

RTP y RTCP

EFORT

<http://www.efort.com>

1 RTP y RTCP

Para transportar la voz o el vídeo sobre IP, se utilizan el protocolo IP (Internet Protocol) a nivel 3 y el protocolo UDP (User Datagram Protocol) a nivel 4. Es cierto que TCP (Transmission Control Protocol) presenta la ventaja de gestionar una transferencia fiable (reenvío de paquetes IP sin error), pero sin embargo es incompatible con un flujo en tiempo real, ya que los mecanismos de TCP previenen una reducción automática de la tasa por parte del emisor en caso de congestión de la red y un aumento lento de vuelta a la tasa nominal (para proteger la red de sobresaltos de emisión que pretendan utilizar el ancho de banda al mismo tiempo).

Pero estos dos protocolos UDP e IP no son suficientes para asegurar el transporte de la voz. De hecho, UDP es un protocolo sin corrección de errores, y en ningún momento se asegura la llegada de paquetes en su orden de emisión. Para el transporte de datos en tiempo real, como la voz o el vídeo, es necesario utilizar dos protocolos suplementarios: RTP (Real-Time Transport Protocol) y RTCP (RTP Control Protocol).

RTP y RTCP son dos protocolos que se sitúan a nivel de aplicación y se utilizan con el protocolo de transporte UDP. RTP y RTCP tanto pueden utilizar el modo Unicast (punto a punto) como el modo Multicast (multipoint).

RTP y RTCP utilizan puertos diferentes. RTP utiliza un número de puerto par, y RTCP el número de puerto impar que sigue a continuación. Cuando una sesión RTP es abierta, al mismo tiempo se abre una sesión RTCP implícita.

Los números de puerto utilizados por RTP y RTCP están comprendidos entre 1025 y 65535. Los puertos RTP y RTCP por defecto son respectivamente 5004 y 5005.

El párrafo 2 introduce el protocolo RTP. El párrafo 3 presenta las entidades Mixer y Translator que pueden ser utilizados en los intercambios de paquetes RTP. El párrafo 4 describe la estructura de paquetes RTP. Finalmente, el párrafo 5 describe en detalle el protocolo RTCP, su funcionamiento y los paquetes RTCP.

2 RTP (Real-time Transport Protocol)

La función de RTP es proporcionar un medio uniforme de transmisión de datos sometidos a limitaciones de tiempo real (audio, vídeo, etc.).

RTP permite:

- identificar el tipo de información transportada,
- añadir marcadores temporales que permitan indicar el instante de emisión del paquete. De esta forma, la aplicación destino podrá sincronizar los flujos y medir los retardos y la fluctuación.
- incluir números de secuencia a la información transportada para detectar la pérdida de paquetes y poder entregar los paquetes a la aplicación destino.

Además, RTP puede ser transportado por paquetes multicast para encaminar conversaciones hacia múltiples destinos.

No obstante, RTP no está concebido para realizar reservas de recursos o controlar la calidad de servicio, ni garantiza la entrega del paquete en recepción.

3 Mixer y Translator

El mixer (*mezclador*) es una aplicación que recibe flujos de datos de varias fuentes denominadas SSRCs (Synchronization Sources), modifica eventualmente el formato y reenvía un único flujo de datos agregado. Al mismo tiempo, en una aplicación de conferencia, probablemente es económico reagrupar, siempre que sea posible, las ondas sonoras que provengan de varias fuentes, ya que, de todas formas, serán mezcladas en el destino. Realizando esta tarea, el mixer se comporta como una fuente particular (SSRC) que ragrupa los datos de varias otras fuentes que estaban SSRCs y que se han convertido en CSRC (*Contributing Source*). Los paquetes emitidos por una fuente cualquiera (terminal, mixer) contienen una información de identificación de la SSRC que las emite. Estos paquetes pueden contener igualmente la identificación de CSRC que son fuentes reales de informaciones cuando se trata de un paquete construido por transformación.

