

DETERMINACIÓN DE UBICACIONES DE EVENTO DE ENTREGA DE MEDIOS PARA TRANSPORTE DE MEDIOS

[0001] Esta solicitud reivindica el beneficio de la Solicitud Provisoria de los Estados Unidos No. **62/276.674**, presentada el 8 de enero de 2016, cuyo contenido completo se incorpora por referencia.

CAMPO TÉCNICO

[0002] La divulgación se refiere al transporte de datos de medios.

ANTECEDENTES

[0003] Se pueden incorporar capacidades de audio y video digitales en una amplia gama de dispositivos, que incluyen televisores digitales, sistemas digitales de transmisión directa, sistemas de transmisión inalámbrica, asistentes digitales personales (PDA), computadoras portátiles o de escritorio, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de juegos de video, consolas de videojuegos, teléfonos celulares o de radio por satélite, dispositivos de teleconferencia de video, y similares.

[0004] Los datos de medios, tales como datos de audio y video, se pueden comprimir para el transporte. La compresión generalmente incluye la codificación de los datos de medios. Después de que los datos de medios se hayan codificado, los datos de video se pueden empaquetar para su transmisión o almacenamiento. Los datos de medios pueden ensamblarse en un archivo multimedia conforme con cualquiera de una variedad de estándares, tales como la Organización Internacional de Normalización de Formato de Archivos de Medios de Base (ISO BMFF) y extensiones de los mismos, tales como AVC.

[0005] Se pueden utilizar diversas técnicas para el transporte de datos de medios a través de una red informática. Por ejemplo, los datos de los medios pueden entregarse a través de un protocolo de difusión o un protocolo de unidifusión. En un protocolo de difusión, un dispositivo servidor envía datos de medios a una pluralidad de dispositivos suscripción de cliente. En un protocolo de unidifusión, un dispositivo cliente solicita los datos de medios desde un dispositivo servidor, y el dispositivo servidor proporciona los datos de medios en respuesta a la solicitud.

SÍNTESIS

[0006] En general, esta divulgación describe las técnicas relacionadas con el soporte de Eventos de Entrega de Medios (MDE) utilizados en, por ejemplo, entrega de

intervalo de bytes consciente de medios mediante ROUTE (ATSC Working Draft: Signaling, Delivery, Synchronization, and Error Protection, 20 de noviembre de 2015) y MMT (ISO/IEC: ISO/IEC 23008-1 2.^a edición, Information technology—High efficiency coding and media delivery in heterogeneous environments—Part 1: MPEG media transport (MMT)) o FLUTE (RFC 6726, que es un subconjunto adecuado de ROUTE. Estas técnicas de administración pueden utilizar además difusión de Streaming Adaptativo Dinámico a través de HTTP (DASH), Formato de Aplicación de Medios Comunes (CMAF), o Segmento similar o métodos de entrega de medios basados en intervalos de bytes conscientes de medios. Esta divulgación también describe técnicas para determinar el etiquetado de tiempo de MDE detectados y técnicas adaptativas para la entrega en la capa física que tienen en cuenta el impacto de entrega de medios temprana con relación a las tramas de capa física alineadas por medios.

[0007] En un ejemplo, un método de transporte de datos de medios incluye, mediante una unidad de envío de protocolo basado en archivos de un dispositivo fuente, recibir un flujo de datos que comprende Segmentos de datos de medios de un segmentador del dispositivo fuente que forma los Segmentos, cada uno de los Segmentos comprende un respectivo archivo individualmente recuperable asociado con un localizador uniforme de recursos (URL) único o un identificador uniforme de recursos (URI), determinar ubicaciones de eventos de entrega de medios (MDE) en el flujo de datos de medios, donde los MDE incluyen datos para al menos una porción de uno de los Segmentos, determinar uno o más requisitos de tiempo de transmisión para los MDE que representan tiempos en los que los MDE han de enviarse a un dispositivo cliente, y proporcionar los MDE y datos que representan los requisitos de tiempo de transmisión a una unidad de envío de capa física del dispositivo fuente de acuerdo con ranuras de entrega disponibles para la unidad de envío de capa física.

[0008] En otro ejemplo, un dispositivo para el transporte de datos de medios incluye un segmentador configurado para formar un flujo de datos que comprende Segmentos de datos de medios, cada uno de los Segmentos comprende un respectivo archivo individualmente recuperable, por ejemplo, formato ISO BMFF asociado con un localizador uniforme de recursos (URL) único o un identificador uniforme de recursos (URI). El dispositivo también incluye una unidad de envío de capa física configurada para entregar eventos de entrega de medios (MDE) a un dispositivo cliente de acuerdo con los requisitos de tiempo de transmisión para los MDE, donde los MDE incluyen datos para al menos una porción de uno de los Segmentos, donde la unidad de envío de capa física está configurada con ranuras de entrega disponibles para la recepción

de datos a entregar. El dispositivo incluye además una unidad de envío de protocolo basado en archivos configurada para recibir el flujo de datos que comprende los Segmentos de datos de medios del Segmentador, determinar ubicaciones de los MDE en el flujo de datos de medios, determinar uno o más de los requisitos de tiempo de transmisión para los MDE que representan tiempos en los que los MDE han de enviarse al dispositivo cliente, y proporcionar los MDE y datos que representan los requisitos de tiempo de transmisión a la unidad de envío de capa física de acuerdo con las ranuras de entrega disponibles para la unidad de envío de capa física.

[0009] En otro ejemplo, un dispositivo para el transporte de datos de medios incluye un segmentador configurado para formar un flujo de datos que comprende Segmentos de datos de medios, cada uno de los Segmentos comprende un respectivo archivo individualmente recuperable asociado con un localizador uniforme de recursos (URL) único o un identificador uniforme de recursos (URI). El dispositivo también incluye una unidad de envío de capa física configurada para entregar eventos de entrega de medios (MDE) a un dispositivo cliente de acuerdo con los requisitos de tiempo de transmisión para los MDE, donde los MDE incluyen datos para al menos una porción de uno de los Segmentos, donde la unidad de envío de capa física está configurada con ranuras de entrega disponibles para la recepción de datos a entregar. El dispositivo incluye además medios para recibir un flujo de datos que comprende Segmentos de datos de medios del segmentador, medios para determinar ubicaciones de los MDE en el flujo de datos de medios, medios para determinar uno o más requisitos de tiempo de transmisión para los MDE que representan tiempos en los que los MDE han de enviarse al dispositivo cliente, y medios para proporcionar los MDE y datos que representan los requisitos de tiempo de transmisión a la unidad de envío de capa física de acuerdo con ranuras de entrega disponibles para la unidad de envío de capa física.

[0010] En otro ejemplo, un medio de almacenamiento legible por computadora que tiene almacenado en el mismo instrucciones que, al ejecutarse, hacen a una unidad de envío de protocolo basado en archivos de un dispositivo fuente recibir un flujo de datos que comprende Segmentos de datos de medios de un segmentador del dispositivo fuente que forma los Segmentos, cada uno de los Segmentos comprende un respectivo archivo individualmente recuperable asociado con un localizador uniforme de recursos (URL) único o un identificador uniforme de recursos (URI), determinar las ubicaciones de los eventos de entrega de medios (MDE) en el flujo de datos de medios, donde los MDE incluyen datos de al menos una porción de uno de los

Segmentos, determinar uno o más requisitos de tiempo de transmisión para los MDE que representan tiempos en los que los MDE han de enviarse a un dispositivo cliente, y proporcionar los MDE y datos que representan los requisitos de tiempo de transmisión a una unidad de envío de capa física del dispositivo fuente de acuerdo con ranuras de entrega disponibles para la unidad de envío de capa física.

[0011] Los detalles de uno o más ejemplos se exponen en las figuras adjuntas y la descripción siguiente. Otras características, objetos y ventajas serán evidentes a partir de la descripción y las figuras, y de las reivindicaciones.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0012] La FIG. 1 es un diagrama de bloques que ilustra un sistema de ejemplo que implementa técnicas para la transmisión de datos de medios sobre una red.

[0013] La FIG. 2 es un diagrama de bloques que ilustra un sistema de ejemplo de componentes de la unidad de recuperación 52 de la FIG. 1 con mayor detalle.

[0014] La FIG. 3 es un diagrama conceptual que ilustra elementos de contenido multimedia de ejemplo.

[0015] La FIG. 4 es un diagrama de bloques que ilustra elementos de un archivo de video de ejemplo, que puede corresponder a un Segmento de una Representación.

[0016] La FIG. 5 es un diagrama de bloques que ilustra un Streaming Adaptativo Dinámico de difusión de ejemplo a través de sistema de infraestructura (DASH) HTTP.

