

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 848 116**

51 Int. Cl.:

H04N 21/2343 (2011.01)

H04N 21/845 (2011.01)

H04N 21/262 (2011.01)

H04N 21/61 (2011.01)

H04L 29/06 (2006.01)

H04L 29/08 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.03.2016** **PCT/US2016/020652**

87 Fecha y número de publicación internacional: **09.09.2016** **WO16141165**

96 Fecha de presentación y número de la solicitud europea: **03.03.2016** **E 16711071 (7)**

97 Fecha y número de publicación de la concesión europea: **04.11.2020** **EP 3266218**

54 Título: **Transmisión basada en formato de archivo con formatos DASH basados en LCT**

30 Prioridad:

04.03.2015 US 201562128380 P

05.03.2015 US 201562128943 P

02.03.2016 US 201615058963

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.08.2021

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)
International IP Administration, 5775 Morehouse
Drive
San Diego, California 92121-1714, US

72 Inventor/es:

STOCKHAMMER, THOMAS;
WALKER, GORDON, KENT y
WANG, YE-KUI

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 848 116 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Transmisión basada en formato de archivo con formatos DASH basados en LCT

5 CAMPO TÉCNICO

[0001] Esta divulgación se refiere al almacenamiento y transporte de datos de vídeo codificados.

ANTECEDENTES

[0002] Las capacidades de vídeo digital se pueden incorporar a una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de radiodifusión directa digital, sistemas de radiodifusión inalámbrica, asistentes personales digitales (PDA), ordenadores portátiles o de escritorio, cámaras digitales, dispositivos de grabación digitales, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos de radio celulares o por satélite, dispositivos de videoconferencia y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como los descritos en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263 o ITU-T H.264/MPEG-4, Parte 10, Codificación Avanzada de Vídeo (AVC), ITU-T H.265/MPEG-H, Parte 2 (también denominada Codificación de Vídeo de Alta Eficiencia (HEVC)) y ampliaciones de dichas normas, para transmitir y recibir información de vídeo digital más eficientemente.

[0003] Samsung Electronics "Object Flow Mapping and Consumption", reunión TSG SA4 n.º 80 del 3 al 7 de agosto, San Francisco: se refiere a las mejoras compatibles con versiones anteriores para abordar los requisitos establecidos por el trabajo de mejoras de FLUTE. El documento WO 2014/063730 se refiere al campo de la entrega de contenido de medios multimedia. El documento WO 2016/018042 A1 se refiere a un procedimiento y aparato para procesar datos de medios transmitidos a través de banda ancha y radiodifusión en un sistema de radiodifusión en el que se combinan banda ancha y radiodifusión.

[0004] Las técnicas de compresión de vídeo realizan predicción espacial y/o predicción temporal para reducir o eliminar la redundancia inherente a las secuencias de vídeo. Para la codificación de vídeo basada en bloques, una trama o un fragmento de vídeo se pueden dividir en macrobloques. Cada macrobloque se puede dividir aún más. Los macrobloques de una trama o un fragmento intracodificados (I) se codifican usando predicción espacial con respecto a macrobloques contiguos. Los macrobloques de una trama o un fragmento intercodificados (P o B) pueden usar predicción espacial con respecto a macrobloques contiguos de la misma trama o fragmento, o predicción temporal con respecto a otras tramas de referencia.

[0005] Después de que se hayan codificado los datos de vídeo, los datos de vídeo se pueden agrupar en paquetes para su transmisión o almacenamiento. Los datos de vídeo se pueden ensamblar en un archivo de vídeo que se ajusta a cualquiera de una variedad de normas, tales como el formato de archivo de medios de base de la Organización Internacional de Normalización (ISO) y ampliaciones del mismo, tales como la AVC.

BREVE EXPLICACIÓN

[0006] Los aspectos de la invención están definidos en las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS**[0007]**

La FIG. 1 es un diagrama de bloques que ilustra un sistema de ejemplo que implementa técnicas para transmitir en continuo datos de medios a través de una red.

La FIG. 2 es un diagrama conceptual que ilustra elementos de contenido multimedia de ejemplo.

La FIG. 3 es un diagrama de bloques que ilustra elementos de un archivo de vídeo de ejemplo, que puede corresponder a un segmento de una representación.

La FIG. 4 es un diagrama conceptual que ilustra un escenario de ejemplo que a menudo surge en la entrega de radiodifusión.

La FIG. 5 es un diagrama conceptual que ilustra un sistema de ejemplo que puede realizar las técnicas de esta divulgación.

La FIG. 6 es un diagrama conceptual que ilustra un sistema de ejemplo que puede realizar las técnicas de esta divulgación.

La FIG. 7 es un diagrama conceptual que muestra diferentes aspectos de la entrada del servicio en el ejemplo

de DASH sobre FLUTE.

La FIG. 8 es un diagrama conceptual que muestra diferentes aspectos de la entrada de servicio para un ejemplo de DASH sobre ROUTE.

La FIG. 9 es un diagrama conceptual que ilustra un conjunto de ejemplos de campos de cabecera de acuerdo con RFC5651 que pueden usarse para transportar datos de acuerdo con las técnicas de esta divulgación.

La FIG. 10 es un diagrama conceptual que ilustra varias opciones para señalar cuando un prefijo de un objeto se puede liberar a la siguiente capa para descodificación.

La FIG. 11 es un diagrama conceptual que ilustra modelos de ejemplo para recepción de datos.

La FIG. 12 es un diagrama conceptual que ilustra un sistema de ejemplo para recibir datos de medios.

La FIG. 13 es un diagrama conceptual que ilustra un procedimiento de envío de ejemplo.

La FIG. 14 es un diagrama conceptual que ilustra un modelo de cliente DASH híbrido de ejemplo.

La FIG. 15 es un diagrama de flujo que ilustra un ejemplo de procedimiento para transportar datos de medios o una presentación de medios mediante sesiones LCT de acuerdo con las técnicas de esta divulgación.

La FIG. 16 es un diagrama de flujo que ilustra otro ejemplo de procedimiento para transportar datos de medios de una presentación de medios mediante sesiones LCT de acuerdo con las técnicas de esta divulgación.

La FIG. 17 es un diagrama de flujo que ilustra otro ejemplo de procedimiento para transportar datos de medios de una presentación de medios mediante sesiones LCT de acuerdo con las técnicas de esta divulgación.

DESCRIPCIÓN DETALLADA

[0008] En general, esta divulgación describe técnicas que permiten la capacidad de utilizar un protocolo unidireccional, por ejemplo, entrega de archivos por transporte unidireccional (FLUTE), entrega de objetos en tiempo real por transporte unidireccional (ROUTE) o entrega de objetos para la codificación en capas asíncrona y protocolos de multidifusión fiables con orientación NACK (FCAST), así como segmentos de transmisión adaptativa dinámica por HTTP (DASH) basándose en el formato de archivo de medios base ISO (ISO BMFF) y posiblemente también otros formatos como transmisión de transporte (TS) MPEG-2, para crear un sistema de entrega de medios de radiodifusión basado en IP que soporte la entrega y reproducción de transmisiones de medios que:

- Deben estar sincronizados entre sí (gestionados por ISO BMFF).
- Deben accederse a los mismos de forma aleatoria (tratados mediante señalización específica en el protocolo de entrega).
- Deben reproducirse de manera que no se produzca un nuevo almacenamiento en búfer ni un estancamiento.
- Deben combinarse con transmisiones de medios que se proporcionan y ofrecen en banda ancha y unidifusión.
- Habilitar la entrega y multiplexación de múltiples programas.
- Habilitar retardos de inicio bajos y cambios de canal rápidos.
- Habilitar el empalme de contenido en el remitente y en el receptor.
- Y proporcionar todas las características de DASH en un sistema de distribución de radiodifusión.

[0009] Más particularmente, las técnicas de esta divulgación permiten la recepción de datos de medios de una o más representaciones a través de sesiones de transporte de codificación en capas (LCT), sin (o antes de) recibir un archivo de manifiesto (como una descripción de presentación de medios (MPD)) para las representaciones. De esta manera, las técnicas de esta divulgación pueden reducir la latencia asociada con el inicio del consumo de los datos de los medios (por ejemplo, realizar el inicio del servicio). Es decir, en ausencia de estas técnicas, un dispositivo cliente necesitaría esperar la recepción del archivo de manifiesto, lo cual puede ocasionar una mala experiencia del usuario (por ejemplo, ver una pantalla negra y/o no escuchar reproducción de audio). El uso de estas técnicas puede reducir la latencia, de modo que se pueda mejorar la experiencia del usuario (por ejemplo, permitiendo una reproducción más rápida de datos de audio y/o vídeo).

[0010] Las técnicas de esta divulgación se pueden aplicar a los archivos de vídeo conformes a los datos de vídeo

encapsulados de acuerdo con cualquiera de entre el formato de archivo de medios de base ISO, el formato de archivo de codificación de vídeo escalable (SVC), el formato de archivo de codificación de vídeo avanzada (AVC), el formato de archivo del Proyecto de colaboración de tercera generación (3GPP) y/o el formato de archivo de codificación de vídeo de múltiples visualizaciones (MVC) u otros formatos similares de archivo de vídeo. Es decir, los segmentos de representaciones pueden formarse de acuerdo con cualquiera de estos diversos formatos. En general, los segmentos representan archivos multimedia que se pueden recibir de forma independiente de las representaciones correspondientes.

[0011] En la transmisión continua HTTP, las operaciones de uso frecuente incluyen HEAD, GET y GET parcial. La operación HEAD recupera una cabecera de un archivo asociado a un localizador de recursos uniforme (URL) o un nombre de recursos uniforme (URN) dado, sin recuperar una carga útil asociada al URL o al URN. La operación GET recupera un archivo entero asociado a un URL o URN dado. La operación GET parcial recibe un rango de bytes como parámetro de entrada y recupera un número continuo de bytes de un archivo, donde el número de bytes corresponde al rango de bytes recibido. Por tanto, se pueden proporcionar fragmentos de película para la transmisión continua del HTTP, porque una operación GET parcial puede obtener uno o más fragmentos de película individuales. En un fragmento de película, pueden existir varios fragmentos de pista de diferentes pistas. En la transmisión continua del HTTP, una presentación de medios puede ser una recopilación de datos estructurados que es accesible para el cliente. El cliente puede solicitar y descargar la información de datos de medios para presentar un servicio de transmisión continua a un usuario.

[0012] En el ejemplo de transmisión en continuo de datos 3GPP usando transmisión en continuo HTTP, pueden existir múltiples representaciones para datos de vídeo y/o audio de contenido multimedia. Como se describe a continuación, diferentes representaciones pueden corresponder a diferentes características de codificación (por ejemplo, diferentes perfiles o niveles de una norma de codificación de vídeo), diferentes normas de codificación o ampliaciones de normas de codificación (tales como ampliaciones de múltiples visualizaciones y/o escalables), o diferentes velocidades de transmisión de bits. El manifiesto de dichas representaciones se puede definir en una estructura de datos de descripción de presentación de medios (MPD). Una presentación de medios puede corresponder a un grupo de datos estructurado que es accesible para un dispositivo cliente de transmisión en continuo HTTP. El dispositivo cliente de transmisión en continuo HTTP puede solicitar y descargar información de datos de medios para presentar un servicio de transmisión en continuo a un usuario del dispositivo cliente. Una presentación de medios se puede describir en la estructura de datos MPD, que puede incluir actualizaciones de la MPD.

[0013] Una presentación de medios puede contener una secuencia de uno o más períodos. Los períodos pueden estar definidos por un elemento *período* en la MPD. Cada período puede tener un atributo *start* en la MPD. La MPD puede incluir un atributo *start* y un atributo *availableStartTime* para cada período. Para servicios en directo, la suma del atributo de *start* del período y del atributo *availableStartTime* de la MPD puede especificar el tiempo de disponibilidad del período en formato de UTC, en particular, el primer segmento de medios de cada representación en el período correspondiente. Para servicios a demanda, el atributo *start* del primer período puede ser 0. Para cualquier otro período, el atributo *start* puede especificar una desviación temporal entre el tiempo de inicio del período correspondiente con respecto al tiempo de inicio del primer período. Cada período se puede ampliar hasta el inicio del siguiente período, o hasta el final de la presentación de medios en el caso del último período. Los tiempos de inicio de período pueden ser precisos. Pueden reflejar la temporización real resultante de la reproducción de los medios de todos los períodos anteriores.

[0014] Cada período puede contener una o más representaciones para el mismo contenido de medios. Una representación puede ser una de un número de versiones codificadas alternativas de datos de audio o vídeo. Las representaciones pueden diferir según el tipo de codificación, por ejemplo, según la velocidad de transmisión de bits, la resolución y/o el códec para los datos de vídeo y la velocidad de transmisión de bits, el lenguaje y/o el códec para los datos de audio. El término representación se puede usar para referirse a una sección de datos de audio o vídeo codificados correspondientes a un período en particular del contenido multimedia y codificados de una forma en particular.

[0015] Las representaciones de un período en particular se pueden asignar a un grupo indicado en la MPD por un atributo indicativo de un conjunto de adaptación al que pertenecen las representaciones. Las representaciones en el mismo conjunto de adaptación en general se consideran alternativas entre sí, en la medida en que un dispositivo cliente puede conmutar entre estas representaciones de manera dinámica y sin interrupciones, por ejemplo, para realizar una adaptación de ancho de banda. Por ejemplo, cada representación de datos de vídeo para un período particular se puede asignar al mismo conjunto de adaptación, de modo que cualquiera de las representaciones se puede seleccionar para descodificar a fin de presentar datos de medios, tales como datos de vídeo o datos de audio, del contenido multimedia para el período correspondiente. El contenido de medios dentro de un período se puede representar mediante una representación cualquiera del grupo 0, si está presente, o bien la combinación de como máximo una representación de cada grupo no cero, en algunos ejemplos. Los datos de temporización para cada representación de un período se pueden expresar con respecto al tiempo de inicio del período.

[0016] Una representación puede incluir uno o más segmentos. Cada representación puede incluir un segmento de inicialización, o cada segmento de una representación puede ser autoinicializador. Cuando está presente, el segmento de inicialización puede contener información de inicialización para acceder a la representación. En general, el segmento de inicialización no contiene datos de medios. Un segmento se puede referenciar de forma única mediante un identificador, tal como un localizador de recursos uniforme (URL), un nombre de recursos uniforme (URN) o un identificador de recursos uniforme (URI). La MPD puede proporcionar los identificadores para cada segmento. En algunos ejemplos, la MPD también puede proporcionar gamas de bytes en forma de un atributo de rango, que puede corresponder a los datos para un segmento dentro de un archivo accesible mediante el URL, el URN o el URI.

[0017] Se pueden seleccionar diferentes representaciones para una recuperación sustancialmente simultánea para diferentes tipos de datos de medios. Por ejemplo, un dispositivo cliente puede seleccionar una representación de audio, una representación de vídeo y una representación de texto temporizado a partir de las cuales se pueden recuperar segmentos. En algunos ejemplos, el dispositivo cliente puede seleccionar conjuntos de adaptación particulares para realizar una adaptación de ancho de banda. Es decir, el dispositivo cliente puede seleccionar un conjunto de adaptación que incluye representaciones de vídeo, un conjunto de adaptación que incluye representaciones de audio y/o un conjunto de adaptación que incluye texto temporizado. De forma alternativa, el dispositivo cliente puede seleccionar conjuntos de adaptación para determinados tipos de medios (por ejemplo, vídeo) y seleccionar directamente representaciones para otros tipos de medios (por ejemplo, audio y/o texto temporizado).

[0018] La FIG. 1 es un diagrama de bloques que ilustra un sistema 10 de ejemplo que implementa las técnicas de esta divulgación para transmitir datos de medios a través de una red. Los diversos elementos del sistema 10 pueden corresponder en general a elementos similares de los ejemplos mostrados en las FIGS. 5 y 6, como se analiza con mayor detalle a continuación. Asimismo, los componentes del dispositivo cliente 40 pueden corresponder en general a los componentes de las FIGS. 11, 12 y 14, como también se analiza con mayor detalle a continuación.

[0019] En el ejemplo de la FIG. 1, el sistema 10 incluye un dispositivo de preparación de contenido 20, un dispositivo servidor 60 y un dispositivo cliente 40. El dispositivo cliente 40 y el dispositivo servidor 60 están acoplados de forma comunicativa por una red 74, que puede comprender Internet. En algunos ejemplos, el dispositivo de preparación de contenido 20 y el dispositivo servidor 60 también pueden estar acoplados por la red 74 u otra red, o pueden estar directamente acoplados de forma comunicativa. En algunos ejemplos, el dispositivo de preparación de contenido 20 y el dispositivo servidor 60 pueden comprender el mismo dispositivo.

[0020] El dispositivo de preparación de contenido 20, en el ejemplo de la FIG. 1, comprende una fuente de audio 22 y una fuente de vídeo 24. La fuente de audio 22 puede comprender, por ejemplo, un micrófono que produce señales eléctricas representativas de datos de audio captados que el codificador de audio 26 va a codificar. De forma alternativa, la fuente de audio 22 puede comprender un medio de almacenamiento que almacena datos de audio previamente registrados, un generador de datos de audio tal como un sintetizador informatizado, o cualquier otra fuente de datos de audio. La fuente de vídeo 24 puede comprender una cámara de vídeo que produce datos de vídeo que el codificador de vídeo 28 va a codificar, un medio de almacenamiento codificado con datos de vídeo previamente registrados, una unidad de generación de datos de vídeo, tal como una fuente de gráficos de ordenador, o cualquier otra fuente de datos de vídeo. El dispositivo de preparación de contenido 20 no está necesariamente acoplado de forma comunicativa al dispositivo servidor 60 en todos los ejemplos, pero puede almacenar contenido multimedia en un medio separado que el dispositivo servidor 60 lee.

[0021] Los datos de audio y vídeo no procesados pueden comprender datos analógicos o digitales. Los datos analógicos se pueden digitalizar antes de que el codificador de audio 26 y/o el codificador de vídeo 28 los codifiquen. La fuente de audio 22 puede obtener datos de audio a partir de un participante que habla mientras el participante que habla está hablando, y la fuente de vídeo 24 puede obtener simultáneamente datos de vídeo del participante que habla. En otros ejemplos, la fuente de audio 22 puede comprender un medio de almacenamiento legible por ordenador que comprende datos de audio almacenados, y la fuente de vídeo 24 puede comprender un medio de almacenamiento legible por ordenador que comprende datos de vídeo almacenados. De esta manera, las técnicas descritas en esta divulgación se pueden aplicar a la transmisión continua en directo y en tiempo real de datos de audio y vídeo, o de datos de audio y vídeo archivados y pre-registrados.

[0022] Las tramas de audio que corresponden a tramas de vídeo son en general tramas de audio que contienen datos de audio que la fuente de audio 22 ha captado (o generado) al mismo tiempo que unos datos de vídeo, que la fuente de vídeo 24 ha captado (o generado), que están contenidos dentro de las tramas de vídeo. Por ejemplo, mientras un participante que habla en general produce, al hablar, datos de audio, la fuente de audio 22 capta los datos de audio, y la fuente de vídeo 24 capta los datos de vídeo del participante que habla al mismo tiempo, es decir, mientras la fuente de audio 22 está captando los datos de audio. Así pues, una trama de audio puede corresponder temporalmente a una o más tramas de vídeo en particular. Por consiguiente, una trama de audio correspondiente a una trama de vídeo corresponde en general a una situación en la que se han captado datos de audio y datos de vídeo al mismo tiempo, y para la que una trama de audio y una trama de vídeo comprenden,

respectivamente, los datos de audio y los datos de vídeo que se han captado al mismo tiempo.

[0023] En algunos ejemplos, el codificador de audio 26 puede codificar una marca de tiempo en cada trama de audio codificada, que representa un tiempo en el que se han registrado los datos de audio para la trama de audio codificada y, de forma similar, el codificador de vídeo 28 puede codificar una marca de tiempo en cada trama de vídeo codificada, que representa un tiempo en el que se han registrado los datos de vídeo para la trama de vídeo codificada. En dichos ejemplos, una trama de audio correspondiente a una trama de vídeo puede comprender una trama de audio que comprende una marca de tiempo y una trama de vídeo que comprende la misma marca de tiempo. El dispositivo de preparación de contenido 20 puede incluir un reloj interno a partir del cual el codificador de audio 26 y/o el codificador de vídeo 28 pueden generar las marcas de tiempo, o que la fuente de audio 22 y la fuente de vídeo 24 pueden usar para asociar datos de audio y vídeo, respectivamente, a una marca de tiempo.

[0024] En algunos ejemplos, la fuente de audio 22 puede enviar datos al codificador de audio 26, correspondientes a un tiempo en el que se han registrado los datos de audio, y la fuente de vídeo 24 puede enviar datos al codificador de vídeo 28, correspondientes a un tiempo en el que se han registrado los datos de vídeo. En algunos ejemplos, el codificador de audio 26 puede codificar un identificador de secuencia en unos datos de audio codificados para indicar un orden temporal relativo de los datos de audio codificados, pero sin indicar necesariamente un tiempo absoluto en el que se han registrado los datos de audio y, de forma similar, el codificador de vídeo 28 también puede usar identificadores de secuencia para indicar un orden temporal relativo de los datos de vídeo codificados. De forma similar, en algunos ejemplos, un identificador de secuencia se puede asignar a, o en cualquier caso correlacionar con, una marca de tiempo.

[0025] El codificador de audio 26, en general, produce un flujo de datos de audio codificados, mientras que el codificador de vídeo 28 produce un flujo de datos de vídeo codificados. Cada flujo de datos individual (ya sea de audio o vídeo) se puede denominar flujo elemental. Un flujo elemental es un componente único codificado digitalmente (y posiblemente comprimido) de una representación. Por ejemplo, la parte de vídeo o audio codificado de la representación puede ser un flujo elemental. Un flujo elemental se puede convertir en un flujo elemental paquetizado (PES) antes de encapsularse dentro de un archivo de vídeo. Dentro de la misma representación, se puede usar un ID de flujo para distinguir los paquetes PES que pertenecen a un flujo elemental de los otros. La unidad básica de datos de un flujo elemental es un paquete de flujo elemental paquetizado (PES). Por tanto, los datos de vídeo codificados corresponden en general a flujos de vídeo elementales. De forma similar, los datos de audio corresponden a uno o más flujos elementales respectivos.

[0026] Muchas normas de codificación de vídeo, tales como ITU-T H.264/AVC y la inminente norma de codificación de vídeo de alta eficiencia (HEVC), definen la sintaxis, la semántica y el proceso de descodificación para flujos de bits sin errores, cualquiera de los cuales se ajusta a un determinado perfil o nivel. Las normas de codificación de vídeo típicamente no especifican el codificador, pero el codificador se ocupa de garantizar que los flujos de bits generados cumplan las normas para un descodificador. En el contexto de las normas de codificación de vídeo, un "perfil" corresponde a un subconjunto de algoritmos, rasgos característicos o herramientas y restricciones que se les aplican. Como se define en la norma H.264, por ejemplo, un "perfil" es un subconjunto de toda la sintaxis del flujo de bits especificada por la norma H.264. Un "nivel" corresponde a las limitaciones del consumo de recursos del descodificador, tales como, por ejemplo, memoria y cálculo del descodificador, que están relacionados con la resolución de las imágenes, la velocidad de bits y la velocidad de procesamiento de bloques. Un perfil se puede señalar con un valor `idc` de perfil (indicador de perfil), mientras que un nivel se puede señalar con un valor `level_idc` (indicador de nivel).

[0027] La norma H.264, por ejemplo, reconoce que, dentro de los límites impuestos por la sintaxis de un perfil dado, todavía es posible requerir una gran variación del rendimiento de los codificadores y descodificadores, dependiendo de los valores adoptados por los elementos de sintaxis en el flujo de bits, tales como el tamaño especificado de las imágenes descodificadas. La norma H.264 reconoce, además, que, en muchas aplicaciones, no es ni práctico ni económico implementar un descodificador capaz de tratar todos los usos hipotéticos de la sintaxis dentro de un perfil en particular. Por consiguiente, la norma H.264 define un "nivel" como un conjunto especificado de restricciones impuestas a los valores de los elementos de sintaxis en el flujo de bits. Estas restricciones pueden ser simples limitaciones de valores. De forma alternativa, estas restricciones pueden adoptar la forma de restricciones sobre combinaciones aritméticas de valores (por ejemplo, la anchura de la imagen multiplicada por la altura de la imagen multiplicada por el número de imágenes descodificadas por segundo). La norma H.264 establece, además, que las implementaciones individuales pueden soportar un nivel diferente para cada perfil soportado.

[0028] Un descodificador que se ajusta a un perfil normalmente soporta todos los rasgos característicos definidos en el perfil. Por ejemplo, como rasgo característico de codificación, la codificación de imágenes B no está soportada en el perfil de línea de base de H.264/AVC, pero está soportada en otros perfiles de H.264/AVC. Un descodificador que se ajusta a un nivel deberá ser capaz de descodificar cualquier flujo de bits que no requiere recursos fuera de las limitaciones definidas en el nivel. Las definiciones de perfiles y niveles pueden ser útiles para la interpretabilidad. Por ejemplo, durante la transmisión de vídeo, se pueden negociar y acordar un par de definiciones de perfil y nivel para una sesión de transmisión completa. Más específicamente, en H.264/AVC, un nivel puede definir limitaciones

en el número de macrobloques que es necesario procesar, el tamaño de la memoria intermedia de imágenes descodificadas (DPB), el tamaño de la memoria intermedia de imágenes codificadas (CPB), el rango de vectores de movimiento vertical, el número máximo de vectores de movimiento para cada dos MB consecutivos y si un bloque B puede tener divisiones de submacrobloque inferiores a 8x8 píxeles. De esta manera, un descodificador puede determinar si el descodificador es capaz de descodificar apropiadamente el flujo de bits.

[0029] En el ejemplo de la FIG. 1, la unidad de encapsulación 30 del dispositivo de preparación de contenido 20 recibe flujos elementales que comprenden datos de vídeo codificados desde el codificador de vídeo 28 y flujos elementales que comprenden datos de audio codificados desde el codificador de audio 26. En algunos ejemplos, el codificador de vídeo 28 y el codificador de audio 26 pueden incluir, cada uno, paquetizadores para formar paquetes PES a partir de datos codificados. En otros ejemplos, el codificador de vídeo 28 y el codificador de audio 26 pueden interactuar, cada uno, con los paquetizadores respectivos para formar paquetes PES a partir de datos codificados. En otros ejemplos más, la unidad de encapsulación 30 puede incluir paquetizadores para formar paquetes PES a partir de datos de audio y de vídeo codificados.

[0030] El codificador de vídeo 28 puede codificar datos de vídeo de contenido multimedia en una variedad de formas, para producir diferentes representaciones del contenido multimedia a diversas velocidades de transmisión de bits y con diversas características, tales como resoluciones de píxeles, velocidades de tramas, conformidad con diversas normas de codificación, conformidad con diversos perfiles y/o niveles de perfiles para diversas normas de codificación, representaciones que tienen una o múltiples visualizaciones (por ejemplo, para reproducción bidimensional o tridimensional), u otras características de ese tipo. Una representación, como se usa en esta divulgación, puede comprender uno de datos de audio, datos de vídeo, datos de texto (por ejemplo, para subtítulos cerrados) u otros datos de este tipo. La representación puede incluir un flujo elemental, tal como un flujo elemental de audio o un flujo elemental de vídeo. Cada paquete PES puede incluir un identificador stream_id que identifica el flujo elemental al que pertenece el paquete PES. La unidad de encapsulación 30 es responsable de ensamblar flujos elementales en archivos de vídeo (por ejemplo, segmentos) de diversas representaciones.

