ESCUELA POLITÉCNICA DEL EJÉRCITO

FACULTAD DE INGENIERÍA ELECTRÓNICA

PROYECTO DE GRADO PARA LA OBTENCIÓN DEL TÍTULO EN INGENIERÍA ELECTRÓNICA

ESTUDIO TECNICO DEL CONTROL DE CALIDAD DE LOS SERVICIOS SOBRE IP

HILDA ESTHER LOZADA PEÑAFIEL

QUITO – ECUADOR

2005

CERTIFICACIÓN

Certificamos que la Srta. Hilda Esther Lozada Peñafiel ha elaborado el proyecto de grado titulado “Estudio técnico del control de calidad de los servicios sobre IP” para la obtención del título en Ingeniería Electrónica, bajo nuestra dirección.

Atentamente,

Ing. Carlos Usbeck W.	Ing. Rodrigo Silva
DIRECTOR	CODIRECTOR

AGRADECIMIENTO

Mi más sincero agradecimiento a todas aquellas personas que siempre han estado a mi lado, en los momentos buenos y también en los difíciles, pero de forma especial a mis padres, Moisés Lozada e Hilda Peñafiel, que con sus consejos y ejemplos me han enseñado a crecer en todos los ámbitos de la vida, siendo su apoyo y presencia la fuerza fundamental para seguir adelante día a día.

A mis amigos y compañeros de curso, que durante los años de estudio y ahora en la vida profesional han sido un apoyo permanente, con ellos aprendí que la amistad es mucho más que compartir buenos momentos, es estar juntos en los momentos difíciles y brindar una mano fraterna.

Un agradecimiento especial a las personas que me han brindado su apoyo en la elaboración de este proyecto de grado con sus conocimientos técnicos y guía, especialmente al Señor Ingeniero Carlos Usbeck W.

Y sobre todas las cosas doy gracias a Dios y a la Virgen, por permitirme culminar uno más de mis sueños.

Atentamente,

Hilda Esther Lozada P.

DEDICATORIA

A mis padres, Moisés e Hilda, guías de amor y ejemplo, a quienes debo la mejor herencia de la vida, la educación.

A mi abuelita Maria Presentación, a quien no pude agradecerle en vida lo suficiente por el tiempo dedicado.

A mi esposo y amigo Mario Oswaldo, quien con mucho cariño y paciencia me ha dado ánimo en los momentos difíciles.

Y en forma especial a mi hija Nathaly Victoria, quien con una solo sonrisa es capaz de borrar todas mis tristezas y darme las fuerzas para continuar.

PRÓLOGO

Este proyecto de grado centra su objetivo principal en ofrecer una visión actualizada de los múltiples protocolos y recomendaciones que se desarrollaron y se encuentran en proceso de mejora, en cuanto a la Calidad de Servicio QoS (Quality of Service) sobre las Redes IP para de forma eficiente y segura poder brindar nuevos servicios de voz, datos y video.

En el primer capítulo: INTRODUCCION, se indicará las nuevas prestaciones que ofrece la Red IP para transportar datos, voz y video simultáneamente, tal como: Fax sobre IP, Voz sobre IP, Multimedia sobre IP, dentro de una simple red, cuya tendencia actual es convertirse en una *Red Multiservicios*. También se indicará los pre-requisitos de la infraestructura de red para que pueda soportar estos servicios.

En el segundo capítulo: VOZ SOBRE IP VoIP (Voice over IP), hablaremos de las características principales, de tal forma de recordar conocimientos que ya han sido materia de otros proyectos de tesis, así como las aplicaciones más comunes. Pero el enfoque principal de este capítulo va dirigido a analizar los tópicos referentes a Calidad de Servicio, como son: Retardo, acumulación de retardos, procesamiento de retardo, red de retardos, compensación de pérdida de paquetes y eco.

En el tercer capítulo: FAX SOBRE IP FoIP (Fax over IP), esta orientado hacia una de las aplicaciones (servicio) que tenemos sobre una red IP que es todo aquello relacionado sobre FAX y los problemas de Calidad de Servicio, adicionalmente también se hablará acerca de las pruebas de Fax sobre IP que deben realizarse con el fin de mantener la compatibilidad y Calidad de Servicio entre una gran variedad de productos de FoIP.

En el cuarto capítulo: MULTIMEDIA SOBRE IP, se detallarán los protocolos que soporta esta tecnología (video, voz y datos), enfocados de forma especial en lo concerniente a las características y prestaciones que ofrecen cada uno de ellos con el objetivo de mejorar la Calidad de Servicio QoS en la Multimedia sobre IP.

En el quinto capítulo: PROTOCOLOS DE CALIDAD DE SERVICIO QoS SOBRE IP, se describe cada una de las tecnologías y los protocolos que se usan actualmente para ofrecer una Calidad de Servicio QoS eficiente sobre una red IP.

INDICE

CAPÍTULO I

INTRODUCCIÓN

El uso de Internet para propósitos comerciales, educativos, profesionales, gubernamentales y personales ha crecido considerablemente desde su nacimiento. Originalmente el Internet fue desarrollado por los físicos del Departamento de Defensa y académicos universitarios como medio para la transferencia eficiente de información a larga distancia.

Hoy en día, el Internet conecta más de 190 millones de usuarios en todo el mundo y esta creciendo a razón de 30 mil nuevos usuarios por mes.
En los Estados Unidos, Canadá, Europa (Oriental y Occidental) y Asia, el Internet esta experimentando su crecimiento más rápido. Solamente en Estados Unidos, más del 69% de la totalidad de computadoras se encuentran enlazadas a la red, según datos de la Sociedad de Internet de Reston, Virginia.

	2000	2002	2004	2006	2008	2010
Usuarios de Internet	600	975	1200	1300	1425	1625
Usuarios de Internet móvil	100	190	250	425	600	800
*** Datos en millones de Usuarios**						

Tabla. 1.1. Tendencia del Mercado de Internet

Con la creación de la Red Ancha Mundial WWW (World Wide Web), el Internet se convirtió en una red global de investigación, negocios y usuarios personales dando como resultado que la red basada en el Protocolo de Internet IP (Internet Protocol) se haya convertido en el factor estándar para la red de datos.

La Red de Telefonía de Conmutación Publica PSTN (Public Switched Telephone Network) y la red de Protocolo de Internet IP son fundamentalmente diferentes en términos de enrutamiento o direccionamiento y desarrollo, pero es posible para las redes a ser conectadas, intercambiar el tráfico de voz, tráfico de datos y actualmente video.
Las computadoras en la red de Protocolo de Internet IP se encuentran todas interconectadas, en otras palabras, la banda ancha es compartida por todos los usuarios activos, provocando que cuando la red esta más ocupada, cada usuario deba permanecer conectado más tiempo (una llamada promedio de los usuarios de Internet es de 30 minutos mientras las llamadas telefónicas normales son mucho menores) lo que provocará que el usuario experimente la degradación del servicio, además de congestionar la red telefónica tradicional.

Esta interconexión de diversas tecnologías de redes, ha incrementado significativamente los problemas concernientes a la **Calidad de Servicio QoS (Quality of Service)**, por lo que se debe proveer de nuevos mecanismos para negociar la calidad entre los usuarios finales y la red.

La Calidad de Servicio QoS se la puede definir como la habilidad de la red de garantizar y mantener ciertos niveles de actuación para cada aplicación de acuerdo a las necesidades especificadas por el usuario, lo cual se logra mediante diversos elementos físicos, estándares y protocolos.

La red IP es considerada como “una red de Mejor Esfuerzo” (best effort) en la transmisión, lo cual había funcionado bien porque la mayoría de las aplicaciones que corren sobre el Protocolo de Internet IP eran de baja prioridad, por ejemplo, aplicaciones de datos de pequeño ancho de banda con una elevada

tolerancia al retraso y a las variaciones del retraso. Sin embargo, esta situación está cambiando rápidamente, las nuevas aplicaciones de la red IP requieren actualmente un ancho de banda garantizado y estrictos requisitos de tiempo.

Este es el reto fundamental, extender las redes IP existentes con capacidades de multiservicio escalables y suministrar las ventajas de las redes IP con la Calidad de Servicio QoS de extremo a extremo que cubra los complejos requerimientos de los Acuerdos de Nivel de Servicio SLA (Service Level Agreements).

1.1 NUEVOS SERVICIOS SOBRE LA RED IP

Actualmente además de los servicios tradicionales que prestaba la red IP como son:

- ***Mensajería Electrónica:*** Se trata fundamentalmente del correo electrónico (e-mail), el cual es uno de los más importantes servicios que presta la red porque mediante este se puede enviar mensajes (de audio o datos), archivos, etc. a otra persona localizada en cualquier parte del planeta de forma privada y casi inmediata.
- ***Recursos Remotos:*** Permite trabajar de forma remota sobre otros equipos (PCs), pudiéndose realizar consultas a bases de datos o trabajar con programas cliente de otros servicios que se ejecutan en la máquina remota.
- ***Transferencia de archivos:*** Es un método para pasar información de una máquina a otra, mediante el uso del Protocolo de Transferencia de Archivos FTP (File Transfer Protocol), que da la posibilidad de conectarse a servidores de FTP.
- ***Integración de información:*** Servicios como Red Ancha Mundial WWW dónde podemos solicitar páginas de información y saltar de un punto a otro, descarga y localización de ficheros, se tiene acceso a los buscadores de información, a centros de debates (chats), etc.

Los nuevos servicios de la red IP, los cuales presentan mayores y mejores beneficios al usuario, debido a su gran flexibilidad y simplicidad, son:

- La telefonía IP o Voz sobre IP VoIP (Voice over IP)
- Mensajería de Fax sobre IP FoIP (Fax over IP)
- Multimedia sobre IP MoIP (Multimedia over IP)

Es decir que dentro de unos años más todo será transformado en bits: televisión, radio, prensa, diagnósticos médicos, etc. y podrá ser transportado mediante la red IP en una forma más potente, sencilla y barata a los terminales de los usuarios. Esta convergencia favorecerá la implantación de estándares en lapsos más breves y la calidad de servicios mejorará con la difusión tecnológica. Estos servicios de detallarán más ampliamente en los capítulos posteriores.

1.2 PRE-REQUISITOS DE LA RED PARA NUEVOS SERVICIOS SOBRE IP

Como ya se menciono antes la tendencia actual es converger todas las redes en una sola y el tremendo éxito de la Internet ha sido la causa de que el protocolo IP haya sido elegido como la capa multiservicio extremo a extremo. El protocolo IP hace de “puente” de varias tecnologías diferentes en la capa física y en la capa de enlace del modelo de Interconexión de Sistemas Abiertos OSI (Open Systems Interconnection) y proporciona una interfaz de red de capa 3 del modelo OSI hacia los servicios y las aplicaciones (conocida como *integración horizontal*).

Así la convergencia de las redes de comunicaciones que actualmente incluyen tecnologías de circuito conmutado y la Red IP crean capacidades que ninguna de las 2 redes solas pueden soportar por si misma, lo cual permite proporcionar nuevos y sorprendentes servicios al usuario final.

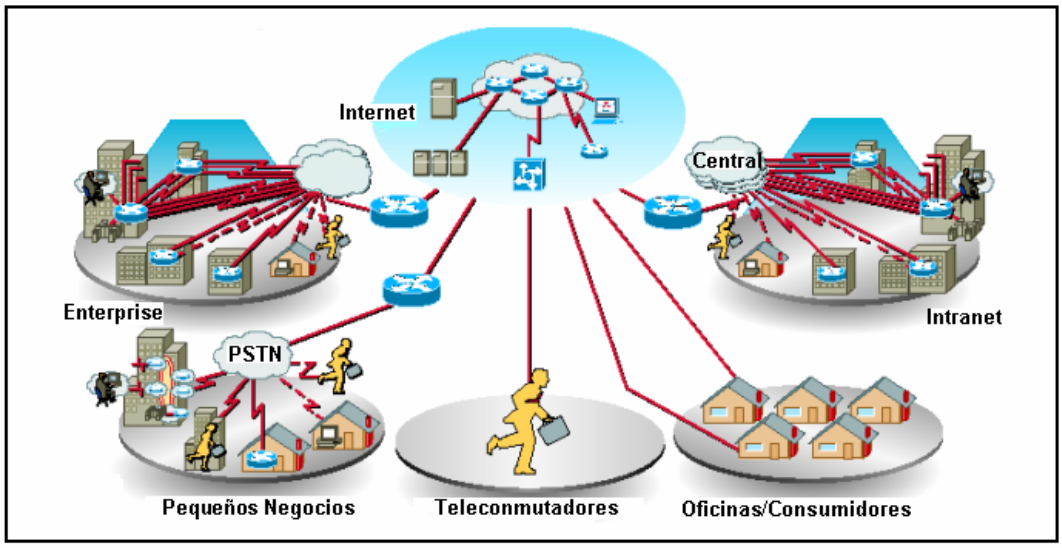


Figura 1.1. Red Multiservicio basada en IP

Pero no todas las redes de circuitos conmutados están preparadas para soportar estos nuevos servicios de la Red IP con calidad, sin embargo, mejoras en dichas redes están siendo una realidad diaria, se trata de implementar:

- La sustitución de cableado Categoría # 5 que permite llevar 100 Base T (estándar de transmisión sobre cable de par trenzado de velocidad de 100 Mbps) por un sistema de cableado Categoría # 6 y/o # 7 que nos permitirá transmitir a velocidades de 1000 Mbps (1Gbps).
- El uso de tecnología de Backbone (actúa como conductor primario del tráfico de datos de la red, comúnmente recibe y manda información a otras redes) de Modo de Transferencia Asincrónica ATM (Asynchronous Transfer Mode), Fast Ethernet o Gigabit Ethernet.
- Sustitución de equipos de networking antiguos por equipos más potentes, tales como routers y equipos de conmutación (switches) que permiten disponer de velocidades de 10/100/1000 Mbps que adicionalmente proporcionan mecanismos de Calidad de Servicio QoS que se necesitan para un transporte en tiempo real.

Adicionalmente se están observando cambios en los siguientes campos:

- Una mayor aceptación e implantación del Protocolo de Transporte de Tiempo Real RTP (Real Time Protocol) y el Protocolo de Reserva de

Recursos RSVP (Resource Reservation Protocol) en los elementos de la red.

- Una continúa proliferación de la tecnología de Red Digital de Servicios Integrados ISDN (Integrated Services Data Network) que puede ser usada para la conexion de Red de Área Ancha WAN (Wide Area Network) con el estándar H.323.

Otros aspectos importantes que se deben tener en la red IP para implantar servicios en tiempo real, es decir los pre-requisitos necesarios para la implementación de los nuevos servicios son:

- Protocolo Punto a Punto PPP (Point to Point Protocol) segmentación de paquetes para controlar retardos en la transmisión al viajar a través de líneas de baja velocidad (por ejemplo usando mecanismos PPP multilínea).
- Redes de Área Local LAN (Local Area Network) basadas en control de flujo para permitir coexistencia de datos en tiempo real y no real en un switch de conexión Ethernet.
- Manejar peticiones del Protocolo de Reserva de Recurso RSVP (Resource Reservation Protocol).
- El costo de servicio debe estar basado en el enrutamiento para las redes IP.
- Donde se conecta con la Red de Telefonía de Conmutación Publica PSTN un conmutador o switch debe soportar el Sistema de Señalización 7 SS7 (Signalling System Number 7). El Sistema de Señalización SS7 se usa eficazmente para fijar llamadas inalámbricas y alámbricas en la Red de Telefonía de Conmutación Pública PSTN y para acceder a los servidores de bases de datos de dicha red. El soporte del Sistema de Señalización SS7 en interruptores de telefonía IP representa un paso importante en la integración de las redes de telefonía PSTN y las redes de datos IP.
- Implementar un grupo de estándares de telefonía (Sistema de Señalización SS7, Estándar H.323) para que los ambientes de telefonía IP y demás redes puedan operar en conjunto con todas sus características.

CAPÍTULO II

VOZ SOBRE IP VoIP (VOICE OVER IP)

Voz sobre IP VoIP es el término genérico que la industria utiliza para referirse a la transmisión de tráfico de voz a través de una red IP. En lugar del término VoIP otros autores prefieren utilizar el de Telefonía IP IPT (IP Telephony).

Antes de seguir adentrándonos en esta novedosa tecnología conviene que aclaremos cierta confusión respecto al concepto Telefonía IP o Voz sobre IP, y es que ninguno de los dos son sinónimos de Telefonía Internet.
La telefonía Internet o los teléfonos Internet son programas (software) que una vez instalados en un ordenador permiten establecer llamadas telefónicas a través de Internet, mientras que voz sobre IP ha sido diseñado para transmitir de forma conjunta señales de audio y de datos a través de cualquier red IP, entre las cuales figura Internet.

Actualmente, en todo el mundo, la voz sobre IP junto con la telefonía móvil son los dos fenómenos que captan mayor interés dentro del mundo de las telecomunicaciones, y prueba de ello es el crecimiento experimentado en el número de usuarios que están utilizando estos servicios.
La utilización de la voz sobre IP como sustituto de la telefonía convencional se debe principalmente a:

- La reducción de costos, puesto que el tráfico de datos se cursa a precios más baratos que el tráfico de voz especialmente en llamadas de larga distancia (tarifación plana).
- Un mejor aprovechamiento de los recursos en lo que se refiere a compresión de datos y supresión de silencios, mejorando la velocidad de transmisión.
- Integración de voz y datos, es decir, lo que se desea es implementar una red multiservicio IP que tiende a homogenizar las redes.
- Aplicaciones multiservicio como son: mensajería unificada, telefonía Intranet, extensiones de la Central Telefónica Privada PBX (Private Branch Exchange) remotas.
- Mejoras en las tecnologías de Calidad de Servicio.
- Crecimiento de Internet e Intranet, con el objeto de extender las aplicaciones de negocios a los proveedores y clientes.

El servicio de voz sobre IP es nuevo en el sentido de que realmente no es un simple sustituto de los servicios existentes (telefonía tradicional).
Un punto muy importante de mencionar es que es impredecible la cantidad de nuevos servicios que pueden surgir cuando uno de los extremos de la llamada es un ordenador o PC que a su vez está sujeto a una evolución tremenda.

Ya en el mundo sé esta utilizando esta tecnología y aunque todavía esta en etapa de desarrollo y no ha logrado una norma estándar ni en equipos ni en reglamentación, existen una gran cantidad de estudios e investigaciones sobre distintos aspectos que envuelven esta tecnología.

El mayor problema que experimenta voz sobre IP VoIP es que a diferencia de los datos, la voz necesita una calidad de servicio QoS mayor para no experimentar un retardo en la transmisión (eco, superposición de la conversación “talker overlap”), variabilidad o fluctuación del retardo llamado comúnmente “jitter” y la pérdida de paquetes por lo que se necesita técnicas que minimicen estos problemas (protocolos, técnicas de compresión de datos y eliminación de silencios, introducción de buffers para eliminar el jitter) de tal forma que garanticen en la medida de lo posible la calidad de servicio QoS.

2.1 APLICACIONES DE VOZ SOBRE IP VoIP

La voz sobre IP VoIP podría ser aplicada a casi cualquier requerimiento de comunicación de voz, en un rango que va desde la simple comunicación interna en una oficina hasta complejos ambientes de teleconferencias multipunto con imagen compartida. Al mismo tiempo, es posible determinar el nivel de calidad de servicio deseado, menor para comunicaciones al interior de la empresa y mayor cuando las comunicaciones van hacia fuera de ésta, por ello, el equipo necesario para voz sobre IP debe tener la flexibilidad necesaria como para acomodarse a los distintos tipos de requerimientos y también para integrarse a los sistemas de telefonía tradicional.

Si bien, al hablar de voz sobre IP, generalmente, se hace relación a una conexión en tiempo real, este concepto es también muy adecuado para servicios de voz del tipo “almacenamiento y envío” (store and forward). Por ejemplo, los mensajes de voz pueden ser fácilmente almacenados en un servidor para ser recuperados posteriormente. Este tipo de servicios abre también un abanico de nuevas posibilidades, como el poder manejar documentos que llevan anotaciones adjuntas de voz. Algunos ejemplos de aplicaciones de VoIP son:

Interconexión de la Internet a la red PSTN: Esta interconexión pede ser alcanzada usando una gateway, ya sea integrada a una Central Telefónica Privada PBX (central para extensiones telefónicas) o puede ser también proporcionada por un elemento separado. Un teléfono basado en PC, por ejemplo, podría tener acceso a la red pública llamando a una gateway en un punto cercano al de destino (procurando minimizar la carga de larga distancia), minimizando los costos y combinando el tráfico de voz y datos.
Los teléfonos ordinarios (normales o inalámbricos) pueden ser mejorados para servir como un elemento de acceso a la Internet, tan bueno como lo provee actualmente la telefonía normal.

El servicio de directorio, por ejemplo, podría ser consultado sobre la Internet mediante la introducción de un nombre y recibir una contestación de voz o texto.

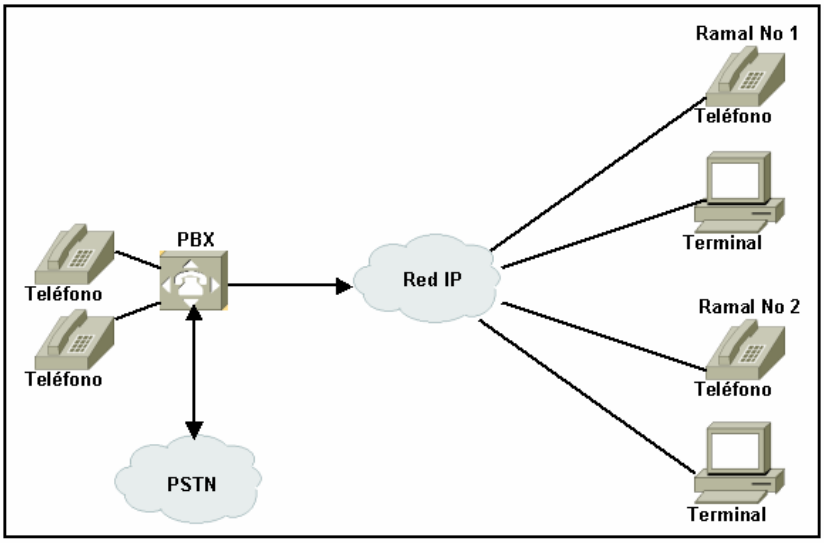


Figura. 2.1. Interconexión de la Red IP con la Red PSTN

Conexión entre oficinas sobre la Intranet Corporativa: Se llama también canalización o comúnmente conocido como “Trunking”. Esta tecnología apuntan a reemplazar las líneas troncales que enlazan las distintas centrales PBX de las sucursales de una empresa, consolidando los enlaces y el equipamiento y también proporciona una forma de acceso remota los servicios de comunicaciones de tanto de voz como de datos de una compañía emulando una extensión remota, lo cual puede resultar muy atractivo para aplicaciones de Centros de Llamadas (Call Centers), mejorando la calidad de información intercambiada. Por ejemplo, un usuario podría navegar por información en línea, antes de realizar la consulta a un operador, una vez en comunicación con el operador, ambos podrían trabajar con un documento compartido a través de la pantalla.

De esta forma se consigue sistemas de una gran calidad en el servicio a ofrecer, además de reducir de forma considerable el costo de líneas telefónicas y de los Distribuidores Automáticos de Llamadas ACD (Automatic Call Distributor).

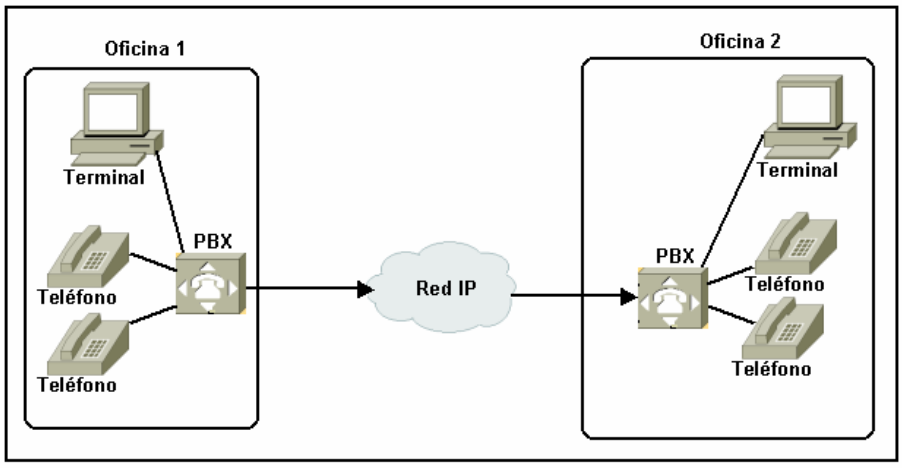


Figura. 2.2. Conexión entre oficinas mediante la Red IP

Los Centros de llamadas WEB: A través de la tecnología de red mundial WEB se trata de construir centros de atención a los clientes de tal forma que un cliente podrá estar navegando por el Internet para consultar cualquier tipo de información y a la vez se conectará automáticamente con un experto de la compañía, quien le resolverá sus dudas o preguntas acerca del producto o sitio consultado. Esta aplicación tiene las siguientes ventajas:

- Al ser la llamada a través de Internet, para el usuario no tiene costo adicional, aprovecha la llamada telefónica que tenía establecida para la comunicación de datos para de esta forma mantener también la comunicación de voz, esto permite a la empresa tener un servicio similar al de las líneas gratuitas 1-800, 1-700.
- El usuario puede mantenerse navegando (en línea) mientras habla con un operador de ventas, haciendo el sistema muy flexible y atractivo.
- El cliente trata con operadores humanos que le podrán asesorar, esta característica mejorará sin lugar a dudas el resultado de un sistema de comercio electrónico haciéndolo más humano.
- El operador puede concretar la venta de manera más fácil, ya que el usuario es bastante reacio a dar los datos de su tarjeta de crédito en una pagina Web por los problemas de seguridad que todos conocen, sin embargo no tendrá ningún inconveniente de dar esos datos verbalmente al

operador de ventas, teniendo el usuario plena garantía de que sus datos están a salvo.

Interconexión o Interworking con Redes Celulares: Los datos de voz en una red celular digital ya están comprimidos y empaquetados para la transmisión sobre aire mediante el teléfono celular. Luego la red IP puede transmitir el paquete de voz celular comprimido y pueden ahorrar una tremenda cantidad de ancho de banda. Una gateway proporciona la función de transcodificación requerida para convertir los datos de voz celular al formato requerido por la red PSTN, de esta forma la red IP hace de puente entre la red celular y la red de telefonía tradicional minimizando de forma considerable el costo, el tiempo y el ancho de banda.

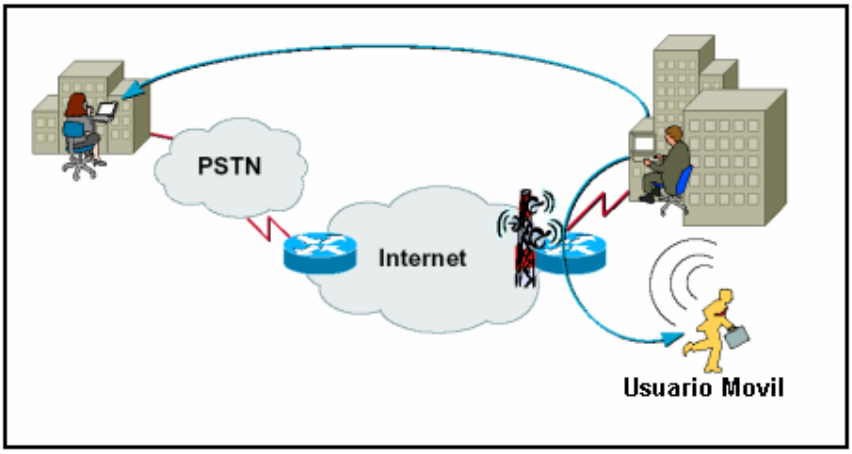


Figura. 2.3. Interworking con la Red Móvil

Doblador de Línea: En la actualidad cuando un usuario se conecta al Internet a través de conexión vía módem por línea analógica, la línea telefónica permanece ocupada durante toda la sesión de navegación, lo que impide la recepción o establecimiento de nuevas llamadas telefónicas. Mediante voz sobre IP, la gateway podrá a la vez que se mantiene la sesión de Internet y enviar una señal al usuario para que éste acepte o rechace las llamadas telefónicas procedentes

del exterior. Y viceversa, el usuario podrá activar el software de teléfono IP instalado previamente en su PC para que en un momento determinado se pueda efectuar una llamada telefónica al exterior a la vez que sigue navegando por la red.

Mensajería Unificada: Voz sobre IP permite la implementación de un sistema que ofrece un tratamiento conjunto a los mensajes de voz y de correo electrónico desde una única aplicación, facilitando enormemente las comunicaciones actuales.

La voz sobre IP VoIP está llamada a causar una revolución en los servicios de telecomunicaciones. Una nueva generación de empresas está surgiendo para dar servicios de telefonía a través de Internet. Los llamados Proveedores de Servicio Telefónico de Internet ITSP (Internet Telephony Service Provider) ya han comenzado a inquietar el mercado con interesantes propuestas de telefonía internacional teléfono a teléfono, PC a teléfono y PC a PC, a muy bajos precios aprovechando sus enlaces Internet. Estas incipientes iniciativas a nivel local, son sólo un ejemplo de la factibilidad real de la voz sobre IP como sistema masivo de comunicaciones y representan la punta del iceberg que está por venir y que presenta una multitud de ventajas en todos los aspectos, especialmente en el económico.

2.2 COMPONENTES DE UNA RED DE VoIP

Una red de voz sobre IP VoIP puede ser implementada completamente mediante cuatro elementos principales que son:

- Terminales de usuario
- Gateway
- Gatekeeper

- Unidad de Control Multipunto MCU (Multipoint Control Unit)

Terminales de usuario: En la actualidad los usuarios cuentan con dos posibilidades, la primera es utilizar el teléfono convencional y la segunda equipar sus PCs con un software específico que les permitirá la comunicación directa a través de la red IP. Así actualmente aparecen los teléfonos IP, que son terminales inteligentes, los cuales permiten convertir la señal analógica de telefonía en un conjunto de paquetes de datos IP, o los vídeo-teléfonos IP, que además de sonido también transmiten vídeo.

Tampoco se deberá esperar mucho tiempo para poder utilizar sistemas o terminales universales IP ya sean en formato de teléfono móvil, PC o TV, lo cual nos abrirán las puertas al mundo de las comunicaciones multimedia a través de redes IP.

Gateway: Es el elemento que permite la conversión de tráfico de voz procedente de una red IP en tráfico de voz de otra arquitectura de red y viceversa. Dada la importancia, la capacidad, el rendimiento, la calidad, que representa la Gateway en una solución voz sobre IP, ésta debe reunir ciertas características imprescindibles para proporcionar un correcto funcionamiento, entre las cuales podemos mencionar las siguientes:

- Los módulos que reciben las llamadas de los terminales del usuario deberán estar equipados con tecnología de Procesador de Señal Digital DSP (Digital Signal Processor) para distribuir las funciones de proceso de las llamadas y así proporcionar al sistema el máximo rendimiento.
- Debe tener un soporte de todos los codificadores y decodificadores, generalmente llamados “Codecs” definidos como estándares.
- Debe tener un soporte de señalización y procesamiento de Tono Dual de Multifrecuencia DTMF (Dual Tone Multi Frequency).
- Debe ser capas de manejar alta densidad de llamadas.

El gatekeeper: Este elemento está principalmente diseñado para controlar las conexiones a través de la red. La admisión de llamadas, la translación de direcciones, la seguridad y el control del uso del ancho de banda, son funciones generalmente desempeñadas por los gatekeepers. Además localiza a las distintas gateways y unidades de control multipunto MCU cuando se necesitan. De esta forma se tiene un control de los accesos, seguridad, movilidad del usuario y tarifación si se da el caso.

Unidad de Control Multipunto (MCU): Las unidades de control MCU permiten el establecimiento de conferencias entre dos o más usuarios. Es responsable de controlar las sesiones y de efectuar la mezcla de los flujos de audio, datos y video.

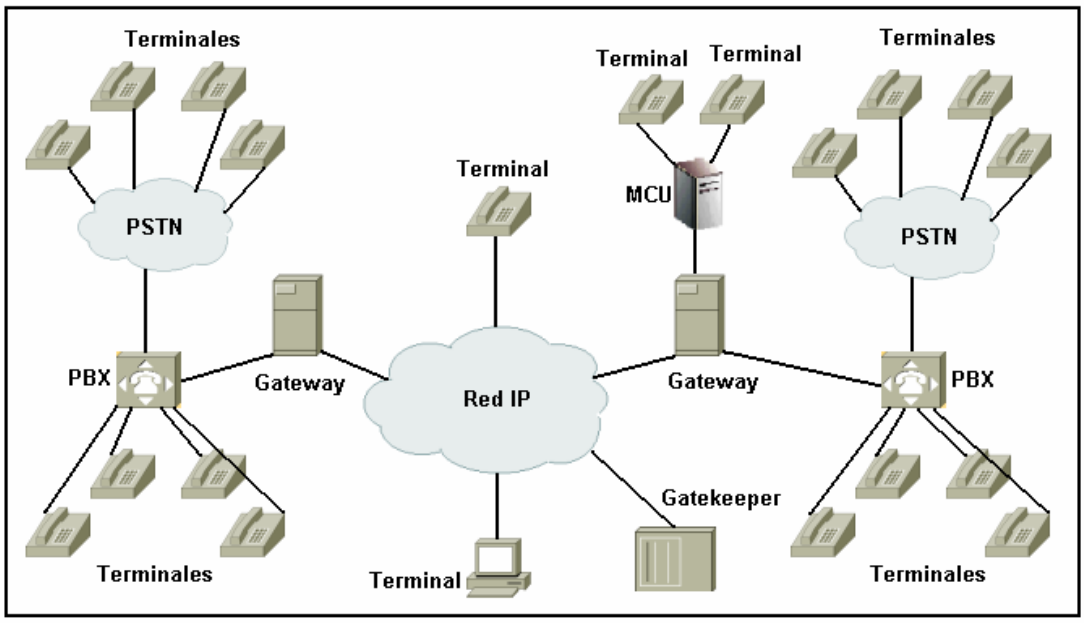


Figura. 2.4. Elementos de una Red de VoIP

Los distintos elementos pueden residir en arquitecturas físicas separadas, o se pueden encontrar con varios elementos conviviendo en la misma arquitectura. De este modo es bastante habitual encontrar juntos Gatekeeper y Gateway.

2.3 CALIDAD DE SERVICIOS DE VoIP

La transmisión de paquetes de voz es similar a la transmisión de un correo electrónico desde el origen hasta el destino. El problema es que en las transmisiones de voz sobre la red IP no está garantizado el éxito de la transmisión, por lo cual a diferencia del correo electrónico cuando este no es legible o se pierde algún paquete, simplemente se solicita la retransmisión del mismo y su recuperación es factible, en el caso de la transmisión de voz esto no es así, dada la necesidad de recibir los paquetes en un determinado orden, la necesidad de asegurar que no haya pérdidas y de conseguir una tasa de transmisión mínima, hacen prácticamente necesaria la implantación de sistemas de calidad de servicio QoS.

Numerosos métodos son usados para superar el ambiente hostil de la red IP y proveer una calidad de servicio aceptable, los aspectos que se quiere mejorar con estos métodos son:

- Retardo o también llamado Latencia
- Variabilidad o fluctuación del retardo “Jitter”
- El Eco
- La congestión
- Pérdida de paquetes
- Errores en la secuencia

Como otras aplicaciones en tiempo real, VoIP debe tener un ancho de banda establecido y es sumamente sensible al retardo por lo una buena ingeniería en la red de terminal a terminal es necesaria para usar satisfactoriamente esta tecnología, en cuento se referente a la calidad de la transmisión de voz, todos los fabricantes y las investigaciones hacen referencia fundamentalmente a tres factores determinantes:

- **Codificadores/decodificadores de Voz:** Los cuales influyen en la digitalización de la voz en paquetes de datos que contienen voz y que serán transmitidos por la red IP, también influyen en el retardo necesario para la descompresión de esos paquetes voz, lo que atribuye muchas veces un retardo añadido a la comunicación.
- **Cancelación del Eco:** Es el requerimiento necesario para una comunicación, ya que elimina de forma automática y en tiempo real posibles ecos, si no lo hiciera así, haría incomprensible la comunicación.
- **Retardo o Latencia:** Es el tiempo necesario para que la voz viaje de un extremo al otro, incluyen los tiempos necesarios para la compresión, transmisión y descompresión. Este tiempo tiende a minimizarse pero jamás podrá ser suprimido. Actualmente los tiempos que se están obteniendo de retardo o latencia giran alrededor de 120 ms.

Para que VoIP sea un reemplazo realista de la red de telefonía PSTN los clientes necesitan recibir la misma calidad de transmisión de voz que reciben con los servicios de teléfono tradicional.

Es decir, las transmisiones de VoIP deben ser inteligibles al receptor, no deben perderse paquetes de voz o estar excesivamente retardados o sufrir una variación de retardo alta. Con el fin de poder alcanzar estos objetivos, se deben cumplir los siguientes requerimientos (claves para hacer pruebas de VoIP):

Retardo o Latencia: Coma ya se menciono anteriormente, la latencia es el tiempo de retraso incurrido en un diálogo. El exceso del retardo hace la conversación inconveniente y no natural. Cada componente en la ruta de la transmisión (emisor, red, receptor) aumenta el retardo en la misma.

La recomendación TG.114 (Tiempo de Transmisión Unidireccional) de la ITU recomienda 150 ms. como máximo deseado de latencia unidireccional para conseguir voz de alta calidad de extremo a extremo.

Y para las llamadas internacionales, el retardo de 300 ms. es aceptable, sobre todo para la transmisión de satélite.

Retardo Unidireccional	Descripción
0 a 150 ms.	Aceptable para la mayoría de las aplicaciones
150 a 400 ms.	Aceptable, provee que los administradores estén alertas del impacto del tiempo de la transmisión en la calidad de las aplicaciones de usuario.
Más de 400 ms.	Inaceptable para los propósitos de planificación de la red, sin embargo, este límite en algunos casos excepcionales puede ser excedido, por ejemplo, en conexiones con satélite.

Tabla. 2.1. Recomendaciones de Retardo de la ITU

El codificador/decodificador (codec) por defecto recomendado por la ITU es: G.729, cuyo retardo normal es de 125 ms., a lo cual se debe adicionar el retardo variable que dependerá del enlace y de la carga de la red.

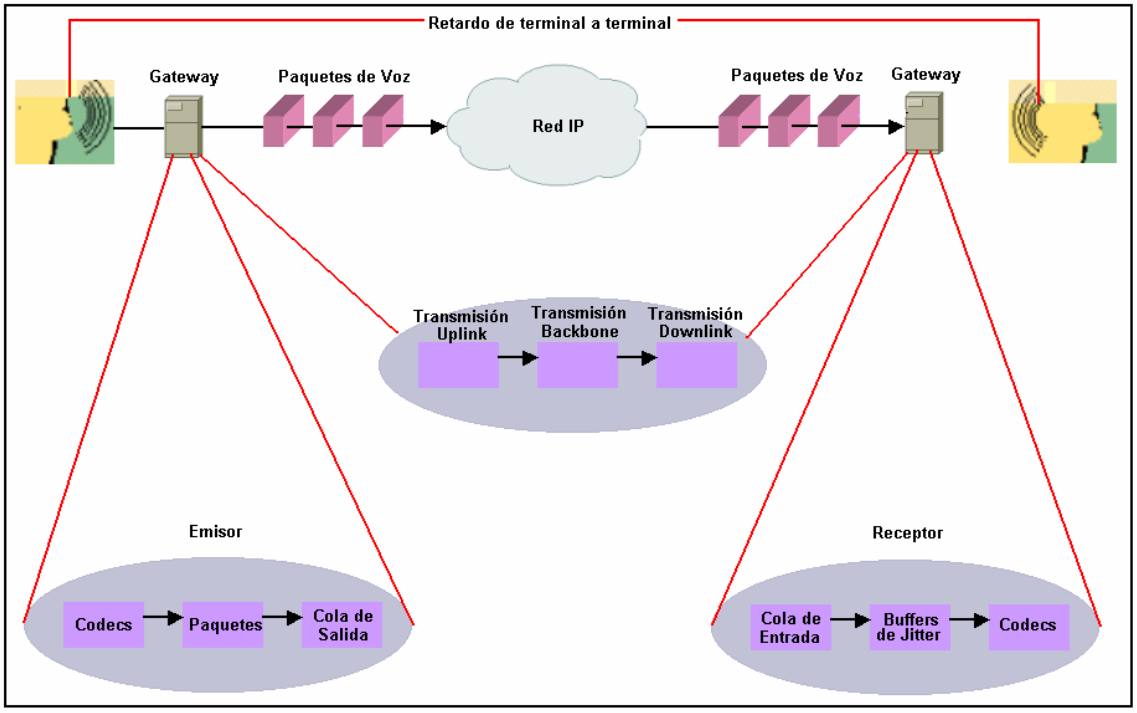


Figura. 2.5. Retardo de Terminal a Terminal

En el gráfico de la figura 2.6 se muestra como la calidad de voz se deteriora con el incremento de la carga de la red, cuando no se han implementado características de calidad de servicio QoS. Mientras que con características de calidad de servicio, la calidad de voz puede ser conservada a pesar de la carga de tráfico de la red.

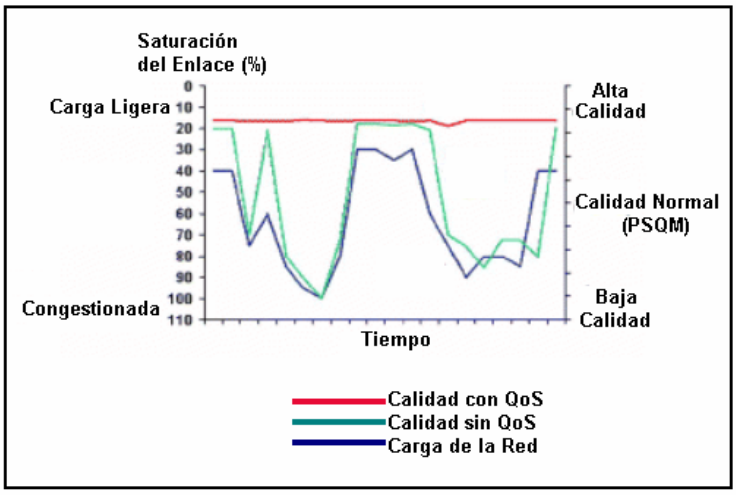


Figura 2.6. Gráfico de Saturación del Enlace vs. Calidad de Servicio

El Eco: El eco incluso está presente en una red de servicio telefónico tradicional, sin embargo, es aceptable cuando el retardo es menor a 50 ms. y el eco es enmascarado por el tono lateral normal que cada teléfono genera.

El eco realmente se vuelve un problema en VoIP porque el retardo casi siempre es mayor que 50 ms. Así, que se debe usar técnicas de cancelación de eco.

Los estándares G.165 y G.168 de la ITU definen los requisitos de la actuación para los canceladores de eco.

La variabilidad o latencia "Jitter": Esta característica cuantifica los efectos de los retardos de los paquetes en la red que llegan al receptor. Los paquetes transmitidos en intervalos iguales desde la gateway del emisor hasta el arribo en la gateway del receptor en intervalos irregulares. La variabilidad o Jitter excesivo hace la conversación cortada y dificultosa de entender.

El Jitter es calculado basado en el tiempo de arribo promedio de los paquetes. Para una alta calidad de la voz, el promedio de arribo al receptor debería ser aproximadamente igual al intervalo de los paquetes en el receptor y la desviación estándar debería ser baja. Con el fin de mejorar esta característica de calidad de servicio QoS se utilizan los buffers de Jitter para contrarrestar los efectos de las fluctuaciones de la red y crear un flujo normal de los paquetes en el terminal receptor, el cual es eficaz en variaciones de retardo menor a 100 ms.

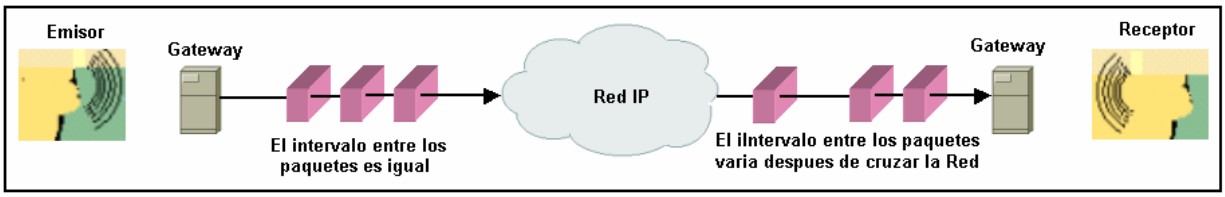


Figura. 2.7. Variabilidad o Jitter

Pérdida de Paquetes: Generalmente ocurre tanto en forma abrupta o en forma periódica debido la congestión de la red. La pérdida periódica no puede estar por encima del 5% al 10% de todos los paquetes de voz transmitidos ya que se degradaría la calidad de voz significativamente.

La pérdida abrupta de paquetes puede también hacer dificultosa la conversación, aunque no es muy frecuente que suceda.

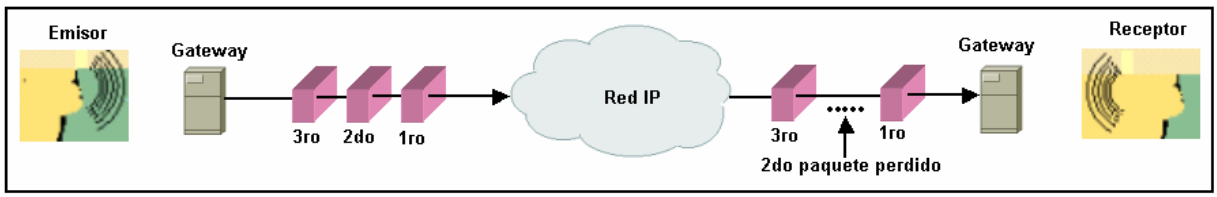


Figura. 2.8. Pérdida de Paquetes

Errores en la secuencia: La congestión en las redes IP puede causar que los paquetes tengan diferentes rutas para encontrar el mismo destino. Los paquetes pueden entonces arribar en desorden y distorsionar la conversación, pero al igual que en la pérdida de paquetes esta falta de secuencia no debe ser mayor al 5 %.