[0017] La FIG. 6 es un diagrama conceptual que ilustra una composición de alto nivel de ejemplo de MDE.

[0018] La FIG. 7 es un diagrama conceptual que ilustra una cadencia de tipo de trama de ejemplo y MDE definidos correspondientes.

[0019] La FIG. 8 es un diagrama conceptual que representa una vista simple del tiempo del sistema.

[0020] La FIG. 9 es un diagrama de bloques que representa una relación de tiempo en una infraestructura de transmisión.

[0021] La FIG. 10 es un diagrama conceptual que ilustra múltiples Segmentos por grupo de imágenes (GOP) con una ubicación escalonada de punto de acceso aleatorio (RAP).

[0022] La FIG. 11 es un diagrama conceptual que ilustra un retardo mínimo para decodificar un tiempo para la reproducción de medios.

[0023] La FIG. 12 es un diagrama conceptual que ilustra sincronización de capa física opcional y metadatos relacionados.

[0024] La FIG. 13 es un diagrama de flujo que ilustra un ejemplo de secuencia de

eventos de adquisición.

[0025] La FIG. 14 es un diagrama de flujo que ilustra otro método de ejemplo de acuerdo con las técnicas de la presente divulgación.

DESCRIPCIÓN DETALLADA

[0026] En general, esta divulgación describe técnicas relacionadas con el transporte de datos de medios. Las técnicas de la presente divulgación pueden realizarse por un dispositivo servidor u otro dispositivo fuente que transmite datos de medios a un dispositivo cliente. En particular, el dispositivo fuente puede incluir un segmentador, una unidad de envío de protocolo basado en archivos, y una unidad de envío de capa física. El segmentador puede corresponder a una unidad que forma Segmentos de datos de medios, donde los Segmentos se ajustan a, por ejemplo, Streaming Adaptativo Dinámico a través de Segmentos HTTP (DASH). Por ejemplo, el segmentador puede encapsular unidades de capa de abstracción de red (NAL) en Segmentos DASH respectivos, donde los Segmentos pueden estar asociados con un localizador uniforme de recursos (URL) distinto o un identificador uniforme de recursos (URI) diferente. La unidad de envío de protocolo basado en archivos puede estar configurada para enviar datos mediante el uso de un protocolo basado en archivos, tal como Entrega de Archivos sobre Transporte Unidireccional (FLUTE) o ROUTE. La unidad de envío de capa física puede enviar datos de medios en la capa física, por ejemplo, mediante el uso de Ethernet.

[0027] De acuerdo con las técnicas de esta divulgación, el segmentador puede proporcionar ambos Segmentos de datos de medios, en forma de un flujo de los Segmentos, a la unidad de envío de protocolo basado en archivos. Por lo tanto, la unidad emisora de protocolo basado en archivos puede recibir un flujo de datos que incluye los Segmentos de datos de medios del segmentador. La unidad de envío de protocolo basado en archivos luego puede determinar ubicaciones de eventos de entrega de medios (MDE) en el flujo de datos de medios. En un ejemplo, la unidad de envío de protocolo basado en archivos puede formar los MDE del flujo de datos de medios. Por ejemplo, la unidad de envío de protocolo basado en archivos puede determinar qué partes del flujo han de incluirse en un MDE en particular, de manera tal que el MDE incluya datos que sean útiles para un decodificador medios (por ejemplo, datos que pueden ser decodificados correctamente por el decodificador de medios, suponiendo que los datos más tempranos se entreguen correctamente).

[0028] Para determinar las ubicaciones de los MDE, la unidad de envío de protocolo basado en archivos puede utilizar técnicas de coincidencia de patrones, en las que la

unidad de envío de protocolo basado en archivos está configurada con uno o más patrones para los datos de medios incluidos en el flujo. Tales patrones pueden representar, por ejemplo, una disposición de datos de medios en un orden de decodificación jerárquica. Por ejemplo, los datos de video pueden estar dispuestos de acuerdo con un cierto número de cuadros intra-predichos (cuadros I), seguido por un cierto número de cuadros inter-predichos unidireccionales (cuadros P), seguido de (o intercalados con) un cierto número de cuadros inter-predichos bidireccionales (cuadros B). La unidad de envío de protocolo basado en archivos puede determinar las ubicaciones de MDE con base en la detección de los cuadros I, P, y B y, por ejemplo, determinar que un MDE incluye una secuencia de video de un primer cuadro I y terminar con un último cuadro B antes de un cuadro I subsiguiente, como un ejemplo. Además, la unidad de envío de protocolo basado en archivos puede determinar que un video MDE corresponde a un intervalo de bytes particular de un Segmento, donde el intervalo de bytes termina en un límite de cuadro. Como otro ejemplo, la unidad de envío de protocolo basado en archivos puede determinar que los MDE de audio incluyen un número fijo de tramas de audio.

[0029] Como otro ejemplo, para determinar las ubicaciones de los MDE (por ejemplo, para formar los MDE), la unidad de envío de protocolo basado en archivos se puede configurar de acuerdo con reglas que definen ubicaciones de los diferentes tipos de datos de medios, tales como datos de audio, datos de video y/o datos de texto temporizado, en el flujo. Las reglas pueden definir, por ejemplo, una disposición de los diversos tipos de datos de medios en el flujo, de manera tal que la unidad de envío de protocolo basado en archivos puede diferenciar entre los tipos de datos de medios y asignar los tipos de datos de medios a los respectivos MDE.

[0030] Además, la unidad de envío de protocolo basado en archivos también puede determinar uno o más requisitos de tiempo de entrega (también denominados requisitos de tiempo de transmisión) para los MDE, cuando los requisitos de tiempo de transmisión representan tiempos en los que los MDE han de enviarse al dispositivo cliente. Por ejemplo, los requisitos de tiempo de transmisión pueden representar tiempos más tempranos y/o más posteriores en los que los MDE han de enviarse al dispositivo cliente. Los MDE puede tener requisitos de tiempo más tempranos donde los MDE se pueden enviar al dispositivo cliente con el fin de, por ejemplo, evitar un desbordamiento de memoria en el dispositivo cliente. Los MDE también puede tener requisitos temporales más posteriores en los que los MDE se pueden enviar al dispositivo cliente con el fin de, por ejemplo, evitar una insuficiencia de memoria

intermedia en el dispositivo cliente.

[0031] La unidad de envío de protocolo basado en archivos luego puede enviar los MDE a la unidad de envío de capa física, junto con datos representativos de los requisitos de tiempo de transmisión. En particular, la unidad de envío de protocolo basado en archivos puede determinar ranuras de entrega disponibles para la unidad de envío de capa física, y enviar los datos de solicitud de MDE y de tiempo de transmisión a la unidad de envío de capa física de manera tal que la unidad de envío de capa física sea capaz de entregar los MDE de acuerdo con los requisitos de tiempo de transmisión.

[0032] El dispositivo fuente analizado con anterioridad puede estar configurado para transmitir datos de acuerdo con un protocolo de transmisión HTTP, tal como DASH. En el flujo HTTP, las operaciones de uso frecuente incluyen HEAD, GET y GET parcial. La operación HEAD recupera una cabecera de un archivo asociado con un determinado localizador uniforme de recursos (URL), nombre uniforme de recursos (URN), o identificador uniforme de recursos (URI), sin recuperar una carga útil asociada a la dirección URL, URN, o URI. La operación GET recupera un archivo completo asociado a un URL, URN, o URI determinada. La operación GET parcial recibe un intervalo de bytes como parámetro de entrada y recupera un número continuo de bytes de un archivo, donde el número de bytes corresponde al intervalo de byte recibido. Por lo tanto, se pueden proporcionar fragmentos de películas para el flujo HTTP, debido a que una operación GET parcial puede obtener uno o más fragmentos de películas individuales. En un fragmento de película, puede haber diversos fragmentos de pista de diferentes pistas. En el flujo HTTP, una presentación de medios puede ser una recolección estructurada de datos que es accesible para el cliente. El cliente puede solicitar y descargar información de datos de medios para presentar un servicio de streaming a un usuario.

[0033] En el ejemplo de streaming de datos 3GPP que utiliza streaming HTTP, puede haber múltiples representaciones de datos de audio y/o video de contenido multimedia. Como se explica a continuación, las diferentes representaciones pueden corresponder a diferentes características de codificación (por ejemplo, diferentes perfiles o niveles de un estándar de codificación de video), diferentes estándares de codificación o extensiones de estándares de codificación (tales como extensiones multivista y/o escalables), o diferentes velocidades de bits. El manifiesto de tales representaciones se puede definir en una estructura de datos de medios de Descripción de Presentación de Medios (MPD). Una presentación de medios puede corresponder a una recolección

estructurada de datos accesible a un dispositivo cliente de streaming HTTP. El dispositivo cliente de streaming HTTP puede solicitar y descargar información de datos de medios para presentar un servicio de streaming a un usuario del dispositivo cliente. Una presentación de medios puede describirse en la estructura de datos de MPD, que puede incluir actualizaciones de los MPD.

