



Proyecto Redes de Computadores I

Protocolo de Voz sobre IP

Fecha 06 de Septiembre 2013

Integrantes Nicolás Oneto Velásquez
Nicolás Ramírez Didier
Max Sottovia Rodríguez

Profesor Agustín J. González

Resumen

El fuerte crecimiento y desarrollo de las redes IP, tanto local como remotamente, sumado a las técnicas avanzadas de digitalización de voz, mecanismos de control y de tráfico, así como el estudio de nuevos estándares para la calidad de servicio en redes IP, han creado un panorama en donde es posible transmitir telefonía sobre IP lo que no significará de ninguna manera la desaparición de las redes telefónicas fijas, sino que necesariamente existirá una transición en donde ambas tecnologías coexistan.

A estas características anteriores debemos sumarle el fenómeno que ha experimentado el Internet en nuestro país en los últimos 10 años, junto con el potencial ahorro económico que este tipo de tecnologías acarrearán, se concluye fácilmente que el protocolo de Voz sobre IP es un tema importantísimo para el desarrollo de las comunicaciones tanto en el ámbito privado como en el público.

Este documento se referirá a las características principales del protocolo de Voz sobre IP, analizando cómo se puede construir su red, sus ventajas y desventajas, así como sus proyecciones y desafíos a futuro como tecnología.

Introducción

La industria de las comunicaciones ha avanzado de forma agigantada en los últimos, digamos, 30 años. El primer gran cambio se produjo con el paso de los sistemas predominantemente analógicos a sistemas digitales y el segundo, ocurrido a finales de la década de los '90, el paso a la telefonía IP.

La migración ha sido relativamente lenta, principalmente porque las ventajas que esta tecnología ofrece no fueron tan “grandes” como cuando se pasó de sistemas analógicos a sistemas digitales. Hoy en día, generalmente son las empresas más grandes con un importante volumen de tráfico entre delegaciones o sedes en el extranjero las que se benefician de la telefonía IP: se reducen los costos de llamados, se implementa una infraestructura de red única, etc.

Una de las grandes ventajas que presenta VoIP es unir en un mismo protocolo dos características importantísimas: la transmisión de voz y la transmisión de datos. Es importante entender el concepto: VoIP no es un servicio sino una tecnología, mediante la cual, con una conexión estándar a internet se puede crear una plataforma para realizar llamadas gratuitas.

Dada la proliferación de las conexiones banda ancha en Chile, las llamadas VoIP que solían ser de baja calidad, superaron este punto. De hecho, muchos operadores han hecho híbridos

en sus sistemas y probablemente alguna de las llamadas que se hacen hoy en día de larga distancia, sean sobre este protocolo.

Marco teorico

Debemos notar la diferencia fundamental que existe entre la telefonía analógica tradicional y el tema que se presenta, VoIP.

La *PSTN* (Public Switched Telephone Network) se basa en un sistema de llamadas muy fiable pero a la vez un poco ineficaz, conocido como circuito de conmutación, concepto que se ha usado en las redes telefónicas desde sus comienzos. Al realizarse una llamada, la conexión se mantiene a lo largo de la misma. Dado que se están conectando dos puntos en ambas direcciones, se llama circuito. Es por esto que la Red Telefónica Conmutada podemos definirla como:

“Se define como el conjunto de elementos constituido por todos los medios de transmisión y conmutación necesarios para enlazar a voluntad dos equipos terminales mediante un circuito físico que se establece específicamente para la comunicación y que desaparece una vez que se ha completado la misma. Se trata por tanto, de una red de telecomunicaciones conmutada.”

En este sistema el funcionamiento de una llamada telefónica es el siguiente:

1. El emisor toma el auricular y escucha un tono de marcación.
2. Marcar el número al que desea llamar.
3. La llamada se direcciona a través del conmutador al nodo local a la parte que se está llamando.
4. Se establece una conexión entre el teléfono y la línea del receptor de la llamada usando varios conmutadores interconectados en el camino.
5. El teléfono en el otro extremo repica, y alguien responde la llamada.
6. La conexión abre el circuito.
7. Se habla de un período de tiempo y, luego se cuelgan los auriculares.
8. Al colgar, se cierra el circuito, liberando la línea y todas las líneas en el medio.

Supongamos que esta conversación duró 5 minutos, esto significa que durante ese tiempo, el circuito permaneció abierto entre los dos teléfonos. Recordemos que en los primeros sistemas telefónicos cada llamada debía tener un cable dedicado que va desde un extremo de la llamada al otro, durante lo que ésta durara.