[0031] La unidad de encapsulación 30 recibe paquetes PES para flujos elementales de una representación desde el codificador de audio 26 y el codificador de vídeo 28 y forma las correspondientes unidades de capa de abstracción de red (NAL) a partir de los paquetes PES. En el ejemplo de la H.264/AVC (codificación de vídeo avanzada), los segmentos de vídeo codificados están organizados en unidades de NAL, que proporcionan una representación de vídeo "apta para redes" dirigida a aplicaciones tales como la videotelefonía, el almacenamiento, la radiodifusión o la transmisión continua. Las unidades NAL se pueden clasificar en unidades NAL de capa de codificación de vídeo (VCL) y unidades NAL no VCL. Las unidades VCL pueden contener el motor de compresión central y pueden incluir datos a nivel de bloque, macrobloque y/o fragmento. Otras unidades NAL pueden ser unidades NAL no VCL. En algunos ejemplos, una imagen codificada en una instancia de tiempo, normalmente presentada como una imagen codificada primaria, puede estar contenida en una unidad de acceso, que puede incluir una o más unidades NAL.

[0032] Las unidades NAL no VCL pueden incluir unidades NAL de conjunto de parámetros y unidades NAL SEI, entre otras. Los conjuntos de parámetros pueden contener información de cabecera a nivel de secuencia (en conjuntos de parámetros de secuencia (SPS)) y la información de cabecera a nivel de imagen que cambia ocasionalmente (en conjuntos de parámetros de imagen (PPS)). Con los conjuntos de parámetros (por ejemplo, PPS y SPS), la información que cambia ocasionalmente no necesita ser repetida para cada secuencia o imagen, de ahí que pueda mejorarse la eficiencia de la codificación. Además, el uso de conjuntos de parámetros puede permitir la transmisión fuera de banda de la información de cabecera importante, evitando la necesidad de transmisiones redundantes para la resistencia a los errores. En los ejemplos de transmisión fuera de banda, las unidades NAL de conjunto de parámetros se pueden transmitir en un canal diferente al de otras unidades NAL, tales como las unidades NAL SEI.

[0033] La información de mejora complementaria (SEI) puede contener información que no es necesaria para descodificar las muestras de imágenes codificadas a partir de las unidades NAL VCL, pero puede ayudar en los procesos relacionados con la descodificación, visualización, resistencia a los errores y otros propósitos. Los mensajes de SEI pueden estar contenidos en unidades NAL no VCL. Los mensajes de SEI son la parte normativa de algunas memorias descriptivas habituales y, por tanto, no siempre son obligatorios para la implementación de descodificadores que cumplen las normas. Los mensajes de SEI pueden ser mensajes de SEI a nivel de secuencia o mensajes de SEI a nivel de imagen. Parte de la información a nivel de secuencia puede estar contenida en mensajes de SEI, tales como mensajes de SEI de información de adaptabilidad a escala en el ejemplo de SVC y mensajes de SEI de información de adaptabilidad a escala de la visualización en MVC. Estos ejemplos de mensajes de SEI pueden transmitir información, por ejemplo, sobre extracción de puntos de funcionamiento y características de los puntos de funcionamiento. Además, la unidad de encapsulación 30 puede formar un archivo de manifiesto, tal como una descripción de presentación de medios (MPD) que describe características de las representaciones. La unidad de encapsulación 30 puede formatear la MPD de acuerdo con un lenguaje de marcado extensible (XML).

[0034] La unidad de encapsulación 30 puede proporcionar datos para una o más representaciones de contenido

multimedia, junto con el archivo de manifiesto (por ejemplo, la MPD), a la interfaz de salida 32. La interfaz de salida 32 puede comprender una interfaz de red o una interfaz para escribir en un medio de almacenamiento, tal como una interfaz de bus serie universal (USB), una grabadora o copiadora de CD o DVD, una interfaz para medios de almacenamiento magnéticos o flash, u otras interfaces para almacenar o transmitir datos de medios. La unidad de encapsulación 30 puede proporcionar datos de cada una de las representaciones de contenido multimedia a la interfaz de salida 32, que puede enviar los datos al dispositivo servidor 60 por medio de transmisión por red o medios de almacenamiento. En el ejemplo de la FIG. 1, el dispositivo servidor 60 incluye un medio de almacenamiento 62 que almacena diversos contenidos multimedia 64, incluyendo cada uno un respectivo archivo de manifiesto 66 y una o más representaciones 68A a 68N (representaciones 68). En algunos ejemplos, la interfaz de salida 32 también puede enviar datos directamente a la red 74.

[0035] En algunos ejemplos, las representaciones 68 se pueden separar en conjuntos de adaptación. Es decir, diversos subconjuntos de representaciones 68 pueden incluir respectivos conjuntos comunes de características, tales como códec, perfil y nivel, resolución, número de visualizaciones, formato de archivo para segmentos, información de tipo de texto que puede identificar un lenguaje u otras características de un texto que se va a visualizar con la representación y/o datos de audio que se van a descodificar y presentar, por ejemplo, mediante altavoces, información de ángulo de cámara que puede describir un ángulo de cámara o una perspectiva de cámara real de una escena para representaciones del conjunto de adaptación, información de calificación que describe la idoneidad del contenido para audiencias en particular, o similares.

[0036] El archivo de manifiesto 66 puede incluir datos indicativos de los subconjuntos de representaciones 68 correspondientes a conjuntos de adaptación en particular, así como características comunes para los conjuntos de adaptación. El archivo de manifiesto 66 también puede incluir datos representativos de características individuales, tales como las velocidades de transmisión de bits, para representaciones individuales de conjuntos de adaptación. De esta manera, un conjunto de adaptación puede hacer posible una adaptación simplificada del ancho de banda de red. Las representaciones de un conjunto de adaptación se pueden indicar usando elementos hijo de un elemento del conjunto de adaptación del archivo de manifiesto 66.

[0037] El dispositivo servidor 60 incluye una unidad de procesamiento de peticiones 70 y una interfaz de red 72. En algunos ejemplos, el dispositivo servidor 60 puede incluir una pluralidad de interfaces de red. Además, uno cualquiera o todos los rasgos característicos del dispositivo servidor 60 se pueden implementar en otros dispositivos de una red de entrega de contenido, tales como encaminadores, puentes, dispositivos proxy, conmutadores u otros dispositivos. En algunos ejemplos, los dispositivos intermedios de una red de entrega de contenido pueden almacenar en memoria caché datos de contenido multimedia 64, e incluir componentes que se ajustan sustancialmente a los del dispositivo servidor 60. En general, la interfaz de red 72 está configurada para enviar y recibir datos por medio de la red 74.

[0038] La unidad de procesamiento de peticiones 70 está configurada para recibir peticiones de red desde dispositivos cliente, tales como el dispositivo cliente 40, para datos del medio de almacenamiento 62. Por ejemplo, la unidad de procesamiento de peticiones 70 puede implementar el protocolo de transferencia de hipertexto (HTTP) versión 1.1, como se describe en RFC2616, "Hypertext Transfer Protocol-HTTP/1.1", de R. Fielding y otros, Grupo de Trabajo de la Red, IETF, junio de 1999. Es decir, la unidad de procesamiento de peticiones 70 puede estar configurada para recibir peticiones GET o GET parciales HTTP y proporcionar datos de contenido multimedia 64 como respuesta a las peticiones. Las peticiones pueden especificar un segmento de una de las representaciones 68, por ejemplo, usando un URL del segmento. En algunos ejemplos, las peticiones también pueden especificar uno o más rangos de bytes del segmento, comprendiendo por tanto peticiones GET parciales. La unidad de procesamiento de peticiones 70 puede estar configurada, además, para atender peticiones HEAD HTTP para proporcionar datos de cabecera de un segmento de una de las representaciones 68. En cualquier caso, la unidad de procesamiento de peticiones 70 puede estar configurada para procesar las peticiones para proporcionar los datos solicitados a un dispositivo solicitante, tal como el dispositivo cliente 40.

[0039] De forma adicional o alternativa, la unidad de procesamiento de peticiones 70 puede estar configurada para entregar datos de medios por medio de un protocolo de radiodifusión o multidifusión, tal como el eMBMS. El dispositivo de preparación de contenido 20 puede crear segmentos y/o subsegmentos DASH, sustancialmente de la misma manera que se ha descrito, pero el dispositivo servidor 60 puede entregar estos segmentos o subsegmentos usando el eMBMS u otro protocolo de transporte de red de radiodifusión o multidifusión. Por ejemplo, la unidad de procesamiento de peticiones 70 puede estar configurada para recibir una petición para unirse a un grupo de multidifusión desde el dispositivo cliente 40. Es decir, el dispositivo servidor 60 puede comunicar una dirección de protocolo de Internet (IP), asociada a un grupo de multidifusión a unos dispositivos cliente, que incluyen el dispositivo cliente 40, asociados a un contenido de medios en particular (por ejemplo, radiodifusión de un acontecimiento en directo). El dispositivo cliente 40, a su vez, puede presentar una petición para unirse al grupo de multidifusión. Esta petición se puede propagar por toda la red 74, por ejemplo, los encaminadores que componen la red 74, de modo que se hace que los encaminadores dirijan el tráfico destinado a la dirección IP asociada al grupo de multidifusión a los dispositivos cliente abonados, tales como el dispositivo cliente 40.

[0040] Como se ilustra en el ejemplo de la FIG. 1, el contenido multimedia 64 incluye el archivo de manifiesto 66,

que puede corresponder a una descripción de presentación de medios (MPD). El archivo de manifiesto 66 puede contener descripciones de diferentes representaciones alternativas 68 (por ejemplo, servicios de vídeo con diferentes calidades) y la descripción puede incluir, por ejemplo, información de códec, un valor de perfil, un valor de nivel, una velocidad de transmisión de bits y otras características descriptivas de las representaciones 68. El dispositivo cliente 40 puede recuperar la MPD de una presentación de medios para determinar cómo acceder a segmentos de las representaciones 68.

[0041] En particular, la unidad de recuperación 52 puede recuperar datos de configuración (no mostrados) del dispositivo cliente 40 para determinar las capacidades de decodificación del decodificador de vídeo 48 y las capacidades de representación de la salida de vídeo 44. Los datos de configuración también pueden incluir cualquiera o todas las preferencias de lenguaje seleccionadas por un usuario del dispositivo cliente 40, una o más perspectivas de cámara correspondientes a las preferencias de profundidad establecidas por el usuario del dispositivo cliente 40 y/o una preferencia de calificación seleccionada por el usuario del dispositivo cliente 40. La unidad de recuperación 52 puede comprender, por ejemplo, un navegador web o un cliente de medios configurados para presentar peticiones GET y GET parcial HTTP. La unidad de recuperación 52 puede corresponder a unas instrucciones de software ejecutadas por uno o más procesadores o unidades de procesamiento (no mostrados) del dispositivo cliente 40. En algunos ejemplos, la totalidad o unas partes de la funcionalidad descrita con respecto a la unidad de recuperación 52 se pueden implementar en hardware, o una combinación de hardware, software y/o firmware, donde se puede proporcionar el hardware requerido para ejecutar instrucciones para software o firmware.

[0042] La unidad de recuperación 52 puede comparar las capacidades de decodificación y representación del dispositivo cliente 40 con las características de las representaciones 68 indicadas por la información del archivo de manifiesto 66. La unidad de recuperación 52 puede recuperar inicialmente al menos una parte del archivo de manifiesto 66 para determinar las características de las representaciones 68. Por ejemplo, la unidad de recuperación 52 puede solicitar una parte del archivo de manifiesto 66 que describe las características de uno o más conjuntos de adaptación. La unidad de recuperación 52 puede seleccionar un subconjunto de representaciones 68 (por ejemplo, un conjunto de adaptación) que tiene características que se pueden satisfacer mediante las capacidades de codificación y representación del dispositivo cliente 40. La unidad de recuperación 52 puede, a continuación, determinar las velocidades de transmisión de bits para las representaciones del conjunto de adaptación, determinar una cantidad actualmente disponible de ancho de banda de red y recuperar segmentos de una de las representaciones que tiene una velocidad de transmisión de bits que se puede satisfacer mediante el ancho de banda de red.

[0043] En general, las representaciones de velocidades de transmisión de bits mayores pueden producir una reproducción de vídeo de mayor calidad, mientras que las representaciones de velocidades de transmisión de bits más bajas pueden proporcionar una reproducción de vídeo de calidad suficiente cuando el ancho de banda de red disponible se reduce. En consecuencia, cuando el ancho de banda de red disponible es relativamente alto, la unidad de recuperación 52 puede recuperar datos de representaciones de velocidades de transmisión de bits relativamente altas, mientras que cuando el ancho de banda de red disponible es bajo, la unidad de recuperación 52 puede recuperar datos de representaciones de velocidades de transmisión de bits relativamente bajas. De esta manera, el dispositivo cliente 40 puede transmitir en continuo datos de medios a través de la red 74 mientras que también se adapta a la disponibilidad cambiante de ancho de banda de red de la red 74.

[0044] De forma adicional o alternativa, la unidad de recuperación 52 puede estar configurada para recibir datos de acuerdo con un protocolo de red de radiodifusión o multidifusión, tal como la multidifusión eMBMS o IP. En dichos ejemplos, la unidad de recuperación 52 puede presentar una petición para unirse a un grupo de red de multidifusión asociado a un contenido de medios en particular. Después de unirse al grupo de multidifusión, la unidad de recuperación 52 puede recibir datos del grupo de multidifusión sin peticiones adicionales emitidas al dispositivo servidor 60 o al dispositivo de preparación de contenido 20. La unidad de recuperación 52 puede presentar una petición para abandonar el grupo de multidifusión cuando ya no se necesitan datos del grupo de multidifusión, por ejemplo, para detener la reproducción o para cambiar canales a un grupo de multidifusión diferente.

[0045] La interfaz de red 54 puede recibir y proporcionar datos de segmentos de una representación seleccionada a la unidad de recuperación 52, que a su vez puede proporcionar los segmentos a la unidad de desencapsulación 50. La unidad de desencapsulación 50 puede desencapsular elementos de un archivo de vídeo en flujos PES constituyentes, desempaquetar los flujos PES para recuperar datos codificados y enviar los datos codificados al decodificador de audio 46 o bien al decodificador de vídeo 48, dependiendo de si los datos codificados forman parte de un flujo de audio o vídeo, por ejemplo, como lo indican las cabeceras de paquetes PES del flujo. El decodificador de audio 46 decodifica datos de audio codificados y envía los datos de audio decodificados a la salida de audio 42, mientras que el decodificador de vídeo 48 decodifica datos de vídeo codificados y envía los datos de vídeo decodificados, que pueden incluir una pluralidad de visualizaciones de un flujo, a la salida de vídeo 44.

[0046] El codificador de vídeo 28, el decodificador de vídeo 48, el codificador de audio 26, el decodificador de

audio 46, la unidad de encapsulación 30, la unidad de recuperación 52 y la unidad de desencapsulación 50 se pueden implementar cada uno como cualquiera de una variedad de circuitos de procesamiento adecuados, según corresponda, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), circuitos lógicos discretos, programa informático, hardware, firmware o cualquier combinación de los mismos. Tanto el codificador de vídeo 28 como el descodificador de vídeo 48 se pueden incluir en uno o más codificadores o descodificadores, ambos de los cuales se pueden integrar como parte de un codificador/descodificador (CÓDEC) de vídeo combinado. Asimismo, tanto el codificador de audio 26 como el descodificador de audio 46 pueden estar incluidos en uno o más codificadores o descodificadores, ambos de los cuales pueden estar integrados como parte de un CÓDEC combinado. Un aparato que incluye un codificador de vídeo 28, un descodificador de vídeo 48, un codificador de audio 26, un descodificador de audio 46, una unidad de encapsulación 30, una unidad de recuperación 52 y/o una unidad de desencapsulación 50 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono móvil.

[0047] El dispositivo cliente 40, el dispositivo servidor 60 y/o el dispositivo de preparación de contenido 20 pueden estar configurados para funcionar de acuerdo con las técnicas de esta divulgación. Con propósitos de ejemplo, esta divulgación describe estas técnicas con respecto al dispositivo cliente 40 y al dispositivo servidor 60. Sin embargo, se deberá entender que el dispositivo de preparación de contenido 20 puede estar configurado para realizar estas técnicas, en lugar (o, además) del dispositivo servidor 60.

[0048] La unidad de encapsulación 30 puede formar unidades NAL que comprenden una cabecera que identifica un programa al cual pertenece la unidad NAL, así como una carga útil, por ejemplo, datos de audio, datos de vídeo o datos que describen el flujo de transporte o de programa al cual corresponde la unidad NAL. Por ejemplo, en H.264/AVC, una unidad NAL incluye una cabecera de 1 byte y una carga útil de tamaño variable. Una unidad NAL que incluye datos de vídeo en su carga útil puede comprender diversos niveles de granularidad de datos de vídeo. Por ejemplo, una unidad NAL puede comprender un bloque de datos de vídeo, una pluralidad de bloques, un fragmento de datos de vídeo o una imagen completa de datos de vídeo. La unidad de encapsulación 30 puede recibir datos de vídeo codificados desde el codificador de vídeo 28 en forma de paquetes PES de flujos elementales. La unidad de encapsulación 30 puede asociar cada flujo elemental a un programa correspondiente.

[0049] La unidad de encapsulación 30 también puede ensamblar unidades de acceso desde una pluralidad de unidades NAL. En general, una unidad de acceso puede comprender una o más unidades NAL para representar una trama de datos de vídeo, así como datos de audio correspondientes a la trama cuando dichos datos de audio están disponibles. Una unidad de acceso incluye en general todas las unidades NAL para una instancia de tiempo de salida, por ejemplo, todos los datos de audio y vídeo para una instancia de tiempo. Por ejemplo, si cada visualización tiene una frecuencia de tramas de 20 tramas por segundo (fps), cada instancia de tiempo puede corresponder a un intervalo de tiempo de 0,05 segundos. Durante este intervalo de tiempo, las tramas específicas para todas las visualizaciones de la misma unidad de acceso (la misma instancia de tiempo) se pueden representar simultáneamente. En un ejemplo, una unidad de acceso puede comprender una imagen codificada en una instancia de tiempo, que se puede presentar como una imagen codificada primaria.

[0050] Por consiguiente, una unidad de acceso puede comprender todas las tramas de audio y vídeo de una instancia temporal común, por ejemplo, todas las visualizaciones correspondientes al tiempo X. Esta divulgación también se refiere a una imagen codificada de una visualización particular como un "componente de visualización". Es decir, un componente de visualización puede comprender una imagen (o trama) codificada para una visualización en particular en un tiempo en particular. Por consiguiente, se puede definir una unidad de acceso que comprende todos los componentes de visualización de una instancia temporal común. El orden de decodificación de las unidades de acceso no necesita ser el mismo que el orden de salida o de visualización.

[0051] Una presentación de medios puede incluir una descripción de presentación de medios (MPD), que puede contener descripciones de diferentes representaciones alternativas (por ejemplo, servicios de vídeo con diferentes calidades) y la descripción puede incluir, por ejemplo, información de códec, un valor de perfil y un valor de nivel. Una MPD es un ejemplo de archivo de manifiesto, tal como el archivo de manifiesto 66. El dispositivo cliente 40 puede recuperar la MPD de una presentación de medios para determinar cómo acceder a fragmentos de película de diversas presentaciones. Los fragmentos de película pueden estar situados en cuadros de fragmento de película (cuadros moof) de archivos de vídeo.

[0052] El archivo de manifiesto 66 (que puede comprender, por ejemplo, una MPD) puede comunicar la disponibilidad de segmentos de representaciones 68. Es decir, la MPD puede incluir información que indica el tiempo de hora de reloj en el cual un primer segmento de una de las representaciones 68 queda disponible, así como información que indica las duraciones de los segmentos dentro de las representaciones 68. De esta manera, la unidad de recuperación 52 del dispositivo cliente 40 puede determinar cuándo está disponible cada segmento, basándose en el tiempo de inicio, así como de las duraciones de los segmentos que preceden a un segmento en particular.

[0053] Después de que la unidad de encapsulación 30 haya ensamblado las unidades de NAL y/o las unidades

de acceso en un archivo de vídeo, basándose en los datos recibidos, la unidad de encapsulación 30 pasa el archivo de vídeo a la interfaz de salida 32 para su salida. En algunos ejemplos, la unidad de encapsulación 30 puede almacenar el archivo de vídeo localmente o enviar el archivo de vídeo a un servidor remoto por medio de la interfaz de salida 32, en lugar de enviar el archivo de vídeo directamente al dispositivo cliente 40. La interfaz de salida 32 puede comprender, por ejemplo, un transmisor, un transceptor, un dispositivo para escribir datos en un medio legible por ordenador tal como, por ejemplo, una unidad óptica, una unidad de medios magnéticos (por ejemplo, una unidad de disquetes), un puerto de bus serie universal (USB), una interfaz de red u otra interfaz de salida. La interfaz de salida 32 emite el archivo de vídeo a un medio legible por ordenador, tal como, por ejemplo, una señal de transmisión, un medio magnético, un medio óptico, una memoria, una unidad flash u otro medio legible por ordenador.

[0054] La interfaz de red 54 puede recibir una unidad NAL o unidad de acceso por medio de la red 74 y proporcionar la unidad NAL o la unidad de acceso a la unidad de desencapsulación 50, por medio de la unidad de recuperación 52. La unidad de desencapsulación 50 puede desencapsular un elemento de un archivo de vídeo en flujos PES constituyentes, desempaquetar los flujos PES para recuperar los datos codificados y enviar los datos codificados al descodificador de audio 46 o al descodificador de vídeo 48, dependiendo de si los datos codificados forman parte de un flujo de audio o vídeo, por ejemplo, como se indica en las cabeceras de paquetes PES del flujo. El descodificador de audio 46 descodifica datos de audio codificados y envía los datos de audio descodificados a la salida de audio 42, mientras que el descodificador de vídeo 48 descodifica datos de vídeo codificados y envía los datos de vídeo descodificados, que pueden incluir una pluralidad de visualizaciones de un flujo, a la salida de vídeo 44.

[0055] Los diversos elementos del sistema 10 (por ejemplo, dispositivo cliente 40, dispositivo de preparación de contenido 20 y/o dispositivo servidor 60) pueden configurarse para realizar las diversas técnicas de esta divulgación. En general, el ejemplo del sistema 10 incluye varios elementos que pueden configurarse para iniciar el consumo de una presentación multimedia DASH basándose en el formato de archivo multimedia base ISO (ISO BMFF), o un archivo ISO BMFF segmentado, para el cual se entregan los segmentos a través del protocolo de entrega de objetos de transporte de codificación en capas (LCT), por ejemplo, el protocolo ROUTE sin la MPD (por ejemplo, archivo de manifiesto 66) y sin la entrega de metadatos asociados entre los identificadores de objetos de transporte (TOI) y las URL (como el FDT, el EFDT, o la cabecera de entidad en GFD o ROUTE) y la señalización de MPD se proporciona mediante otros medios, por ejemplo, la descripción de la instancia de sesión LCT (LSID) en ROUTE, el SDP en FLUTE, la cabecera LCT, el ISO BMFF IS, o similares. Es decir, el dispositivo cliente 40 puede iniciar el consumo realizando el inicio del servicio, que puede incluir el acceso a un punto de entrada del servicio.

[0056] Los elementos del sistema 10 (tales como el dispositivo cliente 40) también pueden configurarse (de forma adicional o alternativa) para consumir por completo un flujo unidireccional, es decir, una distribución de radiodifusión, bajo los supuestos de entrega anteriores.

[0057] Los elementos del sistema 10 también pueden configurarse (de forma adicional o alternativa) para iniciar una presentación multimedia DASH sin la MPD (por ejemplo, archivo de manifiesto 66), pero después del inicio y la reproducción inicial, la MPD puede recibirse y procesarse (por ejemplo, entregarse mediante el dispositivo servidor 60 y recibirse mediante el dispositivo cliente 40) con el fin de obtener información más precisa y combinarla con la entrega de banda ancha/unidifusión.

[0058] La MPD puede contener solo información absolutamente necesaria para la sintonización de la radiodifusión o el cambio de canal, donde se puede aplicar una o más de las siguientes:

- Cuando no se necesita ninguna información de MPD, la MPD puede estar vacía.
- Las copias regulares de MPD (con alguna información que no es absolutamente necesaria, por ejemplo, solo necesaria para algunas mejoras u optimizaciones) se incluyen escasamente en algunos puntos de acceso aleatorio (RAP), y entre dos copias regulares de MPD, una MPD ligera (con solo información absolutamente necesaria) se incluye en cada RAP

[0059] Los elementos del sistema 10 también pueden configurarse (de forma adicional o alternativa) para usar un tiempo de transmisión objetivo de cada paquete que se agrega al paquete ROUTE, reflejando un tiempo en el que el receptor está consumiendo (descodificando y renderizando) los datos relativos a otros datos en la misma sesión de ROUTE, en la que uno o más de los siguientes se aplican al tiempo de transmisión objetivo:

- Por el cual este tiempo se indica en la cabecera LCT.
- Por el cual se señala este tiempo en la cabecera CC.
- Por el cual este tiempo se indica en una cabecera de extensión.

- En el que este tiempo es un tiempo relativo.
- En el que este tiempo es un tiempo absoluto de reloj de pared, como un tiempo de protocolo de tiempo de red (NTP).
- En el que este tiempo es un tiempo relativo y se indica un tiempo de liberación en un paquete, y el paquete solo se libera para consumo cuando el tiempo de liberación es menor o igual que el tiempo de transmisión objetivo.

[0060] Los elementos del sistema 10 también pueden configurarse (de forma adicional o alternativa) para usar un indicador, agregado al paquete de ROUTE, que refleja si el receptor está consumiendo (descodificando y renderizando) los datos contenidos en el paquete cuando está presente, en el que pueden aplicarse uno o más de los siguientes:

- Cuando el indicador es igual a 1, el remitente guarda el paquete y no lo pasa al receptor para su consumo.
- Cuando el indicador es igual a 0, el remitente envía el paquete al receptor para su consumo.
- Se requiere que el valor del indicador sea igual a 1 para el último paquete de un objeto.

[0061] Por tanto, el sistema 10 puede implementar un diseño que sea un enfoque ventajoso por razones de eficiencia de ancho de banda, retardo de puesta en marcha inicial, simplicidad, robustez, extensibilidad y complejidad sin inconvenientes significativos.

[0062] De esta manera, y como se explica con mayor detalle a continuación, el dispositivo de servidor 60 representa un ejemplo de un dispositivo para enviar datos de medios que incluye una interfaz de red para enviar datos de una pluralidad de sesiones de transporte de codificación en capas (LCT), y un procesador configurado para construir una descripción de la instancia de sesión LCT (LSID) que incluye información representativa de una pluralidad de sesiones LCT, con cada una de las sesiones LCT que incluye datos de una respectiva de una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP (DASH), en la que la LSID indica correspondencias entre las sesiones LCT y las representaciones, genera la LSID a través de la interfaz de red y genera datos de las representaciones en las sesiones LCT correspondientes a través de la interfaz de red.

[0063] Asimismo, el dispositivo cliente 40 representa un ejemplo de un dispositivo cliente para recibir datos de medios que incluye uno o más descodificadores de medios configurados para descodificar datos de medios, una interfaz de red configurada para recibir una descripción de la instancia de sesión de transporte de codificación en capas (LCT) (LSID), donde LSID incluye información representativa de una pluralidad de sesiones LCT, con cada una de las sesiones LCT que incluye datos de una respectiva de una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP (DASH) y datos de una o más de las sesiones LCT, y un procesador configurado para iniciar el consumo de una o más de las representaciones de la presentación de medios DASH usando la LSID y sin usar un archivo de manifiesto para la presentación de medios DASH, en el que para iniciar el consumo, el procesador está configurado para recibir, a través de la interfaz de red, paquetes de las sesiones LCT que incluyen partes de datos de una o más de las representaciones, y proporcionar datos de los paquetes a uno o más descodificadores de medios.

