

4 CAPA DE SISTEMA

Como hemos dicho anteriormente, esta capa del estándar MPEG-2 está definida en la parte de la ISO/IEC 13818-1. En ella, se indica la forma de combinar uno o más flujos elementales (*elementary streams* o ES). Un ES es un flujo continuo de datos (audio, vídeo u otro tipo) que contiene toda la información de cada tipo de una sola fuente. Estos ES están formados por Unidades de Acceso, que en el caso de la señal de vídeo serán de tres tipos. Vemos un ejemplo de esto en la figura 4.1.

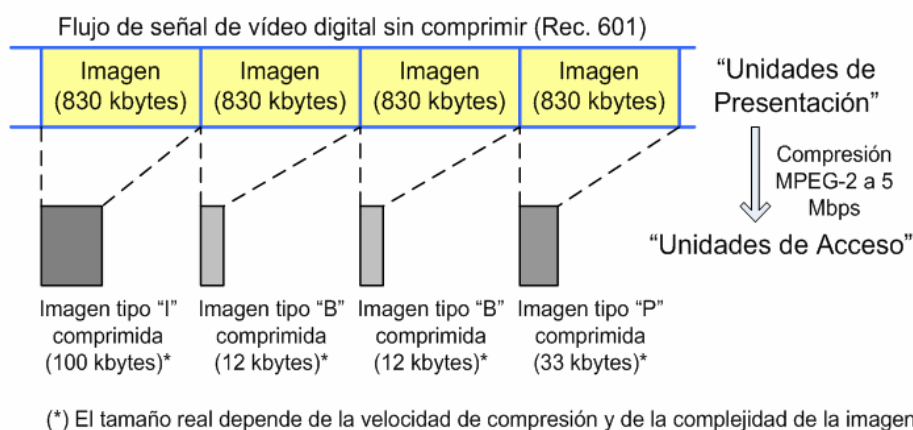


Figura 4.1 Compresión MPEG-2 de la señal de vídeo digital

El hecho de ser flujos continuos no facilita la transmisión y almacenamiento, por lo que se debe definir cómo se procesan los ES para conseguir estructuras más adecuadas, y la sintaxis de estas estructuras. De esta forma podemos dividir los ES y la capa de sistema nos indicará la relación entre ellos y la fuente a la que pertenecen. Además se necesita un sistema de temporización que permita al decodificador proceder correctamente.

4.1 Multiplexaje y sintaxis

Como hemos dicho, en aplicaciones prácticas, los flujos continuos ES que portan la información codificada de audio y vídeo deben romperse en paquetes o estructuras adecuadas. Estas unidades son los en PES *packets*, que se multiplexan en un flujo conjunto llamado flujo paquetizado elemental (*Packetized Elementary Stream* o PES combinando los distintos tipos de datos. Además se les añade una cabecera con

información adicional como podemos ver en la figura 4.2. La carga útil del PES *packet* es una sucesión de bytes de datos tomados secuencialmente del ES original. No hay ningún requerimiento para alinear el comienzo de una Unidad de Acceso y el comienzo de la carga de un PES *packet*. De esta forma, una nueva Unidad de Acceso puede comenzar en cualquier punto de la carga del PES *packet*, y también es posible que varias Unidades de Acceso estén contenidas en el mismo PES *packet*.

La cabecera tiene longitud variable. Podemos destacar algunos de sus campos como son:

- **Stream_id.** Es un byte que identifica el PES *packet*. Hay unos rangos de valores para audio y otros para vídeo, además de los correspondientes a otros tipos de información.
- **Time Stamps.** Son marcas de tiempo que sirven para asegurar la correcta sincronización. Son valores binarios de 33 bits expresados en unidades de 90 kHz. De este campo se hablará de forma más detallada en el apartado 4.4.
- **PES header length.** Indica en número de bytes opcionales hasta alcanzar la carga útil.