En el ejemplo representado en la figura 1, un mixer recibe paquetes de tres fuentes (SSRC). Puede realizar conversiones de formato, mezcla el contenido en función de la aplicación y produce un nuevo paquete RTP que retransmite a su destino.

El mixer provee la sincronización para el nuevo flujo y se identifica como la nueva fuente SSRC.

Las tres fuentes de origen son declaradas como CSRC en el paquete RTP compuesto.

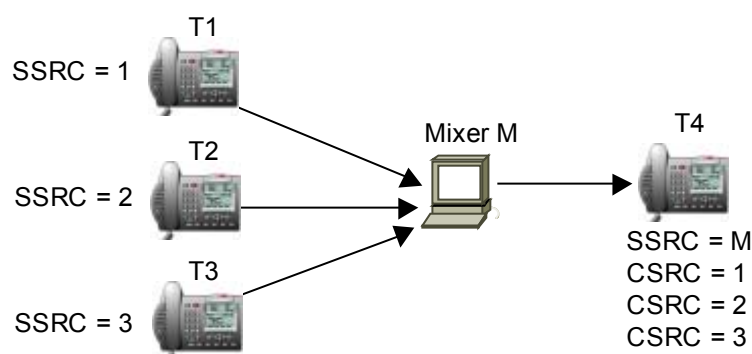


Figura 1: Mixer RTP

El traductor (*traductor*) es una aplicación que transmite paquetes RTP que recibe sin cambiar el identificador de SSRC (al revés que el mezclador). Un traductor permite, por ejemplo, cambiar la codificación de un dato o la tasa, o también tratar problemas de seguridad (*firewalls*) en la frontera de una red privada.

Si un terminal que utiliza una codificación PCM μ -law a 64 kbit/s desea establecer una sesión con un terminal que soporta una codificación G.726 ADPCM a 32 kbit/s, entonces es necesario interconectar los dos terminales a través de un traductor.

Las funciones Mixer y Translator se establecen generalmente en un Gateway.

4 Los datos RTP

RTP transporta las señales audio o vídeo codificadas mediante paquetes RTP que contienen un header RTP (cabecera) seguido de estas señales audio o vídeo

Un paquete RTP pasa por la capa UDP, que le añade una cabecera UDP. El conjunto es traspasado a la capa IP, que agrega una cabecera IP. Entonces, el datagrama IP es encaminado hacia el destino. En recepción, el paquete es entregado a la aplicación adecuada.

4.1 Codificación de señales

Es esencial, teniendo en cuenta del importante número de normas de codificación de señales audio o vídeo, incluir un mecanismo a RTP para permitir al destino conocer la codificación utilizada y así poder decodificar correctamente. RTP realiza esta función mediante el número de tipo de contenido (payload type number) en el header (cabecera) RTP. Los números de tipo de contenido están especificados en el RFC 3551 (RTP Profile Audio and video Conferences with Minimal Control) y se encuentran en la lista de la tabla 1.

Tipo Payload	Codec	Frecuencia (Hz)	Descripción
0	PCMU	8000	ITU G.711 PCM μ -Law audio 64 kbit/s
1	1016	8000	CELP Audio 4.8 kbit/s
2	G721	8000	ITU G.721 ADPCM Audio 32 kbit/s
3	GSM	8000	European GSM Audio 13 kbit/s
5	DVI4	8000	DVI ADPCM Audio 32kbit/s
6	DVI4	16000	DVI ADPCM Audio 64kbit/s
7	LPC	8000	Experimental LPC Audio
8	PCMA	8000	ITU G.711 PCM A-Law audio 64 kbit/s
9	G722	8000	ITU G.722 Audio
10	L16	44100	Linear 16 bit Audio 705,6 kbit/s
11	L16	44100	Linear 16 bit Stereo Audio 1411.2 kbit/s
14	MPA	90000	MPEG-I ou MPEG-II Audio
15	G728	8000	ITU G.728 Audio 16kbit/s
25	CELB	90000	CelB Video
26	JBEG	90000	JBEG Video
28	NV	90000	nv Video
31	H261	90000	ITU H.261 Video
32	MPV	90000	MPEG-I et MPEG-II Video
33	MP2T	90000	MPEG-II transport stream Video