[0064] La FIG. 2 es un diagrama conceptual que ilustra elementos del contenido multimedia 102 de ejemplo. El contenido multimedia 102 puede corresponder al contenido multimedia 64 (FIG. 1), o a otro contenido multimedia almacenado en el medio de almacenamiento 62. En el ejemplo de la FIG. 2, el contenido multimedia 102 incluye una descripción de presentación de medios (MPD) 104 y una pluralidad de representaciones 110A-110N. La representación 110A incluye datos de cabecera 112 y segmentos 114A a 114N (segmentos 114) opcionales, mientras que la representación 110N incluye datos de cabecera 122 y segmentos 124A- 124N (segmentos 124) opcionales. La letra N se usa para designar el último fragmento de película en cada una de las representaciones 110A, 110N por conveniencia. En algunos ejemplos, pueden existir diferentes números de fragmentos de película entre las representaciones 110A, 110N.

[0065] La MPD104 puede comprender una estructura de datos separada de las representaciones 110A-110N. La MPD104 puede corresponder al archivo de manifiesto 66 de la FIG. 1. Del mismo modo, las representaciones 110A-110N pueden corresponder a las representaciones 68 de la FIG. 1. En general, la MPD104 puede incluir datos que en general describen características de las representaciones 110A-110N, tales como características de codificación y representación, conjuntos de adaptación, un perfil al que corresponde la MPD104, información de tipo de texto, información de ángulo de cámara, información de calificación, la información de modo truco (por ejemplo, información indicativa de representaciones que incluyen subsecuencias temporales) y/o información para recuperar períodos remotos (por ejemplo, para inserción de publicidad dirigida en el contenido de medios durante la reproducción).

[0066] Los datos de cabecera 112, cuando están presentes, pueden describir características de los segmentos 114, por ejemplo, localizaciones temporales de puntos de acceso aleatorio (RAP, también denominados puntos de acceso de flujo (SAP)), cuáles de los segmentos 114 incluyen puntos de acceso aleatorio, desviaciones de bytes a puntos de acceso aleatorio dentro de los segmentos 114, localizadores de recursos uniformes (URL) de los segmentos 114 u otros aspectos de los segmentos 114. Los datos de cabecera 122, cuando están presentes, pueden describir características similares para los segmentos 124. Adicionalmente o de forma alternativa, dichas características pueden estar incluidas por completo dentro de la MPD104.

[0067] Los segmentos 114, 124 incluyen una o más muestras de vídeo codificadas, cada una de las cuales puede incluir tramas o fragmentos de datos de vídeo. Cada una de las muestras de vídeo codificadas de los segmentos 114 puede tener características similares, por ejemplo, requisitos de altura, anchura y ancho de banda. Dichas características pueden describirse por datos de la MPD104, aunque dichos datos no se ilustren en el ejemplo de la FIG. 2. La MPD104 puede incluir características como se describen en la especificación 3GPP, con la adición de cualquiera o toda la información señalizada descrita en esta divulgación.

[0068] Cada uno de los segmentos 114, 124 puede estar asociado a un único localizador de recursos uniforme (URL). Por tanto, cada uno de los segmentos 114, 124 puede ser independientemente recuperable usando un protocolo de red de transmisión en continuo, tal como DASH. De esta manera, un dispositivo de destino, tal como el dispositivo cliente 40, puede usar una petición GET HTTP para recuperar los segmentos 114 o 124. En algunos ejemplos, el dispositivo cliente 40 puede usar peticiones GET parciales HTTP para recuperar rangos de bytes específicos de los segmentos 114 o 124.

[0069] La FIG. 3 es un diagrama de bloques que ilustra elementos de un archivo de vídeo 150 de ejemplo, que puede corresponder a un segmento de una representación, tal como uno de los segmentos 114, 124 de la FIG. 2. Cada uno de los segmentos 114, 124 puede incluir datos que se ajustan sustancialmente a la disposición de datos ilustrada en el ejemplo de la FIG. 3. Se puede decir que el archivo de vídeo 150 encapsula un segmento. Como se ha descrito anteriormente, los archivos de vídeo, de acuerdo con el formato de archivo de medios de base ISO, y las ampliaciones del mismo, almacenan los datos en una serie de objetos, denominados "cuadros". En el ejemplo de la FIG. 3, el archivo de vídeo 150 incluye el cuadro de tipo de archivo (FTYP) 152, el cuadro de película (MOOV) 154, los cuadros de índices de segmento (sidx) 162, los cuadros de fragmento de película (MOOF) 164 y el cuadro de acceso aleatorio de fragmento de película (MFRA) 166. Aunque la FIG. 3 representa un ejemplo de archivo de vídeo, se deberá entender que otros archivos de medios pueden incluir otros tipos de datos de medios (por ejemplo, datos de audio, datos de texto temporizado o similares) que están estructurados de forma similar a los datos del archivo de vídeo 150, de acuerdo con el formato de archivo de medios de base ISO y sus ampliaciones.

[0070] El cuadro de tipo de archivo (FTYP) 152 describe en general un tipo de archivo para el archivo de vídeo 150. El cuadro de tipo de archivo 152 puede incluir datos que identifican una especificación que describe un mejor uso para el archivo de vídeo 150. El cuadro de tipo de archivo 152 puede estar colocada de forma alternativa antes del cuadro MOOV154, los cuadros de fragmento de película 164 y/o el cuadro MFRA166.

[0071] En algunos ejemplos, un segmento, tal como el archivo de vídeo 150, puede incluir un cuadro de actualización de MPD (no mostrada) antes del cuadro FTYP152. El cuadro de actualización de MPD puede incluir información que indica que se va a actualizar una MPD correspondiente a una representación que incluye el archivo de vídeo 150, junto con información para actualizar la MPD. Por ejemplo, el cuadro de actualización de MPD puede proporcionar un URI o URL para un recurso que se va a usar para actualizar la MPD. Como otro ejemplo, el cuadro de actualización de MPD puede incluir datos para actualizar la MPD. En algunos ejemplos, el cuadro de actualización de MPD puede seguir inmediatamente a un cuadro de tipo de segmento (STYP) (no mostrada) del archivo de vídeo 150, donde el cuadro STYP puede definir un tipo de segmento para el archivo de vídeo 150. La FIG. 7, analizada con mayor detalle a continuación, proporciona información adicional con respecto al cuadro de actualización de MPD.

[0072] El cuadro de MOOV154, en el ejemplo de la FIG. 3, incluye el cuadro de cabecera de película (MVHD) 156, el cuadro de pista (TRAK) 158 y uno o más cuadros de ampliación de película (MVEX) 160. En general, el cuadro de MVHD156 puede describir características generales del archivo de vídeo 150. Por ejemplo, el cuadro de MVHD156 puede incluir datos que describen cuándo se creó inicialmente el archivo de vídeo 150, cuándo se modificó por última vez el archivo de vídeo 150, una escala de tiempo para el archivo de vídeo 150, una duración de reproducción para el archivo de vídeo 150 u otros datos que describen en general el archivo de vídeo 150.

[0073] El cuadro de TRAK158 puede incluir datos para una pista del archivo de vídeo 150. El cuadro de TRAK158 puede incluir un cuadro de cabecera de pista (TKHD) que describe las características de la pista correspondiente al cuadro de TRAK158. En algunos ejemplos, el cuadro TRAK158 puede incluir imágenes de vídeo codificadas, mientras que, en otros ejemplos, los cuadros de vídeo codificado de la pista pueden estar incluidas en fragmentos de película 164, a las cuales se puede hacer referencia mediante unos datos del cuadro TRAK158 y/o los cuadros SIDX162.

[0074] En algunos ejemplos, el archivo de vídeo 150 puede incluir más de una pista. Por consiguiente, el cuadro

de MOOV154 puede incluir un número de cuadros de TRAK igual al número de pistas del archivo de vídeo 150. El cuadro de TRAK158 puede describir las características de una pista correspondiente del archivo de vídeo 150. Por ejemplo, el cuadro de TRAK158 puede describir información temporal y/o espacial para la pista correspondiente. Un cuadro de TRAK similar al cuadro de TRAK158 del cuadro MOOV154 puede describir características de una pista de conjunto de parámetros, cuando la unidad de encapsulación 30 (FIG. 2) incluye una pista de conjunto de parámetros en un archivo de vídeo, tal como el archivo de vídeo 150. La unidad de encapsulación 30 puede señalar la presencia de mensajes de SEI a nivel de secuencia en la pista de conjunto de parámetros dentro del cuadro de TRAK que describe la pista de conjunto de parámetros.

[0075] Los cuadros de MVEX160 pueden describir características de correspondientes fragmentos de película 164, por ejemplo, para señalar que el archivo de vídeo 150 incluye fragmentos de película 164, además de los datos de vídeo incluidos dentro del cuadro de MOOV154, si los hubiera. En el contexto de la transmisión continua de datos de vídeo, las imágenes de vídeo codificadas pueden estar incluidas en los fragmentos de película 164 en lugar de en el cuadro de MOOV154. En consecuencia, todas las muestras de vídeo codificadas pueden estar incluidas en fragmentos de película 164, en lugar de en el cuadro de MOOV154.

[0076] El cuadro de MOOV154 puede incluir un número de cuadros de MVEX160 igual al número de fragmentos de película 164 del archivo de vídeo 150. Cada una de los cuadros de MVEX160 puede describir características de un fragmento correspondiente de los fragmentos de película 164. Por ejemplo, cada cuadro de MVEX puede incluir un cuadro de cabecera de ampliación de película (MEHD) que describe una duración temporal para el cuadro correspondiente de los fragmentos de película 164.

[0077] Como se indica anteriormente, la unidad de encapsulación 30 puede almacenar un conjunto de datos de secuencia en una muestra de vídeo que no incluye datos de vídeo codificados reales. Una muestra de vídeo puede corresponder en general a una unidad de acceso, que es una representación de una imagen codificada en una instancia de tiempo específica. En el contexto de la AVC, la imagen codificada incluye una o más unidades NAL VCL que contienen la información para construir todos los píxeles de la unidad de acceso y otras unidades NAL no VCL asociadas, tales como mensajes de SEI. Por consiguiente, la unidad de encapsulación 30 puede incluir un conjunto de datos de secuencia, que puede incluir mensajes de SEI a nivel de secuencia, en uno de los fragmentos de película 164. La unidad de encapsulación 30 puede señalar, además, la presencia de un conjunto de datos de secuencia y/o de mensajes de SEI a nivel de secuencia con la presencia de estos en uno de los fragmentos de película 164 dentro de una de los cuadros de MVEX160 correspondiente al uno de los fragmentos de película 164.

[0078] Los cuadros de SIDX162 son elementos opcionales del archivo de vídeo 150. Es decir, los archivos de vídeo que se ajustan al formato de archivo 3GPP u otros formatos de archivo de este tipo, no incluyen necesariamente cuadros de SIDX162. De acuerdo con el ejemplo del formato de archivo 3GPP, se puede usar un cuadro de SIDX para identificar un subsegmento de un segmento (por ejemplo, un segmento contenido dentro del archivo de vídeo 150). El formato de archivo 3GPP define un subsegmento como "un conjunto autónomo de uno o más cuadros de fragmento de película consecutivos con un(os) cuadro(s) de datos de medios correspondiente(s), y un cuadro de datos de medios que contiene datos a los que se hace referencia mediante un cuadro de fragmento de película debe seguir a esos cuadro de fragmento de película y preceder al siguiente cuadro de fragmento de película que contiene información sobre la misma pista". El formato de archivo 3GPP también indica que un cuadro de SIDX "contiene una secuencia de referencias a subsegmentos del (sub)segmento documentado por el cuadro. Los subsegmentos a los que se hace referencia son contiguos en el tiempo de presentación. De forma similar, los bytes a los que un cuadro de índice de segmento hace referencia siempre son contiguos dentro del segmento. El tamaño al que se hace referencia da el recuento del número de bytes en el material al que se hace referencia".

[0079] Los cuadros de SIDX162 en general proporcionan información representativa de uno o más subsegmentos de un segmento incluido en el archivo de vídeo 150. Por ejemplo, dicha información puede incluir tiempos de reproducción en los que comienzan y/o terminan los subsegmentos, desviaciones de bytes para los subsegmentos, si los subsegmentos incluyen (por ejemplo, comienzan con) un punto de acceso de flujo (SAP), un tipo para el SAP (por ejemplo, si el SAP es una imagen de actualización de descodificador instantánea (IDR), una imagen de acceso aleatorio limpio (CRA), una imagen de acceso de enlace roto (BLA) o similares), una posición del SAP (en términos de tiempo de reproducción y/o desviación de bytes) en el subsegmento y similares.

[0080] Los fragmentos de película 164 pueden incluir una o más imágenes de vídeo codificadas. En algunos ejemplos, los fragmentos de película 164 pueden incluir uno o más grupos de imágenes (GOP), cada uno de los cuales puede incluir un número de imágenes de vídeo codificadas, por ejemplo, tramas o imágenes. Así mismo, como se describe anteriormente, los fragmentos de película 164 pueden incluir conjuntos de datos de secuencia en algunos ejemplos. Cada uno de los fragmentos de película 164 puede incluir un cuadro de cabecera de fragmento de película (MFHD, no mostrada en la FIG. 3). El cuadro MFHD puede describir características del fragmento de película correspondiente, tales como un número de secuencia para el fragmento de película. Los fragmentos de película 164 pueden estar incluidos por orden de número de secuencia en el archivo de vídeo 150.

[0081] El cuadro de MFRA166 puede describir puntos de acceso aleatorio dentro de fragmentos de película 164 del archivo de vídeo 150. Esto puede ayudar a realizar modos de avance, tales como realizar búsquedas hasta

ubicaciones temporales en particular (es decir, tiempos de reproducción) dentro de un segmento encapsulado por el archivo de vídeo 150. El cuadro de MFRA166 en general es opcional y no necesita estar incluida en los archivos de vídeo, en algunos ejemplos. Del mismo modo, un dispositivo cliente, tal como el dispositivo cliente 40, no tiene necesariamente que hacer referencia al cuadro de MFRA166 para descodificar y visualizar correctamente los datos de vídeo del archivo de vídeo 150. El cuadro MFRA166 puede incluir un número de cuadros de acceso aleatorio de fragmento de pista (TFRA) (no mostradas) igual al número de pistas del archivo de vídeo 150 o, en algunos ejemplos, igual al número de pistas de medios (por ejemplo, pistas sin indicaciones) del archivo de vídeo 150.

[0082] En algunos ejemplos, los fragmentos de película 164 pueden incluir uno o más puntos de acceso de flujo (SAP), tales como imágenes IDR. Del mismo modo, el cuadro de MFRA166 puede proporcionar indicaciones de ubicaciones dentro del archivo de vídeo 150 de los SAP. En consecuencia, se puede formar una subsecuencia temporal del archivo de vídeo 150 a partir de los SAP del archivo de vídeo 150. La subsecuencia temporal también puede incluir otras imágenes, tales como tramas P y/o tramas B que dependen de los SAP. Las tramas y/o fragmentos de la subsecuencia temporal pueden estar dispuestos dentro de los segmentos de modo que las tramas/fragmentos de la subsecuencia temporal que dependen de otras tramas/fragmentos de la subsecuencia pueden descodificarse apropiadamente. Por ejemplo, en la disposición jerárquica de los datos, los datos usados para la predicción de otros datos también pueden estar incluidos en la subsecuencia temporal.

[0083] La FIG. 4 es un diagrama conceptual que ilustra un escenario de ejemplo que a menudo surge en la entrega de radiodifusión. El escenario se puede describir de la forma siguiente:

- Se distribuyen varios servicios (típicamente en diferentes canales).
- Los servicios tienen componentes de radiodifusión y pueden tener componentes de unidifusión/banda ancha adicionales.
- Típicamente, los servicios incluyen diferentes componentes, por ejemplo, vídeo, audio, subtítulos, audio alternativo, vídeo alternativo, etc., que se describen mediante metadatos.
- En la radiodifusión, cada componente puede enviarse en una sesión de transporte diferente, de modo que los componentes se pueden seleccionar/filtrar.
- Los componentes y servicios se multiplexan a nivel de paquete.

[0084] En el ejemplo de la FIG. 4, hay dos estaciones de radiodifusión de medios separadas 180A, 180B. La estación 180A proporciona datos de arranque 182, incluida la LSID184 y los fragmentos de señalización relacionados 186A-186N (fragmentos de señalización 186). La estación 180A también proporciona datos de audio 188 (incluidos los datos de audio 188A y 188B mostrados en la FIG.4), datos de vídeo 190 (incluidos los datos de vídeo 190A y 190B en la FIG.4), datos de texto temporizados 192, datos de capa de mejora de vídeo 194 (incluidos los datos de capa de mejora de vídeo 194A y 194B mostrados en la FIG. 4), y los datos de lenguaje de audio alternativo 196 (incluyendo los datos de lenguaje de audio alternativo 196A y 196B mostrados en la FIG. 4). En un ejemplo, la estación 180A proporciona datos de arranque 182 a través del protocolo de datagramas de usuario (UDP) usando el Identificador de sesión de transporte (TSI) 0, datos de audio 188 a través de UDP usando TSI1, datos de vídeo 190 a través de UDP usando TSI2, datos de capa de mejora de vídeo 194 a través de banda ancha y datos de lenguajes de audio alternativos 196 a través de HTTP.

[0085] La estación 180B, en este ejemplo, proporciona datos de arranque 200, incluyendo LSID202 y fragmentos de señalización relacionados 204A-204N (fragmentos de señalización 204). La estación 180B también proporciona datos de audio 206 (incluidos los datos de audio 206A y 206B mostrados en la FIG. 4), los datos de vídeo 208 (incluidos los datos de vídeo 208A y 208B en la FIG. 4) y los datos de audio de accesibilidad 210 (incluidos los datos de lenguaje de audio de accesibilidad 210A y 210B mostrados en la FIG. 4). En un ejemplo, la estación 180A proporciona datos de arranque 182 a través del protocolo de datagramas de usuario (UDP) usando el Identificador de sesión de transporte (TSI) 0, datos de audio 188 a través de UDP usando TSI1, datos de vídeo 190 a través de UDP usando TSI2, datos de capa de mejora de vídeo 194 a través de banda ancha y datos de lenguaje de audio alternativo 196 a través de unidifusión (por ejemplo, de acuerdo con HTTP).

[0086] En general, las estaciones 180A, 180B pueden incluir dispositivos similares al dispositivo de preparación de contenido 20 y al dispositivo de servidor 60 de la FIG. 1. En particular, las estaciones 180A, 180B pueden incluir dispositivos separados entre sí que realizan una funcionalidad similar a la atribuida al dispositivo de preparación de contenido 20 y/o al dispositivo servidor 60. De forma adicional o alternativa, las estaciones 180A, 180B pueden preparar datos de medios y proporcionar los datos de medios preparados a una red de entrega de contenido (CDN), afiliadas de estaciones locales, compañías de cable u otros distribuidores de contenido, que pueden distribuir los datos de medios a través de radiodifusión de red o unidifusión o transmisión por aire.

[0087] Aunque no se muestra en la FIG. 4, los flujos de datos (es decir, datos de arranque 182, 200, datos de audio 188, 206, datos de vídeo 190, 208, datos de texto temporizados 192, datos de capa de mejora de vídeo 194,

datos de lenguajes de audio alternativos 196 y audio de accesibilidad 210) pueden organizarse en segmentos, como segmentos de inicialización (IS), incluidos datos de configuración estáticos y una secuencia de segmentos de medios que contienen muestras codificadas de datos de medios. Los segmentos pueden contener suficiente información para la temporización, de modo que los datos se pueden descodificar y presentar junto con otras representaciones para una reproducción sincronizada. Por ejemplo, los segmentos de los datos de audio 188 y los datos de vídeo 190 pueden incluir datos de sincronización de manera que el audio y el vídeo se sincronizan durante la reproducción.

[0088] Los segmentos siguen reglas específicas de acuerdo con, por ejemplo, DASH. Es decir, los segmentos son objetos de datos, y el dispositivo de preparación de contenido 20 o el dispositivo servidor 60 pueden generar segmentos independientemente de otros segmentos. Los segmentos siguen en orden de descodificación, es decir, todos los datos contenidos en segmentos con un número más bajo tienen tiempos de descodificación más bajos (es decir, antes) que los segmentos con números de segmento más altos. Además, el contenido puede empalmarse en determinadas posiciones. Esto da como resultado un nuevo período, para el cual cada representación recibirá un nuevo IS. Hay otros casos en los que se puede proporcionar un nuevo IS, pero la cronología es continua. Este puede ser el caso, por ejemplo, si el IS incluye nuevos metadatos.

[0089] La FIG. 5 es un diagrama de bloques que ilustra componentes de ejemplo de la unidad de recuperación 52 del dispositivo cliente 40 de la FIG. 1. En este ejemplo, la unidad de recuperación 52 incluye la unidad de multidifusión IP220, la unidad de configuración estática 222, la unidad de recepción de ROUTE 224, la memoria caché 226, la unidad de servidor proxy 228 y el cliente DASH 230. En general, el cliente DASH 230 puede recuperar segmentos de una o más representaciones a través de la red 74 (por ejemplo, desde el dispositivo servidor 60) usando peticiones de unidifusión. De forma adicional o alternativa, la unidad de recuperación 52 puede utilizar la unidad de multidifusión IP 220 para suscribirse a una multidifusión IP, para recibir datos de multidifusión IP que incluyen segmentos de una o más representaciones. La unidad 220 de multidifusión IP puede pasar los segmentos 232 recibidos a la unidad 224 de recepción de ROUTE, que puede almacenar los segmentos recibidos en la memoria caché 226. El cliente DASH 230 puede entonces solicitar los segmentos de la unidad 228 de servidor proxy. De esta manera, el cliente DASH 230 puede solicitar segmentos desde una dirección IP del dispositivo servidor 60, o desde una dirección de dispositivo principal del dispositivo cliente 40 que incluye la unidad de recuperación 52. En cualquier caso, el cliente DASH 230 puede recuperar los segmentos usando peticiones de GET HTTP o GET parcial. Por tanto, la unidad de recuperación 52 en general puede recibir segmentos mediante unidifusión o multidifusión IP.

[0090] Si los datos están disponibles a través de unidifusión, entonces los datos disponibles se describen en el archivo de manifiesto (por ejemplo, la descripción de presentación de medios (MPD)) utilizando conceptos DASH como Presentación de medios, períodos, conjuntos de adaptación, representaciones y segmentos. En implementaciones básicas para la entrega de radiodifusión unidireccional de presentaciones de medios DASH, se propone distribuir la MPD y los segmentos como objetos de datos regulares utilizando un protocolo de entrega de objetos como ROUTE, FLUTE o FCAST. El receptor de objetos puede entonces usar los objetos recuperados para proporcionar una presentación de medios desde un servidor local a un cliente DASH como se muestra en las FIGS. 1 y 5 (usando ROUTE como ejemplo, pero estas técnicas también son aplicables a FLUTE o FCAST). Los segmentos se distribuyen de manera que las URL en la MPD se asignan a los objetos entregados en el protocolo de entrega de objetos, por ejemplo, a través de una tabla de entrega de archivos (FDT), una FDT extendida (EFDT) o como parte del objeto mediante cabeceras de entidad HTTP.

[0091] El remitente (por ejemplo, el dispositivo servidor 60 o un dispositivo similar) entrega los objetos de manera que los segmentos se reciben en el dispositivo cliente 40 antes de los respectivos tiempos de disponibilidad de los segmentos anunciados en la MPD. El dispositivo de servidor 60 (o el dispositivo de preparación de contenido 20) empaqueta los segmentos y multiplexa los paquetes en consecuencia. Los objetos pueden estar separados por diferentes identificadores de objetos de transporte (TOI) y la asignación a las URL se proporciona como se mencionó anteriormente. En una configuración de envío más sofisticada, todos los segmentos de una representación se entregan en una sesión dedicada, por ejemplo, utilizando un puerto de protocolo de datagramas de usuario (UDP) dedicado o una sesión de transporte de codificación en capas (LCT) dedicada identificada por un identificador de sesión de transporte único (TSI). Esta asignación permite al receptor seleccionar o ignorar sesiones enteras basándose en la selección de representaciones por parte del cliente DASH 230.

[0092] Sin embargo, el suministro de una presentación multimedia a través de la radiodifusión puede crear algunos desafíos para el remitente, ya que el remitente puede necesitar predecir el tiempo de entrega y procesamiento para establecer correctamente los tiempos de disponibilidad del segmento en la MPD. Si, por alguna razón, el sistema de distribución tiene más o menos retardo de lo esperado, el rendimiento del cliente DASH 230 u otro participante puede verse afectado negativamente. El inicio de la reproducción puede ser demasiado pronto o demasiado tarde. La reproducción que se inicia demasiado pronto puede provocar un bloqueo de la reproducción de medios y la reproducción que se inicia demasiado tarde puede dar como resultado un cambio de canal más lento y una latencia de extremo a extremo aumentada de lo que sería posible de otra manera. Otro problema es que, antes de que el cliente DASH 230 haga uso de los segmentos y los medios contenidos, el cliente DASH 230 puede necesitar recibir el FDT (o una estructura de datos equivalente) y la MPD. Para los puntos de acceso

aleatorio frecuentes, esto puede aumentar la sobrecarga de los metadatos.

[0093] Como se analizó con mayor detalle anteriormente y posteriormente, respectivamente, las FIGS. 1 y 6 ilustran ejemplos de infraestructuras de envío. Más particularmente, la FIG. 6 es un diagrama de bloques que ilustra un ejemplo de una arquitectura de emisor básica 240 que proporciona diferentes opciones de envío de acuerdo con las técnicas de esta divulgación. En este ejemplo, la arquitectura del remitente 240 incluye codificador de medios fuera de línea de velocidad de transmisión de bits múltiple 244, codificador de medios en vivo de velocidad de transmisión de bits múltiple 248, unidad de inserción de publicidad (ad) 252, unidad de autoría fuera de línea DASH 250, unidad de autoría en vivo DASH 256, servidor de emulación en vivo HTTP 258, servidor activo HTTP 260 y servidor ROUTE 262. En este ejemplo, el codificador de medios fuera de línea de velocidad de transmisión de bits múltiple 244 codifica uno o más conjuntos de contenido de fuente de audio/vídeo (A/V) 242, mientras que el codificador de medios en vivo de velocidad de transmisión de bits múltiple 248 codifica uno o ambos contenidos de fuente de A/V 246A y A/V contenido en vivo 246B. El contenido de fuente A/V 246A puede representar contenido pregrabado, mientras que el contenido A/V en vivo 246B puede representar contenido en vivo que se captura sobre la marcha, por ejemplo, un evento de noticias en vivo, un evento deportivo u otro evento en vivo de este tipo.

[0094] El codificador de medios fuera de línea de velocidad de transmisión de bits múltiple 244 proporciona contenido multimedia codificado a la unidad de autoría fuera de línea DASH 250. La unidad de autoría fuera de línea de DASH 250 puede preparar en general datos de medios codificados para su transporte a través de, por ejemplo, DASH. Por ejemplo, para datos de vídeo, la unidad de autoría fuera de línea DASH 250 puede preparar segmentos que incluyen un conjunto de tramas de vídeo codificadas, que pueden incluir una trama de vídeo de punto de acceso aleatorio (RAP). La unidad de inserción de anuncios 252 puede preparar contenido publicitario para su inserción en puntos apropiados de un flujo de medios preparado por el servidor de emulación en vivo HTTP 258. Asimismo, la unidad de autoría fuera de línea de DASH 250 puede proporcionar segmentos preparados de una o más representaciones de medios al servidor de emulación en vivo HTTP 258.

[0095] En general, el servidor de emulación en vivo HTTP 258 emula un servidor de medios en vivo, como el servidor en vivo HTTP 260. Por ejemplo, el servidor de emulación en vivo HTTP 258 puede indicar tiempos de disponibilidad de segmento para segmentos de diversas representaciones (por ejemplo, representaciones de audio y vídeo, así como representaciones de texto temporizadas en algunos ejemplos). El servidor de emulación en vivo HTTP 258 puede proporcionar los segmentos al servidor ROUTE 262, que puede emitir los segmentos en momentos particulares a través de ROUTE usando multidifusión IP, basándose en la información de pistas y temporización 254. De forma adicional o alternativa, el servidor ROUTE 262 puede recibir contenido DASH en vivo desde el servidor en vivo HTTP 260, que recibe contenido DASH de la unidad de autoría en vivo DASH 256, que puede preparar el contenido DASH en vivo a partir de datos de medios codificados recibidos desde el codificador 248 de medios en vivo de velocidad de transmisión de bits múltiple. Aunque en este ejemplo se muestra un único codificador 248 de medios en vivo de velocidad de transmisión de bits múltiple, debe entenderse que se pueden emplear una pluralidad de codificadores para codificar datos de medios en vivo, en algunos ejemplos.