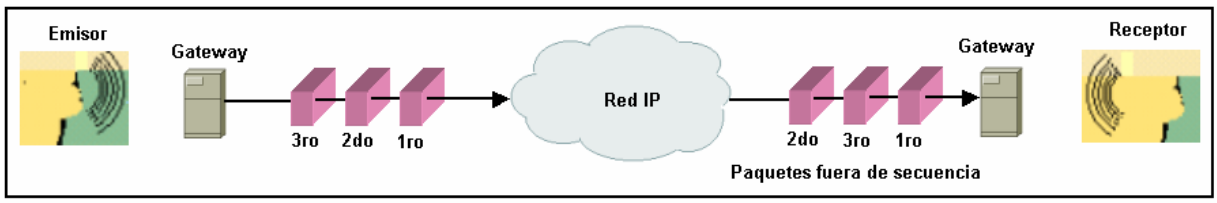


Figura. 2.9. Errores en la secuencia de los Paquetes

En general se puede decir que: VoIP sólo puede garantizar la transmisión de voz de alta calidad solo si los paquetes de voz, para la señalización y los canales de audio, tienen prioridad sobre de otro tipo de tráfico, además, debe garantizarse cierto ancho de banda y requerimientos de retardo o latencia y jitter.

2.4 CARACTERÍSTICAS PARA PROPORCIONAR QoS EN VoIP

La calidad de servicio QoS debe proporciona un mejoramiento en la red para proveer las siguientes características, que son:

- Suporte de Ancho de banda dedicado
- Mejoramiento de las características de pérdida de paquetes
- Evitar y manejar la congestión de la red
- Proporcionar un ordenamiento del tráfico de la red
- Establecimiento de las prioridades del tráfico por la red

La calidad de servicios para voz sobre IP discute varios conceptos y características, entre ellas tenemos:

- Ancho de Banda suficiente

- Clasificación de paquetes
- Mecanismos de Cola de espera
- Fragmentación e Interpolación
- Formación del Tráfico
- Compresión del encabezamiento RTP
- Servicios Diferenciados para VoIP
- Protocolo de Reserva de Recurso RSVP

2.4.1 Ancho de Banda suficiente

Antes de considerar aplicar cualquiera de las características de calidad de servicio, QoS, debe primero proporcionarse el ancho de banda necesario a la red para soportar el tráfico de voz en tiempo real.
Por ejemplo, una llamada de VoIP de G.711 de 80 Kbps (64Kbps de carga más 16 Kbps encabezamiento) será pobre sobre un enlace de 64 Kbps porque por lo menos 16 Kbps de los paquetes (20%) se perderán (se asume que ningún otro tráfico está fluyendo sobre el enlace).

Así que se debe proporcionar el ancho de banda suficiente para tráfico de voz, lo cual es de vital importancia para mejorar las características de QoS y la determinación del mismo depende entre otras cosas del codec utilizado y el número de muestras por paquete. Una forma de optimizar el uso del ancho de banda es:

- Utilizar el codec adecuado con una alta factibilidad de compresión.
- Minimizar el tamaño de las cabeceras que encapsulan los datos usando protocolos como: el Protocolo RTP Comprimido cRTP (Compress Real Time Protocol), el cual reduce los 40 bytes de cabecera IP/UDP/RTP a 2 o 4 bytes, permitiendo un ahorro del ancho de banda.
- Reservar el ancho de banda necesario para aplicaciones en tiempo real mediante el Protocolo de Reserva de Recurso RSVP.

2.4.2 Clasificación de Paquetes

La base para proporcionar cualquier QoS se basa en la habilidad de un dispositivo de la red para identificar y agrupar paquetes específicos. Este proceso de identificación se llama **clasificación de paquetes**.

Después de que un paquete ha sido clasificado, el paquete necesita ser marcado para establecer los bits designados en el encabezamiento IP.
Los dispositivos de la red usan la dirección de origen y de destino IP en el encabezamiento de IP o los números de puertos del Protocolo de Datagrama de Usuario UDP (User Datagram Protocol) de origen y de destino en el encabezamiento de UDP para identificar los paquetes de VoIP, este proceso se conoce como **clasificación estática**.

Además de la clasificación estática que involucra la información de los encabezamientos de la capa de red o capa de transporte del modelo OSI, se puede usar un mecanismo como el Protocolo de Reservación de Recurso RSVP para una **clasificación dinámica**.

El Protocolo RSVP usa paquetes de señalización H.245 para determinar cual puerto del Protocolo UDP de voz usar para la conversación. Entonces se establece una lista de acceso dinámico para identificar el tráfico de VoIP y los lugares del tráfico dentro de una cola reservada.

La clasificación del paquete puede ser un proceso intensivo, que debe ocurrir en el borde de la red. Porque cada punto de conexión entre los terminales de la red llamados “Hop”, a través de los cuales pasa la información de un punto al siguiente, todavía necesita hacer una determinación en el tratamiento que un paquete debe recibir, se necesita tener un método de clasificación más simple y eficaz en el centro de la red. Esta clasificación más simple se logra a través de

marcar o determinar el byte de Tipo de Servicio ToS (Type of Service) en el encabezamiento de IP.

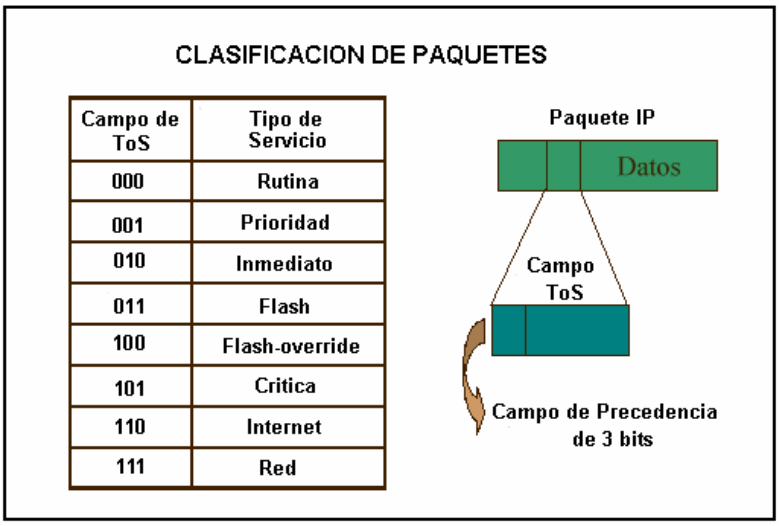


Figura. 2.10. Campo del Byte de Tipo de Servicio ToS

Los tres bits más significativos del byte de tipo de servicio ToS son llamados los bits de “precedencia IP”, la mayoría de las aplicaciones y proveedores de servicios determinan y reconocen estos 3 bits.

El marcando esta evolucionando a fin de que los seis bits más significativos del byte de Tipo de Servicio ToS, llamado Punto de Código de Servicios Diferenciados DSCP (Differentiated Services Code Point), pueda ser usado para definir las clases de Servicios Diferenciados DS (Differentiated Services).

Después que cada punto de conexión en la red es habilitado para clasificar e identificar los paquetes de VoIP, tanto con la información de dirección de red como con el byte ToS, estos puntos de conexión pueden proporcionar a cada paquete de VoIP con los requerimientos de QoS. En este punto, se debe establecer técnicas especiales para proporcionar la cola de prioridad asegurando que los paquetes largos de datos no interfieran con la transmisión de datos de

voz, y reducir los requerimientos del ancho de banda por la compresión de los 40 bytes del encabezamiento IP/UDP/RTP a 2 o 4 bytes.

Los dispositivos de la red usan varios criterios para poder clasificar al tráfico en un cierto número de clases. Por consiguiente los nodos deberían marcar los paquetes tan rápido como ellos hayan sido identificados y clasificados.
Si un nodo puede fijar el bit de precedencia IP o bits de Punto de Código de Servicios Diferenciados DSCP en el byte de ToS del encabezamiento de IP tan rápido como identifique el tráfico como tráfico de VoIP, entonces todos los otros nodos de la red pueden realizar la clasificación basados en estos bits.
El marcado es el proceso del nodo de fijar los siguientes bits:

- Tres bits de Precedencia IP en el byte ToS IP.
- Seis bits de DSCP en el byte ToS IP.
- Tres bits de Conmutación de Etiquetas Multiprotocolo MPLS (Multiprotocol Label Switching).
- Tres bits Ethernet 802.1p de Clase de Servicio CoS (Class of Service).
- Un bit de Probabilidad de Pérdida de Celda (CLP Cell Loss Probability).

En la mayoría de las redes de IP, el marcando de los bits de precedencia IP o bits de DSCP debe ser suficiente para identificar el tráfico como tráfico de voz sobre IP.

2.4.3 Mecanismo de Cola de espera

Varios algoritmos y protocolos priorizan y localizan los recursos para limitar la trama de datos cuando un paquete de voz esta presente en la red a través de mecanismos de cola de espera normales, entre los cuales se enlistan los siguientes:

Mecanismo de Cola de espera "Primero en entrar, primero en salir" FIFO (First-in, first-out): En el cual los paquetes llegan y salen de la cola de espera en

exactamente el mismo orden. La configuración es simple y rápida pero este mecanismo no proporciona servicios de prioridad o ancho de banda garantizado.

Mecanismo de Cola de espera de Prioridad PQ (Priority Queuing): El tráfico es clasificado en colas de espera de prioridad alta, media, normal y baja, siendo el tráfico de prioridad alta servido primero y así sucesivamente. La desventaja de este mecanismo es que el tráfico de prioridad alta puede hambrear el ancho de banda de la cola de prioridad más baja, es decir no existe ancho de banda garantizado

Mecanismo de Cola de espera de Cliente CQ (Custom Queuing): El tráfico es clasificado en colas de espera múltiples con limites de cola de espera configurable. Los límites de la cola de espera son calculados basados en tamaño promedio del paquete, la Unidad de Transmisión Máxima MTU (Maximum Transmission Unit) y el porcentaje de ancho de banda a ser asignado. Los límites de la cola de espera (en números de bytes) son adecuados a cada cola y proporcionan por consiguiente un ancho de banda asignado. La limitación de esta clase cola de espera es que ningún servicio de prioridad es posible y la configuración es relativamente difícil.

Actualmente, después de que todo el tráfico ha sido clasificado en clases basadas en los requisitos de la calidad de servicio QoS, se utilizan mecanismos de cola de espera de rendimiento inteligente, para proporcionar un ancho de banda garantizado y prioridad de servicio a través de red IP, estos mecanismos son:

- Cola de espera de Carga Exacta WFQ (Weighted Fair Queuing)
- Cola de espera de Retardo Bajo LLQ (Low Latency Queuing)

2.4.3.1 Mecanismo de Cola de espera Carga Exacta WFQ (Weighted fair Queuing)

La utilización de este mecanismo de control de congestión esta directamente asociado al retardo causado por la cola de espera o cola de retardo. Con su utilización, no importará la cantidad de tráfico de datos entrante porque existirá una línea de paquetes de datos y otra de paquetes de voz siendo cada paquete serializado por tiempo. De este modo, la voz siempre tendrá un lugar garantizado en la transmisión.

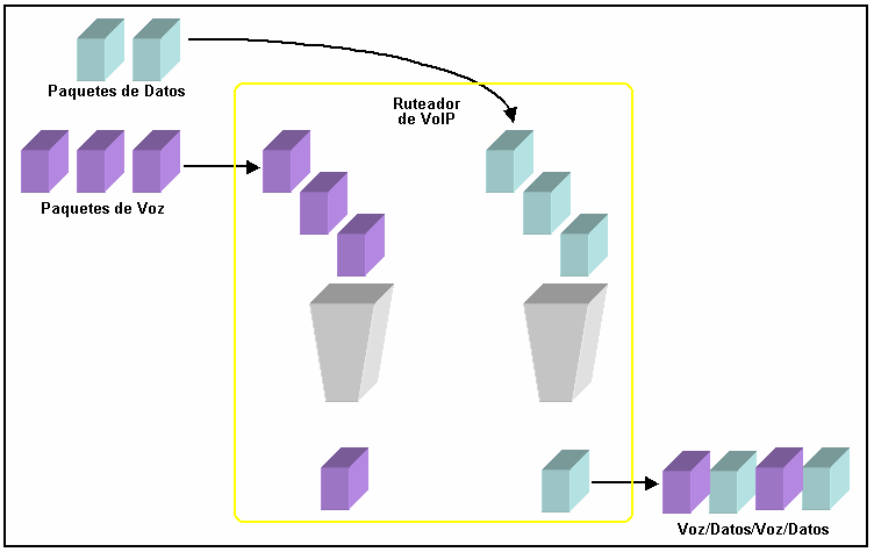


Figura. 2.11. Cola de Carga Exacta WFQ

Este mecanismo también atiende otro factor necesario para el transporte de voz: Prioridad del tráfico de voz frente al tráfico de datos, además minimiza y controla el retardo, esto es necesario para garantizar la entrega de los paquetes de voz.

El protocolo RSVP, el cual es utilizado para garantizar que existirá un ancho de banda mínimo para la comunicación entre el emisor y receptor, es usado conjuntamente con el mecanismo de la cola de espera WFQ por lo que la serialización de los paquetes es programada y guiada por los requisitos del protocolo RSVP.

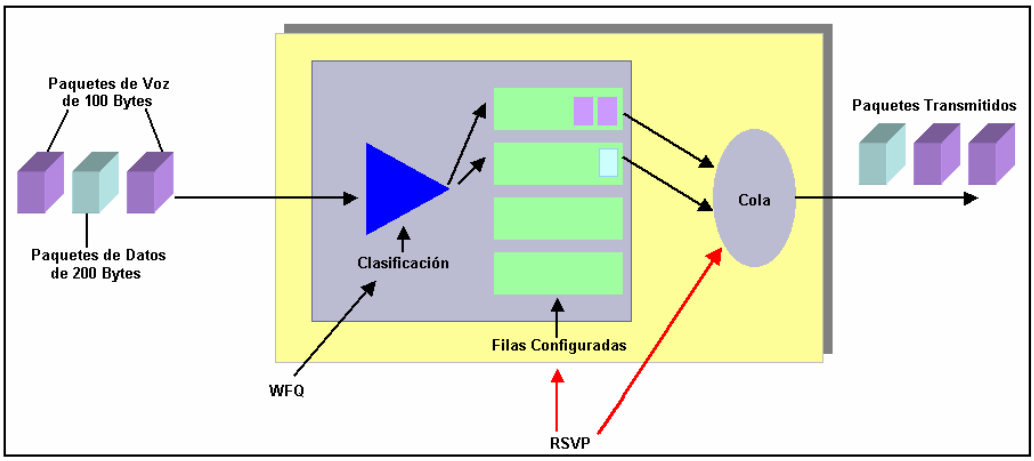


Figura. 2.12: Clasificación de Paquetes con WFQ y RSVP

2.4.3.2 Mecanismo de Cola de espera de Retardo Bajo LLQ (Low Latency Queuing)

Una cola de prioridad se requiere para VoIP, se puede usar cualquier mecanismo del cola eficazmente que proporcione prioridad alta a VoIP, pero la cola de espera de retardo bajo LLQ es la más recomendada porque es flexible y fácil de manejar. La cola de retardo bajo LLQ usa el método de configuración de Control de Cola de espera de Mensajes MQC (Message Queuing Control) para proporcionar prioridad a ciertas clases y mantener el ancho de banda mínimo garantizado para otras clases.

Durante los períodos de congestión, la cola de prioridad es monitoreada en la tasa configurada para que el tráfico de prioridad no monopolice todo el ancho de banda disponible (esto impediría garantizar del ancho de banda para otras clases). Si se usa la cola de espera LLQ correctamente, el tráfico entrante en la cola de prioridad nunca debería exceder la tasa configurada.

El mecanismo de LLQ también permite colas de espera de profundidad ser especificadas para determinar cuando el router debe descartar los paquetes si son demasiados en espera en cualquier clase de cola.

Hay también una clase predefinida que se usa para determinar el tratamiento de todo el tráfico no clasificado por una clase configurada.

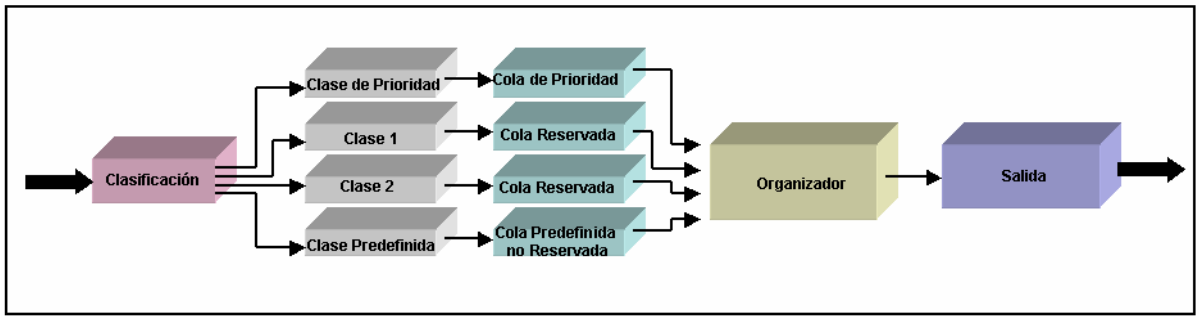


Figura. 2.13. Funcionamiento de la Cola de Retardo Bajo LLQ

En la figura, todo el tráfico saliente de una interfase o subinterfase es clasificado usando Control de Cola de espera de Mensajes MQC primero. Hay cuatro clases:

- Una clase de prioridad alta
- Dos clases de ancho de banda garantizado
- Una clase predefinida

El tráfico de clase de prioridad alta se coloca en una cola de espera de prioridad y el tráfico de las clases de ancho de banda garantizado se coloca en las colas reservadas.

El tráfico de la clase predefinida puede colocarse en una cola reservada o puede ponerse en una cola predefinida no reservada, donde cada flujo conseguirá aproximadamente una parte igual de ancho de banda no reservado y disponible.

El servicio del organizador de colas es proporcionar las facilidades respectivas de tal forma que el tráfico de cola de prioridad sea el primero en salir, a menos que exceda el ancho de banda de prioridad configurado y este ancho de banda sea requerido por una cola reservada (es decir, hay congestión). Las colas reservadas son servidas según su ancho de banda reservado que el organizador usa para calcular la carga.

La carga se usa para determinar cuan a menudo una cola reservada es servida y cuantos bytes son servidos al mismo tiempo.
Los servicios del organizador son basados en el algoritmo de cola de espera de carga exacta WFQ.

Si las colas de prioridad están llenas porque la tasa de transmisión de tráfico de prioridad es mayor que el ancho de banda de prioridad configurado, los paquetes al final de la cola de prioridad serán desechados solo si ningún ancho de banda no reservado no está disponible.
Ninguna de las colas reservadas se restringe al ancho de banda configurado si un mayor ancho de banda está disponible.
Los paquetes que violan el ancho de banda garantizado y prioridad son desechados durante los períodos de congestión.

2.4.3.3 Otros Mecanismos de Cola de espera de QoS

Muchos otros métodos de cola de espera están disponibles y pueden usarse para proporcionar una prioridad alta al tráfico de VoIP.

Mecanismos de Cola	Descripción	Beneficios	Limitaciones
Clase basada en WFQ CBWFQ (Class-Based WFQ)	El Control de Cola de Mensaje MQC es usado para clasificar el tráfico, luego el tráfico es colocado en colas de ancho de banda reservado o en colas no reservadas asignadas. El organizador de servicios de colas esta basado en la carga garantizando el ancho de banda.	Similar a la cola de espera LLQ excepto que no hay cola de prioridad. La configuración es muy simple y además proporciona un ancho de banda garantizado.	No hay servicio de prioridad.
Cola de Prioridad WFQ PQ-WFQ, (Priority Queue WFQ)	También llamada prioridad IP RTP. Un simple comando de interfase es usado para proporcionar el servicio de prioridad en todos los paquetes del protocolo UDP destinados a los números de puerto sin un rango específico.	Proporciona el servicio de prioridad a los paquetes del protocolo RTP.	El tráfico restante es tratado con WFQ. El tráfico de RTCP no es priorizado. No hay capacidad de garantizar el ancho de banda.

Tabla. 2.2. Mecanismos (Software) disponibles de cola de espera

2.4.4 Fragmentación e Interpolación

Porque las transmisiones de VoIP son sumamente sensibles al retardo, deben interpolarse paquetes de VoIP o deben insertarse entre los fragmentos de paquete de datos. Aun cuando la cola de espera esta trabajando muy bien y priorizando el tráfico de voz, hay momentos cuando la cola de espera de prioridad esta vacía y un paquete de otra clase se repara.

Si un paquete de voz de prioridad llega a la cola de salida que mientras estos paquetes están reparándose, el paquete de VoIP podría esperar una

cantidad sustancial de tiempo antes de enviarse. Se puede asumir que un paquete de VoIP necesitará esperar detrás de un paquete de los datos y que el paquete de datos puede ser, a lo sumo, de igual tamaño a la unidad MTU (1500 bytes para interfase serial y 4470 bytes para las interfases serial de gran velocidad), así se puede calcular el tiempo de espera basado en velocidad del enlace. Por ejemplo, para una velocidad de enlace de 64 Kbps y tamaño de unidad MTU de 1500 bytes, tenemos:

$$\text{Retardo de Serialización} = \frac{(1500 \text{ bytes} * 8 \text{ bits/byte})}{(64{,}000 \text{ bits/sec})} = 187.5 \text{ ms}$$

Por consiguiente, un paquete de VoIP puede necesitar esperar a 187.5 ms antes de que pueda enviarse si se encuentra detrás de un solo paquete de datos de 1500 bytes en un enlace de 64 Kbps.

Normalmente se envían paquetes de VoIP cada 20 ms., con un retardo de extremo a extremo de 150 ms. Pero se necesita un mecanismo para asegurar que el tiempo de transmisión sea menos de 10 ms. Cualquier paquete que tiene más de 10 ms de retardo de serialización necesita ser fragmentado en pedazos cortos o fragmentos de 10 ms. Un fragmento de 10 ms. es el número de bytes que pueden enviarse sobre un enlace en 10 ms. Se puede calcular el tamaño del fragmento usando la velocidad del enlace, como se indica a continuación:

$$\text{Tamaño de fragmentación} = \frac{(0.01 \text{ segundos} * 64{,}000 \text{ bps})}{(8 \text{ bits/byte})} = 80 \text{ bytes}$$

Toma 10 ms. para enviar un paquete de 80 bytes o fragmento sobre un encale de 64 Kbps.

En enlaces de velocidad bajas donde un paquete de 10 ms es más pequeño que la unidad MTU, la fragmentación es necesaria.

Pero la fragmentación por si sola es insuficiente, porque si el paquete de VoIP debe esperar detrás de todos los fragmentos de un solo paquete de los datos grande, se tardará más allá del límite de retardo de extremo a extremo.

El paquete de VoIP debe interpolarse o debe insertarse entre los fragmentos de paquete de datos.

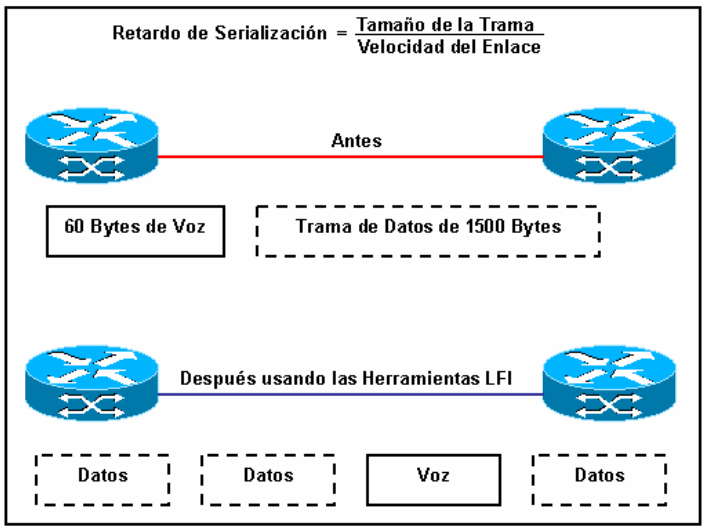


Figura. 2.14. Fragmentación e Interpolación del paquete de VoIP

La tabla 2.3 muestra los tamaños recomendados de los fragmentos para varias velocidades de enlaces basados en la regla de 10 ms.

Velocidad del Enlace (Kbps)	Tamaño de fragmento (Bytes)
56	70
64	80
128	160
256	320
512	640
768	960
1024	1280
1536	1920

Tabla. 2.3. Tamaño de fragmentación y velocidad del enlace

NOTA: La fragmentación no es requerida si el tamaño del fragmento es más largo que el tamaño del enlace de la unidad MTU. Por ejemplo, para un enlace T1 (Línea de Transmisión implementada por AT&T con una velocidad de 1.544

Mbps) con una MTU de 1500 bytes, el tamaño del fragmento es 1920 bytes, así la fragmentación no es requerida.

Además el tamaño de fragmentación de paquete de datos nunca debe ser menor que el tamaño del paquete de VoIP y nunca debe fragmentarse paquetes de VoIP en porque se puede causar establecimientos de numerosas llamadas y problemas de calidad.

2.4.5 Formación del Tráfico

La formación de tráfico es un mecanismo de QoS usado para enviar tráfico en ráfagas cortas en una tasa de transmisión configurada.
La formación del Tráfico nos proporciona las siguientes características:

- Control del uso del ancho de banda disponible
- Establecimiento de mediciones del tráfico
- Regularización del flujo del tráfico para evitar congestión
- Limitación de la cantidad del tráfico y de la perdida de paquetes
- Priorización del tráfico en tiempo real

Los escenarios normalmente están compuestos de un hub y un emisor de red (spoke), donde la velocidad de enlace del hub es más alta que cualquiera de las velocidades de los enlaces remotos.

En algunos casos, la suma de las velocidades del los enlaces remotos es más alta que la velocidad de enlace de hub y esto causa sobresuscripción. Sin la formación del tráfico, el hub podría intentar enviar tráfico en altas tasas a los sitios remotos causando a la red una eliminación de tráfico arbitrariamente.

Por consiguiente, se necesita controlar tasas de transmisión en la red para que se pueda controlar que paquetes se desechan y que paquetes reciben servicio de prioridad.

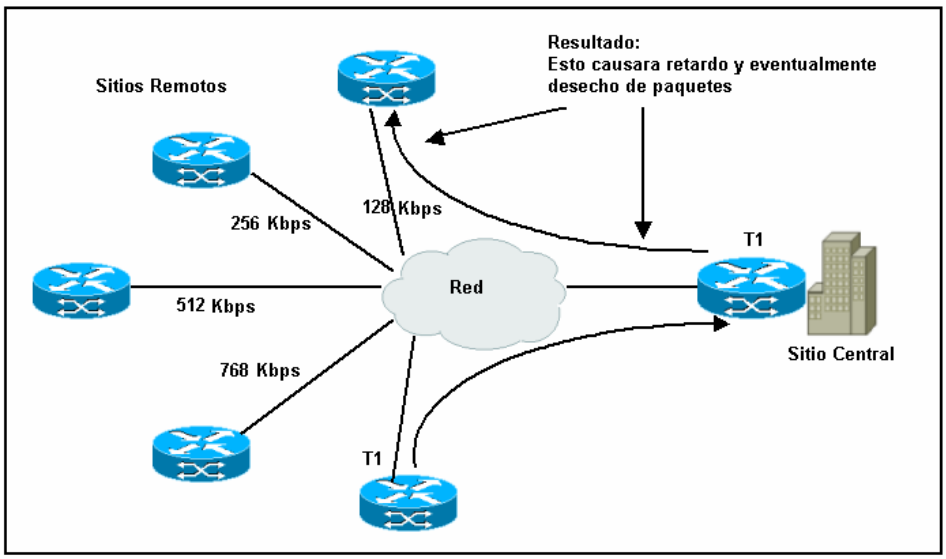


Figura. 2.15. Red de sin Formación de Tráfico

2.4.6 Compresión del encabezamiento IP RTP

La compresión del encabezamiento IP RTP reduce el tamaño de 40 bytes (IP+UDP+RTP) a 2 a 4 bytes, de esta forma se reduce el ancho de banda requerido por la llamada de voz en enlaces punto a punto. El encabezamiento es comprimido en un extremo del enlace y descomprimió en el otro extremo.

Otro nombre del estándar para esta técnica es Protocolo de Transporte en Tiempo Real Comprimido CRTP (Compress Real Time Protocol).

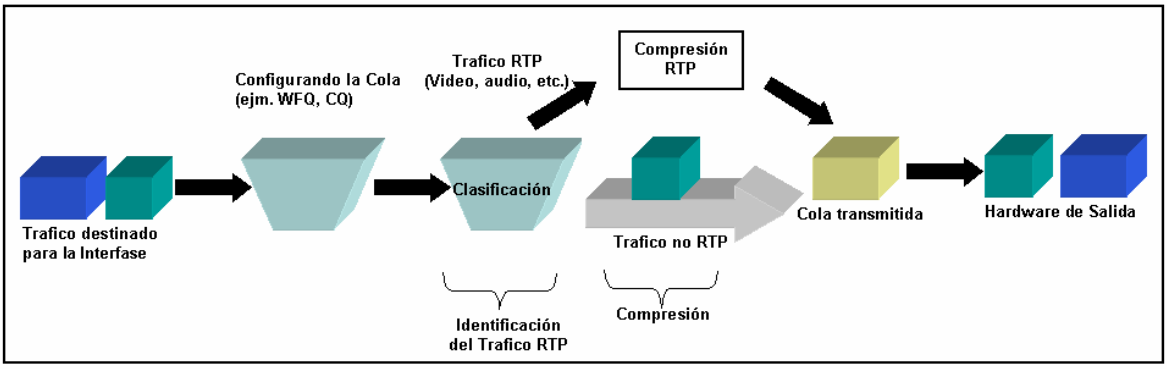


Figura. 2.16. Funcionalidad de la compresión del encabezamiento RTP

La compresión del protocolo RTP es recomendada en enlaces de bajas velocidades donde el ancho de banda es escaso y hay pocas llamadas de VoIP.

2.4.7 Servicios Diferenciados para VoIP

Las primeras redes de IP fueron basadas en el modelo de *servicio de mejor esfuerzo*, por lo cual el retardo, variabilidad del retardo o jitter, pérdida del paquetes y asignación del ancho de banda era imprevisible. Al usar el modelo del mejor esfuerzo, los proveedores de servicio no tienen ningún medio de ofrecer Acuerdos de Nivel de Servicios SLAs a sus clientes, a parte de sobreprovisionar la red para tratar con las horas de tráfico pico.

La arquitectura de Servicios Diferenciados DS (Differentiated Services) del modelo QoS proporciona un mecanismo escalable para clasificar los paquetes en grupos o clases que tienen similares requerimientos de QoS y entonces proporcionar el tratamiento requerido a estos grupos en cada punto de conexión en la red. La escalabilidad viene del hecho de que los paquetes son clasificados en los bordes de la red y se marcan apropiadamente para que los routers centrales en la red puedan proporcionar QoS simplemente basada en la clase servicios diferenciados DS. Los seis bits más significativos de byte de Tipo de

Servicio ToS se usan para especificar la clase de servicios diferenciados DS, el Punto de Código de Servicios Diferenciados DSCP (Differentiated Services Code Point) define estos seis bits. Los dos bits siguientes en el byte IP ToS actualmente se encuentran sin uso.

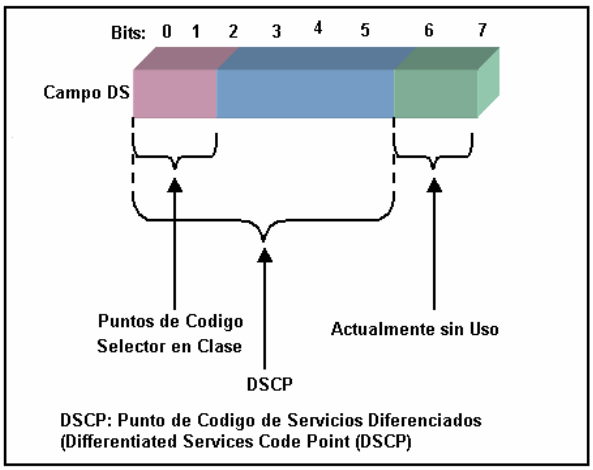


Figura. 2.17. Encabezamiento IP de Servicios Diferenciados

La tabla 2.4 muestra los valores del bit IP mapeado a DSCP.

Clase de Precedencia IP	**Valor de Bit IP Precedencia**	*Bits DSCP*	**Clase DSCP**
5	101	101000	Reenvío Apresurado
4	100	100000	Reenvío Seguro 4
3	011	011000	Reenvío Seguro 3
2	010	010000	Reenvío Seguro 2
1	001	001000	Reenvío Seguro 1
0	000	000000	Mejor Esfuerzo

Tabla. 2.4. Precedencia IP para mapeado de DSCP

Los próximos dos bits serán usados para definir la preferencia de eliminación. Por ejemplo, si el tráfico en Clase 4 (100) excede una cierta tasa, los paquetes en exceso podrían remarcarse para que la preferencia de eliminación se levante en lugar de realizarse.
Si la congestión fuera a ocurrir en la red, los primeros paquetes en ser desechados serían aquellos que poseen una preferencia alta de eliminación. Estos mecanismos permiten a la capa de red tomar decisiones inteligentes del tráfico durante los períodos de congestión. Los Servicios Diferenciados DS permiten el funcionamiento similar sobre una red de IP.

El sexto bit debe ponerse a 0 para indicar a los dispositivos de la red que las clases han sido fijadas según el estándar de DS.
La arquitectura de DS define un conjunto de acondicionadores de tráfico que se usan para limitar el tráfico en una región de DS y ponerlo en las clases de DS apropiadas. Los metros, marcadores, formadores (shapers) y terminadores son todos los acondicionadores de tráfico.

El Comportamiento por Salto PHB (Per hop behavior) describe lo que una clase de servicios DS deben experimentar en términos de pérdida, retardo y jitter. El comportamiento por salto PHB determina cómo el ancho de banda se asigna, cómo el tráfico se restringe y cómo se desechan paquetes durante congestión.

Se definen tres tipos de comportamientos por salto PHB en servicios DS basados en el comportamiento del reenvío requerido y estos son:

- Las clases de Mejor Esfuerzo: bits de selección de clase se ponen a 000.
- PHB de Reenvío Seguro: bits de selección de clase se ponen a 001, 010, 011 o 100.
- PHB de Reenvío Apresurado: bits de selección de clase se ponen a 101.

El estándar de Reenvío Seguro AF (Assured Forwarding) especifica cuatro clases de ancho de banda garantizado y describe el tratamiento que cada uno debe recibir. También especifica la preferencia de niveles de terminación y produce un total de 12 posibles clases AF, como se muestra en tabla.

Niveles de preferencia de terminación	Clase AF1	Clase AF2	Clase AF3	Clase AF4
Precedencia de eliminación baja	001010	010010	011010	100010
Precedencia de eliminación media	001100	010100	011100	100100
Precedencia de eliminación alta	001110	010110	011110	100110

Tabla. 2.5. Posibles Clases de Reenvío Seguro

Se puede usar clases de reenvío seguro para el tráfico de datos que no requieran tratamiento de prioridad y es principalmente basado en el protocolo TCP.

El reenvío apresurado EF (Expedited Forwarding) se usa para aplicaciones sensibles al retardo que requieren un ancho de banda garantizado. Un reenvío apresurado EF marca el servicio de prioridad reservando cierta cantidad mínima de ancho de banda que puede ser usada por el tráfico de prioridad alta.
En el reenvío apresurado EF, la tasa de salida debe ser mayor o igual a la suma de la tasa de ingreso para que no haya congestión en los paquetes marcados por EF. Se implementa el reenvío EF usando la cola de prioridad de la cola de espera LLQ. El ancho de banda constante se garantiza para el tráfico que pertenece a la clase de EF, pero al mismo tiempo si hay congestión, se eliminan los paquetes de no conformidad que exceden la tasa de prioridad especificada para asegurar que los paquetes en otras colas pertenecientes a las clases diferenciadas tengan el ancho de banda deseado.

Los valores de DSCP recomendados para el reenvío EF son 101110 (46). Por consiguiente, si los dispositivos de IP en la red pueden reconocer el precedente de IP o DSCP para la clasificación y propósitos de marcando, se puede proporcionar QoS de extremo a extremo.

2.4.8 Protocolo de Reservación de Recurso RSVP

Protocolo de Reservación de Recurso RSVP es una aplicación de la arquitectura de Servicios Integrados IS (Integrated Services) para QoS. Cuando VoIP fue introducido, el protocolo RSVP se vio inmediatamente como un componente importante que mantendría el control de admisión y QoS para los flujos de VoIP. Sin embargo, debido a la manera en que se integraron el protocolo RSVP y el estándar H.323 previamente no proporciona control de admisión ni QoS adecuado para flujos de voz.

Se han hecho varios perfeccionamientos manejar estas limitaciones y el protocolo RSVP puede usarse ahora implementando el Control de Admisión de Llamada CAC (Call Admission Control) y para señalar un QoS deseado que proporcionará una buena calidad de voz extremo a extremo, incluso en la presencia de congestión. La aplicación inicial del protocolo RSVP para VoIP tenía dos limitaciones:

- El control de admisión de llamada CAC no podría ser implementado con el protocolo RSVP porque el proceso de reservación no se sincronizaba con la señalización de llamada de voz. Una llamada seguía aun cuando la reservación de RSVP había fallado o no se había completado.
- Una reservación del protocolo RSVP exitosa no podría proporcionar calidad de voz buena durante los períodos de congestión de la red. RSVP creó un flujo de cola de espera por tráfico reservado dentro del sistema de cola de espera Carga Exacta WFQ y confió en ese sistema para garantizar un retardo limitado. Sin embargo, la cola de espera WFQ era incapaz en algunos casos para mantener un retardo limitado aceptable de la voz.

2.4.8.1 Control de Admisión de Llamada CAC (Call Admisión Control)

El control de Admisión de Llamada CAC es un concepto que se aplica solo al tráfico de voz, no al de datos.

Si un flujo de datos se subscribe a un enlace particular de la red, las decisiones de cola de espera, almacenamiento y eliminación de paquetes resuelve la congestión. El tráfico extra es retardado hasta que la interfase llegue a ser viable para enviar el tráfico o si el tráfico es eliminado, hasta que el protocolo o el usuario inicie una pausa requerida para la retransmisión de la información.

Los mecanismos de control de admisión de llamada CAC extienden las capacidades de las herramientas de QoS para proteger el tráfico de voz.

La figura muestra porque el control CAC es necesario. Si el enlace de acceso de la red WAN entre dos centrales PBX tiene el ancho de banda para llevar solamente 2 llamadas de VoIP, aceptando una tercera llamada se deterioraría la calidad de voz de las tres llamadas.

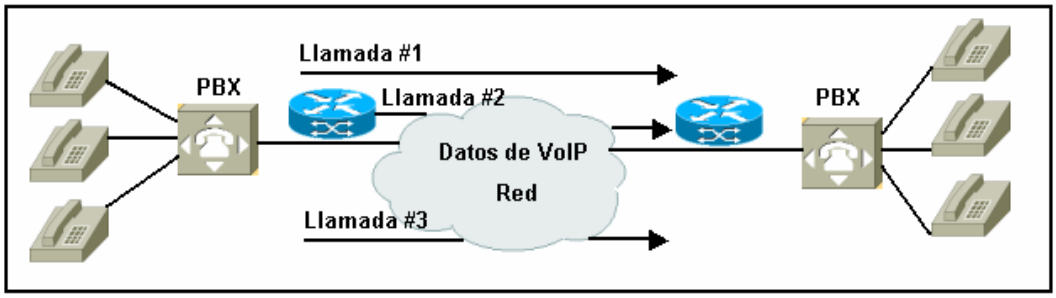


Figura. 2.18. Red de VoIP sin CAC

La razón de este deterioro es que los mecanismos de cola de espera proporcionan el orden, no el control CAC, de esta forma si los paquetes excedentes a la tasa admisible son recibidos, estos paquetes son puestos al final

de la cola. No existiendo la capacidad en los mecanismos de cola de espera para identificar cual paquete IP pertenece a cada llamada, por lo que cualquier paquete excedente será eliminado. Así, las tres llamadas experimentarán pérdida de paquetes.

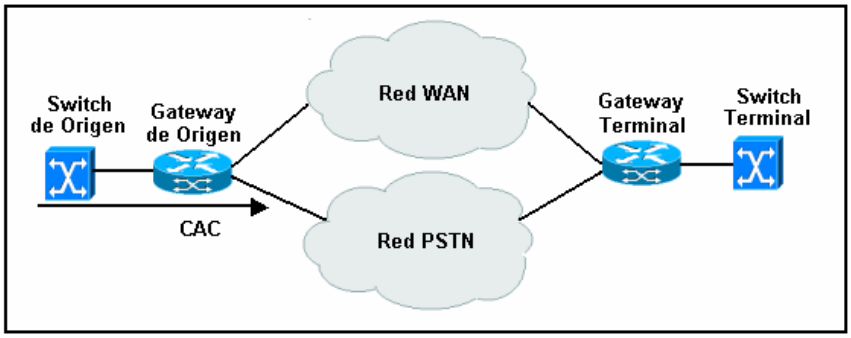


Figura. 2.19. Red de VoIP con CAC

La figura ilustra el punto en el cual la decisión de control de admisión de llamada CAC es alcanzada por la gateway terminal para permitir que una llamada continué.

2.4.8.2 El protocolo RSVP para CAC

El protocolo RSVP usando para el control de admisión de llamada CAC de VoIP requiere la sincronización de la señalización de establecimiento de llamada y la señalización de RSVP. Esta sincronización garantiza que el timbre del teléfono se realice solo después de los recursos para la llamada han sido reservados. Además proporciona a la gateway de voz el control de que acción tomar antes de que el establecimiento de la llamada se traslade a la fase de alerta si la reservación falla o que no pueda ser completada dentro de un periodo de tiempo predefinido.

Una llamada de voz activará dos reservaciones del protocolo RSVP porque la reservación y mecanismos de control de admisión proporcionados por RSVP son unidireccional.

Cada gateway de la voz es responsable de iniciar y mantener la reservación hacia otras gateways de voz.

La figura 2.20 muestra la secuencia de paquetes intercambiados entre las gateways durante el establecimiento de la llamada si el protocolo de RSVP es usado para la reservación del recurso.

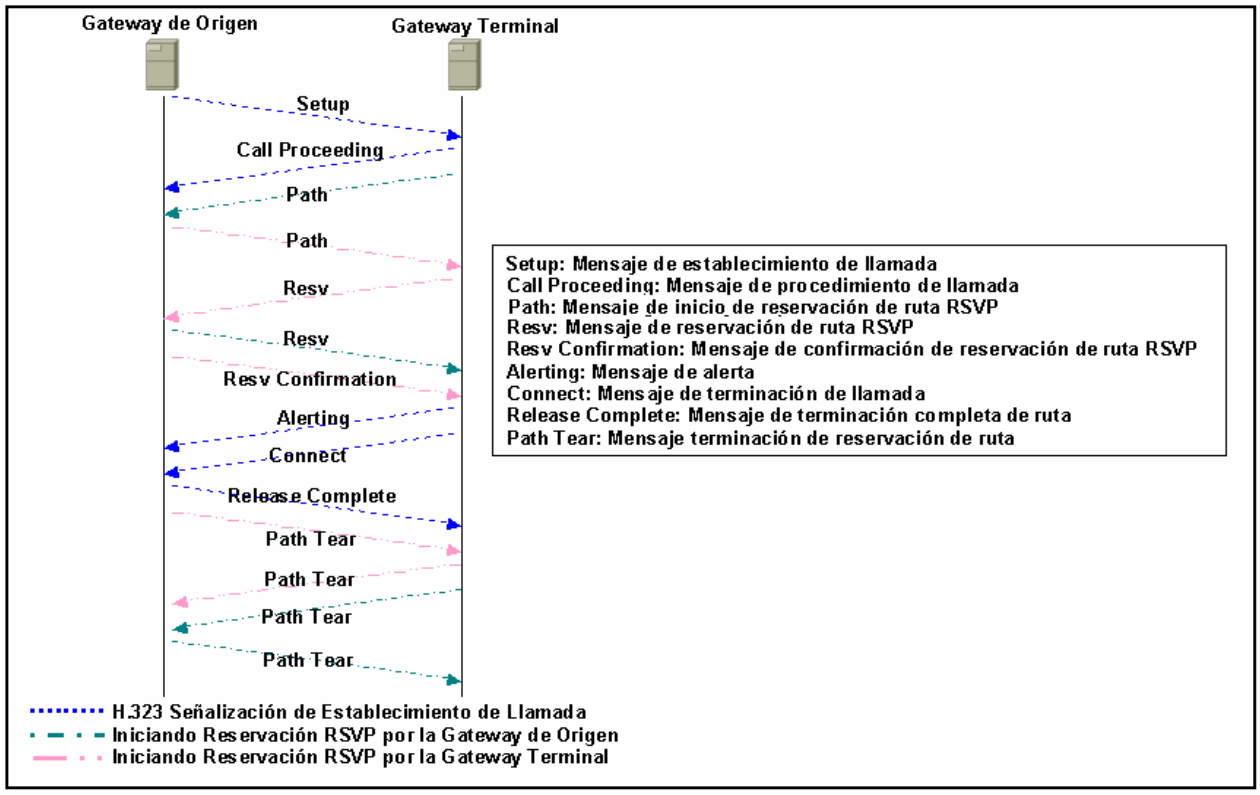


Figura 2.20. Establecimiento de llamada con RSVP

En la figura, una gateway de origen inicia una llamada hacia una gateway terminal. La gateway de origen envía un mensaje de establecimiento “Setup” a la gateway terminal para comenzar la llamada. Ese mensaje de establecimiento lleva la QoS que la gateway de origen considera aceptable para la llamada. La

gateway terminal responde con un mensaje de procedimiento de llamada “Call Proceding”.
La gateway de origen y terminal inician la reservación requerida enviando un mensaje de ruta “path” RSVP. El paquete que fluye de ambas reservaciones es independiente a menos que una de ellas falle.
La gateway terminal bloquea el proceso de establecimiento de llamada esperando por los resultados de la reservación. La gateway terminal controla la decisión de admisión de llamada y necesita ser notificado de que las reservaciones en ambas las direcciones tuvieron éxito. La gateway terminal descubre que su reservación tuvo éxito cuando recibe el mensaje de reservación “Resv” RSVP. La gateway terminal detecta que la reservación de la gateway de origen tuvo éxito cuando recibe un mensaje de confirmación de reservación “Resv Confirmation” RSVP. En este punto, la gateway terminal permite el establecimiento de la llamada continua y envía un mensaje de alerta “Alerting”, una vez a la gateway de origen notifica que el lado llamado está en estado de alerta. Una desconexión normal se comienza cuando un mensaje de terminación de llamada “Connect” completo se envía después de que la llamada esta conectada. En este punto, las gateways terminan sus reservaciones enviando mensajes de terminación de reservación “Release Complete” y terminación de ruta “Path release” RSVP.