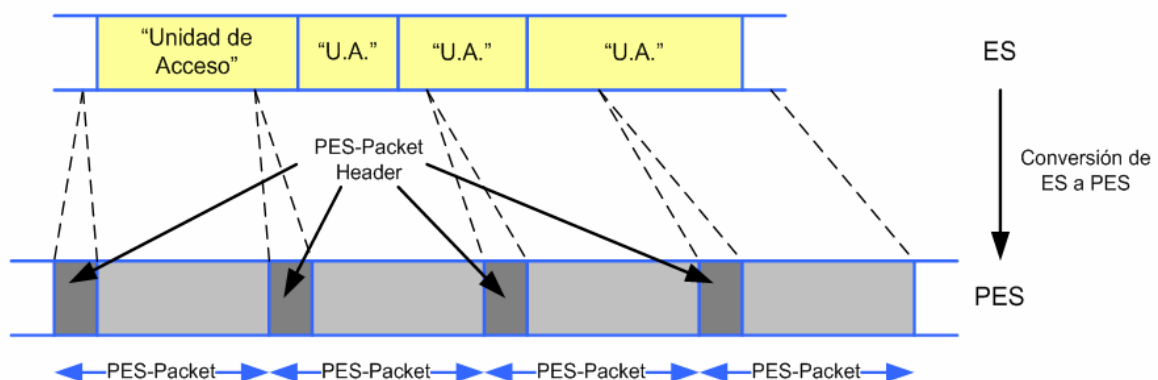


Figura 4.2 Conversión de ES a PES

El estándar especifica dos formas de encapsulado que generan dos tipos diferentes de flujos de datos a nivel de sistema, que son el Flujo de Transporte (*Transport Stream* o TS) y Flujo de Programa (*Program Stream* o PS). Cada tipo es óptimo en diferentes clases de aplicaciones. El primero (TS) se usa en medios de transmisión no fiables en los que pueden perderse paquetes, mientras que el segundo (PS), se usa para medios de

almacenamiento digitales o medios fiables y es muy similar a la capa de sistema de MPEG-1.

En la figura 4.3 podemos ver un esquema general del proceso de generación y decodificación de estos dos tipos de flujos. El vídeo y el audio son codificados según lo especificado en la capa de vídeo y la capa de audio, respectivamente, definidas en el estándar en sus correspondientes partes. Los ES resultantes se dividen y empaquetan, añadiendo cabeceras, para producir PES. Estas unidades básicas de la capa de sistema se combinan con información del sistema para formar los flujos de programa (PS) y de transporte (TS). La forma de combinar y dividir los PES depende del flujo del que vayan a formar parte.

En el caso de Flujos de Programa, se multiplexan todos los componentes del programa (vídeo, audio, datos, etc) y se incorpora el reloj del sistema, pero solamente se transmite la información correspondiente a un único programa audiovisual. En cambio, en el caso de Flujos de Transporte existe la posibilidad de multiplexar varios programas audiovisuales.