[0096] La arquitectura del remitente 240 puede proporcionar contenido a petición como un servicio pseudo-en vivo y/o un servicio en vivo generado usando DASH. Además, la arquitectura del remitente 240 puede incluir herramientas que soporten la inserción de anuncios (anuncios), como la unidad de inserción de anuncios 252. Otra opción es agregar algo de robustez a DASH, por ejemplo, enviando múltiples períodos para permitir la resincronización usando el nivel MPD y/o el nivel de segmento. El contenido puede estar disponible a través de un servicio HTTP de modo que los clientes que están conectados a través de unidifusión (por ejemplo, a través de la red 74) puedan acceder a todo o partes del servicio a través de unidifusión. Un caso de uso de ejemplo es la distribución a través de un protocolo de entrega de archivos. Es decir, la arquitectura del remitente 240 puede entregar los objetos DASH generados (MPD, segmentos u otros objetos de datos) a través de un protocolo de entrega de archivos, por ejemplo, FLUTE, FCAST, transporte de medios MPEG (MMT) entrega de archivos genérica (GFD) o ROUTE. En el caso de la transmisión en vivo, el uso de ROUTE puede ser preferente, ya que ROUTE soporta capacidades en tiempo real y asigna los objetos de manera apropiada a los paquetes entregados a través de multidifusión IP y potencialmente a continuación a través de una capa física de radiodifusión como DVB-T2 con encapsulación de flujo genérico (GSE) o una tecnología definida por ATSC3.0.

[0097] Típicamente, para el servicio de multidifusión de radiodifusión multimedia (MBMS), MBMS mejorado (eMBMS) y ATSC3.0, se considera un enfoque de arriba hacia abajo para la entrada del servicio:

- Puede haber un punto de entrada de servicio, por ejemplo, una página HTML-5 que contiene una MPD o una MPD en sí.
- El servicio de transmisión puede ser descrito por una MPD que contiene múltiples componentes de medios, cada uno documentado en uno o más conjuntos de adaptación para una selección adecuada, incluidos metadatos, clasificación, accesibilidad, etc.
- Cada conjunto de adaptación puede contener una o varias representaciones.

- Algunas representaciones están disponibles en radiodifusión y crean un servicio básico.
 - Otras están disponibles en unidifusión para enriquecer y mejorar la presentación.
- 5 • Para arrancar el servicio, la MPD es necesaria y la MPD controla el tiempo y la reproducción.

[0098] Se han considerado como datos de acceso aleatorio los siguientes:

- Para DASH sobre FLUTE como se define en MBMS:

10 ○ Se supone que ciertos fragmentos de metadatos, a saber, el paquete de descripción de servicios de usuario (USDB) y el protocolo de descripción de sesión (SDP), son estáticos y están en la memoria caché.

15 ○ Los datos que deben adquirirse en tiempo real son los siguientes: Instancia FDT, MPD, IS y segmento de medios comenzando con un punto de acceso aleatorio (RAP) de grupo cerrado de imágenes (GOP). El requisito de RAP de GOP cerrado es una restricción de las implementaciones en la actualidad. La recopilación de datos necesarios se denomina grupo.

20 ○ El primer paquete del FDT es el primer paquete de un grupo, es decir, este paquete debe ser recibido y a continuación los siguientes paquetes restantes que están asociados a este grupo.

25 ○ Para poder entregar los datos al cliente DASH para su reproducción, es necesario recibir al menos un segmento completo, es decir, solo después de la recepción completa el segmento está "disponible" y se puede programar para que lo solicite el cliente DASH. El cliente DASH a continuación progresará y programará la reproducción del segmento generando un búfer apropiado o siguiendo las instrucciones del retardo de presentación sugerido en la MPD.

 ○ Tenga en cuenta que este procedimiento es equivalente al DASH basado en segmentos sobre ROUTE.

30 **[0099]** La FIG. 7 es un diagrama conceptual que ilustra ejemplos de diferentes aspectos de la entrada de servicios en el ejemplo de DASH sobre FLUTE. En particular, la FIG. 7 ilustra una entrada de servicio 280 de ejemplo, en la que la dirección y el puerto del protocolo de Internet (IP) junto con la referencia de URL MPD pueden usarse para recuperar datos. En particular, la entrada de servicio 280 ilustra que un dispositivo cliente, como el dispositivo cliente 40, puede usar la tabla de entrega de archivos (FDT) 288 para recuperar MPD 284, y a continuación usar FDT 288 y MPD 284 para recuperar el segmento de inicialización 282 y el segmento de medios 286.

35 **[0100]** La FIG. 7 también ilustra la dependencia de procesamiento 290 entre estos diversos elementos. En particular, FDT 288 es independiente, MPD 284 depende de FDT 288, el segmento de inicialización 282 depende de FDT 288 y MPD 284, y el segmento de medios 286 depende de cada uno de FDT 288, MPD 284 y el segmento de inicialización 282.

40 **[0101]** La FIG. 7 ilustra además una orden de envío típica 292 de izquierda a derecha. Debido a la dependencia de procesamiento 290, de acuerdo con el orden de envío típico 292, el FDT 288 se envía antes de la MPD 284, que se envía antes del segmento de inicialización 282, que se envía antes de uno o más paquetes de segmento de medios 286A, 286B (que forman parte del segmento de medios 286).

- Para DASH sobre ROUTE:

50 ○ En este caso, se asume que la descripción de la instancia de sesión LCT (LSID) está disponible en la memoria caché de un dispositivo receptor, es decir, el receptor tiene información sobre las sesiones programadas y lo que se asigna a cada una de las sesiones de transporte LCT.

55 ○ Los datos a adquirir en tiempo real pueden incluir los siguientes: Instancia EFDT, MPD, IS y segmento de medios comenzando con un RAP de GOP cerrado. El requisito de RAP de GOP cerrado es una restricción de las implementaciones en la actualidad.

 La recopilación de datos necesarios se denomina grupo. En la discusión de esta divulgación, se propone la eliminación del EFDT y la MPD.

60 ○ El EFDT es el primer paquete de un grupo, es decir, es necesario recibir este paquete y a continuación los siguientes paquetes restantes que están asociados a este grupo.

65 ○ En un modo de funcionamiento, sería factible el modo de recepción basado en segmentos como se analizó anteriormente. Sin embargo, dado que ROUTE soporta la entrega basada en eventos de entrega de medios (MDE), el cliente DASH y/o el motor de procesamiento ISO BMFF pueden iniciar la reproducción antes en el momento en que se recibe un prefijo suficientemente grande del segmento de

medios. En este caso, para poder entregar los datos al cliente DASH para su reproducción, el MDE debe estar disponible y puede programarse para que lo solicite el cliente DASH. A continuación, el cliente DASH progresará y programará la reproducción del prefijo de los segmentos de medios para acelerar el retardo inicial. Lo que es relevante y debe tenerse en cuenta es la forma en que la disponibilidad y los tiempos de descodificación (en general idénticos) se obtienen correctamente.

[0102] La FIG. 8 es un diagrama conceptual que muestra otro conjunto de aspectos de ejemplo de una entrada de servicio 300 en un ejemplo de entrega de datos DASH sobre ROUTE. En este ejemplo, un FDT extendido (EFDT) 302 contiene una asignación de TOI a URL y el tipo de contenido para MPD306, el segmento de inicialización 304 y el segmento de medios 310 (formado por el evento de entrega de medios del segmento de medios (MDE) 308A y el resto del segmento de medios 308B). Desde un punto de visualización de procesamiento (es decir, como se muestra en la dependencia de procesamiento 312), el acceso al segmento de medios 310 con los datos de medios depende del segmento de inicialización 304, MPD306 y EFDT302. El segmento de inicialización 304 depende del procesamiento de MPD306 y EFDT302. Y el procesamiento de MPD306 depende del procesamiento de EFDT302. Para permitir tal procesamiento, la orden de envío 314 de paquetes puede ser como se muestra en la FIG. 7, es decir, EFDT302, a continuación MPD306, a continuación el segmento de inicialización 304 y a continuación los paquetes de segmento de medios (por ejemplo, el segmento de medios MDE308A, seguido por el resto del segmento de medios 308B).

[0103] La recepción específica de ROUTE se ha documentado de la siguiente manera:

- La LSID describe la combinación de IP/puerto y diferentes sesiones de LCT.
- Los diferentes componentes de medios representados en archivos ISO BMFF segmentados están asociados y asignados de forma única a una combinación de IP/puerto/TSI.
- El servicio se describe mediante una URL MPD
- El receptor ROUTE recupera objetos para un componente/servicio del canal.
- En la recepción de los objetos, el receptor ROUTE asigna una URL a cada objeto usando el FDT/EFDT (o posiblemente el modo entidad).
- El cliente DASH obtiene la URL de MPD, la obtiene de la memoria caché local y comienza a consumir los segmentos basándose en las URL en la MPD.
- El tiempo es controlado por el cliente MPD y DASH: los segmentos no tienen tiempo asignado.

[0104] En un enfoque alternativo, se puede aplicar lo siguiente:

- La señalización de la capa inferior proporciona información suficiente para iniciar el servicio sin la MPD.
- Si se agrega unidifusión o si es necesaria una selección más precisa y se consulta la MPD, la MPD aún puede ser una MPD unificado, pero se considera que no es necesario para el inicio. La MPD puede incluso configurarse de manera que para solo radiodifusión, no se utilice MPD y, además, no sea necesaria la vinculación de URL. De esta manera, estas técnicas pueden evitar la necesidad de FDT, EFDT u otros medios para vincular URL a objetos.

[0105] Esta divulgación describe un diseño de ejemplo similar a este enfoque, que puede proporcionar ventajas por razones de eficiencia del ancho de banda, puesta en marcha inicial, simplicidad, robustez, extensibilidad y complejidad, sin desventajas significativas. Para mantener la compatibilidad con los servicios existentes y abordar diferentes casos de uso, se puede considerar el siguiente enfoque:

- La señalización del servicio proporciona una indicación sobre cuál de los siguientes enfoques se pueden tomar:
 1. El receptor necesita la MPD incluso para el inicio, es decir, el cliente DASH no debe iniciarse y el inicio del servicio no debe realizarse sin la MPD. Esto básicamente replicaría el enfoque actual.
 2. El proveedor de servicios proporciona suficiente señalización para que el inicio sin la MPD sea posible y esté permitido, pero la MPD también se proporciona y describe ofertas y alternativas de contenido más preciso.
 3. El proveedor de servicios no proporciona ningún MPD, el servicio está completamente descrito por la señalización de la capa inferior y no se habilita ninguna información enriquecida. Esto evitaría un servicio híbrido de radiodifusión/banda ancha.

- Servicios de ejemplo para los diferentes casos:

1. Servicio encriptado con información detallada en MPD, utilizando el primer enfoque.
2. Servicio A/V gratuito con componentes de unidifusión que pueden enriquecer el servicio mediante el segundo enfoque.
3. Servicio A/V en abierto simple sin necesidad de MPD, utilizando el tercer enfoque.

[0106] Los siguientes problemas pueden ocurrir con el enfoque de arriba hacia abajo que distribuye la MPD como punto de entrada:

- Los formatos de medios DASH se distribuyen a través de una distribución de radiodifusión.
- Para iniciar la sesión, es necesario recibir varios objetos de datos (señalización de capa inferior, descripción de la sesión, FDT/EFDT/procedimientos equivalentes, MPD, IS y al menos una parte del segmento de medios) para acceder aleatoriamente a los datos y para programar la reproducción.
- La selección de Medios y el control de tiempo está en la MPD, por lo que se requiere recibir la MPD antes de la puesta en marcha y recibir los metadatos para identificar la MPD.
- Otro problema es que la MPD, si no se modifica, debe generarse en el remitente de manera que se pueda predecir la sincronización en el receptor.
- Otro problema es que la MPD debe enviarse con cada punto de acceso aleatorio en cada componente para acceder rápidamente al vídeo o audio.
- Otro problema más es que se necesita FDT o EFDT para poder asignar las URL.
- Por último, la MPD y los metadatos suelen formatearse de acuerdo con XML, y la MPD unificada describe todo el servicio, incluidos todos los componentes, así como todos los datos entregados de unidifusión/banda ancha. Esto puede hacer que la MPD sea innecesariamente grande, mientras que al principio solo se puede requerir la información de radiodifusión.

[0107] Las técnicas de esta divulgación pueden superar uno o todos los problemas analizados anteriormente.

[0108] Por ejemplo, un dispositivo receptor (por ejemplo, el dispositivo cliente 40) puede configurarse para funcionar (es decir, recibir, procesar y decodificar datos de medios) sin EFDT, FDT y/o MPD en un modo unidireccional/de radiodifusión. En particular, el dispositivo cliente 40 puede recibir versiones alternativas de la información de estas estructuras de datos que de otro modo serían necesarias para la puesta en marcha y/o el funcionamiento permanente mediante otros medios.

[0109] La FIG. 9 es un diagrama conceptual que ilustra los campos de la cabecera LCT320 de acuerdo con RFC5651 que pueden usarse para transportar datos de acuerdo con las técnicas de esta divulgación. En particular, la cabecera 320 de LCT incluye el campo de versión 322, indicador de control de congestión 324, campo de indicación específica de protocolo (PSI) 326, indicador de identificador de sesión de transporte (S) 328, indicador de identificador de objeto de transporte (O) 330, indicador de media palabra (H) 332, campo reservado (RES) 334, indicador de cerrar sesión (A) 336, indicador de cerrar objeto (B) 338, campo de longitud de cabecera LCT (HDR_LEN) 340, campo de punto de código 342, información de control de congestión 344, identificador de sesión de transporte 346, identificadores de objeto de transporte 348 y campo de extensiones de cabecera 350.

[0110] Los valores del campo de versión 322, indicador de control de congestión 324, campo de indicación específica de protocolo (PSI) 326, indicador de identificador de sesión de transporte (S) 328, indicador de identificador de objeto de transporte (O) 330, indicador de media palabra (H) 332, campo reservado (RES) 334, indicador de cerrar sesión (A) 336, indicador de cerrar objeto (B) 338, campo de longitud de cabecera LCT (HDR_LEN) 340, campo de punto de código 342, información de control de congestión 344, identificador de sesión de transporte 346 e identificador de objeto de transporte 348 puede establecerse de acuerdo con RFC5651, y de acuerdo con las técnicas de esta divulgación, como se describe a continuación. El campo 350 de extensión de cabecera puede establecerse de acuerdo con las técnicas de esta divulgación.

[0111] Las siguientes funcionalidades están disponibles y posiblemente se pueden utilizar para soportar la entrega de información que de otro modo se proporciona en la MPD y/o EFDT. Los campos que se pueden usar para transportar datos de acuerdo con las técnicas de esta divulgación están marcados en cursiva a continuación:

- *LSID* (como se define en el borrador de la especificación ROUTE de ATSC3.0)

○ *Uso del descriptor de contenido en la LSID para señalar la asignación de propiedades del componente a la sesión de transporte con el fin de seleccionar la sesión de transporte que se entregará a la capa de medios.*

○ *Uso de la asignación de puntos de código*

• **Campos de cabecera ROUTE/LCT (consulte RFC5651 y FIG.9)**

○ *Información de control de congestión (CCI) 344:32, 64, 96 o 128 bits. Se utiliza para transportar información de control de congestión. Por ejemplo, la información de control de congestión podría incluir números de capa, números de canal lógico y números de secuencia. Este campo es opaco a los efectos de esta especificación, por lo que puede utilizarse de forma arbitraria.*

○ *TSI: Identificador de sesión de transporte (TSI) 346:0, 16, 32 o 48 bits. El TSI identifica de forma única una sesión entre todas las sesiones de un remitente en particular. El TSI tiene el alcance de la dirección IP del remitente y, por lo tanto, la dirección IP del remitente y el TSI identifican de manera única la sesión. Aunque un TSI junto con la dirección IP del remitente siempre identifica de forma única una sesión, si el TSI se incluye o no en la cabecera LCT depende de lo que se utilice como valor de TSI. Si el transporte subyacente es UDP, entonces el número de puerto de origen UDP de 16 bits PUEDE servir como TSI para la sesión.*

○ *Identificador de objeto de transporte (TOI) 348:0, 16, 32, 48, 64, 80, 96 o 112 bits. Este campo indica a qué objeto dentro de la sesión pertenece este paquete. Por ejemplo, un remitente puede enviar varios archivos en la misma sesión, usando TOI = 0 para el primer archivo, TOI = 1 para el segundo, etc. Como otro ejemplo, el TOI puede ser un identificador global único del objeto. que se transmite de varios remitentes al mismo tiempo, y el valor TOI puede ser el resultado de una función hash aplicada al objeto.*

○ *Punto de código 342: hay 8 bits disponibles para señalar diferentes propiedades del objeto contenido, así como la relación con el paquete actual. Un identificador opaco que se pasa a la decodificación de la carga útil del paquete para transmitir información sobre el códec que se utiliza para la carga útil del paquete. La asignación entre el punto de código y el códec real se define por sesión y se comunica fuera de banda como parte de la información de descripción de la sesión. El uso del campo CP es similar al campo Tipo de carga útil (PT) en las cabeceras RTP como se describe en [RFC3550].*

○ *Indicadores de cabecera específicos en la cabecera LCT*

▪ *PSI 326: El uso de estos bits, si los hay, es específico para cada instanciación de protocolo que usa el bloque de construcción LCT. Si no se define un uso específico de instanciación de protocolo de estos bits, entonces un remitente DEBE ponerlos a cero y un receptor DEBE ignorar estos bits.*

▪ *RES334: propiedad de LCT, por lo que no se utiliza actualmente*

• **Cabecera de extensión ROUTE/LCT (consulte RFC5651 y FIG. 9 ¡Error! Fuente de referencia no encontrada.).**

○ *Las extensiones de cabecera 350 pueden usarse en LCT para contener campos de cabecera opcionales que no siempre se usan o tienen un tamaño variable. Entre los ejemplos del uso de extensiones de cabecera se incluyen:*

▪ *Versiones de tamaño extendido de campos de cabecera ya existentes.*

▪ *Información de autenticación del remitente y el destinatario.*

▪ *Transmisión de información de tiempos.*

• **Segmento de inicialización (cabecera de película) de ISO BMFF (consulte ISO/IEC14496-12)**

○ *Los metadatos de una presentación se almacenan en el único cuadro de película que se encuentra en el nivel superior de un archivo. La cabecera de película permite agregar información adicional relacionada con los aspectos específicos de los medios del componente, por ejemplo:*

▪ *cabecera de medios, información general sobre los medios*

- manejador, declara el tipo (manejador) de medio
- contenedor de información multimedia

• Fuera de ROUTE, LCT e ISO BMFF

○ información sobre la capa física (FIT)

○ información en la capa de presentación/aplicación que ejecuta el servicio.

[0112] Las funcionalidades relevantes de la MPD se resumen a continuación y cómo abordarlas de acuerdo con las técnicas de esta divulgación se discuten con mayor detalle a continuación.

[0113] Tiempos de disponibilidad y anclaje del tiempo de presentación: El tiempo de disponibilidad se señala de manera que el tiempo de disponibilidad del objeto o la parte MDE del objeto se indica en la cabecera. El anclaje del tiempo de presentación es tal que cuando los datos están disponibles para el cliente ISO BMFF, el cliente comienza a descodificar y ejecutar la reproducción inmediatamente. Los detalles de la temporización se analizan a continuación.

[0114] Tipo de presentación-perfil, estática, dinámica, etc.: No es necesario configurar estos parámetros para la radiodifusión, pero los datos pueden seguir ciertas reglas. Se puede definir un perfil. El tipo se considera estático. El perfil detallado y la definición de formato de medios se proporcionan a continuación.

[0115] Descripciones de ancho de banda y búfer: Estos parámetros no son relevantes para la distribución de radiodifusión, ya que la sincronización está determinada por la llegada de los datos. Sin embargo, con el fin de iniciar un breve tiempo de búfer mínimo, se debe considerar este aspecto. Este problema se analiza a continuación.

[0116] Profundidad del búfer de cambio de tiempo: Este parámetro no es relevante para la radiodifusión, ya que el búfer de desplazamiento de tiempo está determinado por la información de recepción. Sin embargo, se puede considerar que existen algunas instrucciones del remitente sobre cuánto tiempo se deben almacenar los datos en el dispositivo cliente (es decir, el dispositivo receptor).

[0117] Empalmar y restablecer información utilizando períodos: Para esto, la señalización puede ser necesaria. Un período cambia el segmento de inicialización y restablece la cronología. Es necesario que esta información se transmita y que la reproducción del inicio del período en relación con los datos anteriores se programe correctamente. A continuación se proporcionan detalles sobre cómo se puede lograr esto.

[0118] Conjunto de adaptación y metadatos de representación para selección/rechazo: Para la distribución basada en ROUTE, la selección de Conjunto de adaptación y representaciones debe ocurrir en el nivel de sesión de transporte LCT. Una vez que se selecciona una sesión de transporte, estos datos se envían para descodificar y procesar a la capa ISO BMFF. La capa ISO BMFF, basada en la información en el segmento de inicialización, aún puede seleccionar, rechazar o filtrar algunos datos, pero esto no es relevante para el análisis aquí. Para seleccionar la sesión adecuada, se utiliza la información de la LSID. Este tema se analiza con más detalle a continuación.

[0119] Relación de representaciones (conmutable, dependencia, etc.) e información de conmutación perfecta: En la mayoría de las aplicaciones, solo se distribuye una única representación por conjunto de adaptación en la radiodifusión. Sin embargo, si se da el caso de que se radiodifundan varias representaciones, los aspectos individuales de cada representación se pueden proporcionar en la LSID como se describe a continuación. Si se requiere un cambio continuo entre las diferentes representaciones, esto requeriría la MPD, pero solo cuando se realiza el cambio de representación, y la información requerida para el cambio continuo entre las diferentes representaciones también se puede señalar en el segmento de inicialización. En ciertos casos, por ejemplo, cuando se utiliza codificación en capas, esta información también se proporciona en el segmento de inicialización. La dependencia de las sesiones LSID se indica en la LSID, y a continuación se describen detalles adicionales.

[0120] Desviación temporal de presentación: La desviación temporal de presentación señala el primer tiempo de presentación de la representación en el período, es decir, proporciona un anclaje. Esta información debe replicarse utilizando información en las cabeceras de ROUTE. Este tema se analiza con más detalle a continuación.

[0121] Ubicación y URL de los segmentos de medios y de inicialización: La MPD describe la ubicación y el enlace de objetos a las estructuras MPD abordando los siguientes problemas. Dice qué objeto es (segmento de inicialización, segmento de medios, otros tipos) y lo pone en contexto, por ejemplo, se describe la secuencia de segmentos de medios y proporciona al cliente DASH la información sobre cómo utilizar los objetos de datos. La MPD apunta a las URL y, mediante esto, se establece un vínculo entre la información de la MPD y los archivos, la transmisión de medios. Además, el (E)FDT proporciona la asignación de TOI a URL. Esta relación e indicación de

tipo debe ser replicada por la entrega de ROUTE y para ello se utiliza la cabecera de ROUTE y es necesario un uso estricto de las cabeceras TSI, TOI y ROUTE para cumplir con este propósito. También se pueden tener en cuenta ciertos pedidos de envío. Otro aspecto que se puede considerar es cómo un IS se relaciona con los segmentos de medios. Las restricciones y los detalles se analizan a continuación.

5

[0122] Duración de los segmentos de medios: En general, la duración de los segmentos de medios se usa para calcular los tiempos de disponibilidad del segmento y con el propósito de buscar. De lo contrario, la duración no es relevante. Entonces, en general, la información no es necesaria para su radiodifusión.

10

[0123] Información detallada sobre la protección de contenido: La protección del contenido se puede señalar en el segmento de inicialización. Si se van a señalar problemas complejos, la MPD puede considerarse necesaria. Pero, en general, la información del segmento de inicialización se considera suficiente.

15

[0124] Señalización de flujo de eventos: Las representaciones pueden llevar un flujo de eventos que puede tener que ser analizado por el cliente DASH. Si los flujos de eventos se consideran relevantes, pueden indicarse en el descriptor de contenido LSID. De forma alternativa, la MPD se puede utilizar para comunicar flujos de eventos, pero dichos eventos no son relevantes para el inicio. Los detalles adicionales sobre la señalización de flujos de eventos en banda en la LSID se analizan a continuación.

20

[0125] Información del programa de datos auxiliares, métricas: La MPD puede transportar información que sea relevante para ciertas operaciones, como información del programa, inicio de recopilación de métricas u otros medios. Esta información no es crítica en tiempo real y se puede proporcionar, si es necesario, en una MPD que se entrega con una frecuencia más lenta o que se proporciona a través de unidifusión.

25

[0126] Se consideran los siguientes aspectos para establecer un modelo de temporización y búfer:

- Cada paquete tiene un tiempo de transmisión objetivo (TTT) similar al definido en RFC2250 para la entrega basada en RTP de MPEG-2 TS. El tiempo de transmisión objetivo puede:

30

- ser solo información de sugerencia del remitente, o

- señalizarse en la cabecera de control de congestión del paquete LCT como marca de tiempo NTP (los 32 bits del medio) o como un reloj de 90 kHz similar a MPEG-2 TS o como cualquier otro reloj con la escala de tiempo definida en la LSID, por ejemplo, el reloj de los medios contenidos.

35

- Si está presente y es conocido en el receptor, el TTT proporciona información sobre cuándo exactamente enviar los datos a la siguiente capa, es decir, a la capa ISO BMFF.

40

- Se supone que el manejador ISO BMFF inicia la descodificación de inmediato y procesa los datos instantáneamente una vez que se liberan del nivel de ROUTE.

- Se consideran diferentes formas sobre cómo señalar la hora al receptor cuando puede entregar los datos:

45

- Una cabecera de extensión en el primer paquete del grupo RAP (el IS típicamente) y el primer paquete del segmento de medios indica cuánto tiempo deben retenerse los datos en el receptor ROUTE hasta que se liberen al siguiente nivel. Este tiempo se llama tiempo de liberación (RT). RT está en la misma escala de tiempo que TTT y puede compararse con las marcas de tiempo TTT (no con la hora real). Al aumentar TTT, el RT no debe disminuir. Si el tiempo RT en la cabecera de extensión excede el tiempo TTT más grande del objeto/segmento actual, entonces los datos contenidos en este paquete deben retenerse hasta que se alcance el tiempo TTT en el siguiente segmento o en cualquier segmento futuro. Específicamente, el receptor puede actuar de la siguiente manera:

50

- Si se recibe un paquete con un tiempo RT, entonces estos datos se pueden retener hasta que se reciba un paquete con TTT que exceda el RT.

55

- Si se recibe un paquete sin un tiempo de RT, entonces los datos del objeto contenidos pueden liberarse junto con los datos anteriores, donde "precedente" es el orden de los objetos en el aumento de start_offset y el aumento de los números TOI.

60

Si el TTT está en uso y el receptor observa una fluctuación significativa entre el TTT y el tiempo de recepción real, esta fluctuación debe compensarse para evitar subflujos de la memoria intermedia.

- Por ejemplo, un dispositivo puede agregar un retardo adicional o la infraestructura de transmisión agrega un retardo de inicio adicional.

65

- Opción 2: En este caso, se utiliza el tiempo absoluto de reloj de pared para programar la descodificación.