Si por lo menos una reservación de RSVP falla, se puede configurar una gateway de voz para tomar las siguientes acciones:

- La gateway de voz puede informar la falla de la llamada al usuario o al conmutador que entregaron la llamada.
- La llamada puede ser redirigida a través de otro camino.
- La llamada puede conectarse con QoS de mejor esfuerzo.

Esta última conducta es posible porque la gateway terminal sabe que QoS es aceptable para la llamada de su propia configuración y el valor incluido por la gateway de origen en el mensaje de establecimiento.

Los beneficios de usar RSVP sólo pesan más que los costos donde el ancho de banda es limitado.

2.4.8.3 El protocolo RSVP con Cola de espera LLQ

Los flujos que le piden una QoS particular mediante el protocolo RSVP pueden aprovecharse de las alternativas de la cola de espera disponible en LLQ que tiene dos componentes principales:

- Una cola de espera QP
- Una Clase Basada en WFQ CBWFQ.

Las aplicaciones más recientes de RSVP confiaron en WFQ para reunir los requisitos de QoS para el tráfico sensible al retardo. Una cola reservada con una carga baja fue creada cuando la reservación de RSVP fue instalada. Sin embargo, WFQ no podría reunir los requisitos de retardo para el tráfico de voz y las llamadas de voz usando RSVP no podían aprovecharse de las características de PQ disponibles a lo largo de la cola de espera LLQ.

La figura muestra la estructura LLQ para una interfase donde el tráfico es clasificado en colas diferentes que usan varios métodos, incluso RSVP.

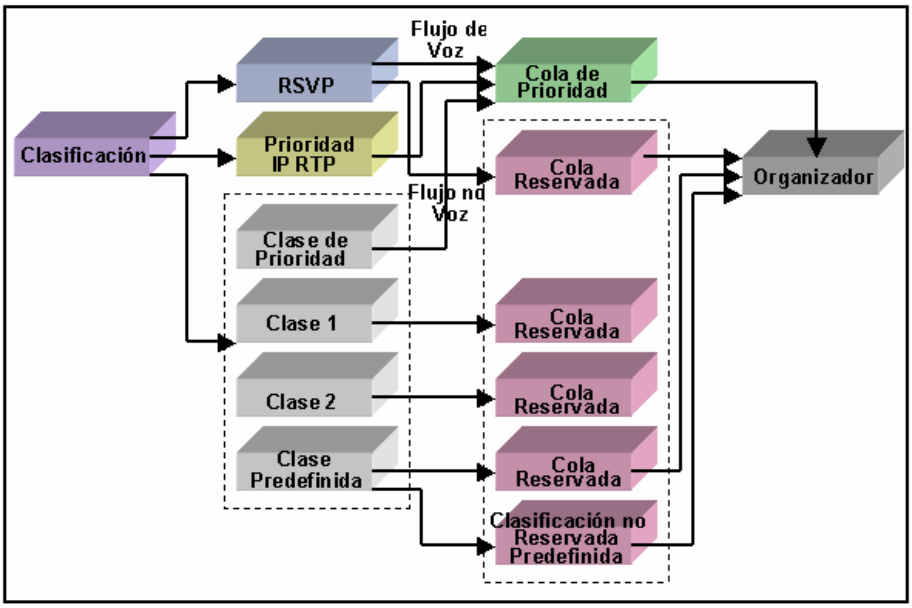


Figura. 2.21. RSVP soportado para LLQ

Una de las implicaciones más importantes del soporte de RSVP para LLQ es que le permite clasificar tráfico de la voz basado en sus características de tráfico en lugar del protocolo UDP y número del puerto.

El funcionamiento apropiado de LLQ confía en la suposición de que la cola de prioridad sólo es usada por el tráfico bien educado, como la voz, que tiene una tasa predecible y un tamaño de ráfaga muy bajo.

La clasificación basada en el protocolo UDP y el número de puerto podría permitir ráfagas o tráfico no critico en la cola de prioridad que podría afectar la calidad de llamadas de voz existentes y desarrollo del tráfico que usa el sistema de WFQ.

RSVP soporta prioridades LLQ de paquetes pero no cuida de la señalización de la voz. No puede ser posible comenzar nuevas llamadas durante los períodos de congestión pesada debido a la pérdida de paquetes de señalización.

2.5 PROBLEMAS DE QoS

2.5.1 El Retardo

Uno de los elementos claves de calidad de servicio que el usuario percibe es el retardo de punto a punto. El retardo puede ser afectado por:

- Retardo de la trama: Es la cantidad de tiempo representado en el paquete de voz.
- Retardo del Codec: Los codificadores de voz tienen cierto retardo inherente.
- Retardo de empaquetamiento: Un terminal o Gateway tendrá retardo pasando los paquetes de voz a través de su apuntador IP e inyectándolo a la red IP.
- Retardo del tránsito: Los paquetes de voz transportados a través de la red IP experimentarán un retardo al tiempo de transmisión de paquete a través de cada enlace y también retardos procedentes de los routers dentro de la red.

El retardo causa dos problemas fundamentales que son:

- El eco
- La superposición de la comunicación (talker overlap).

El eco: Es causado por las reflexiones de las señales de la voz del emisor desde el equipo de teléfono del extremo lejano (receptor) retornando al oído del emisor. El eco se vuelve un problema significante cuando el retardo del viaje completo es mayor que 50 ms. por lo que se deben usar técnicas de cancelación de eco. Los

estándares G.165 y G.168 definen los requerimientos de funcionamiento de los canceladores de eco.

Los canceladores de eco comparan los datos de voz desde la red IP con los datos de voz que están siendo transmitidos a la red. El eco desde la red de teléfonos es eliminado por un filtro digital en el camino de transmisión en la red IP. Debido a que se percibe el eco como un problema de calidad significativo, la voz sobre sistemas de paquetes debe direccionar la necesidad para el control del eco e implementar algunos medios de cancelación de eco.

La superposición de la conversación: Este problema se vuelve significativo si el retardo unidireccional es mayor a 250 ms. El retardo de extremo a extremo es, por consiguiente, la mayor restricción y requerimiento de manejo para incrementar la calidad de servicio a través de una red IP.

Las fuentes de retardo de extremo a extremo en una aplicación de VoIP incluyen:

- Acumulación de Retardos
- Retardo del procesamiento
- Retardo de la red

2.5.1.1 Acumulación de Retardos

Este retardo es causado por la necesidad de coleccionar una trama de muestras de voz para ser procesadas por el codificador de voz. Se relaciona al tipo de codificador de voz usado y varía desde un solo tiempo de la muestra (125 ms.) a muchos milisegundos. En la siguiente tabla se presenta los codificadores de voz normales y sus tiempos de trama:

CODIFICADOR	TIEMPO
G.726	125 ms
G.728	2.4 ms
G.729	10 ms
G.723.1	30 ms

Tabla. 2.6. Codificador y tiempo de trama

2.5.1.2 Retardo del procesamiento

Este retardo es causado por el proceso real de codificación y acumulación de las muestras codificadas en un paquete para la transmisión sobre la red IP.
El retardo de la codificación es una función del tiempo de ejecución de procesador y del tipo de algoritmo usado. A menudo, múltiples tramas del codificador de voz se acumulan en solo paquete para reducir el paquete de red overhead. Por ejemplo, tres tramas de G.729 equivalen a 30 milisegundos de dialogo y pueden acumularse y empaquetarse en un solo paquete.

2.5.1.3 Retardo de la red

Este retraso es causado por el medio físico y los protocolos usados para transmitir los datos de voz y por los buffers usados para eliminar el jitter del paquete en el lado del receptor. La red de retardo es una función de la capacidad de los enlaces y el proceso que ocurren como el tránsito de los paquetes.
Los buffers de jitter agregan retardo, el cual puede ser una parte significante del retardo global, cuando las variaciones de retardo del paquete son tan altas como 70 a 100 ms. en algunas redes de IP.

2.5.2 Variabilidad del retardo “Jitter”

El problema de retardo es incrementado también por la necesidad de eliminar la variabilidad o jitter. Eliminar el jitter requiere acumular paquetes y mantenerlos el tiempo suficiente para permitir a los paquetes más lentos llegar al mismo tiempo para ser ejecutados en la sucesión correcta, pero esto causa un retardo adicional.

Las dos metas contradictorias de minimización del retardo y eliminación del jitter han engendrado varios esquemas para adaptar el tamaño de buffer de jitter para emparejar los requisitos de variación de tiempo de eliminación del jitter de red. Esta adaptación tiene la meta explícita de minimizar el tamaño y retardo del buffer de jitter, mientras al mismo tiempo impide un mínimo de flujo (underflow) del buffer causado por el jitter.

Dos mecanismos para adaptar el tamaño de buffer de jitter se detallan a continuación. La selección del mecanismo dependerá del tipo de red en que los paquetes de están cruzando.

El primer mecanismo es medir la variación del nivel del paquete en el buffer de jitter sobre un período de tiempo y incrementalmente adaptar del tamaño del buffer para emparejar el jitter calculado. Este mecanismo trabaja mejor con redes que proporcionan una actuación del jitter consistente con el tiempo, como redes de ATM.

El segundo mecanismo es contar el número de paquetes que llegan tarde y crean una proporción de estos paquetes al número de paquetes que se procesan con éxito. Esta proporción es luego usada para ajustar el buffer de jitter para establecer la proporción del paquete tardío aceptable. Este mecanismo trabaja

mejor con las redes con muy inconstante de intervalos de arribo de paquetes como redes de IP.

Además de las técnicas descritas, la red debe configurarse y debe manejarse para proporcionar retardo mínimo y jitter mínimo y debe habilitarse una QoS consistente.

2.5.3 Pérdida de paquetes

Los paquetes perdidos pueden ser un problema muy severo, dependiendo del tipo de red de paquetes que está usándose. Debido a que las redes IP no garantizan la calidad de servicio, estas normalmente exhibirán una incidencia mucho más alta de paquetes de voz perdidos que las redes de ATM.

En redes IP actuales, todas las tramas de voz se tratan como las tramas de datos. Bajo las cargas pico y congestión, las tramas de voz serán desechadas de igual forma que las de datos. Sin embargo, las tramas de datos, no son sensibles al tiempo y se desechan paquetes que pueden ser apropiadamente corregidos a través del proceso de retransmisión. Los paquetes de voz perdidos, sin embargo, no pueden ser reparados de esta manera.

Algunos esquemas usados por software de paquetes de voz para direccionar el problema de tramas perdidas son:

Interpolación: Los paquetes del diálogo perdidos por re-ejecución del último paquete recibido durante el intervalo cuando se suponía que el paquete perdido estaba agotado, es un método simple que llena el tiempo entre las tramas del diálogo no continuo. Esta técnica funciona bien cuando la incidencia de tramas

perdidas es poco frecuente, no funciona bien si hay varios paquetes perdidos seguidos o un estallido de paquetes perdidos.

Redundancia: Es el envío de información redundante de acuerdo a la utilización del ancho de banda. Este método de copias básicas y envío de n paquetes de información de voz junto con el paquete (n+1), tiene la ventaja de ser capaz de corregir exactamente el paquete perdido, sin embargo, esta aproximación usa más ancho de banda y también crea un retardo mayor.

Codificador de Voz: Generalmente se usa un método híbrido con un codificador de voz de ancho de banda mucho más bajo para proporcionar información redundante llevada a lo largo de (n+1) paquetes. Esto reduce el problema del ancho de banda extra requerido pero no resuelve el problema de retardo.

CAPÍTULO III

FAX SOBRE IP FoIP (Fax over IP)

A pesar del gran crecimiento del servicio de e-mail durante los años noventa, el fax es un medio común de comunicaciones comerciales. La facilidad de uso, compatibilidad internacional y la entrega inmediata ha contribuido a la aceptación mundial del fax. Hoy, se envían mensajes de fax mediante la red PSTN. Las llamadas son facturadas por minuto y incurren en los cargos más altos en el área comercial e internacional. Estos mismos mensajes de fax pueden ser entregados usando la red IP con un ahorro significativo. Usando el Internet para entregar el tráfico de fax, los documentos pueden ser entregados sin la necesidad de llamadas de larga distancia.

La idea de fusionar el tráfico de fax y redes de IP no es nueva. Lo que es revolucionario, sin embargo, es la capacidad para entregar los mensajes de fax internacionalmente sin incurrir en cargos de teléfono de larga distancia.

Sin embargo, una razón por la que el uso del fax sobre IP FoIP no se ha popularizado aún, es la forma en que las empresas manejan los faxes: la responsabilidad de enviar faxes suele distribuirse en varios departamentos y contar con presupuestos diversos. Además, muchas organizaciones ven en el fax sobre IP FoIP sólo una pequeña parte de una estrategia global de convergencia de red que estaría basada en la telefonía IP.

Muchos administradores están ocupados ideando la implantación de sistemas de proyectos importantes, la integración de sus sistemas de voz, datos y las redes de fax tiene una prioridad relativamente baja en su lista de prioridades.

Inicialmente, el envío de faxes sobre IP se concibió como una forma de recortar los costos de las llamadas de larga distancia resultantes del uso del fax normal, sin embargo ahora se habla mucho menos del ahorro en los costos y más de las ventajas y flexibilidad de este servicio.

3.1 APLICACIONES DE FAX SOBRE IP

Hay tremendas oportunidades para reducir costos en las llamadas de transmisión de fax sobre las redes IP. Los datos de fax en su forma original son digitales, sin embargo, se modulan y se convierte a analógicos para la transmisión sobre la red PSTN. Esta forma analógica usa 64 Kbps de ancho de banda en ambas direcciones.

Un equipo de fax sobre IP invierte esta conversión analógica, transmitiendo datos digitales sobre la red de paquetes y luego reconvirtiendo los datos digitales a analógicos para la máquina del fax receptor. Este proceso de la conversión reduce el ancho de banda global requerido para enviar el fax, porque la forma digital es mucho más eficaz y la transmisión del fax es half duplex, es decir, sólo una dirección es usada en cualquier momento. La tasa más alta para una transmisión del fax es 14.4 Kbps en una dirección.

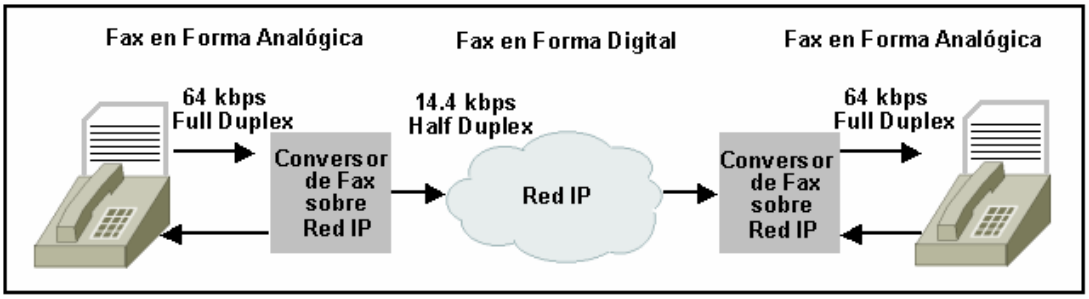


Figura. 3.1. Proceso de conversión de FoIP

La figura 3.2 muestra una aplicación de FoIP, la cual es una configuración de red de una compañía con numerosas oficinas que quieren usar la red IP, en lugar de la red PSTN, para proporcionar acceso a todos los faxes de la oficina central.

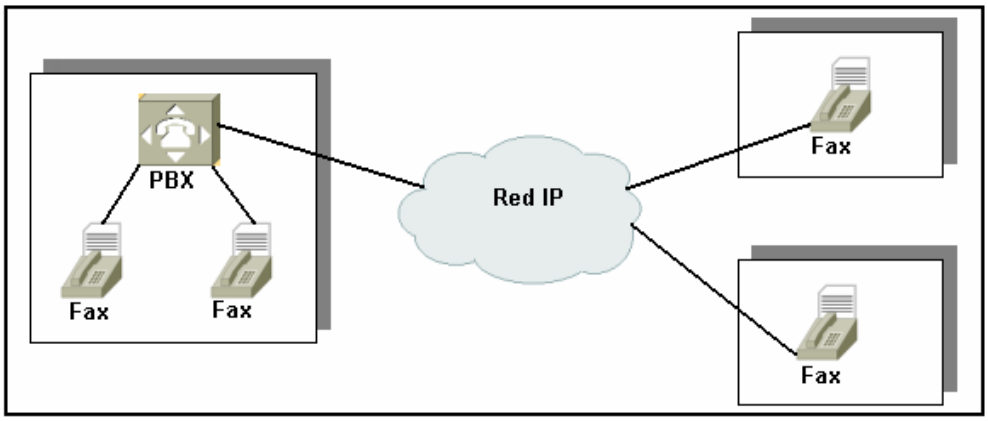


Figura. 3.2. Aplicación de FoIP

La transmisión de tráfico de FoIP requiere la integración del fax con la informática y la infraestructura de red. Varias aplicaciones de Fax de Internet se describen a continuación:

3.1.1 E-mail a las Gateways de Fax

La unión más reciente de fax y del Internet hacen posible que los mensajes de e-mail sean dirigidos a máquinas de fax normales. Usando esta aplicación, los usuarios del e-mail pueden dirigirse a los destinatarios del e-mail y destinatarios del fax en un solo mensaje.

Se soportan mensajes de texto simples, así como: Formato de Archivo de Imagen de Etiqueta TIFF (Tag Image File Format) y documentos adjuntos compatibles con las Extensiones de Correo Electrónico Multipropósito MIME (Multipurpose Internet Mail Extensions).

Con el rápido crecimiento de usuarios de Internet, este tipo de servicio se ha vuelto más atractivo aun.

3.1.2 Gateways de Fax de Internet

Las gateways avanzadas dirigirán el tráfico hacia y fuera del Internet y servirán como una base para habilitar un fax de Internet. Las gateways se desarrollaron en dos configuraciones:

- Un Equipo de Premisas de Cliente CPE (Client Premise Equipment), donde las gateways de sitio permitirán a una compañía enviar tráfico de fax sobre las redes internas e Internet.
- Gateways de fax de Internet que se desarrollan en los centros de comunicaciones, como Proveedores de Fax de Internet, o compañías de telecomunicaciones. A un centro de comunicaciones, la entrega de fax se proporcionará como un servicio adicional usando el Internet en lugar de PSTN para la entrega.

Con suerte, las comunicaciones entre las gateways serán basadas en estándares abiertos y esto permitirá la interoperabilidad y el intercambio fiable de mensajes de fax. En cada circunstancia, los costos menores de direccionamiento asegurarán que se minimicen costos de comunicaciones. Las gateways de Fax de Internet serán accesibles a maquinas de fax normales llamados “standalone” (que funcionan sin necesidad de un equipo agregado y sin estar conectado a la red) así como a clientes de fax desde sus PCs, permitiendo que todo el tráfico de fax sea entregado vía Internet.

3.1.3 Universal In-box

La tarea de manejar mensajes es tiempo consumido y involucra a menudo procesos dispares para recuperar e-mail, correos de voz y documentos de fax. En esta área, la convergencia promete tener un impacto significante en el flujo del trabajo. Se han desarrollado herramientas para manejar fax, e-mail y correo de voz desde una sola aplicación, eliminando la necesidad de supervisar sistemas múltiples.

Universal In-box también puede ser una herramienta poderosa para los usuarios móviles. Usando el Internet para la recuperación, todos los mensajes (fax, e-mail o voz) que puede entregarse en cualquier parte en el mundo. Una gateway de destino guarda mensajes hasta que un usuario necesita recuperarlos. Los usuarios se conectan a la gateway vía Internet. Se transmiten entonces los mensajes al usuario. El acceso al Internet está disponible casi a nivel mundial a través de los numerosos Proveedores de Servicio de Internet ISP (Internet Service Provider). Además, muchas compañías mantienen el acceso de la red sin restricción para el uso del empleado. Una vez conectado al Internet, varios métodos de recuperación son concebibles, incluso los e-mails del multimedia y protocolos soportados en WWW.

3.1.4 Servidor de demanda WWW/Fax

A pesar del crecimiento fenomenal del Internet, el uso no es todavía tan universal como el teléfono. Una integración potencial de fax y el Internet hace accesible información sobre páginas Web desde las máquinas de fax. Los servidores de demanda de fax con capacidades de Internet pueden recuperar páginas Web basadas en selecciones de “tono de toque”. De esta manera, los proveedores de información pueden mantener los datos en un servidor Web, mientras distribuyen esa información al público que tienen acceso a máquinas de fax , pero no proporcionan un acceso al Internet completo. Habilitando el envio de documentos de Lenguaje para Escribir Hiper Textos HTML (Hyper Text Markup Language).

Las nuevas generaciones de servidores de demanda de fax permitieran a las compañías tener un almacenamiento simple de documentos que pueden ser accedidos por cualquier método (Web o Fax) que los usuarios encuentran conveniente.

3.1.5 Uso del Fax en Información y Administración

Una vez que la dirección sabe que sectores de la compañía gastan más por servicios de fax, es posible controlar estos gastos. Típicamente, estos costos se entregan en informes de teléfono regulares y hacen difícil diferenciar entre el uso de fax y uso de voz (teléfono normal). Por direccionamiento del tráfico de fax a través de los servidores y la red corporativa, el uso es supervisado fácilmente, permitiendo un control más rígido de los gastos de la empresa.

3.1.6 Entrega segura

La seguridad de la información es una preocupación seria que se debe tener en cuenta con el método usado para enviar los documentos. Mientras se dirige mucha atención al problema de seguridad de Internet, la entrega de fax sobre IP tiene el potencial para aumentar la seguridad del documento.

Las transmisiones de fax sobre de las líneas del teléfono normales son transmitidas en forma digital comprimida y pueden descifrarse fácilmente. Así, las máquinas de fax son a menudo compartidas y los documentos pueden ser leídos por cualquier número de personas. Al contrario de esto, el fax sobre IP envían un fax direccionado que puede proporcionar entrega directamente en una caja privada de un destinatario. El acceso puede restringirse con el uso de una contraseña. Para los mensajes verdaderamente privados, la criptografía puede transformar un mensaje de fax digital en un mensaje codificado para la entrega por la red. Solo el destinatario puede descifrar el fax a ver, incrementando enormemente la seguridad de este sistema de envió.

3.1.7 Entrega garantizada y recibo de nunca ocupado

Cuando un fax se envía a través del Internet, las opciones de la entrega son variadas. Si el número de teléfono de destino está ocupado, el mensaje de fax podría recibirse y guardarse para una entrega posterior.
El servidor de fax podría manejar cualquier remarcado y proporcionar la información al remitente con respecto al estado de la entrega.

3.1.8 Transmisión Múltiple de Fax

Frecuentemente el mismo documento debe entregarse a múltiples destinatarios. Este puede ser un proceso tedioso en muchas máquinas de fax normales. Cuando se conecta a una gateways de fax de Internet, las capacidades adicionales están disponibles. Por ejemplo, pueden mantenerse una lista de distribución y pueden guardarse en el servidor y simplificarse la tarea de entrega de mensajes de fax.

3.2 DESCRIPCIÓN GENERAL DE LA OPERACIÓN DE FAX

Enviar un fax es mucho más complejo que marcar un número telefónico y enviar una imagen. La unidad llamada debe primero confirmar que la llamada está siendo contestada por un dispositivo de fax y no por un módem de datos, una contestadora o un ser humano. Después de determinar que dos dispositivos de fax están comunicados, estos dispositivos deben entonces intercambiar información para determinar las capacidades que soportan cada uno como velocidad de transmisión de datos, resolución de imágenes, esquemas de compresión de datos, tamaño de papel, etc. Después, deben ponerse de acuerdo respecto a un subconjunto de estas capacidades que ambos soporten.

A continuación, debe evaluarse la conexión telefónica para determinar la velocidad práctica máxima de los datos disponibles. Finalmente se envía la imagen fax. Sin embargo, es posible que el ruido o la distorsión corrompan la imagen durante su tránsito por la red de conexión. Para verificar esta posibilidad, el dispositivo receptor debe evaluar la imagen y enviar un mensaje de aceptación indicando si la imagen fue recibida en forma correcta o si tiene un número inaceptable de errores.

La estructura básica del sistema de un dispositivo fax típico se muestra en la figura.

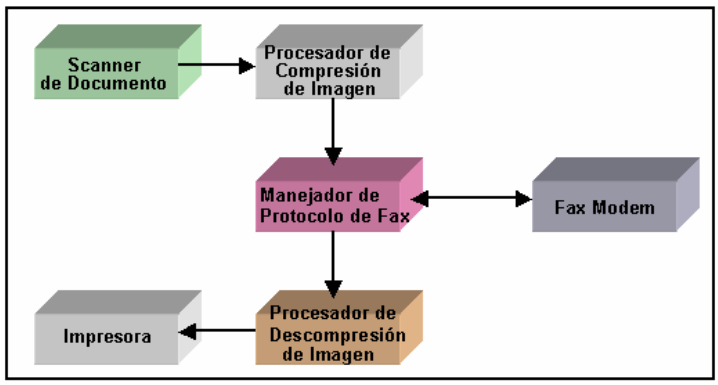


Figura. 3.3. Estructura Básica de un Sistema de Fax

La secuencia para la transmisión de un fax simple procede de la siguiente manera:

1. La imagen a transmitir se escanea primero para crear un mapa de bits en el cual los puntos negros y blancos son representados por unos y ceros.
2. Este mapa de bits se comprime para eliminar datos redundantes y después de transfiere a un manejador de protocolo, el cual es un conjunto de rutinas de software que maneja el intercambio de información de capacidad y transmisión de imagen.
3. Los datos son transferidos finalmente a un módem para su transmisión a través de la red.

De manera similar, las imágenes recibidas son extraídas de las secuencias del mensaje, se descomprimen y se imprimen o se almacenan.

3.2.1 Llamada de Fax sobre la red PSTN

Las máquinas de fax en uso actual implementan las recomendaciones de los protocolos de la ITU:

- Protocolo T.30
- Protocolo T.4

El protocolo T.30 describe el formato de datos de no página, como mensajes que se usan para la negociación de capacidades, mientras el protocolo T.4 describe el formato de datos de imagen de página.
El protocolo T.30 y T.4 han evolucionado substancialmente con el tiempo y son ahora bastante complejos, dado que ellos intentan describir el ambiente de un conjunto evolucionado de máquinas del fax.

La transmisión de datos del fax de extremo a extremo es acompañada por dos acciones:

- ✓ **Negociación:** Para asegurar que los datos examinados puedan entregarse al destinatario.
- ✓ **Confirmación de la entrega:** Proporciona al emisor la convicción de que los datos finales se han recibido y se han procesado.

Todas las máquinas de fax usan el protocolo V.21 (300 baudios) para la fase de la negociación de transmisión del fax. La fase de traslado de página se negocia a velocidades más altas (V.17, V.27, y así sucesivamente).

Las llamadas de fax sobre la red PSTN son divididas en cinco fases. Este ejemplo asume que la llamada es cumplida sin errores.
El procedimiento se pone algo más complicado si los errores ocurren. Las cinco fases son:

- Establecimiento de la llamada

- Control e intercambio de capacidades
- Transferencia de la página
- Señalización de extremo de página y de multipágina
- Terminación de llamada

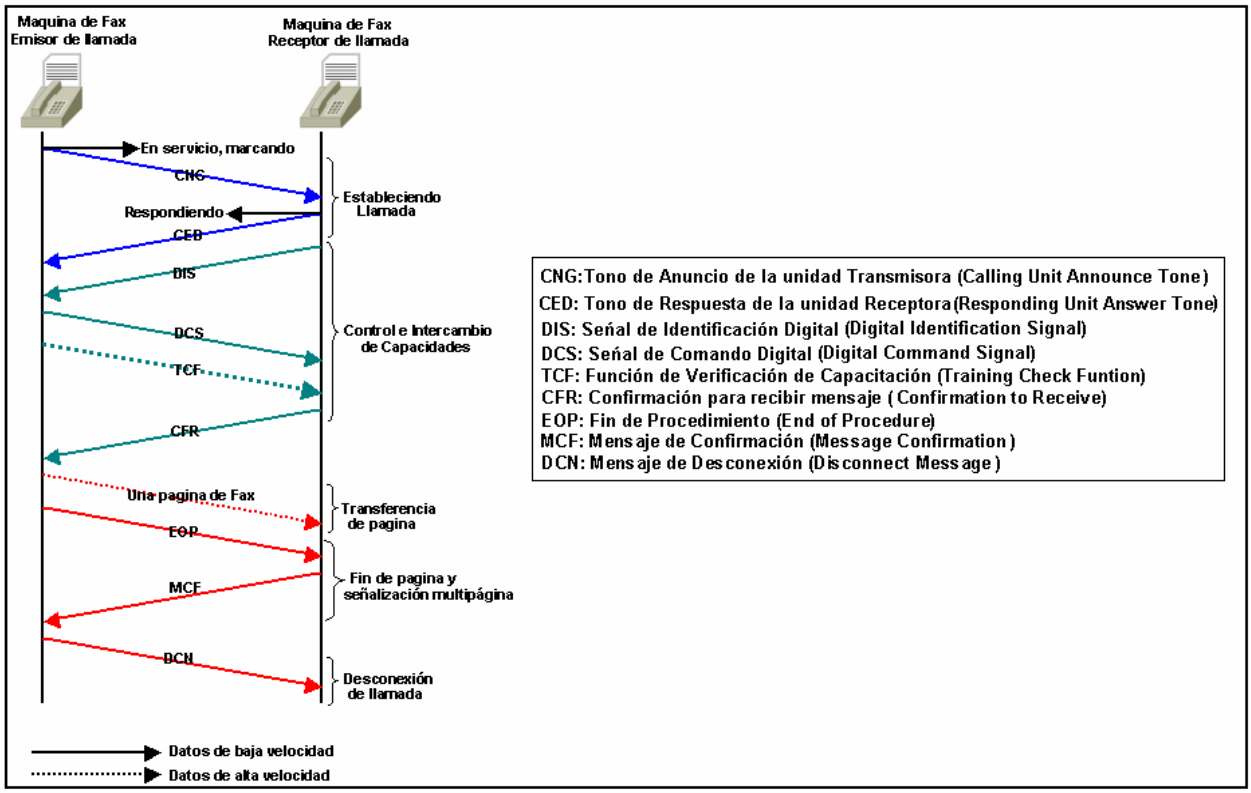


Figura. 3.4. Flujo de llamada de Fax

3.2.1.1 Establecimiento de llamada

La llamada de fax se establece a través de un proceso manual, de acuerdo con el cual alguien marca una llamada y pone la máquina en el modo de fax o por procedimientos automáticos, sin que ninguna interacción humana sea requerida. En ambos casos, el contestador de fax automático devuelve un tono de respuesta llamado Tono de Respuesta de la Unidad Receptora CED (Responding Unit Answer Tone), que es el tono alto que se escucha cuando se llama a una máquina de fax. Si la llamada se marca automáticamente, la estación emisora de la llamada también indicará la realización de la llamada de fax con un tono

llamado Tono de Anuncio de la Unidad Llamada CNG (Calling Unit Announce Tone), el cual es corto, el tono empieza inmediatamente después de que el número se marca. Estos tonos se generan para permitirle al participante humano comprender que una máquina de fax está presente en el otro extremo de la llamada.

3.2.1.2 Control e Intercambio de Capacidades

El control y la fase de intercambio de capacidades se usa para identificar las capacidades de la máquina de fax al otro extremo de la llamada. También negocia las condiciones aceptables para la llamada. El intercambio de mensajes de control a lo largo de la llamada del fax son enviados usando un modo de modulación de velocidades bajas (300 bps). Cada mensaje de control permite condicionar el canal de comunicación para la transmisión fiable.

La máquina de fax llamada empieza el procedimiento enviando un mensaje de Señal de Identificación Digital DIS (Digital Identification Signal) que contiene las capacidades de la máquina del fax. Por ejemplo, las capacidades que podrían identificarse en este mensaje son: el soporte de la tasa de señalización de datos, resolución de imagen seleccionada, modo de compresión de datos seleccionado, tamaño de imagen seleccionado, etc..
Una vez la máquina de fax receptora recibe el mensaje DIS, determina las condiciones para la llamada examinando su propia tabla de capacidades. La máquina receptora responde con la Señal de Comando Digital DCS (Digital Command Signal) que define las condiciones de la llamada.

El módem de gran velocidad se usará en la próxima fase de la llamada del fax para transferir datos de la página. La máquina de fax receptora envía un campo de Función de Verificación de Capacitación TCF (Training Check Funtion)

a través del sistema de modulación para verificar el entrenamiento y asegurar que el canal este activo según la tasa de transmisión de datos aceptada.
La máquina del fax llamada responde con un mensaje de Confirmación para Recibir CFR (Confirmation to Receiver) que indica que todas las capacidades y la velocidad de modulación han sido confirmadas y la página del fax puede enviarse.

3.2.1.3 Transferencia de página

El módem de gran velocidad se usa para transmitir los datos de la página que se han escaneado y comprimido. Se usa el estándar T.4 de la ITU para estructurar los datos de la página para la transmisión sobre el canal.

3.2.1.4 Señalización de extremo de página y de multipágina

Después de que la página se ha transmitido con éxito, la máquina de fax receptora envía un mensaje de Fin de Procedimiento EOP (End of Procedure) si la llamada del fax está completa y todas las páginas se han transmitido.
La máquina emisora responderá con un Mensaje de Confirmación MCF (Message Confirmation) para indicar que el mensaje se ha recibido con éxito y que está listo recibir más páginas.

3.2.1.5 Terminación de la Llamada

La fase de descargo o terminación es la fase final de la llamada en la que la máquina receptora envía un Mensaje de Desconexión DCN (Disconnect Message). Mientras el mensaje de DCN es una indicación positiva de que la llamada de fax finaliza, no es una indicación fiable, pues la máquina de fax se puede desconectar prematuramente sin enviar el mensaje de DCN.

3.2.2 Llamada de Fax sobre la red IP

El proceso de una llamada de fax sobre IP es similar a la llamada de fax sobre la red PSTN en las etapas que se deben emplear, la diferencia ocurre en el hardware que se maneja y los nuevos protocolos que se imponen para el trabajo de fax sobre IP. Normalmente el documento a transmitir se envía al servidor de fax que inicia la llamada telefónica (como es vía Internet la llamada es simplemente local) a la máquina de fax de destino. Así, el usuario envía el documento al servidor de fax. En este caso, el servidor de fax no marca el número de teléfono, más bien, avisa a una gateway de fax de Internet cercana a la máquina de fax de destino. El fax se transmite por el Internet a la gateway de fax remota que comienza una llamada telefónica local para entregar el fax. De esta manera, pueden entregarse documentos de fax a cualquier destino, internacional o local, con llamadas sólo locales.

La entrega mediante el Internet no se limita a faxes que se originen en computadoras, los fax de máquinas de fax normales pueden transmitir por el Internet también. Actualmente, se conectan máquinas de fax directamente a la red del teléfono.

Los números de teléfono permitieron a varios portadores locales y de larga distancias, iniciar las conexiones y entregar faxes. Para lograr esta entrega por el Internet, los servidores de fax podrían coleccionar tráfico desde la maquina de fax normal y dirigirlo al Internet para la entrega. Desde la perspectiva del usuario, nada ha cambiado. Así como antes, el usuario marca el número de teléfono de destino, los servidores de fax de Internet interpretan los dígitos marcados, no la red PSTN. El servidor de fax usará los dígitos marcados para localizar un servidor remoto apropiado y comenzar una conexión de Internet. Luego, el Fax se transmitirá vía Internet al servidor del fax remoto. La entrega del documento requiere sólo una llamada telefónica local cuando el destino está cercano a un punto de Internet.

Usando métodos como éstos, el Internet se vuelve una alternativa viable a la red PSTN para la entrega de documentos de fax. Actualmente existen dos métodos para el envió en redes de FoIP:

- Método de almacenamiento y envió
- Método de tiempo real

La diferencia primaria entre estos dos métodos esta en la entrega y en la confirmación de recepción. La ITU y el IETF están trabajando para continuar mejorando el estándar de red T.38 de FoIP en tiempo real así como el estándar de red de almacenamiento y envió T.37 de FoIP. Ambos T.37 y T.38 fueron aceptados por el ITU en junio, 1998.
Además, T.38 es el protocolo de transmisión de fax seleccionado para H.323.

3.2.2.1 Método de Almacenamiento y envió y el estándar T.37

Las gateways de almacenamiento y envió de fax pueden tomar llamadas desde máquinas de fax normal, convirtiéndolos en mensajes de e-mail y enviándolos sobre Internet. Otra gateway de almacenamiento y envió de fax en el extremo terminal de la llamada recibe el mensaje de e-mail y lo convierte en un mensaje de fax, el cual es entregado a una máquina del fax normal.

El método de almacenamiento y envió generalmente envía el fax como un dato adjunto del e-mail. Una extensión para el Protocolo de Transferencia de Correo Electrónico Simple SMTP (Simple Mail Transfer Protocol) para fax llamado Extensión de Correo Electrónico de Internet Multiproposito MIME (Multipurpose Internet Mail Extensions), esta disponible para este servicio de fax del IETF. Las imágenes de fax son adjuntadas a los encabezamientos de e-mail y son codificados en el Formato de Archivo de Imagen de Etiqueta TIFF (Tag Image File Format).

La ITU-T desarrolló el protocolo T.30 y otros estándares del fax y ha adoptado el protocolo SMTP MIME para el fax como parte de un nuevo estándar de la ITU-T llamado T.37. Este estándar define el almacenamiento y envió de fax por e-mail y ha aprobado el modo simple (RFC 2305). El modo simple restringe la codificación en formato TIFF a s-profile (subconjunto negro y blanco mínimo de TIFF para el fax), limitando la transmisión de fax a sólo los formatos de máquina de fax más populares. Estos formatos usa la compresión de imagen Modificada Huffman MH (Modified Huffman) con estándar o resolución fina.

En terminología T.37, las gateways de fax pueden enviar faxes a una máquina del fax conectada a una máquina convencional conectada a la red PSTN o a otra gateway de fax sobre de una red IP.
La gateway de origen (emisor) es llamada gateway en rampa (on-ramp), la gateway terminal (receptor) es llamada gateway fuera de rampa (off-ramp).

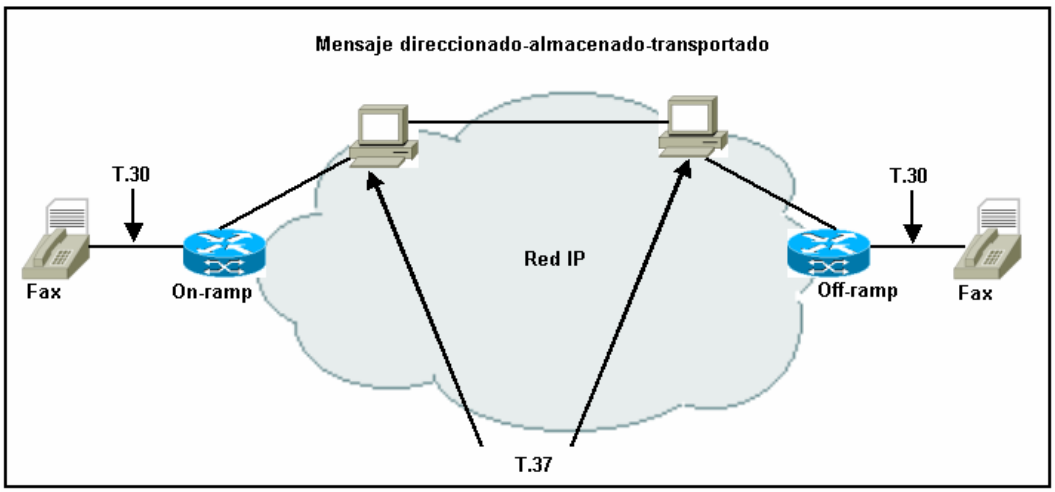


Figura. 3.5. Método de almacenamiento y envió de fax

3.2.2.2 Método de tiempo real y el estándar T.38

Las gateways de fax de tiempo real pueden entregar un fax a una máquina del fax remota mientras la máquina de fax emisora todavía está procesando las páginas del fax. La confirmación de la entrega es el proceso de la última página sin un mensaje del error y es inmediata.

El estándar T.38 define los protocolos de la red IP usados en los dispositivos del fax de Internet y las gateways de fax IP.
Las gateways de fax IP proporcionan las siguientes funciones:

- Demodulación de las señales de fax T.30 entrantes en la gateway emisora
- Traducción de los signos de fax T.30 en los paquetes del Protocolo de Fax de Internet IFP (Internet Fax Protocol) T.38.
- Intercambio de los paquetes del Protocolo de Fax de Internet IFP entre las gateways emisora y receptora T.38.
- Traducción paquetes del Protocolo de Fax de Internet IFP T.38 después de las señales T.30 en la gateway receptora.

La recomendación de la ITU-T, T.38, usa dos aplicaciones, las cuales difieren en su habilidad de tratar con el retardo de la red IP y son:

- Fax relay
- Fax de tiempo real con Spoofing

Fax Relay

Con Fax Relay, la gateway recibe una señal del fax analógico y lo demodula en su forma digital usando un fax-módem. El demodulado digital es entonces empaquetado y enviado sobre de la red IP. Al extremo receptor, la gateway de fax remodula los paquetes de fax digitales en señales de fax T.30 analógicas para ser enviada a la máquina de fax de destino a través de un módem de la gateway.

La red de retardo se vuelve un factor importante cuando el servicio de fax de tiempo real es desarrollado. Las redes de datos privados se han puesto a punto para el tráfico de VoIP, el retraso se ha reducido a menos de 500 ms de extremo a extremo y el fax relay puede usarse eficazmente.
Si los retrasos son demasiado largos, como en las gateways empleadas sobre el Internet, donde el retraso no está bajo el control del administrador directo, el fax de tiempo real con spoofing debe usarse.

Fax de tiempo real con Spoofing

Las técnicas de Spoofing son usadas para extender la tolerancia de retraso de máquinas del fax. Estas técnicas agregan al protocolo T.30 usado por máquinas del fax la capacidad de aguardar en línea más allá de sus intervalos de interrupción. El protocolo T.30 spoofing y el envio datos redundantes, son usados para proporcionar una tolerancia de jitter del paquete de imagen. Spoofing y jitter permiten a las máquinas del fax tolerar el retraso de la red sin perder comunicación. Estos mecanismos son suficientes para enviar fax sobre del Internet. El estándar T.38 define protocolos diferentes, dependiendo del mecanismo de transporte de fax en tiempo real.

Transporte UDP/IP

El Protocolo de Datagrama de Usuario UDP es un protocolo rápido pero inestable. La velocidad de UDP le permite ser empleado para el fax de tiempo real sin la necesidad de spoofing. Además, el protocolo T.38 proporciona dos métodos para mejorar la fiabilidad del mecanismo de transporte UDP.
Un método usa la redundancia de los datos de la imagen, el otro usa un esquema de Corrección del Error de Envió FEC (Forward Error Correction).

Transporte TCP/IP

Aunque el Protocolo de Control de Transmisión TCP (Transmisión Control Protocol) agrega fiabilidad a través del uso de los mecanismos de verificación de error en cada router, este transmite también retrasos. T.38 especifican un protocolo simple para transporte por TCP que no incluye ningún chequeo de error.

El modelo de servicio de fax en tiempo real sobre IP T.38 se ilustra en la figura.

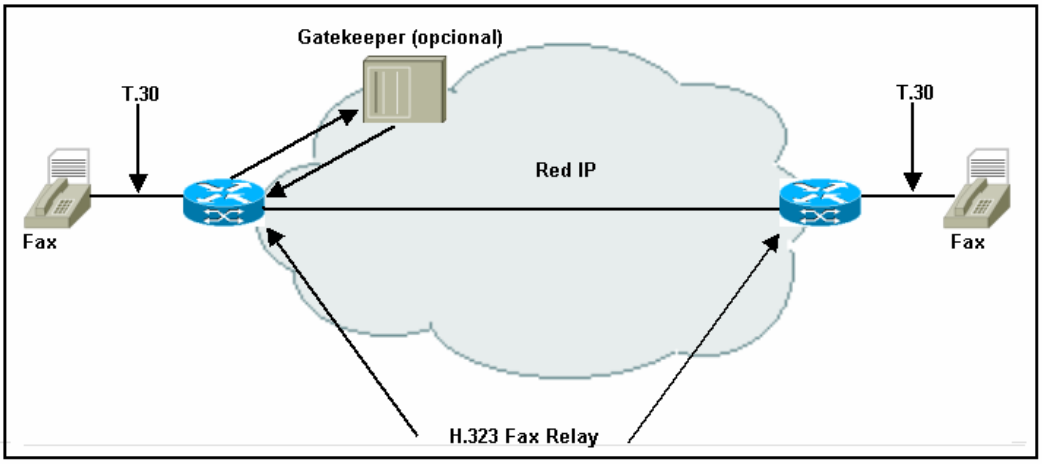


Figura. 3.6. Método de Tiempo Real

3.3 CALIDAD DE SERVICIO DE FoIP

Las ventajas de reducir costos y ahorrar el ancho de banda en las redes de fax sobre IP son asociadas con algunos problemas de calidad de servicio QoS y que pueden afectar la fiabilidad de la transmisión del fax. Los problemas mas comunes en FoIP son:

- Sincronización o “Timing”
- Retardo o “Jitter”

- Compensación de Pérdida de Paquetes

3.3.1 Sincronización o “Timing”

Un problema mayor en la aplicación de redes de fax sobre IP FoIP es la sincronización inexacta de los mensajes, este problema es causado por el retardo a través de la red.

El retardo de paquetes de fax a través de una red IP causa que la sincronización, la cual debería ser precisa, que es requerida por varias partes del protocolo de fax, sea inestable y puedan producir pérdidas de llamadas.

Hay dos fuentes de retardo de extremo a extremo:

- Retardo de la red
- Retardo del procesamiento.

3.3.1.1 Retardo de la red

Al igual que en la voz sobre IP, este problema es causado por el medio físico y los protocolos que se usan para transmitir los datos del fax y también por los buffers usados para remover el jitter en el extremo receptor.

Este retardo es una función de la capacidad de los enlaces en la red y del procesamiento que ocurre durante el tránsito de los paquetes en la red. Los buffers de jitter agregan retardo cuando remueven la variación de retardo de cada paquete. Este retardo puede ser una parte significativa del retardo global cuando las variaciones de retardo del paquete pueden ser tan altas como de 70 a 100 ms en las redes IP.

3.3.1.2 Retardo del procesamiento

Es causado por el proceso de demodulación y colección de la información digital del fax en un paquete para la transmisión sobre la red IP. El retardo de la codificación es una función del tiempo de ejecución de procesador y de la cantidad de datos almacenada antes de enviar un paquete a la red.
Los datos de baja velocidad, por ejemplo, son normalmente enviados con un solo byte por paquete, el tiempo para almacenar un byte de información en 300 bps es 30 ms.