Hoy en día las llamadas hechas sobre la red de telefonía tradicional son, obviamente, más eficientes. La voz es digitalizada y junto con las miles de otros se pueden combinar en un cable de fibra óptica por gran parte del recorrido.

Estas llamadas se transmiten a una tasa fija de 64 kbps en cada dirección, para un total de tasa de transmisión de 128 kbps.

Sabemos que:

$$1 \text{ kb} = 8 \text{ KB}$$

Por lo tanto en cada segundo de llama estaremos transmitiendo 16 Kbyte. En nuestra llamada de 5 minutos habremos transmitido:

$$5 * 60 [s] * 16 [Kbyte/s] = 4800 \text{ Kbytes}$$

Lo que es muy cercano a los 5 MB.

Ahora consideremos que cuando estamos en una conversación, la mitad del tiempo habla una persona, y la mitad del tiempo habla la otra, es decir, simplemente haciendo esta consideración podríamos reducir la transferencia a la mitad.

Para hacer este análisis aún más exhaustivo, es importante notar que durante una conversación en gran parte del tiempo ninguna de las dos personas habla, es decir, hay silencio.

Entonces nos hacemos la pregunta: ¿Y si sólo enviamos bits en los momentos en que no hay silencio, y cuando lo hay, no enviamos nada? Esto aumentaría notablemente la eficiencia del sistema.

La **conmutación de paquetes** mantiene la conexión abierta constantemente, en cambio el **intercambio de paquetes** que se utiliza en VoIP sólo abre una conexión, suficientemente extensa para enviar una pequeña porción de información llamada paquete, y funciona de la siguiente manera:

- El dispositivo que envía divide la información en pequeños paquetes, con una dirección en cada uno indicando a los dispositivos de red donde enviar los mismos.
- Adentro de cada paquete hay una porción de la información que se está enviando, la voz.
- El dispositivo emisor envía un paquete al router más cercano y se olvida del mismo.
- El router cercano envía el paquete a otro router que se encuentre más cerca del destino, ese router se lo envía a otro que se encuentra todavía más cerca del destino, ese a otro más cerca, y así sucesivamente.
- Cuando la computadora receptora finalmente recibe los paquetes (que pueden haber tomado caminos completamente diferentes para haber llegado ahí). Usa las instrucciones contenidas en los paquetes para rearmar los datos en su estado original.

El intercambio de paquetes es muy eficiente. Deja a la red enviar los paquetes a lo largo de las rutas menos congestionadas. También libera a las computadoras de forma que estas pueden también aceptar información proveniente de otras computadoras.

¿Cómo es un paquete VoIP?

La tecnología que estamos describiendo requiere de dos clases de protocolos: un protocolo de señalización, el cual mencionaremos posteriormente, y un protocolo para llevar los paquetes de voz.

El Real-Time Protocol (RTP) lleva la transmisión de voz y es de uso común en los proveedores de redes IP.

Un teléfono IP genera un paquete de voz cada 10, 20, 30 o 40 [ms] dependiendo del proveedor. Los 10 y 40 [ms] de voz digitalizada pueden ser comprimidos e incluso encriptados, esto no es importante para el protocolo, ya que toma muchos paquetes para llevar una palabra.

La demora teléfono a teléfono tiene que ser limitada. Cuanto más corto sea el retraso de creación de paquetes, la red puede tolerar más retraso en la llamada VoIP. Entre más cortos sean los paquetes, estos causarán un problema menor si el paquete se pierde. Paquetes cortos requieren más ancho de banda, sin embargo, debido a la sobrecarga de paquete mayor. Los paquetes más largos, que contienen más bytes de voz reducen los requisitos de ancho de banda, pero produce un retraso en la construcción y son más difíciles de solucionar si se pierde.

El formato de un paquete RTP se puede ver a continuación en la Figura 1.

Ethernet Trailer	Digitized Voice	RTP Header	UDP Header	IP Header	Ethernet Header
---------------------	--------------------	---------------	---------------	--------------	--------------------

Figura 1: Paquete RTP.

Cada uno de estos paquetes contienen parte de una palabra hablada digitalizada. El tipo de paquete es de 50 paquetes por segundo para 20 ms y 33.3 paquetes por segundo para los paquetes de voz 30ms. La voz se transmite en estas tasas fijas. El campo de voz digitalizada puede contener tan poco como 10 bytes de voz comprimida o tanto como 320 bytes de voz sin comprimir.