Al objeto se le asigna una hora de reloj de pared en la que el objeto se va a mover al motor ISO BMFF. Este caso es más relevante cuando es importante la reproducción sincronizada entre diferentes usuarios. Con el aumento del tiempo de reloj de pared, el RT debe aumentar. Específicamente, se espera que el receptor actúe de la siguiente manera:

- Si se recibe un paquete con una hora RT, estos datos se pueden retener hasta que se alcance la hora del reloj de pared documentada en RT.
- Si se recibe un paquete sin un tiempo de RT, entonces los datos del objeto contenidos pueden liberarse junto con los datos anteriores, donde precedente es el orden de los objetos que está aumentando start_offset y aumentando los números TOI.
 - Opción 3: Un bit de cabecera en la cabecera ROUTE se establece en 10 para indicar que los datos contenidos aún no se pueden liberar al siguiente nivel. Solo si el bit está desactivado, activado (es decir, igual a 1), los datos se pueden entregar y enviar. Si el último paquete del objeto (indicado por el indicador B que se establece en la cabecera LCT) todavía tiene el bit establecido en 10, entonces este se puede liberar solo cuando el objeto está completo. Sin embargo, tenga en cuenta que siempre se puede entregar un objeto completo. De forma alternativa, este bit de cabecera tiene la obligación de ser igual a 1 para el último paquete de cada objeto.
- Si el TTT está en uso y el receptor observa una fluctuación significativa entre el TTT y el tiempo de recepción real, esta fluctuación debe compensarse para evitar subflujos de la memoria intermedia.
 - Por ejemplo, un dispositivo puede agregar un retardo adicional o la infraestructura de transmisión agrega un retardo de inicio adicional.

[0127] Se consideran tres tipos diferentes de recepción:

1. Recepción basada en MDE: En un primer caso, la reproducción en el receptor se produce lo más rápido posible basándose en los prefijos de los archivos ISO BMFF.
2. En un segundo caso, solo los segmentos completos deben pasar al siguiente nivel. Los segmentos completos son solicitados por el cliente DASH o se publican de acuerdo con RT y TTT o la comparación del reloj de pared. Típicamente, el receptor ROUTE hace que el segmento esté disponible basándose en la información de RT. El cliente DASH envía peticiones basadas en la información de MPD. El contenido se debe crear y entregar de manera que la hora de inicio de disponibilidad del segmento no sea anterior a la hora de publicación.
3. Anclaje de reloj de pared: En un tercer caso, la presentación de los datos se retiene durante algún tiempo para permitir la sincronización con otros medios, por ejemplo, entregados fuera del sistema DASH. El RAP debe publicarse de acuerdo con estos datos.

Sin embargo, el principio en los tres casos sigue siendo el mismo; puede ser simplemente que las unidades de datos que se publican sean MDE o segmentos completos.

[0128] La FIG. 10 es un diagrama conceptual que ilustra varias opciones para señalar cuando un prefijo de un objeto se puede liberar a la siguiente capa para descodificación. Las siguientes opciones de ejemplo, correspondientes a algunas de las diversas opciones analizadas con mayor detalle a continuación, se explican brevemente a continuación y se ilustran en la FIG. 10:

1. Primera opción de ejemplo 360: Para este ejemplo, la FIG. 10 ilustra una serie de paquetes que incluyen un paquete RAP 366 y los paquetes 368A, 368B y 374. Cada uno de estos paquetes contiene un tiempo de transmisión objetivo (TTT) respectivo que básicamente coincide con el tiempo de descodificación a nivel del sistema. En este ejemplo, cada paquete RAP 366 y los paquetes 368A, 368B contienen TTT1 362, mientras que el paquete 374 incluye TTT2 370. El programador utiliza la información TTT para programar la entrega. Además, el paquete RAP 366 incluye un tiempo de liberación (RT) 364 en el mismo dominio que la información TTT, y RT 364 determina cuándo el primer paquete del grupo (el segmento de inicialización (IS), típicamente), o cualquier otro paquete, se puede liberar a la siguiente capa para su procesamiento inmediato. El RT puede agregarse una vez para un objeto o puede agregarse varias veces para un objeto. El RT puede aumentar junto con el aumento del tiempo de envío dentro de un objeto para liberar objetos de una manera más similar a la transmisión. En este ejemplo, el paquete RAP 366, el paquete 368A y el paquete 368B se liberan cuando $TTT1 < RT\ 364 < TTT2$.
2. Segunda opción de ejemplo 380: Para este ejemplo, la FIG. 10 ilustra una serie de paquetes que incluyen un paquete RAP 384 y paquetes 386, 388, 392. El paquete RAP 384 está asociado con RT 382. Se supone que el paquete RAP 384 se recibe en el momento NTP1, el paquete 386 se recibe en NTP2, el paquete 388 se

recibe en NTP3 y el paquete 392 se recibe en NTPX. En este caso, solo se agrega un tiempo de liberación en NTP para el objeto. Esto es similar a un tiempo de disponibilidad de segmento en la MPD, es decir, se permite reenviar el objeto después de este tiempo. Por tanto, el paquete RAP 384 puede liberarse en el tiempo 390 cuando $NTP > RT$ 382. El aspecto beneficioso es que esto no requiere señalización de temporización adicional, pero el remitente debe tener en cuenta el retardo de entrega y la fluctuación al establecer este tiempo de liberación en NTP, o existe el problema de que cualquier retardo inesperado en la entrega ocasione problemas ya que el objeto no se puede presentar a tiempo (lo cual da lugar a retardos en el inicio) o el objeto se recibe con retardo de modo que se pierde la cronología. La opción es beneficiosa para el propósito cuando se usa una sincronización de reloj de pared.

3. Tercer ejemplo, opción 400: Para este ejemplo, la FIG. 10 ilustra una serie de paquetes que incluyen el paquete RAP 404 y los paquetes 406A, 406B, 412. En este caso, el paquete RAP 404 contiene un indicador de liberación (RF) 402 (o "indicador de mantenimiento") que incluye información que indica después de qué tiempo los datos se pueden liberar a la siguiente capa. En este ejemplo, $RF = 0$ para el paquete RAP 404 y los paquetes 406A, 406B. Por tanto, después del tiempo NTP 410, el paquete RAP 404 y los paquetes 406A, 406B pueden liberarse, porque el tiempo en el tiempo NTP 410 es $RF = 1$ 408 (asociado con el paquete 412). Este enfoque es bastante simple, pero puede encontrar un problema, ya que podría no permitir la señalización a través de los límites de los segmentos. No obstante, la opción 400 debe tenerse en cuenta, ya que simplifica el funcionamiento del receptor ROUTE.

[0129] La FIG. 11 es un diagrama conceptual que ilustra dos modelos de ejemplo para recibir y almacenar en memoria intermedia los datos recibidos. La unidad de recuperación 52 del dispositivo cliente 40 puede incluir uno o más componentes configurados de acuerdo con cualquiera de estos ejemplos. En el primer modelo de ejemplo 432, el receptor de ROUTE y la memoria intermedia de salida 426 reciben paquetes, tales como el paquete RAP 420 y los paquetes 422, 424. El receptor ROUTE y el búfer de salida inician y programan la descodificación y presentación con el descodificador ISO BMFF 430. El receptor ROUTE y el búfer de salida pasan los datos recibidos en forma de un flujo ISO BMFF al búfer ISO BMFF 428. Basándose en el horario establecido por el receptor ROUTE y la memoria intermedia de salida 426, el descodificador ISO BMFF 430 obtiene datos de la memoria intermedia ISO BMFF 428, a continuación descodifica y presenta los datos obtenidos.

[0130] En el segundo modelo de ejemplo 452, el receptor de ROUTE 446 recibe paquetes, tales como el paquete RAP 440 y los paquetes 442, 444. El receptor ROUTE almacena en búfer los paquetes recibidos en la salida ROUTE y el búfer ISO BMFF 448, e inicia y programa la descodificación y presentación con el descodificador ISO BMFF 450. El descodificador ISO BMFF 450 obtiene datos de la salida ROUTE y del búfer ISO BMFF 448, y descodifica y presenta los datos obtenidos.

[0131] Basándose en la descripción anterior, los siguientes aspectos pueden ser relevantes:

- El tiempo de disponibilidad se proporciona, por ejemplo, de acuerdo con una de las tres opciones de la FIG. 10 anterior, por ejemplo, una de las opciones 360, 380 o 400. Este es el momento en que se lanza la primera parte del segmento al cliente ISO BMFF.
- El tiempo de presentación es tal que la descodificación y la reproducción se inician instantáneamente en el receptor basándose en la información proporcionada. Tenga en cuenta que este es solo un modelo de datos y que el búfer entre el receptor ROUTE y el receptor ISO BMFF puede ser compartido. Dos modelos de ejemplo (432, 452) se muestran en la FIG. 11. El paquete RAP (por ejemplo, el paquete RAP 420 o el paquete RAP 440) puede contener información en el MDE que compensa el inicio del período y la desviación temporal de presentación, es decir, el inicio de la presentación más temprana en el segmento de medios puede ser instantáneo.
- La verificación del búfer y los retardos de reproducción inicial pueden indicarse mediante una de las tres opciones de ejemplo (360, 380, 400) anteriores.
- Hay diferentes casos que se pueden considerar en el siguiente modo operativo (a continuación se proporciona un ejemplo de señalización de los diferentes tipos). Los diferentes casos a considerar incluyen:
 - Los paquetes de segmentos de medios regulares simplemente se pasan al búfer ISO BMFF sin ninguna programación en el receptor ROUTE. El remitente garantiza que el retardo de reproducción inicial proporciona una reproducción perfecta sin falta de datos del búfer. Estos paquetes pueden contener RT también para programar la liberación, pero típicamente esto no es necesario, ya que pueden liberarse de inmediato, ya que la reproducción es programada por la capa ISO BMFF.
 - Los segmentos de inicialización redundantes son ignorados por el receptor y se descartan y no se pasan al descodificador ISO BMFF.
 - Los segmentos de inicialización aumentados se proporcionan al descodificador ISO BMFF. Sin

embargo, se debe informar al decodificador que los medios son continuos en el tiempo y que solo ha cambiado parte de la información no esencial. Sin embargo, si no existe dicha API, se puede realizar un reinicio y se puede reprogramar la reproducción.

- 5 ○ Límite del período de contenido: En este caso, se debe reenviar el IS y se aplica una nueva programación completa.

10 Tenga en cuenta que esto puede dar lugar a casos en los que el búfer ISO BMFF con algunas muestras se vacíe o no haya datos disponibles durante un breve período de tiempo que deba ser manejado por el decodificador. Esto es idéntico a las operaciones DASH normales.

- Tenga en cuenta que la operación anterior puede ser posible porque el remitente debe cumplir con los requisitos de envío que se describen a continuación con mayor detalle.

15 **[0132]** Si el remitente y el receptor están configurados para usar la entrega/recepción basada en segmentos, entonces se puede agregar información a la señalización de temporización, como se describe a continuación. Sin embargo, en el caso de que la recepción basada en segmentos se realice basándose en una entrega basada en MDE, entonces el problema es más complicado, porque el receptor necesita entender cómo hacer uso de la información de tiempo. Deben considerarse las siguientes opciones:

- 20 • La MPD está disponible para programar y señalar los tiempos de disponibilidad del segmento.
- 25 • Se agrega un atributo a la LSID para cada componente que indica el retardo adicional en los tiempos del tiempo de los medios (y TTT) para soportar la recepción basada en segmentos. Típicamente, ese tiempo es la duración máxima del segmento. Tenga en cuenta que el tiempo no está en tiempo de reloj de pared, sino en tiempo de transmisión, por lo que si se realiza una entrega basada en ráfagas, entonces el retardo adicional de entrega de segmento puede ser marginalmente mayor que el de la recepción basada en MDE.

30 **[0133]** Si el remitente y el receptor están configurados para utilizar anclaje de reloj de pared, es decir, descodificación basada en reloj de pared, entonces se puede agregar información a la información de señalización de temporización que se describe a continuación. Sin embargo, puede haber casos en los que se señale la entrega basada en MDE, pero el receptor necesita sincronizar la descodificación y la reproducción con la hora del reloj de pared. También se pueden considerar las siguientes opciones:

- 35 • La MPD está disponible para programar y señalar los tiempos de disponibilidad del segmento, así como el retardo de presentación sugerido.
- 40 • Se agrega un atributo a la LSID para cada componente que indica la asignación del TTT o el tiempo de descodificación al tiempo del reloj de pared.
- 45 • Se crea una cabecera de extensión para LCT que proporciona esta información, es decir, la asignación a la programación de reproducción del reloj de pared.

45 **[0134]** Para permitir la selección de los medios entregados en una sesión de transporte LCT basada en la información de la LSID, se pueden agregar a la LSID muchos elementos y atributos que se asignan a un conjunto de adaptación, excepto las representaciones. Específicamente, esto puede incluir parte o toda la siguiente información (para obtener más detalles, consulte ISO/IEC23009-1, cláusula 5.3.4 (Conjunto de adaptación) y 5.3.7 (Atributos comunes)):

- 50 • Un identificador que usa el elemento @id.
- 55 • Una asociación de grupo que utiliza el atributo @group.
- 60 • el lenguaje descrito por el atributo @lang.
- 65 • el tipo de componente de medios descrito por el atributo @contentType.
- la relación de aspecto de la imagen descrita por el atributo @par.
- la propiedad de rol como se describe en los elementos de rol.
- la propiedad de accesibilidad descrita por los elementos de accesibilidad.
- la propiedad de punto de visualización como se describe en los elementos de punto de visualización.
- la propiedad de calificación como se describe en los elementos de calificación.

- Información sobre las propiedades del segmento (acceso aleatorio, etc.): consulte los experimentos básicos de SISSI.
- 5 • Atributo @profiles para el conjunto de adaptación.
- El suministro de @width y @height especifica el tamaño de presentación visual horizontal y vertical del tipo de medio de vídeo en una cuadrícula determinada por el atributo @sar.
- 10 • @sar especifica la relación de aspecto de muestra del tipo de componente de medio de vídeo.
- @framerate: especifica la frecuencia de tramas de salida (o en el caso de entrelazado, la mitad de la velocidad del campo de salida) del tipo de medio de vídeo.
- 15 • @audiosamplingRate: frecuencia de muestreo máxima del tipo de componente de medios de audio.
- @mimeType: especifica el tipo MIME de la concatenación del segmento de inicialización, si está presente, y todos los segmentos de medios consecutivos en la representación.
- 20 • @codecs: especifica los códecs presentes dentro de la representación. Los parámetros del códec también pueden incluir la información de perfil y nivel cuando corresponda.
- @scanType: especifica el tipo de escaneo del material de origen del tipo de componente de medios de vídeo.
- 25 • FramePacking: especifica la información de la disposición de empaquetado de tramas del tipo de componente de medios de vídeo.
- AudioChannelConfiguration: especifica la configuración del canal de audio del tipo de componente de medios de audio.
- 30 • ContentProtection: especifica información sobre los esquemas de protección de contenido utilizados para las representaciones asociadas.
- EssentialProperty: especifica información sobre el elemento contenedor que el autor de la presentación multimedia considera esencial al seleccionar este componente.
- 35 • SupplementalProperty: especifica información complementaria sobre el elemento contenedor que se puede utilizar para procesar el componente.
- 40 • InbandEventStream: especifica la presencia de un flujo de eventos en banda en las representaciones asociadas.

[0135] Toda esta información se puede utilizar para la selección en el nivel de sesión de transporte LCT. Si no se proporciona ninguna selección, todos los flujos se pueden enviar al manejador ISO BMFF para su procesamiento. Esta información puede proporcionarse para el filtrado y la selección tempranos a nivel de sesión de transporte.

[0136] La FIG. 12 es un diagrama conceptual que ilustra un conjunto de ejemplo de componentes de un dispositivo receptor 460, que incluye una unidad de selección y filtrado 462, un receptor de ROUTE 464, otros procesadores de objetos 466, procesador de medios ISO BMFF 468 y preprocesador ISO BMFF 470. El dispositivo receptor 460 puede corresponder al dispositivo cliente 40, donde la unidad de recuperación 52 puede incluir la unidad de selección y filtrado 462 y el receptor de ROUTE 464, mientras que la unidad de desencapsulación 54 puede corresponder al preprocesador ISO BMFF 470, el procesador de medios ISO BMFF 468 y otros procesadores de objetos 466.

[0137] En este ejemplo, el receptor de ROUTE 464 proporciona a la LSID descriptores de contenido a la unidad de selección y filtrado 462. Basándose en la información, la unidad de selección y filtrado 462 puede seleccionar los componentes apropiados que deberían enviarse a los procesadores (por ejemplo, preprocesador ISO BMFF 470, procesador de medios ISO BMFF 468 y otro procesador de objetos 466).

[0138] Al aplicar el principio anterior, se puede garantizar la extensibilidad con nuevos esquemas y extensiones. El sistema se mantiene compatible con las implementaciones basadas en DASH. Además, el procesador de medios ISO BMFF 468 también puede seleccionar o rechazar ciertos flujos de medios en su nivel.

[0139] Las cabeceras LCT se pueden utilizar de forma flexible para señalar determinadas propiedades. Sin embargo, el uso de las cabeceras debe ser obligatorio para ROUTE. Esto se analiza a continuación.

[0140] En una LSID que transporta flujos de medios (indicado por el tipo de contenido xxx/mp4), la Tabla 1 proporciona una configuración del campo de punto de código. Esto permite que el emisor y el receptor funcionen sin el FDT y la MPD.

5

TABLA1: Asignación de puntos de código para flujos de medios con tipo de contenido xxx/mp4

Punto de código	Semántica
0	Tipo de objeto descrito en FDT, sin fragmentar (el paquete completo contiene IS, sin cabecera ROUTE)
1	Tipo de objeto descrito en FDT, cabecera ROUTE de 32 bits
2	Nuevo IS, discontinuidad de la cronología, sin fragmentar (es decir, el paquete completo contiene IS, sin cabecera ROUTE)
3	Nuevo IS, cronología modificada, fragmentada, cabecera ROUTE de 32 bits
4	Nuevo IS, discontinuidad de la cronología, sin fragmentar (es decir, el paquete completo contiene IS, sin cabecera ROUTE)
5	Nuevo IS, cronología continua, fragmentada, cabecera ROUTE de 32 bits
6	IS redundante, no fragmentado (es decir, el paquete completo contiene IS, sin cabecera ROUTE)
7	IS redundante, fragmentado, cabecera ROUTE de 32 bits
8	Segmento de medios sin fragmentar
9	Segmento de medios fragmentado
10	Objeto en modo de entidad sin fragmentar
11	Objeto en modo de entidad fragmentado
12 – 127	Reservado para la asociación de objetos estáticos
128 – 255	Asignación dinámica de puntos de código a través de LSID

[0141] Las siguientes cabeceras LCT se pueden utilizar en ROUTE sin MPD:

10

- El TSI señala una única representación basada en ISO BMFF
- El TOI se utiliza para señalar los objetos. Los segmentos de medios dentro de un período de contenido deben enviarse con un número TOI creciente, es decir, los objetos se incrementan en 1. Solo en períodos de contenido esto puede cambiar. Los TOI de segmento de medios solo pueden usar el espacio para el que el primer bit del TOI está establecido en 1. Los segmentos no multimedia pueden usar el espacio TOI para el que el primer bit se establece en 0.
- El punto de código se utiliza como se documenta en la sección 3.6.2.
- Los bits de PSI se utilizan de la siguiente manera:
 - Si el primer bit se establece en 0, se aplica la señalización de liberación basada en el tiempo:
 - Si el segundo bit se establece en 0, el campo de control de congestión no se utiliza.
 - Si el segundo bit se establece en 1, entonces el campo de control de congestión contiene una marca de tiempo en una frecuencia de reloj de 90 kHz.
 - Si el primer bit se establece en 1, entonces
 - El segundo bit señala el indicador de liberación del objeto.
- El campo de control de congestión puede utilizarse para señalar la marca de tiempo en una frecuencia de reloj de 90 kHz.

15

20

25

30

35

[0142] Se definen las siguientes cabeceras de extensión LCT de ejemplo:

- Tiempo de liberación inicial en TTT: Especifica el tiempo de publicación más temprano de los datos contenidos en este paquete en tiempo TTT. Sería adecuada una cabecera de extensión de 32 bits de tamaño.
- Tiempo de liberación inicial en NTP: Especifica el tiempo de liberación exacto de los bytes iniciales actuales del objeto en tiempo NTP. Nuevamente, 32 bits pueden ser adecuados, proporcionando los 32 bits intermedios de una marca de tiempo NTP.

40

45

[0143] Se supone que los procedimientos de envío de la opción 1 se eligen a continuación. Las variantes de las

opciones 2 y 3 quedan en estudio y pueden utilizarse junto con las técnicas que se describen a continuación u otras técnicas. El siguiente comportamiento de envío se considera sin utilizar la MPD para una determinada presentación multimedia:

- 5 • Los fragmentos de LSID se pueden proporcionar mediante arranque.
 - Los fragmentos de LSID pueden actualizarse para señalar cualquier cambio en la oferta del servicio.
 - Los fragmentos de LSID describen todos los componentes de radiodifusión del servicio.
- 10 ○ Se puede agregar un descriptor de contenido para describir las propiedades de cada componente de los medios. El descriptor de contenido debe proporcionar información suficiente para que el receptor de ROUTE pueda seleccionar o rechazar ciertos componentes.
- 15 • En general, solo se distribuye una única representación para cada componente (conjunto de adaptación). Sin embargo, si se distribuyen diferentes representaciones para el mismo conjunto de adaptación, entonces la LSID puede contener información suficiente para diferenciar los componentes, por ejemplo, por la clasificación de calidad, la resolución espacial, el parámetro de códecs necesario, etc.
- 20 • Se puede utilizar una sesión LCT (identificada por un TSI) para cada representación. Las sesiones LCT se multiplexan a nivel de paquete. El TTT debe utilizarse de forma coherente en las sesiones de LCT multiplexadas.
- 25 • Cada paquete puede contener un tiempo de transmisión objetivo sobre una base de 90 kHz. Este tiempo expresa el tiempo de entrega objetivo. Preferentemente, se usa el tiempo de decodificación de los medios contenidos en ISO BMFF.
- 30 • El siguiente procedimiento de envío se puede aplicar para representaciones dentro de una sola sesión LCT:
 - Cualquier objeto puede enviarse en la sesión LCT. Si el objeto no es un segmento de medios, entonces el bit más significativo en el TOI de puede ser 0. Si el objeto es un segmento de medios, entonces el bit más significativo en el TOI puede ser 1. Un objeto puede identificarse mediante la asignación estática en la Tabla 1, o puede definirse un punto de código en la LSID.
 - 35 ○ Suponga que un segmento de medios tiene un tiempo de publicación específico. El IS correspondiente al segmento de medios puede indicar el mismo tiempo de liberación. Tenga en cuenta que el tiempo de liberación solo se puede enviar en el IS si el segmento de medios sigue inmediatamente.
 - 40 ○ Los segmentos de inicialización se envían como objetos mediante el protocolo ROUTE. Para IS, solo se puede usar el número TOI que comienza con un 0 en el bit más significativo.
 - El tipo de segmento de inicialización se anuncia en el punto de código de acuerdo con la Tabla 1.
 - Si el IS encaja en un solo paquete de ROUTE, entonces se puede usar el punto de código 2, 4 o 6. Si el IS no cabe en un solo paquete de ROUTE, entonces se pueden usar los puntos de código 3, 5 o 7.
 - Si el IS se envía primero o la cronología no es continua, entonces se puede señalar el punto de código 2 o 3. El IS debe utilizar un nuevo TOI. El primer paquete del IS puede proporcionar un número aleatorio para TTT. Sin embargo, en el último caso y en una transmisión continua, el TTT debería continuar expresando el horario de reproducción.
 - 50 ▪ Si el IS se repite para soportar el acceso aleatorio, entonces se puede señalar el punto de código 6 o 7. El TOI puede permanecer igual. El TTT puede ser continuo y no decreciente.
 - 55 ▪ Si el IS está actualizado, pero la cronología es continua, entonces se usa el punto de código 4 o 5. El IS debe utilizar un nuevo TOI. El TTT puede ser continuo y no decreciente.
 - 60 ▪ Junto con un IS, se puede enviar una cabecera de extensión que indica el tiempo de liberación más temprano del IS a la siguiente capa en la misma escala que el TTT. Si no está presente, el tiempo de liberación es inmediato.
 - 65 ○ Los segmentos de medios (MS) se envían como objetos mediante el protocolo ROUTE. Para MS solo se puede usar el número TOI que comienza con un 1 en el bit más significativo.
 - El tipo de MS se anuncia en punto de código de acuerdo con la Tabla 1.
 - Si el MS encaja en un solo paquete de ROUTE, se puede utilizar el punto de código 8. Si el MS no

encaja en un solo paquete de ROUTE, entonces puede usarse el punto de código 9.

- Pueden ser necesarias algunas reglas adicionales.

[0144] La FIG. 13 es un diagrama conceptual que ilustra un procedimiento de envío de ejemplo para una serie de paquetes de ejemplo. En este ejemplo, una primera representación incluye el segmento de inicialización 480 y dos segmentos de medios 482, 484, que se enviarán seguidos de una segunda representación con un nuevo segmento de inicialización 486 y un segmento de medios posterior 488. Cada uno de los segmentos usa TSI = 1. Es decir, se utiliza esta sesión LCT. El envío de los paquetes en orden se muestra en la FIG. 13:

- El primer paquete que se enviará (paquete de segmento de inicialización 490A) corresponde al segmento de inicialización 480, siendo el TTT el tiempo de descodificación de la primera muestra en el MS A. Se selecciona un TOI para el paquete de segmento de inicialización 490A, y el CP se establece en 2 para indicar una nueva cronología. El tiempo de liberación del segmento de inicialización 480 se establece en un valor que es ligeramente mayor que A para indicar cuándo comenzará la presentación.
- El segundo paquete (paquete de segmento de medios 492A) es el primer paquete del segmento de medios 482, y el segmento de medios 482 está fragmentado. Por tanto, CP = 9. Se utiliza un nuevo TOI para indicar un nuevo objeto. Se utiliza el mismo TTT que el utilizado para el segmento de inicialización 480. Como los datos se van a liberar junto con el segmento de inicialización 480, no se envía ningún nuevo RT.
- Para el segundo paquete (paquete de segmento de medios 492B) del segmento de medios 482, se establece un nuevo tiempo de transmisión (asumiendo que contiene un tiempo de descodificación posterior) con $C > A$. Si $B < C$, esto daría como resultado la liberación del segmento de inicialización 480 y el paquete de segmento de medios 492A a la siguiente capa.
- Para permitir el acceso aleatorio, el segmento de inicialización 480 se enviará de forma redundante (CP = 4) con un TTT que coincida con el del paquete de segmento de medios 494A del segmento de medios 484. Se establece un tiempo de liberación en la escala del TTT.
- Entonces, el segmento de medios 484 se envía en dos paquetes (paquetes de segmentos de medios 494A, 494B) de la misma manera que el segmento de medios 482. Los tiempos se utilizan para liberar los paquetes. Se utiliza un nuevo TOI, que es uno más que el TOI del segmento de medios 482.
- Con el inicio de un nuevo período, se enviará un nuevo segmento de inicialización 486 (en forma de paquete de segmento de inicialización 496), que por lo tanto está marcado con CP = 2. Se utiliza un nuevo TOI para el segmento de inicialización 486. El TTT y el RT de la nueva representación deben ser una continuación del TTT anterior para instruir la secuencia y el tiempo de reproducción.
- Posteriormente, el segmento de medios 488 puede enviarse como uno o más paquetes, incluido el paquete de segmento de medios 498

[0145] En un ejemplo alternativo, se puede utilizar una MPD ligera. En este ejemplo, la señalización y las operaciones de temporización y almacenamiento en memoria intermedia son las mismas que las anteriores, mientras que no se toman otros cambios anteriores para la señalización de información MPD sin temporización a través de otros medios. En cambio, la MPD se cambia para permitir que contenga solo la información absolutamente necesaria para los diferentes escenarios (incluso para la sintonización y el cambio de canal en la radiodifusión) mientras que no hay otra información. En el extremo, cuando no se necesita ninguna información de MPD, puede estar vacía. Cuando solo un subconjunto de la información es necesario y está presente, la MPD es ligera y, por lo tanto, se evita una sobrecarga innecesaria. El remitente puede optar por colocar copias MPD regulares de forma escasa en algunos RAP, y entre dos copias MPD regulares, colocar una MPD ligera en cada RAP

[0146] Esto puede dar como resultado cualquiera o todos los siguientes:

- Inicio de MPD que son simples, pero redundantes una vez que se recibe la MPD completa
- Los MPD solo contendrían la información del componente
- El tiempo se ignoraría ya que es impulsado por la capa inferior.
- No se especifican URL para segmentos, ya que se entregan a través del protocolo de objeto.
- Básicamente, el identificador de contenido mencionado anteriormente se asignaría a un objeto separado.