[0034] Una presentación de medios puede contener una secuencia de uno o más períodos. Los períodos pueden ser definidos por un elemento *Period* 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 tiempo real, la suma del atributo *start* del período y el atributo MPD *availableStartTime* puede especificar el tiempo de disponibilidad del período en formato UTC, en particular, el primer Segmento de medios de cada representación en el período correspondiente. Para los servicios bajo demanda, el atributo *start* del primer período puede ser 0. Para cualquier otro período, el atributo *start* puede especificar un desplazamiento entre el tiempo de inicio del Período correspondiente en relación con el tiempo de inicio del primer Período. Cada período puede extenderse hasta el comienzo del próximo Período, o hasta el final de la presentación de los medios en el caso del último Período. Los tiempos de inicio del período pueden ser precisos. Pueden reflejar el tiempo real resultante de reproducir los medios de todos los períodos anteriores.

[0035] 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 video. Las representaciones pueden diferir por los tipos de codificación, por ejemplo, por la tasa de bits, resolución, y/o códec para los datos de video y tasa de bits, idioma, y/o códec para datos de audio. El término representación puede utilizarse para referirse a una sección de datos de audio o video codificados correspondientes a un período particular del contenido multimedia y codificados de una manera particular.

[0036] Las representaciones de un período particular, pueden asignarse a un grupo indicado por un atributo en la MPD indicativo de un conjunto de adaptación al que pertenecen las representaciones. Las representaciones en el mismo conjunto de adaptación generalmente se consideran alternativas entre sí, dado que un dispositivo cliente puede cambiar en forma dinámica y sin problemas entre estas representaciones, por ejemplo, para llevar a cabo la adaptación del ancho de banda. Por ejemplo, cada representación de datos de video para un período particular puede asignarse al mismo conjunto de adaptación, de manera tal que cualquiera de las

representaciones se pueda seleccionar para la decodificación para presentar datos de medios, tales como datos de video o datos de audio, de los contenidos multimedia para el período correspondiente. El contenido de los medios dentro de un período puede representarse por cualquier representación de grupo 0, si está presente, o la combinación de, como máximo, una representación de cada grupo distinto de cero, en algunos ejemplos. Los datos de tiempo para cada representación de un período pueden expresarse en relación con el tiempo de inicio del período.

[0037] Una representación puede incluir uno o más Segmentos. Cada representación puede incluir un Segmento de inicialización, o cada Segmento de una representación puede ser de autoinicialización. Cuando está presente, el Segmento de inicialización puede contener información de inicialización para el acceso a la representación. En general, el Segmento de inicialización no contiene datos de medios. Un Segmento puede ser referenciado en forma única por un identificador, tal como un localizador uniforme de recursos (URL), nombre uniforme de recursos (URN), o identificador uniforme de recursos (URI). la MPD puede proporcionar los identificadores para cada Segmento. En algunos ejemplos, la MPD también puede proporcionar intervalos de bytes en la forma de un atributo de *intervalo*, que puede corresponder a los datos para un Segmento dentro de un archivo accesible por URL, URN, o URI.

[0038] Diferentes representaciones se pueden seleccionar para la recuperación sustancialmente simultánea para diferentes tipos de datos de medios. Por ejemplo, un dispositivo cliente puede seleccionar una representación de audio, una representación de video, y una representación de texto temporizado para recuperar Segmentos. En algunos ejemplos, el dispositivo cliente puede seleccionar conjuntos de adaptación particulares para llevar a cabo la adaptación del ancho de banda. Es decir, el dispositivo cliente puede seleccionar un conjunto de adaptación que incluye representaciones de video, un conjunto de adaptación que incluye representaciones de audio, y/o un conjunto de adaptación que incluyen texto temporizado. Alternativamente, el dispositivo cliente puede seleccionar conjuntos de adaptación para ciertos tipos de medios (por ejemplo, video), y seleccionar directamente representaciones para otros tipos de medios (por ejemplo, audio y/o texto temporizado).

[0039] La FIG. 1 es un diagrama de bloques que ilustra un ejemplo de sistema 10 que implementa técnicas para la transmisión de datos de medios sobre una red. En este ejemplo, el sistema 10 incluye un dispositivo de preparación de contenidos 20, dispositivo servidor 60, y dispositivo cliente de dispositivo 40. El dispositivo cliente 40 y

el servidor 60 están acoplados comunicativamente por la red 74, que puede comprender Internet. En algunos ejemplos, el dispositivo de preparación de contenidos 20 y el dispositivo servidor 60 también se pueden acoplar por la red 74 u otra red, o pueden estar directamente acoplados comunicativamente. En algunos ejemplos, el dispositivo de preparación de contenidos 20 y el dispositivo servidor 60 pueden comprender el mismo dispositivo.

[0040] El dispositivo de preparación de contenidos 20, en el ejemplo de la FIG. 1, comprende fuente de audio 22 y fuente de video 24. La fuente de audio 22 puede comprender, por ejemplo, un micrófono que produce señales eléctricas representativas de los datos de audio capturados a codificar por el codificador de audio 26. Alternativamente, la fuente de audio 22 puede comprender un medio de almacenamiento que almacena datos de audio pregrabados un generador de datos de audio tal como un sintetizador automatizado, o cualquier otra fuente de datos de audio. La fuente de video 24 puede comprender una cámara de video que produce datos de video a codificar por el codificador de video 28, un medio de almacenamiento codificado con datos de video pregrabados, una unidad de generación de datos de video tal como una fuente de infografías, o cualquier otra fuente de datos de video. El dispositivo de preparación de contenidos 20 no está necesariamente acoplado de forma comunicativa al dispositivo servidor 60 en todos los ejemplos, pero puede almacenar contenido multimedia en un medio de separado que es leído por el dispositivo servidor 60.

[0041] Los datos de audio y video sin procesar pueden comprender datos analógicos o digitales. Los datos analógicos pueden ser digitalizados antes de codificarse por el codificador de audio 26 y/o el codificador de video 28. La fuente de audio 22 puede obtener datos de audio de un participante que habla mientras que el participante que habla está hablando, y la fuente de video 24 puede obtener simultáneamente datos de video del participante que habla. En otros ejemplos, la fuente de audio 22 puede comprender un medio de almacenamiento legible por computadora que comprende datos de audio almacenados, y la fuente de video 24 puede comprender un medio de almacenamiento legible por computadora que comprende datos de video almacenados. De esta manera, las técnicas descritas en la presente divulgación se pueden aplicar al streaming en tiempo real, de datos de audio y de video en vivo o a datos de audio y video archivados, pregrabados.

[0042] Las tramas de audio que corresponden a los cuadros de video son generalmente tramas de audio que contienen datos de audio que se capturaron (o

generaron) mediante la fuente de audio 22 simultáneamente con los datos de video capturados (o generados) mediante la fuente de video 24 que está contenida dentro de los cuadros de video. Por ejemplo, mientras un participante que habla generalmente produce datos de audio al hablar, la fuente de audio 22 capta los datos de audio, y la fuente de video 24 capta los datos de video del participante que habla al mismo tiempo, es decir, mientras la fuente de audio 22 captura los datos de audio. Por lo tanto, una trama de audio puede corresponder temporalmente a uno o más cuadros de video particulares. En consecuencia, una trama de audio correspondiente a un cuadro de video generalmente corresponde a una situación en la que datos de audio y datos de video se capturaron al mismo tiempo y para la que una trama de audio y un cuadro de video comprenden, respectivamente, los datos de audio y los datos de video capturados al mismo tiempo.

[0043] 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 que se grabaron los datos de audio para la trama de audio codificada, y de manera similar, el codificador de video 28 puede codificar una marca de tiempo en cada cuadro de video codificada que representa un tiempo donde se grabaron los datos de video para el cuadro de video codificado. En tales ejemplos, una trama de audio correspondiente a un cuadro de video puede comprender una trama de audio que comprende una marca de tiempo y un cuadro de video que comprende la misma marca de tiempo. El dispositivo de preparación de contenidos 20 puede incluir un reloj interno del que el codificador de audio 26 y/o codificador de video 28 pueden generar las marcas de tiempo, o que la fuente de audio 22 y la fuente de video 24 pueden utilizar para asociar los datos de audio y video, respectivamente, con una marca de tiempo.