3.3.2 Retardo o "Jitter"

Los problemas de retardo son compuestos por la necesidad de eliminar el jitter, la sincronización de un paquete causado por la red que atraviesa el paquete. Un método para remover el jitter es almacenar paquetes y mantenerlos más tiempo hasta que los paquetes más lentos lleguen y sean colocados a tiempo en la sucesión correcta, sin embargo, se causa un retardo adicional.
En la mayoría de los protocolos de FoIP, un tiempo adicional es incorporado en el paquete para asegurar que la información sea colocada en el momento apropiado.

3.3.3 Compensación de Perdida de Paquetes

Los paquetes perdidos pueden ser el problema más severo, por ejemplo en una aplicación de VoIP, la pérdida de paquetes puede ser direccionada reenviando los últimos paquetes o usando otros métodos de interpolación.
En una aplicación de FoIP, sin embargo, tiene inconvenientes más severos en la pérdida de datos, más cuando el protocolo de fax puede fallar.

Este problema varía, dependiendo del tipo de máquina de fax usada y si el modo de corrección del error se habilita. Dos esquemas pueden ser usados por el software de FoIP para direccionar los problemas de tramas perdidas y estos son:

- La repetición de la información en las tramas subsecuentes para que el error pueda ser corregido por el mecanismo de “playout” del receptor.
- Uso de un protocolo de corrección de error como el protocolo TCP para transportar los datos de fax a expensas del retardo añadido.

3.4 PRUEBAS DE FAX SOBRE IP

La continúa proliferación de productos y servicios basados en la tecnología de fax, particularmente aquellos productos y servicios que usan redes IP y la Internet, junto con servicios basados en voz, hacen cada vez más necesario realizar pruebas de fax. Los sistemas y redes de fax deben ser probados en cuanto a cumplimiento de las normas, el rendimiento de la red, la localización y la solución de problemas.

Debido a que las máquinas de fax son muy fáciles de usar (se coloca el papel en la guía, se marca el número y se oprime la tecla enviar), muchas personas asumen que todo el proceso es muy simple y que las pruebas no representan un aspecto importante. La verdad es que la compatibilidad de comunicación entre los dispositivos de fax es un problema mayúsculo por varias razones:

- La tecnología de fax sigue evolucionando, incorporando un gran número de funciones en forma regular y los nuevos dispositivos de fax deben ser compatibles no solo con los modelos más recientes sino con la base instalada existente de máquinas fax.
- Muchas compañías diferentes con muchos métodos de diseño diferentes fabrican este tipo de productos y estos productos deben comunicarse entre sí.

- Las tecnologías de comunicaciones y estructuras de redes están cambiando a un ritmo acelerado y todas ellas deben poder manejar tanto tráfico de fax como de voz.

3.4.1 Por qué son necesarias las Pruebas de Fax

Los nuevos productos de fax deben ser diseñados para ser compatibles con la base instalada existente de sistemas de fax fabricados por compañías diferentes. Deben ser compatibles con productos de fax de una gran diversidad de proveedores que tienen características exclusivas de rendimiento. Deben adaptarse a tecnologías de fax que cambian con el paso de los años y deben ser capaces de manejar nuevos servicios mejorados.
Las pruebas son un requisito vital para lograr y mantener dicha compatibilidad, debido a que los portadores y los operadores de redes deben ser capaces de desplazar grandes y crecientes volúmenes de tráfico de fax a través de redes cada vez más complejas, los problemas de transmisión y compatibilidad deben ser analizados y corregidos rápidamente.

La eficiencia de la transmisión, es decir, el rendimiento, debe ser mejorada constantemente utilizando los esquemas más efectivos para compresión de datos, con el fin de reducir el costo por llamada y deben identificarse y compensarse de inmediato todos los factores que impiden una transmisión eficiente, tales como acumulación innecesaria de imágenes.

Además, la calidad de la transmisión debe mejorarse continuamente. La realización de pruebas completas en forma permanente es casi obligatoria para lograr estos objetivos. Con el advenimiento de las redes de voz sobre IP VoIP, la detección de tráfico de fax en gateways VoIP y transmisión de fax vía redes IP son componentes importantes para el éxito de servicios basados en IP en tiempo real.

Las expectativas de transporte y calidad de fax en redes IP son muy distintas a las de voz. Para manejar fax se han desarrollado protocolos Fax IP especiales para las transmisiones, método de almacenamiento y envió (protocolo T.37) y de tiempo real (protocolo T.38) en redes de FoIP. Los sistemas que utilizan estos nuevos protocolos no solamente requieren pruebas detalladas durante su desarrollo e implantación, sino que también requieren pruebas continuas durante su operación para garantizar una calidad de servicio QoS aceptable. Los aspectos de prueba que deben ser considerados incluyen los siguientes:

- Problemas de interoperatividad entre equipos de distintos proveedores en el lado IP.
- Problemas de compatibilidad entre nuevos protocolos Fax IP y redes telefónicas convencionales que usan protocolos fax estándar más antiguos.
- Limitaciones de la red IP, tales como pérdida de paquetes, jitter y extensas demoras que con frecuencia deben ser compensadas modificando la transmisión por fax en las redes telefónicas, estos problemas (la mayoría relacionados con la sincronización) son manejados típicamente en las gateways de la red mediante diversas técnicas de compensación, las que incluyen la saturación de las tramas de fax con información vacía o el envio de tramas falsas para mantener en línea al fax receptor, estas técnicas pueden tener un impacto considerable en la duración y precisión de las transmisiones por fax.

3.4.2 Consideraciones para Pruebas de Sistemas de Fax

Los sistemas de fax deben ser probados con el fin de asegurar que cumplan con los estándares de la industria y para verificar su compatibilidad con la base instalada existente de dispositivos de fax muchos de los cuales se desvían de los estándares de varias formas diferentes.

Los procedimientos de operación de un fax son definidos por estándares publicados por la ITU Internacional, comités integrados por representantes de los fabricantes de máquinas de fax, compañías de telecomunicaciones y agencias gubernamentales que desarrollan estos estándares, algunos de los cuales se listan a continuación:

Estándares de Modulación Métodos de Codificación de Bits para Transmisión de Imágenes y Datos de Control:

- ✓ **V.21** Define la modulación usada para la transmisión de mensajes de control a 300 bps en un fax convencional
- ✓ **V.17, V.29, V.27 ter** Define la modulación usada para transmisión de datos de imágenes a velocidades de hasta 14.4 Kbps
- ✓ **V.34 fax** Define la modulación y los procedimientos usados para la transmisión de imágenes a velocidades de hasta 33.6 Kbps
- ✓ **V.8** Define la modulación y los procedimientos para el intercambio inicial de datos para configurar un fax V.34

Estándares de Protocolo Procedimientos y Secuencias de Mensajes

- ✓ **T.30** Define los procedimientos y secuencias de mensajes usados para comunicaciones por fax

Estándares de Compresión de Datos Procedimientos usados para Comprimir Imágenes para Transmisión por Fax

- ✓ **T.4** Define los algoritmos usados para compresión de datos unidimensional y bidimensional
- ✓ **T.6** Define los algoritmos usados para compresión de imágenes en modo corrección de errores

El propósito de estos estándares es asegurar que todos los dispositivos de fax sigan un procedimiento común con el fin de que dispositivos fabricados por distintas compañías puedan comunicarse de manera confiable entre ellos.

Desafortunadamente, el cumplimiento con los estándares varía de un fabricante a otro. Los estándares de la ITU son recomendaciones, no leyes, y están sujetos a la interpretación (o mal interpretación) y en ocasiones por diversas razones técnicas, los fabricantes deliberadamente pasan por alto estos estándares. Se requieren pruebas extensas para asegurar el cumplimiento con los estándares y un manejo sin problemas de desviaciones comunes.

3.4.3 Aspectos relativos a la prueba de redes

No solamente los fabricantes de sistemas de fax necesitan probar sus productos. Los fabricantes de equipos para redes y los operadores de redes también requieren contar con capacidades para prueba de fax por las siguientes razones:

- Las demoras de la red pueden entrar en conflicto con las limitaciones de sincronización T.30.
- En necesario efectuar pruebas para detectar conflictos de sincronización de la red y los usuarios para la instrumentación de soluciones.
- Los errores y la distorsión inducidos por la red pueden degradar la calidad.
- Las pruebas constantes muestran cómo pueden reducirse los errores y mejorar la calidad.
- Esquemas de compresión ineficientes y conglomeraciones innecesarias pueden reducir el rendimiento de la red.
- Las pruebas son necesarias para comparar el desempeño de la red contra un desempeño óptimo e identificar las causas de un menor rendimiento.
- Los gateways y servidores pueden sobrecargarse y causar la pérdida de llamadas.

- La generación de llamadas a granel puede reflejar los efectos de tráfico pesado de fax y medir limitaciones del desempeño.

Las pruebas son necesarias para medir el desempeño de la red y permitir a los usuarios solucionar problemas y afinar las operaciones.

3.4.4 Desviaciones comunes a partir de los estándares de la ITU

Los estándares ITU para fax especifican la secuencia de operaciones en una llamada de fax, incluyendo los mensajes a intercambiar, el formato de dichos mensajes, los procedimientos para probar la validez de los mensajes, procedimientos para recuperación de errores, sincronización, entre otros.

La mayoría de los sistemas de fax se desvían de los estándares ITU de una forma u otra. Con el fin de ser compatibles con la base instalada de dispositivos fax en todo el mundo, los sistemas y las redes deben tolerar muchas de las anomalías y violaciones los estándares que encontrarán constantemente. Algunas de las anomalías y desviaciones de los estándares más comunes se describen a continuación.

3.4.4.1 Desviaciones de sincronización típicas

Además del formato de los mensajes, el estándar T.30 también especifica diversos requisitos de sincronización. Por ejemplo, después de marcar un número, se supone que la unidad transmisora debe escuchar una respuesta en 35 segundos antes de darse por vencida. Estos períodos de tiempo muerto “time out” son los siguientes:

- ✓ **T1 35 ± 5 segundos:** El tiempo en que dos unidades de fax intentan identificarse entre sí.
- ✓ **T2 6 ± 1 segundos:** Un tiempo muerto usado para comenzar la secuencia para cambiar los parámetros de transmisión.
- ✓ **T3 10 ± 5 segundos:** Un tiempo muerto usado en el manejo de interrupciones complejas.
- ✓ **T4 60 ± 5 segundos** Un tiempo muerto usado en el modo de corrección de errores.

Estos tiempos muertos con frecuencia son mal interpretados. Además son ignorados en forma rutinaria y en muchos casos así es como debe ser. Por ejemplo, después de establecer una llamada, se supone que la unidad transmisora debe esperar 35 segundos antes de darse por vencida.
Si la unidad receptora no responde al primer timbre o si una máquina contestadora está conectada a la línea, o si existen varias demoras en la red, entonces la demora antes de la respuesta puede ser mucho mayor a 35 segundos.

Las unidades de fax que soportan el modo de corrección de errores pueden responder a un mensaje de apretón de manos posterior a la imagen con un mensaje Receptor no Listo RNR (Receiver Not Ready). La unidad transmisora consulta a la unidad receptora de fax con un mensaje Receptor Listo RR (Receiver Ready). Si la unidad receptora sigue ocupada (por ejemplo, si está imprimiendo), repetirá el mensaje de Receptor no Listo RNR.

De acuerdo con el estándar T.30, esta secuencia de mensajes RR/RNR RR/RNR puede ser repetida hasta el final de T4 (60 ± 5 segundos). Sin embargo, muchos sistemas de fax ignorarán el tiempo muerto y continuarán la secuencia indefinidamente a menos que el usuario la cancele manualmente. Todos los tiempos muertos están sujetos a alteración y, en algunos casos, a un uso

inadecuado. Se requieren pruebas extensas para verificar que las desviaciones encontradas con mayor frecuencia sean manejadas en forma adecuada.

3.4.4.2 Problemas de brechas entre Portadores

El estándar T.30 especifica un periodo de silencio de 75 ± 20 ms entre señales que utilizan distintos modos de modulación, por ejemplo, entre el final de un mensaje de Señal de Comando Digital DCS y el comienzo de una transmisión posterior a la imagen. Este requisito con frecuencia es violado, particularmente durante el período de silencio entre un mensaje de Señal de Comando Digital DCS y la Función de Verificación de Capacitación TCF. Muchos sistemas de fax en forma rutinaria extienden la duración de este periodo de silencio a mucho más de 100 ms. Desafortunadamente, si este periodo es demasiado prolongado, puede interferir con el procedimiento de recuperación de error del mensaje cuando el mensaje DCS ha sido corrompido por paquetes perdidos o ruido en la línea.

Todos los sistemas deben ser probados para asegurar que permanezcan dentro de los límites prescritos de T.30 y que en caso de exceder estos límites no se vea limitada su capacidad para recuperarse de condiciones de error.

3.4.4.3 Otras variaciones de sincronización

También se requieren pruebas para determinar la habilidad de un sistema de fax para manejar variaciones en la duración de las pausas entre repeticiones de mensajes no reconocidos y también en las pausas entre la recepción de un comando de intercambio y el comienzo de la respuesta a dicho comando. Con el fin de reducir el tiempo total de transmisión, muchos sistemas de fax comienzan a enviar un mensaje de respuesta antes de recibir el final del comando.

3.4.4.4 Anomalías en el tono de respuesta

El estándar T.30 establece que el dispositivo de fax receptor debe enviar un tono de respuesta de 2100 Hz durante aproximadamente 3 segundos antes de enviar el primer mensaje de apretón de manos. Algunas máquinas de fax envían un tono de 1850 Hz, algunas envían un tono de 1100 Hz y algunas omiten totalmente el tono de respuesta y simplemente comienzan con el primer mensaje de apretón de manos. Son necesarias pruebas para verificar que una unidad transmisora de fax complete la llamada, incluso en ausencia de un tono de respuesta.

3.4.4.5 Otras desviaciones de los estándares ITU

Existen otras muchas anomalías y desviaciones encontradas con frecuencia por las cuales es necesario probar los sistemas, incluyendo las siguientes:

- Desviaciones en la secuencia de la trama
- Variaciones en las secuencias del preámbulo y la bandera
- Uso inadecuado de Señal de Paginas Múltiples EOM (Multipage Signal)
- Secuencias inusuales de caída de velocidad de datos
- Algoritmos de detección de patrones comunes de entrenamiento
- Desviaciones de transmisión de imágenes
- Uso de tono con protección contra eco
- Atenuación de imagen y líneas cortas
- Uso del mensaje Reentrenamiento Positivo RTP (Retrain Positive) / Reentrenamiento Negativo RTN (Retrain Negative)
- Líneas de larga duración
- Secuencias de desconexión no estándares
- Uso del mensaje de desconexión DCN

3.4.5 Tipos de clientes para Pruebas de Fax

Muchos tipos de empresas requieren pruebas de sistemas de fax. Los fabricantes de máquinas de fax, fax módems, productos relacionados con fax y compañías de software, fabricantes de equipo para redes y operadores de redes requieren realizar pruebas de fax. Estos requerimientos incluyen no solamente pruebas de diagnóstico durante el desarrollo de los productos, sino también pruebas en producción y de control de calidad, pruebas para soporte a clientes y (para los operadores de redes en particular) mediciones de calidad de servicio QoS en forma continúa. Los tipos de clientes típicos incluyen a los siguientes:

- Fabricantes de máquinas de fax
- Fabricantes de tarjetas fax para PC
- Fabricantes de módems
- Compañías de software para fax
- Compañías de fax celulares
- Proveedores de gateways
- Proveedores de equipo para redes
- Operadores de redes

3.4.6 Áreas de aplicación para pruebas de fax

La prueba de sistemas de fax es una necesidad no solamente en laboratorios de ingeniería en donde se crean productos de fax, sino también en muchas otras áreas. Las áreas de aplicación para pruebas de fax incluyen las siguientes:

- Desarrolladores de Productos para Fax
- Pruebas en tiempo real para visualizar las operaciones en detalle
- Dispositivo de experimentación para enviar y recibir faxes

Aseguramiento de la Calidad

- Prueba del peor caso
- Pruebas de regresión usando comparación contra llamadas de referencia estándar
- Generación de reportes a la medida para documentar los resultados

Pruebas en Producción

- Pruebas de cumplimiento con T.30 e interoperatividad
- Comparación contra llamadas de referencia estándar para detectar desviaciones en el desempeño
- Generación de reportes adecuados a las necesidades de producción

Soporte a Clientes

- Requerimientos para enviar y recibir faxes en los sistemas del cliente
- Pruebas y análisis en tiempo real de los problemas del cliente
- Operaciones de fax celular
- Simulación de problemas celulares únicos cancelaciones por portador, etc.

Pruebas del Servidor Fax

- Generación de grandes volúmenes de llamadas simultáneas para pruebas de carga

Pruebas de Redes

- Pruebas extremo a extremo para registrar y analizar efectos de demora
- Monitoreo en tiempo real para mostrar los efectos de ruido y distorsión
- Prueba de gateways
- Pruebas de cumplimiento e interoperatividad T.30
- Generación de llamadas simultáneas para pruebas de carga.

CAPÍTULO IV

MULTIMEDIA SOBRE IP MoIP (Multimedia over IP)

Desde la introducción del primer sistema de multimedia hasta nuestros días muchas cosas han cambiado desde el punto de vista tecnológico y de aplicación a las necesidades de los usuarios. Las primeras soluciones estaban basadas en tecnologías propietarias y por lo tanto, no permitían la comunicación entre sistemas de diferentes fabricantes. Desde aquel primer concepto, se ha pasado a sistemas mucho más flexibles, económicos y que pueden adaptarse a las diferentes necesidades de los usuarios, según sean sus aplicaciones y capacidad económica.

Actualmente la mayor aplicación de la multimedia es la videoconferencia entre usuarios que supone la comunicación audiovisual entre los mismos y opcionalmente la provisión de facilidades para el intercambio de datos, tales como: prestaciones telemáticas típicas, intercambio de ficheros, compartición de aplicaciones ofimáticas, etc.

Tradicionalmente, las Redes IP se utilizaban para la transmisión de datos, pero conforme las aplicaciones tienden a ser multimedia y los sistemas de comunicaciones en vez de ser elementos independientes y aislados para atender un determinado tipo de comunicación, son servidores de un conjunto más complejo y se tiende a transmitir cualquier tipo de información sobre los medios existentes, con la adopción de ciertos estándares y la incorporación de algunos

elementos es posible enviar datos, voz y vídeo mediante la red IP con la gran ventaja y ahorro que supone el utilizar la infraestructura existente.

Sin embargo, mientras los datos no son sensibles al retardo, la alteración del orden en que llegan los paquetes o la pérdida de alguno de ellos, ya que en el extremo lejano se reconstruyen, la voz y la imagen necesitan transmitirse en tiempo real, siendo especialmente sensibles a cualquier alteración que se pueda darse en sus características, por lo que requieren de redes IP que ofrezcan un alto grado de servicio y garanticen el ancho de banda necesario además de la utilización de los protocolos necesarios para la transmisión en tiempo real como son:

- H.323 Protocolo de Multimedia
- H.225 Protocolo de Señalización de Llamada
- H.245 Protocolo de Control de Llamada
- Protocolo de Inicio de Sesión SIP (Session Initiation Protocol)
- Protocolo de Control de Gateway Media MGCP (Media Gateway Control Protocol)

Los cuales serán analizados a continuación.

4.1 ESTÁNDAR H.323

El protocolo H.323 es una familia de estándares definidos por la Unión de Telecomunicación Internacional ITU, la cual especifica los componentes, protocolos y procedimientos que proporcionan los servicios para las comunicaciones multimedia, audio, video y comunicación de datos, en tiempo real sobre redes basadas en paquetes, tal como: la Red LAN e Internet pública que generalmente no proporciona una calidad de servicio QoS garantizada.

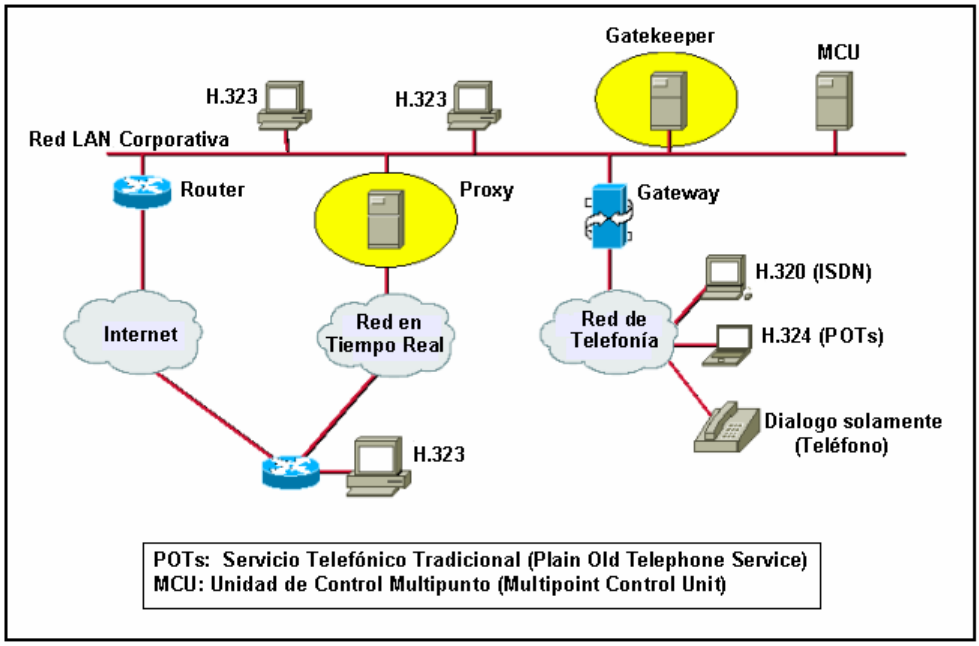


Figura. 4.1. Infraestructura H.323

Aunque se hable del H.323 como de un estándar, la ITU lo considera una recomendación, que está abierta a la interpretación de diferentes fabricantes, lo cual se puede considerar como una ventaja puesto que deja en libertad a los fabricantes para implementar capacidades que cumplan con los requerimientos de aplicaciones especiales. Con la compatibilidad de H.323, los productos y las aplicaciones multimedia de distintos fabricantes pueden interoperar, permitiéndoles a los usuarios comunicarse sin preocuparse de la compatibilidad del hardware. Como resultado, H.323 se considera “clave” para los productos del mercado de consumo basado en redes, negocios, entretenimiento y las aplicaciones profesionales.

El H.323 se fundamenta en las especificaciones del estándar H.320 y comprende varios documentos relacionados que describen los terminales, equipos, servicios e interacciones. Muchos de los componentes del H.320 se incluyen en el H.323, por lo que se lo puede considerar como una extensión del H.320. El estándar H.323 fue diseñado específicamente con las siguientes ideas:

- Basarse en los estándares existentes, incluyendo H.320, el Protocolo de Tiempo Real RTP y Q.931, entre otros.

- Incorporar algunas de las ventajas que las redes de conmutación de paquetes ofrecen para transportar datos en tiempo real.
- Solucionar la problemática que plantea el envío de datos en tiempo real sobre redes de conmutación de paquetes.

En resumen podemos decir que H.323 es un estándar amplio y comprensivo en su alcance, flexible y práctico en continua implementación y desarrollo, utilizado con gran énfasis en la comunicación multimedia sobre redes basadas en paquetes, especialmente IP.

En Enero de 1996, un grupo de fabricantes de soluciones de redes y de ordenadores propuso la creación de un nuevo estándar al Sector de Telecomunicaciones de la Unión de Telecomunicación Internacional ITU-T (International Telecommunication Union – Telecommunications Sector) para incorporar videoconferencia en la red LAN. Inicialmente, las investigaciones se centraron en las redes LAN, pues éstas son más fáciles de controlar, sin embargo, con la expansión del Internet, el grupo hubo de contemplar todas las redes IP dentro de una única recomendación, lo cual marcó el inicio del H.323.

En Mayo de 1997, el Grupo 15 de la ITU redefinió el estándar H.323 como la recomendación para "los sistemas multimedia de comunicaciones en aquellas situaciones en las que el medio de transporte sea una red de conmutación de paquetes que no pueda proporcionar una calidad de servicio garantizada".
El alcance de H.323 puede resumirse en las siguientes categorías:

- Soporte para conferencia multipunto y punto a punto
- La interoperabilidad de la Internet
- Las capacidades del cliente heterogéneas
- Codificadores de audio y video
- Manejo y soporte de contabilidad
- La seguridad

- Los servicios suplementarios

Soporte para conferencia multipunto y punto a punto

Las conferencias de H.323 pueden establecerse entre dos o más clientes sin ningún software de control multipunto especializado o hardware. Sin embargo, cuando la Unidad de Control Multipunto MCU (Multipoint Control Unit) es usada, H.323 soporta una topología flexible para las conferencias de multipunto. Una conferencia de multipunto puede centralizarse donde los nuevos participantes pueden unirse a otros en la conferencia. Éste es una aproximación a la topología de árbol flexible.

La interoperabilidad de la Internet

Los clientes de H.323 son interoperables con los clientes de la red de circuito conmutador SCN (Switching Circuit Network) tal como aquellos basados en las recomendaciones H.320 para la red ISDN, H.321 para la red de ATM y H.324 para la red PSTN y redes inalámbricas. Se puede decir simplemente que H.323 mantiene un protocolo común para que los productos de comunicaciones ofrecidos por vendedores diferentes puedan trabajar juntos, es decir, interoperar. Las aplicaciones funcionan transparentemente y son independientes de la red, plataforma y aplicación.

Las capacidades heterogéneas del cliente

Un cliente de H.323 debe soportar la comunicación de audio (obligatoria), el video y de los datos (opcional). Esta heterogeneidad y la flexibilidad no hacen a los clientes incompatibles. Durante el establecimiento de la llamada las capacidades son intercambiadas y la comunicación establecida basándose en el denominador común más bajo.

Codificadores de audio y video

H.323 especifica un codificador de audio y video. Sin embargo no hay ninguna restricción en el uso de otros codificadores y dos clientes pueden estar de acuerdo en cualquier codificador, el cual debe ser soportado por los dos.

Manejo y soporte de contabilidad

Las llamadas de H.323 pueden restringirse en una red basándose en:

- El número de llamadas en progreso.
- Limitaciones de ancho de banda.
- Restricciones de tiempo.

Usando estas políticas el gerente de la red pueden manejar el tráfico de H.323, además este estándar también proporciona medios de contabilidad que pueden usarse para propósitos de facturación.

La seguridad

La recomendación H.235 especifica los requisitos de seguridad para las comunicaciones de H.323. Se proporcionan cuatro servicios de seguridad:

- La autenticación
- La integridad
- La privacidad
- No rechazo

La autenticación es proporcionada por el control de admisión de puntos terminales. Esta es manejada por el gatekeeper que administra la zona.
La integridad de los datos y la privacidad son proporcionadas por la encriptación.
El no rechazo asegura que ningún punto terminal puede negar que fue participe en la llamada. Esta característica es también proporcionada por los servicios del gatekeeper.

Para implementar estos servicios de seguridad de H.235 puede usarse los estándares existentes como Seguridad de IP IPSec (IP Security) y Seguridad de Capa de Transporte TLS (Transport Layer Security).

Los servicios suplementarios

H.323 es un puente de conexión entre la red de telefonía tradicional y la red IP, que presenta un gran potencial para los nuevos servicios y aplicaciones ya que se aprovechan las capacidades de ambas redes. Estos servicios pueden agregar valor a los servicios de teléfono tradicionales como la transferencia de llamadas, mensajería integrada (e-mail, correo de voz, facsímil, mensajería instantánea, etc).

H.323 proporciona una arquitectura flexible para los servicios suplementarios a través de las recomendaciones de la serie de H.450.x.
H.450 adopta una arquitectura jerárquica para el desarrollo de nuevos servicios. Los Nuevos servicios pueden ser desarrollados por los usuarios finales combinando uno o más servicios básicos
En resumen, se puede decir, que H.323 reconoce el gran potencial para aplicaciones basadas en la telefonía IP y multimedia y mantiene un estructura básica para desarrollar tales servicios como son:.

- Transferencia de llamada
- Reenvío de llamada
- Mensajería integrada

Otros estándares de referencia en H.323 como son:

- H.225 especifica los procedimientos para el control de llamada incluyendo la señalización de la llamada, registro de la llamada y admisiones, formato de paquete de medios de comunicación, y sincronización.
- H.245 especifica los procedimientos para el intercambio de capacidad, negociación del canal, es decir los mensajes de abierto y cerrado de los canales y otros comandos, requisitos e indicaciones, y control de flujo.
- H.450.x es la serie de los procedimientos para los servicios suplementarios, definición de la señalización y los procedimientos utilizados para proveer servicios similares a los telefónicos.
- H.246 especifica los procedimientos para la interoperabilidad de los terminales a través de las gateways.
- H.235 proporciona los procedimientos de la seguridad y encriptación para proveer autentificación y cifrado en sistemas H.323.
- T.120 especifica conferencias de datos en tiempo real, punto a punto y multipunto. Provee interoperabilidad en los niveles de aplicación, de transporte y de red.
- H.332 para la provisión de conferencias a gran escala basadas en H.323.

Otros estándares de referencia incluyen codificadores de los medios de comunicación para audio (como, G.711, G.723.1, G.729, G.728 y G.722) y video (como, H.261 y H.263).

4.1.1 La arquitectura de H.323

Coma ya se menciono anteriormente, H.323 especifica los componentes, protocolos y procedimientos para la comunicación multimedia multipunto y punto a punto en tiempo y también establece las pautas de la interoperabilidad para la comunicación entre redes habilitadas con H.323 y la familia H.32X basado en los estándares de la conferencia.

En general H.323 implementa una aplicación donde se requieren cuatro entidades lógicas o componentes, estos son:

- Los terminales H.323
- Las gateways
- Los gatekeepers o gatekeepers
- Las Unidades de Control Multipunto MCU (Multipoint Control Unit).

Los terminales, gateways y las unidades de control multipunto MCUs son colectivamente conocidos como puntos terminales.

Aunque una red H.323 puede establecerse sólo con terminales, los otros componentes son esenciales para proporcionar mayor utilidad práctica a los servicios.

La figura muestra un esquema general de una Internet habilitada con H.323 con todos los componentes necesarios.

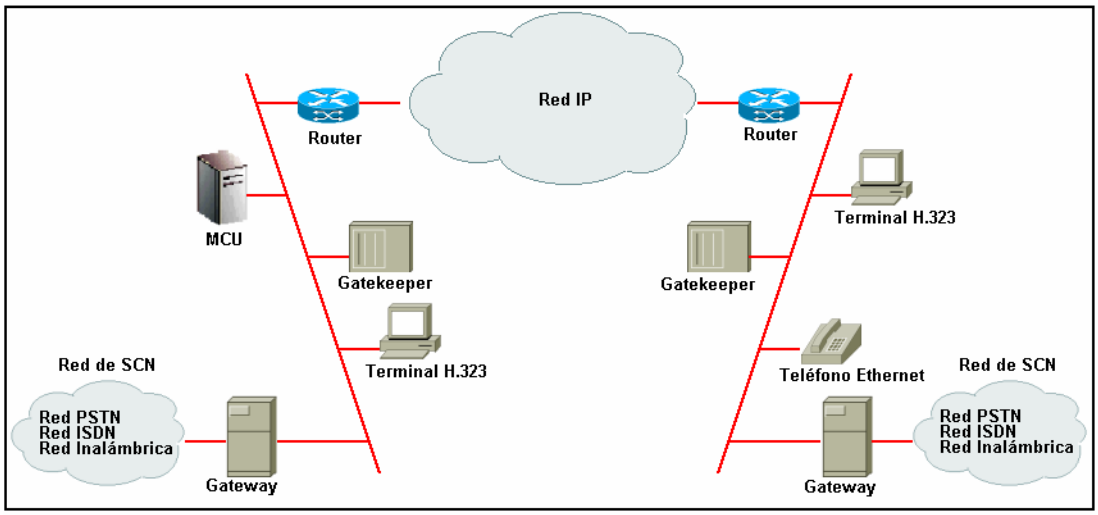


Figura. 4.2. El esquema de Internet H.323

4.1.1.1 Terminal H.323

Un terminal es dónde las cadenas de datos de H.323 y la señalización se originan y finaliza.

Puede ser un PC multimedia con una pila H.323 o un dispositivo como un teléfono IP de Bus de Serie Universal USB (Universal Serial Bus), en otras palabras, un dispositivo autosuficiente que ejecute H.323 y las aplicaciones multimedia.

Un terminal debe soportar comunicación de audio, video y el soporte de comunicación de datos es optativo.

El terminal H.323 proporciona comunicaciones bidireccionales en tiempo real interactuando con otro terminal H.323, gateway o Unidad de Control Multipunto MCU.

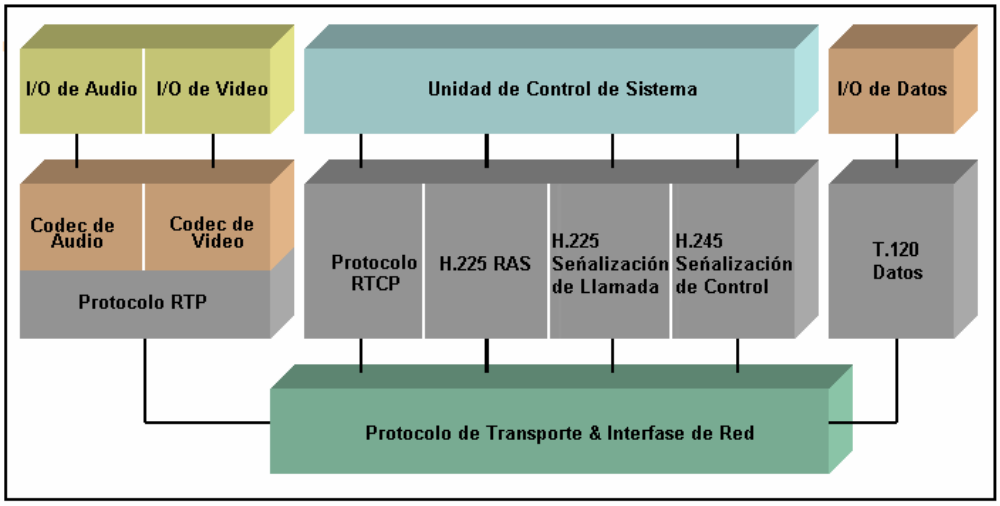


Figura. 4.3. La Pila Protocolar del terminal H.323

4.1.1.2 Gateway

Una gateway de H.323 es un componente optativo en una red H.323. Sin embargo, cuando la comunicación se requiere entre diferente redes, una gateway se necesita en la interfase como un puente entre las redes que se van a comunicar. Una gateway de H.323 proporciona la conectividad entre un terminal de H.323 y una red no H.323. A través de la provisión de gateways en H.323 es posible para los terminales de H.323 interoperar. Por ejemplo, es posible para un terminal de H.323 establecer una conferencia con terminales basados en H.320 o H.324 a través de una gateway apropiada o entre un terminal de H.323 y la red de circuitos conmutados SCN.

Es decir que la gateway proporciona la característica de “interoperabilidad” en el estándar H.323, haciendo viable la integración de servicios aun con plataformas diferentes. Una gateway proporciona las siguientes funciones:

- La translación de formato de los datos
- La translación de la señalización de control
- La translación del codificador de audio y video
- Establecimiento de llamada y terminación en ambos lados de la red.

Dependiendo del tipo de red en la que la translación se requiere la gateway pueden soportar H.310, H.320, H.321, H.322, o puntos terminales de H.324.
Las gateways si cursan información de usuario deben soportar en el Protocolo en Tiempo Real RTP (Real Time Protocol), Protocolo de Datagrama de Usuario UDP (User Datagram Protocol) y el Protocolo de Internet IP.

En resumen se puede decir que el propósito de la gateway es reflejar transparentemente las características de un extremo en la red IP a otro en una red conmutada y viceversa. Una gateway no se requiere, sin embargo, para la comunicación entre dos terminales en una red de H.323.

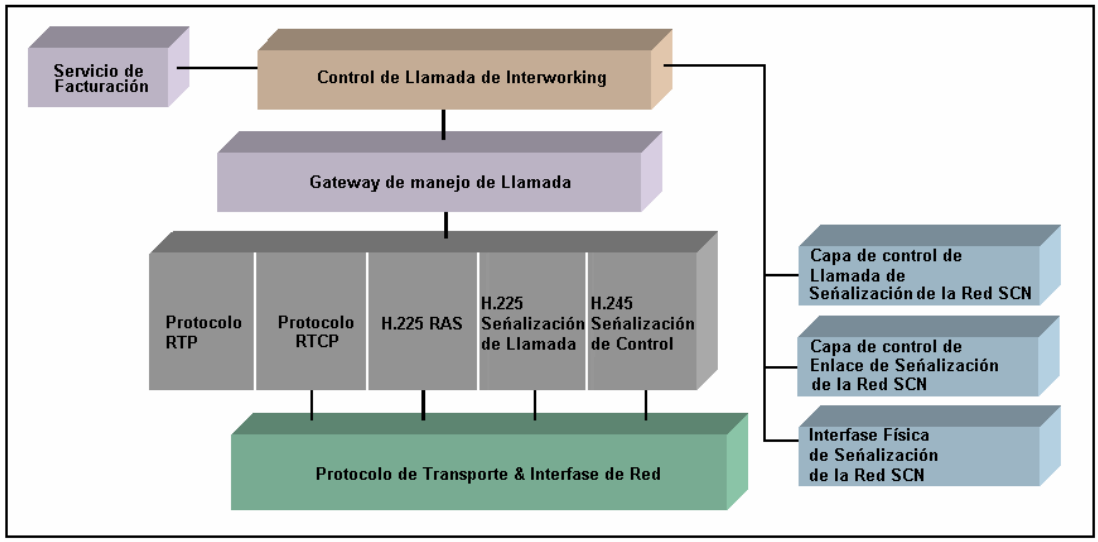


Figura. 4.4. La Pila Protocolar de Gateway

4.1.1.3 Gatekeeper

Un gatekeeper puede ser considerado el cerebro de la red de H.323 y es el más complejo, podemos decir que son entidades de control y señalización que

proporcionan los servicios que no pueden ser descentralizados e implementados en los puntos terminales (terminales, gateways, unidades MCUs). En otras palabras, es el punto focal para todas las llamadas dentro de la red de H.323. Aunque no son obligatorios, los gatekeepers proporcionan servicios importantes tales como:

- Control de acceso y administración de recursos
- Autorización de llamadas
- Traducción de direcciones de transporte entre direcciones IP, alias y números E.164
- Manejo del ancho de banda
- Manejo de zonas H.323
- Autenticación de terminales y gateways
- Grabación de detalles de Llamadas CDR (Call Details Record)
- Facturación y cobro

Un gatekeeper es muy útil para asegurar fiabilidad de las comunicaciones comercialmente factibles. Cuando un gatekeeper existe todos los puntos terminales deben registrarse con él. Los mensajes de control de los puntos terminales registrados se direccionan a través de él.

Para desarrollar las funciones anteriormente mencionadas, entre el gatekeeper y un punto terminal se emplea el Protocolo de Registro, Admisión y Estado RAS (Registration Admission and Status) sobre el Protocolo de Datagrama de Usuario UDP. La señalización de control de llamada puede o no estar soportada a través del gatekeeper. Si ésta es soporta, entonces el gatekeeper debe soportar el protocolo Q.931 (sobre el Protocolo de Control de Transmisión TCP) según establece H.225.

Un gatekeeper y sus puntos terminales definen una zona de H.323, de manera que en entornos de redes LANs es suficiente un gatekeeper, pero en

otros entornos de redes IP no basta con uno solamente, sino que se requerirán varios, cada uno definiendo una zona de H.323. Lógicamente, entre gatekeepers se requerirá comunicación, comportándose de cierta manera como un conmutador virtual.

Si bien el gatekeeper no es obligatorio su empleo en un entorno de H.323 sí, posibilitando el empleo más eficientemente de la plataforma H.323, por ejemplo, mediante el direccionamiento de llamadas a su través.

Un gatekeeper proporciona varias funciones básicas (obligatorias) a todos los puntos terminales en su zona. Estas funciones incluyen:

- ✓ **La traslación de dirección:** Es el método por el cual un alías de la dirección, por ejemplo, dirección de e-mail, es trasladada a una dirección de transporte. Un gatekeeper mantiene una base de datos para la traslación entre los alías, como son los números de teléfono internacionales y direcciones de la red.
- ✓ **La admisión y control de acceso de puntos terminales:** Este control puede ser basado en la disponibilidad de ancho de banda, limitaciones en el número de llamadas simultáneas de H.323 o el registro privilegiado de puntos terminales. El control de admisión es una manera de limitar el acceso de H.323 a una red o a una conferencia particular usando los mensajes Registro, Admisión y Estado RAS definidos en la recomendación H.225.
- ✓ **Manejo de la zona de H.323:** Se refiere a las funciones anteriores para los terminales, gateways y unidades MCUs registrados en su zona de control.
- ✓ **Manejo del ancho de banda:** Los administradores de la red pueden manejar en ancho de banda especificando, las limitaciones en el número de llamadas simultáneas y limitar la autorización de terminales específicos para poner las llamadas en momentos especificados. Un gatekeeper también maneja la asignación del ancho de banda al punto terminal de H.323 usando los mensajes de registro, admisión y estado RAS.

- ✓ **La capacidad de direccionamiento:** Un gatekeeper puede dirigir todas las llamadas originadas o terminadas en su zona. Esta capacidad proporciona numerosas ventajas:
  - ❖ Puede mantenerse información de contabilidad de llamadas que pueden ser usada para la facturación y propósitos de seguridad.
  - ❖ Un gatekeeper puede redireccionar una llamada a una gateway apropiada basado en la disponibilidad del ancho de banda. El redireccionamiento puede usarse para desarrollar servicios avanzados como el direccionamiento móvil, reenvío de llamada y correo de voz.

Adicionalmente el gatekeeper proporciona funciones no básicas (opcionales) tales como:

- ➢ **Señalización de control de llamada:** En conferencias punto a punto el gatekeeper puede procesar mensajes Q.931
- ➢ **Autorización de llamadas:** Sobre la base de la especificación Q.931 el gatekeeper puede rechazar llamadas, por ejemplo, acceso restringido a y desde determinados terminales o gateways, restringir el acceso durante ciertos períodos de tiempo, etc.
- ➢ **Manejo de llamadas:** El gatekeeper puede, mediante una tabla propia, disponer del estado de las llamadas H.323, pudiendo saber el estado de un terminal dado y el "gasto" de ancho de banda asociado a las llamadas en curso y proceder en consecuencia respecto a las nuevas solicitudes.

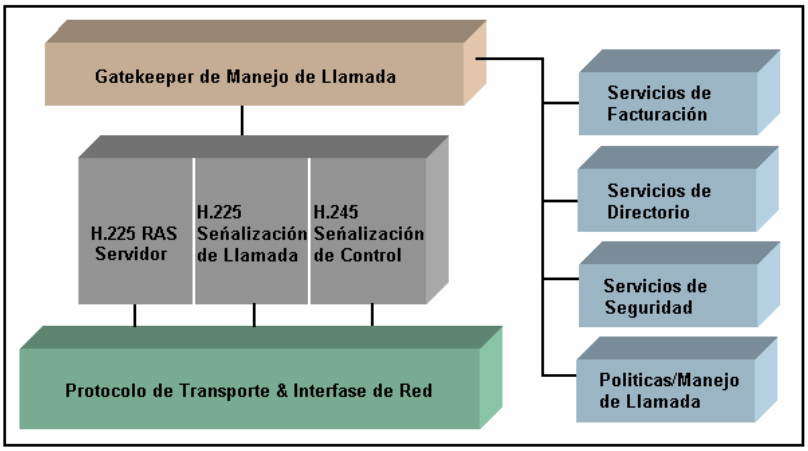


Figura. 4.5. La Pila Protocolar del Gatekeeper

4.1.1.4 Unidad de Control Multipunto MCU (Multipoint Control Unit)

Una Unidad de Control Multipunto MCU habilita la conferencia entre tres o más puntos terminales, la cual se forma de dos partes:

- Un Controlador Multipunto MC (Multipoint Controller) obligatorio
- Un Procesador Multipunto MP (Multipoint Processor) opcional

En el caso más simple, una unidad MCU puede estar formada por un controlador multipunto MC únicamente. La unidad MCU puede combinarse con un terminal, gateway o gatekeeper y es un componente optativo de una red H.323. La unidad MCU es requerida en una conferencia de multipunto centralizada donde cada terminal establece una conexión punto a punto con la unidad MCU.

La unidad MCU determina las capacidades de cada terminal y envía stream media mixta. En el modelo descentralizado de conferencia multipunto, el controlador multipunto MC asegura la compatibilidad de comunicación pero las stream medias son multicast y el mezclando se realiza en cada terminal.

Controlador Multipunto MC (Multipoint Controller)

Un controlador multipunto MC es una entidad de H.323 que proporciona las capacidades de negociación entre todos los terminales para conseguir la comunicación y puede controlar los recursos de la conferencia tales como el vídeo multicast. El controlador multipunto MC proporciona una localización centralizada para el establecimiento de la llamada multipunto. La señalización de control y de llamada son dirigidas a través del controlador multipunto MC para que puedan determinarse las capacidades de los puntos terminales y los parámetros de comunicación negociados. Un MC también puede usarse en una llamada punto a punto, la cual puede extenderse después en una conferencia de multipunto. Otro

trabajo útil del MC es determinar stream de unicast o multicast de audio y video dependiendo de la capacidad de la red subyacente y la topología de la conferencia multipunto.

Procesador Multipunto MP (Multipoint Processor)

Un procesador multipunto MP es la entidad de H.323 cuyo hardware y software especializado mezclan, conmutan y procesan el audio, vídeo y/o los datos de los participantes en una conferencia multipunto. El procesador multipunto MP puede procesar una única secuencia multimedia o varias simultáneamente, dependiente del tipo de conferencia soportada.

4.1.1.5 Proxy H.323

Es un servidor proxy con soporte H.323 que proporciona acceso a los usuarios de una red segura a otra utilizando la información que cumpla las recomendaciones del estándar H.323. Un proxy H.323 se comporta como dos extremos H.323 pasando mensajes de establecimiento de llamadas e información en tiempo real a un destino situado en la parte segura de un firewall (sistema diseñado para prevenir el acceso no autorizado a o desde una red). Puede estar integrado con otro dispositivo de seguridad o entidades H.323 como gateways, firewalls, etc.

4.1.1.6 Zona de H.323

Una zona de H.323 es un conjunto de todos los terminales, gateways y unidades MCU manejadas por un solo gatekeeper. Una zona incluye un terminal por lo menos y puede incluir gateways o unidades MCU pero sólo un gatekeeper.

Una zona puede ser independiente de la topología de la red y puede comprenderse de segmentos de la red múltiples que se conectan usando routers u otros dispositivos.