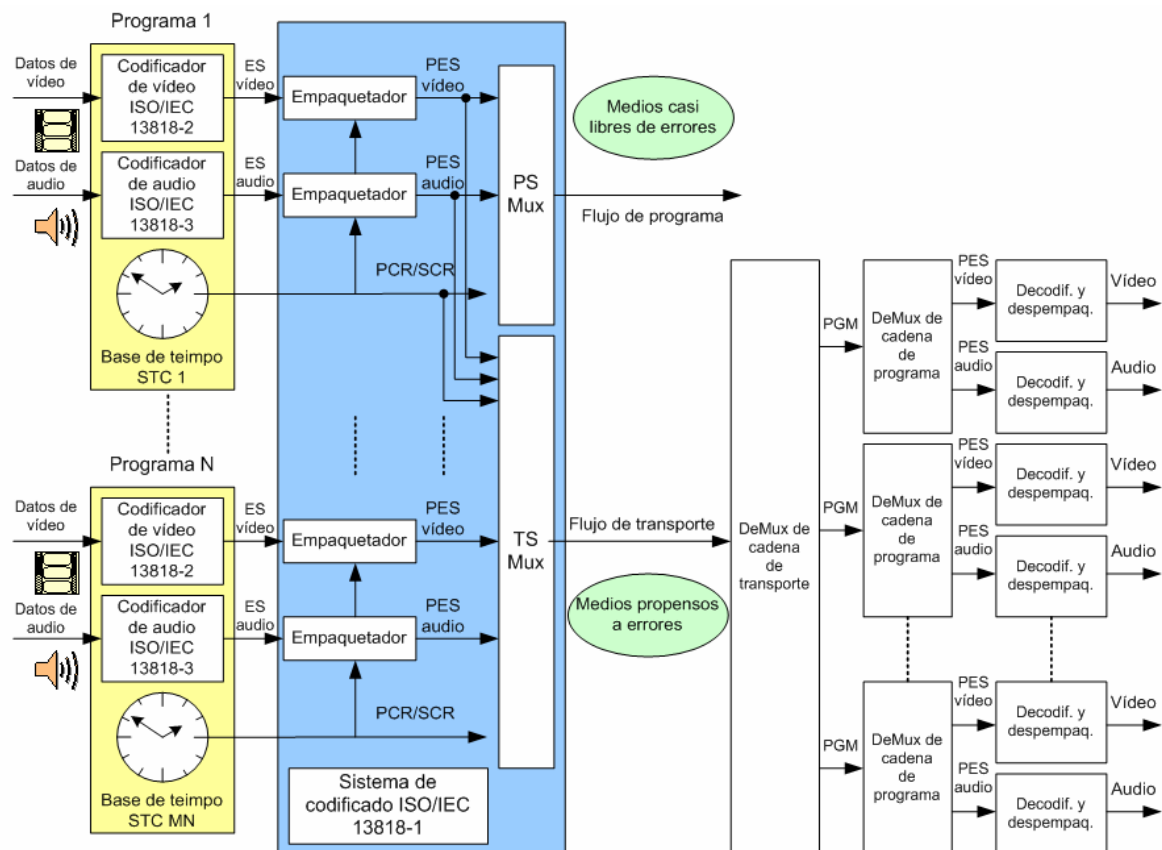


Figura 4.3 Esquema del proceso de generación y decodificación del PS y TS

Una vez en el receptor, el primer paso es extraer la información de sistema para poder identificar los datos recibidos y generar los distintos ES. Una vez regenerados, se almacenan en buffers, donde esperan para ser representados en el instante indicado por el sistema de temporización.

Las estructuras del PS y del TS no siguen estrictamente un modelo por capas, pero aun así es posible y razonable convertir uno en el otro, gracias al uso común de los PES. Todos los campos necesarios en un PS, no están contenidos en un TS ni viceversa y por lo tanto deben ser derivados durante el proceso de conversión.

4.2 Program Stream

El *Program Stream* (PS) es similar al flujo MPEG-1, pero utiliza una nueva sintaxis para poder añadir nuevas funcionalidades a la vez que mantiene la compatibilidad con MPEG-1 (un decodificador MPEG-2 debe poder decodificar un flujo MPEG-1). El PS contiene el equivalente a un solo programa audiovisual y es el resultado de multiplexar uno o más ES con un base de tiempo común.

Estructuralmente, un PS está formado por “*packs*”, que a su vez comprenden una cabecera del *pack* o “*pack-header*”, opcionalmente una cabecera de sistema o “*system-header*”, y un número indeterminado de PES *packets* (de longitud variable y relativamente grandes), tomados de los ES que componen el programa audiovisual. El PS puede contener hasta 32 flujos de audio, 16 flujos de video, 16 flujos de datos y varios flujos con información variada para el funcionamiento interno. Estas limitaciones vienen impuestas por el campo de *Stream_id* de la cabecera de los PES *packets*. Un esquema de la formación del PS está representado en la figura 4.4.

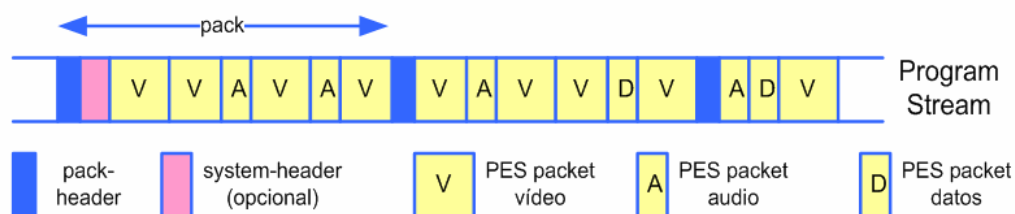


Figura 4.4 Estructura del Program Stream

No existe una especificación acerca de la longitud de cada *pack*, pero existe la restricción de que debe aparecer un *pack-header* al menos cada 0,7 segundos, puesto que esta cabecera contiene importante información para la sincronización denominada Referencia del Reloj de Sistema (*System Clock Reference* o SCR).