Tabla 1: Los tipos de payload para la codificación de señales audio y vídeo

4.2 Codificación PCM

El código del sonido con una calidad de tipo telefónico está basado en el PCM μ -law o, en ocasiones, a-law. La norma μ -law era la que estaba en vigor en la red telefónica americana, y la norma a-law en la red telefónica europea. En esta norma, el sonido está codificado en 14 bits con una frecuencia de 8000 Hertz, ya que cada muestra es convertida a un tamaño de 8 bits mediante una tabla de traducción. Esta es la diferencia entre las tablas de traducción que distingue los dos códigos μ -law y a-law. Esta frecuencia de muestreo proporciona una frecuencia audible que cubre la banda entre 300 Hertz y 3500 Hertz. Con muestras de 8 bits, el flujo generado llega a una tasa de 64 kbit/s.

4.3 Codificación ADPCM

Una compresión ADPCM por codificación diferencial ha sido propuesta para reducir la tasa generada. Ésta pasa a 32 kbit/s. Esta compresión altera muy débilmente la calidad sonora, y como máximo limita la dinámica de los agudos.

4.4 Formato del paquete RTP

La cabecera de un paquete RTP (RTP header) está obligatoriamente constituida de 12 octetos, y eventualmente seguida de una lista de identificadores de fuentes contributivas CSRCs en el caso de un mezclador. Esta cabecera precede el « payload » que representa los datos útiles (Figura 2).

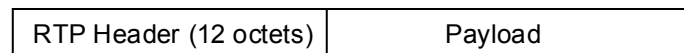


Figura 2: Formato del paquete RTP

- Version (V) (2 bits): Este campo indica el número de versión RTP utilizada. La versión actual es la versión 2.
- Padding (P) (1 bit): Si vale 1, este campo significa que el campo de datos (payload) tiene una parte de relleno. Recordemos que la longitud de datos debe ser un múltiple de 4 octetos, lo que hace necesarios los octetos de relleno, sobretodo para el último paquete. El último octeto de esta parte de relleno indica el número de octetos de relleno que se deben ignorar.
- Extension (X) (1 bit): Si vale 1, este campo indica que la cabecera fija tiene una parte de cabecera suplementaria.
- CSRC Count (CC) (4 bits): Este campo contiene el número de identificantes CSRC que siguen la cabecera fija, es decir, el número de fuentes contributivas ligadas a este paquete.
- Marker (M) (1 bit): Se trata de un bit de señalización. Su significado depende de los datos transportados.
- Payload type (PT) (7 bits): Este campo identifica el tipo de contenido (audio, vídeo, etc.) que representa el tipo de codificación de información transportada en el paquete. La lista se presenta en la tabla 8.1.
- Sequence Number (16 bits): El valor de este campo se incrementa en 1 por cada paquete RTP enviado, pero su valor inicial es aleatorio. Este campo permite detectar paquetes RTP perdidos.
- TimeStamp (32 bits): Un protocolo como RTP utiliza marcas temporales para datar los paquetes emitidos. Estas informaciones son la base de los cálculos que permiten evaluar el retardo y la fluctuación introducidos por un sistema de comunicación. La pertinencia de estas informaciones depende enteramente de la precisión y de la sincronización de los relojes utilizados en las maquinas que van a basar sus cálculos en estos valores.
- Synchronization Source (SSRC) (32 bits): Este campo identifica la fuente que ha producido el paquete. Al principio de una sesión, cada participante escoge un número de SSRC. En este caso se habla de sincronización ya que la escala de tiempo establecida por la fuente en estos paquetes va a servir de referencia a los receptores para restituir la información correctamente.
- Contributing source (CSRC): De 0 a 15 instancias de este campo pueden estar presentes en la cabecera del paquete RTP. El número está indicado por el campo CC. Cuando un flujo RTP es el resultado de una agregación por un mezclador RTP, la lista de SSRCs que han aportado su contribución es añadida en la cabecera de los paquetes RTP del flujo resultante (Figura 3).