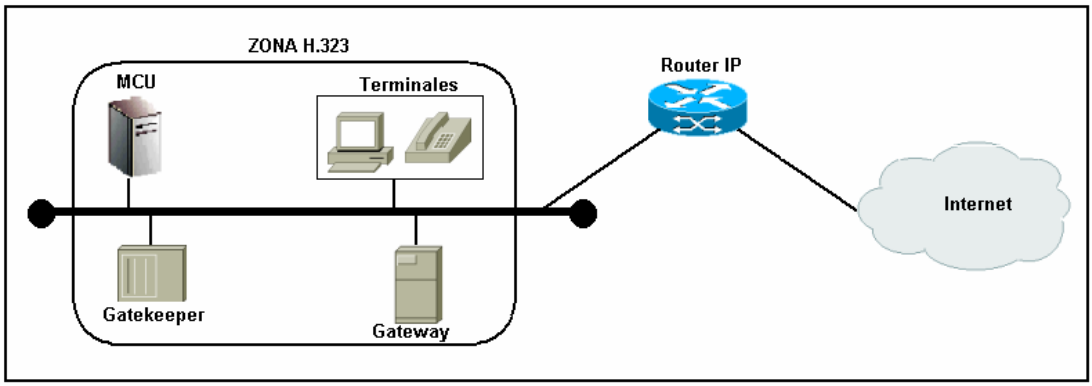


Figura. 4.6. Zona H.323

4.1.2 Tipos de Conferencias Multipunto H.323

Se definen la conferencia multipunto como una llamada entre tres o más usuarios. Una conferencia multipunto H.323 involucra a dos terminales, es decir un terminal y una unidad MCU o un terminal y una gateway. Durante estas conferencias, el control de la llamada y operaciones de los medios de comunicación pueden volverse considerablemente más complejas que durante una conferencia de punto a punto simple. La coordinación y notificación de participantes entrando y dejando una conferencia junto con la clasificación de los medios de comunicación requieren la presencia de por lo menos un controlador multipunto MC.

Tenemos 2 modelos de conferencias multipunto, las cuales difieren en su manejo de los medios de comunicación en tiempo real y son:

- Conferencia Multipunto Descentralizada

- Conferencia Multipunto Centralizada

4.1.2.1 Conferencia Multipunto Descentralizada

Los terminales H.323 pueden recibir más de un canal de audio y vídeo simultáneamente, en estos casos, los terminales H.323 pueden necesitar realizar tareas de mezcla y conmutación para presentarle al usuario la señal de vídeo adecuada en cada momento.

Una conferencia multipunto descentralizada es aquella en la que los terminales participantes envían en modo multicast (emisión simultánea de varios canales) sus señales de audio y vídeo a todos los demás terminales. No hay una unidad MCU involucrada en esta tarea, lo cual se convierte en una ventaja puesto que este es un recurso potencialmente caro y limitado.

Los terminales son los responsables de:

- Absorber las secuencias recibidas de audio
- Seleccionar una o más de las secuencias recibidas para mostrarlas.

En este caso, no se requiere el Procesador Multipunto MP de audio y vídeo. La mayor desventaja de este tipo de conferencia multipunto es la ineficacia si más de cuatro entidades participan en la conferencia.

Figura. 4.7. Conferencia Multipunto Descentralizada

4.1.2.2 Conferencia Multipunto Centralizada

El modelo de conferencia multipunto centralizado opera en la misma forma como una conferencia basada en circuito (por ejemplo, H.320). En este modelo, todo el audio y el video se transmiten a una unidad MCU central que mezcla audios múltiples, selecciona el video correspondiente y retransmite el resultado a todos los participantes, es decir, todos los terminales participantes se comunican en modo punto a punto con una unidad MCU. El controlador multipunto MC que están en la unidad MCU centraliza y administra la conferencia. El procesador multipunto MP, también incluido en la unidad MCU, procesa las señales de audio, vídeo y/o datos, devolviendo a cada terminal la secuencia procesada. Las ventajas que presenta son:

- Los puntos terminales que participan en este tipo de conferencia no requieren ser muy poderosos.
- Cada punto terminal sólo tiene que codificar y descodificar el conjunto enviado por la unidad MCU.
- La unidad MCU puede proporcionar una conferencia especializada o de alto desarrollo.

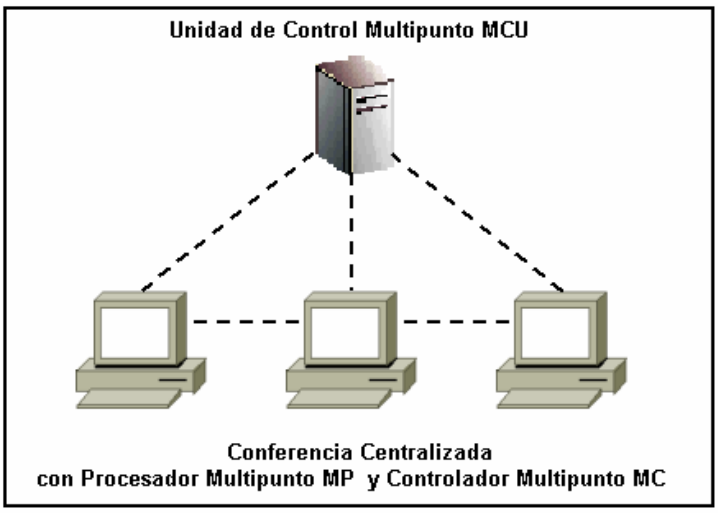


Figura. 4.8. Conferencia Multipunto Centralizada

4.1.3 La Pila Protocolar de H.323

La pila protocolar de H.323 (figura 4.3) se ejecuta sobre la capa de transporte y capa de la red. Si la red subyacente es basada en el Protocolo IP entonces el audio, el video y los paquetes de registro, admisión y estado RAS de H.225.0 usan el protocolo de datagrama de usuario UDP para el transporte mientras los paquetes de datos y de control son transportados usando el protocolo de control de transmisión TCP fiable.

A continuación una descripción más detallada de los elementos de la pila protocolar de H.323.

4.1.3.1 Codificadores/Decodificadores de audio

Un codificador/decodificador, generalmente llamado “codec”, de audio pone en código la señal de audio del micrófono para la transmisión en el terminal emisor de H.323 y descodifica el código de audio recibido que se envía en el terminal receptor de H.323.

Debido a que el audio es el servicio mínimo proporcionado por el estándar de H.323, todos los terminales de H.323 deben tener un codec de audio como soporte.

H.323 especifica una serie de codificadores de audio en una tasa desde 5.3 has 64 Kbps.

CODECS DE AUDIO	
G.711	Utiliza la técnica de Modulación del Código de Pulso PCM (Pulse Code Modulation) para digitalizar la señal de voz. La tasa de transmisión es de 64 Kbps.
G.722	Usa una variante de la técnica Modulación del Código de Pulso Diferencial Adaptativo ADPCM (Adaptative Differential Pulse Code Modulation) denominada Sub Modulación del Código de Pulso Diferencial Adaptativo Sub-ADPCM (Sub-Band Adaptative Differential Pulse Code Modulation) para digitalizar señales de 7 KHz a una tasa de transmisión de 48, 54 y 64 Kbps.
G.723.1	Opera a una frecuencia de 5.3 y 6.3 Kbps, con una mayor calidad para una taza más alta, este estándar ha surgido como la base para la interoperabilidad entre redes de diversas plataformas.
G.728	Utiliza la técnica de Bajo Retardo de Predicción Lineal Excitada de Código LD-CELD (Low Delay Codebook Excited Linear Prediction), la cual es una técnica híbrida de codificación por vocalización y codificación por forma de onda. La señal de voz que digitaliza esta limitada a 4 KHz a una tasa de 16 Kbps.
G.729	Usa una técnica denominada Estructura Algebraica Conjugada de Predicción Lineal Excitada de Código AS-ACELP (Conjugate Structure Algebraic Codebook Excited Linear Prediction) para codificar una señal analógica a una señal digital a 8 Kbps.

Tabla. 4.1. Estándares de Codecs de Audio en H.323

4.1.3.2 Codificadores/Decodificadores de video

Un codec de video pone en código el video de la cámara para la transmisión en el terminal emisor de H.323 y descodifica el video recibido que se envía al display de video en el terminal receptor de H.323.

Debido a que H.323 especifica el soporte de video como optativo, el soporte de codec de video también es optativo.

H.323 especifica dos codificadores de videos fundamentales: H.261 y H.263, pero pueden usarse otros codificadores con tal de que ambos terminales están de acuerdo y lo soporten.

CODECS DE VIDEO	
H.261	Codificador de video para los servicios audiovisuales en 64 Kbps.
H.263	Codificador de video para la comunicación de tasa de bits baja sin la pérdida de calidad.

Tabla. 4.2. Estándares de Codecs de Video en H.323

El estándar H.261 soporta dos formatos de video, que son:

- El formato de Intermedio común CIF (Common Intermediate Format) con una resolución de 352x288 píxeles.
- El formato de Intermedio común cuádruple QCIF (Quarter Common Intermediate Format) que tiene una resolución de 176x144 píxeles.

El estándar H.263 soporta los formatos de videos:

- El sub formato de Intermedio común cuádruple sub-QCIF (Quarter Common Intermediate Format) que tiene una resolución de 128x96 píxeles.
- El formato de Intermedio común cuádruple QCIF (Quarter Common Intermediate Format) que tiene una resolución de 176x144 píxeles.
- El formato de Intermedio común CIF (Common Intermediate Format) con una resolución de 352x 244.
- El formato de Intermedio común 4CIF con una resolución de 702 x 576 píxeles.
- El formato de Intermedio común 16CIF con una resolución de 1408 x 1152 píxeles.

Los tres primeros son obligatorios mientras los dos siguientes son opcionales. A través del formato de intermedio común cuádruple QCIF de H.263 se hace compatible con H.261.

4.1.3.3 Estándar H.225 de Registro, Admisión y Estado RAS

El Registro, Admisión y Estado RAS es el estándar entre los puntos terminales (los terminales y gateways) y los gatekeepers. El estándar RAS H.225 es sólo necesario cuando un gatekeeper existe, se usa para:

- El descubrimiento del gatekeeper
- El registro del punto terminal

- El control de admisión
- Los cambios del ancho de banda
- Los procedimientos desconexión entre los puntos terminales y los gatekeepers

Un canal RAS es usado para intercambiar los mensajes RAS. Este canal de señalización se abre entre un punto terminal y un gatekeeper para el establecimiento de cualquier otro canal.

4.1.3.4 Estándar H.225 de Señalización de Llamada

La señalización de llamada H.225 se usa para establecer una conexión entre dos puntos terminales de H.323 (terminales y gateways). Esto se logra intercambiando los mensajes protocolares de H225 en el canal de señalización de llamada fiable. Por ejemplo, los mensajes protocolares de H.225 son llevados sobre el protocolo TCP en un red H.323 basada en IP.

Los mensajes de H.225 son intercambiados entre los puntos terminales si no hay ningún gatekeeper en la red de H.323. Cuando un gatekeeper existe en la red, los mensajes de H.225 se intercambian directamente entre los puntos terminales o entre los puntos terminales después de que son direccionados a través del gatekeeper.

En el primer caso la señalización de llamada es directa.

El segundo caso se llama señalización de llamada direccionada por el gatekeeper.

El método escogido se decide por el gatekeeper durante el intercambio de mensajes RAS.

4.1.3.5 Estándar H.245 de Señalización de Control

La flexibilidad de H.323 requiere que los puntos terminales negocien para determinar los seteos compatibles antes de que los enlaces de comunicación de audio, video, y/o datos puedan ser establecidos.

La señalización de control de H.245 consiste en el intercambio de extremo a extremo de los mensajes de H.245 entre los puntos terminales de H.323. La implementación del control de H.245 es obligatoria en todo los puntos terminales.

4.1.3.6 Protocolo de Transporte en Tiempo Real RTP

El estándar H.323 emplea el protocolo RTP para transportar voz, vídeo y datos, proporcionando además la información de control necesaria para la recuperación en tiempo real de los medios en el receptor. El protocolo RTP no garantiza reserva de recursos ni da garantía de la calidad de servicio QoS, su papel fundamental es actuar como una interfase entre aplicaciones de tiempo real y los protocolos de la capa de transporte, sin establecer que protocolo de ésta siendo usado, aunque típicamente se soporta sobre el protocolo UDP.
El protocolo de transporte en tiempo real RTP proporciona:

- El tipo de identificación de carga útil
- La secuencia de numeración
- La entrega de monitoreo

Mientras que el protocolo de datagrama de usuario UDP proporciona:

- Servicios de múltiplexación
- Servicios de chequeo

4.1.3.7 Protocolo de Control de Transporte en Tiempo Real RTCP

El protocolo de control RTCP es el colega del protocolo RTP que proporciona los servicios de control. La función primaria del protocolo RTCP es proporcionar la regeneración en la calidad de la distribución de los datos. Otras funciones del protocolo RTCP incluyen transportar un identificador de transporte de nivel para un recurso del protocolo RTP, llamado nombre canónico, el cual es usado por los receptores para sincronizar audio y video.

4.1.3.8 Protocolo de Conferencia de Datos T.120

Se requiere la capacidad de conferencia de datos en tiempo real para actividades como:

- Aplicación compartida
- El whiteboard compartido
- La transferencia de archivos
- La transmisión de fax
- La mensajería instantánea

El estándar T.120 proporciona esta capacidad optativa a H.323. T.120 es un protocolo de comunicación de datos en tiempo real diseñado específicamente para las necesidades de conferencia entre varios clientes a través de diferentes redes.

T.120 proporciona varias ventajas sobre la transmisión de los datos regulares como son:

- ✓ **Soporte de conferencia multipunto:** T.120 soporta la entrega de datos multipunto que habilita las actividades de colaboración de grupo. La unidad

MCU maneja el mezclando y conmutación de datos de una manera similar que para el video y audio.

- ✓ **Independencia de la plataforma y de la red:** T.120 opera en la parte superior de la capa de transporte de red, así es transparente e independiente del hardware de la red y software usado.
- ✓ **Interoperabilidad:** T.120 es la referencia de todas los estándares de conferencia de H.32X, asegurando un grado alto de interoperabilidad al nivel de aplicación.
- ✓ **Soporte Multicast:** T.120 soporta multicast de datos en las redes. Este soporte es flexible, el mezclado de unicast y multicast también es posible durante una conferencia.

Otros beneficios que proporciona T.120 es la capacidad de corrección de error sobre la capa transporte de red que asegura la entrega fiable, posee una arquitectura escalable y extensible con provisiones para añadir nuevas aplicaciones que se aprovechan de la entrega de los datos fiable y eficaz en tiempo real.

4.1.4 Llamada de H.323

El establecimiento de una llamada H.323 se lleva a cabo en tres fases:

- ✓ **Fase RAS:** Es la fase de intercambio de mensajes entre gatekeeper y punto terminal, para traslación de direcciones, autorización o rechazo de las llamadas y manejo del ancho de banda, básicamente.
- ✓ **Fase Q.931:** En esta fase se realiza el intercambio de mensajes entre punto terminal para el establecimiento de conexiones lógicas.
- ✓ **Fase H.245:** Es la fase de intercambio de mensajes entre puntos terminales para “acordar” el intercambio de información del usuario.

A continuación se describen los pasos involucrados en la creación de una llamada de H.323, establecimiento de la comunicación y terminación de la

llamada en forma mas detallada. Se supondrá que la red contiene dos terminales de H.323, T1 y T2, conectados a un gatekeeper y se asume además que la señalización de llamada es directa y se usa el protocolo RTP.

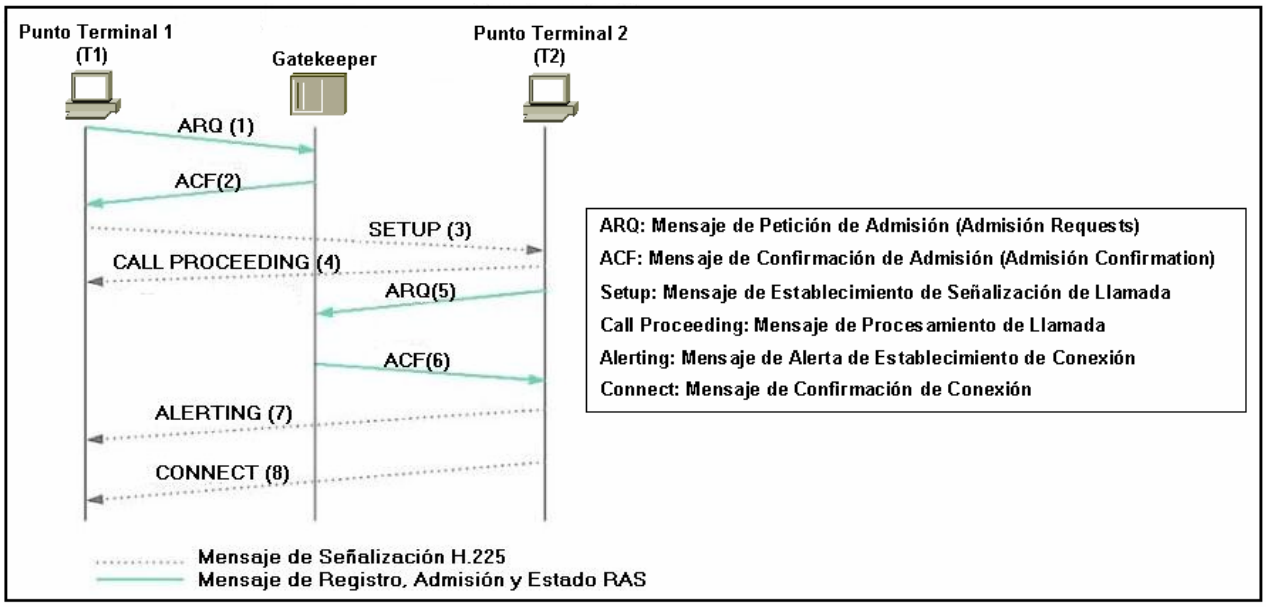


Figura. 4.9. El Establecimiento de una llamada de H.323

El punto terminal T1 le envía el mensaje de Petición de Admisión ARQ (Admisión Request) en el canal de RAS al gatekeeper para el registro. El punto terminal T1 pide el uso de señalización de la llamada directa.

El gatekeeper confirma la admisión del punto terminal T1 enviando un Mensaje de Confirmación de Admisión ACF (Admisión Confirmation) a T1. El gatekeeper indica en el mensaje de Confirmación de Admisión ACF que T1 puede usar la señalización de la llamada directa.

El punto terminal T1 envía un mensaje de establecimiento de señalización de llamada (Setup) de H.225 al punto terminal T2 que pide una conexión.

El punto terminal T2 responde con un mensaje de procedimiento de llamada (Call Proceeding) de H.225 al punto terminal T1.

Ahora T2 tiene que registrar con el gatekeeper. Le envía el mensaje de Petición de Admisión ARQ al gatekeeper en el canal de RAS.

El gatekeeper confirma el registro enviándole el mensaje de Confirmación de

Admisión ACF RAS al punto terminal T2.

T2 alerta a T1 del establecimiento de conexión enviando un mensaje de alerta (Alerting) de H.225.

Entonces T2 confirma el establecimiento de conexión enviando un mensaje de conexión (Connect) de H.225 a T1, y la llamada se establece.

El proceso continua con el flujo de señalización de control de H.323 que utiliza los mensajes H.245 para el establecimiento del canal de comunicación entre los puntos terminales T1 y T2.

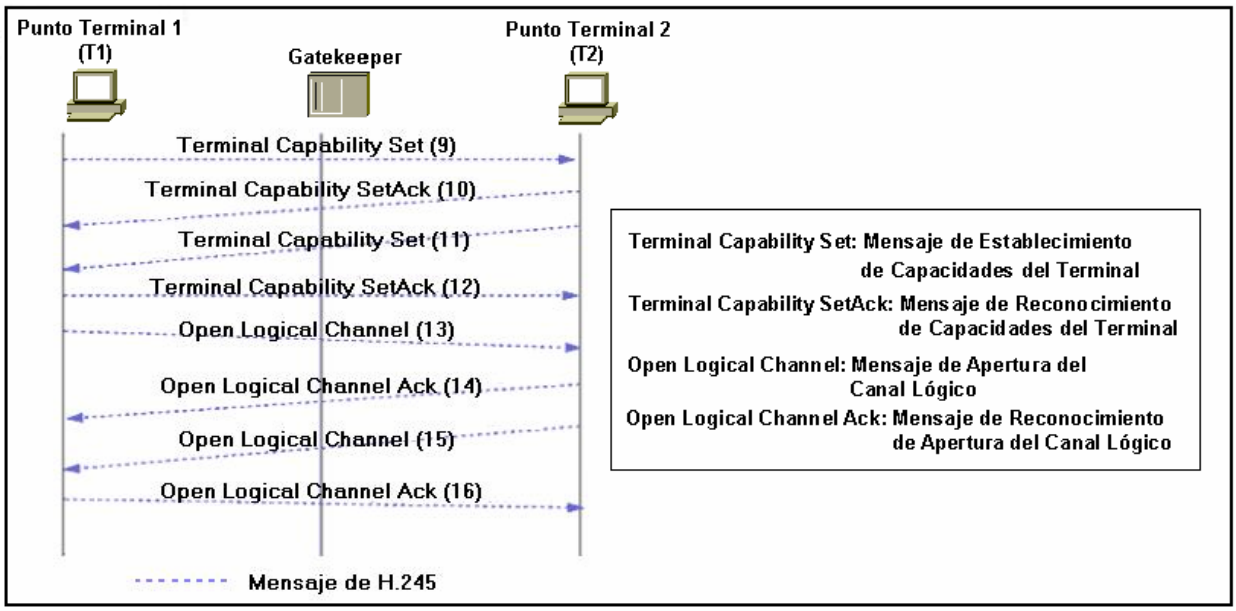
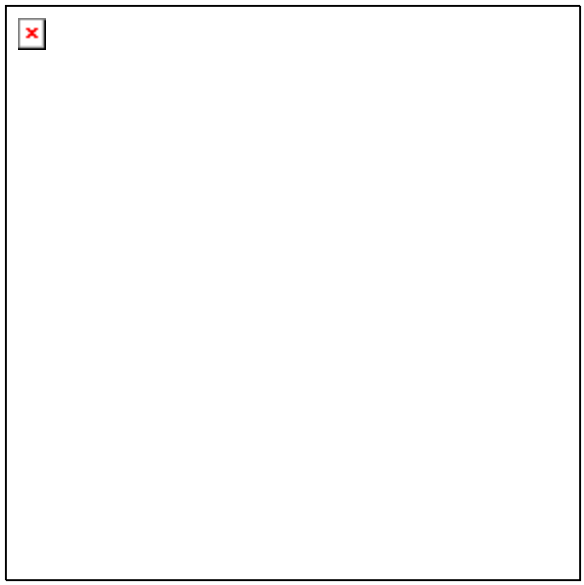


Figura. 4.10. El flujo de la Señalización de Control de H.323



Los canales de control de H.245 se establecen entre los puntos terminales T1 y T2. El punto terminal T1 le envía el mensaje de Establecimiento de Capacidades del Terminal (Terminal Capability Set) de H.245 al punto terminal T2 para intercambiar sus capacidades.

El punto terminal T2 reconoce las capacidades del punto terminal T1 enviándole el mensaje de Mensaje de Reconocimiento de Capacidades del Terminal (Terminal Capability SetAck) de H.245.

El punto terminal T2 intercambia sus capacidades con el punto terminal T1 enviándole el mensaje de Establecimiento de Capacidades del Terminal (Terminal Capability Set) de H.245.

El punto terminal T1 reconoce las capacidades de T2 enviándole el mensaje de Reconocimiento de Capacidades del Terminal (TerminalCapabilitySetAck) de H.245.

Así el punto terminal T1 abre un canal de comunicación con T2 enviando un mensaje de Apertura del Canal Lógico (Open Logical Channel) de H.245. La dirección de transporte del canal del protocolo de Control de Transporte en Tiempo Real RTCP es incluido en el mensaje.

T2 reconoce el establecimiento del canal lógico unidireccional de T1 a T2 enviando un mensaje de Reconocimiento de Apertura del Canal Lógico (Open Logical Channel Ack) de H.245. Incluido en el mensaje de reconocimiento de las direcciones de transporte del protocolo de Transporto en Tiempo Real RTP asignada por T2 para ser usado por el T1

para enviar las cadenas de los medios de comunicación del protocolo RTP y la dirección del protocolo RTCP recibida desde T1 antes.

Entonces, T2 abre un canal de los medios de comunicación con T1 enviando un mensaje de Apertura de Canal Lógico (Open Logical Channel) de H.245. La dirección de transporte del canal del protocolo RTCP es incluido en el mensaje.

T1 reconoce el establecimiento del canal lógico unidireccional de T2 a T1 enviando un mensaje de Reconocimiento de Apertura del Canal Lógico (Open Logical Channel Ack) de H.245. Incluido en el mensaje de reconocimiento de las direcciones de transporte del protocolo RTP asignada por T1 para ser usado por el T2 para enviar las cadenas de los medios de comunicación del protocolo RTP y la dirección del protocolo RTCP recibida antes desde T2.

Ahora las cadenas de comunicación bidireccionales están establecidas. partir de este punto los paquetes pueden ser enviados a través del protocolo RTP con el control que se realiza por medio del protocolo RTCP. La figura ilustra el flujo de paquetes y el flujo de control del protocoloRTCP.

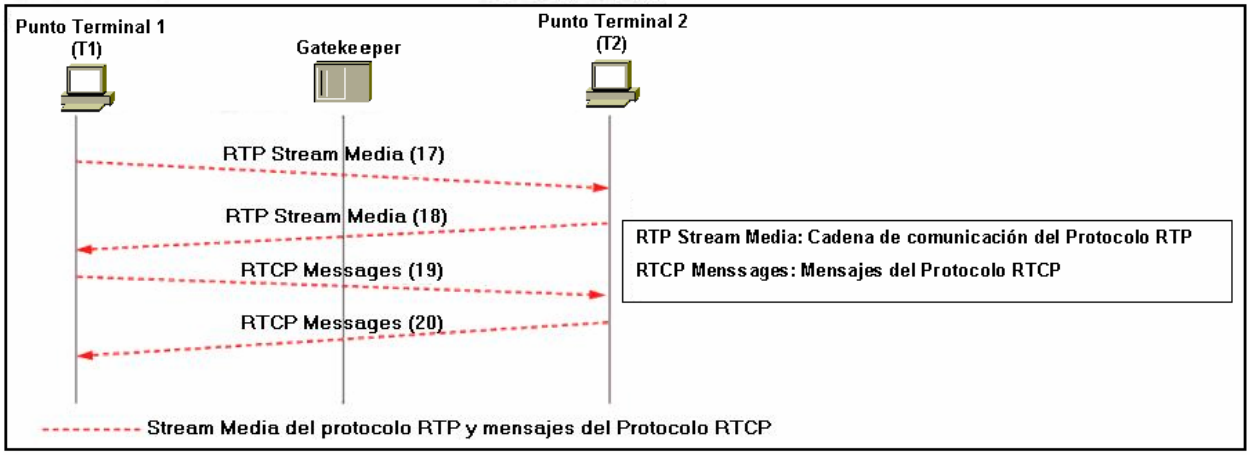


Figura. 4.11. Flujos de control de medios de comunicación de H.323

El punto terminal T1 envía una cadena de comunicación del protocolo

RTP al punto terminal T2.

El punto terminal T2 envía una cadena de comunicación del protocolo RTP al punto terminal T1.

El punto terminal T1 envía los mensajes del protocolo RTCP al punto terminal T2.

El punto terminal T2 envía los mensajes del protocolo RTCP al punto terminal T1.

Después se termina el intercambio de información entre los puntos terminales T1 y T2 para finalizar la llamada.

Este procedimiento involucra el cambio de mensajes H.225, H.245 y mensajes RAS como se indica en la figura, la cual ilustra los flujos de descargo de llamada.

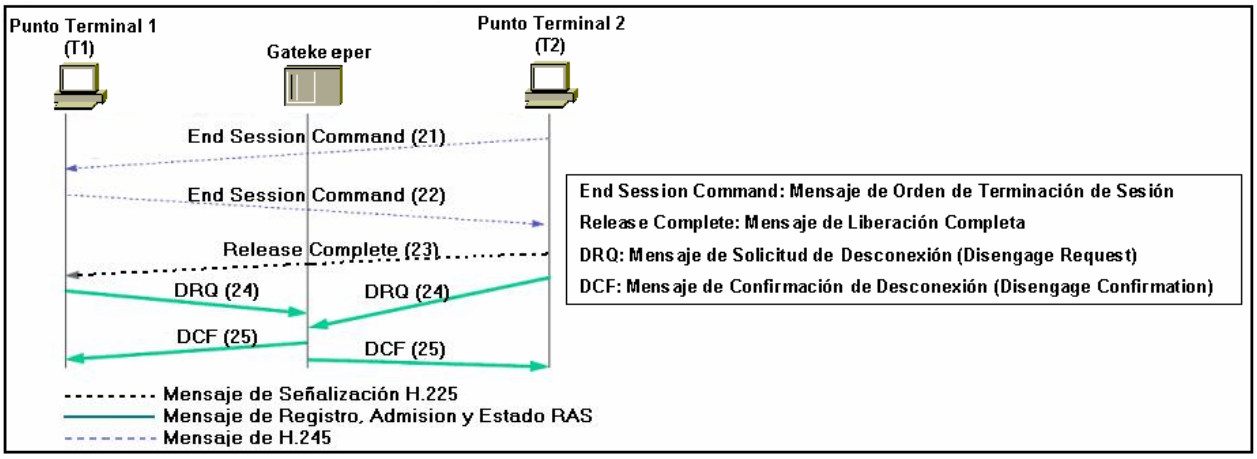
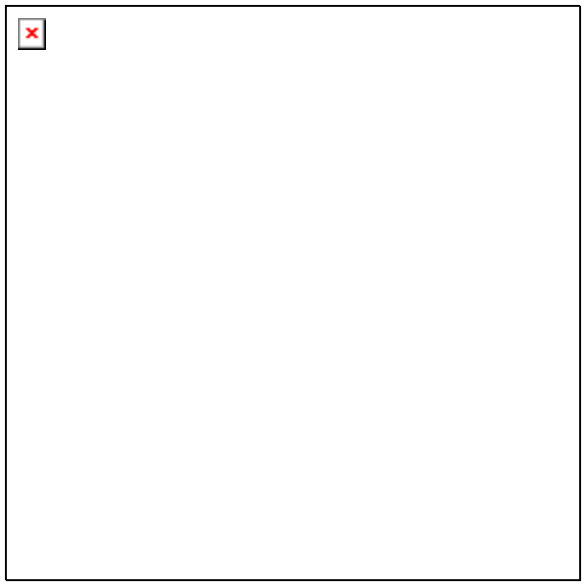


Figura. 4.12. Flujo de terminación de la llamada de H.323



El punto terminal T2 inicia la desconexión de la llamada. Enviando un mensaje de Orden de Terminación de Sesión (End Session Command) de H.245 al punto terminal T1.

El punto terminal T1 confirma la desconexión de la llamada enviando un mensaje de Orden de Terminación de Sesión (End Session Command) H.245 al punto terminal T2.

El punto terminal T2 completa la desconexión de la llamada enviando el mensaje de Liberación Completa (Release Complete) de H.225 a T1.

Los puntos terminales T1 y T2 se desconectan del guardabarrera enviándole un mensaje de Solicitud de Desconexión DRQ (Disengage Request) de RAS.

El guardabarrera desconecta a los puntos terminales T1 y T2 y confirma la desconexión enviando los mensajes de Confirmación de Desconexión DCF (Disengage Confirmation) a los puntos terminales T1 y T2.

4.1.5 Calidad de Servicio QoS

La calidad de video en las redes públicas como el Internet es pobre. Incluso el audio sobre Internet no es comparativamente favorablemente al de la red

PSTN. La calidad de servicio QoS es determinada por el retardo, la fluctuación del retardo o Jitter, la pérdida de paquetes, además, el retardo en el establecimiento de la llamada es otro parámetro que afecta los niveles de QoS requeridos.

H.323 es un protocolo de la capa más alta y puede usar cualquier QoS que se construye en las capas más bajas. Por ejemplo, en las redes de IP, H.323 puede hacer uso de la tecnología de Servicios Integrados IS (Integrated Services), la tecnología de Servicios Suplementarios DS (Differentiated Services), el protocolo RSVP. El desarrollo de codecs más eficaz también jugará un papel importante en el mejoramiento de la calidad.

A continuación describiremos algunas de las características que han mejorado la calidad de servicios en el estándar H.323

Retardo en el establecimiento de Llamada: H.323 usan procedimientos de señalización H.225/Q.931 para establecer una conexión entre el emisor y receptor de llamada. Dependiendo si un gatekeeper esta siendo usado o no, establecer una llamada con H.323 versión 1, puede tomar aproximadamente 6 a 7 tiempos de viaje completo porque el establecimiento de la llamada necesita conexión TCP primero y luego la conexión de llamada. En muchos casos el emisor no tiene que esperar por una contestación antes de enviar el próximo mensaje. H.323 versión 2 con Conexión Rápida (Fast Connect) puede usar un solo intercambio y puede reducir el retardo a 2.5 tiempos de viaje completo RTTs (Round Trip Times). H.323 versión 3 reduce esto a 1.5 RTTs. H.323 versión 3 establece una conexión de UDP y una conexión de TCP casi simultáneamente y proporciona un mecanismo eficaz para cerrar la conexión de TCP si el establecimiento de UDP tiene éxito. Si el establecimiento de UDP falla, TCP puede encargarse inmediatamente.

Actualmente H.323 versión 4 se utiliza en las videoconferencias, y el tiento RTT es aun menor lo cual representa una mejora en la Calidad de servicio en la red IP.

Soporte de Calidad de Servicio QoS para Flujos Multimedia: Los gatekeepers en H.323 proporcionan un rico conjunto de funciones de control y manejo, incluyendo la translación de dirección, control de admisión, control de ancho de banda y manejo de la zona.

El control de admisión determina si la red tiene recursos suficientes para soporta la QoS requerido para una llamada y acepta o rechaza la llamada como consecuencia. A fin de realizar el control de admisión, el protocolo debe manejar el dirección del ancho de banda, dirección de llamada y control del ancho de banda.

H.323 versión 3 pueden ofrecer algunos Servicios Diferenciados basados en negociación de parámetros QoS (tasa de bit, retardo y jitter). En la iniciación de una llamada, un terminal puede pedir uno de tres clases de servicio definido: "Servicio Garantizado", "Servicio Controlado", y "Servicio No especificado".

H.323 versión 4 define nuevos procedimientos para permitir el uso del protocolo RSVP cuando no se este usando Conexión Rápida.

Pérdida de paquetes: H.323 versión 1 y versión 2 son basadas en un protocolo de transporte fiable. El uso de TCP simplificaría la máquina de estado para el protocolo de control de llamada, puesto que tiene su propio control de flujo, control de ventana, y mecanismos del retransmisión para asegurar la fiabilidad. H.323 versión 3 y versión 4 especifican sus propias políticas de retransmisión para el emisor y receptor para soportar TCP y UDP, lo cual simplifica de manera la perdida de paquetes.

Detección del lazo de envió: Un tipo común de error son los lazos de reenvío de llamada, los cuales pueden ocurrir sobre todo cuando gatekeepers múltiples están envueltos en el establecimiento de una llamada. H.323 versión 3 definen un

campo de Valor de Ruta (Path Value) para indicar el número máximo de gatekeepers que el mensaje de señalización debe cruzar antes de eliminarse. Usando el campo de Valor de Ruta pueden reducir la tasa de ocurrencia del lazo. Cuando un lazo ocurre sin saber los nombres de gatekeepers, los mensajes de la señalización no se detendrán hasta localizar el valor de ruta. Además, cuando la configuración de la red cambia, el valor de ruta puede necesitar ser cambiado posiblemente.

Tolerancia a la falla: H.323 versión 3 proporcionan tolerancia a la falla mejor por gatekeepers redundantes y puntos terminales. Durante el registro, un gatekeeper puede indicar gatekeepers alternados al punto terminal del registro que pueden usarse en caso de una falla del gatekeeper primario. Igualmente, un punto terminal puede indicar un backup, dirección de transporte redundante o alternada. Esto permite al punto terminal para tener una interfase de la red secundaria o un punto terminal de H.323 secundario como un backup.

Soporte para un número grande de Dominios: El intento inicial de H.323 fue para redes LANs, de tal forma que no fue diseñado inherentemente para direccionamiento de áreas grandes. El concepto de una zona fue añadido para acomodar el direccionamiento de área grandes. Se definen procedimientos para la localización de usuarios a través de las zonas por los nombres de e-mail. El anexo G define comunicación entre los dominios administrativos y describe métodos para permitir la resolución de dirección, autorización de acceso y uso de reporte entre los dominios administrativos. En búsquedas de multidominio, no hay ninguna manera fácil de realizar la detección del lazo.

Soporte para un número grande de Llamadas: Aunque inicialmente no fue soportado, el control de llamada H.323 puede ser implementado en una manera sin estados. Una gateway puede usar mensajes definidos en H.225 para ayudar el gatekeeper realizando el equilibrio de la carga a través de las gateways.

Capacidad de Multicast: H.323 tienen una especificación separada, H.332, para incluir oyentes del multicast. Bruscamente, los puntos terminales declaran que ellos son multicast capaz al controlador MC. El controlador MC decide entonces si usar multicast e informa a los puntos terminales de las direcciones para usar. Como una invitación, normalmente un miembro de la conferencia le pide al controlador MC que invite a alguien más. Sin embargo, el mensaje de facilidad (facility) permitiría a un punto terminal transferir una llamada en una conferencia.

Interoperabilidad con otros Protocolos de Señalización: Para soportar los servicios de la telefonía tradicional, los protocolos de señalización tienen que soportar el Sistema de Señalización 7 de ISDN. SS7 realizan la señalización fuera de banda en soporte del establecimiento, facturación, direccionamiento y las funciones del intercambio de información de la Red de Telefonía de Conmutación Publica PSTN. Hay dos especificaciones de la señalización disponible en SS7 para las diferentes interfases:

- Q.931 usada para la Interfase de Usuario a Red UNI (User toNetwork Interfase)
- ISUP usado para la Interfase de la Red a Red NNI (Network to Network Interfase)

H.323 acepta la aproximación de circuito conmutador tradicional basado en los protocolos ISDN/Q.931. Q.931, similar a los mensajes de señalización son usados en los procedimientos de H.323 y hacen más fácil interoperar con ISDN/Q.931. Sin embargo, los mensajes de establecimiento de llamada de H.323 son sólo un subconjunto de estos en SS7/ISUP. Porque no hay ninguna norma establecida por relevar de los mensajes SS7/ISUP sobre una red H.323, H.323 pueden trasladar sólo una parte de los mensajes de SS7 en la conversión.

La familia H.32x de recomendaciones ofrece un estándar especifico para interoperar con otras redes de circuito conmutados, por ejemplo, H.320 para ISDN y B-ISDN, H.324 para GSTN. Dentro de estos estándares, la interoperabilidad por gateways esta bien definida.

Planificación de capacidad: Antes de poner el tráfico de video en una red, es necesario asegurar un adecuado ancho de banda para todas las aplicaciones requeridas. Empezar calculando los requerimientos mínimos de ancho de banda para cada aplicación (por ejemplo, voz, video, y datos). La suma representa el requisito de ancho de banda mínimo para cualquier enlace dado, y no debe consumir más del 75 por ciento del ancho de banda total disponible en ese enlace. Esta regla del 75 por ciento asume que algún ancho de banda se requiere para tráfico overhead, tal como actualización del protocolo de enrutamiento y capa 2, así como las aplicaciones adicionales, como e-mail y tráfico de HTTP.

Clasificación: Para proporcionar las garantías de QoS apropiadas para el tráfico, los dispositivos de la red necesitan poder identificar tal tráfico.
El modelo de Servicios Diferenciados (DiffServ) de QoS usa los valores del Punto de Código de Servicios Diferencias DSCP (DiffServ Code Point) para separar tráfico en clases. El modelo de Servicios Diferenciados DiffServ define dos conjuntos de valores de DSCP:

- ✓ **Reenvío Apresurado EF (Expedited Forwarding):** Proporciona un solo valor de DSCP (101110) que da el nivel más alto de servicio a los paquetes marcados de la red.
- ✓ **Reenvío Aseguro AF (Assured Forwarding):** Proporciona cuatro clases, cada una con tres niveles de precedencia de eliminación.

4.2 ESTÁNDAR H.225

El objetivo general de este protocolo es proporcionar un medio para sincronizar paquetes que hacen uso de las facilidades fundamentales de transporte/red basadas en paquetes. No requiere que todos los medios de comunicación y control sean mezclados en un solo stream, el cual será luego empaquetado.

Se puede decir que el estándar H.225 describe los métodos por los cuales audio, video, datos y información de control son asociados, codificados y empaquetados para ser transportados entre los equipos de H.323 y las redes basadas en paquetes como la red IP.

Como parte las transmisiones de video y audio, no podemos hablar de H.225 como un estándar separado puesto que mantiene una estrecha relación con H.323, H.225 usa la forma de paquete especificado por el IETF, las especificaciones del protocolo RTP y el protocolo RTCP con el objeto de cumplir con las siguientes tareas:

- ✓ **Trama Lógica:** Define como las tramas protocolares (empaquetar) de audio y video, se insertan en bits de datos (paquetes) para transportarse sobre un canal de comunicación seleccionado.
- ✓ **Numeración de secuencia:** Determina el orden de los paquetes de datos transportador sobre un canal de comunicación.
- ✓ **Detección de error:** Después de iniciada una llamada, uno o más conexiones RTP o RTCP son establecidas. Streams múltiples permiten a H.225 enviar y recibir diferentes tipos de comunicación simultáneamente, cada uno de ellos con sus propios números de secuencia de trama y opciones de calidad de servicio. Con RTP y RTCP, los nodos de recepción sincronizan los paquetes recibidos en el orden apropiado, por lo que el usuario escuchara u observara la información en forma correcta.

4.2.1 H.225 Señalización de Llamada

La señalización de llamada H225 es un requerimiento básico necesario para establecer y terminar una llamada entre dos puntos terminales. El estándar H.225.0 usa un subconjunto del Protocolo de Señalización Q.931 que fue inicialmente desarrollado para la señalización de las Redes ISDN.

El protocolo Q.931 define como cada capa de H.323 interactúa con capas iguales, por lo que los participantes pueden interoperar con formatos acordados.

Este protocolo reside dentro de H.225 como parte del control de llamada de H.323, Q.931 es enlace protocolar de capas para establecimiento de conexiones y trama de datos, proporcionando un método de definición de canales lógicos dentro de un canal más grande.

Los mensajes Q.931 contienen un discriminador protocolar que identifica cada mensaje como único, con el valor de referencia de la llamada y el tipo de mensaje. La capa de H.225 luego especifica como estos mensajes son recibidos y procesados.

La señalización de llamada H.225 es enviada directamente entre los puntos terminales cuando no existe un gatekeeper (señalización de la llamada Directa), pero cuando una gatekeeper existe en la red los mensajes (figura 4.9) tienen que ser direccionados por medio de el (señalización de llamada indirecta direccionada por el gatekeeper).

4.2.1.1 Señalización de la Llamada Directa

Durante la confirmación de admisión, el gatekeeper indica que los puntos terminales pueden intercambiar los mensajes de señalización de llamada directamente. Los puntos terminales intercambian la señalización de la llamada en el canal de señalización de llamada.

4.2.1.2 Señalización de la Llamada direccionada por el gatekeeper

Los mensajes de admisión se intercambian entre los puntos terminales y el gatekeeper en los canales de RAS. El gatekeeper recibe los mensajes de señalización de llamada en el canal de señalización de llamada desde un punto terminal y los dirige al otro punto terminal en el canal de señalización de llamada de otro punto terminal.
La estructura de H.225 sigue el estándar Q.931 como se muestra en la siguiente ilustración:

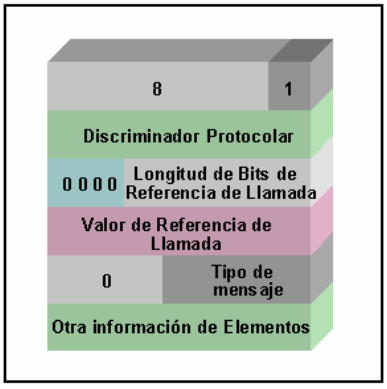


Figura. 4.13. Estructura del encabezado Q.931

Discriminador Protocolar: Divide los mensajes de llamada de red-usuario de otros mensajes.

Valor de Referencia de Llamada: Identifica la llamada o facilita el registro o cancelación de requerimiento en la interfase de red- usuario.
Tipo de mensaje: identifica la función del mensaje enviado.
Otra información de Elementos: Contenidos de los mensajes.

Mensaje de Registro, Autentificación y Estado RAS de H.225

El estándar H.225 también incluye los mensajes de Registro, Autentificación y Estado RAS (esencialmente un protocolo de gastekeeper), el cual define las comunicaciones entre los puntos terminales a un gatekeeper, este tipo de mensajes es solo necesario cuando existe el gatekeeper. A diferencia de la señalización de llamada H.225 y H.245, H.225 RAS usa un transporte no confiable (canal) para la entrega. En una red IP, los mensajes H.225 RAS usan el Protocolo de Datagrama de Usuario UDP.

Desde que los mensajes RAS son transmitidos en un canal no confiable, H.225 recomienda tiempos de espera y reintentos de cuenta para varios mensajes. En un punto terminal o gatekeeper que no pueda responder a un requerimiento dentro del tiempo de espera determinado se puede usar el mensaje de Solicitud en Progreso RIP (Request in Progress) para indicar que esta en proceso aun la solicitud. Un punto terminal o gatekeeper recibiendo el mensaje RIP resetea los tiempos de espera y los reintentos de cuenta.
Las comunicaciones H.225.0 RAS incluyen:

- ✓ **El descubrimiento de Gatekeeper:** Es usado por los puntos terminales para localizar sus Gatekeepers, es decir, en cual deben registrarse. Un punto terminal, el cual necesita encontrar la dirección de transporte de este gatekeeper será multicast un mensaje de Requerimiento de Gatekeeper GRQ. Uno o más gatekeepers podrían reemplazarse con el mensaje de Confirmación de Gatekeeper GCF conteniendo la dirección de transporte del gatekeeper.

- ✓ **Registro del Punto Terminal:** Un solo gatekeeper existe para todos los puntos terminales que deben registrarse con el. Esto es necesario debido a que los gatekeepers necesitan conocer el alias y direcciones de transporte de todos los puntos terminales en su zona para direccionar las llamadas.
- ✓ **Localización del Punto Terminal:** Los gatekeepers usan este mensajer parta localizar los puntos terminales con una dirección especificada. Este proceso es requerido, por ejemplo, cuando el gatekeeper actualiza la base de datos de las direcciones de transporte (alias).