[0044] En algunos ejemplos, la fuente de audio 22 puede enviar datos al codificador de audio 26 correspondientes a un tiempo donde se graban los datos de audio, y la fuente de video 24 puede enviar datos al codificador de video 28 correspondiente a un tiempo donde se grabaron los datos de video. En algunos ejemplos, el codificador de audio 26 puede codificar un identificador de secuencia en datos de audio codificados para indicar un ordenamiento temporal relativo de los datos de audio codificados, pero sin indicar necesariamente un tiempo absoluto en el cual se grabaron los datos de audio, y de manera similar, el codificador de video 28 también puede usar identificadores de secuencia para indicar un ordenamiento temporal relativo de los datos de video codificados. De manera similar, en algunos ejemplos, un identificador de secuencia se puede asignar o de otra manera correlacionar con una marca de

tiempo.

[0045] El codificador de audio 26 generalmente produce un flujo de datos de audio codificados, mientras que el codificador de video 28 produce un flujo de datos de video codificados. Cada flujo individual de datos (ya sea de audio o video) puede denominarse un flujo elemental. Un flujo elemental es un componente codificado digitalmente único, (posiblemente comprimido) de una representación. Por ejemplo, la parte codificada de video o audio de la representación puede ser un flujo elemental. Un flujo elemental puede convertirse en un flujo elemental empaquetado (PES) antes de ser encapsulado dentro de un archivo de video. Dentro de la misma representación, una ID de flujo puede utilizarse para distinguir los paquetes PES que pertenecen a un flujo elemental del otro. La unidad básica de datos de un flujo elemental es un paquete de flujo elemental en paquetes (PES). Por lo tanto, los datos de video codificados generalmente corresponden a flujos de video elementales. De manera similar, los datos de audio corresponden a una o más flujos elementales respectivos.

[0046] En el ejemplo de la FIG. 1, la unidad de encapsulación 30 del dispositivo de preparación de contenidos 20 recibe flujos elementales que comprenden datos de video codificados desde el codificador de video 28 y flujos elementales que comprenden datos de audio codificados de codificador de audio 26. En algunos ejemplos, el codificador de video 28 y el codificador de audio 26 pueden incluir cada uno empaquetadores para formar paquetes PES de datos codificados. En otros ejemplos, el codificador de video 28 y el codificador de audio 26 pueden tener una interfaz con empaquetadores respectivos para la formación de paquetes PES de datos codificados. En otros ejemplos más, la unidad de encapsulación 30 puede incluir empaquetadores para la formación de paquetes PES a partir de datos de audio y video codificados.

[0047] El codificador de video 28 puede codificar datos de video de contenido multimedia en una variedad de maneras, para producir diferentes representaciones del contenido multimedia en diferentes velocidades de bits y con diferentes características, tal como resolución de píxeles, frecuencia de cuadro, conformidad con varios estándares de codificación, conformidad con varios perfiles y/o niveles de perfiles para diferentes estándares de codificación, representaciones que tiene una o múltiples vistas (por ejemplo, para reproducción en dos dimensiones o tres dimensiones), u otras tales características. Una representación, tal como se utiliza en la presente divulgación, puede comprender uno de datos de audio, datos de video, datos de texto (por ejemplo, para subtítulos opcionales), u otros tales datos. La representación puede

incluir un flujo elemental, tal como un flujo elemental de audio o un flujo elemental de video. Cada paquete PES puede incluir un `stream_id` que identifica el flujo elemental al que pertenece el paquete PES. La unidad de encapsulación 30 es responsable del montaje de flujos elementales en archivos de video (por ejemplo, Segmentos) de diversas representaciones.

[0048] La unidad de encapsulación 30 recibe paquetes PES para los flujos elementales de una representación de codificador de audio 26 y codificador de video 28 y forma unidades de capa de abstracción de red correspondientes (NAL) de los paquetes PES. En el ejemplo de H.264/AVC (Codificación Avanzada de Video), los Segmentos de video codificados se organizan en unidades NAL, que proporcionan una representación de video “de red amigable” que da cabida a aplicaciones tales como telefonía de video, almacenamiento, difusión o streaming. Las unidades NAL se pueden categorizar en unidades NAL de Capa de Codificación de Video (VCL) y unidades NAL no VCL. Las unidades VCL pueden contener el motor de compresión del núcleo y pueden incluir bloques, macrobloques, y/o datos de nivel de porciones. Otras unidades NAL pueden ser unidades NAL no VCL. En algunos ejemplos, una imagen codificada en un instante de tiempo, normalmente presentada como una imagen codificada principal, puede estar contenida en una unidad de acceso, que puede incluir una o más unidades NAL.

[0049] Las unidades NAL no VCL pueden incluir conjuntos de parámetros de unidades NAL y unidades SEI NAL, entre otros. Los conjuntos de parámetros pueden contener información de cabecera de nivel de secuencia (en conjuntos de secuencias de parámetros (MSF)) e información de cabecera de nivel de imagen de cambio con poca frecuencia (en conjuntos de parámetros de imagen (PPS)). Con los conjuntos de parámetros (por ejemplo, PPS y SPS), la información de cambio con poca frecuencia no necesita ser repetirse para cada secuencia o imagen, por lo tanto, la eficiencia de codificación puede mejorarse. 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 capacidad de recuperación de error. En los ejemplos de transmisión fuera de banda, los conjuntos de parámetros de unidades NAL se pueden transmitir en un canal diferente de otras unidades NAL, tal como unidades SEI NAL.

[0050] La Información de Mejora Suplementaria (SEI) puede contener información que no es necesaria para la decodificación de las muestras de imágenes codificadas de unidades NAL VCL, pero puede asistir en los procesos relacionados con la

decodificación, visualización, capacidad de recuperación de error, y otros propósitos. Los mensajes SEI pueden estar contenidos en unidades no NAL VCL. Los mensajes SEI son la parte normativa de algunas especificaciones de estándar, y por lo tanto no siempre son obligatorios para la implementación del estándar compatible con el decodificador. Los mensajes SEI pueden estar a nivel de secuencia de mensajes SEI o nivel de imagen de mensajes SEI. Cierta información de nivel de secuencia puede estar contenida en los mensajes SEI, tal como mensajes SEI de información de escalabilidad en el ejemplo de SVC y mensajes SEI de información escalabilidad de vista en MVC. Estos mensajes SEI de ejemplo pueden llevar información sobre, por ejemplo, la extracción de puntos de operación y características de los puntos de operación. Además, la unidad de encapsulación 30 puede formar un archivo de manifiesto, tal como un descriptor de presentación de medios (MPD) que describe características de las representaciones. La unidad de encapsulación 30 puede formatear la MPD acuerdo con el lenguaje de marcado extensible (XML).

[0051] 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 escritura en un medio de almacenamiento, tal como una interfaz de bus de serie universal (USB), una grabadora o copiadora de CD o DVD, una interfaz para medios de almacenamiento magnéticos o de 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 a través de la red de transmisión o medio de almacenamiento. En el ejemplo de la FIG. 1, el dispositivo servidor 60 incluye un medio de almacenamiento 62 que almacena varios contenidos multimedia 64, que incluyen cada uno un respectivo archivo de manifiesto 66 y una o más representaciones 68A-68N (representaciones 68). En algunos ejemplos, la interfaz de salida 32 también puede enviar los datos directamente a la red 74.

[0052] En algunos ejemplos, las representaciones 68 pueden separarse en conjuntos de adaptación. Es decir, varios subconjuntos de representaciones 68 pueden incluir respectivos conjuntos de características comunes, tales como códec, perfil y nivel, resolución, número de vistas, formato de archivo para los Segmentos, información del tipo de texto que pueden identificar un idioma u otras características del texto a visualizar con los datos de representación y/o audio a decodificar y presentar, por

ejemplo, mediante altavoces, información de ángulo de cámara que puede describir una perspectiva del ángulo de la cámara o cámara de mundo real de un escena para representaciones en el conjunto de adaptación, información de calificación que describe el contenido adecuado para un público determinado, o similares.

[0053] El archivo de manifiesto 66 puede incluir datos indicativos de los subconjuntos de representaciones 68 correspondientes a determinados conjuntos de adaptación, así como características comunes de los conjuntos de adaptación. El archivo de manifiesto 66 también puede incluir datos representativos de las características individuales, tales como tasas de bits, para las representaciones individuales de los conjuntos de adaptación. De esta manera, un conjunto de adaptación puede proporcionar una adaptación de ancho de banda de red simplificada. Las representaciones en un conjunto de adaptación se pueden indicar mediante el uso de elementos secundarios de un elemento de conjunto de adaptación de archivo de manifiesto 66.