Estructura de una red VoIP

La estructura de una red de voz sobre IP es la misma que se maneja en Internet. Las aplicaciones, los medios de transporte, la organización del ruteo sobre la red, los modos de enlace y la transmisión de la señal por los medios físicos deben formar parte del modelo OSI. La ventaja de la red VoIP es que no importa el tipo de aplicación mientras ésta pueda transformar su información en paquetes, segmentos, datagramas y finalmente bits.

Como ya hemos mencionado el protocolo que se utiliza para la capa de transporte es el RTP en datagramas tipo UDP sobre IP. Se ha escogido éste sobre el TCP dado que, como vimos durante el curso, TCP es caracterizado por ser un protocolo donde se deben recibir señales de reconocimiento (ACK) por parte del receptor antes de enviar el siguiente segmento. A ello se le conoce como protocolo orientado a conexión y ofrece seguridad a la transmisión y recepción de los paquetes aunque introduce retardos en la comunicación. UDP funciona distinto, se le denomina no orientado a la conexión ya que simplemente se encarga de enviar el paquete y no requiere de ningún reconocimiento.

La voz debe ser enviada en tiempo real, con la menor cantidad de retardos posibles. Una recomendación de la ITU-T (ITU Telecommunication Standardization Sector) es que estos retardos no sean mayores a 150 [ms] en una dirección.

En la Figura 2 se diagrama la estructura de una red IP.

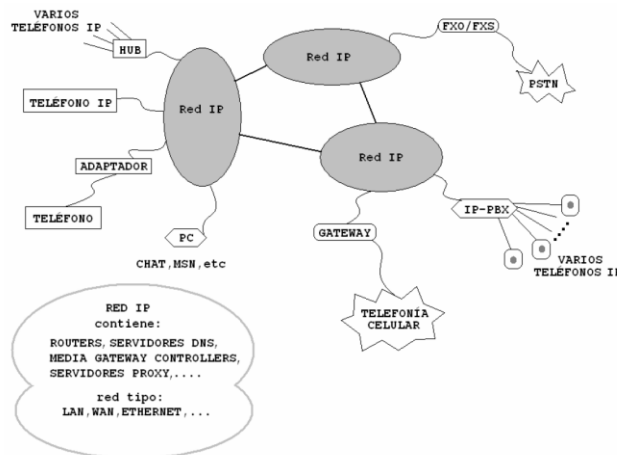


Figura 2: Estructura red IP.

En la Figura 2 existen los elementos FXO y FXS, el primero conecta VoIP con Centrales de la PSTN (Red Pública) y el segundo convierte de VoIP a canal analógico a nivel línea de abonado.

Aplicación

Dada la dificultad de montar una red IP real (PBX) nos limitamos a usar la popular aplicación Viber, para poder visualizar los paquetes VoIP en una comunicación entre el computador y un dispositivo móvil.

Es importante notar que el paquete RTP no se puede identificar ya que hace algunos meses esta aplicación fue *hackeada* justamente porque se podían leer los paquetes y aplicar *sniffing*, es decir, a partir del mismo Wireshark era posible reconstruir una conversación entre dos personas. La aplicación fue actualizada y hoy en día esta información se encuentra encriptada.

The image shows a Wireshark capture of Viber RTP packets. The top part is a packet list table with columns for Time, Source IP, Destination IP, Protocol, and Length. Several packets are highlighted in red. Below the list, the details pane for packet 1641 is expanded, showing metadata such as arrival time, epoch time, and frame length. The frame length is 177 bytes (1416 bits). At the bottom, the raw packet data is shown in hexadecimal and ASCII.

