sobre una misma plataforma IP.

La telefonía sobre IP abre un espacio muy importante dentro del universo que es Internet. Es la posibilidad de estar comunicados a costos más bajos dentro de las empresas y fuera de ellas, es la puerta de entrada de nuevos servicios apenas imaginados y es la forma de combinar una página de presentación de Web con la atención en vivo y en directo desde un Centro de llamada, entre muchas otras prestaciones. Lentamente, la telefonía sobre IP ha ido ganando terreno y muchos desean tenerla.

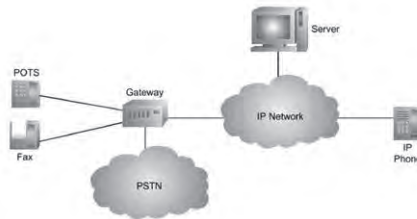
VoIP o Voz sobre IP o Telefonía IP como suelen llamarle los autores Alberto Escudero Pascual y Louise Berthilson (Pascual & Berthilson, 2006), son los mecanismos para enviar las llamadas telefónicas sobre redes de paquetes IP (Protocolo de Internet), como parte de la convergencia de servicios e infraestructuras de voz, datos y video. Esto significa, que la

voz puede viajar por una red de paquetes concebida originalmente para datos, ya sea en una red de transporte o de acceso. La principal ventaja sería el aprovechamiento de una red para brindar múltiples servicios y la simplificación de la gestión de la red, además de ahorro de ancho de banda. Para llevar a cabo las llamadas telefónicas se hace a través de un muestreo y digitalización de la voz, para después encapsularlas en paquetes, bajo el control de diferentes protocolos de señalización.

COMPONENTES PRINCIPALES DE VOIP

Los principales componentes de una red VoIP se muestran en la figura 1. El Gateway o Pasarela convierte las señales desde las interfaces de telefonía tradicional a VoIP. Un teléfono IP es un terminal que tiene soporte VoIP nativo y puede conectarse directamente a una red IP. Por terminales se entienden los equipos de los usuarios que pueden ser teléfonos IP, computadoras con software que realiza la emulación de teléfonos IP o pasarelas de media residenciales.

Fig. 1 Principales componentes de VoIP (Boullet, 2000).



El servidor provee el manejo y funciones administrativas para soportar el enrutamiento de llamadas a través de la red. En un sistema basado en el estándar H.323, el servidor es conocido como un Gatekeeper. En un sistema SIP, el servidor es un servidor SIP. En un sistema basado en MGCP o H.248/MEGACO, el servidor es un Agente de Llamadas o Softswitch. Finalmente, la red IP provee conectividad entre todos los terminales e implementa el direccionamiento, el enrutamiento y la señalización. Esta red puede pertenecer a un operador de telecomunicaciones, una Red de Área Local o Internet.

Los sistemas están definidos mediante protocolos de señalización estándares, empleados tanto para el control de las llamadas como de los equipos que se encargan de realizarlas. Estos son la familia de protocolos H.323 de la ITU-T, el protocolo SIP (Protocolo de Inicio de Sesión) del IETF (Fuerza de Tarea de Ingeniería para Internet) y el H.248 o

MEGACO, una evolución del MGCP (Protocolo de Control de Pasarelas de Media), desarrollado en conjunto por el IETF y la ITU-T. Los dos primeros se basan en el enfoque de equipos IP "inteligentes" o arquitectura distribuida y el tercero en equipos IP controlados por una inteligencia centralizada (Controlador de Pasarelas de Media o Agente de Llamadas), estos sólo tratan las funciones de codificación y tratamiento de los paquetes (Boullet, 2000).

H323 fue el primer protocolo abierto de telefonía IP. Surge como una evolución de otros protocolos de videoconferencia y telecomunicaciones. Define entre sus componentes al **Terminal** como el dispositivo encargado de iniciar o recibir una llamada. El **Gatekeeper** o controlador de dominio proporciona funciones como la traducción de direcciones, el control del ancho de banda, la administración de zonas y la autorización de llamadas entre otras. Finalmente el **Gateway** o convertidor de medios permite la conversión de redes H323 a otras redes como la Red Pública Conmutada (PSTN, por sus siglas en inglés) y la Red Digital de Servicios Integrados (RDSI).