[0147] En este momento, se pueden utilizar segmentos que se ajusten al perfil en vivo ISO BMFF. Las

optimizaciones para un formato de medios mejorado para tales aplicaciones se analizan en Stockhammer *et al.*, "LOW LATENCY VIDEO STREAMING [TRANSMISIÓN DE VÍDEO DE BAJA LATENCIA]", Solicitud Provisional de EE. UU. 62/114 423, presentada el 10 de febrero de 2015.

5 **[0148]** Para una oferta de servicios híbridos, la MPD puede ser importante. Se puede utilizar una MPD unificada que describa la radiodifusión y el contenido distribuido de banda ancha, mediante el cual:

- La MPD se entrega como un objeto en la sesión ROUTE o la MPD se proporciona a través de banda ancha solo a través de un enlace.
- 10 • Un EFDT se entrega como un objeto en la sesión ROUTE o el EFDT se proporciona a través de banda ancha solo a través de un enlace.
- El EFDT documenta la asignación entre el TOI en la entrega del objeto y la URL que se puede ver en la MPD.
- 15 • A continuación, la MPD controla la sincronización, es decir, los segmentos de radiodifusión ya no se insertan en el cliente ISO BMFF, sino que el cliente DASH controla la sincronización. Sin embargo, los detalles son específicos de la implementación.

20 **[0149]** La FIG. 14 es un diagrama conceptual que ilustra un modelo de cliente DASH híbrido de ejemplo. La unidad de recuperación 52 del dispositivo cliente 40 de la FIG. 1 puede configurarse de acuerdo con el ejemplo de la FIG. 14. En este ejemplo, el modelo de cliente DASH híbrido incluye el receptor ROUTE y el búfer de salida 506 y el cliente DASH 512. De acuerdo con el modelo de cliente DASH híbrido, un dispositivo cliente puede recibir
25 segmentos (por ejemplo, paquete RAP500 y paquetes 502, 504) mediante transmisión ROUTE usando el receptor ROUTE y el búfer de salida 506 o mediante transmisión de unidifusión desde la red 514 usando el cliente DASH 512. Al recibir segmentos, el receptor ROUTE y el búfer de salida 506 proporcionan los segmentos al búfer ISO BMFF 508 o el cliente DASH 512 proporciona los segmentos al búfer ISO BMFF 508.

30 **[0150]** Hay dos tipos de implementaciones de ejemplo para tipos de recepciones para radiodifusión:

- Recepción sin MPD: El receptor ROUTE y el búfer de salida 506 reenvían directamente los datos al búfer ISO BMFF 508 para descodificarlos mediante el descodificador ISO BMFF 510 y reproducirlos.
- 35 • Recepción basada en MPD: El receptor ROUTE y el búfer de salida 506 genera una MPD que incluye toda la información, de modo que el cliente DASH 512 recupera los datos del receptor ROUTE y el búfer de salida 506 (por ejemplo, actuando como un servidor proxy) y almacena los datos en el búfer ISO BMFF 508 para descodificar acabar.

40 **[0151]** Ambas versiones son posibles y son una opción de implementación. Además, se puede implementar un cliente híbrido que utilice radiodifusión y banda ancha de la manera que se muestra en la FIG. 14. En este caso, el receptor ROUTE y la memoria intermedia de salida 506 proporcionan datos recibidos a la memoria intermedia ISO BMFF 508 después de recibir datos del flujo de radiodifusión. Una vez que se agrega la banda ancha, la MPD crea la relación. El cliente DASH 512 asigna URL a los TOI y, por lo tanto, puede cambiar a banda ancha basándose
45 en la información de la MPD.

[0152] La FIG. 15 es un diagrama de flujo que ilustra un ejemplo de procedimiento para transportar datos de medios o una presentación de medios mediante sesiones LCT de acuerdo con las técnicas de esta divulgación. El ejemplo de la FIG. 15 se explica con respecto al dispositivo servidor 60 y el dispositivo cliente 40 de la FIG. 1. Sin embargo, se debe entender que otros dispositivos se pueden configurar para implementar estas técnicas. Por ejemplo, los dispositivos de las FIGS. 4, 5, 6, 11, 12 o 14 pueden configurarse para realizar estas técnicas.

[0153] Inicialmente, el dispositivo servidor 60 puede obtener datos para una pluralidad de representaciones (550) de una presentación multimedia. Por ejemplo, el dispositivo de servidor 60 puede recibir contenido preparado desde el dispositivo de presentación de contenido 20 (FIG. 1). De forma alternativa, el dispositivo servidor 60 puede incluir codificadores de medios (como el codificador de audio 26 y el codificador de vídeo 28) y/o una unidad de encapsulación (como la unidad de encapsulación 30) para preparar las representaciones para su distribución. La pluralidad de representaciones puede corresponder a representaciones 68 de contenido multimedia 64 de la FIG. 1. Además, el dispositivo servidor 60 puede obtener un archivo de manifiesto, tal como una MPD, para la pluralidad de representaciones (por ejemplo, el archivo de manifiesto 66 de la FIG. 1).

[0154] El dispositivo servidor 60 puede entonces asignar las representaciones a sesiones LCT (552). En general, cada una de las sesiones de LCT puede transportar datos de una de las representaciones, de modo que existe una relación de uno a uno entre las sesiones de LCT y las representaciones. Además, el dispositivo servidor 60
65 puede construir una LSID para las sesiones LCT indicando una correspondencia entre las representaciones y las sesiones LCT (554). La LSID puede señalar, por ejemplo, relaciones entre TSI para las sesiones LCT e

identificadores de representación (por ejemplo, anchos de banda para las representaciones). Como se indicó anteriormente, el dispositivo servidor 60 también puede construir la LSID para describir combinaciones de IP y puerto para las diversas sesiones LCT.

[0155] El dispositivo de servidor 60 puede construir además la LSID para incluir datos que se incluirían convencionalmente en el archivo de manifiesto, como, por ejemplo, atributos que incluyen cualquiera o todos los @id, @group, @lang, @contentType, @par (aspecto de imagen ratio), elementos de rol, elementos de accesibilidad, elementos de punto de visualización, elementos de valoración, propiedades de segmento de los segmentos de las representaciones, @profiles, @width y @height, @sar, @framerate, @audiosamplingRate, @mimeType, @codecs, @scanType, FramePacking, AudioChannelConfiguration, ContentProtection, EssentialProperty, Supplemental-Property y/o InbandEventStream, como se analizó anteriormente.

[0156] Además, el dispositivo servidor 60 puede construir paquetes de las sesiones LCT para incluir una cabecera LCT de acuerdo con, por ejemplo, el ejemplo de la FIG. 9. Los paquetes pueden incluir datos de segmentos de las representaciones. La información de la cabecera puede indicar, por ejemplo, TSI y/o TOI de representaciones y segmentos correspondientes.

[0157] El dispositivo servidor 60 puede enviar entonces la LSID y los datos de las representaciones a través de las correspondientes sesiones LCT (556). El dispositivo cliente 40 puede recibir la LSID para las sesiones LCT (558). Aunque no se muestra en el ejemplo de la FIG. 15, el dispositivo servidor 60 también puede enviar el archivo de manifiesto periódicamente, por ejemplo, con ciertos puntos de acceso aleatorios de las representaciones. En particular, en este ejemplo, se supone que el dispositivo cliente 40 recibe la LSID antes de recibir el archivo de manifiesto (por ejemplo, entre archivos de manifiesto). Sin embargo, de acuerdo con las técnicas de esta divulgación, el dispositivo cliente 40 puede acceder a datos de medios de una o más de las sesiones LCT (y, por lo tanto, las representaciones correspondientes) usando la LSID, sin (o antes) de recibir el archivo de manifiesto.

[0158] Por ejemplo, el dispositivo cliente 40 puede determinar las correspondencias entre las sesiones LCT y las representaciones. El dispositivo cliente 40 también puede determinar características de las representaciones usando datos señalizados en la LSID. Por ejemplo, el dispositivo cliente 40 puede determinar cuál de las representaciones coincide con las características de codificación y reproducción soportadas por elementos del dispositivo cliente 40, como el descodificador de audio 46, el descodificador de vídeo 48, la salida de audio 42 y la salida de vídeo 44. Basándose en las características soportadas y las características de codificación y reproducción de las representaciones, el dispositivo cliente 40 puede seleccionar las representaciones usando la LSID (560). Por ejemplo, el dispositivo cliente 40 puede seleccionar aquellas representaciones que tienen características de codificación y reproducción soportadas por los elementos del dispositivo cliente 40.

[0159] El dispositivo cliente 40 puede además extraer paquetes de las sesiones LCT de las representaciones seleccionadas (562). Cada paquete puede corresponder a un segmento de la representación. Cada segmento de la representación puede transmitirse en forma de uno o más paquetes. Como se analizó anteriormente con respecto a los diversos ejemplos de la FIG. 10, los paquetes pueden incluir información que indique cuándo pueden liberarse los paquetes. Así, después de que se hayan recibido todos los paquetes para un segmento, la unidad de recuperación 52 del dispositivo cliente 40 puede enviar un segmento reconstruido a la unidad de desencapsulación 50, que finalmente puede desencapsular los datos de los medios y enviar los datos de los medios a los descodificadores apropiados, por ejemplo, el descodificador de audio 46 o descodificador de vídeo 48. De esta manera, el dispositivo cliente 40 puede enviar datos de medios desde los paquetes a los descodificadores apropiados (564) para descodificar y, en última instancia, presentarlos.

[0160] De esta manera, el procedimiento de ejemplo de la FIG. 15 representa un ejemplo de un procedimiento que incluye la determinación de una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP (DASH) a partir de una descripción de la instancia de sesión de transporte de codificación en capas (LCT) (LSID), en el que la LSID incluye información representativa de una pluralidad de sesiones LCT, con cada una de las sesiones LCT que incluye datos de una respectiva de las representaciones, e iniciando el consumo de una o más de las representaciones de la presentación de medios DASH usando la LSID y sin usar un archivo de manifiesto para la presentación de medios DASH, en el que iniciar el consumo comprende recibir paquetes de las sesiones LCT que incluyen partes de datos de una o más de las representaciones, y proporcionar datos de los paquetes a un descodificador de medios.

[0161] El procedimiento de ejemplo de la FIG. 15 también representa un ejemplo de un procedimiento que incluye la construcción de una descripción de la instancia de sesión de transporte de codificación en capas (LCT) (LSID) que incluye información representativa de una pluralidad de sesiones LCT, incluyendo cada una de las sesiones LCT datos de una respectiva de una pluralidad de representaciones de una presentación de medios transmisión adaptativa dinámica por HTTP (DASH), en la que la LSID indica correspondencias entre las sesiones LCT y las representaciones, dando salida a la LSID y dando salida a datos de las representaciones en las sesiones LCT correspondientes.

[0162] La FIG. 16 es un diagrama de flujo que ilustra otro ejemplo de procedimiento para transportar datos de

medios de una presentación de medios mediante sesiones LCT de acuerdo con las técnicas de esta divulgación. En este ejemplo, el dispositivo cliente 40 determina inicialmente una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP (DASH) a partir de una descripción de la instancia de sesión de transporte de codificación en capas (LCT) (LSID), en la que la LSID incluye información representativa de una pluralidad de sesiones LCT, con cada una de las sesiones LCT que incluye datos de una respectiva de las representaciones (580). A continuación, el dispositivo cliente 40 inicia el consumo de una o más de las representaciones de la presentación de medios DASH usando la LSID y sin usar un archivo de manifiesto para la presentación de medios DASH (582). En particular, cuando se inicia el consumo, el dispositivo cliente 40 recibe paquetes de las sesiones LCT que incluyen partes de datos de una o más de las representaciones (584), y proporciona datos de los paquetes a un descodificador de medios (586).

[0163] La FIG. 17 es un diagrama de flujo que ilustra otro ejemplo de procedimiento para transportar datos de medios de una presentación de medios mediante sesiones LCT de acuerdo con las técnicas de esta divulgación. En este ejemplo, el dispositivo servidor 60 construye una descripción de la instancia de sesión de transporte de codificación en capas (LCT) (LSID) que incluye información representativa de una pluralidad de sesiones LCT, incluyendo cada una de las sesiones LCT datos de una respectiva de una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP (DASH), en la que la LSID indica correspondencias entre las sesiones LCT y las representaciones (590). A continuación, el dispositivo servidor 60 genera la LSID (592) y genera datos de las representaciones en las correspondientes sesiones LCT (594).

[0164] En uno o más ejemplos, las funciones descritas se pueden implementar en hardware, software, firmware o en cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en, o transmitir por, un medio legible por ordenador como una o más instrucciones o código, y ejecutar mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador que correspondan a un medio tangible, tales como medios de almacenamiento de datos, o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder, en general, a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o a (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser medios disponibles cualesquiera a los que se puede acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0165] A modo de ejemplo, y no de limitación, dichos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento en disco óptico, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que se pueda usar para almacenar el código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador. Además, cualquier conexión recibe apropiadamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otra fuente remota usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas están incluidos en la definición de medio. Sin embargo, se deberá entender que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, están dirigidos a medios de almacenamiento no transitorios tangibles. Los discos, como se usan en el presente documento, incluyen el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos reproducen normalmente datos de forma magnética y otros discos reproducen los datos de forma óptica con láseres. Las combinaciones de lo anterior también se deben incluir dentro del alcance de los medios legibles por ordenador.

[0166] Las instrucciones se pueden ejecutar por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices lógicas programables por campo (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o de programa informático dedicados configurados para la codificación y la descodificación, o incorporarse en un códec combinado. Además, las técnicas se podrían implementar por completo en uno o más circuitos o elementos lógicos.

[0167] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluyendo un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). En esta divulgación se describen diversos componentes, módulos o unidades para destacar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no se requiere necesariamente su realización mediante diferentes unidades de hardware. En su lugar, como se describe

anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec o proporcionar mediante un grupo de unidades de hardware interoperativas, que incluya uno o más procesadores como se describe anteriormente, junto con software y/o firmware adecuados.

- 5 **[0168]** Se han descrito diversos ejemplos. Estos y otros ejemplos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

1. Un procedimiento de recepción de datos de medios, comprendiendo el procedimiento:

5 determinar una pluralidad de representaciones de una presentación de medios de transmisión adaptativa
dinámica por HTTP, DASH, a partir de una descripción de una instancia de sesión de transporte de
codificación en capas, LCT, LSID, recibida en el que la LSID incluye información representativa de una
pluralidad de sesiones LCT, con cada una de las sesiones LCT que incluye datos de una respectiva de
10 las representaciones y en el que la LSID indica correspondencias entre las sesiones LCT y las
representaciones; e

iniciar el consumo (560) de una o más de las representaciones de la presentación de medios DASH
utilizando la LSID, en el que el inicio del consumo comprende:

15 determinar las correspondencias entre las sesiones LCT y las representaciones basadas en la LSID;

recibir (562) un primer conjunto de datos que incluye paquetes de las sesiones LCT que incluyen
partes de datos de una o más de las representaciones hasta un primer tiempo de reproducción,
antes de recibir un archivo de manifiesto;

20 proporcionar datos (564) de los paquetes a un decodificador de medios;

recibir un archivo de manifiesto después de recibir el primer conjunto de datos; y

25 recibir un segundo conjunto de datos, diferente del primer conjunto de datos, de la presentación de medios
DASH utilizando el archivo de manifiesto, teniendo el segundo conjunto de datos tiempos de reproducción
siguientes al primer tiempo de reproducción.

2. El procedimiento según la reivindicación 1, que comprende además:

30 determinar al menos una de las características de codificación o características de reproducción de las
representaciones de la presentación de medios DASH a partir de uno o más descriptores de contenido de
la LSID; y

35 seleccionar una o más de las representaciones basándose en las características de codificación
determinadas o características de reproducción.

3. El procedimiento según la reivindicación 2, en el que una o más características de codificación o
características de reproducción incluyen una o más de códec, información de accesibilidad, calidad,
40 resolución espacial, punto de visualización, clasificación, un atributo de perfil de un conjunto de adaptación,
relación de aspecto de muestra, frecuencia de tramas, frecuencia de muestreo de audio, tipo de mímica, tipo
de escaneo, información de empaquetado de tramas, configuración del canal de audio, preparación de
contenido, propiedad esencial, propiedad complementaria o flujo de eventos en banda.

4. El procedimiento según la reivindicación 1, que comprende además usar el archivo de manifiesto para
combinar la entrega de radiodifusión y unidifusión de datos de la presentación de medios DASH.

5. El procedimiento según la reivindicación 1, en el que el archivo de manifiesto es una descripción de
presentación de medios, MPD.

6. El procedimiento según la reivindicación 1, en el que al menos una de las una o más representaciones incluye
un segmento de inicialización y uno o más segmentos de medios formateados de acuerdo con un formato de
segmento DASH, y en el que los paquetes que comprenden datos para el segmento de inicialización o los
55 segmentos de medios comprenden además cabeceras LCT.

7. El procedimiento según la reivindicación 1, que comprende además usar campos de identificador de sesión
de transporte, TSI, de cabeceras LCT de los paquetes de descripción de sesiones LCT para determinar
correspondencias entre las sesiones LCT y las representaciones.

8. El procedimiento según la reivindicación 1, que comprende además determinar tiempos de liberación para
datos de paquetes de las sesiones LCT a partir de al menos uno de indicación específica de protocolo, PSI,
bits de cabeceras LCT de los paquetes o cabeceras de extensión de las cabeceras LCT de los paquetes.

9. El procedimiento según la reivindicación 1, que comprende además determinar tiempos de transmisión
objetivo para paquetes de las sesiones LCT a partir de información de control de congestión de cabeceras
65 LCT de los paquetes.

10. Un dispositivo (40) para recibir datos de medios de recepción, comprendiendo el dispositivo:

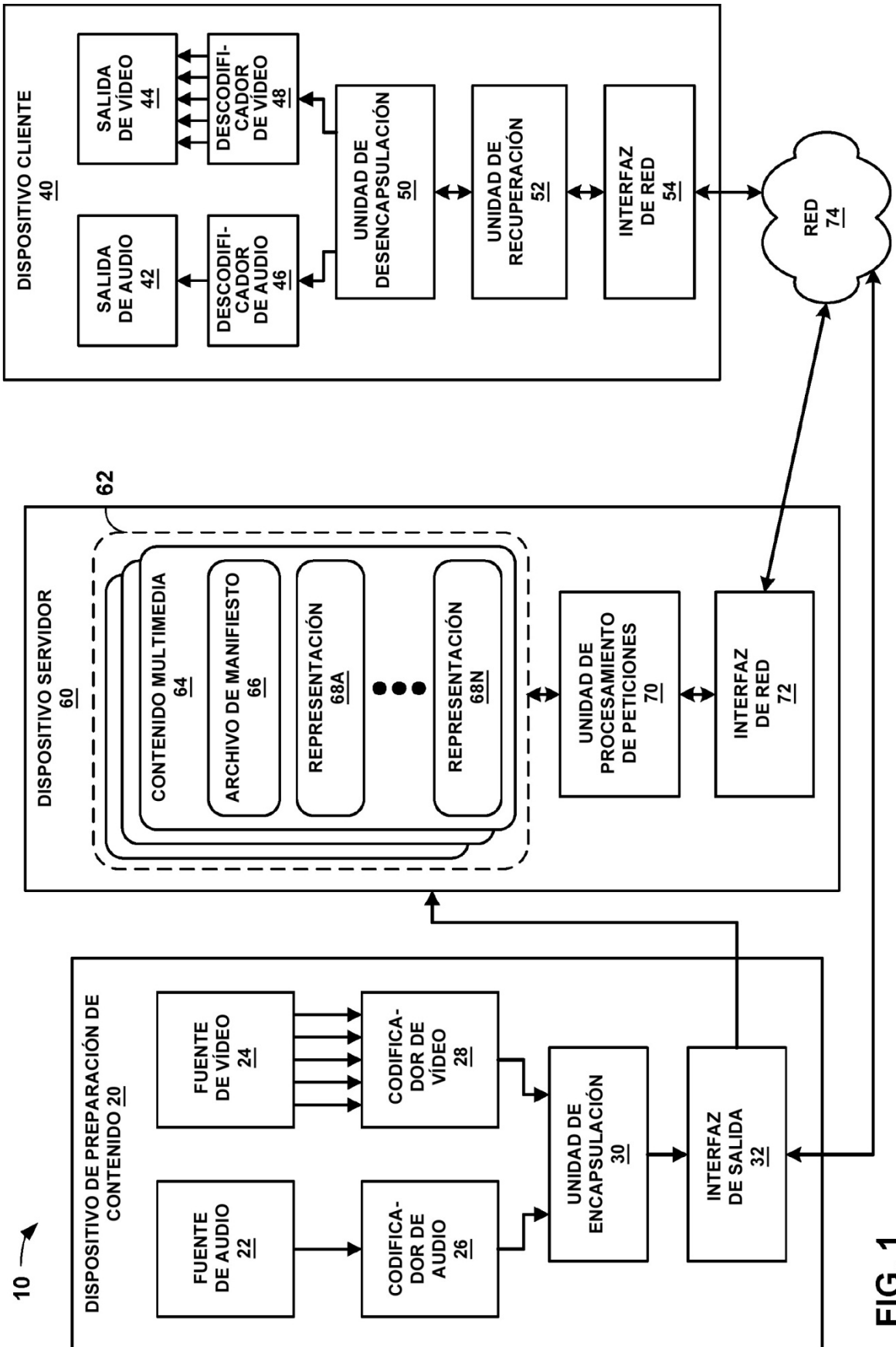
- 5 medios (54) para determinar una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP, DASH, a partir de una descripción de una instancia de la sesión de transporte de codificación en capas, LCT, LSID, en el que la LSID incluye información representativa de una pluralidad de sesiones LCT, con cada una de las sesiones LCT que incluye datos de una respectiva de las representaciones y en el que la LSID indica correspondencias entre las sesiones LCT y las representaciones; y
- 10 medios (52) para iniciar el consumo de una o más de las representaciones de la presentación de medios DASH usando la LSID, en el que los medios para iniciar el consumo comprenden:
- 15 determinar las correspondencias entre las sesiones LCT y las representaciones basadas en la LSID;
- medios (52) para recibir un primer conjunto de datos que incluyen paquetes de las sesiones LCT que incluyen partes de datos de una o más de las representaciones hasta un primer tiempo de reproducción antes de recibir un archivo de manifiesto;
- 20 medios (50) para proporcionar datos de los paquetes a un decodificador de medios, en el que los medios (52) para recibir están dispuestos para recibir un archivo de manifiesto después de recibir el primer conjunto de datos; y
- 25 medios para recibir un segundo conjunto de datos, diferente del primer conjunto de datos, de la presentación de medios DASH utilizando el archivo de manifiesto, teniendo el segundo conjunto de datos tiempos de reproducción siguientes al primer tiempo de reproducción.

11. Un procedimiento de envío de datos de medios, comprendiendo el procedimiento:

- 30 construir (554) una descripción de instancia de sesión de transporte de codificación en capas, LCT, LSID, que incluye información representativa de una pluralidad de sesiones LCT, incluyendo cada una de las sesiones LCT datos de una respectiva de una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP, DASH, en el que la LSID indica correspondencias entre las sesiones LCT y las representaciones;
- 35 emitir (556) la LSID; y
- emitir un primer conjunto de datos que incluyen los datos (556) de las representaciones en las sesiones LCT correspondientes hasta un primer tiempo de reproducción, antes de emitir un archivo de manifiesto;
- 40 emitir un archivo de manifiesto después de emitir el primer conjunto de datos; y
- 45 emitir un segundo conjunto de datos, diferente del primer conjunto de datos, de la presentación de medios DASH en respuesta a una o más peticiones basadas en el archivo de manifiesto, teniendo el segundo conjunto de datos tiempos de reproducción después del primer tiempo de reproducción.