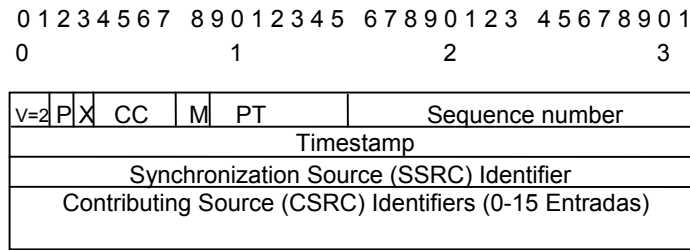


Figura 3: Cabecera RTP

5 RTCP (RTP Control Protocol)

El protocolo RTCP está basado en transmisiones periódicas de paquetes de control para todos los participantes en la sesión.

Es un protocolo de control de los flujos RTP, que permite transportar informaciones básicas de los participantes de una sesión y de la calidad de servicio.

Existen cinco tipos distintos de paquetes RTCP para cada tipo de información:

- SR (Sender Report) contiene las estadísticas de transmisión y de recepción para los participantes que son emisores activos.
- RR (Receiver Report) contiene estadísticas de recepción para los participantes que no son emisores activos pero sí receptores de una sesión.
- SDES (Source Description) describe la fuente: nombre, email, teléfono, etc.
- BYE permite a una estación indicar el fin de su participación en una sesión.
- APP es un paquete de señalización específico a una aplicación.

El control de flujo RTP se realiza guardando una evaluación del número de participantes en una sesión (fuentes y receptores). A partir de esta evaluación, se calcula un intervalo de tiempo que sirve de periodo de recurrencia en la difusión de informaciones SR o RR, según el caso. Globalmente, los algoritmos de control limitan el volumen de las informaciones de control transmitidas (los datos RTCP, por lo tanto) a un 5% del volumen global de los intercambios de la sesión. En este volumen, el 25% está reservado a las informaciones de las fuentes (mensajes SR). De esta forma, se garantiza una posibilidad de gestionar grupos de gran tamaño del punto de vista del volumen de información intercambiada. Cuanto más elevado es el número de participantes, menos precisa es la visión que tiene cada participante del estado de la red.

Si se considera una sesión audio con dos participantes, los paquetes RTCP pueden ser emitidos cada 5 segundos, mientras que para 4 participantes son emitidos cada 10 segundos. Los paquetes que se transmiten con más frecuencia son SR y RR.

5.1 Paquete RTCP Sender Report (SR)

Los participantes en una sesión que a la vez emiten y reciben paquetes RTP, utilizan paquetes RTCP SR. El formato de este paquete está presentado en la figura 4.

El paquete SR contiene un header (cabecera), informaciones sobre el emisor, un cierto número de bloques de informes de recepción y opcionalmente una extensión específica al perfil.

La estructura es similar a la de un paquete RTP.

La cabecera contiene los campos siguientes:

- Version (V) (2 bits): Este campo indica el número de versión RTCP. El valor de la versión actual del protocolo RTCP es 2 (10).
- Padding (P) (1 bit): Cuando vale 1, este campo indica que hay un relleno cuyo tamaño es indicado por el último octeto.

- Reception report count (RC) (5 bits): Este campo precisa el número de informes de recepción contenidos en el paquete SR, considerando un informe para cada fuente. Por lo tanto, se pueden incluir hasta un máximo de 31 informes en el paquete SR.
- Packet type (PT) (8 bits): Este campo indica el tipo de paquete; se trata de un paquete SR, representado por el valor 200.
- Length (16 bits): Este campo indica la longitud total del paquete en palabras de 32 bits (cabecera y relleno comprendidos).