4.3 ESTÁNDAR H.245

El estándar H.245 proporciona los mecanismos de control de la llamada que permiten la compatibilidad en los terminales de H.323 para conectarse unos con otros. H.245 especifica la señalización, control de flujo y canalización de mensajes, peticiones y comandos. La trama de H.245 permite la selección de códigos y capacidades de negociación dentro de H.323. La tasa de bits, la tasa de tramas, el formato de cuadros de selección de algoritmos son algunos de los elementos negociados por H.245.

Los mensajes de control de H.245 son llevados sobre los canales de control de H.245, estos mensajes de control llevan información relacionada a:

- ➢ Intercambio de capacidades.
- ➢ Apertura y cerrando de canales lógicos usados para llevar streams medias.
- ➢ Los mensajes de control de flujo.
- ➢ Los controles generales y mensajes.

4.3.1 Intercambio de las capacidades

El intercambio de capacidades es un proceso que usa los mensajes de los terminales de comunicación para proporcionar sus capacidades de transmisión y recepción. Las capacidades de transmisión describen la habilidad del terminal de transmitir la stream media. Las capacidades de recepción describen la habilidad **de un terminal para recibir y procesar la stream media entrante.**

Cada punto terminal graba sus capacidades de recepción y transmisión, por ejemplo, codificadores, tasa de bits, etc., en un mensaje y lo envía a otro punto terminal.

4.3.2 Señalización del canal lógico

Un canal lógico lleva la información de un punto terminal a otro punto terminal (en el caso de una conferencia de punto a punto) o los puntos terminales múltiples (en el caso de una conferencia punto a multipunto). H.245 proporciona los mensajes para abrir o cerrar un canal lógico (bidireccional).

4.3.3 Mensajes de control de flujo

Estos mensajes proporcionan la regeneración a los puntos terminales cuando se encuentran problemas de comunicación.

4.3.4 Controles generales y mensajes

Varios controles y mensajes pueden ser usados durante una llamada como un comando para establecer el codificador en la recepción del punto terminal

cuando cambia su codificador. Los mensajes de control de H.245 también pueden ser direccionados a través de un gatekeeper.

4.4 PROTOCOLO DE INICIO DE SESIÓN SIP (SESSION INITIATION PROTOCOL)

El protocolo de Inicio de Sesión SIP, es un nuevo protocolo del nivel de control de aplicación que maneja la señalización y el control de llamadas que forman parte de las especificaciones de IETF para establecer llamadas y conferencias multimedia en tiempo real sobre redes IP que esta basado en ASCII. Cada llamada o sesión puede incluir diferentes tipos de datos tales como: vídeo y audio, aunque actualmente la mayoría de las direcciones de extensión SIP dirige comunicación de audio.

SIP ha sido basado en protocolos de Internet tradicionales como el Protocolo de Transferencia de Hipertexto HTTP (Hypertext Transfer Protocol) y el Protocolo de Transferencia de Correo Simple SMTP (Simple Mail Transfer Protocol) por consiguiente el protocolo SIP es un protocolo de capa de aplicación que opera sobre UDP o TCP, es decir es un protocolo neutral en relación con los protocolos de las capas inferiores. Como un protocolo basado en texto, SIP usa un modelo cliente-servidor con solicitudes generadas por el cliente y enviadas al servidor que las responde.

Este protocolo es un estándar abierto y escalable que ha sido diseñado para ser un protocolo de propósito universal, sin embargo, otros protocolo son necesarias para crear un protocolo en términos funcionales de interoperabilidad.
Como otros protocolos de voz sobre IP y multimedia sobre IP, SIP esta diseñado para direccionar las funciones de señalización y manejo de sesión dentro de una red IP. La señalización permite llevar la información de la llamada a través de los

límites de la red. El manejo de sesión proporciona la habilidad de controlar los atributos de una llamada del extremo a extremo.

SIP permite la movilidad personal dependiendo de la capacidad para alcanzar una llamada tripartita en una simple, direcciones independientes de la localización, a continuación enlistamos algunas de las capacidades que proporciona SIP:

- ✓ **Localización del punto terminal designado:** SIP soporta la resolución de dirección, mapeado del nombre y redirección de llamada. Para asegurarse que la llamada se realizarse sin tener en cuenta la localización, cada usuario se identifica a través de una Dirección Global de Documentos de Internet URL (Uniform Resource Locator) jerárquica para construir en base a los elementos de la red el número de teléfono de un usuario o nombre del host, por ejemplo, SIP: user@company.com. Debido a esta similitud, las direcciones URLs SIP son fáciles de asociar con la dirección de e-mail de un usuario.
- ✓ **Determinación de las capacidades de los medios de comunicación del punto terminal designado y negociación de las características:** Mediante el Protocolo de Descripción de Sesión SDP (Session Description Protocol), SIP determina el "nivel más bajo" de servicios comunes entre los puntos terminales. Se establecen conferencias usando sólo las capacidades de los medios de comunicación que pueden ser soportados por todos los puntos terminales. Por ejemplo, una llamada entre un usuario de voz de teléfono móvil y dos usuarios de video estarían de acuerdo en soportar las características de voz solamente. Cuando el usuario del teléfono móvil deja la llamada, los participantes restantes pueden renegociar las características para activar las comunicaciones de video.
- ✓ **Disponibilidad del punto terminal designado:** Si la llamada no puede completarse porque el punto terminal designado no esta disponible, SIP determina si la llamada tripartita esta ya en el teléfono o no contestó en

el número asignado de timbres. Entonces retorna un mensaje que indica el por qué el punto terminal designado no estaba disponible.

- ✓ **Establece una sesión entre el punto de origen y el punto terminal designado:** Si la llamada puede completarse, SIP establece una sesión entre los puntos terminales. SIP también soporta cambio en media llamada como la adición de otro punto terminal a la conferencia o el cambio de una característica de los medios de comunicación o del codec.
- ✓ **Manejo o direccionamiento de la llamada participante:** Durante una llamada, un participante puede traer a otros usuarios en la llamada o conexiones de transferencia, manteniendo la conexión o cancelando la conexión. Durante una transferencia de llamada, SIP simplemente establece una sesión entre el emisor y un nuevo punto terminal (especificado por la llamada tripartita transferida) y termina la sesión entre el emisor y la llamada tripartita transferida. Al final de una llamada, SIP termina las sesiones entre todas las partes de la llamada.

El protocolo SIP provee los mecanismos de protocolo necesarios de modo que sistemas terminales y servidores proxy puedan proporcionar los siguientes servicios:

- ➢ Reenvío de llamadas (call forwarding), incluyendo: llamadas 700, 800, 900, reenvío de llamadas no contestadas, reenvío de llamadas ocupadas, reenvío de llamadas incondicionales y otros servicios de translación de dirección.
- ➢ Entrega del número del emisor y receptor de la llamada, donde los números puede estar en cualquier esquema de nombre (preferiblemente exclusivo).
- ➢ Movilidad personal, por ejemplo, la capacidad para completar una llamada de tripartita a una simple y localización independiente de la dirección cuando el usuario cambia los terminales.
- ➢ Tipo de negociación terminal y selección: el emisor puede dar una opción de cómo completar la llamada tripartita, por ejemplo, mediante telefonía de Internet, teléfono móvil, servicio contestador automático, etc.

- Capacidad de negociación terminal.
- Autentificación del emisor y receptor de la llamada.
- La facturación y supervisión de transferencia de llamadas.
- Invitaciones para conferencias de multicast.

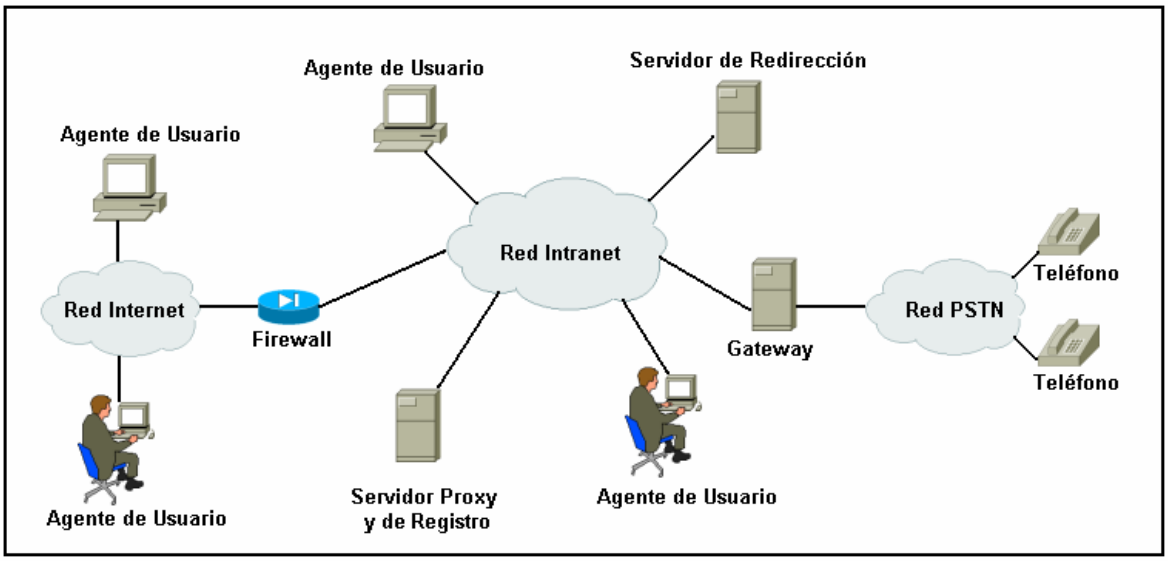


Figura. 4.14. Red SIP

Tanto SIP como H.323 definen mecanismos para direccionamiento de llamadas, intercambio de capacidades, control de medios de comunicación y servicios suplementarios, pero SIP es un protocolo que promete escalabilidad, flexibilidad y facilidad de implementación cuando se construyen sistemas complicados.

4.4.1 Arquitectura del Protocolo SIP

SIP es un protocolo par a par por lo que la arquitectura básica es un par servidor-usuario. Estos pares en una llamada son llamados Agentes del Usuario UAs (User Agents). Un agente del usuario puede funcionar como:

- ✓ **Cliente de Agente de Usuario UAC (User Agent Client):**Una aplicación del cliente que comienza la solicitud del SIP.

- ✓ **Servidor de Agente de Usuario UAS (User Agent Server):** Una aplicación del servidor que comunica al usuario cuando una solicitud del protocolo SIP se recibe y se devuelve una contestación en nombre del usuario.

Las entidades principales en el protocolo SIP son:

- El Agente de Usuario UA (User Agent)
- El servidor proxy
- El servidor de redirección
- El servidor registrador

4.4.1.1 Agente de Usuario UA

Los agentes de usuario UA o puntos terminales de SIP, como se mencionó anteriormente funcionan como clientes UACs al iniciar las solicitudes mediante un mensaje de invitación y como servidores UASs al responder a las solicitudes.

El mensaje de invitación contiene la descripción de la llamada que informa el receptor que tipo de medios de comunicación el emisor puede aceptar y donde desea enviar los datos. El Agente del Usuario está normalmente envuelto en algún tipo de interacción con el usuario, pero también puede usarse para el correo de voz, reenvió de llamadas y servicios similares sin involucrar al usuario. Los agentes del usuario se comunican directamente con otros agentes del usuario o mediante un servidor intermedio.

Los clientes del SIP incluyen:

- ✓ **Teléfonos:** Pueden actuar como un UAS o UAC. Softphones (PCs que tiene capacidades telefónicas instaladas) y los teléfonos de IP SIP pueden iniciar una solicitud SIP y pueden responder a las demandas.
- ✓ **Gateways:** Proporcionan el control de la llamada y muchos servicios más, el más común una función de traslación entre los punto terminales

de conferencia de SIP y otros tipos de terminales. Esta función incluye traslación entre los formatos de transmisión y entre los procedimientos de comunicaciones. Además, la gateway traduce los codecs del video y audio y realiza el establecimiento de la llamada y desbloquea ambos lados de la red.

4.4.1.2 Servidor proxy

Los servidores proxy se ocupan de reenviar las solicitudes y respuestas SIP para el establecimiento y liberación de llamadas de VoIP, adicionando los medios necesarios para garantizar que ese diálogo de señalización entre solicitante y solicitado se desarrolle por la misma vía, es decir, que los mensajes de señalización SIP de ida y vuelta sigan la misma ruta.
Es un programa “intermediario” que actúa como servidor y como cliente, pues respecto al emisor de llamada se comporta como servidor y respecto al receptor de llamada como cliente.
Un servidor proxy puede ser:

- Con estado (stateful)
- Sin estado (stateless).

Un proxy con estado “stateful” mantiene o retienen la información de la llamada durante todo el tiempo en que esta se esta realizando mientras un servidor proxy sin estado “stateless” procesa un mensaje SIP y luego se olvida de todo lo referente a la llamada en cuestión hasta que el próximo mensaje SIP asociado a la misma llamada llegue.

La implementación sin estado provee buena escalabilidad, pues los servidores no requieren mantener información a cerca del estado de la llamada una vez que la transacción ha sido procesada. Además, esta solución es muy robusta dado que el servidor no necesita “recordar” nada en relación con una llamada.

Sin embargo, no todas las funcionalidades pueden ser implementadas en un servidor proxy sin estado, por ejemplo, las funcionalidades relativas a la contabilización y facturación de las llamadas pueden requerir funcionalidades de proxy con estado, de manera que se le pueda seguir el rastro a todos los mensajes y estados de una comunicación.

4.4.1.3 Servidor de redirección

Los servidores de redirección como su nombre indica redireccionan las solicitudes de llamadas al procesar los mensajes y retornan la dirección (o direcciones) del receptor. Numerosas conexiones pueden tener lugar hasta alcanzar el destino final.
En caso contrario rechazan la llamada, enviando una respuesta de error. Desarrollan una funcionalidad similar al Gatekeeper H.323 cuando en la solución ITU se emplea el modelo de llamada directo.

4.4.1.4 Servidor registrador

Un servidor registrador es un servidor que acepta la solicitud de registro que contiene la dirección SIP y la dirección IP asociada, es decir, garantizan el mapeado entre direcciones SIP y direcciones IP, posibilitando el registro correspondiente a la localización actual de los usuarios, esto es, “seguir el rastro” de los usuarios, pues por diferentes razones las direcciones IP de éstos puede cambiar. También se les denomina servidores de localización, pues son utilizados por los servidores proxy y de redirección para obtener información respecto a la localización o localizaciones posibles del receptor de llamada.

El agente de usuario envía un mensaje de registro al servidor registrador y este almacena la información de registro en un servicio de localización mediante un protocolo no SIP. Una vez que la información es almacenada, el servidor le envía la contestación apropiada de regreso al agente del usuario. Un servidor registrador es típicamente co-localizado con un servidor proxy o servidor de redirección.

La información registrada en estos servidores, no es permanente, requiere ser refrescada, el registro correspondiente será borrado. En consecuencia, para mantener la información de registro, el usuario necesita ser refrescado periódicamente.

La función principal de los servidores de SIP es proporcionar la resolución del nombre y localización del usuario, cuando es improbables revocar la dirección IP o el nombre del host que realizo las llamadas. Usando un recordatorio, el agente de usuario de la llamada puede identificar un servidor específico para resolverse la información de dirección de llamada realizada.

4.4.2 Protocolo de Transporte de SIP

El protocolo SIP hace referencia mínima sobre el protocolo de transporte, puede usar directamente cualquier datagrama o protocolo de stream, con la única restricción que una solicitud SIP entera o la contestación tiene que ser entregada por completo o nada. El protocolo SIP puede así ser usado con el protocolo UDP o el protocolo TCP en redes IP.

4.4.3 Codificación de Mensajes

A diferencia de otros protocolos como Q.931 y H.323, SIP es un protocolo basado en texto, este diseño fue escogido minimizar el costo de acceso. Las

estructuras de datos necesarias en los encabezamientos SIP caen dentro de la categoría del parámetro-valor, posiblemente con un solo nivel de sub-parámetros, por lo que los mecanismos de codificación de datos genéricos como la Notación de Sintaxis Abstracta 1 ASN.1 (Abstract Syntax Notation 1) no ofrecen ninguna ventaja funcional.

Los protocolos basados en texto pueden ser difíciles de dividirse debido a su estructura irregular. SIP intenta evitar esto manteniendo una estructura común de todos los campos del encabezamiento permitiendo una división genérica. SIP fue diseñado para un conjunto de caracteres independientes, así que cualquier campo puede contener cualquier carácter de la Organización Internacional de Estandarización ISO (International Organization for Standardization) 10646. Puesto que SIP opera en un canal limpio de 8 bits, el dato binario no tiene que ser codificado.

4.4.4 Direccionamiento e Identificación

Para ser invitado e identificado la llamada tripartita tiene que tener un nombre. Puesto que es la forma más común de direccionamiento es la de usuario de Internet, el protocolo SIP escogió un identificador similar al de e-mail, de la siguiente forma: "user@domain", "user@host", "user@IP_Address" o "phone_number@gateway".

Donde el “user” puede ser un nombre de usuario o una dirección E.164 y el identificador o dominio, pueden referirse al nombre del host del usuario en ese momento, a las direcciones de e-mail o el nombre de un servicio de traslación de nombre de un dominio específico.

Las direcciones de la forma "phone_number@gateway" designa números de teléfono de la Red de Telefonía de Conmutación General GSTN (General Switching Telephony Network) accesibles mediante la gateway.

El SIP usa estas direcciones como parte de la dirección URL SIP, por ejemplo: "sip:usuario1@ejemplo.com". Esta dirección URL puede ponerse bien en una página de Web, para que pulsando el botón se inicie una llamada a esa dirección, similar al enlace de e-mail URL.

4.4.5 Bifurcación

El protocolo SIP difiere de los otros protocolos de señalización en que permite a una solicitud de llamada ser bifurcada, es decir, un servidor puede enviar dos o más solicitudes a diferentes destinos basado en la solicitud entrante, en seguida o en secuencia, si una solicitud anterior falla. Esta característica soporta un número de servicios telefónicos avanzados, tal como: el reenvío de llamada para correo de voz, la Distribución de la Llamada Automática ACD (Automatic Call Distribution) y localización del usuario, donde el mismo número puede timbrar en casa y en el trabajo, por ejemplo.

4.4.6 Protocolo de Descripción de Sesión SDP (Session Description Protocol)

El Protocolo de Descripción de Sesión SDP es un protocolo para describir sesiones de audio, vídeo y multimedia. El protocolo SIP, el Protocolo de Control de Gateway Media MGCP (Media Gateway Control Protocol), el Protocolo de Anuncio de Sesión SAP (Session Announcement) y el protocolo Streaming en Tiempo Real RTSP (Real Time Streaming Protocolo) todos usan el protocolo de descripción de sesión SDP.

El protocolo SDP incluye la siguiente información:

- Nombre de la sesión y propósito.
- Tiempo que la sesión esta activa.
- Los medios de comunicación que comprenden la sesión como: el tipo de datos, el protocolo de transporte, el formato. Para las sesiones multicast IP, las direcciones multicast de destino de stream y puerto de transporte de destino de stream. Para las sesiones unicast IP, la dirección remota, el puerto de transporte para la dirección del contacto y semántica de la dirección e información del puerto.
- Información de cómo recibir esos medios de comunicación por ejemplo, direcciones, puertos, formatos y así sucesivamente.

La información adicional sería:

- La información sobre el ancho de banda ha ser usado por la conferencia.
- Información del contacto para la persona responsable de la conferencia.

4.4.7 Funcionamiento básico del protocolo SIP

SIP es un protocolo simple, basado en ASCII que usa las solicitudes y contestaciones para establecer comunicación entre varios componentes en la red y establecer una conferencia finalmente entre dos o más puntos terminales.
Los usuarios en una red SIP son identificados por direcciones SIP únicas, por ejemplo: userID@gateway.com. y deben registrarse con un servidor registrador usando las direcciones SIP asignadas, el servidor registrador proporciona esta información al servidor de redirección en la solicitud.

Cuando un usuario inicia una llamada, una solicitud de SIP es enviada al servidor de SIP (proxy o redirección). La solicitud incluye la dirección del emisor y del receptor de llamada La localización del usuario final puede registrarse dinámicamente con el servidor de SIP. El servidor de redirección puede usar uno o más protocolos para localizar al usuario final, debido a que el usuario final

puede acceder al sistema en más de una estación y además porque el servidor de redirección a veces puede tener información inexacta.

Si la solicitud está pasando por un servidor proxy de SIP, el servidor proxy probará cada una de las direcciones hasta localizar al usuario final. Pero si la solicitud está pasando por un servidor de redirección, este le envía todas las direcciones al emisor de llamada en el campo de encabezamiento de la contestación.

Los mensajes de solicitud SIP para establecer una llamada son:

- ✓ **Mensaje de Invitación "Invite":** Invita a un usuario o servicio a participar en una sesión. El cuerpo del mensaje contiene una descripción de la sesión.
- ✓ **Mensaje de confirmación "Ack":** Confirma que el cliente emisor ha recibido una respuesta final desde un servidor a una solicitud, por ejemplo, a la solicitud de invitación, reconociendo la respuesta como adecuada. **Mensaje de Opciones "Options":** Posibilita descubrir las capacidades del receptor de la llamada.
- ✓ **Mensaje de despedida "Bye":** Finaliza una llamada o una solicitud de llamada. Puede ser enviado por el agente emisor o por el agente receptor.
- ✓ **Mensaje de cancelación "Cancel":** Cancela una solicitud pendiente, pero no afecta una solicitud completada. Este método finaliza una solicitud de llamada incompleta.
- ✓ **Mensaje de registro "Register":** Se utiliza este método como un servicio de localización que registra la localización actual de un usuario. También es necesario cuando hay varios servidores SIP en un mismo host, en cuyo caso solo uno de los servidores puede usar el número de puerto predeterminado.

Las posibles respuestas a estas solicitudes son códigos de tres enteros e indican el resultado de comprender y satisfacer una solicitud. Las respuestas SIP son:

- ✓ **1xx "Informativo":** Solicitud recibida, que continua para procesar la solicitud.

- ✓ **2xx "Sucesos":** La acción fue recibida de forma adecuada, comprendida y aceptada.
- ✓ **3xx "Redireccionado":** Más acciones deben ser consideradas para completar la solicitud.
- ✓ **4xx "Error de cliente":** La solicitud contiene mal la sintaxis o no puede ser resuelta en este servidor.
- ✓ **5xx "Error de servidor":** El servidor ha errado en la resolución de una solicitud aparentemente válida.
- ✓ **6xx "Fallo global":** La solicitud no puede ser resuelta en servidor alguno.

Las respuestas 2xx, 3xx, 4xx, 5xx y 6xx son "respuestas finales" y terminan la transacción SIP. En cambio, las respuestas 1xx son "respuestas provisionales", y no terminan la transacción SIP.

Por ejemplo, la figura muestra cómo el servidor proxy trabaja para conectar dos agentes de usuario UAs.

El usuario UNO en el host "emisor.com" que actúa como cliente UAC quiere invitar al usuario DOS. El cliente UAC obtiene la dirección IP para DOS buscada en el Servicio de Nombres de Dominio DNS (Domain Name Service). Esto produce en el servidor "ejemploproxy", una solicitud de invitación (mensaje INVITE) generada y enviada a servidor. Como se menciono antes, el mensaje de invitación contiene una descripción de la sesión mediante el protocolo SDP que informa el receptor de la llamada qué tipo de medios de comunicación el emisor de llamada puede aceptar y donde desea enviar los datos de los medios de comunicación. El servidor acepta la invitación y busca la localización actual del usuario DOS en el servidor de localización.

El servidor de localización devuelve host "receptor.com"en el que el usuario DOS se encuentra. El servidor proxy genera y le envía una solicitud de invitación al host "receptor.com".

El servidor UAS en el host "receptor.com" pregunta al usuario DOS si él quiere la invitación o no. La aceptación es devuelta al servidor proxy. El servidor proxy envía la aceptación al emisor original UNO. La aceptación es confirmada por un

mensaje de reconocimiento ACK tanto para el servidor proxy como para el usuario DOS. Con la confirmación de aceptación de los dos participantes, una sesión es establecida entre el emisor y el receptor de la llamada. El protocolo RTP generalmente es usado para la comunicación entre el emisor y receptor.

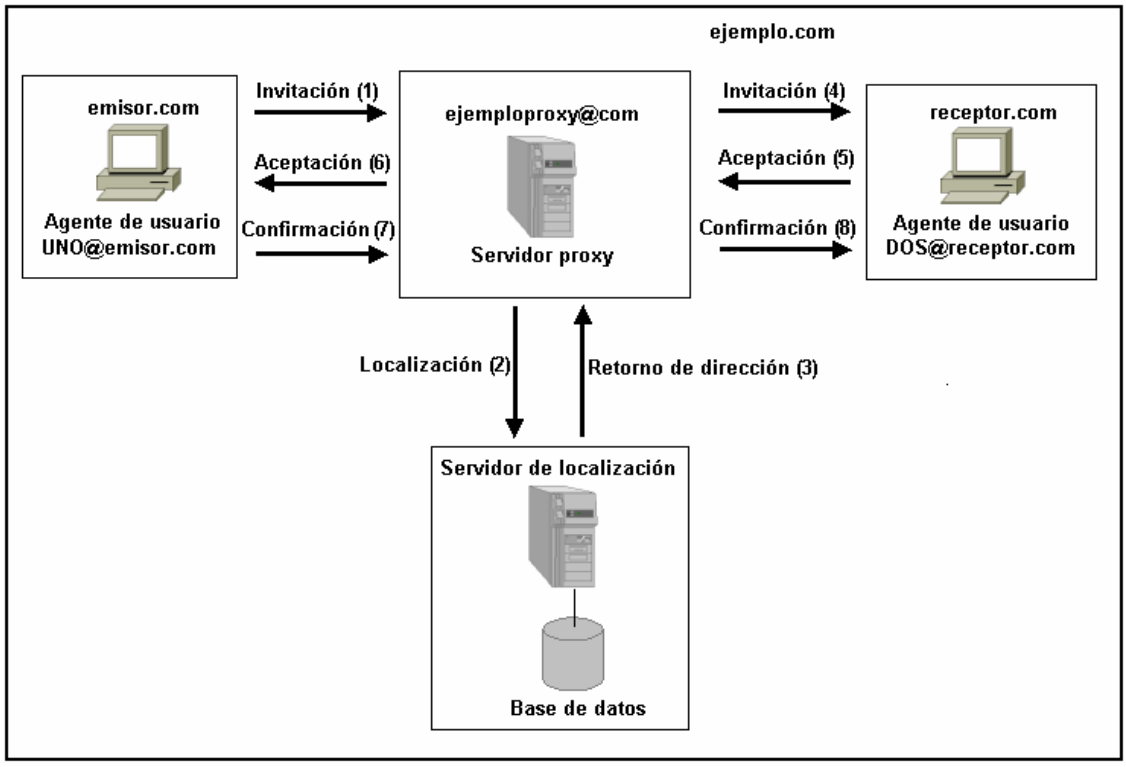


Figura. 4.15. Ejemplo de uso del Servidor proxy

Si la llamada se quiere establecer por medio del servidor de redirección, se sigue el mismo procedimiento del servidor proxy:

1. El usuario UNO invita al usuario DOS por medio del servidor de redirección. “ejemploredirección” mediante una solicitud de invitación.
2. El servidor de redirección acepta la invitación y busca la localización actual del usuario DOS en el servidor de localización, al igual que en el caso anterior.
3. El servidor de localización devuelve host “receptor.com”en el que el usuario DOS se encuentra.

4. El servidor de redirección le envía un mensaje al emisor de la llamada con la dirección del usuario DOS.
5. El emisor de la llamada envía un mensaje de reconocimiento ACK al servidor de redirección para reconocer el mensaje anterior.
6. Entonces el emisor genera y envía una nueva solicitud de invitación a la nueva dirección.

De similar manera que en el caso anterior, el servidor UAS en el host "receptor.com" pregunta al usuario DOS si él quiere la invitación o no, y devuelve la aceptación al servidor UAS del usuario UNO en caso afirmativo. La aceptación es confirmada por un mensaje de reconocimiento ACK.

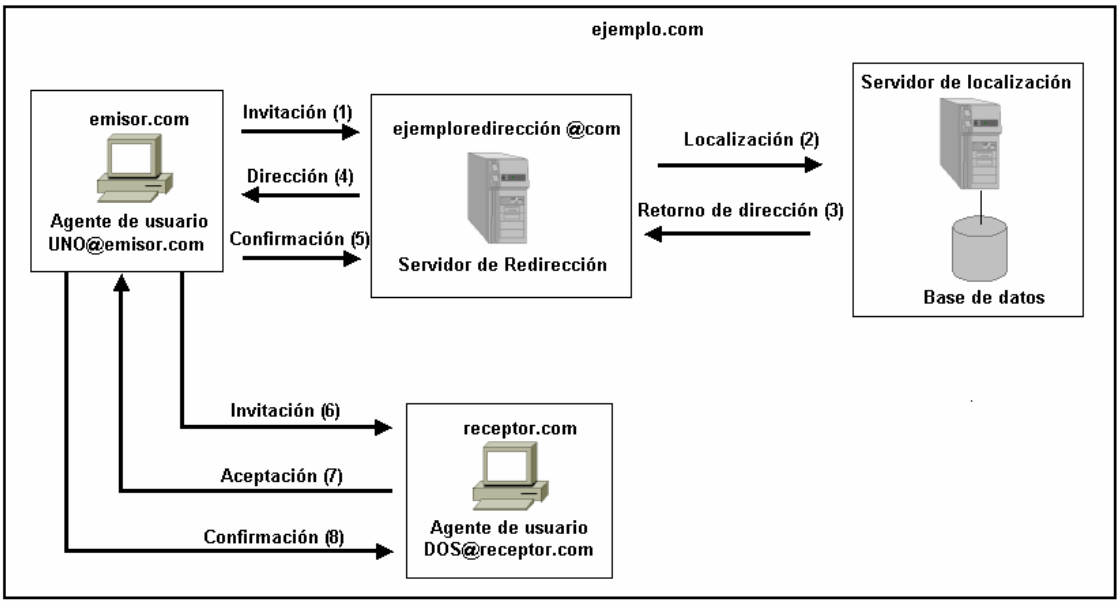


Figura. 4.16. Ejemplo de uso del Servidor de redirección

Durante esta conferencia, cualquier usuario puede invitar a otro usuario o puede cambiar la sesión por diferente invitando UA para cambiar su localización.

4.4.8 Calidad de Servicio QoS

Como ya se ha mencionado anteriormente la calidad de servicio QoS es determinada por el retardo, la fluctuación del retardo o Jitter y la pérdida de paquetes.

A continuación describiremos algunas de las características que han mejorado la calidad de servicios en el protocolo de inicio de sesión SIP:

Retardo en el establecimiento de Llamada: El procedimiento de establecimiento de llamada SIP es similar al de H.323, sin embargo, si el establecimiento de llamada del protocolo UDP falla a diferencia de H.323 que establecía una conexión de UDP y TCP casi simultáneamente, SIP opera UDP y TCP secuencialmente, esto incrementa el retardo en el establecimiento de llamada cuando una conexión UDP no es viable.

Pérdida de paquetes: La fiabilidad de los mensajes de SIP es lograda considerando que el cliente retransmite las solicitudes cada 0.5 segundos hasta que un reporte en progreso o un estado final sea recibido. El servidor simplemente retransmite la respuesta final original hasta que un mensaje de reconocimiento ACK sea recibido. El cliente retransmite el mensaje ACK para cada mensaje final, este es el método utilizado para evitar la pérdida de paquetes.

Detección del lazo de envió: Un tipo común de error son los lazos de reenvió de llamada, los cuales pueden ocurrir sobre todo cuando múltiples servidores SIP están siendo usados en la llamada. SIP proporciona algoritmos de detección similar al usado en el Protocolo de Gateway de Borde BGP (Border Gateway Protocol) para prevenir la búsqueda del lazo. Este algoritmo trabaja a través del campo de encabezamiento, antes que un servidor proxy direccione cualquier solicitud, este chequea el campo. Si encuentra el nombre propio el lazo debe ocurrir, pero si el nombre no esta en la lista del campo entonces transmite la solicitud a otro servidor proxy o punto terminal.

Soporte para un número grande de Llamadas: SIP soporta n a n escalamientos entre UAs y servidores. SIP toma menos ciclos para generar mensajes de señalización, en consecuencia un servidor podría teóricamente manejar más transacciones.

Seguridad: SIP soporta la autentificación del emisor y receptor de llamada mediante mecanismos http. Criptográficamente asegura la autentificación y encriptación soportado de conexión a conexión mediante SSL/TSL (¿?), pero SIP podría usar cualquier capa de transporte o mecanismo de seguridad similar a HTTP.

4.5 PROTOCOLO DE CONTROL DE GATEWAY MEDIA MGCP (MEDIA GATEWAY CONTROL PROTOCOL)

Entre los protocolos actuales y nuevos de sistemas de comunicaciones multimedia, el protocolo de control de gateway media MGCP constituye una opción compatible y complementaria para el control de gateways, ofreciendo ventajas esenciales en aspectos de interés como son: reducción del costo de las operaciones, optimización de las redes y la gestión de los cambios.

El Protocolo de Control de Gateway Media MGCP, fue diseñado para controlar las gateway de redes de telefonía IP que están construidas descomponiendo las gateway de voz sobre IP, las soluciones de voz sobre IP basadas en el protocolo MGCP separan la inteligencia del control de llamada (señalización) y el manejo del medio. Más específicamente, el protocolo MGCP es usado por los elementos de control de llamada externos llamados Controladores de Gateway Media MGCs (Media Gateway Controllers) o agentes de llamada para controlar Gateway Media MGs.

En particular, MGCP debe su origen a la confluencia del Protocolo de Control de Gateway Simple SGCP (Simple Gateway Control Protocol) y el Protocolo de Internet Control de Dispositivo IPDC (Internet Protocol Device Control).

MGCP es un protocolo complementario para el protocolo SIP y H.323. En el modelo del protocolo MGCP, el controlador de gateway media MGC maneja el procesamiento de la llamada con la red IP mediante comunicaciones con un dispositivo de señalización IP como un gatekeeper de H.323 o un servidor de SIP y con la red PSTN vía una gateway de señalización opcional.

Usando H.323, el controlador MGC implementan las capas de "señalización" de H.323 y se presenta a sí mismo como un "gatekeeper H.323" o como uno o más "puntos terminales de H.323". Dentro del protocolo MGCP, las gateways medias MGs se enfoca en la función traslación de la señal de audio, la conversión del trabajo entre el transporte de la señal audio en circuitos telefónicos y paquetes de datos transportados en rede IP.

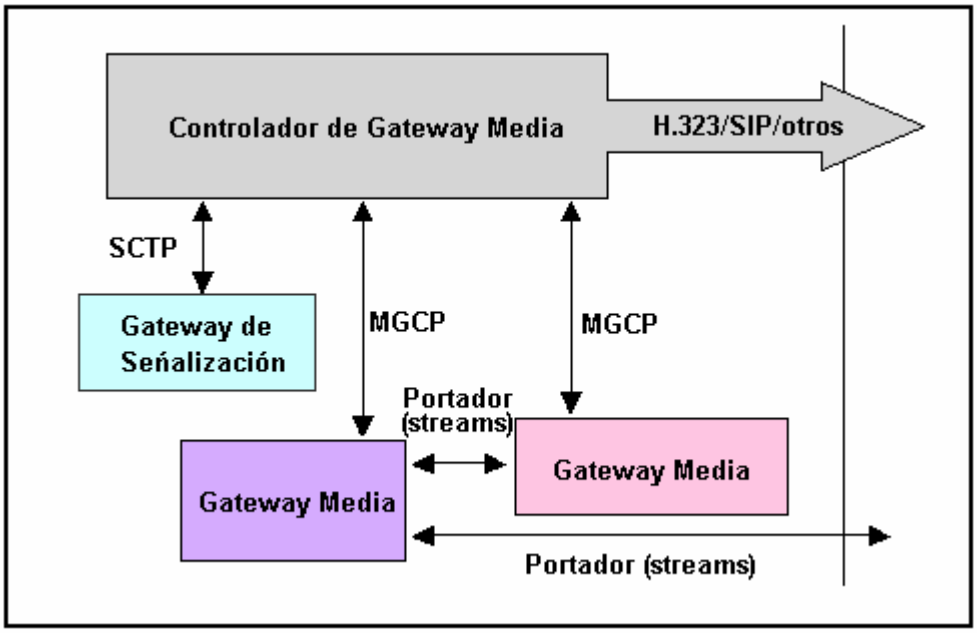


Figura. 4.17. Interworking del protocolo MGCP

4.5 ARQUITECTURA DE MGCP Y ENTIDADES

MGCP es un protocolo maestro-esclavo con un acoplamiento entre la MG que actúa como un punto terminal y el MGC como servidor. Similarmente a los protocolos SIP y H.323 confirma una variedad de protocolos existentes tal como el protocolo SDP para describir los aspectos medios de la llamada y los protocolos RTP o RTCP (usados por MGs) para manejar el transporte en tiempo real.

En la arquitectura del protocolo MGCP, el servidor de MGC o "agente de llamadas" es obligatorio y maneja las llamadas y conferencias y soporte de servicios. El punto terminal de MG es ignorante de las llamadas y conferencias pero no mantiene el estado de las llamadas. Las MGs ejecutan comandos enviados por los agentes de llamadas de MGC. MGCP asume que el agente de llamadas sincronizará cada uno enviando comandos lógicos a MGs bajo su control pero no define un mecanismo para sincronizar los agentes de llamadas.
Las entidades principales de MGCP son:

- Puntos Terminales
- Conexiones

MGCP supone un modelo de conexión donde la construcción básica de puntos terminales y conexiones son usados para establecer la ruta de voz entre los participantes de la llamada.

Los Puntos Terminales: Son fuentes de datos y pueden ser virtuales y físicos. La creación del punto terminal físico requiere la instalación del hardware mientras el programa es suficiente para crear un punto terminal virtual. Una interfaz en una gateway que finaliza en una troncal conectada a un conmutador de red PSTN es un ejemplo de un punto terminal físico. Una fuente de audio en un servidor de contención de audio es un ejemplo de un punto terminal virtual.

Las conexiones: Pueden ser tanto punto a punto o multipunto. Un conexión punto a punto asocia dos puntos terminales. Una de vez que estas asociaciones se establecen para los dos puntos terminales, la transferencia de datos entre estos puntos terminales pueden iniciarse. Una conexión de multipunto se establece por conectar el punto terminal a una sesión de multipunto.
Las conexiones pueden ser establecidas sobre varios tipos de redes de portador: la transmisión de paquetes de audio usando el protocolo RTP y UDP sobre una red TCP/IP. Para conexión punto a punto y conexión de multipunto los puntos terminales pueden estar en gateways separadas o en la misma gateway.

Las señales de control para operaciones del protocolo MGCP enviadas desde el MGC al MG y los eventos enviados desde la MG al MGC. Los conceptos de señal y eventos son usados para establecer y finalizar llamadas. Un MGC empieza las transacciones para manejar o configurar los puntos terminales de MG usando los comandos del protocolo MGCP. MGs envía respuestas a la solicitud de transacción del MGC usando cualquier notificación o reiniciando el comando.
Las señales y los eventos necesarios para soportar la función de telefonía específica o tipo de punto terminal son agrupados en paquetes de señales/eventos. Los paquetes definidos en MGCP incluyen:

- Paquete de Media Genérico
- Paquete de Multifrecuencia de Doble Tono “DTMF”
- Paquete de Troncal
- Paquete de Línea
- Paquete de Microteléfono
- Paquete de RTP
- Paquete de Servidor de Anuncio

Los comandos MGCP tienen como tarea controlar las operaciones de las MGs en relación con la creación y liberación de “conexiones”, e informar al controlador

MGC de los eventos que acontezcan en los puntos terminales asociados a las MGs. En total, el protocolo MGCP dispone de ocho comandos:

- Creación de conexión “CreateConnection”
- Modificación de conexión “ModifyConnection”
- Suprimir conexión “DeleteConnection”
- Notificación de Respuesta “NotificationRequest”
- Notificación “Notify”
- Auditoria de punto terminal “AuditEndpoint”
- Auditoria de conexión “AuditConnection”
- Reinicio en progreso “RestartInProgress”

Todos los comandos están compuestos de un encabezamiento del comando y pueden contener otros parámetros, por ejemplo, un descriptor de sesión, que no es más que un conjunto de parámetros que activa un punto terminal para generar y reconocer el formato apropiado de los medios.

Otros parámetros para los comandos MGCP son: Identificador de llamada “Call ID”, identificador de conexión “Connection ID”, etc.

El protocolo MGCP desarrolla sus funciones mediante una secuencia de comandos y respuestas obligatorias. El controlador MGC se responsabiliza del envío de los comandos a los puntos terminales y recibe un mensaje de reconocimiento por cada comando.

Las respuestas consisten de un encabezamiento y pueden tener parámetros adicionales. Algunas respuestas pueden contener también un descriptor de sesión, como en el proceso de negociación de parámetros para establecer una llamada. Las respuestas son codificadas en base a números enteros, cada uno representan a un mensaje por ejemplo:

- 200: ejecución normal
- 401: teléfono descolgado
- 500: terminación desconocida

Los encabezamientos y descriptores de la sesión son codificados en un conjunto de líneas de texto, separados por un retorno de transporte y una línea alimentación del carácter o opcionalmente una sola línea de alimentación del carácter. Los encabezamientos están separados del descriptor de la sesión por una línea vacía. MGCP usa un identificador de la transacción para correlacionar las órdenes y las contestaciones, los cuales tienen valores entre 1 y 999999999.
Una entidad de MGCP no puede reusar un identificador de la transacción hasta 3 minutos después de la realización de la orden anterior en el que el mismo identificador fue usado.
El encabezamiento de la orden está compuesto de:

- Una línea del comando, identificando la acción pedida, el identificador de la transacción, el punto terminal hacia en el que la acción es requerida y la versión protocolar de MGCP.
- Un conjunto de líneas del parámetro, compuesto del nombre del parámetro seguido por un valor del parámetro.

La línea de comando está compuesta de:

- El nombre de la solicitud.
- El identificador de transacción.
- El nombre del punto terminal que debe ejecutar la orden.
- La versión protocolar.

Estos cuatro parámetros se codifican como cadenas de caracteres de ASCII imprimibles, separados por espacios en blancos.

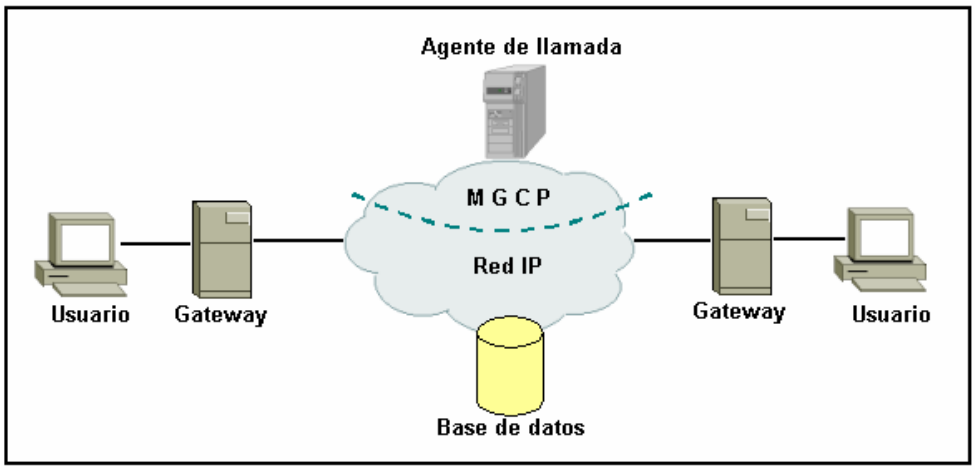


Figura. 4.18. La red IP y el protocolo MGCP

4.6 PROTOCOLO DE CONTROL DE GATEWAY MEDIA MEGACO (MEDIA GATEWAY CONTROL)

Existe una gran confusión entre el protocolo MEGACO o H.248 y el protocolo MGCP debido al nombre de los mismos, pero existen diferencias muy grandes entre los dos.

El protocolo MEGACO representa un método de control de gateway simple que abarca todas las aplicaciones de gateway, además de cualidades de simplicidad, eficiencia, flexibilidad y eficacia en función del costo, que hacen de MEGACO un estándar convincente para su uso en las redes IP de la próxima generación.

MEGACO es el protocolo estándar oficial de la industria para interfase entre agentes de llamada externos llamados Controladores de Gateway Media MGC (Media Gateway Controllers) y Gateway Media MG. MEGACO ofrece mejoras claves en comparación con MGCP tales como:

- Soporta multimedia y conferencia multipuntos mejorando los servicios.
- Perfecciona la sintaxis para un procesamiento de mensaje de semántica mas eficiente.
- Opciones de transporte por medio del protocolo TCP y el protocolo UDP.
- Permite cualquiera codificación de texto o binaria.
- Formalizó un proceso de extensión para mejorar la funcionalidad.

- Amplia la definición de “paquetes”.

4.6.1 Arquitectura de MEGACO y Entidades

El protocolo MEGACO tiene la misma arquitectura del protocolo MGCP, sus comandos o órdenes son similares a los comandos del MGCP, sin embargo, la diferencia principal entre estas dos implementaciones es que con MEGACO los comandos aplicados a terminaciones relativas a un contexto, en lugar de conexiones individuales, como sucede con MGCP. Las conexiones se logran obtener colocando dos o más terminaciones dentro de un contexto común. Este es el concepto de un contexto que facilita el soporte de multimedia y llamadas de conferencia. El Contexto puede ser visto como un puente mixto que soporta múltiples streams medias para mejorar los servicios de multimedia. Los paquetes MEGACO incluyen más detalles que los paquetes de MGCP, definen propiedades adicionales y estadísticas junto con información de eventos e información de señal que puede ocurrir en las terminaciones.

En el protocolo MEGACO los elementos funcionales para el servicio de multimedia sobre IP son:

- Gateway Media GM (Media Gateway).
- Controlador de Gateway Media MGC (Media Gateway Controller).
- Puntos terminales.

Gateway Media GM: Recibe los streams de medios desde la red de circuitos conmutados SCN u otro origen no IP, paqueta los datos y los entrega a la red IP. Realiza la operación inversa cuando los streams de medios fluyen desde la red IP. Las gateways MG realizan un control muy limitado de la llamada, mediante ellas se transfieren las señales de medios, esto es, la información de usuario, de manera que le son comunes funciones tales como el cambio del formato de los

datos y otras. Según su función específica, o su ubicación funcional, las MGs se pueden clasificar así:

- MGs residenciales (entre teléfonos y red IP)
- MGs troncales (entre redes PSTN y red IP)
- MGs de acceso (entre PBX´s y red IP), etc.