12. Un dispositivo (60) para enviar datos de medios, comprendiendo el dispositivo:

- 50 medios (62) para construir una descripción de instancia de sesión de transporte de codificación en capas, LCT, LSID, que incluye información representativa de una pluralidad de sesiones LCT, incluyendo cada una de las sesiones LCT datos de una respectiva de una pluralidad de representaciones de una presentación de medios de transmisión adaptativa dinámica por HTTP, DASH, en el que la LSID indica correspondencias entre las sesiones LCT y las representaciones;
- 55 medios (72) para emitir la LSID; y
- medios (72) para emitir un primer conjunto de datos que incluyen los datos de las representaciones en las sesiones LCT correspondientes hasta un primer tiempo de reproducción, antes de emitir un archivo de manifiesto, en el que los medios (72) para emitir datos están dispuestos para emitir un manifiesto archivo después de emitir el primer conjunto de datos; y
- 60 medios (72) para emitir un segundo conjunto de datos, diferente del primer conjunto de datos, de la presentación de medios DASH en respuesta a una o más peticiones basadas en el archivo de manifiesto, teniendo el segundo conjunto de datos tiempos de reproducción después del primer tiempo de reproducción.
- 65



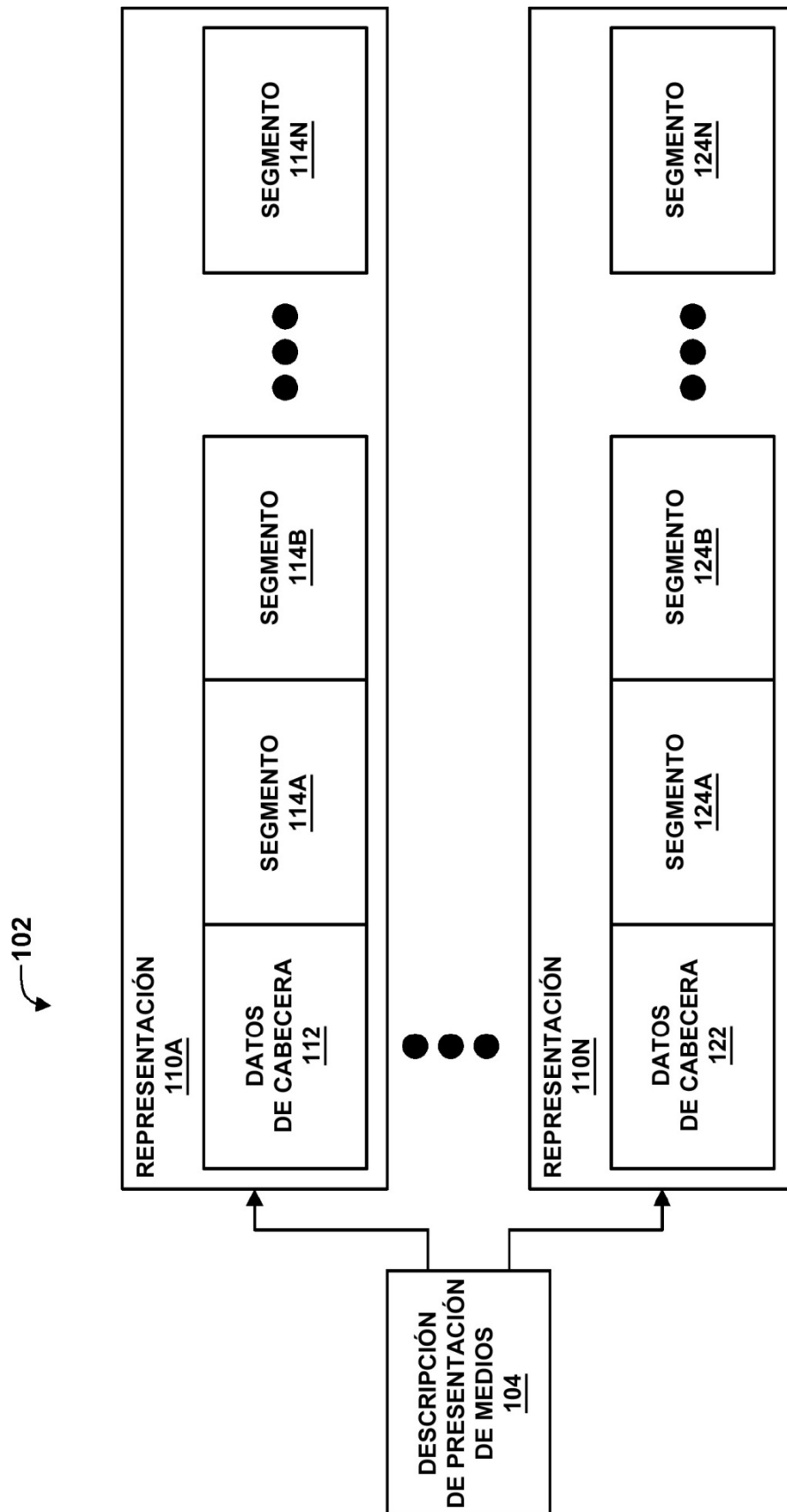


FIG. 2

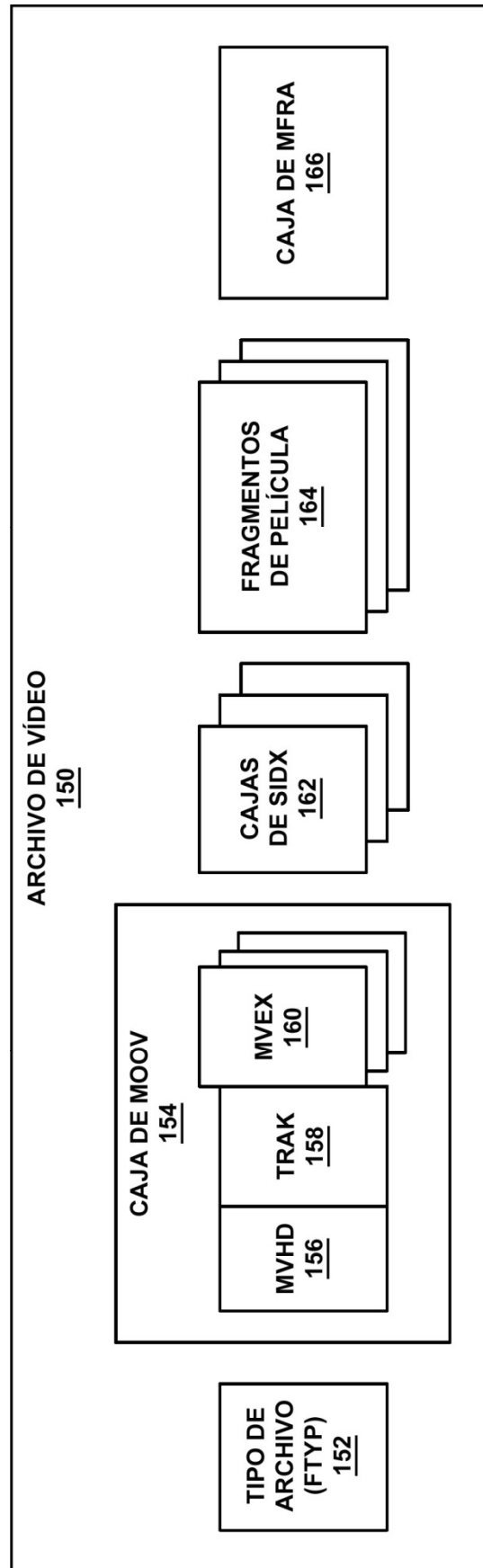


FIG. 3

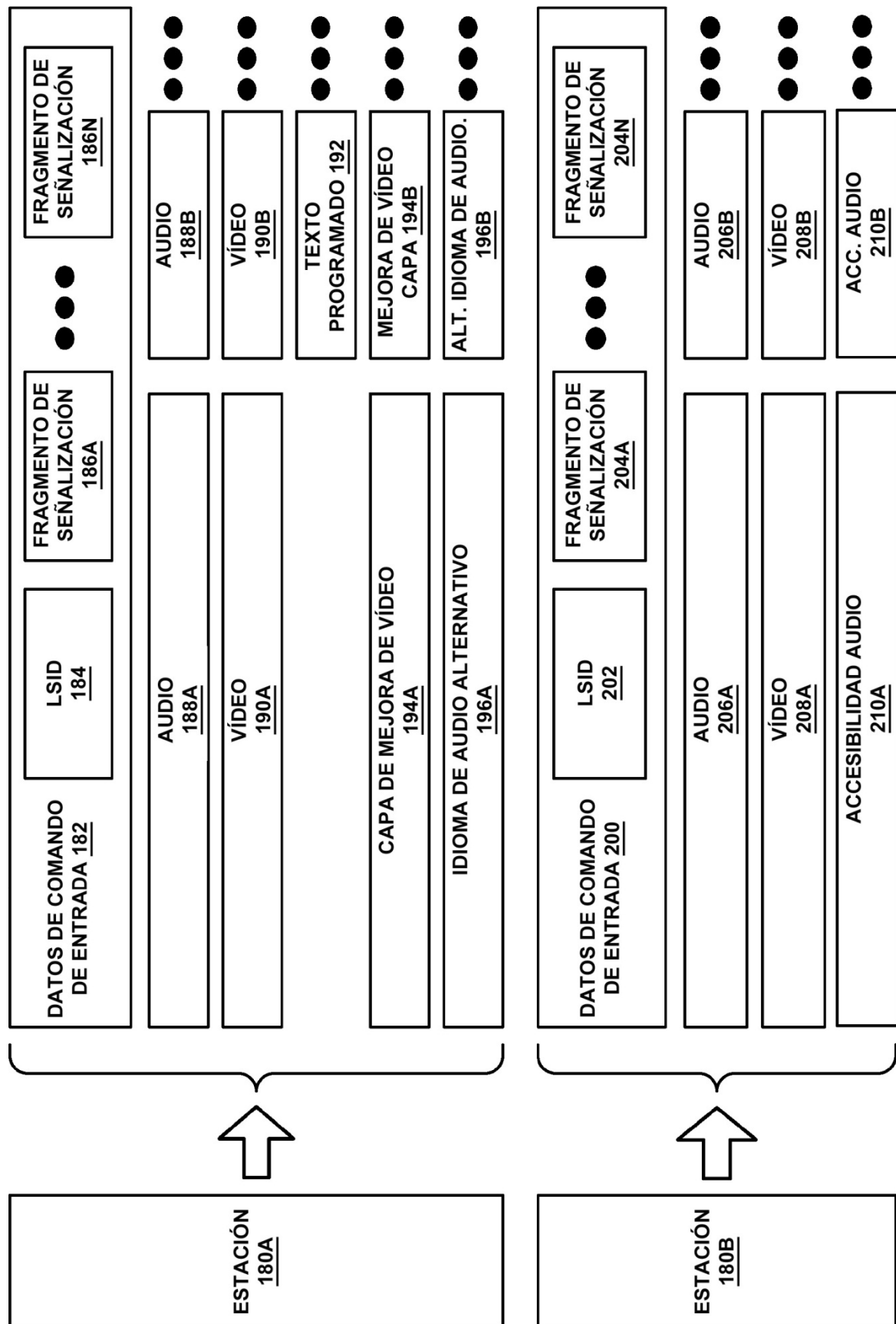


FIG. 4

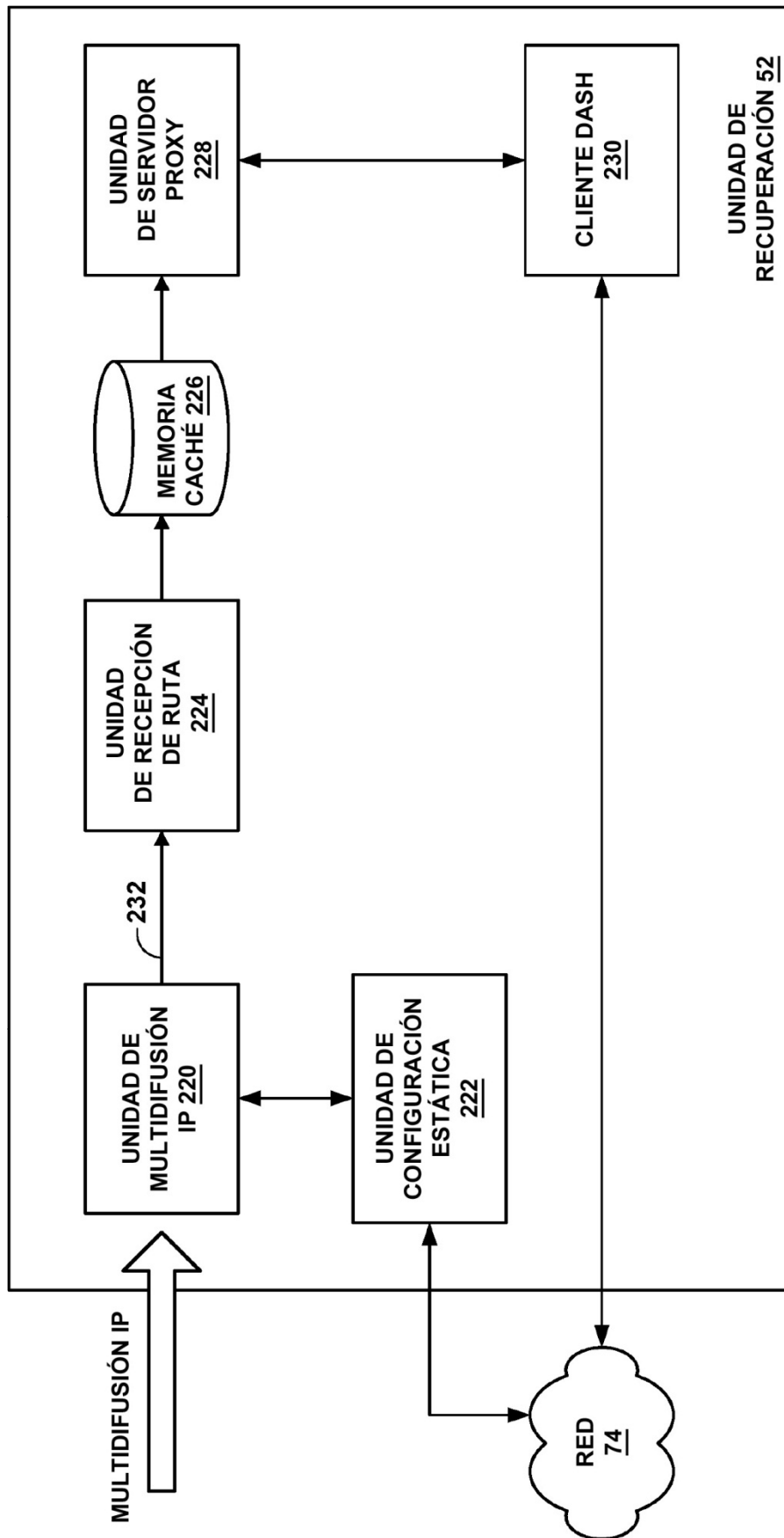


FIG. 5

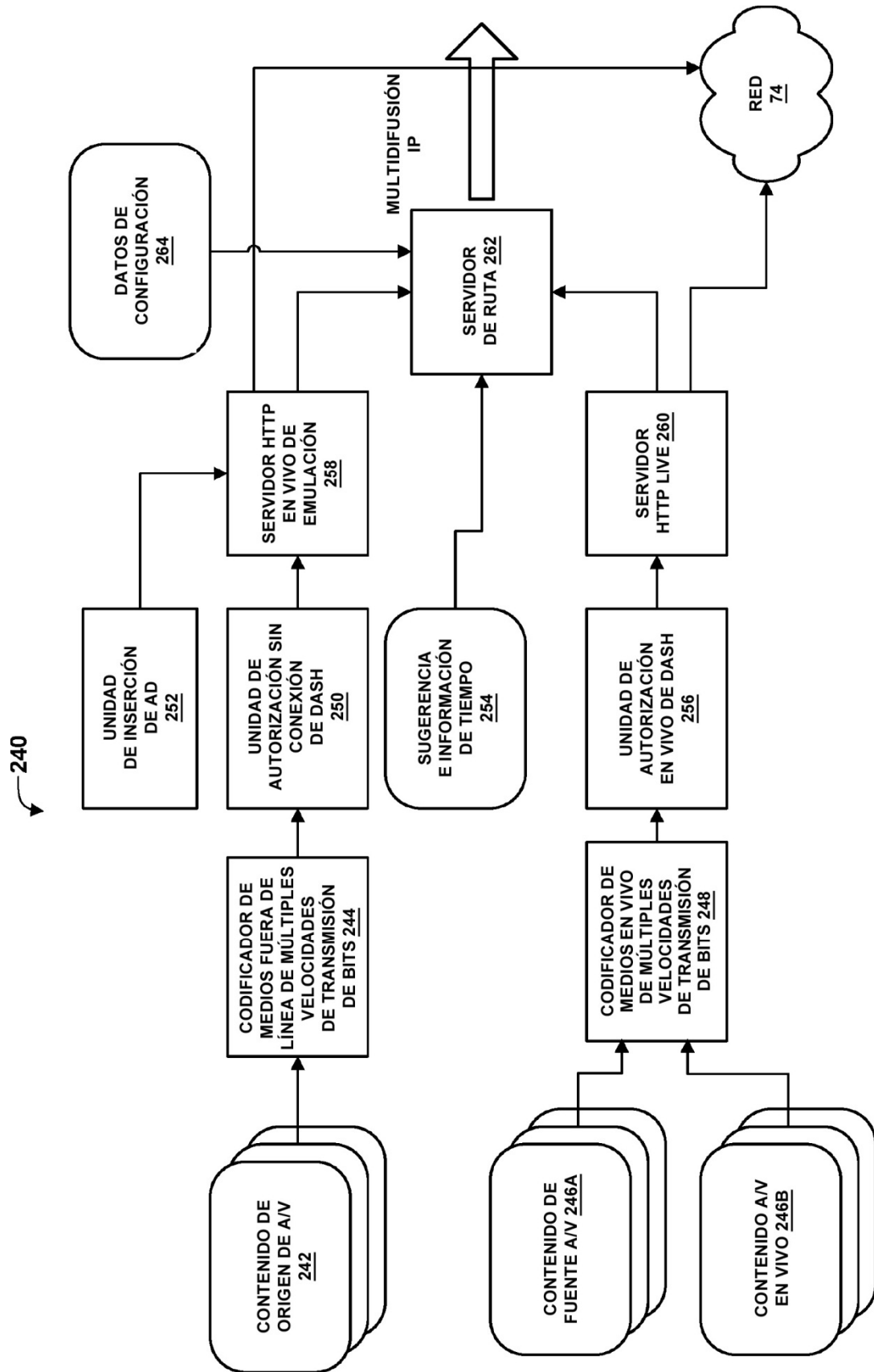


FIG. 6

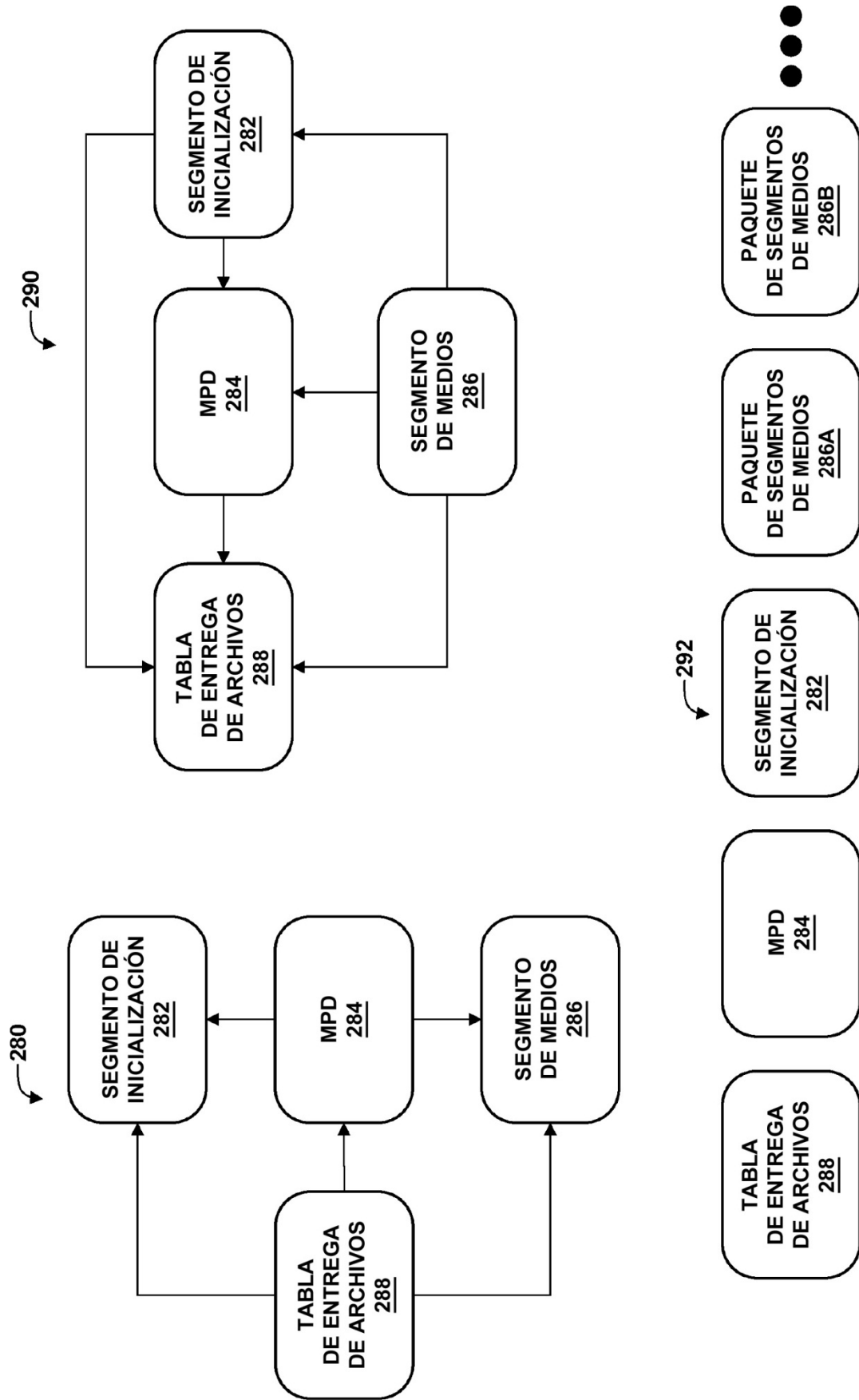
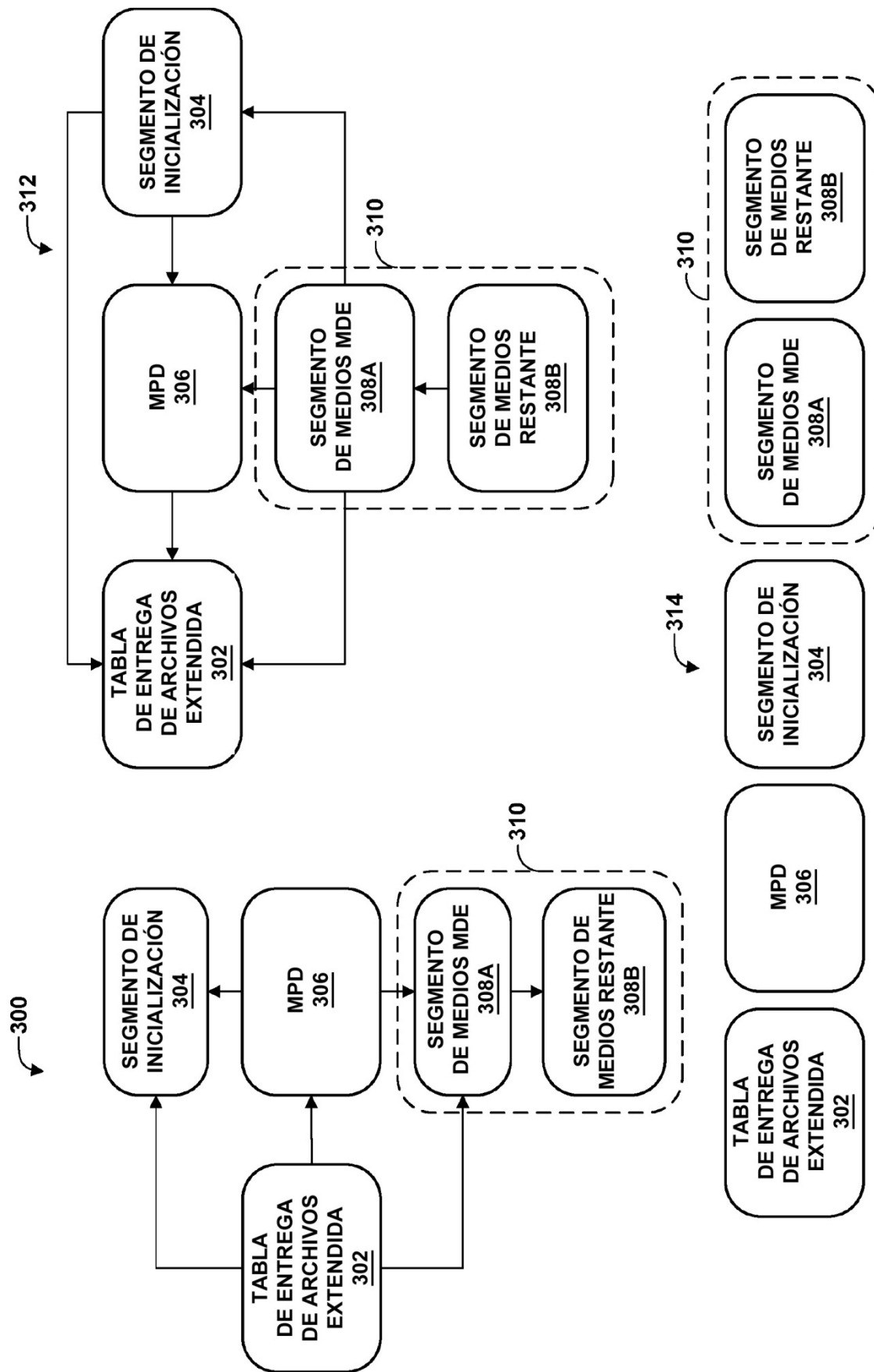