En el campo opcional *system-header* de un *pack* se incluye un sumario de las características del PS como por ejemplo la velocidad binaria máxima, en número de Es de vídeo y audio que lo componen, información complementaria de temporización.

Además de los PES *packets* de audio y vídeo se deben destacar dos tipos de paquetes importantes que aparecen en el PS que son el *Program Stream Map* y el *Program Stream Directory*, que establecen la descripción y relación entre cadenas elementales que transporta cada paquete así como un directorio de dónde aparecen las imágenes I en la secuencia.

El PS está concebido para su empleo en entornos libres de errores, puesto que es bastante vulnerable a ellos. Hay dos razones para esto: En primer lugar, el PS comprende una sucesión de paquetes relativamente largos y de longitud variable. Como se ha visto, cada paquete comienza con unas cabeceras que contienen información importante para su recuperación, por lo que cualquier error en dichas cabeceras puede provocar la pérdida entera del paquete. A su vez, como el paquete de un PS puede contener muchos kbytes de datos, la pérdida de un solo paquete puede representar la pérdida o corrupción de una trama de vídeo completa.

En segundo lugar, la diferente longitud de los paquetes implica que un decodificador no puede predecir donde termina un paquete y comienza otro, por lo que se ve obligado a leer e interpretar el campo incluido en la cabecera que informa de la longitud del paquete. Si este campo que identifica la longitud del paquete se corrompe por un error, el decodificador perderá el sincronismo con el flujo de datos, resultando como mínimo la pérdida de un paquete.

Por ejemplo se puede usar para aplicaciones basadas en software y entornos libres de errores, como aplicaciones multimedia sobre CD-ROM. La longitud de los paquetes está normalmente entre un margen de 1 a 2 Kbytes para poder encajar en los sectores de los discos (típicamente 2 Kbytes). Sea como sea, los paquetes pueden ser de una longitud máxima de 64 Kbytes.

El PS mantiene características que ya incorporaba el MPEG-1, como el control de los buffers del decodificador, información de temporización, campos para asistir al análisis

del PS después de un acceso aleatorio y campos para permitir la sincronización de los ES. Pero además incluye características que no son soportadas por el MPEG-1, tales como la encriptación de datos, la asignación de diferentes niveles de prioridad a los paquetes, información para asistir al alineamiento de los paquetes, indicación de copyright, indicación de *fast forward* (avance rápido), *fast reverse* (retroceso rápido), un campo opcional para testear el funcionamiento de la red y una numeración opcional de las secuencias de paquetes.

4.3 Transport Stream

El *Transport Stream* (TS) difiere significativamente del *Program Stream* (PS). El TS puede contener uno o múltiples programas audiovisuales y cada uno con su base de tiempo independiente. Es el resultado de multiplexar los ES que constituyen los diferentes programas mediante el uso de paquetes de longitud fija (y más pequeños que en el caso del PS), los *TS packets*, cuya longitud es de 188 bytes, 4 bytes de encabezamiento y 184 bytes de carga útil. En algunos casos, después de los 4 bytes de encabezamiento encontramos un Campo de Adaptación, para rellenar el exceso de espacio disponible. Además de esta función también transporta el Reloj de Referencia del Programa (*Program Clock Reference* o PCR). Vemos un ejemplo de la formación de los *TS packets* en la figura 4.5.

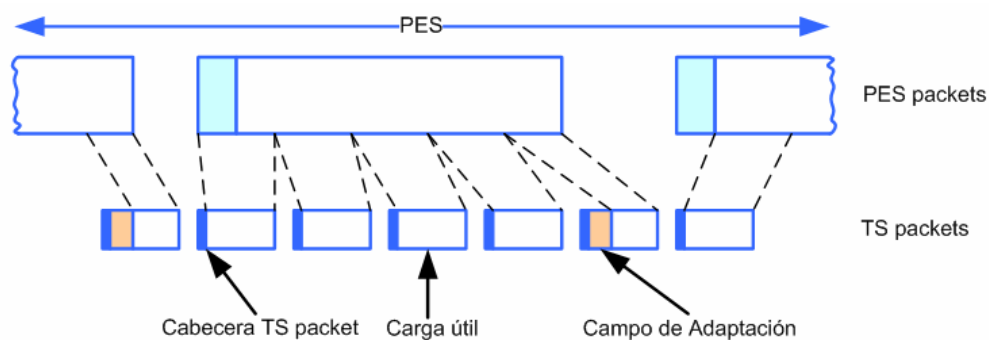


Figura 4.5 División de un PES en TS packets

El proceso de formación de los paquetes de transporte está sujeto al cumplimiento de dos condiciones:

- El primer byte de cada PES *packet* debe ser el primer byte de la carga útil de un TS *packet* (de ahí la necesidad del Campo de Adaptación)

- Un TS *packet* solamente puede contener datos tomados de un PES *packet*.

Entre las informaciones contenidas en la cabecera (que tiene longitud fija) están las siguientes:

- **Sync Byte.** Fija el comienzo del TS y permite la sincronización de la transmisión. Tiene un valor fijo (47 en hexadecimal) que no es único dentro del paquete.
- **Packet Identifier (PID).** Este campo tiene una longitud de trece bits es usado para identificar el canal. Contiene la información requerida para encontrar, identificar y reconstruir programas.
- **Adaptation Field.** Es opcional y de longitud variable. Cuando se usa proporciona información adicional sobre los datos transportados y causa una reducción de estos datos para mantener una tasa constante.

La pequeña longitud de estos paquetes es adecuada para procesos basados en hardware y permite ofrecer la robustez necesaria para la transmisión de video y audio comprimido, a través de canales ruidosos y propensos a introducir errores, tales como las redes de cable o de satélites.

Actualmente el TS está diseñado para soportar muchas funcionalidades tales como el multiplexaje asíncrono de programas, un acceso rápido al programa deseado cuando se cambia de canal, una correcta sincronización de los ES en la reproducción y el control de los buffers del decodificador durante la inicialización y la reproducción, tanto para programas con bitrate constante como variable.

4.4 Temporización

Con objeto de conseguir sincronización entre codificador y decodificador, obtener un presentación sincronizada entre los ES de audio y video, posibilitar el reordenamiento de las imágenes (recordamos que el orden de transmisión no era el mismo que el orden de display) y la anulación de los retardos aleatorios, en la capa de sistema se especifica un modelo de temporización. El retardo producido entre codificador y decodificador es la suma del tiempo transcurrido en la codificación, buffer de codificador, multiplexaje, transmisión o almacenamiento, demultiplexaje, buffer de decodificador, decodificación y presentación.

Este modelo se basa en la inserción de marcas de tiempo (time stamps) en la cabecera de los PES packets y de referencias de reloj en la cabecera de los TS packets. Los time stamps y las referencias de reloj son muestras o valores del reloj que interviene en el proceso de codificación y multiplexaje de cada programa.

Cada programa tiene asignado un reloj que recibe el nombre de *System Time Clock* (STC). Este reloj es la base de tiempo del programa asociado. En un TS cada programa lleva asociado un solo STC, pero un STC puede estar asociado a uno o varios programas a la vez. En cambio, en un PS, solamente se transporta un solo programa y por lo tanto solo existe un único STC.

Existen tres tipos de muestras de reloj que son las siguientes:

- **SCR** (*Sytsstem Clock Reference*) o **PCR** (*Program Clock Reference*). Estos permiten la regeneración, en el decodificador, del reloj de referencia del programa escogido. PCR es la terminología que se usa en los TS, ya que no existe un único reloj de sistema, sino uno por cada programa.
- **DTS** (*Decoding Time Stamp*). Indican en que instante se debe decodificar cada imagen.
- **PTS** (*Presentation Time Stamp*). Indican en que instante se debe presentar cada imagen o muestra de audio.

Los DTS y los PTS son una de las partes más importantes del modelo de temporización, ya que si por ejemplo en codificación bidireccional, una imagen tendría que ser decodificada un tiempo antes de que se muestre para permitirle actuar como referencia para una imagen B, el DTS indicaría el momento que una imagen debe ser decodificada y el PTS indicaría cuando tiene que ser presentada a la salida del decodificador.