Las informaciones sobre el emisor consisten en los campos siguientes:

- SSRC of sender (32 bits): Este campo precisa la identificación de la fuente específica del emisor.
- NTP timestamp (64 bits): La representación del tiempo utilizado por NTP (Network Time Protocol) es bastante simple: una fecha está codificada en 64 bits y medida en segundos desde las 0h del 1 de enero de 1900. La parte entera de esta fecha expresada en segundos está codificada en 32 bits de peso fuerte (palabra más significativa), y la parte fraccionaria (fracción de segundo) en los 32 bits de peso débil (palabra menos significativa). Esta representación garantiza una precisión de aproximadamente 200 pico-segundos, probablemente suficiente para la mayor parte de aplicaciones. El problema de **acercamiento a cero** de esta representación llegará en el 2036. Entonces, será necesario definir una nueva versión del protocolo antes de esta fecha.

RTP timestamp (32 bits): Este campo indica el mismo tiempo que el que indica NTP Timestamp precedente, pero utilizando las mismas unidades que las utilizadas para especificar el valor timestamp en los paquetes RTP.

- Sender's packet count (32 bits): Este campo indica el número total de paquetes RTP transmitido por el emisor desde el inicio de la sesión. Está reinicializado en el contexto de una sesión si el emisor cambia de identificador SSRC.
- Sender's octet count (32 bits): Este campo indica el número total de octetos RTP (sólo son considerados los octetos de datos de usuario y no los octetos de cabecera o de relleno) transmitidos por el emisor desde el principio de la sesión. También se reinicializa si el emisor cambia de identificador de SSRC.

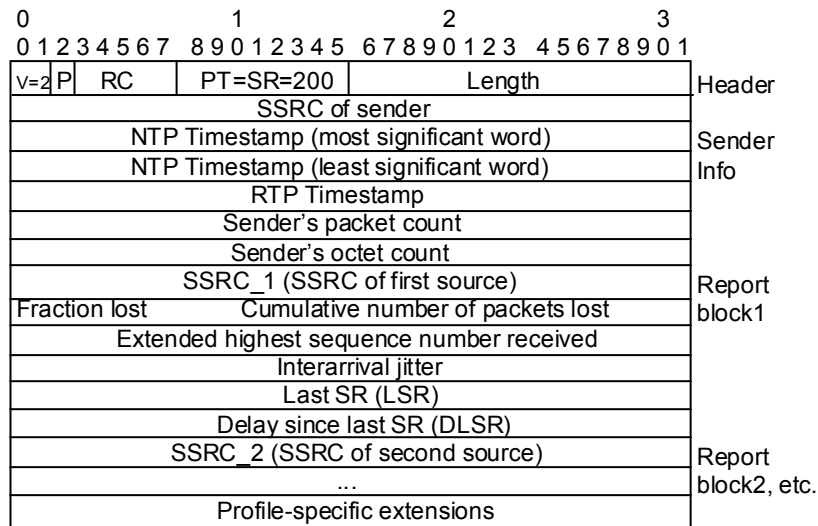


Figura 4: Paquete RTCP Sender Report

Uno o varios bloques de informes de recepción (RR, Receiver Report) siguen la información del emisor. Proporcionan a los otros participantes de la sesión la información referente al número de paquetes RTP que han sido emitidos o recibidos con éxito por el emisor del paquete SR.

Los campos siguientes están incluidos en cada bloque RR (Figura 5).

- SSRC_n (32 bits): Este campo especifica la identificación de la fuente en la sesión, que se refiere a los datos incluidos en el bloque RR.
- Fraction lost (8 bits): Este campo indica la fracción de paquetes RTP perdidos desde el último informe emitido por este participante. La fracción representa la relación entre el número de paquetes perdidos y el número de paquetes esperados. El número de paquetes perdidos puede ser deducido a partir del análisis del número de secuencia (Sequence Number) de cada paquete RTP recibido.
- Cumulative number of packets lost (24 bits): Este campo indica el número total de paquetes RTP de la fuente en cuestión, que han sido perdidos desde el principio de la sesión RTP.
- Extended highest sequence number received (32 bits): Este campo especifica el número de secuencia del último paquete RTP recibido desde esta fuente SSRC_n.
- Interarrival jitter (32 bits): Este campo informa de la variación del retardo de transmisión de los paquetes RTP.
- Last SR Timestamp (LSR) (32 bits): Este campo representa los 32 bits del medio del campo NTP Timestamp utilizado en el primer paquete SR recibido desde la fuente en cuestión. De esta forma, considera los 16 bits de peso débil de la parte entera de esta fecha (segundos) y los 16 bits de peso fuerte de la parte fraccionaria (fracciones de segundo). Si ningún paquete RTCP SR aun no ha sido recibido, entonces el valor de este campo es igual a 0.
- Delay Since Last SR (DLSR) (32 bits): Este campo representa el retardo expresado en unidades de 1/65536 segundos entre el instante de recepción del último paquete SR de la fuente SSRC_n y el instante de emisión de este bloque RR. Si ningún paquete SR aun no ha sido recibido de la fuente SSRC_n, el valor del campo DLSR se sitúa a 0.