El acceso a la especificación de los protocolos H32X no es gratuito, el estándar se ha de comprar, siguiendo la política de distribución de estandarización ISO. Aún así, el ITU-T ha publicado algunos estándares de manera libre (Rosa, 2009).

El otro protocolo que distingue el proceso de señalización es el de inicio de sesión (SIP, por sus siglas en inglés). En la propia RFC 3261 se expresa:

"SIP es un protocolo de control de la capa de aplicación que puede establecer, modificar, y terminar sesiones multimedia como llamadas telefónicas por Internet. SIP puede también invitar participantes a sesiones existentes, como conferencias."

SIP soporta cinco funciones de las comunicaciones multimedia:

Localización del usuario: determina el lugar del usuario final para la comunicación.

Disponibilidad del usuario: determina la voluntad del llamado de unirse a una comunicación.

Capacidades del usuario: determina los medios que se usuario final.

Sesión de configuración: "llamar", establecimiento de los parámetros de sesión para llamante y llamado.

Gestión de la sesión: incluyendo transferencia y terminación de las sesiones, modificación e invocación de servicios.

A diferencia de H323, SIP está disponible de manera pública y gratuita en la web de los RFC. SIP está basado en el protocolo HTTP, del cual toma el modelo general y algunos códigos de respuesta. Cualquier paquete SIP se escribe en texto plano, y puede observarse mediante cualquier analizador de red (Rosa, 2009).

Una vez establecido el protocolo de señalización, se garantiza en la red IP el esquema de direccionamiento requerido para identificar el origen y destino de las llamadas. También es empleado para asociar las clases de servicio a cada una de las llamadas dependiendo de la prioridad. El enrutamiento por su parte se encarga de determinar el mejor camino a seguir por el paquete desde la fuente hasta el destino y transporta la información a través de la red de la manera más eficiente. Mientras que el proceso de señalización alerta a las estaciones terminales y a los elementos de la red su estado y la responsabilidad inmediata que tienen al establecer una conexión.

CALIDAD DE SERVICIO

La UIT (Unión Internacional de Telecomunicaciones) define la calidad de servicio como: "El efecto colectivo del desempeño de los servicios que determina el grado de satisfacción del usuario con respecto a un servicio". El objetivo de la VoIP es emular la calidad de las llamadas telefónicas en la RTPC (Red Telefónica Pública Conmutada), con la convergencia de la voz y los datos en las redes IP. El problema es que las redes IP utilizan un servicio no orientado a conexión y con el mejor esfuerzo, por lo que los paquetes pueden llegar en orden diferente a como se enviaron y siguiendo rutas diferentes, siendo eliminados en las colas de los enrutadores si hay congestión.

Las aplicaciones en tiempo real como la voz son afectadas por la pérdida de paquetes, la demora, el *Jitter* o variación de la demora entre llegadas de los paquetes y el eco. Estos fenómenos degradan la calidad vocal, afectada principalmente por el retraso y distorsión boca a oído (Vleeschauer, Janssen, Petit, & Poppe, 2000). El retraso boca a oído para las llamadas empaquetadas tiene su mayor contribución en dos inestabilidades, el eco y la pérdida de interactividad (Büchli, Moffaert, Petit, Vleeschauer, & Janssen, 2001). A la pérdida de interactividad, contribuyen el proceso de codificación, empaquetado, propagación por la red, encolado, servicio, retraso en la supresión del *jittering* y la decodificación.

El eco puede aparecer en las híbridas de cuatro a dos hilos en la red telefónica local, en teléfonos analógicos o en la Pasarela Residencial en las instalaciones de los abonados. La distorsión surge a causa de la compresión de la voz, o la pérdida de paquetes durante el transporte debido a congestión o en el buffer para suprimir el *jittering*, principalmente.