Controlador de Gateway Media MGC: Controla el registro y control de recursos en las gateway medias, pudiendo incluso disponer de la capacidad para autorizar el uso de estos recursos bajo cierta política. Para propósitos de transporte de señalización, el controlador MGC puede actuar como punto de origen y terminación para protocolos de la red SCN, como por ejemplo la Parte de Usuario de ISND ISUP (ISDN User Part) de SS7.

En un modelo de control de llamada centralizado, casi toda la inteligencia del sistema está en los controladores MGC y una pequeña parte de ésta en las gateway MG, de ahí que sea adecuado para terminales donde se disponen de poca inteligencia como teléfonos convencionales, etc.

El principio de funcionamiento del modelo MEGACO está basado en el paradigma maestro-esclavo, donde el controlador MGC y las MG dialogan sobre la base del protocolo MGCP. De manera que casi todo el control de la llamada está soportado en el MGC, siendo éste un elemento central de control, y las MG son los elementos funcionales que median entre las redes IP y los terminales y/o otras redes.

Puntos terminales: Los puntos terminales son las fuentes y los terminales de la información de usuario, pudiéndose clasificar así:

- Puntos terminales físicos: enlace troncal, nodo físico, teléfono, etc.
- Puntos terminales virtuales: módulo software sobre un punto terminal físico.

CAPITULO V

PROTOCOLOS DE SERVICIO QoS SOBRE IP

El Internet está cambiando cada aspecto de nuestra vida, detrás de todo este éxito esta: el Protocolo de Internet, IP, el cual fue diseñado para proporcionar servicios de mejor esfuerzo con cualquier medio de comunicación y plataforma del sistema. Esta popularidad actualmente ha cambiado el paradigma de "IP sobre todo," a "todo sobre IP" para manejar la multitud de aplicaciones existentes actualmente, por lo que la red de IP debe ser diseñada para proporcionar los requisitos de calidad de servicio QoS.

La Calidad de Servicio tiene tres ámbitos de aplicación que son:

- Operadores de Redes de Telecomunicación
- Centros proveedores de servicio
- Usuarios (residenciales, negocios, académicos)

En los operadores de Redes de Telecomunicación la calidad de servicio es:

- Evitar situaciones de congestión en los nodos de la Red
- Proporcionar mecanismos para clasificar el tráfico
- Entregar los paquetes al destino conservando la tasa de envío y el perfil de la información
- Asignar prioridades en función de los contratos de tráfico con los clientes
- Realizar un encapsulado óptimo de las aplicaciones en las unidades del transporte

- Atender las demandas y reclamaciones de los usuarios con la mayor eficiencia posible

En los centro proveedores de servicios en cambio la calidad de servicio se ve representada en:

- Dimensionar los recursos de Red de forma óptima en función del número de suscriptores.
- Proporcionar un conjunto de aplicaciones que satisfagan el perfil de los usuarios abonados al mismo.
- Ofrecer a un determinado conjunto de usuarios un grado de servicio acordado aún en casos de congestión.
- Actuar de forma rápida y eficiente frente a las incidencias que se produzcan con los suscriptores.
- Disponer de un contrato al que se puedan ajustar los usuarios en el que se reflejen las aplicaciones que estos pueden utilizar y el nivel de servicio que van a obtener.

Y para los usuarios del Internet, la calidad de servicio se puede ver en:

- Obtener un tiempo de conexión y un nivel de servicio de acuerdo con la tarifa que se abona al proveedor.
- Mejorar su productividad (negocios), calidad de enseñanza (académicos) o simplemente proporcionar contenidos de interés (residencial) mediante el acceso a la Red.
- Disponer de un servicio de atención al cliente que responda rápida y eficientemente en ocasiones de necesidad.
- Obtener un acceso más fiable, ágil y sencillo que le permita sacar el máximo partido de las ventajas que ofrece la Sociedad de las tecnologías de la Información

El diseño original de Internet contemplaba un servicio de mejor esfuerzo, es decir, sin ninguna garantía de QoS. A partir de 1994 aproximadamente, de aplicaciones de videoconferencia y multimedia en tiempo real en Internet, muy

sensibles a situaciones de congestión y retardo, iniciaron el interés de adaptar los protocolos de Internet para ofrecer algún tipo de QoS.

En realidad ya había algo de QoS desde el principio en IPv4, pues en la cabecera del datagrama se encontraba el campo denominado Tipo de Servicio ToS (Type Of Service) de ocho bits de los cuales los tres primeros representaban una prioridad (denominada precedencia) que permitía marcar los datagramas según su importancia, marcado que permitía establecer en principio políticas o prioridades en la transmisión de los datagramas por la red.

Aunque la prioridad representa una cierta calidad de servicio, por cuanto permite clasificar los datagramas en categorías, no es capaz en general de ofrecer una garantía estricta, al estilo de la ofrecida por ATM, donde es posible reservar un ancho de banda determinado para un circuito, aplicación o flujo determinado, asignándole un circuito virtual CBR, por ejemplo.

Para facilitar la calidad de servicio QoS de extremo a extremo en una red IP, el IETF ha propuesto varias tecnologías y protocolos como son:

- Servicios Integrados IS (Integrated Services) o “Intserv”
- Protocolo de Reserva de Recursos RSVP (Resource Reservartion Protocol)
- Servicios Diferenciados DS (Differentiated Services) o “Diffserv”
- Protocolo de Transporte en Tiempo Real RTP (Real Time Procotol)
- Protocolo de Control en Tiempo Real RTCP (Real Time Control Protocol)
- Conmutación de Etiquetas Multiprotocol MPLS (Multiprotocol Label Switching)

Los mecanismos de QoS proporcionan un conjunto de herramientas que el administrador de redes puede utilizar para **administrar** el uso de recursos de red de una forma controlada y eficaz. Como resultado, se obtendrá un servicio mejor

en las aplicaciones y a usuarios de misiones críticas, al mismo tiempo que se va frenando el ritmo al que es necesario aumentar la capacidad. En resumen, QoS puede ayudar a mejorar el servicio a los usuarios de la red, al mismo tiempo que reduce los costos de ofrecer dichos servicios.

Los beneficios que se esperan como resultado de la implantación de QoS son:

- ✓ **Mejor rendimiento de aplicaciones de misiones críticas a través de vínculos de WAN:** QoS permite al administrador de la red favorecer el tráfico de misiones críticas para que sean inmunes a la congestión de los vínculos de WAN. Esto se puede conseguir con un costo mínimo para las aplicaciones menos significativas y competitivas. La solución QoS es parecida a proporcionar carriles especiales para cubrir ciertas necesidades en autopistas muy transitadas. El tráfico de estas misiones críticas se desvía a estos "carriles".
- ✓ **Controlar las repercusiones del tráfico multimedia en la red:** Las aplicaciones de transmisión multimedia generan grandes volúmenes de tráfico de protocolo de datagrama de usuario UDP. Este tráfico no es muy partidario de las redes en el sentido de que no "da marcha atrás" en caso de congestión. A consecuencia de las posibles repercusiones de este tipo de tráfico en recursos de red, los administradores de redes prohíben o limitan la implementación de aplicaciones multimedia en sus redes. Los mecanismos de QoS permiten al administrador de la red controlar las repercusiones de estas aplicaciones en la red.
- ✓ **Compatibilidad multimedia:** QoS se puede aplicar para garantizar una calidad de servicio específica a determinadas aplicaciones de medios de secuencias. En este caso, QoS permite convergencia real de redes de multimedia y de datos. Entre las ventajas que ofrece esta convergencia se puede destacar la telefonía IP utilizable con el ahorro de costos proporcional.

5.1 CÓMO FUNCIONA QOS

Las aplicaciones generan tráfico a ritmos variables y requieren normalmente que la red pueda transportar tráfico al ritmo que las aplicaciones lo han generado. Así mismo, las aplicaciones son más o menos tolerantes a retrasos de tráfico en la red y a variaciones de los mismos. Algunas aplicaciones pueden tolerar cierto grado de pérdida de tráfico, mientras que otras no. Si dispusiéramos de recursos de red infinitos, todo el tráfico de las aplicaciones podría transportarse al ritmo requerido, sin latencia y sin pérdida de paquetes, sería el ambiente ideal. Sin embargo, los recursos de red no son infinitos. Como consecuencia, hay partes de la red en las que los recursos no pueden responder a la demanda.

Las redes están construidas mediante la unión de dispositivos de red, tales como modificadores y enrutadores. Estos dispositivos se intercambian el tráfico entre ellos mediante interfaces. Si la velocidad en la que el tráfico llega a una interfaz es superior a la velocidad en la que la interfaz puede enviar tráfico al siguiente dispositivo, se produce una congestión. De esta forma, la capacidad de una interfaz para enviar tráfico constituye un recurso de red fundamental. Los mecanismos de QoS funcionan al establecer preferencias en la asignación de este recurso en favor de cierto tráfico. Para poder realizar esta acción, es necesario, en primer lugar, identificar tráficos diferentes. El tráfico que llega a los dispositivos de red se separa en distintos *flujos* mediante el proceso de **clasificación de paquetes**. El tráfico de cada flujo se envía a una *cola* en la interfaz de reenvío. Las colas de cada interfaz se *gestionan* de acuerdo con algunos algoritmos. El algoritmo de administración de cola determina la velocidad a la que se reenvía el tráfico de cada cola.
De este modo, se determinan los recursos que se asignan a cada cola y a los flujos correspondientes.

Para proporcionar QoS en redes, es necesario configurar y proporcionar a los dispositivos de red lo siguiente:

- Información de clasificación por la que los dispositivos separan el tráfico en flujos.
- Colas y algoritmos de administración de cola que controlan el tráfico de los diferentes flujos.

Nos referiremos a ambos como mecanismos de control de tráfico. Los mecanismos de control del tráfico por separado no resultan útiles. Deben proporcionarse o configurarse a través de muchos recursos de una forma coordinada que proporcione servicios de un extremo a otro en una red. Para proporcionar servicios útiles, son necesarios tanto los mecanismos de control de tráfico como los mecanismos de provisión y configuración.

5.2 ARQUITECTURAS Y PROTOCOLOS DE QOS

5.2.1 Arquitectura de Servicios Integrados IS (Integrated Services) o "Intserv"

Hacia 1995 había una tendencia en Internet a desarrollar un nuevo modelo para mejorar la calidad de servicio y fruto de esta filosofía es la denominada arquitectura de Servicios Integrados IS (Integrated Services).

La arquitectura de Servicios Integrados Intserv es una estructura que permite transmitir servicios integrados, siendo prioridad de esta tecnología ofrecer conexiones de alta calidad a clientes exigentes pero con costos superiores. La idea fundamental de la arquitectura de servicios integrados Intserv, radica en que las aplicaciones son vistas como flujos dentro de Internet y por cada flujo se deberá crear un estado ("soft state") en cada uno de los routers por donde fluyen estos flujos. En estos estados se realizará la reserva de los recursos necesarios para ofrecer QoS a las aplicaciones.

Las herramientas necesarias para el operador de la red si se desea Intserv son:

- Especificación del servicio
- Mecanismo de reservación del recurso
- Mecanismo de facturación

Normalmente, aunque no de forma necesaria, Intserv se asocia con el protocolo de señalización RSVP e implica una reserva de recursos en la red para cada flujo de información de usuario, así como el mantenimiento en la red (en los routers) de un estado para cada flujo, esto es, mantenimiento de la "reserva".
Esto conduce a un considerable tráfico de señalización y ocupación de recursos en cada router para cada flujo, con la consiguiente complejidad en el hardware, al margen del aporte que esta señalización hace a la congestión de la red, siendo Intserv una solución no escalable, no es una solución adecuada para grandes entornos como Internet porque aumenta considerablemente el numero de usuarios, aunque si lo es para entornos más limitados y también para redes de acceso al backbone.

El segundo problema que presenta Intserv es: el nivel de carga. El promedio del nivel de carga que mantiene es bajo, debido a que la predicción del trafico es difícil difiriendo mucho de ocupado a desocupado en horas diferentes.
Y el último problema es el costo, el cual es elevado en lo que se refiere a manejo, políticas y facturación en la red.

En la arquitectura IntServ ocupa un papel fundamental el concepto de flujo. Entendemos por flujo un tráfico continuo de datagramas relacionados entre sí que se produce como consecuencia de una acción del usuario y que requiere una misma calidad de servicio. Un flujo es unidireccional y es la entidad más pequeña a la que puede aplicarse una determinada calidad de servicio. Los flujos pueden agruparse en clases; todos los flujos de una misma clase reciben la misma calidad de servicio.

En IPv4 los flujos se identifican por las direcciones de origen y destino, el puerto de origen y destino (a nivel de transporte) y el protocolo de transporte utilizado (TCP o UDP).

En IPv6 la identificación puede hacerse de la misma forma que en IPv4, o alternativamente por las direcciones de origen y destino y el valor del campo Etiqueta de Flujo. Aunque el campo Etiqueta de Flujo en IPv6 se definió con este objetivo la funcionalidad aún no se ha implementado en la práctica.

En la arquitectura IntServ se definen tres tipos de servicio:

- ✓ **Servicio Garantizado:** Garantiza un caudal mínimo y un retardo máximo. Cada router del trayecto debe ofrecer las garantías solicitadas, aunque a veces esto no es posible por las características del medio físico (por ejemplo en Ethernet compartida).
- ✓ **Servicio de Carga Controlada:** Este servicio debe ofrecer una calidad comparable a la de una red de datagramas poco cargada, es decir en general un buen tiempo de respuesta, pero sin garantías estrictas. Eventualmente se pueden producir retardos grandes.
- ✓ **Servicio Best Effort:** Este servicio no tiene ninguna garantía.

Como ya se menciono, en la arquitectura IntServ es necesario contar con un protocolo que cree, mantenga y elimine recursos y este protocolo, generalmente, es el protocolo de Reserva de Recursos RSVP y que se encuentra presente tanto en los hosts extremos como en los routers. El protocolo RSVP está pensado fundamentalmente para tráfico multicast, ya que este tipo de tráfico es especialmente adecuado para la distribución de flujos de audio y vídeo en tiempo real que requieren unas condiciones estrictas de calidad de servicio. Sin embargo nada impide la utilización de RSVP en tráfico unicast. En una emisión multicast los usuarios pueden apuntarse o borrarse del grupo multicast de forma dinámica y sin advertencia previa, por ejemplo, en una red se emiten de forma multicast

diversos programas simultáneamente (equivalente a canales de televisión) y los usuarios desde sus hosts van continuamente haciendo “zapping” de un canal a otro, en un momento dado los usuarios que estén viendo un determinado canal forman un grupo multicast, pero el grupo puede cambiar con rapidez. Suponiendo que todos los programas se emiten desde el mismo host, este host será la raíz del árbol de expansión (spanning tree) de la emisión multicast; para cada programa multicast que se emite hay un conjunto de receptores que configuran un árbol de expansión diferente; esto es tarea del protocolo de routing multicast, no de RSVP.

Por tanto a partir de aquí supondremos resuelta esa parte del problema. El primero de los receptores del programa provoca la creación por parte del protocolo de routing del árbol de expansión y envía un mensaje de reserva hacia el emisor empleando el encaminamiento del camino inverso que hemos visto al hablar de routing multicast. Cada router por el que pasa el mensaje de reserva toma nota del ancho de banda solicitado y lo reserva, o bien devuelve un mensaje de error si no hay capacidad disponible. Si todo va bien al final del proceso el receptor ha reservado el ancho de banda necesario en todo el camino hasta la raíz del árbol.

Cuando aparece en la red un segundo receptor de esa misma emisión multicast envía su mensaje de reserva, pero la reserva sólo se efectuará en aquella parte del trayecto (o rama del árbol) que no sea común con el primer receptor y no haya sido por tanto ya reservada por éste. De esta forma se asegura un uso óptimo de la red, no reservando caudal dos veces en el mismo enlace, a la vez que se evita por completo la congestión (suponemos que RSVP no realiza sobresuscripción, es decir que no asigna recursos por encima de la capacidad disponible).

Es evidente, que aunque se trate de un protocolo Internet RSVP es un protocolo orientado a conexión, ya que los routers tienen que guardar una cierta

información de estado de cada flujo para el que se efectúe la reserva, algo equivalente a un circuito virtual.

En Intserv se definen un conjunto de mensajes para crear, mantener y eliminar los estados en cada nodo de la red, para reservar recursos a solicitud de una aplicación. Una aplicación en el host transmisor envía un mensaje Path hacia el host receptor para especificar los requerimientos de tráfico y definir el trayecto que seguirán los paquetes de datos. Cuando el mensaje Path llega al host receptor, este envía hacia el host transmisor un mensaje Resv para realizar la reserva de recursos en cada nodo de la red definido por el mensaje Path. Si en un nodo, al recibir un mensaje Path o Resv no puede interpretar (por ejemplo, tipo de reserva no definida), este nodo deberá generar un mensaje de error PathErr o ResvErr respectivamente. Cuando los mensajes Path y Resv son recibidos por los hosts correspondientes, los paquetes de datos serán enviados a la red. Estos mensajes deberán ser enviados a la red cada 30 segundos para refrescar las reservas asignadas. Cuando una reserva es realizada, esta puede ser informada al host receptor con el envío del mensaje ResvConf. Una vez finalizado el envío de datos, los estados creados y refrescados en cada nodo por los mensajes Path y Resv deberán ser eliminados por los mensajes PathTear y ResvTear respectivamente.

5.2.2 Protocolo de Reserva de Recursos RSVP (Resource Reservartion Protocol)

El protocolo de Reserva de Recursos RSVP es parte de un esfuerzo grande para reforzar la arquitectura del Internet actual con soporte de calidad de servicios siendo este un protocolo señalización que posibilita:

- Dar a las aplicaciones un modo uniforme para solicitar determinado nivel de QoS.
- Encontrar una forma de garantizar cierto nivel de QoS.
- Proveer autentificación.

RSVP es un protocolo que se desarrolla entre los usuarios y la red, y entre los diferentes nodos (*routers*) de la red que soportan este protocolo. Consiste en hacer “reservas” de recursos en dichos nodos para cada flujo de información del usuario, con la consecuente ocupación de los mismos. Esto requiere, lógicamente, intercambio de mensajes RSVP entre dichos entes funcionales, así como “mantener” estados de reserva en cada nodo RSVP. De manera que tanto la solicitud de las reservas, como el mantenimiento de éstas durante la comunicación, y la posterior cancelación, implican el intercambio de mensajes de señalización, lo que representa un tráfico considerable cuando de entornos como Internet se trata.

La necesidad de reservar los recursos de red difiere del tráfico de datos versus el tráfico en tiempo real, como sigue:
El tráfico de datos raramente necesita reservar banda ancha porque las interedes mantienen servicios de datagrama para el tráfico de los datos. Esta conmutación de paquete asíncrono puede o no necesitar garantías de calidad de servicio. Los controles de extremo a extremo entre remitentes del tráfico de datos y los receptores aseguran la transmisión adecuada de estallidos de información.
El trafico en tiempo real (es decir, voz o información de video) experimenta problemas al operar sobre los servicios del datagrama. Porque el tráfico en tiempo real envía un flujo casi constante de información, por lo que alguna garantía debe darse con tal de que el servicio entre los host en tiempo real no varíe para no correr el riesgo de la ruptura irrecuperable de la información que está enviándose.

El protocolo RSVP ofrece dos tipos de servicios: servicio de carga controlada y servicio garantizado.

- ✓ **Servicio de carga controlada:** Aunque no está muy bien definido, se entiende en general que la pérdida de paquetes debe ser muy baja o nula.
- ✓ **Servicio garantizado:** Se basa en solicitar determinado ancho de banda y cierta demora de tránsito máxima.

De los dos tipos de servicios que RSVP soporta, el más adecuado para aplicaciones con requerimientos de tiempo real es el servicio garantizado, aunque es más complejo de implementar que el servicio de carga controlada.

5.2.2.1 Mecanismos de funcionamiento

Para realizar la reservación de los recursos en la red, RSVP requiere de mensajes básicos, los cuales son:

- ✓ **Mensajes de Path:** Son generados por la fuente de mensajes de usuario necesitados de garantía de QoS, e indica las características de éstos en cuanto a recursos que necesita. La ruta que deben seguir estos mensajes es la misma que siguen los datos de usuario, para lo cual se requiere previamente un “diálogo” entre el proceso RSVP y el proceso de routing, pues dicha ruta quien la determina es el protocolo de routing, de lo contrario para nada serviría RSVP. En su paso por cada router RSVP los mensajes PATH´s se actualizan y se retransmiten, consistente esto en poner la dirección IP del router que lo actualiza y reenvía. Cada router RSVP también almacena la dirección del router anterior. Así, con los mensajes paths se posibilita indicar al receptor, o receptores, no solo las características del tráfico de usuario, sino también la ruta por donde debe solicitar las correspondientes reservas de recursos. Los routers que no soporten RSVP transfieren transparentemente los mensajes paths.
- ✓ **Mensajes de Resv:** Son producidos por el receptor (o receptores) de los flujos de información de usuario, como “respuesta” a los mensajes paths, y solicitan a la red (a los routers RSVP) las correspondientes reservas de recursos para soportar la comunicación con cierta QoS, fluyendo hasta la fuente del stream de datos de usuario, es decir, en sentido upstream. Con la información de ruta que suministran previamente los mensajes paths, los mensajes resvs dirigen las solicitudes de reservas a los routers RSVP apropiados, esto es, por donde fluirán los streams de datos. Los mensajes resvs especifican el ancho de banda mínimo que se requiere para obtener determinada demora en un stream de datos específico.

Además es posible efectuar reservas compartidas, esto es, una misma reserva aplicable a varios streams de datos de usuario. Estas reservas de recursos en los routers RSVP de la red se materializan mediante estados de reserva “soft-states” en dichos routers, estados que requieren para mantenerse de “refrescamientos” periódicos, por lo que durante toda la comunicación se necesita “señalizar” para mantener las reservas previamente efectuadas. En consecuencia, esto conlleva a cierta señalización “permanente” durante la fase de transferencia de información de usuario, con la consiguiente carga de tráfico que implica.

También la reserva de recursos extremo a extremo que posibilita RSVP será válida si, y solo si, la congestión y demora que introduzcan los routers no RSVP no es significativa.

Otros mensajes del protocolo que también son generados en el protocolo RSVP son:

- ✓ **Pathtear:** Son mensajes generados por la fuente de datos de usuario para eliminar los estados paths en todos los routers RSVP. Siguen la misma ruta que los mensajes paths. También pueden ser originados por cualquier nodo cuando se agota el timeout del estado path.
- ✓ **Resvtear:** Son generados por los receptores para borrar los estados de reserva en los routers RSVP, por tanto viajan en el sentido upstream. Pueden ser también originados por nodos RSVP al agotarse el timeout del estado de reserva de los mismos.
- ✓ **Patherrr:** Viajan en sentido upstream hacia el emisor siguiendo la misma ruta que los mensajes paths, y notifican errores en el procesamiento de mensajes paths, pero no modifican el estado del nodo por donde ellos pasan en su “viaje” hacia la aplicación emisora.
- ✓ **Resverr:** Notifican errores en el procesamiento de mensajes resv, o notifican la interrupción de una reserva. Se transfieren en la dirección downstream hacia el receptor o receptores apropiados.

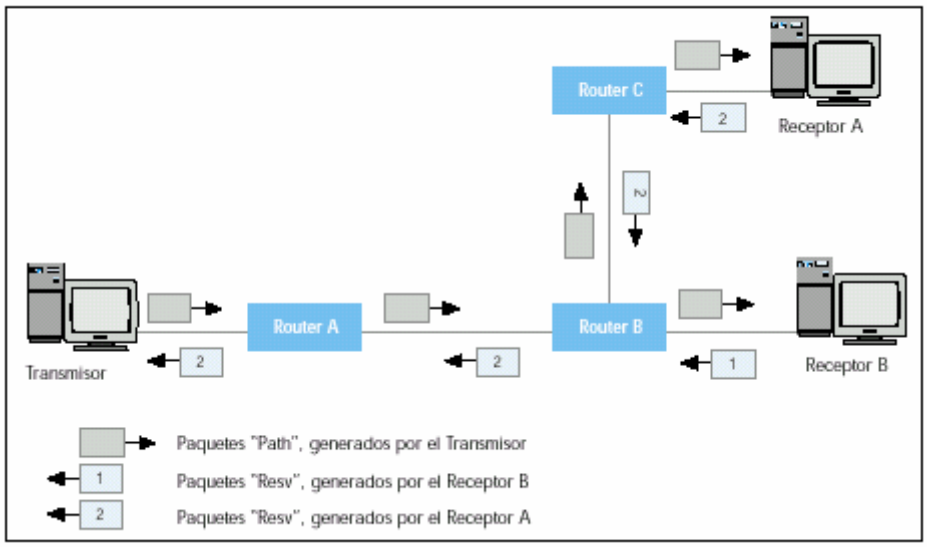


Figura. 5.1. Intercambio básico de mensajes

En la figura se muestra de forma muy simplificada el intercambio de mensajes RSVP, específicamente mensajes paths y resvs entre un emisor y dos receptores (A y B), indicándose que la reserva representada por el mensaje resv2 prevalece sobre la reserva representada por el mensaje resv1, de manera que esto sugiere que la reserva solicitada por el receptor A es mayor que la solicitada por el receptor B. Esto es, la reserva “mayor” prevalece sobre la reserva “menor”, así el router B sólo solicita al router A la mayor de las dos solicitudes de reservas a él llegadas desde el router C (originada por el receptor A) y desde el receptor B.

Esto es una característica de RSVP. Estas solicitudes de reserva conducen a que en cada router RSVP se establezca un estado de reserva, sofá state, es decir, una reserva en cada router es un estado con un determinado timeout, que debe ser refrescada periódicamente por los receptores, de lo contrario vence el timeout y se deshace la correspondiente reserva, con la consecuente generación de un mensaje resvtear. La liberación de recursos reservados mediante RSVP se puede materializar de diferentes maneras, así la solicitud para dar baja a determinada reserva puede ser originada:

- por el emisor,

- por el receptor, o
- por un nodo de la red.

Por parte del emisor o de un receptor acontece cuando así lo decide la aplicación correspondiente, en cuyo caso esto se produce mediante la generación de un mensaje pathtear o un mensaje resvtear, respectivamente. Por parte de un nodo se lleva a cabo cuando vence el timeout correspondiente del estado path o del estado de reserva, lo que origina la emisión de un mensaje pathtear o un mensaje resvtear, respectivamente.

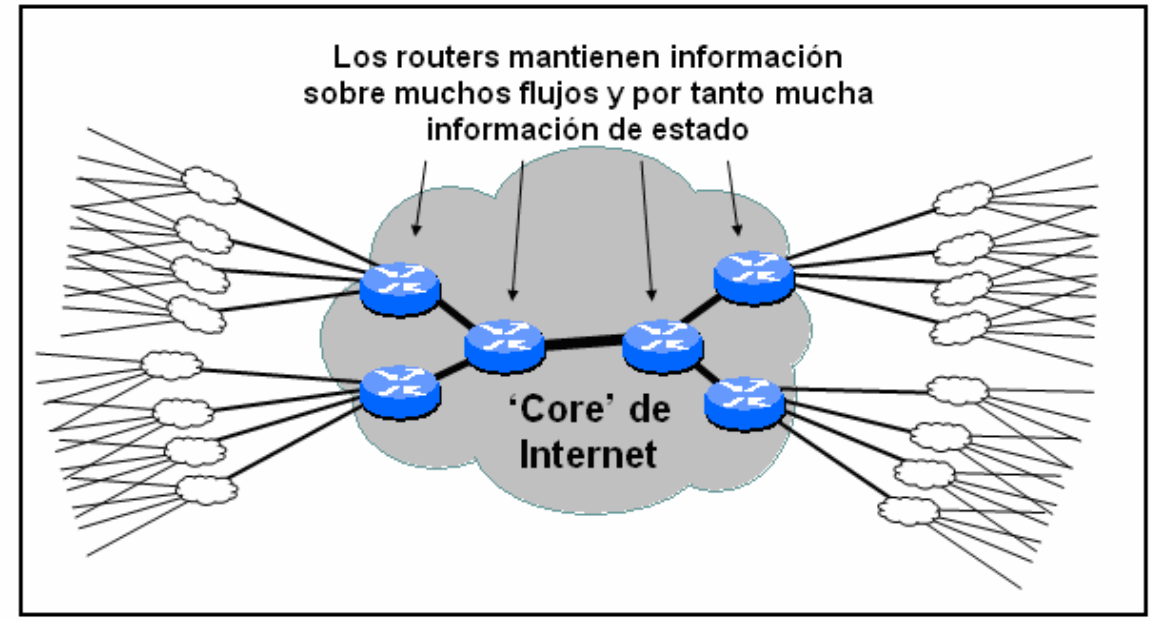


Figura 5.2. Problema de escalabilidad de RSVP

Cabe señalar también que los routers que estén en capacidad de implementar RSVP han de incorporar cuatro elementos:

- ✓ **Control de admisión:** Comprueba si la red tiene los recursos suficientes para satisfacer la petición. Equivalente al Control de Admisión de Conexión CAC (Connection Admission Control) de ATM.
- ✓ **Control de políticas:** Determina si el usuario tiene los permisos adecuados para la petición realizada (por ejemplo si tiene crédito

disponible). La comprobación se puede realizar consultando una base de datos mediante el protocolo COPS (Common Open Policy Service).

- ✓ **Clasificador de paquetes:** Clasifica los paquetes en categorías de acuerdo con la QoS a la que pertenecen. Cada categoría tendrá una cola y un espacio propio para buffers en el router.
- ✓ **Organizador de paquetes**: Organiza el envío de los paquetes dentro de cada categoría (cada cola).

5.2.3 Arquitectura de Servicios Diferenciados DS "Diffserv" (Differentiated Services)

Dado que la arquitectura Intserv seguía sin resolver el problema de la calidad de servicio en Internet hacia 1997 apareció un modelo alternativo denominado Arquitectura de Servicios Diferenciados DS (Differentiated Services). La idea básica de DiffServ consiste en que cada paquete lleva escrito un código que indica a que clase pertenece; se supone que los routers saben el tratamiento que han de dar a cada una de las clases posibles, por lo que no han de mantener ninguna información sobre conexiones o flujos concretos, el número de clases posibles es limitado e independiente del número de hosts o de flujos que pasan por los routers, por lo que la arquitectura DiffServ es escalable. En realidad DiffServ 'reinventó' hasta cierto punto el olvidado campo ToS de IPv4, que incluía entre otras cosas el subcampo precedencia.

La arquitectura de Servicios Diferenciados es un mecanismo de tratamiento del tráfico por acumulación apropiado para grandes redes enrutadas. Estas redes pueden transportar varios miles de conversaciones. Por tanto, no resulta práctico tratar el tráfico por conversación individual. Servicios Diferenciados Diffserv define un campo en los encabezados IP de los paquetes, conocido como Punto de Código de Servicios Diferenciados DSCP (Differentiated Services Code Point), este código es todo lo que necesitamos para identificar una clase de tráfico . Los

host o los enrutadores que envían tráfico a una red de servicios diferenciados diffserv marcan cada paquete transmitido con el valor DSCP. Los enrutadores de una red diffserv utilizan DSCP para clasificar paquetes y para aplicar un comportamiento de cola específico basado en los resultados de la clasificación. El tráfico de varios flujos con requisitos de QoS parecidos se marca con el mismo DSCP, al agregar el flujo a una cola común o al programar el comportamiento. Define y utiliza diferentes tipos de routers. Esta diferenciación no es la misma en los diferentes nodos, sino depende de si se trata de un nodo interior o un nodo frontera. En consecuencia, la red con nodos Diffserv no establece ni mantiene estados de las conexiones por flujos de paquetes. Es una solución escalable, más apropiada para grandes entornos como Internet. Puede ser "fácilmente" implementada en las redes IP existentes.

En la arquitectura definida por Diffserv aparece nodos frontera DS de entrada y salida, así como nodos DS internos. Este conjunto de nodos definen el dominio Diffserv y presenta un tipo de políticas y grupos de comportamiento por salto PHB (Per Hop Behavior) que determinarán el tratamiento de los paquetes en la red.

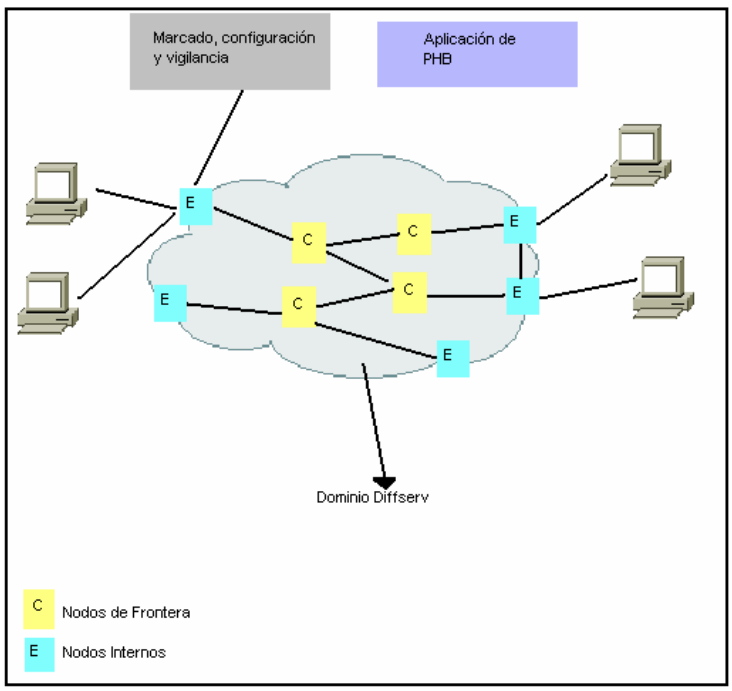


Figura. 5.3. Arquitectura de Servicios Diferenciados

Debemos tener en cuenta que un dominio Diffserv puede estar formado por más de una red, de manera que el administrador será responsable de repartir adecuadamente los recursos de acuerdo con el contrato de servicio SLA (Service Level Agreement) entre el cliente y el proveedor del servicio.

Veamos a continuación los diferentes tipos de nodos DS:

- ✓ **Nodos de frontera o externos DS:** Será necesario realizar diferentes funciones como el acondicionamiento de tráfico entre los dominios Diffserv interconectados. De esta manera debe clasificar y establecer las condiciones de ingreso de los flujos de tráfico en función de: dirección IP y puerto (origen y destino), protocolo de transporte y DSCP, este clasificador se conoce como Clasificador de Multicampo MF (Multi-Field Classifier). Una vez que los paquetes han sido marcados adecuadamente, los nodos internos deberán seleccionar el PHB definido para cada flujo de datos. Los nodos DS de entrada serán responsables de asegurar que el tráfico de entrada cumple los requisitos de algún Acuerdo de Acondicionamiento de Tráfico TCA (Traffic Conditioning Agreement), que es un derivado del SLA, entre los dominios interconectados. Por otro lado los nodos DS de salida deberán realizar funciones de Acondicionamiento de Tráfico TC (Traffic Conformation) sobre el tráfico transferido al otro dominio DS conectado.
- ✓ **Nodos internos DS:** Podrá realizar limitadas funciones de Acondicionamiento de Tráfico TC, tales como remarcado de DSCP. Los nodos DS internos solo se conectan a nodos internos o a nodos de frontero o externos de su propio dominio. A diferencia de los nodos de frontera o externos para la selección del PHB solo se tendrá en cuenta el campo DSCP, conocido como Clasificador BA (Behavior Aggregate Classifier).

Como se explico anteriormente para escribir la información sobre QoS de cada datagrama se utiliza un campo de un byte en la cabecera DS que se estructura así:

Subcampo	Longitud (bits)
Punto de Código de Servicios Diferenciados DSCP (Differentiated Servcies CodePoint)	6
Notificación de Congestión Explicita ECN (Explicit Congestion Notification)	2

Tabla. 5.1. Estructura del campo de Servicios Diferenciados

DSCP nos permite definir en principio hasta 2^6 = 64 posibles categorías de tráfico, aunque en la práctica se utilizan bastante menos. Los valores de DSCP se dividen en los tres grupos siguientes:

Punto de código	Posibles Valores	Uso
Xxxyy0	32	Estándar
xxxx11	16	Local/Experimental
xxxx01	16	Reservado

Tabla. 5.2. Grupos de puntos de código del campo DS

Así pues, de momento se contemplan 32 posibles categorías de datagramas, correspondientes a los cinco primeros bits del campo DS.

Se han definido tres tipos de Servicios Diferenciados Diffserv:

- Servicio de Reenvío Seguro AFS (Assured Forwarding Service)
- Servicio de Reenvío Apresurado EFS (Expedited Forwarding Service).
- Servicio de Mejor Esfuerzo (Best Effort)

Siendo los dos primeros de QoS garantizada.

5.2.3.1 Servicios de Reenvío Seguro AFS (Assured Forwarding Service)

Los paquetes se etiquetan con “alta prioridad”,es decir asegura un trato preferente, aunque no se garantiza caudales, retardos, ancho de banda, etc.. Se posibilita una QoS superior al servicio tradicional de mejor esfuerzo de Internet. Se definen cuatro clases de servicios posibles pudiéndose asignar a cada clase una cantidad de recursos en los routers (ancho de banda, espacio en buffers, etc.). La clase se indica en los tres primeros bits del DSCP. Para cada clase se definen tres categorías de descarte o “dropping” de paquetes (probabilidad alta, media y baja) que se especifican en los dos bits siguientes (cuarto y quinto).

El servicio diferenciado de Reenvío Seguro define un método por el que Clasificador Agregado de Ambiente BA puede dar condiciones de reenvió diferentes. En estos nodos, la selección del Comportamiento del Punto de Control PHB (Per Hop Behavior) es realizada sólo analizando el contenido del punto de código de servicios diferenciados DSCP.

Algunos proveedores de servicios Internet ISPs ofrecen servicios denominados ‘olímpicos’ con categorías denominadas oro, plata y bronce (o tiempo-real, negocios y normal), asignándose el ancho de banda así:

- Oro: El tráfico en esta categoría se asigna 50 por ciento del ancho de banda disponible.
- Plata: El tráfico en esta categoría se asigna 30 por ciento del ancho de banda disponible.
- Bronce: El tráfico en esta categoría se asigna 20 por ciento del ancho de banda disponible.

Más allá, el comportamiento del punto de control de Reenvío Seguro PHB AF define cuatro clases de Reenvío seguro AF: AF1, AF2, AF3, y AF4. Cada clase se asigna una cantidad específica de espacio del buffer y ancho de banda de la

interfase, según los Acuerdos de Nivel de Servicio SLA con el proveedor de servicio.

Dentro de cada clase Reenvío seguro AF, se puede especificar tres valores de Precedente de Eliminación dP (drop precedence): 1, 2, y 3.

El comportamiento del punto de control PHB de Reenvió Seguro puede expresarse como se muestra en el ejemplo siguiente:

En este ejemplo, “n” representa los números de las clases AF (1, 2, o 3) y “y” representa los valores de dP (1, 2, o 3) dentro de la clase de Reenvío Seguro AFn.

En casos de congestión de tráfico de red, si los paquetes en una clase AF particular (por ejemplo, AF1) necesita ser eliminado, se eliminara los paquetes en la clase AF1 según la pauta siguiente:

dP(AFny) >= dP(AFnz) >= dP(AFnx)

donde:

- dP (AFny): Es la probabilidad de eliminar los paquetes de la clase de Afny, en otras palabras y denotando el dP dentro de una clase de AFn.

En el ejemplo siguiente, se eliminaran paquetes en la clase AF13 antes que los paquetes en la clase AF12 que a su vez se eliminara antes de los paquetes en la clase AF11:

dP(AF13) >= dP (AF12) >= dP(AF11)

El método de dP castiga flujos de tráfico dentro de un BA particular que excede el ancho de banda asignado. Los paquetes en estos flujos podrían ser remarcados por un policer a un precedente de eliminación más alto.

Una clase de AFx puede ser denotada por el valor de DSCP, xyzab0, donde los xyz pueden ser 001, 010, 011, o 100, y el ab representa el valor del dP.

Valor DSCP y correspondiente Precedente de Eliminación	Clase 1	Clase 2	Clase 3	Clase 4
Precedente de Eliminación Bajo	001010	010010	011010	100010
Precedente de Eliminación Medio	001100	010100	011100	100100
Precedente de Eliminación Alto	001110	010110	011110	100110

Tabla. 5.3. Listas del valor de DSCP y los dP correspondientes cada clase de PHB AF.

En el servicio AF el proveedor puede aplicar políticas de tráfico al usuario, y si el usuario excede lo pactado el proveedor puede eliminar datagramas, o bien aumentar la precedencia de la eliminación.

5.2.3.1 Servicio de Reenvío Apresurado EFS (Expedited Forwarding Service)

Equivale a una línea arrendada virtual, por lo que se garantiza cierto ancho de banda y reducida demora de cola. Emula un circuito. Este servicio es el de mayor calidad. El valor del subcampo DSCP relacionado con este servicio es '101110'.

El comportamiento del Punto de Conexión del Reenvío Apresurado PHB EF, es un ingrediente importante de los Servicios diferenciados DiffServ, los suministros este tipo de servicio robusto proporcionando pérdida baja, retardo bajo, jitter bajo y el servicio del ancho de banda seguro.

El servicio de Reenvío Apresurado EF puede ser implementado usando una Cola de Prioridad PQ (Priority Queuing), junto con tasa limitada en la clase.

Cuando se implemento en una red de DiffServ, PHB EF proporciona una línea arrendada virtual, o servicio del premium. Para una eficacia óptima, sin embargo, PHB EF debe ser reservado sólo para las aplicaciones más críticas porque, en casos de congestión de tráfico, no es factible tratar todos o la mayoría el tráfico como prioridad alta. PHB EF está idealmente preparado para las aplicaciones

como VoIP que requiere ancho de banda bajo, ancho de banda garantizado, retardo bajo y jitter bajo.

5.2.3.2 Servicio de Mejor Esfuerzo (Best Effort)

Este servicio se caracteriza por tener a cero los tres primeros bits del DSCP. En este caso los dos bits restantes pueden utilizarse para marcar una prioridad, dentro del grupo 'best effort'. En este servicio no se ofrece ningún tipo de garantías.

5.2.3.3. Componentes de Servicios Diferenciados

Un nodo Servicios diferenciados DS es, en principio, una combinación de cinco módulos funcionales, aunque no todo router DS tiene que contener la totalidad de éstos:

- Clasificador de tráfico
- Medidor de tráfico (Traffic Meter)
- Marcador de paquetes (Packet Markers)
- Conformador (Shapers)
- Droppers

Clasificador de tráfico

La clasificación del paquete usa un descriptor de tráfico (por ejemplo, el DSCP) para categorizar un paquete dentro de un grupo específico a fin de definir ese paquete. Después de que el paquete ha sido definido (es decir, clasificado), el paquete es luego accesible para el manejo de QoS en la red.
Usando la clasificación del paquete, se puede dividir el tráfico de la red en niveles de prioridad múltiples o clases de servicio. Cuando descriptores de tráfico se usan

para clasificar el tráfico, la fuente está de acuerdo en adherir a los términos contraídos y la red promete una calidad de servicio. El policers de tráfico y formadores de tráfico usan el descriptor de tráfico del paquete (es decir, la clasificación del paquete) para asegurar adhesión a ese acuerdo.

- ✓ **Clasificador de Agregados de Comportamiento (BA):** Selecciona paquetes basádose exclusivamente en el campo DS
- ✓ **Clasificador MultiCampo (MF):** Selecciona paquetes en base a varios campos: protocolo, puerto, direcciones IP

Medidor de tráfico (Traffic Meter)

Mide las propiedades temporales de los paquetes como son la velocidad y tamaño de ráfagas que se realiza en los bordes de un dominio de Diffserv.
Los acondicionadores de tráfico realizan las funciones de formación del tráfico y mantenimiento del orden para asegurar que el tráfico que entra en el dominio de DiffServ se adapta a las reglas especificadas por el Acuerdo de Acondicionamiento del Tráfico TCA (Traffic Conditioning Agreement), y obedece el servicio que proporciona la política del dominio.

Marcador de paquetes (Packet Markers)

El marcado del paquete se relaciona a la clasificación del paquete. El marcado del paquete le permite clasificar un paquete basado en un descriptor de tráfico específico (como el valor de DSCP). Esta clasificación puede usarse para aplicaciones de usuario basadas en servicios diferenciados y asociar un paquete con un grupo de QoS local.
La asociación de un paquete con un grupo de QoS local les permite a los usuarios asociar a un grupo ID con un paquete. El grupo ID puede ser usado para clasificar paquetes en los grupos de QoS basados en el prefijo, sistema autónomo, y cadena de la comunidad. Un usuario puede establecer hasta 64 valores de DSCP y 100 grupos de marcado de QoS.

Conformador (Shapers)

Establece cierta demora para uno o más paquetes de un stream. El manejo de congestión (o organización) se logra a través de la ordenación del tráfico y encolamiento del tráfico. Cuando hay congestión de la red, un mecanismo de planificación como la Cola CBWFQ se usa para proporcionar ancho de banda garantizado a las diferentes clases de tráfico.

Droppers

Las técnicas de anulación de congestión supervisan las cargas de tráfico de red en un esfuerzo por anticipar y evitar la congestión en los cuellos de botella de la red comunes. La anulación de congestión se logra a través eliminación de paquetes. Entre los más comúnmente usados mecanismos de anulación de congestión esta la Detección Temprana al Azar de Carga WRED (Weighted Random Early Detection).

Con WRED y Servicios Diferenciados, se tiene la opción de permitir WRED usar el valor de DSCP cuando WRED calcula la probabilidad de eliminación de un paquete.

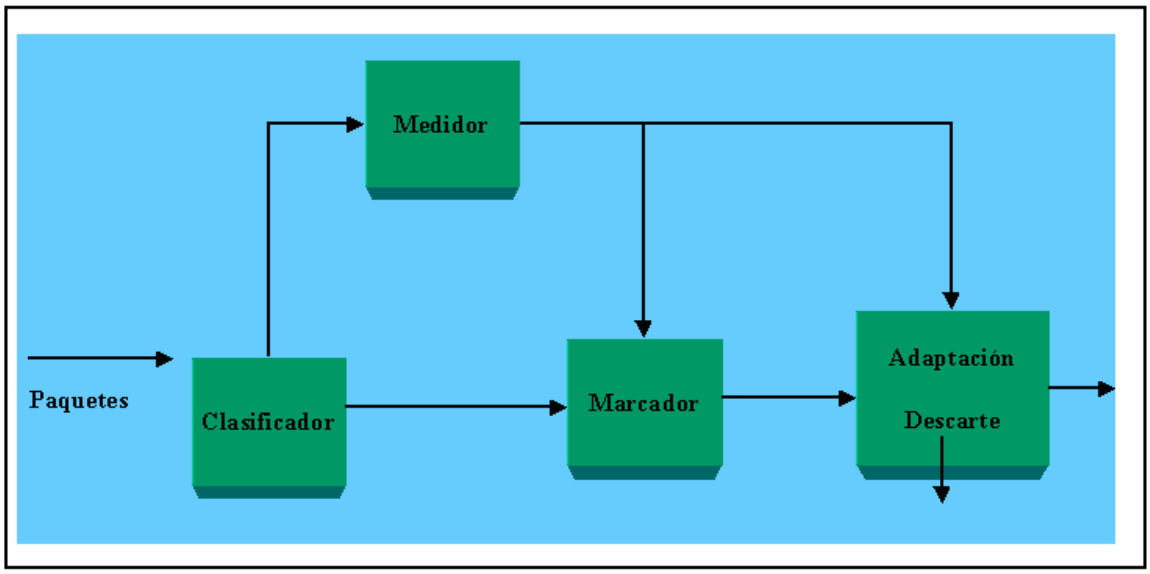


Figura. 5.4 Elementos de un nodo de Servicios Diferenciados

Los tipos de routers en redes Diffserv se clasifican así:

Router de primera conexión (First Hop Router): Es el router más próximo al host emisor de paquetes. Los flujos de paquetes son clasificados y marcados acorde a la etiqueta SLA (Service Level Agreement). Es responsable de que el tráfico esté acorde con el ancho de banda del perfil.