Time	Source IP	Destination IP	Protocol	Length	Info
1626	32.265030000	192.168.1.113	192.168.1.104	RTP	63 Unknown RTP version 3
1627	32.265190000	192.168.1.104	190.44.175.94	RTP	63 Unknown RTP version 0
1628	32.294566000	192.168.1.104	190.44.175.94	RTP	119 Unknown RTP version 0
1629	32.323307000	192.168.1.113	192.168.1.104	RTP	65 Unknown RTP version 3
1630	32.325021000	192.168.1.104	190.44.175.94	RTP	139 Unknown RTP version 0
1631	32.355455000	192.168.1.104	190.44.175.94	RTP	118 Unknown RTP version 0
1632	32.385872000	192.168.1.104	190.44.175.94	RTP	145 Unknown RTP version 0
1633	32.400842000	192.168.1.113	192.168.1.104	RTP	65 Unknown RTP version 3
1634	32.416496000	192.168.1.104	190.44.175.94	RTP	111 Unknown RTP version 0
1635	32.446817000	192.168.1.104	190.44.175.94	RTP	117 Unknown RTP version 0
1636	32.477483000	192.168.1.104	190.44.175.94	RTP	194 Unknown RTP version 0
1637	32.503644000	192.168.1.113	192.168.1.104	RTP	65 Unknown RTP version 3
1638	32.507722000	192.168.1.104	190.44.175.94	RTP	180 Unknown RTP version 0
1639	32.538507000	192.168.1.104	190.44.175.94	RTP	122 Unknown RTP version 0
1640	32.538855000	192.168.1.113	192.168.1.104	RTP	110 Unknown RTP version 3
1641	32.569155000	192.168.1.104	190.44.175.94	RTP	177 Unknown RTP version 0
1642	32.578656000	192.168.1.104	190.44.175.94	RTP	63 Unknown RTP version 0
1643	32.578780000	192.168.1.104	54.225.248.236	RTP	62 Unknown RTP version 0

Frame 1641: 177 bytes on wire (1416 bits), 177 bytes captured (1416 bits) on interface 1

Interface id: 1
WTAP_ENCAP: 1
Arrival Time: Sep 5, 2013 18:12:24.576077000 Hora est. Sudamérica Pacifico
[Time shift for this packet: 0.000000000 seconds]
Epoch Time: 1378419144.576077000 seconds
[Time delta from previous captured frame: 0.030300000 seconds]
[Time delta from previous displayed frame: 0.030300000 seconds]
[Time since reference or first frame: 32.569155000 seconds]
Frame Number: 1641
Frame Length: 177 bytes (1416 bits)
Capture Length: 177 bytes (1416 bits)
[Frame is marked: False]
[Frame is ignored: False]
[Protocols in frame: eth:ip:udp:rtp]

0000 00 25 86 d0 33 5a 00 13 e8 2e 74 89 08 00 45 00 .%.3Z...t...E.
0010 00 a3 09 d5 00 00 80 11 00 da c0 a8 01 68 be 2ch.
0020 5f 5e 76 55 73 8a 00 86 58 5e 13 70 80 67 07 73

Figura 3: Análisis Wireshark sobre Viber.

Como podemos apreciar en la Figura 3, no se puede ver el tipo de RTP, es decir, no podemos tener mucha información acerca del paquete dadas las razones que se han explicado previamente.

También vemos, como se describió, que el largo del paquete es de aproximadamente 177 bytes, podemos ver como estos largos varían, dependiendo de cómo se vaya desarrollando la información que se transmite (voz).

Conclusiones

Una de las ventajas más importantes que nos ofrece VoIP, tanto como para un ambiente de una pequeña como también para una gran organización, es el costo, ya que en la mayoría de los casos este es mucho más barato que su equivalente análogo debido a que se ocupan los mismos recursos tanto como para transmitir datos como para transmitir voz.

VoIP también es muy fácil de complementar con casi cualquier otro servicio que ocupe Internet, y además hay que considerar, desde el punto de vista del usuario, que se puede realizar una llamada VoIP desde cualquier lugar en que exista una conexión a internet, dado que los datos se transmiten como paquetes sobre la red. Particularmente nos encontramos con el caso de una persona que está constantemente cambiándose de lugar por temas de trabajo, no es simple tener asociado un número fijo, para este caso VoIP presta una utilidad gigantesca, considerando el alcance que tiene Internet hoy en día.

Es justamente ésta también una de sus desventajas, dado que requiere internet, y en países menos desarrollados no es tan simple acceder a una conexión banda ancha.

Referencias

- [1] <http://www.voipforo.com>
- [2] http://es.wikipedia.org/wiki/Voz_sobre_Protocolo_de_Internet
- [3] <http://www.telefoniavozip.com>
- [4] Memoria Jesús Recuerda Hueso, Universidad de Granada, España. (2012)
- [5] Apuntes Sistemas de Telecomunicaciones II, Walter Grote, UTFSM.

Anexo

Protocolos de señalización

La red de voz sobre IP es muy grande y eficiente. Como ya se dijo, es una red de paquetes conmutados. La voz es empaquetada y enviada por diferentes rutas, siempre buscando la ruta más corta y/o menos congestionada. Es ahí cuando los protocolos se hacen importantes.