[0054] El dispositivo servidor 60 incluye una unidad de procesamiento de solicitud 70 e interfaz de red 72. En algunos ejemplos, el dispositivo servidor 60 puede incluir una pluralidad de interfaces de red. Además, cualquiera o todas las características del dispositivo servidor 60 se pueden implementar en otros dispositivos de una red de entrega de contenido, tal como enrutadores, puentes, dispositivos proxy, interruptores, 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 conforman sustancialmente a los del dispositivo servidor 60. En general, la interfaz de red 72 está configurada para enviar y recibir datos a través de la red 74.

[0055] La unidad de procesamiento de solicitud 70 está configurada para recibir solicitudes de red de los dispositivos cliente, tal como el dispositivo cliente 40, para los datos de medio de almacenamiento 62. Por ejemplo, la unidad de procesamiento de solicitud 70 puede implementar el protocolo de transferencia de hipertexto (HTTP) versión 1.1, como se describe en RFC 2616, "Hypertext Transfer Protocol – HTTP/1.1", por R. Fielding *et al.*, Network Working Group, IETF, junio de 1999. Es decir, la unidad de procesamiento de solicitud 70 puede estar configurada para recibir solicitudes de HTTP GET o GET parcial y proporcionar datos de contenido multimedia 64 en respuesta a las solicitudes. Las solicitudes pueden especificar un Segmento de una de las representaciones 68, por ejemplo, por el uso de URL o URI del Segmento. En algunos ejemplos, las solicitudes también pueden especificar uno o más intervalos de

bytes del Segmento, que comprenden, por lo tanto, solicitudes de GET parciales. La unidad de procesamiento de solicitud 70 además puede estar configurada para prestar servicio a las solicitudes HTTP HEAD para proporcionar datos de cabecera de un Segmento de una de las representaciones 68. En cualquier caso, la unidad de procesamiento de solicitud 70 puede estar configurada para procesar las solicitudes para proporcionar datos solicitados a un dispositivo de solicitud, tal como un dispositivo cliente 40.

[0056] Adicional o alternativamente, la unidad de procesamiento de solicitud 70 se puede configurar para entregar datos de medios a través de un protocolo de difusión o multidifusión, tal como eMBMS o DVB-T/T2. El dispositivo de preparación de contenidos 20 puede crear Segmentos DASH y/o subsegmentos sustancialmente de la misma manera descrita, pero el dispositivo servidor 60 puede entregar estos Segmentos o subsegmentos mediante el uso de eMBMS, DVB-T/T2 u otro protocolo de transporte de red de difusión o multidifusión. Por ejemplo, la unidad de procesamiento de solicitud 70 puede estar configurada para recibir una solicitud de unión de un grupo de multidifusión al dispositivo cliente 40. Es decir, el dispositivo servidor 60 pueden anunciar una dirección de protocolo de Internet (IP) asociada con un grupo de multidifusión a los dispositivos cliente, que incluyen dispositivo cliente 40, asociada con contenido de medios particular (por ejemplo, una emisión de un evento en directo). El dispositivo cliente 40, a su vez, puede presentar una solicitud para unirse al grupo de multidifusión. Esta solicitud puede ser propagada a lo largo de la red 74, por ejemplo, enrutadores que componen la red 74, de manera tal que los enrutadores dirijan el tráfico destinado a la dirección IP asociada con el grupo de multidifusión al que se suscriben los dispositivos cliente, tal como el dispositivo cliente 40.

[0057] Además, la interfaz de red 72 del dispositivo servidor 60 puede estar configurada para realizar las técnicas de la presente divulgación. Adicional o alternativamente, la unidad de procesamiento de solicitud 70 y la interfaz de red 72 pueden estar configuradas para realizar las técnicas de la presente divulgación. Por fines de explicación, se supone que la interfaz de red 72 incluye subcomponentes que no se muestran explícitamente en la FIG. 1, tales como los diversos ejemplos que se analizan en mayor detalle más abajo. Por ejemplo, la interfaz de red 72 puede incluir un emisor, un programador, un generador de tramas, y un amplificador excitador. Además, en este ejemplo, la unidad de encapsulación 30 representa un ejemplo de un segmentador (es decir, una unidad que forma Segmentos de datos de medios, donde

los Segmentos representan archivos individuales asociados con URL/URI respectivos). En particular, la unidad de encapsulación 30 puede encapsular MDE en Segmentos, donde cada segmento puede incluir uno o más MDE.

[0058] De acuerdo con las técnicas de la presente divulgación, la unidad de encapsulación 30 puede proporcionar un flujo de Segmentos de datos de medios al dispositivo servidor 60, que en última instancia dirige el flujo de Segmentos a la interfaz de red 72 para la transmisión al dispositivo cliente 40. El flujo de Segmentos puede corresponder a una representación, tal como una de las representaciones 68. Además, la unidad de encapsulación 30 puede proporcionar datos que representan requisitos de tiempo de transmisión para MDE incluidos en los Segmentos al dispositivo servidor 60, que, nuevamente, en última instancia dirige los requisitos de tiempo de transmisión a la interfaz de red 72. Por ejemplo, los requisitos de tiempo de transmisión pueden especificarse dentro del archivo de manifiesto 66. Los requisitos de tiempo de transmisión generalmente indican tiempos más tempranos y/o más posteriores en los que los MDE se entregan al dispositivo cliente 40. Por lo tanto, la interfaz de red 72 puede enviar los MDE (por ejemplo, Segmentos o intervalos de bytes de Segmentos) al dispositivo cliente 40 de conformidad con los requisitos de tiempo de transmisión. Por ejemplo, la interfaz de red 72 puede entregar los MDE de manera tal que los MDE no lleguen antes que los tiempos de transmisión más tempranos y/o no más tarde que los tiempos de transmisión más posteriores en el dispositivo cliente 40.

[0059] Como se ilustra en el ejemplo de la FIG. 1, el contenido multimedia 64 incluye un archivo de manifiesto 66, que puede corresponder a una descripción de presentación de medios (MPD). El archivo de manifiesto 66 puede contener descripciones de representaciones alternativas diferentes 68 (por ejemplo, servicios de video con diferentes cualidades) y la descripción puede incluir, por ejemplo, información códec, un valor de perfil, un valor de nivel, una tasa 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 los Segmentos de representaciones 68.

[0060] En particular, la unidad de recuperación 52 puede recuperar datos de configuración (no se muestra) del dispositivo cliente 40 para determinar las capacidades de decodificación del decodificador de video 48 y la representación de las capacidades de salida de video 44. Los datos de configuración pueden también incluir cualquiera o todos de una preferencia de idioma seleccionada 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 configurado para enviar solicitudes de HTTP GET y GET parciales. La unidad de recuperación 52 puede corresponder a instrucciones de software ejecutadas por uno o más procesadores o unidades de procesamiento (no se muestra) del dispositivo cliente 40. En algunos ejemplos, la totalidad o porciones de la funcionalidad descrita con respecto a la unidad de recuperación 52 pueden implementarse en hardware, o una combinación de hardware, software, y/o firmware, donde el hardware requerido puede proporcionarse para ejecutar instrucciones para software o firmware.

[0061] La unidad de recuperación 52 puede comparar las capacidades de decodificación y representación de dispositivo cliente 40 con las características de las representaciones 68 indicadas por la información del archivo de manifiesto 66. La unidad de recuperación de 52 inicialmente puede recuperar al menos una porción del archivo de manifiesto 66 para determinar las características de las representaciones 68. Por ejemplo, la unidad de recuperación 52 puede solicitar una porción del archivo de manifiesto 66 que describe 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 pueden ser satisfacerse por las capacidades de codificación y representación del dispositivo cliente 40. La unidad de recuperación 52 luego puede determinar tasas de bits para las representaciones en el conjunto de adaptación, determinar una cantidad actualmente disponible de ancho de banda de red, y recuperar Segmentos de una de las representaciones que tienen una tasa de bits que puede satisfacerse por el ancho de banda de red.

[0062] En general, las representaciones de tasa de bits mayores pueden proporcionar una reproducción de video de mayor calidad, mientras que las representaciones de tasa de bits menores pueden proporcionar reproducción de video de calidad suficiente cuando el ancho de banda disponible de la red disminuye. Por consiguiente, cuando el ancho de banda de red disponible es relativamente alto, la unidad de recuperación 52 puede recuperar datos de representaciones de tasa de bits relativamente alta, mientras que cuando el ancho de banda de red disponible es bajo, la unidad de recuperación 52 puede recuperar datos de representaciones de tasa de bits

relativamente baja. De esta manera, el dispositivo cliente 40 puede transmitir datos multimedia a través de la red 74 y al mismo tiempo adaptarse a cambios en la disponibilidad de ancho de banda de la red 74.