Router de Ingreso (Ingrees Router): Se sitúan en los puntos de entrada al backbone Diff-Serv (dominio DS), efectuando la clasificación de los paquetes en base al campo DS o en base a múltiples campos de la cabecera de éstos.

Router de Salida (Egress Router): Se ubican en los puntos de salida de redes DiffServ (dominio DS), controlando el tráfico. Efectúan la clasificación de paquetes en base solo al campo DS de las cabeceras.

Router Interior (Interior router): Tienen la misión de “sumar” flujos, realizar la clasificación DS y reenvío de paquetes. Se sitúan dentro del backbone DS (dominio DS).

Como ya se menciono, el campo DS es una incorporación reciente en la cabecera IP. Anteriormente existía en su lugar un campo denominado TOS o ‘Tipo de Servicio’ que tenía la siguiente estructura:

Figura. 5.5. Estructura del campo 'Tipo de servicio'

El subcampo 'Precedencia' permitía especificar una prioridad entre 0 y 7 para el datagrama (7 es la máxima prioridad). Este campo es en cierto modo el antecesor del campo DS. A continuación se encontraba un subcampo compuesto por cuatro bits que actuaban como indicadores o 'flags' mediante los cuales el usuario podía indicar sus preferencias respecto a la ruta que seguiría el datagrama. Los flags, denominados D, T, R y C permitían indicar si se prefería una ruta con servicio de bajo retardo (D=Delay), elevado rendimiento (T=Throughput), elevada fiabilidad (R=Reliability) o bajo costo (C=Cost). El campo ToS ha sido muy impopular: el subcampo precedencia se ha implementado muy raramente en los routers. En cuanto a los flags D,T,R,C prácticamente no se han utilizado y su inclusión en la cabecera IP ha sido muy criticada. Estos problemas facilitaron evidentemente la 'transformación del campo ToS en el DS, aunque existen todavía routers en Internet que interpretan este campo con su antiguo significado de campo ToS.

A continuación analizamos los problemas de implantación de Diffserv:

- Dentro del modelo Diffserv es necesario indicar que este no asegura de manera determinista que los flujos de tráfico consigan determinados parámetros QoS, como puede hacer ATM a través de circuitos. Diffserv permite la creación de agregaciones de tráfico, lo que nos ofrece "cierta probabilidad de QoS", de manera que un proveedor puede integrar las conexiones pertenecientes a diferentes VPNs dentro de un mismo agregado, recibiendo todas ellas las mismas prestaciones a nivel de red. De esta manera el tratamiento que recibieran podría ser diferente del que consiguen usuarios con acceso gratuito a Internet. Sin embargo conseguir que el modelo Diffserv se ponga en funcionamiento requerirá un gran trabajo de ingeniería de red así como un dimensionamiento adecuado para

alcanzar determinados parámetros QoS como puede ser un caudal o retardo asegurados.

- Resulta interesante estudiar la situación en que se encuentra actualmente Internet, donde debemos atravesar diferentes ISPs para alcanzar nuestro destino. En este caso el valor del byte DS puede ser modificado en cualquier equipo intermedio según las políticas de tráfico y diferentes contratos SLA tengan entre estos ISPs. De esta forma, una calidad extremo a extremo sólo será alcanzable cuando todos los elementos involucrados en la cadena (dominios Diffserv) actúen según las mismas políticas.
- Además resulta especialmente curioso que en Diffserv el principal beneficiario de la reserva de QoS será el destino, siendo el origen el que debe pagar por conseguir ese trato diferenciado de su tráfico. De esta forma surgen conflictos por ejemplo en la descarga de audio-streaming, donde el que pagaría sería el servidor en lugar del usuario receptor.
- Analizando el modelo Diffserv parece lógico que alcanzar un destino más lejano resulte más caro que otro más cercano donde se necesiten atravesar menos ISPs. En consecuencia el costo de enviar un paquete será diferente en función del camino que deba atravesar. Esto puede suponer una complicación a la hora de ofrecer el servicio y tarifarlo. Sin embargo, este mismo problema aparecía en el nacimiento de Internet, donde también resultaba más caro enviar un paquete cuantos más ISPs tuviese que atravesar, y sin embargo, por el momento los proveedores de acceso han decidido ofrecer una tarifa fija independientemente del destino de los paquetes. Parece lógico entonces pensar que de alguna manera nuestro proveedor de acceso a Internet nos tarifará adecuadamente teniendo en cuenta que los mensajes deberán ser tratados adecuadamente en los diferentes ISPs (con los costos que ello suponga).
- Por otro lado, el modelo Diffserv no permite lograr QoS en la red de acceso. Cuando se habla de QoS extremo a extremo en Diffserv típicamente se hace referencia a QoS entre los routers de frontera entre origen y destino. No obstante, la solución no presenta muchas dificultades ya que se supone que el usuario tiene mayor control sobre la red de

acceso, quedando un dimensionamiento adecuado en manos del usuario final.

- Otra de las consecuencias del modelo Diffserv es que la reserva de QoS es unidireccional. En muchos casos esto no planteará ningún problema. Sin embargo consideremos el establecimiento de una conexión TCP. Si la reserva es unidireccional los paquetes ACK que viajen en sentido contrario tendrán el tratamiento normal (best-effort) de paquetes, lo que podría llevar a que la QoS final conseguida se limitase a la de los paquetes ACK (que limitan el manejo de la ventana de transmisión).
- Finalmente, el modelo Diffserv plantea ciertos problemas a la hora de decidir quien es el encargado de marcar la QoS en los paquetes. Si se desease que el usuario pudiese elegir personalmente el tratamiento deseado, entonces sería necesario modificar de alguna forma las aplicaciones y/o la pila de protocolos. Considerando poco deseable esta opción, ya que limitaría el acceso a esta tecnología, otra posibilidad consiste en crear algún sistema de comunicación con nuestro proveedor de acceso que nos permita indicar nuestros gustos de QoS en función del servicio.

A partir del estudio de estas limitaciones que presenta el modelo Diffserv podemos distinguir algunas aplicaciones o servicios como las más interesantes para este modelo.

- En primer lugar los servicios basados en suscripción cobran especial importancia. Debido al hecho de ser el origen el encargado de realizar la reserva, servicios como Pagar por ver (Video en demanda), canales de radio, canales de televisión, etc., podrían aparecer en el modelo de negocio de redes Diffserv. En este caso el proveedor de contenidos recibiría cierto ingreso por cada evento distribuido. Y sería el mismo el encargado de seleccionar la calidad de servicio que recibirían los usuarios.
- Siguiendo el mismo razonamiento, el principal problema es poner de acuerdo a origen y destino para alcanzar un acuerdo en la QoS deseada. Este problema desaparecería en el caso de redes virtuales privadas VPNs donde el origen y el destino pertenecen a la misma organización, de

manera que comparten los mismos criterios sobre QoS. De esta manera, el desarrollo de Diffserv podría presentar un especial interés de cara a la creación de VPNs sobre una red IP.

Por otro lado, existen una serie de aplicaciones con determinados requisitos de QoS donde el desarrollo e implementación de alguna tecnología de QoS en la actual Internet podría suponer su despegue. A continuación podemos ver algunos ejemplos:

- Resulta de especial interés en las videoconferencias, donde incluimos cualquier tipo de escenario VoIP. Este servicio podría representar una buena fuente de ingresos en el modelo Diffserv. El desarrollo de este tipo de comunicaciones no ha tenido éxito hasta la fecha por la falta de algún tipo de provisión de QoS, pero la llegada de Diffserv podría suponer el despegue definitivo de estos servicios.
- Otro caso interesante podrían ser los juegos online. Suele tratarse de aplicaciones que no requieren un gran ancho de banda, pero si importantes requisitos de retardo. La existencia de una gran plataforma de videojuegos está supeditada a la provisión de QoS.

Beneficios

Los beneficios de implementar Servicios Diferenciados incluyen:

- Reducir la carga en dispositivos de la red y fácilmente escalar como la red crece.
- Permitir a clientes para mantener cualquier esquema de priorización de ToS de la Capa 3 existente que puede estar en uso.
- Permitir a clientes mezclar dispositivos DiffServ con cualquiera equipo que habilita ToS existente en uso.
- Aligerar los cuellos de botella a través del manejo eficaz de recursos de red.

5.2.4 Protocolo de Transporte en Tiempo Real RTP (Real Time Procotol)

El Protocolo de Transporte en Tiempo Real RTP (Real Time Protocol) es un protocolo basado en IP que provee soporte para el transporte de datos en tiempo real como flujos de audio y video, que se encuentran disponibles desde 1996. Incluye: la identificación de carga útil, la numeración secuencial, la medición de tiempo y el reporte de la calidad.

Entre sus funciones se encuentran: la memorización de datos, la simulación de distribución interactiva, el control y mediciones de aplicaciones.

Está diseñado para trabajar en conjunto con el protocolo auxiliar de Control en Tiempo Real RTCP para obtener información sobre calidad de la transmisión y participantes de la sesión.

El RTP trabaja en capa 4 y sobre el protocolo de datagrama de usuario UDP, de forma que posee funciones de control para la detección de error y la posibilidad de multiplexación de puertas (puerto UDP), por ser un protocolo de transporte no orientado a la conexión, no ofrece confiabilidad, por lo que no generará retransmisiones que puedan congestionar la red (para datos en tiempo real, la confiabilidad no es tan importante como la entrega rápida). Las sesiones de protocolo RTP pueden ser multiplexadas. Para ello se recurre a un doble direccionamiento mediante las direcciones IP y el número de puerto en UDP. Sobre RTP se disponen de protocolos de aplicación del tipo H.320/323 para vídeo y voz.

El protocolo RTP funciona en conjunto con RSVP (capa 3) para la reservación de ancho de banda y asegurar de esta forma la QoS del tipo garantizada. La QoS del tipo diferenciada se logra mediante la priorización de

tráfico que puede adoptar dos alternativas. En IP se pueden asignar diversas alternativas de prioridad para formar una cola de espera en routers.

Las características principales de RTP son:

1. **Timestamping u hora de ocurrencia:** El emisor setea el timestamp de acuerdo al instante en que el primer octeto del paquete fue muestreado. El timestamp aumenta mientras se completa un paquete. Por otro lado, el receptor usa estas marcas para reconstruir el sincronismo original para reproducir la secuencia a la frecuencia correcta.
2. **Secuenciación:** Debido a la necesidad de entregar los paquetes en orden (UDP no provee esta característica) RTP incorpora un número de secuencia que además sirve para la detección de paquetes perdidos).
3. **Identificación de fuente:** Permite conocer al receptor de la información de dónde provienen los datos.
4. **Identificador de tipo de carga útil "payload":** Especifica el formato de datos de la carga de RTP, referente al formato de codificación, de manera que la aplicación receptora pueda saber como reproducirlo.

El protocolo RTP además provee transporte para direcciones unicast y multicast. El paquete de RTP incluyen un encabezado fijo y el payload de datos, el protocolo de control RTCP utiliza el encabezamiento del RTP y ocupa el campo de carga útil.

Un protocolo conocido como Protocolo de Compresión del Encabezamiento en Tiempo Real RTP-HC (Real-Time Protocol Header Compression) permite la compresión del encabezado para mejorar la eficiencia del enlace en paquetes de corta longitud en la carga útil. Se trata de reducir los 40 Bytes de encabezado en RTP/UDP/IP a una fracción de 2 a 5 Bytes, eliminando aquellos que se repiten en todos los paquetes. Como los servicios de tiempo real generalmente trabajan con

paquetes pequeños y generados en forma periódica se procede a formar un encabezado de longitud reducida que mejore la eficiencia de la red.

5.2.5 Protocolo de Control en Tiempo Real RTCP (Real Time Control Protocol)

El Protocolo de Control RTCP (Real Time Control Protocol) permite completar a RTP facilitando la comunicación entre extremos para intercambiar información y monitorear de esta forma la calidad de servicio y obtener información acerca de los participantes en la sesión. RTCP se fundamenta en la transmisión periódica de paquetes de control a todos los participantes en la sesión usando el mismo mecanismo de RTP de distribución de paquetes de datos. El protocolo UDP dispone de distintas puertas (puerto UDP) como mecanismo de identificación de protocolos.

La función primordial de RTCP es la de proveer una realimentación de la calidad de servicio. Se relaciona con el control de congestión y flujo de datos. El RTCP involucra varios tipos de mensajes, por ejemplo:

- ✓ **Reporte de Recepción RR (Receiver Report):** Es enviado por los receptores y contiene información sobre la calidad de la entrega de datos, incluyendo último número de paquete recibido, número de paquetes perdidos y timestamps para calcular el retardo entre el emisor y el receptor.
- ✓ **Reporte de Entrega SR (Send report):** Es enviado por el emisor y además de contener información similar a los mensajes RR, incorpora datos sobre sincronización, paquetes acumulados y número de bytes enviados.

- ✓ **Descripción de la Fuente SDES (Source Description):** Para un identificador de nivel de transporte denominado Nombre Canónico CNAME (Canonical Name), contiene in formación que describe al emisor.
- ✓ **Adios (Bye):** Para indicar el final de la participación en la conexión.
- ✓ **Aplicación de funciones especificas APP(Application specific functions):** Para aplicaciones específicas, por ahora es experimental, esta reservado para aplicaciones futuras.

El mensaje reporte de entrega SR, uno de los más interesantes, disponen de 3 secciones bien diferenciadas:

- ➢ Los primeros 8 Bytes se refieren a un encabezado común.
- ➢ La segunda parte de 20 Bytes permite la evaluación de diferentes parámetros (retardo, jitter, eficiencia de datos, etc).
- ➢ La tercera parte de 24 Bytes lleva reportes que han sido obtenidos desde el último reporte informado. Incluye los siguientes reportes: cantidad total de paquetes RTP perdidos y a la proporción de los mismos; la cantidad de paquetes recibidos y el jitter entre paquetes; el horario del último paquete recibido y el retardo de transmisión del mismo.

5.2.6 Conmutación de Etiquetas Multiprotocolo MPLS (Multiprotocol Label Switching)

Una de las soluciones propuestas por IETF, con el objetivo de proporcionar calidad de servicio QoS a una red de datos es la Conmutación de Etiquetas Multiprotocolo MPLS (MultiProtocol Label Switching), se trata de un estándar de arquitectura multinivel, capaz de soportar cualquier tipo de tráfico e independiente del nivel de transporte de datos sobre el que se apoya, capaz de ofrecer una gran eficiencia a la hora de realizar la transmisión de paquetes de un extremo a

otro de la red MPLS gracias a la combinación de la flexibilidad del nivel de red IP con los beneficios propios de un modelo de red orientado a conexión.

En una red IP tradicional, un router conmuta los paquetes de una interfaz de entrada a una interfaz de salida; además, actualiza la información de enrutamiento.

Para enviar los paquetes, debe examinar la cabecera del paquete IP para cada paquete. Estas dos funciones, envío y enrutamiento, tienen lugar en cada salto que realiza un paquete para cada uno de los paquetes que atraviesan la red. Lo que se busca con MPLS a este respecto es llevar las funciones de enrutamiento únicamente a los equipos exteriores del dominio MPLS, de forma que en el interior de dicho dominio no sea necesario realizar labores de enrutamiento, sino sólo de conmutación mediante la consulta de unas etiquetas añadidas a cada paquete en el momento de entrada al dominio.

Es decir, MPLS es una estructura especifica que proporciona la designación, direccionamiento, envío y conmutación de flujos de tráfico a través de la red. Realiza lo siguiente:

- Especifica los mecanismos para manejar los flujos de tráfico como flujos entre el hardware diferente, las máquinas o incluso los flujos entre las aplicaciones diferentes.
- Permanece independiente de los protocolos de capa 2 y capa 3.
- Proporciona medios para trazar direcciones de IP a simple, etiquetas de longitud fija usada por el reenvió de paquetes diferentes y tecnologías de conmutación de paquete.
- Interfaces para protocolo de direccionamiento existentes, como el protocolo de Reservación de Recurso RSVP.
- Soporta IP, ATM, y protocolos de la capa 2 de Frame Relay.
- Separa las funciones de control y envío.
- Envía paquetes a través de IP a gran velocidad de un modo simplificado y eficiente.

En MPLS, la transmisión de los datos ocurre en Rutas de conmutación de Etiquetas LSPs (Label Switched Paths). Las rutas de conmutación de etiquetas LSPs son una sucesión de etiquetas en cada uno y todos los nodos a lo largo de la ruta desde la fuente al destino. Las rutas de conmutación de etiquetas LSPs son establecidas por cualquier anticipo para la transmisión de los datos (control-manejo) o en la detección de un cierto flujo de datos (datos-manejo). Las etiquetas, estando debajo de identificadores de un protocolo específico, hay algunos protocolos de distribución de etiquetas usados actualmente, como el Protocolo de Distribución de Etiqueta LDP (Label Distribution Protocol) o el protocolo RSVP o plataformas en protocolos de direccionamiento como el Protocolo de Gateway de Borde BGP (Border Gateway Protocol). Cada paquete de datos encapsula y lleva las etiquetas durante su jornada desde la fuente al destino. El conmutador de alta velocidad de datos es posible porque las etiquetas de longitud fija son insertadas en cada inicio del paquete o célula y pueden ser usadas por hardware para cambiar paquetes rápidamente entre los enlaces.

MPLS es una solución versátil para direccionar los problemas enfrentados por la red como son: velocidad, escalabilidad, el manejo de calidad de servicio QoS e ingeniería de tráfico.

5.2. 6.1 Elementos básicos de la arquitectura MPLS

MPLS ha surgido como una solución elegante para reunir el manejo de ancho de banda y requisitos de servicio para la próxima generación de redes basadas en IP, podemos decir que los elementos básicos son:

- ✓ **Router de conmutación de etiqueta LSRs (Label Switched Router) :** Nodo MPLS capaz de enviar paquetes de nivel 3 nativos. El LSR puede ser interior o extremo. Los LSR extremo añaden o eliminan etiquetas. Los interiores sustituyen unas etiquetas por otras.
- ✓ **Etiqueta:** Es un identificador corto, de longitud fija y con significado local empleado para identificar un FEC. Un paquete puede tener una o más

etiquetas apiladas (jerarquía). Cuando un paquete atraviesa dominios interiores a otros dominios, es cuando se produce el apilamiento de etiquetas. El LSR siempre consultará la etiqueta de nivel superior. La etiqueta se añade, de forma general, entre las cabeceras de nivel 2 y 3 (Ethernet, PPP...).

- ✓ **Clases de equivalencia de Reenvío FEC (Forwarding Equivalence Class)**: Agrupación de paquetes que comparten los mismos atributos (dirección destino, VPN..) y requieren el mismo servicio (multicast, QoS...). El FEC se asigna en el momento en que el paquete entra a la red. Todos los paquetes que forman parte del FEC, siguen un mismo LSP.
- ✓ **Ruta de conmutacíon de etiquetas LSP (Label Switched Path):** Ruta a través de uno o más LSRs en un nivel de jerarquía que sigue un paquete de un FEC en particular. Esta ruta puede establecerse tanto mediante protocolos de enrutamiento como manualmente.

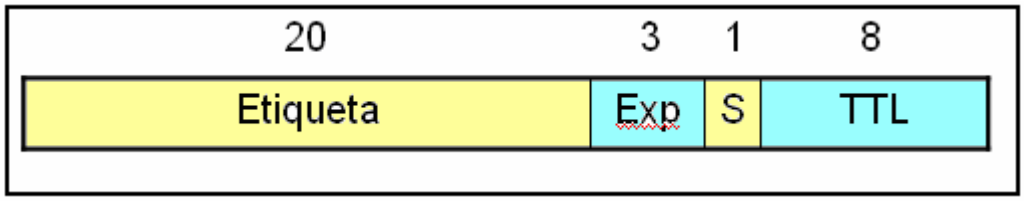


Figura. 5.6 Formato de la etiqueta MPLS

- Etiqueta: La etiqueta propiamente dicha que identifica una FEC (con significado local)
- Exp: Bits para uso experimental; una propuesta es transmitir en ellos información de DiffServ
- S: Apilamiento, vale 1 para la primera entrada en la pila (la más antigua), cero para el resto.
- TTL: Tiempo de vida, es el contador del número de saltos. Este campo reemplaza al TTL de la cabecera IP durante el viaje del datagrama por la red MPLS.

Los LSR pueden ser a su vez de varios tipos:

LSR Interior: El que encamina paquetes dentro de la red MPLS. Su misión es únicamente cambiar las etiquetas para cada FEC según le indica su LIB

LSR Frontera de ingreso: Los que se encuentran en la entrada del flujo a la red MPLS (al principio del LSP). Se encargan de clasificar los paquetes en FECs y poner las etiquetas correspondientes.

LSR Frontera de egreso: Los que se encuentran a la salida del flujo de la red MPLS (al final del LSP). Se encargan de eliminar del paquete la etiqueta MPLS, dejándolo tal como estaba al principio.

5.2.6.1 Familia Protocolar de MPLS

5.2.6.1.1 Distribución de Etiquetas LDP (Label Distribution Protocol)

En 2 routers de conmutación de etiquetas LSR, MPLS deben estar de acuerdo en el significado de las etiquetas para enviar el tráfico entre y a través de ellos. El Protocolo de Distribución de Etiqueta LDP es un nuevo protocolo que define un conjunto de procedimientos y mensajes por cuál uno router de conmutación de etiquetas LSR informa al otro de los enlaces de la etiqueta que ha hecho.

El router de conmutación de etiquetas LSR usa este protocolo para establecer la ruta de conmutación de etiqueta a través de una red trazada por la información de direccionamiento de la capa de red directamente para las rutas de conmutación de la capa de enlace de datos. Estas rutas de conmutación de etiquetas LSPs pueden tener como punto terminal en un vecino, como reenvió de conexión a conexión de IP, o puede tener como punto terminal a un nodo de salida de red, habilitando la conmutación mediante todos los nodos intermedios. Una Clase Equivalente de Reenvió FEC (Forwarding Equivalence Class) es asociado con cada ruta de conmutación de etiqueta LSP creado. Esta clase

equivalente de reenvió FEC especifica qué paquetes son direccionados a es ruta LSP.

Dos routers de conmutación de etiquetas LSRs que usan el protocolo de distribución de etiquetas LDP para intercambiar información de direccionamiento de etiquetas son conocidos como pares del protocolo LDP y ellos tienen una sesión de LDP entre ellos, en otras palabras, el protocolo es bidireccional.
Hay 4 clases de mensajes de LDP:

- Los mensajes de descubrimiento
- Los mensajes de sesión
- Los mensajes de anuncio
- Los mensajes de notificación.

Usando los mensajes de descubrimiento, los router LSRs anuncian su presencia en la red enviando mensajes de saludo periódicamente. Este mensaje de saludo es transmitido como un paquete de UDP. Cuando una nueva sesión debe establecerse, el mensaje de saludo se envía sobre TCP. Aparte del mensaje de descubrimiento, todos los otros mensajes se envían sobre TCP.

Los mensajes de notificación señalan errores y otros eventos de interés. Hay 2 tipos de mensajes de notificación:

- Las notificaciones del error: Estas señales de errores fatales y causa de la terminación de la sesión.
- Las notificaciones de advertencia: Estas son usadas para pasar información del router LSR sobre la sesión del protocolo LDP o el estado de algunos mensaje previamente recibidos desde el par.

Todos los mensajes del protocolo LDP tienen una estructura común que usa un esquema de codificación de Valor de Longitud de Tipo TLV (Type Lenght Value). Este valor de longitud de tipo TLV de codificación se usa para codificar mucha de

la información llevada en mensajes del protocolo LDP. La parte del valor de un objeto codificado con el valor TLV, puede contener uno o más TLVs.

Los mensajes son enviados como PDUs del protocolo LDP. Cada PDU puede contener más de un mensaje del protocolo LDP. Cada PDU es un encabezamiento de LDP seguido por uno o más mensajes de LDP:

La estructura del encabezamiento de LDP se muestra abajo:

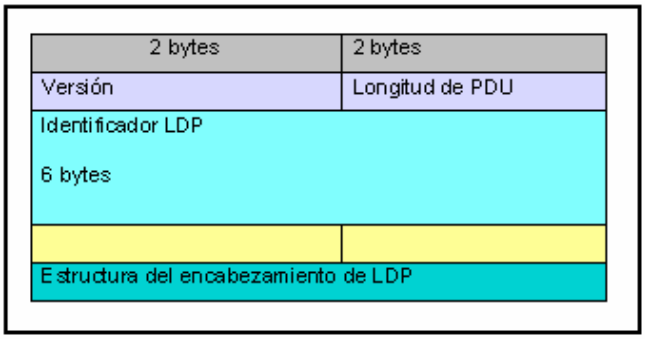


Figura. 5.7. Estructura del encabezamiento LDP

- **Versión:** El número de la versión protocolar.
- **Longitud de PDU:** La longitud total del PDU que excluye la versión y la longitud del campo de PDU.
- **Identificador de LDP:** Este campo únicamente identifica el espacio de la etiqueta del router LSR enviando para el que este PDU aplica. Los primeros 4 octetos codifican la dirección de IP asignada para el router LSR.

Mensajes de LDP

Todos los mensajes de LDP tienen el siguiente formato:

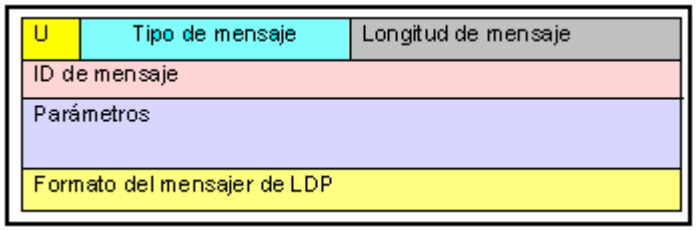


Figura. 5.8. Formato de los mensajes LDP

- **U:** El bit de U es un bit del mensaje desconocido.
- **Tipo del mensaje:** Los tipos del mensaje existentes son:

0x001	Notificación
0x100	Saludo
0x200	Inicialización
0x201	Mantener con vida
0x300	Dirección
0x301	Retirar dirección
0x400	Direccionamiento de etiqueta
0x401	Solicitud de etiqueta
0x404	Solicitud abortada de etiqueta
0x402	Retirar etiqueta
0x403	Terminación de etiqueta
Valor predefinido	Nombre de mensaje desconocidos

Tabla. 5.4. Mensajes de LDP

- **Longitud del mensaje:** La longitud en octetos del ID del mensaje, parámetros obligatorios y parámetros opcionales.
- **ID del mensaje:** Un valor de 32 bits identificaba el mensaje.
- **Parámetros:** Los parámetros contienen el TLVs. Hay parámetros obligatorios y opcionales. Algunos mensajes no tienen ningún parámetro obligatorio y algunos no tienen ningún parámetro opcional. El formato de TLV es:

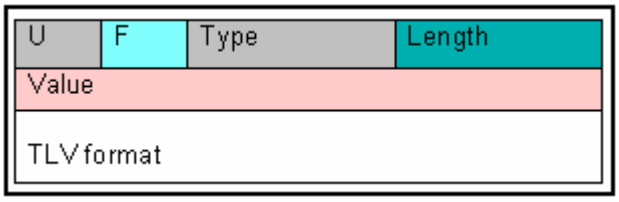


Figura. 5.9 Formato del TLV

- **U:** El bit de U es un bit de TLV desconocido.
- **F:** Envío de un bit de TLV desconocido.
- **Tipo:** Codifica cómo el campo de Valor será interpretado.
- **Longitud:** Especifica la longitud del campo de Valor en octetos.
- **Valor:** La cadena del octeto de octetos de longitud que codifica la información para ser interpretada como especificada por el campo del Tipo.

Los siguientes parámetros son opcionales de TLV:

100	Fec
101	Lista de dirección
103	Cuenta por conexión
104	Vector de ruta
200	Etiqueta Genérica
201	Etiqueta ATM
202	Etiqueta de Frame Relay
300	Estado
301	Estado extendido
302	PDU de regreso
303	Mensaje de regreso
400	Parámetros de saludo común
401	Dirección de transporte
402	Numero de secuencia configurado
500	Parámetros de sesión común
501	Parámetros de sesión de ATM
502	Parámetros de sesión de Frame Relay
600	ID de mensaje de solicitud de etiqueta

Tabla. 5. 5 Parámetros del TLV

5.2.6.1.2 Restricción basada en LDP CR-LDP (Constraint based LDP)

Contiene extensiones de LDP para extender sus capacidades. Esto permite extender la información usando el establecimiento de rutas más allá de lo que está disponible para el protocolo de la asignación de ruta.
CR-LDP es igual a LDP pero tiene los parámetros de TLV adicionales siguientes.

Valor	Parámetro
821	LSPID
822	ResCls
503	Parámetros de sesión optima
800	Direccionamiento explicito
801-804	ER-conexión TLVS
810	Parámetros de trafico
820	Prioridad
823	Direccionamiento de fijo
910	Tipo de interfase óptica
920	Descarte pista óptica
930	Etiqueta óptica
940	Establecimiento de llamada

Tabla. 5.6. Parámetros adicionales de TLV

5.2.6.1.3 Protocolo de la extensión de trafico RSVP-TE (Traffic Extension)

Es una adición al protocolo de RSVP con extensiones especiales que le permite establecer rutas ópticas en una red óptica ágil.

El protocolo de RSVP define una sesión como un flujo de datos con un destino particular y un protocolo de la capa de transporte. Sin embargo, cuando se combinan RSVP y MPLS, un flujo o sesión pueden definirse con mayor flexibilidad y generalidad. El nodo del ingreso de una ruta de conmutación de etiqueta LSP usa varios métodos para determinar qué paquetes son asignados a una etiqueta particular. Una vez una etiqueta es asignada a un conjunto de paquetes, la etiqueta define el flujo eficazmente a través del protocolo LSP. Nosotros nos referimos a semejante LSP como un túnel de LSP porque el tráfico a través de él es no transparente para nodos intermedios a lo largo de la ruta de conmutación de la etiqueta. Una nueva Sesión de RSVP, emisor y objetos de especificación del filtro, llamados IPv4 túnel de LSP y IPv6 túnel de LSP que ha sido definido para soportar la característica de túnel de LSP. La semántica de estos objetos, desde la perspectiva de un nodo a lo largo de la ruta de conmutación de etiqueta, es que el tráfico perteneciente al túnel de LSP es identificado solamente en base de los paquetes que llegan de la conexión anterior POP (previous hop) con un valor de etiqueta particular asignado por este nodo a los emisores del upstream para la sesión. De hecho, los IPv4 y IPv6 que aparece en el nombre del objeto solo denota que la dirección de destino es un a dirección IPv4 o IPv6. Al referirse genéricamente a estos objetos, el calificador de Túnel de LSP es usado.

En algunas aplicaciones es útil asociar conjuntos de túneles de LSP. Esto puede ser útil durante las operaciones de redireccionamiento o extendiendo una troncal de tráfico sobre rutas múltiples. En la aplicación de ingeniería de tráfico, son llaman túneles de ingeniería de trafico (TE Tunnels). Para habilitar la identificación y asociación de los túneles de LSP, dos identificadores son llevados. Un ID de túnel es parte del objeto de la Sesión. El objeto de la Sesión únicamente define un túnel de ingeniería de trafico. El emisor y objetos de la especificación de filtro llevan un ID de LSP. El objeto del emisor (o especificación de filtro), junto con el objeto de la Sesión, únicamente identifica un túnel de LSP.

CONCLUSIONES Y RECOMENDACIONES

- El despliegue de los servicios multimedia en tiempo real basados en la QoS sobre las redes IP es uno de los retos más excitantes al que los proveedores de servicios y de tecnología están actualmente enfrentándose, puesto que esto implica un paradigma tecnológico de desplazamiento en la integración en red de los paquetes. Uno de los principales retos técnicos es el mantener todo junto, es decir, integrar las diferentes tecnologías y las funciones de datos, control y gestión.
- En este proyecto de grado se han expuesto de forma resumida los mecanismos para obtener calidad de servicio QoS en comunicaciones con tráfico con requerimientos de tiempo real a través de redes IP.
- La evolución, desarrollo y aplicación de estas técnicas de señalización se encuentra evolucionando y en experimentación cada día.
- Se evidencia, a diferencia de las redes clásicas modo circuito, la necesidad de señalización para garantizar QoS, bien mediante IntServ, DiffServ o mediante una combinación de ambos.
- La falta de equipamiento y de terminales de cliente han obligado a retrasar el acondicionamiento físico-técnico, obstante, los protocolos y arquitecturas se encuentran en un estado de especificación muy avanzado, que permite tener una visión bastante nítida de cual será la evolución tecnológica y la naturaleza de los nuevos servicios integrados multimedia a los que dará soporte.
- Puede decirse que IP ya está preparado para dar respuesta a este nuevo entorno de convergencia y movilidad para el que no fue diseñado originalmente, con soluciones que ya están siendo desplegadas por los operadores, con protocolos nuevos, disponible en las nuevas versiones del sistema operativo de una gran cantidad de equipos de comunicaciones.
- Los mecanismos QoS proporcionan un servicio mejorado a usuarios de red, al mismo tiempo que permiten al administrador de la red administrar

recursos de red de forma eficaz, tanto en mecanismos de control del tráfico como mecanismos de provisión y configuración, para ofrecer de forma simultánea garantías de alta calidad y un uso eficaz de los recursos de red (producto de alta calidad/eficacia).

- Actualmente los principales centros de investigación están centrando sus esfuerzos en mejorar la arquitectura actual de Internet para que ofrezca QoS a aplicaciones cada más exigente de ancho de banda, bajo retraso (delay) y mínimo variación de retraso (jitter).
- La arquitectura de Servicios Diferenciados tiene como objetivo proporcionar una arquitectura que posibilite una discriminación de servicios escalables en Internet, y presenta mejores condiciones para ser implantados en redes con grandes volúmenes de tráfico, en lugar de la arquitectura de Servicios Integrados.
- Un nuevo enfoque más amplio del protocolo MPLS esta surgiendo, independiente de la red, la razón se debe a que es necesario mejorar la conmutación de los paquetes y mejorar los criterios de establecimiento del camino que debe seguir estos paquetes. No siempre el camino más corto es el más adecuado para ofrecer QoS a una aplicación.
- Un adecuado conocimiento de la terminología de comunicación e inglés técnico ayudarán a las personas interesadas en la información existente en Internet acerca de este tema.

BIBLIOGRAFIA

http://www.cisco.com, Fax Services, Session Initiation Protocol, MPLS
http://www.alcatel.com, QoS in the Internet
http://www.trilluium.com, H. 323
http://www.AligentTechologies.com , Tecnologia y Prueba de Fax
http://www.cs.columbia.edu/~hgs/internet, Internet Technical Resources
http://www.fokus.gmd.de/research/cc/glone/projects/ipt, Internet Telephony
http://www.etsi.org/tiphon European Telecommunications Standards Institute
http://itel.mit.edu, Internet & Telecomms Convergence Consortium
http://www.mplsforum.org/ MPLS Forum:
http://www.monografias.com, Protocolos de Señalización

INDICE DE FIGURAS

INDICE DE TABLAS

TERMINOLOGIA

ABREVIATURA	SIGNIFICADO
ACD	**Distribución de la Llamada Automática (Automatic Call Distribution)**
ACD*	**Distribuidores Automáticos de Llamadas (Automatic Call Distributor)**
ACF	**Confirmación de Admisión (Admisión Confirmation)**
ADPCM	**Modulación del Código de Pulso Diferencial Adaptativo (Adaptative Differential Pulse Code Modulation)**
AF	**Reenvío Seguro (Assured Forwarding)**
AFS	**Servicio de Reenvío Seguro (Assured Forwarding Service)**
APP	**Aplicación de Funciones Especificas (Application specific functions):**
ARQ	**Petición de Admisión (Admisión Request)**
AS-ACELP	**Estructura Algebraica Conjugada de Predicción Lineal Excitada de Código (Conjugate Structure Algebraic Codebook Excited Linear Prediction)**
ASN.1	**Notación de Sintaxis Abstracta 1 (Abstract Syntax Notation 1)**
ATM	**Modo de Transferencia Asincrónica (Asynchronous Transfer Mode)**
BGP	**Protocolo de Gateway de Borde (Border Gateway Protocol)**
CAC	**Control de Admisión de Llamada (Call Admission Control)**
CAC*	**Control de Admisión de Conexión (Connection Admission Control)**
CC	**Centros de Llamadas (Call Centers)**
CDR	**Grabación de detalles de Llamadas (Call Details Record)**
CED	**Tono de Respuesta de la Unidad Receptora (Responding Unit Answer Tone)**
CFR	**Confirmación para Recibir (Confirmation to Receiver)**
CIF	**Formato de Intermedio común (Common Intermediate Format)**
CLP	**Probabilidad de Pérdida de Celda (Cell Loss Probability)**
CNAME	**Nombre Canónico (Canonical Name)**
CNG	**Tono de Anuncio de la Unidad Llamada (Calling Unit Announce Tone)**
CoS	**Clase de Servicio (Class of Service)**
CPE	**Equipo de Premisas de Cliente (Client Premise Equipment)**
CQ	**Mecanismo de Cola de espera de Cliente (Custom Queuing)**
CRTP	**Protocolo de Transporte en Tiempo Real Comprimido (Compress Real Time Protocol)**
DCF	**Confirmación de Desconexión (Disengage Confirmation)**
DCN	**Mensaje de Desconexión (Disconnect Message)**
DCS	**Señal de Comando Digital (Digital Command Signal)**
DIS	**Señal de Identificación Digital (Digital Identification Signal)**
DNS	**Servicio de Nombres de Dominio (Domain Name Service**

dP	**Precedente de Eliminación (drop precedence)**
DRQ	**Solicitud de Desconexión (Disengage Request)**
DS	**Servicios Diferenciados (Differentiated Services)**
DSCP	**Punto de Código de Servicios Diferenciados (Differentiated Services Code Point)**
DSP	**Procesador de Señal Digital (Digital Signal Processor)**
DTMF	**Tono Dual de Multifrecuencia (Dual Tone Multi Frequency)**
ECN	**Notificación de Congestión Explicita (Explicit Congestion Notification)**
EF	**Reenvío Apresurado (Expedited Forwarding**
EFS	**Servicio de Reenvío Apresurado (Expedited Forwarding Service).**
EOM	**Señal de Paginas Múltiples (Multipage Signal)**
EOP	**Fin de Procedimiento (End of Procedure)**
FEC	**Clase Equivalente de Reenvió (Forwarding Equivalence Class)**
FEC*	**Corrección del Error de Envió (Forward Error Correction)**
FIFO	**Mecanismo de Cola de espera "Primero en entrar, primero en salir" (First-in, first-out)**
FoIP	**Fax sobre IP (Fax over IP)**
FTP	**Protocolo de Transferencia de Archivos (File Transfer Protocol)**
GSTN	**Red de Telefonía de Conmutación General (General Switching Telephony Network)**
HTML	**Lenguaje para Escribir Hiper Textos (Hyper Text Markup Language)**
HTTP	**Protocolo de Transferencia de Hipertexto (Hypertext Transfer Protocol)**
IP	**Protocolo de Internet (Internet Protocol)**
IPDC	**Protocolo de Internet Control de Dispositivo (Internet Protocol Device Control)**
IPSec	**Seguridad de IP (IP Security)**
IPT	**Telefonía IP (IP Telephony)**
IS	**Servicios Integrados (Integrated Services)**
ISDN	**Red Digital de Servicios Integrados (Integrated Services Data Network)**
ISND ISUP	**Parte de Usuario de (ISDN User Part)**
ISO	**Organización Internacional de Estandarización (International Organization for Standardization)**
ITSP	**Proveedores de Servicio Telefónico de Internet (Internet Telephony Service Provider)**
ITU-T	**Unión de Telecomunicación Internacional (International Telecommunication Union – Telecommunications Sector)**
LAN	**Redes de Area Local (Local Area Network)**
LD-CELD	**Bajo Retardo de Predicción Lineal Excitada de Código (Low Delay Codebook Excited Linear Prediction)**

LDP	**Protocolo de Distribución de Etiqueta (Label Distribution Protocol)**
LDP CR-LDP	**Restricción basada en (Constraint based LDP)**
LLQ	**Cola de Espera de Retardo Bajo (Low Latency Queuing)**
LSPs	**Rutas de conmutación de Etiquetas (Label Switched Paths)**
LSRs	**Router de conmutación de etiqueta (Label Switched Router)**
MC	**Controlador Multipunto (Multipoint Controller)**
MCF	**Mensaje de Confirmación (Message Confirmation)**
MCU	**Unidad de Control Multipunto (Multipoint Control Unit)**
MF	**Clasificador de Multicampo (Multi-Field Classifier)**
MG	**Gateway Media (Media Gateway)**
MGC	**Controlador de Gateway Media (Media Gateway Controller)**
MGCP	**Protocolo de Control de Gateway Media (Media Gateway Control Protocol)**
MH	**Modificada Huffman (Modified Huffman)**
MIME	**Extensión de Correo Electrónico de Internet Multiproposito (Multipurpose Internet Mail Extensions)**
MoIP	**Multimedia sobre IP (Multimedia over IP)**
MP	**Procesador Multipunto (Multipoint Processor)**
MPLS	**Conmutación de Etiquetas Multiprotocolo (Multiprotocol Label Switching)**
MQC	**Control de Cola de espera de Mensajes (Message Queuing Control)**
MTU	**Unidad de Transmisión Máxima (Maximum Transmission Unit)**
NNI	**Interfase de la Red a Red (Network to Network Interfase)**
OSI	**Interconexión de Sistemas Abiertos (Open Systems Interconnection)**
PBX	**Central Telefónica Privada (Private Branch Exchange)**
PCM	**Código de Pulso (Pulse Code Modulation)**
PHB	**Comportamiento por Salto (Per hop behavior)**
PPP	**Protocolo Punto a Punto (Point to Point Protocol)**
PQ	**Mecanismo de Cola de espera de Prioridad (Priority Queuing)**
PSTN	**Red de Telefonía de Conmutación Publica (Public Switched Telephone Network)**
QCIF	**Formato de Intermedio común Cuádruple (Quarter Common Intermediate Format)**
QoS	**Calidad de Servicio (Quality of Service)**
RAS	**Protocolo de Registro, Admisión y Estado (Registration Admission and Status)**
RIP	**Solicitud en Progreso (Request in Progress)**
RNR	**Receptor no Listo (Receiver Not Ready)**
RR	**Receptor Listo (Receiver Ready)**
RR*	**Reporte de Recepción (Receiver Report)**

RSVP	**Protocolo de Reserva de Recursos (Resource Reservation Protocol)**
RTN	**Reentrenamiento Negativo (Retrain Negative)**
RTP	**Protocolo de Transporte de Tiempo Real (Real Time Protocol)**
RTP*	**Reentrenamiento Positivo (Retrain Positive)**
RTP-HC	**Protocolo de Compresión del Encabezamiento en Tiempo Real (Real-Time Protocol Header Compression)**
RTSP	**Protocolo Streaming en Tiempo Real (Real Time Streaming Protocolo)**
SAP	**Protocolo de Anuncio de Sesión (Session Announcement)**
SCN	**Red de Circuito Conmutador (Switching Circuit Network)**
SDES	**Descripción de la Fuente (Source Description):**
SDP	**Protocolo de Descripción de Sesión (Session Description Protocol)**
SGCP	**Protocolo de Control de Gateway Simple (Simple Gateway Control Protocol)**
SIP	**Protocolo de Inicio de Sesión (Session Initiation Protocol)**
SLA	**Acuerdos de Nivel de Servicio (Service Level Agreements).**
SMTP	**Protocolo de Transferencia de Correo Electrónico Simple (Simple Mail Transfer Protocol)**
SR	**Reporte de Entrega (Send report):**
SS7	**Sistema de Señalización 7 (Signalling System Number 7)**
Sub-ADPCM	**Sub Modulación del Código de Pulso Diferencial Adaptativo (Sub-Band Adaptative Differential Pulse Code Modulation)**
Sub-QCIF	**Sub formato de Intermedio común cuádruple (Quarter Common Intermediate Format)**
TC	**Acondicionamiento de Tráfico (Traffic Conformation)**
TCA	**Acuerdo de Acondicionamiento de Tráfico (Traffic Conditioning Agreement)**
TCF	**Función de Verificación de Capacitación (Training Check Funtion)**
TCP	**Protocolo de Control de Transmisión (Transmisión Control Protocol)**
TCS	**Reconocimiento de Capacidades del Terminal (Terminal Capability SetAck)**
TIFF	**Formato de Archivo de Imagen de Etiqueta (Tag Image File Format)**
TLS	**Seguridad de Capa de Transporte (Transport Layer Security)**
TLV	**Valor de Longitud de Tipo (Type Lenght Value)**
ToS	**Tipo de Servicio (Type of Service)**
UA	**Usuario (User Agent)**
UAC	**Cliente de Agente de Usuario (User Agent Client)**
UAs	**Agentes del Usuario (User Agents)**
UAS	**Servidor de Agente de Usuario (User Agent Server)**
UDP	**Protocolo de Datagrama de Usuario (User Datagram Protocol)**
UNI	**Interfase de Usuario a Red (User toNetwork Interfase)**
URL	**Dirección Global de Documentos de Internet (Uniform Resource Locator)**
USB	**Bus de Serie Universal (Universal Serial Bus)**

VoIP	**Voz sobre IP (Voice over IP)**
WAN	**Red de Área Ancha (Wide Area Network)**
WFQ	**Cola de Espera de Carga Exacta (Weighted Fair Queuing)**
WRED	**Detección Temprana al Azar de Carga (Weighted Random Early Detection).**
WWW	**Red Ancha Mundial (World Wide Web)**

NOTA: **(*) Abreviación idéntica con significado diferente**