5.2 Paquete RTCP Receiver Report (RR)

El paquete RCTP RR (Receiver Report) es emitido por un participante en una sesión que recibe paquetes RTP pero que no emite. El formato del paquete está representado en la figura 5. Tiene una estructura similar al paquete RTCP SR, pero indica el valor 201 para el campo payload type y no incluye información específica del emisor.

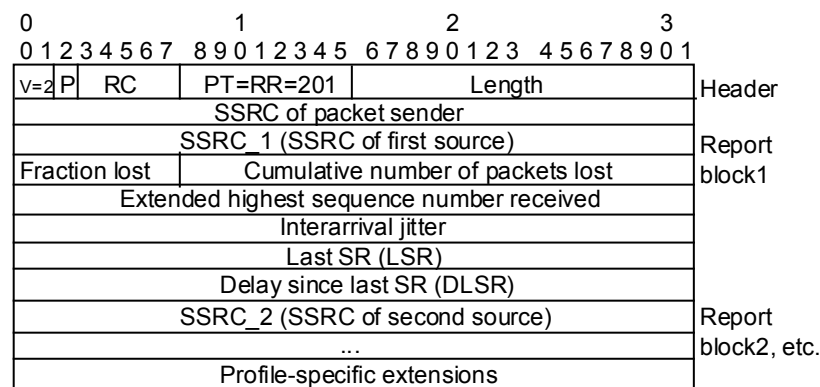


Figura 5: Paquete RTCP Receiver Report

5.3 Paquete RTCP Source Description (SDES)

El paquete RTCP SDES (Source Description) permite identificar los participantes en una sesión y proporcionar informaciones sobre los participantes. La estructura del paquete SDES está compuesta de un header (cabecera) y de ningún, uno o varios chunks, que describen cada una de las fuentes identificada en el chunk. (Figura 6).

La cabecera está formada por:

- Los campos Version Padding y Length, que tienen el mismo significado que los del paquete RTCP SR.
- El campo Packet Type (PT) en 8 bits, que indica un paquete SDES, representado por el valor 202.
- El campo Source Count (SC) en 5 bits que indica el número de chunks SSRC/CSRC contenidos en el paquete SDES.

Cada Chunk contiene un valor de SSRC o de CSRC seguido de uno o varios identificadores y de informaciones que describen este SSRC o CSRC. Estas informaciones se denominan "SDES ítems" y pueden incluir datos como el nombre, la dirección Email, el número de teléfono, la localización del usuario, etc.

Los terminales deben enviar un paquete SDES al principio de la sesión para que cada participante sea explícitamente identificado. El mixer combina los paquetes SDES de distintos participantes y produce un paquete compuesto de tantos Chunks como paquetes SDES de origen.

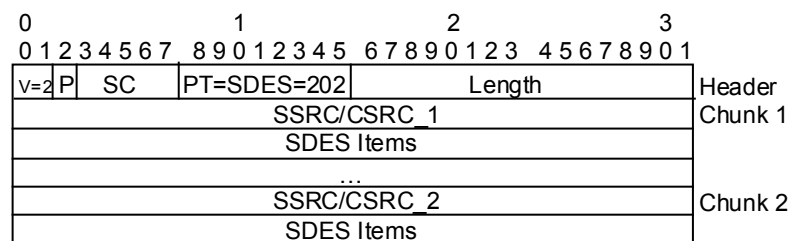


Figura 6: Paquete RTCP Source Description

5.4 Paquete RTCP BYE

El paquete RTCP BYE permite indicar que una fuente (un participante en la sesión) ya no está activa.