[0063] Adicional o alternativamente, la unidad de recuperación 52 puede estar configurada para recibir datos de acuerdo con un protocolo de difusión o red de multidifusión, tal como eMBMS o multidifusión IP. En tales ejemplos, la unidad de recuperación 52 puede presentar una solicitud para unirse a un grupo de red de multidifusión asociada con contenido de medios particular. Después de unirse al grupo de multidifusión, la unidad de recuperación 52 puede recibir datos del grupo de multidifusión sin solicitudes adicionales emitidas al dispositivo del servidor 60 o el dispositivo de preparación de contenidos 20. La unidad de recuperación 52 puede enviar una solicitud para abandonar el grupo de multidifusión cuando los datos del grupo de multidifusión ya no sean necesarios, por ejemplo, para detener la reproducción o para cambiar de canal a un grupo de multidifusión diferente.

[0064] La interfaz de red 54 puede recibir y proporcionar datos de segmentos de una representación seleccionada a la unidad de recuperación 52, que puede a su vez proporcionar los segmentos a una unidad de descapsulación 50. La unidad de descapsulación 50 puede desencapsular elementos de un archivo de video en flujos de PES constituyentes, desempaquetar los flujos de PES para recuperar los datos codificados, y enviar los datos codificados al decodificador de audio 46 o decodificador de video 48, en función de si los datos codificados son una porción de un flujo de audio o de video, por ejemplo, como se indica por 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 video 48 decodifica los datos de video codificados y envía los datos de video decodificados, que pueden incluir una pluralidad de vistas de un flujo, a la salida de video 44.

[0065] El codificador de video 28, el decodificador de video 48, el codificador de audio 26, el decodificador de audio 46, la unidad de encapsulación 30, la unidad de recuperación 52, y unidad de descapsulación 50 pueden implementarse como cualquiera de una variedad de circuitos de procesamiento adecuados, según corresponda, tal como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados de aplicación específica (ASIC), matrices de puertas programables (FPGA), circuitos de lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cada uno del codificador de video 28

y el decodificador de video 48 pueden incluirse en uno o más codificadores o decodificadores, cualquiera de los cuales puede estar integrado como parte de un codificador/decodificador de video combinado (CÓDEC). Del mismo modo, cada uno de codificador de audio 26 y decodificador de audio 46 pueden incluirse en uno o más codificadores o decodificadores, cualquiera de los cuales puede estar integrado como parte de un CÓDEC combinado. Un aparato que incluye el codificador de video 28, el decodificador de video 48, el codificador de audio 26, el decodificador de audio 46, la unidad de encapsulación 30, la unidad de recuperación 52, y/o unidad de descapsulación 50 puede comprender un circuito integrado, un microprocesador, y/o dispositivo de comunicación inalámbrica, tal como un teléfono celular.

[0066] El dispositivo cliente 40, el dispositivo servidor 60, y/o el dispositivo de preparación de contenidos 20 pueden estar configurados para operar de acuerdo con las técnicas de la presente divulgación. Para los propósitos de ejemplo, esta divulgación describe estas técnicas con respecto al dispositivo cliente 40 y el dispositivo servidor 60. Sin embargo, debe comprenderse que el dispositivo de preparación de contenidos 20 puede estar configurado para llevar a cabo estas técnicas, en lugar del (o además del) dispositivo servidor 60.

[0067] La unidad de encapsulación 30 puede formar unidades NAL que comprenden una cabecera que identifica a un programa al que pertenece la unidad NAL, así como una carga útil, por ejemplo, datos de audio, datos de video, o datos que describen el flujo de transporte o programa al que 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 incluye datos de video en su carga útil puede comprender varios niveles de granularidad de datos de video. Por ejemplo, una unidad NAL puede comprender un bloque de datos de video, una pluralidad de bloques, una porción de datos de video, o una imagen completa de datos de video. La unidad de encapsulación 30 puede recibir datos de video codificados del codificador de video 28 en forma de paquetes PES de flujos elementales. La unidad de encapsulación 30 puede asociar cada flujo elemental con un programa correspondiente.

[0068] La unidad de encapsulación 30 puede también ensamblar unidades de acceso de una pluralidad de unidades NAL. En general, una unidad de acceso puede comprender una o más unidades NAL para representar un cuadro de datos de video, así como datos de audio correspondientes a la trama cuando tales datos de audio están disponibles. Una unidad de acceso incluye en general todas las unidades NAL para un ejemplo de tiempo de salida, por ejemplo, todos los datos de audio y de video

para una instancia de tiempo. Por ejemplo, si cada vista tiene una frecuencia de cuadro de 20 cuadros por segundo (fps), entonces cada instante de tiempo puede corresponder a un intervalo de tiempo de 0,05 segundos. Durante este intervalo de tiempo, los cuadros específicos para todas las vistas de la misma unidad de acceso (el mismo instante de tiempo) se pueden representar simultáneamente. En un ejemplo, una unidad de acceso puede comprender una imagen codificada en un instante de tiempo, que puede presentarse como una imagen codificada principal.

[0069] En consecuencia, una unidad de acceso puede comprender todos los cuadros de audio y video de una instancia temporal común, por ejemplo, todas las vistas correspondientes para el tiempo X. Esta divulgación también se refiere a una imagen codificada de una vista particular tal como un “componente de vista”. Es decir, un componente de vista puede comprender una imagen codificada (o marco) para una vista particular en un momento particular. En consecuencia, una unidad de acceso se puede definir como comprendedora de todos los componentes de vista de una instancia temporal común. El orden de decodificación de unidades de acceso no tiene que ser necesariamente igual al orden de salida o visualización.

[0070] 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 video con diferentes cualidades) y la descripción pueden incluir, por ejemplo, información códec, un valor de perfil, y un valor de nivel. una MPD es un ejemplo de un archivo de manifiesto, tal como archivo de manifiesto 66. El dispositivo cliente 40 puede recuperar la MPD de una presentación multimedia para determinar cómo acceder a fragmentos de película de varias presentaciones. Los fragmentos de película pueden estar situados en cajas de fragmentos de película (cajas moof) de archivos de video.

[0071] El archivo de manifiesto 66 (que puede comprender, por ejemplo, una MPD) puede anunciar la disponibilidad de segmentos de representaciones 68. Es decir, la MPD puede incluir información que indica el tiempo de reloj de pared al que se torna disponible un primer segmento de una de las representaciones 68, así como información que indica las duraciones de 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, con base en la hora de inicio, así como las duraciones de los segmentos precedentes a un segmento particular.

[0072] Después de que la unidad de encapsulación 30 ha ensamblado las unidades NAL y/o unidades de acceso en un archivo de video con base en los datos recibidos,

la unidad de encapsulación 30 pasa el archivo de video a la interfaz de salida 32 para la salida. En algunos ejemplos, la unidad de encapsulación 30 puede almacenar el archivo de video de forma local o enviar el archivo de video a un servidor remoto a través de la interfaz de salida 32, en lugar de enviar el archivo de video directamente a un dispositivo cliente 40. La interfaz de salida 32 puede comprender, por ejemplo, un transmisor, un transceptor, un dispositivo para la escritura de datos a un medio legible por computadora tal como, por ejemplo, una unidad óptica, una unidad de medios magnético (por ejemplo, una unidad de disco flexible), un puerto de bus de serie universal (USB), una interfaz de red, u otra interfaz de salida. La interfaz de salida 32 da salida al archivo de video a un medio legible por computadora, tal como, por ejemplo, una señal de transmisión, un medio magnético, un medio óptico, una memoria, una unidad flash, u otro medio legible por computadora.

[0073] La interfaz de red 54 puede recibir una unidad NAL o unidad de acceso a través de la red 74 y proporcionar la unidad NAL o unidad de acceso a la unidad de descapsulación 50, a través de la unidad de recuperación 52. La unidad de descapsulación 50 puede desencapsular elementos de un archivo de video en flujos PES constituyentes, desempaquetar los flujos PES para recuperar los datos codificados, y enviar los datos codificados al decodificador de audio 46 o decodificador de video 48, en función de si los datos codificados son parte de un flujo de audio o de video, por ejemplo, como se indica por PES cabeceras de los paquetes 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 video 48 decodifica los datos de video codificados y envía los datos de video decodificados, que pueden incluir una pluralidad de vistas de un flujo, a la salida de video 44.

[0074] Determinados elementos del ejemplo de la FIG. 1 pueden estar configurados para realizar las técnicas de la presente divulgación. Por ejemplo, la interfaz de red 72 y/o unidad de procesamiento de solicitudes 70 puede estar configurada para funcionar de forma independiente o en conjunto para llevar a cabo las técnicas de la presente divulgación.