El Jitter puede determinarse mediante la expresión (Tosi, 1987):

Donde:

$$J = 1/n \sum_{i=1}^n (f_i - \bar{f}) \quad [\text{Hz}]$$

$$J \% = (J/\bar{f}) 100$$

n = número de ondas consideradas.

f_i = frecuencia de cada una de las ondas sucesivas.

f = frecuencia promedio de las n ondas consideradas.

Por su parte, la cantidad de eco viene dada por la pérdida de eco, que es la relación entre el nivel de la señal original y el nivel de la señal con eco expresada en decibeles. El eco puede ser molesto, al hacer inentendible la conversación. Utilizando un controlador de eco puede reducirse notablemente la cantidad de eco (Büchli, et al., 2001).

Existe relación entre la compresión y la calidad. Mientras más se comprime la señal de voz en el códec, menor es la calidad. Es esencial elegir el códec vocal apropiado. Las características del códec incluyen la calidad básica (es decir, sin degradación) y el desempeño con degradación, tal como ruido de fondo y paquetes perdidos. Para evitar una degradación excesiva por la transcodificación, es necesario controlar en dónde ocurren las transcodificaciones y qué combinaciones de códecs se usa (Cedillo & Gabelloni, 2004).

La mejor manera de evitar paquetes tardíos y perdidos es proyectar la red intentando excluir o minimizar las demoras y otros factores que contribuyan a estas. Esto significa que el control de la congestión (llamado control de admisión de llamadas) debe estar en servicio para evitar que se llenen las colas de los encaminadores, lo cual causa variaciones en el retardo, y posiblemente un desborde.

PAQUETIZACIÓN DE LA VOZ. EL PROCESO

Una vez que la llamada ha sido establecida, mediante alguno de los protocolos de señalización la voz será digitalizada y transmitida a través de la red en tramas IP.

El software del módulo de paquetización de los terminales IP o las p-arelas entre la RTPC y la red IP, es generalmente implementado en un Procesador Digital de Señales el cual toma muestras de voz a una frecuencia determinada, dichas muestras digitales, se codifican y comprimen mediante códecs (codificadores decodificadores) normalizados internacionalmente por el ITU-T, como G.723.1, G.729, G.726, G.728 o G.711, para diferentes algoritmos de compresión y razón de bits. Cada códec tiene una sensibilidad de pérdida de paquetes, esta depende del acuerdo de pérdida de paquetes implementada en el códec (Instruments, 1998).

En dependencia de la arquitectura para servicio VoIP escogida este software además debe hacer cancelación de eco, eliminación del jitter, sincronización de reloj, detección de actividad de voz, y detectar si la llamada procede de un fax o un módem.

A continuación aparece una tabla que relaciona alguno de los códecs, el algoritmo de compresión que utilizan, la razón de bits, la calidad intrínseca y la demora añadida con respecto al Códec G.711, utilizado en la telefonía tradicional.

Tabla 1 Algunos de los códecs utilizados en la telefonía tradicional (Büchli, et al., 2001).

Código Estándar	Algoritmo	Razón de datos	Calidad Intrínseca	Demora añadida
G.711	PCM (Pulse Code Modulation)	64 Kbps	Excelente	No
G.726	ADPCM (Adaptive Differential Pulse Code Modulation)	16,24,32,40 kbps	Buena (40) Aceptable (24)	Muy Baja
G.728	LD-CELP (Low Delay Code Excited Linear Prediction)	16 kbps	Buena	Baja
G.729	CS-ACELP (Conjugate Structure Algebraic CELP)	8 kbps	Buena	Baja
G.723.1	MP-MLQ (Multi-Pulse Maximun Likelihood Quantization)	6,3 kbps 5,3 kbps	Buena	Alta
	ACLPL (Algebraic Code Excited Linear Prediction)	6,3 kbps 5,3 kbps	Buena	Alta

La tabla 2 muestra la familia de protocolos H.323, pero el encapsulamiento de los códecs de audio y video es común a todas las arquitecturas. Los paquetes de voz comprimidos se encapsulan en el protocolo de transporte de media RTP (Protocolo de Tiempo Real), que a su vez hace compresión de cabecera de protocolo para optimizar ancho de banda, este a su vez se encapsula en el espacio de carga del UDP (Protocolo de Datagramas de Usuario) de la capa de transporte TCP/IP, y después se encapsula en el protocolo IP, y este a su vez en el protocolo de la capa de enlace. El protocolo RTCP (Protocolo de Control de Tiempo Real) controla los canales RTP y detecta las situaciones de congestión en la red y toma acciones correctoras.