Los campos Versión, Padding y Length tienen el mismo significado que los del paquete RTCP SR.

El campo Packet Type (PT) indica el tipo de paquete RTCP. Se trata de un paquete BYE, representado por el valor 203.

El campo Source Count (SC) indica el número de identificadores SSRC/CSRC incluidos en este paquete.

Si un paquete BYE es recibido por un mixer, lo retransmite con los identificadores SSRC/CSRC intercambiados. Si por el contrario, el mixer debe dejar de funcionar, entonces emitirá un paquete BYE con una lista de todas las fuentes contributivas de las que se encarga, así como su propio identificador SSRC.

Finalmente, el campo opcional Reason for leaving indica, bajo la forma de una cadena de caracteres, el motivo de envío del paquete BYE. Este campo es precedido por su tamaño.

A diferencia del paquete SDES, no es posible mezclar varios paquetes BYE en un solo paquete BYE compuesto.

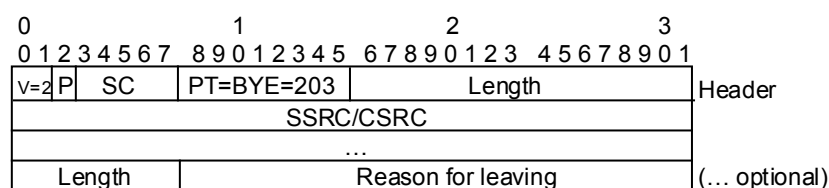


Figura 7: Paquete RTCP BYE

5.5 Paquete RTCP APP

El paquete RTCP APP es un paquete de señalización específico de las aplicaciones (Figura 8).

Los campos Versión, Padding y Length tienen el mismo significado que los del paquete RTCP SR:

El campo subtype (5 bits) puede ser utilizado para definir un sub-tipo que identifica un conjunto de paquetes APP.

El campo name afecta a un nombre único (4 caracteres) para el sub-tipo.

Los datos que dependen de la aplicación están incluidos en el campo "Application-dependent data".

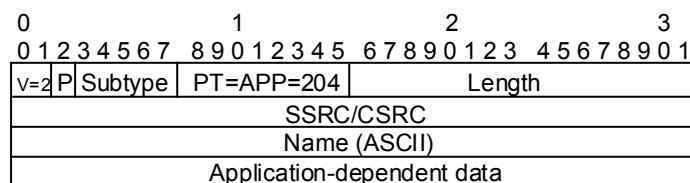


Figura 8: Paquete RTCP APP

6 Conclusión

Los protocolos RTP y RTCP están adaptados para la transmisión de datos en tiempo real. No obstante, funcionan con una estrategia de extremo a extremo y, por tanto, no pueden controlar el elemento principal de la comunicación: la red. Sin embargo, cualesquiera sean los esfuerzos de adaptación de los emisores, o los medios establecidos por los receptores, es en el centro de la red donde se generan los disfuncionamientos críticos. El protocolo Internet ha sido expresamente pensado para aportar inteligencia en los sistemas de extremidad. Es esta simplicidad que ha conducido al éxito de Internet. El protocolo RSVP (Resource Reservation Protocol) definido por el IETF ha sido desarrollado para remediar estas disfuncionalidades y así mejorar las transmisiones en tiempo real.

Los protocolos RTP y RTCP se utilizan principalmente para videoconferencia, donde los participantes son a su turno, emisores o receptores. Para el transporte de voz, permiten una transmisión correcta en las redes adecuadas. Es decir, redes que implementen una calidad de servicio adaptada (Redes IP con MPLS y DiffServ). También es posible basarse en redes bien dimensionadas (ancho de banda, determinismo de las capas sub-adyacentes, etc.), de tipo LAN de empresa.

Referencias

RFC 3550, H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", Julio 2003.