Los dos protocolos que más se usan en las redes IP son H.323 y SIP, los cuales se explican a continuación:

Existen bastantes diferencias entre estos dos protocolos de VoIP. H.323 es un estándar de la ITU-T mientras que SIP, es un estándar más nuevo del IETF. Ambos protocolos realizan las mismas tareas básicas de telefonía (establecimiento de llamada y señalización de su inicio, tonos de marcación y terminación), así como la señalización de características de su mantenimiento, identificación y transferencia de llamadas. Sin embargo, mientras que en las redes H.323 estas tareas dependen de un servidor central (con terminales "tontos" en los extremos), SIP asume un esquema más descentralizado, desplazando cierta inteligencia hacia los clientes (teléfonos, PC, dispositivos inalámbricos, etc.).

Las principales diferencias podemos resumirlas en:

- H.323 se presentó como una evolución de SS7, diseñado para el control de la señalización en redes de conmutación de circuitos. Por el contrario, SIP está más cercano a HTTP, empleado en Internet, paradigma de red de paquetes. De cara al futuro, es mejor decantarse por SIP.
- En ambos casos, los flujos de información multimedia se transportan haciendo uso de RTP, por lo que la elección de un protocolo de control u otro no influye de manera directa en la calidad con que se ofrecen los servicios.
- H.323 es mucho más complejo que SIP. Tiene cientos de mensajes diferentes codificados en binario. en SIP, por el contrario, los mensajes son de texto y muy sencillos tanto en su sintaxis como en su semántica. Por lo tanto, H3.23 complica el trabajo a los desarrolladores, así como a los administradores de redes a la hora de localizar problemas.
- La arquitectura cliente/servidor de SIP es más fácil de implementar, al igual que sus mecanismos de seguridad y de gestión. H323 envía muchos mensajes a la red, con el riesgo potencial de crear congestión. Además, resulta difícil de personalizar.
- SIP es más fácilmente extensible y, por lo tanto, se puede adaptar mejor a las necesidades futuras de los usuarios. H.323 presenta un mayor número de limitaciones en este sentido.

Todo el mundo coincide en afirmar que el futuro está en SIP, que además es empleado por IMS (IP Multimedia Subsystem). El problema es que actualmente no ha alcanzado la madurez de H.323. Por ello, la mayoría de las soluciones comerciales se basan a día de hoy en H.323.

A manera de ejemplo se muestra el flujo de la llamada de la figura 3.

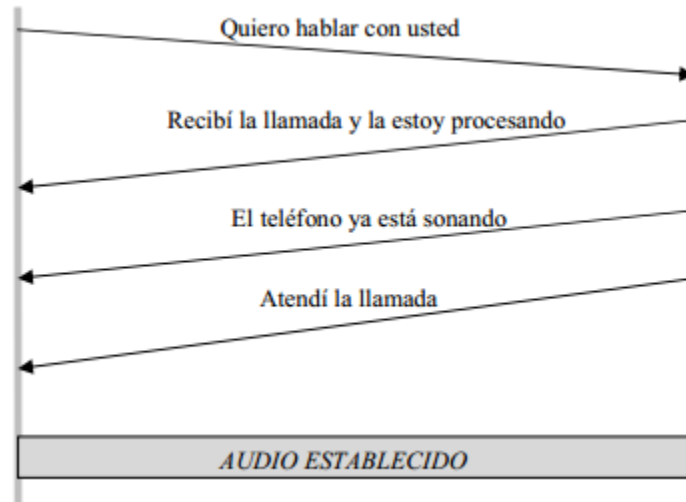


Figura 3: Flujo de llamada.

En el caso de H.323, se utilizan los mensajes que ya habían sido definidos en la norma Q.931 para ISDN. Como estos mensajes no fueron originalmente definidos para VoIP, carecen de parámetros relacionados con el "mundo" IP. Por ello, se definen en el protocolo H.225 los parámetros IP y son transportados sobre el protocolo Q.931 en el campo UUIE (User to User Information Element). H.225/Q.931 corre sobre TCP, por lo que requiere un establecimiento de conexión inicial. En este tipo de llamadas, el llamante trabaja como cliente, y el llamado como servidor en el puerto TCP 1720.

En el caso de SIP, utiliza sus propios mensajes que fueron definidos desde el inicio en un entorno IP. Es importante notar que SIP suele correr sobre UDP por ello aparece el mensaje de ACK confirmando el mensaje de 200OK. Al igual que en H.323, trabaja con un modelo cliente servidor donde el llamante es el cliente y el llamado un servidor en el puerto 5060 de UDP.