[0075] La FIG. 2 es un diagrama de bloques que ilustra un conjunto de ejemplo de componentes de la unidad de recuperación 52 de la FIG. 1 con mayor detalle. En este ejemplo, la unidad de recuperación 52 incluye una unidad de middleware eMBMS 100, un cliente DASH 110, y una aplicación de medios 112.

[0076] En este ejemplo, la unidad de middleware eMBMS 100 además incluye una unidad de recepción eMBMS 106, una memoria caché 104, y una unidad de servidor

102. En este ejemplo, la unidad de recepción eMBMS 106 está configurada para recibir datos a través de eMBMS, por ejemplo, de acuerdo con la Entrega de Archivos sobre Transporte Unidireccional (FLUTE), que se describe en T. Paila *et al.*, “FLUTE—File Delivery over Unidirectional Transport”, Network Working Group, RFC 6726, noviembre de 2012, disponible en <http://tools.ietf.org/html/rfc6726>. Es decir, la unidad de recepción eMBMS 106 puede recibir archivos a través de la emisión de, por ejemplo, un dispositivo servidor 60, que puede actuar como un BM-SC.

[0077] A medida que la unidad de middleware eMBMS 100 recibe los datos de los archivos, la unidad de middleware eMBMS puede almacenar los datos recibidos en la memoria caché 104. La memoria caché 104 puede comprender un medio de almacenamiento legible por computadora, tal como una memoria flash, un disco duro, RAM, o cualquier otro medio de almacenamiento adecuado.

[0078] La unidad de servidor local 102 puede actuar como un servidor para el cliente DASH 110. Por ejemplo, la unidad de servidor local 102 puede proporcionar un archivo MPD u otro archivo de manifiesto a cliente DASH 110. La unidad de servidor local 102 puede anunciar los tiempos de disponibilidad para los Segmentos en el archivo MPD, así como hipervínculos a partir de los cuales los se pueden recuperar Segmentos. Estos hipervínculos pueden incluir un prefijo de la dirección localhost correspondiente al dispositivo cliente 40 (por ejemplo, 127.0.0.1 para IPv4). De esta manera, el cliente DASH 110 puede solicitar Segmentos de unidad de servidor local 102 mediante el uso de solicitudes HTTP GET o GET parciales. Por ejemplo, para un Segmento disponible del enlace <http://127.0.0.1/rep1/seg3>, el cliente DASH 110 puede construir una solicitud HTTP GET que incluye una solicitud para <http://127.0.0.1/rep1/seg3>, y presentar la solicitud a la unidad de servidor local 102. La unidad de servidor local 102 puede recuperar datos requeridos de la memoria caché 104 y proporcionar los datos a cliente DASH 110 en respuesta a tales requisitos.

[0079] La FIG. 3 es un diagrama conceptual que ilustra elementos de contenido multimedia de ejemplo 120. El contenido multimedia 120 puede corresponder a un contenido multimedia 64 (FIG. 1), u otro contenido multimedia almacenado en un medio de almacenamiento 62. En el ejemplo de la FIG. 3, el contenido multimedia 120 incluye una descripción de presentación de medios (MPD) 122 y una pluralidad de representaciones 124A-124N (representaciones 124). La representación 124A incluye datos de cabecera opcionales 126 y Segmentos 128A-128N (segmentos 128), mientras que la representación 124N incluye datos de cabecera opcionales 130 y Segmentos 132A-132N (segmentos 132). La letra N se utiliza para designar el último

fragmento de película en cada una de las representaciones 124 como una cuestión de conveniencia. En algunos ejemplos, puede haber diferentes números de fragmentos de películas entre las representaciones 124.

[0080] MPD 122 puede comprender una estructura de datos separada de las representaciones 124. MPD 122 puede corresponder al archivo de manifiesto 66 de la FIG. 1. Asimismo, las representaciones 124 puede corresponder a representaciones 68 de la FIG. 2. En general, MPD 122 puede incluir datos que describen generalmente características de representaciones 124, tal como las características de codificación y renderización, conjuntos de adaptación, un perfil al que corresponde MPD 122, información de tipo de texto, información de ángulo de la cámara, información de calificación, información de modo de truco (por ejemplo, información indicativa de las representaciones que incluyen subsecuencias temporales), y/o información para la recuperación de períodos remotos (por ejemplo, para la inserción de anuncios dirigidos en contenido de medios durante la reproducción).

[0081] Los datos de cabecera 126, cuando están presentes, pueden describir características de Segmentos 128, por ejemplo, ubicaciones temporales de puntos de acceso aleatorio (RAP, también denominados puntos de acceso de flujo (SAP)), cuyos Segmentos 128 incluyen puntos de acceso aleatorio, desplazamientos de bytes aleatorios de puntos de acceso dentro de los Segmentos 128, localizadores uniformes de recursos (URL) o URI de los Segmentos 128, u otros aspectos de los segmentos 128. Los datos de cabecera 130, cuando están presentes, pueden describir características similares para los segmentos 132. Adicional o alternativamente, tales características pueden estar completamente incluidas dentro de MPD 122.

[0082] Los segmentos 128, 132 incluyen uno o más muestras de video codificado, cada una de las cuales puede incluir cuadros o porciones de datos de video. Cada una de las muestras de video codificado de segmentos 128 puede tener características similares, por ejemplo, requisitos de altura, ancho, y ancho de banda. Tales características pueden describirse por los datos de MPD 122, aunque estos datos no se ilustran en el ejemplo de la FIG. 3. MPD 122 puede incluir características de acuerdo con lo descrito en la Especificación 3GPP, con la adición de cualquiera o toda la información señalada que se describe en esta divulgación.

[0083] Cada uno de los Segmentos 128, 132 pueden estar asociados con un localizador uniforme de recursos (URL) único o URI. Por lo tanto, cada uno de los Segmentos 128, 132 puede ser independientemente recuperable mediante el uso de un protocolo de red de streaming, tal como DASH. De esta manera, un dispositivo

objetivo, tal como el dispositivo cliente 40, puede utilizar una solicitud HTTP GET para recuperar Segmentos 128 o 132. En algunos ejemplos, el dispositivo cliente 40 puede utilizar solicitudes HTTP GET parciales para recuperar intervalos de bytes específicos de Segmentos 128 o 132.

[0084] La FIG. 4 es un diagrama de bloque de elementos de un ejemplo de archivo de video 150, que puede corresponder a un Segmento de una Representación, tal como uno de los segmentos 114, 124 de la FIG. 3. Si bien el ejemplo de la FIG. 4 ilustra un archivo de video, se debe comprender que un archivo de audio u otro archivo pueden incluir datos similares a los de archivo de video 150. Cada uno de los segmentos 128, 132 puede incluir datos que se ajustan sustancialmente a la disposición de los datos ilustrados en el ejemplo de la FIG. 4. Puede decirse que el archivo de video 150 encapsula un Segmento. De acuerdo con lo descrito con anterioridad, los archivos de video de acuerdo con el formato de archivo multimedia de base ISO y sus extensiones almacenan datos en una serie de objetos, denominados “cajas”. En el ejemplo de la FIG. 4, el archivo de video 150 incluye la caja de tipo de archivo (FTYP) 152, la caja de película (MOOV) 154, la caja de índice de segmento (sidx) 162, la caja de fragmento de película (MOOF) 164, y la caja de fragmento de película acceso aleatorio (MFRA) 166. Si bien la FIG. 4 representa un ejemplo de un archivo de video, debe comprenderse que otros archivos de medios pueden incluir otros tipos de datos de medios (por ej., datos de audio, datos de texto temporizado, o similares) estructurados de manera similar a los datos de archivo de video 150, de acuerdo con el formato de archivo multimedia de base ISO y sus extensiones.

[0085] La caja de tipo de archivo (FTYP) 152 describe en general un tipo de archivo para archivo de video 150. La caja de tipo de archivo 152 puede incluir datos que identifican una especificación que describe un mejor uso para tipo de video 150. La caja de tipo de archivo 152 alternativamente puede estar colocada antes de la caja MOOV 154, las cajas de fragmentos de película de 164, y/o una caja MFRA 166.

[0086] En algunos ejemplos, un Segmento, tal como archivo de video 150, puede incluir una caja de actualización MPD (no se muestra) antes de la caja FTYP 152. La caja de actualización MPD puede incluir información que indique que una MPD correspondiente a una representación que incluye archivo de video 150 está siendo actualizada, junto con la información para la actualización de la MPD. Por ejemplo, la caja de actualización MPD puede proporcionar URI o URL de un recurso que se utiliza para actualizar la MPD. Como otro ejemplo, la caja de actualización MPD puede incluir datos para la actualización de la MPD. En algunos ejemplos, la caja de actualización

de la MPD puede seguir inmediatamente a una caja de tipo de segmento (STYP) (no se muestra) del archivo de video 150, donde la caja STYP puede definir un tipo de segmento para el archivo de video 150. La FIG. 7, que se analiza en mayor detalle a continuación, proporciona información adicional con respecto a la caja de actualización de MPD.