Tabla 2 Familia de Protocolos H.323

Audio	Video	RAS	H.245	Q.931
Codecs	Codecs	(H.225.0)		(H.225.0)
RTP/RTCP				
UDP			TCP	
IP				

A continuación un ejemplo de encapsulamiento de voz utilizando el códec G.729:

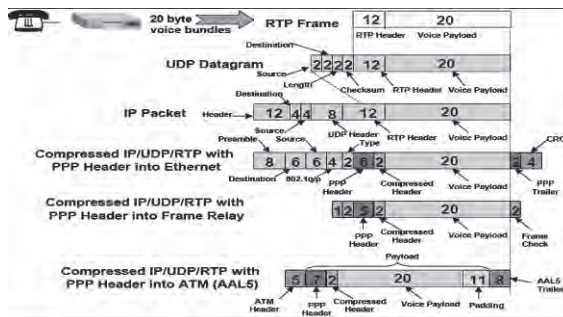


Fig. 2 Ejemplo de encapsulamiento de voz utilizando el códec G.729 (Tyre & Britt, 2008)

Las muestras de voz que ocupan 20 bytes de datos se encapsulan en la carga útil del Protocolo en Tiempo Real (RTP), el mismo posee una cabecera de longitud 12 bytes. Esta trama RTP, se encapsula en un datagrama del Protocolo de Datagrama de Usuario (UDP), posee mayor

velocidad que el Protocolo de Control de la Transmisión (TCP), debido a que es más importante evitar las demoras y las retransmisiones en el caso de que se pierdan algunos paquetes. Este cuenta con 8 bytes de cabecera asignado a los puertos origen y destino, la suma de chequeo y la longitud del datagrama. El RTP emplea los puertos bien conocidos del UDP, en cambio el Protocolo de Control en Tiempo Real (RTCP), utiliza los menos conocidos. El datagrama se encapsula en un paquete IPv4, que tiene 20 bytes de cabecera, utilizando 4 bytes para dirección origen y destino respectivamente. Seguidamente se comprime la cabecera para transmitirla sobre la trama de la Capa de Enlace, en el ejemplo aparece el Protocolo Punto a Punto (PPP) sobre una trama Ethernet, con 2 bytes de cabecera IP/UDP/RTP comprimidos, 6 bytes de cabecera PPP, y 2 bytes de cola. La trama Ethernet posee 26 bytes de cabecera y 4 bytes para el código de chequeo de errores.

CONSIDERACIONES DE RENDIMIENTO

A la hora de seleccionar el tamaño de los paquetes de voz, se establece un estrecho compromiso con el rendimiento. Un incremento en el volumen de la carga útil por paquete, influye directamente en el crecimiento del impacto provocado por las pérdidas de paquetes, disminuyendo a su vez la calidad de audio en la recepción. La operación de paquetización experimenta el aumento del retardo intrínseco y el crecimiento de la dimensión de los buffers a causa del empaquetado de grandes paquetes de voz. La ITU-T establece dentro de las recomendaciones para el uso del códec G.711 (tradicionalmente utilizado en la RTPC) una pérdida de tráfico permisible de 4-16 ms debido a su imperceptibilidad. Perder tráfico de 32-64 ms afecta notablemente el rendimiento.

MONITOREO VoIP

La tecnología VoIP emplea un Ancho de Banda compartido al igual que otras aplicaciones tradicionales como la mensajería instantánea y el correo electrónico. Sin embargo, al funcionar en tiempo real la calidad VoIP se afecta drásticamente por los parámetros de red como retrasos, inestabilidades y pérdidas de paquetes. De aquí la necesidad de contar con un sistema de monitoreo que garantice la supervisión constante de los principales indicadores de calidad VoIP.