FIG. 7



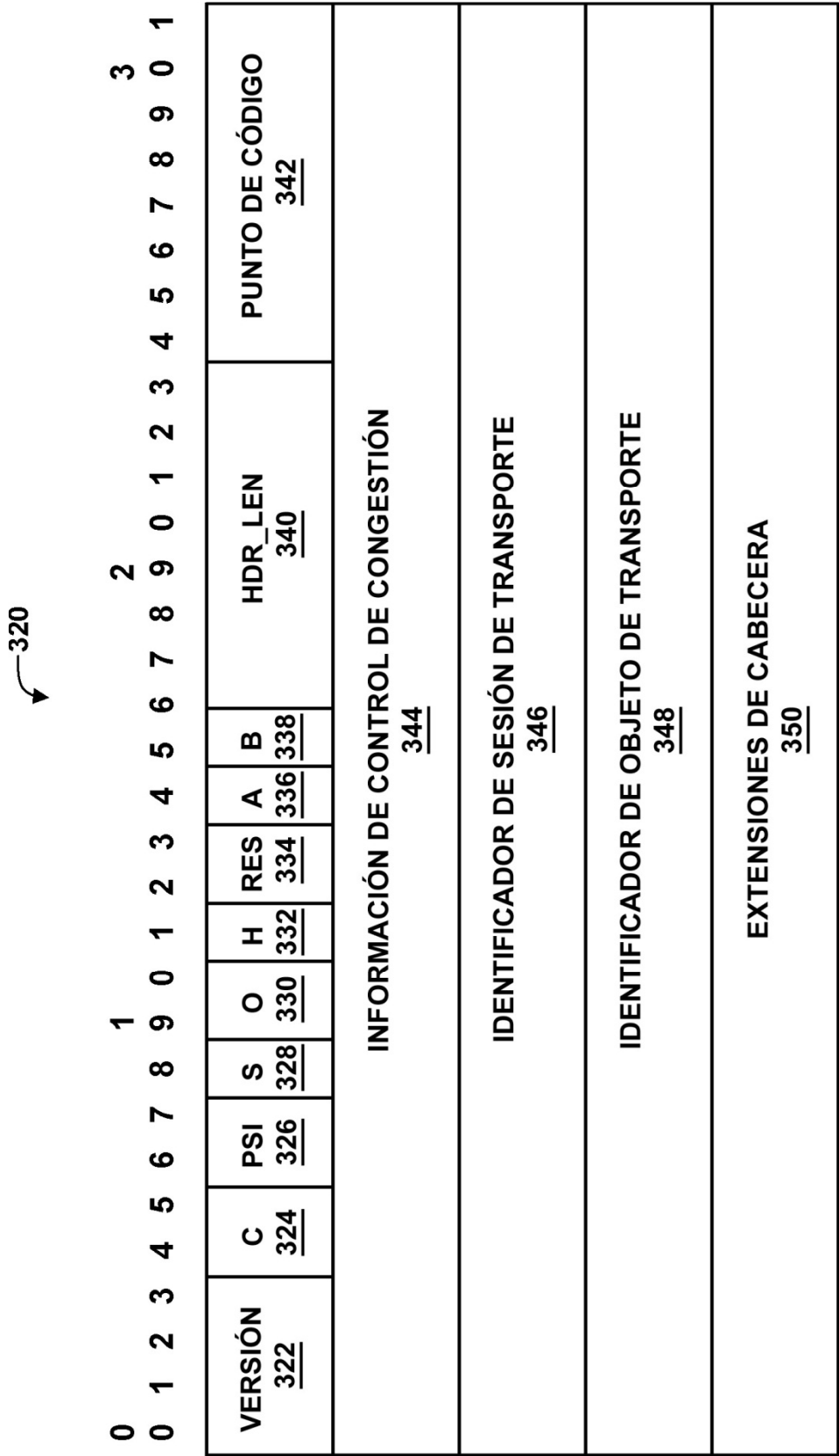


FIG. 9

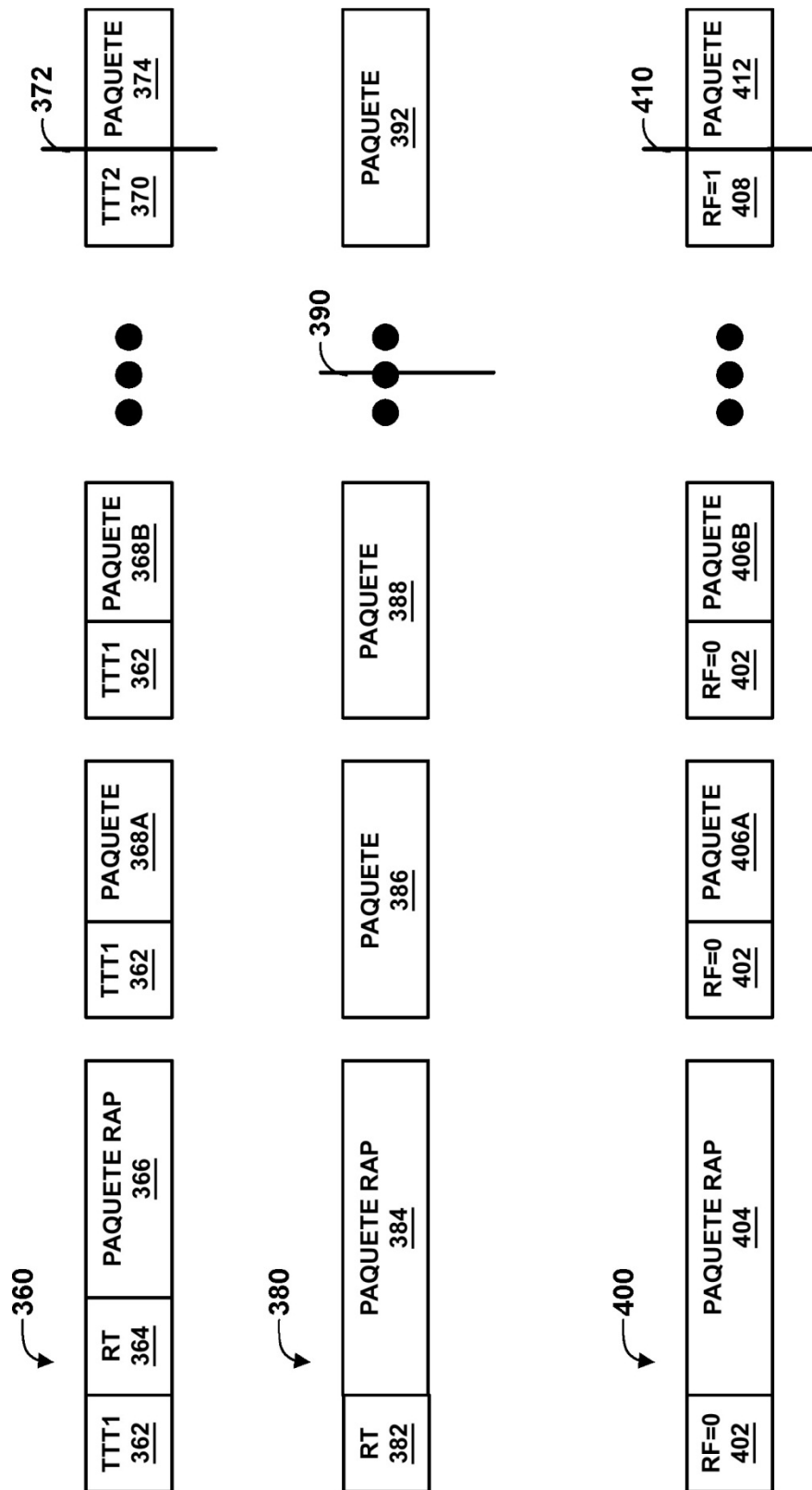


FIG. 10

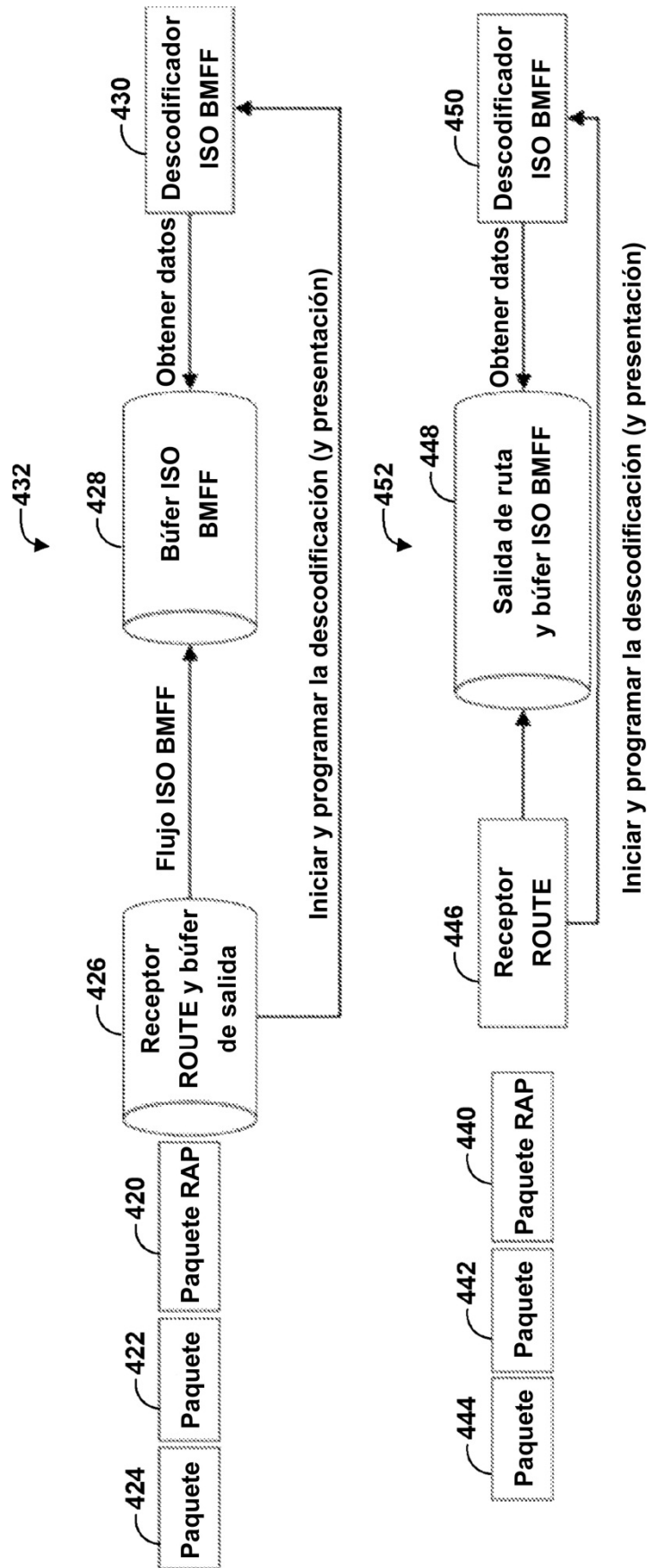


FIG. 11

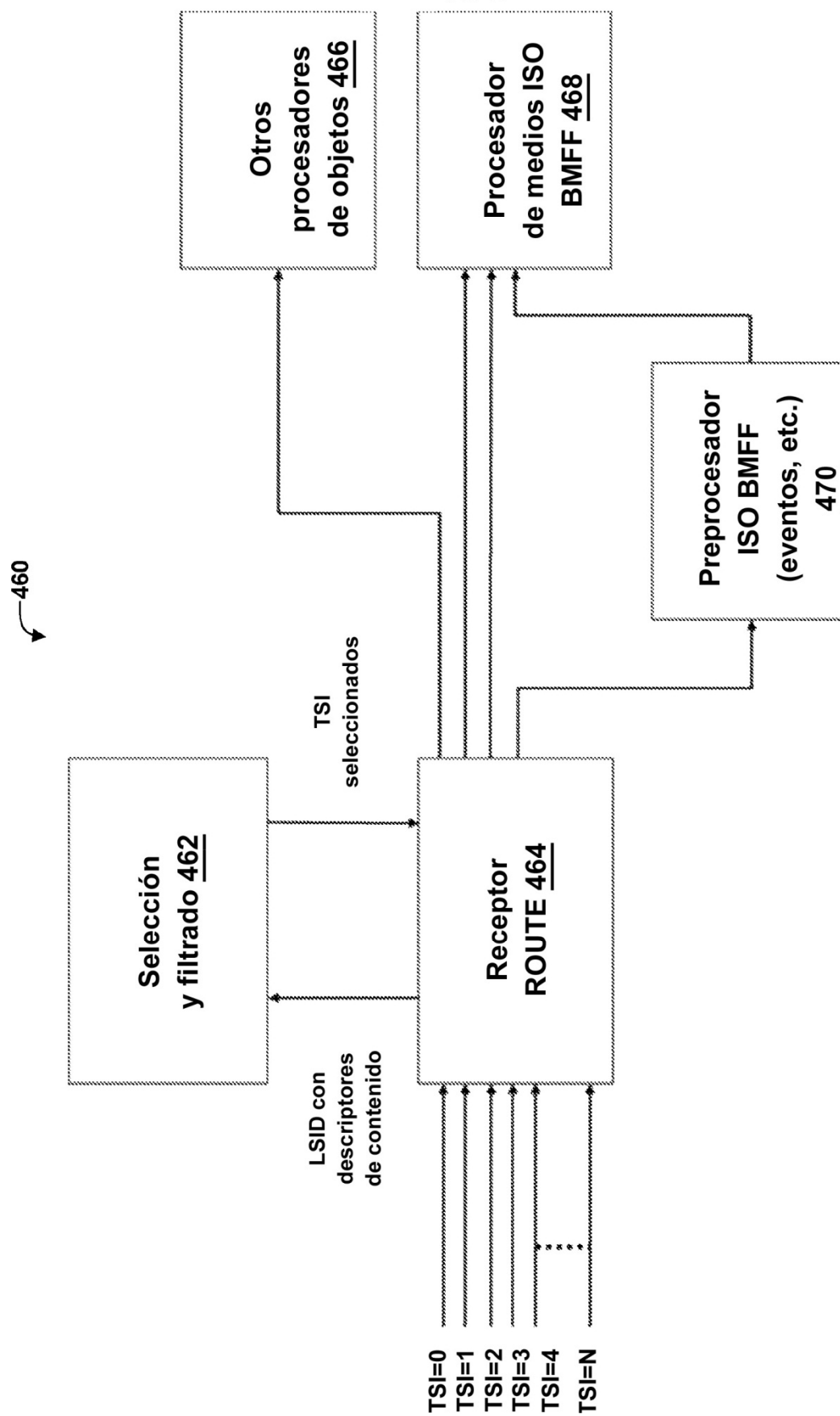


FIG. 12

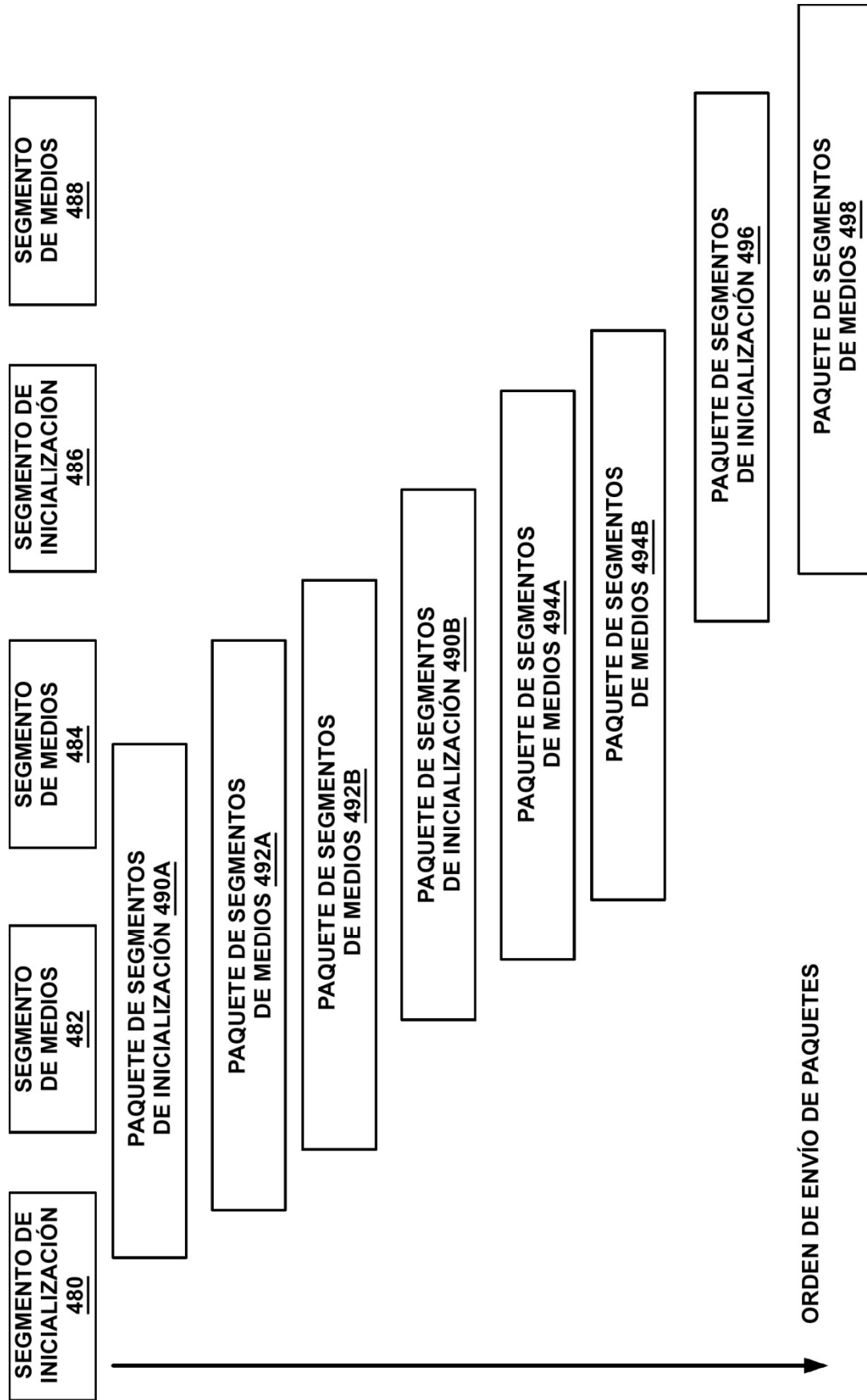


FIG. 13

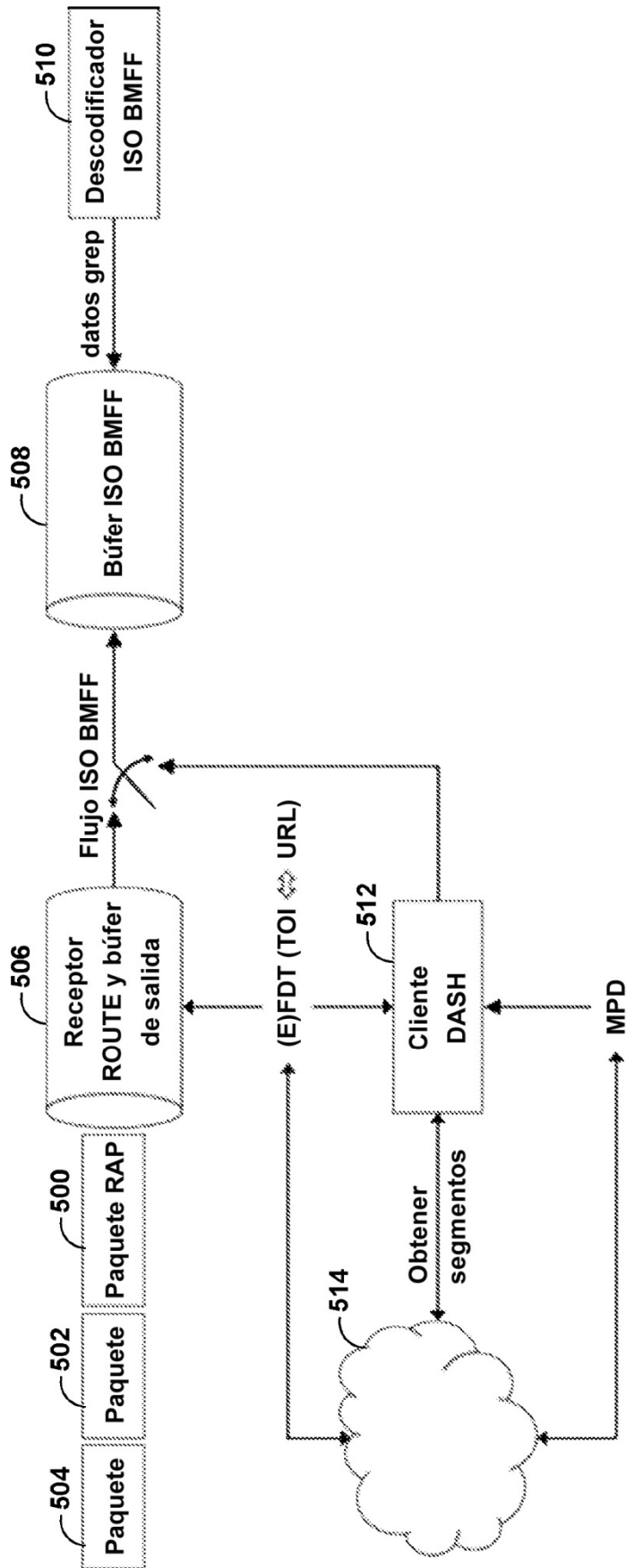


FIG. 14

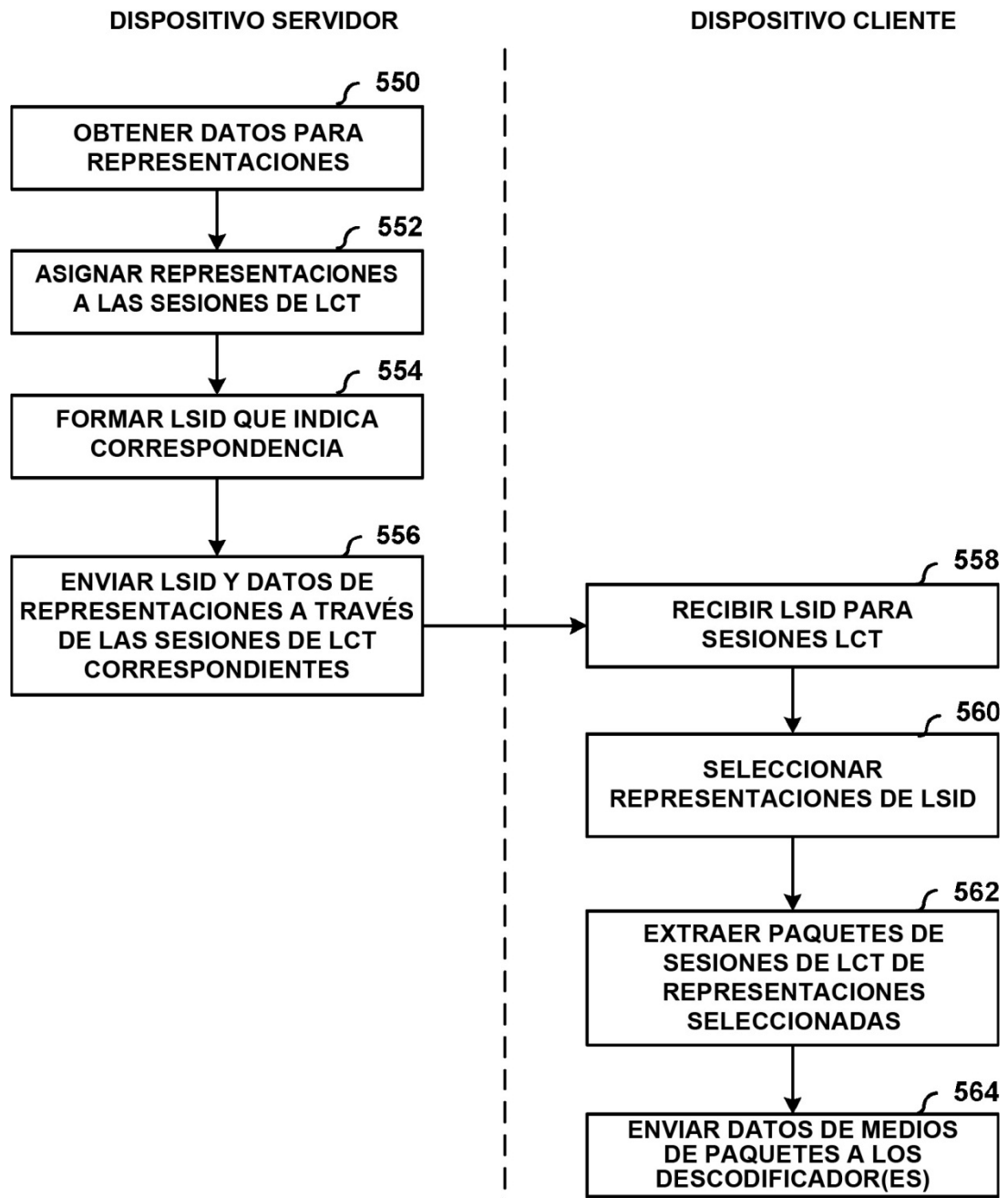


FIG. 15

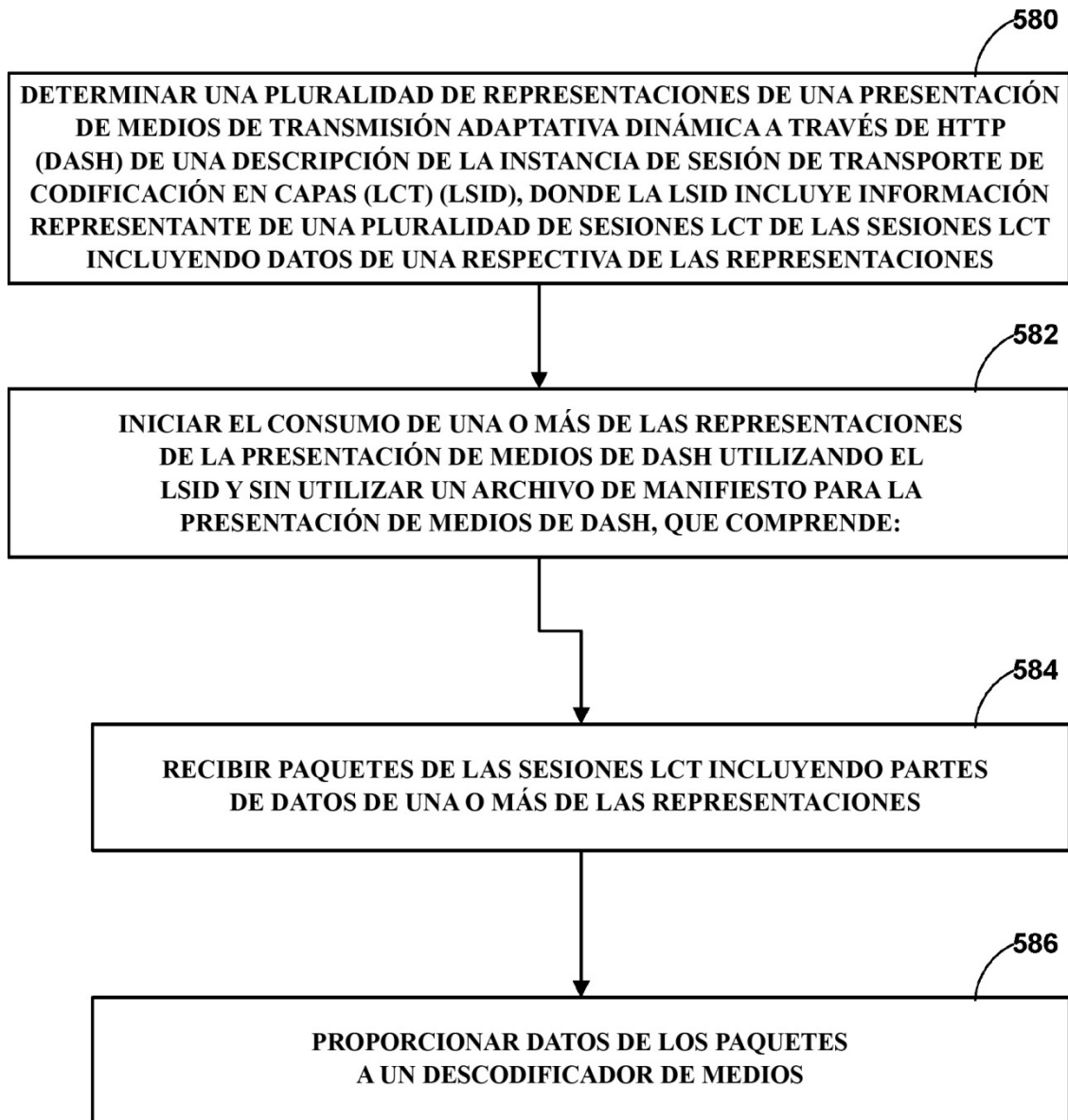


FIG. 16

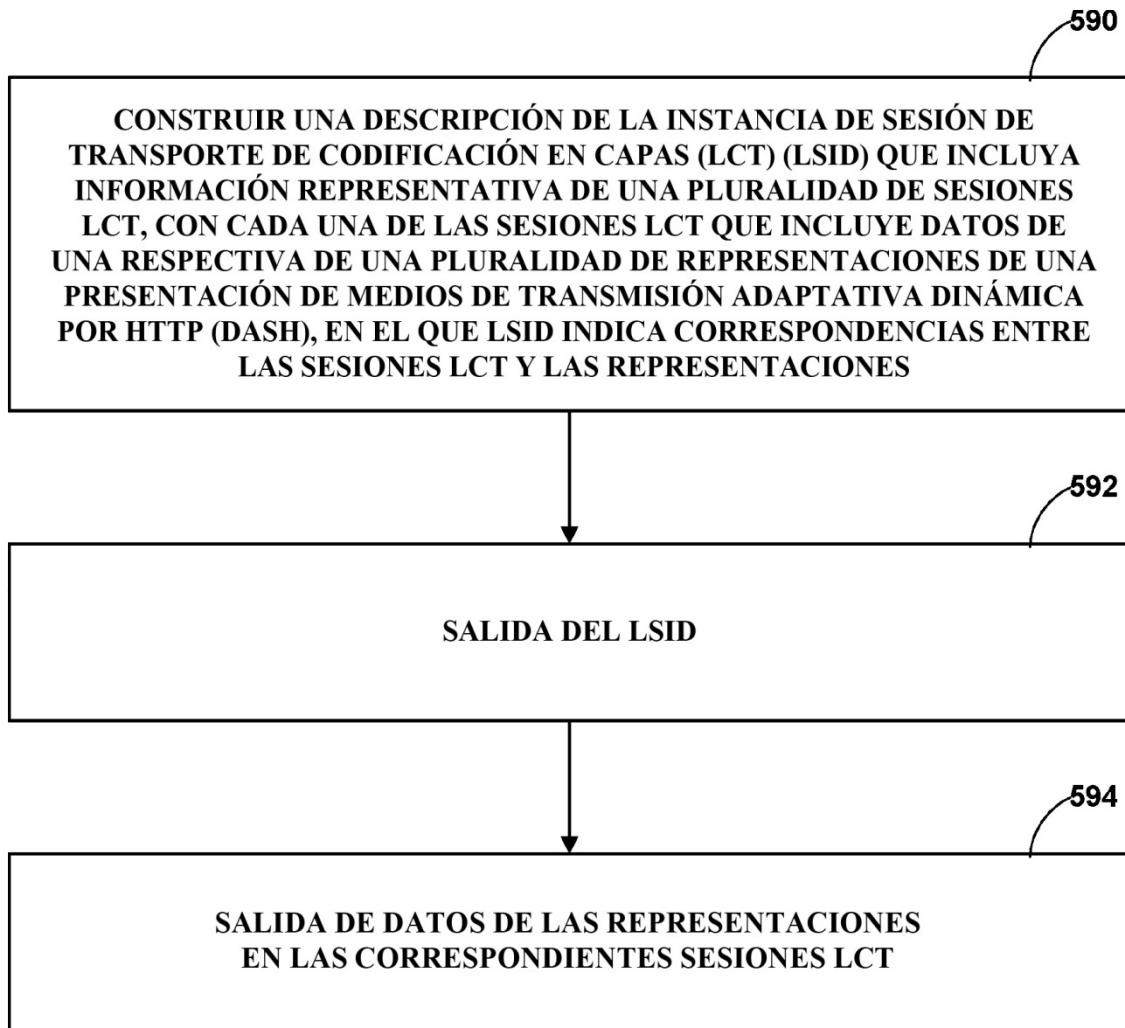


FIG. 17