Entre las principales herramientas utilizadas con este fin se pueden mencionar Nagios y VQManager, pero no son las únicas. A continuación se presenta una tabla comparativa entre algunas de ellas.

Tabla 3 Comparativa entre herramientas de monitoreo

Características	Nagios	VQManager	Monit	Hobbit	Munin
Interfaz web	x	x	x	x	x
Alertas y notificaciones	x		x	x	
Generación de reportes	x	x			
Gráficas estadísticas	x	x	x	x	x
Facilidad de uso	x	x	x	x	x
Escalable y robusto	x			x	
Autenticación de usuarios	x				
Usado para redes locales	x	x	x	x	x
Licencia libre	x		x	x	x
Orientado a VoIP		x	x		

CONCLUSIONES

La VoIP es un factor determinante en la convergencia de los servicios y redes, utilizando como elemento aglutinador el protocolo IP. Pero se debe tener en cuenta los problemas que afectan a la voz a la hora de reproducir la fidelidad de las llamadas telefónicas en una red por conmutación de paquetes. La pérdida de interactividad y la degradación de la calidad vocal son las aristas más visibles del problema. La encapsulación de protocolos estándares es un elemento fundamental para la transmisión de la voz a través de los diferentes niveles de red. Las limitaciones del ancho de banda empleado para la transmisión de voz y datos requieren de la utilización de herramientas de monitoreo. Deben ser generadas las notificaciones oportunas acerca del funcionamiento de aquellos sistemas, aplicaciones y servicios que empleen las redes IP como soporte para la transmisión de voz.

RECOMENDACIONES

La migración hacia la tecnología VoIP está marcada por la situación económica de regiones, países, instituciones y personas. Resulta muy útil profundizar sobre esta temática, particularizando en las necesidades mínimas de capacitación y equipamiento para tales propósitos. Puede ser incorporado a este estudio el análisis sobre distintos métodos aplicados a otras regiones, para el aprovechamiento de la tecnología existente.

REFERENCIAS

- Boulet, M. (2000, 23/01/08). Voz sobre IP en las PABX OmniPCX 4400 y OmniOffice de Alcatel. *Revista de Telecomunicaciones ALCATEL*, 7-11.
- Büchli, M., Moffaert, A. V., Petit, G., Vleeschauwer, D. D., & Janssen, J. (2001, 23/01/08). Calidad de las llamadas transportadas sobre una red de acceso DSL. *Revista de Telecomunicaciones de Alcatel*, Nº. 2, 111-115.
- Cedillo, E., & Gabelloni, E. (2004). Estudios sobre características de la voz basadas en redes que usan IP. *CITEL*.
- Instruments, T. (1998). Voice Over Packet, from http://focus.ti.com/pdfs/vf/bband/vop_white_paper.pdf
- Pascual, A. E., & Berthilson, L. (2006). VoIP para el desarrollo. Una guía para crear una infraestructura de voz en regiones en desarrollo. 1-46. Retrieved from http://comunidad.asterisk-es.org/introduccion_voip.pdf
- Rosa, F. S. (2009). Introducción a la telefonía IP utilizando estándares., 37. Retrieved from <https://forja.rediris.es/docman/view.php/555/853/Intro-voip-uca.pdf>
- Tosi, O. (1987). Medición Computarizada de Jitter y Shimmer (Vol. VII, pp. 1). Available from <http://www.elsevier.es/sites/default/files/elsevier/pdf/309/309v07n01a13152725pdf001.pdf>
- Tyre, J., & Britt, R. (2008, 23/01/08). Performance Characteristics of Voice over IP Networks. *Nortel Application Note AN00SIPT03 v1.0*.
- Vleeschauwer, D. D., Janssen, J., Petit, G., & Poppe, F. (2000, 23/01/08). Límites de calidad para el transporte de paquetes de voz. *Revista de Telecomunicaciones de Alcatel*, 1er trimestre, 19-23.