[0087] La caja MOOV 154, en el ejemplo de la FIG. 4, incluye la caja de cabecera de película (MVHD) 156, la caja de pista (TRAK) 158, y una o más cajas de extensión de película (MVEX) 160. En general, la caja MVHD 156 puede describir las características generales de archivo de video 150. Por ejemplo, la caja MVHD 156 puede incluir datos que describen cuándo se creó originalmente un archivo de video 150, cuándo se modificó por última vez un archivo de video 150, una escala de tiempo para el archivo de video 150, una duración de la reproducción de archivos de video 150, u otros datos que generalmente describen archivos de video 150.

[0088] La caja TRAK 158 puede incluir datos para una pista de archivo de video 150. La caja TRAK 158 puede incluir una caja de cabecera de pista (TKHD) que describe las características de la pista correspondiente a la caja TRAK 158. En algunos ejemplos, la caja TRAK 158 puede incluir imágenes de video codificado, mientras que, en otros ejemplos, las imágenes de video codificado de la pista pueden incluirse en fragmentos de película 164, que pueden ser referenciados por los datos de imagen de la caja TRAK 158 y/o las cajas sidx 162.

[0089] En algunos ejemplos, el archivo de video 150 puede incluir más de una pista. De acuerdo con ello, la caja MOOV 154 puede incluir un número de cajas TRAK igual al número de pistas en archivo de video 150. La caja TRAK 158 puede describir las características de una pista correspondiente del archivo de video 150. Por ejemplo, la caja TRAK 158 puede describir información temporal y/o espacial para la pista correspondiente. Una caja TRAK similar a la caja TRAK 158 de la caja MOOV 154 puede describir las características de una pista de conjunto de parámetros, cuando la unidad de encapsulación 30 (FIG. 3) incluye una pista de conjunto de parámetros en un archivo de video, tal como un archivo de video 150. La unidad de encapsulación 30 puede señalar la presencia de nivel de secuencia de mensajes SEI en la pista de conjunto de parámetros dentro de la caja TRAK que describe la pista de conjunto de parámetros.

[0090] Las cajas MVEX 160 pueden describir las características de los fragmentos de película correspondientes 164, por ejemplo, para indicar que el archivo de video 150 incluye fragmentos de película 164, además de datos de video incluidos dentro de la

caja MOOV 154, en caso de haberlos. En el contexto de streaming de datos de video, pueden incluirse imágenes codificadas de video en los fragmentos de película 164 en lugar de en la caja MOOV 154. En consecuencia, todas las muestras de video codificado se pueden incluir en los fragmentos de película de 164, en lugar de en la caja MOOV 154.

[0091] La caja MOOV 154 puede incluir un número de cajas MVEX 160 igual al número de fragmentos de película 164 en archivo de video 150. Cada una de las cajas MVEX 160 puede describir las características de un fragmento de película 164 correspondiente. Por ejemplo, cada imagen de MVEX puede incluir una caja de cabecera de extensión de película (MEHD) que describe una duración temporal del fragmento de película correspondiente 164.

[0092] Como se señaló con anterioridad, la unidad de encapsulación 30 puede almacenar una secuencia de datos establecidos en una muestra de video que no incluye datos de video codificados reales. Una muestra de video puede corresponder generalmente a una unidad de acceso, que es una representación de una imagen codificada en un instante de tiempo específico. En el contexto de AVC, la imagen codificada puede incluir una o más unidades VCL NAL que contienen la información para construir todos los píxeles de la unidad de acceso y otras unidades no VCL NAL asociadas, tal como mensajes SEI. En consecuencia, la unidad de encapsulación 30 puede incluir un conjunto de datos de secuencia, que puede incluir mensajes SEI de nivel de secuencia, en uno de los fragmentos de película 164. La unidad de encapsulación 30 además puede señalar la presencia de mensajes SEI de datos de secuencia y/o nivel de secuencia presentes en uno de los fragmentos de película 164 dentro de una de las cajas MVEX 160 correspondiente a uno de los fragmentos de película 164.

[0093] Las cajas SIDX 162 son elementos opcionales del archivo de video 150. Es decir, los archivos de video que cumplen con el formato de archivo 3GPP, u otros formatos de archivo, no necesariamente incluyen cajas SIDX 162. De acuerdo con el ejemplo del formato de archivo 3GPP, una caja SIDX puede utilizarse para identificar un subsegmento de un Segmento (por ejemplo, un segmento de contenido dentro de un archivo de video 150). El formato de archivo 3GPP define un subsegmento como “un conjunto autónomo de una o más cajas consecutivas de fragmento de película con la caja correspondiente de Datos de Medios y una Caja de Datos de Medios que contiene datos referenciados por una Caja de Fragmento de Película deben seguir tal caja de Fragmento de Película y preceder a la siguiente caja de Fragmento de Película

que contiene información sobre la misma pista”. El formato de archivo 3GPP también indica que una caja SIDX “contiene una secuencia de referencias a subsegmentos del (sub)segmento documentado por la caja. Los subsegmentos referenciados son contiguos en el tiempo de presentación. De manera similar, los bytes referenciados por una caja de Índice de Segmento son siempre contiguos dentro del segmento. El tamaño referenciado da la cuenta del número de bytes en el material referenciado”.

[0094] Las cajas SIDX 162 generalmente proporcionan información representativa de una o más subsegmentos de un segmento incluido en el archivo de video 150. Por ejemplo, tal información puede incluir tiempos de reproducción a los que los subsegmentos comienzan y/o terminan, desplazamientos 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 refresco de decodificador instantáneo (IDR), una imagen de acceso aleatorio limpio (CRA) , una imagen de acceso a enlace roto (BLA), o similares), una posición del SAP (en términos de tiempo de reproducción y/o desplazamiento de bytes) en el subsegmento, y similares.

[0095] Los fragmentos de película 164 pueden incluir una o más imágenes de video codificado. 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 video codificado, por ejemplo, cuadros o imágenes. Además, como se describió con anterioridad, los fragmentos de película 164 pueden incluir conjuntos de datos de secuencias en algunos ejemplos. Cada uno de los fragmentos de película 164 puede incluir una caja de fragmento de película de cabecera (MFHD, que no se muestra en la FIG. 4). La caja MFHD puede describir características del fragmento de película correspondiente, tal como un número de secuencia para el fragmento de película. Los fragmentos de película 164 se pueden incluir en el orden de número de secuencia en el archivo de video 150.

[0096] La caja MFRA 166 puede describir puntos de acceso aleatorios dentro de fragmentos de película 164 de archivo de video 150. Esto puede ayudar con la realización de modos de truco, tal como la realización pretendida en ubicaciones temporales particulares (es decir, tiempos de reproducción) dentro de un segmento de encapsulado por archivo de video 150. La caja MFRA 166 es generalmente opcional y no debe figurar en los archivos de video, en algunos ejemplos. Del mismo modo, un dispositivo cliente, tal como el dispositivo cliente 40, no necesariamente tiene que hacer referencia a la caja MFRA 166 para decodificar correctamente y exhibir datos de

archivo de video 150. La caja MFRA 166 puede incluir un número de cajas de acceso aleatorio a fragmento de pista (TFRA) (no se muestra) igual al número de pistas de archivo de video 150, o en algunos ejemplos, igual al número de pistas de medios (por ejemplo, pistas no indicadas) de archivo de video 150.

[0097] En algunos ejemplos, los fragmentos de película 164 pueden incluir uno o más puntos de acceso de flujo (SAP), tal como imágenes IDR. Del mismo modo, la caja MFRA 166 puede proporcionar indicaciones de lugares dentro de archivo de video 150 de los SAP. Por consiguiente, una subsecuencia temporal de archivo de video 150 puede estar formada de SAP de archivo de video 150. La subsecuencia temporal también puede incluir otras imágenes, tal como cuadros P y/o cuadros B que dependen de SAP. Los cuadros y/o porciones de la subsecuencia temporal pueden estar dispuestos dentro de los segmentos de manera tal que los cuadros/porciones de la subsecuencia temporal que dependen de otras imágenes/porciones de la subsecuencia puede ser decodificados correctamente. Por ejemplo, en la disposición jerárquica de los datos, los datos utilizados para la predicción para otros datos pueden también incluirse en la subsecuencia temporal.

[0098] La FIG. 5 es un diagrama de bloques que ilustra un sistema de infraestructura de difusión DASH de ejemplo 180. El sistema 180 incluye una unidad de gestión del sistema 198, una memoria intermedia de medios de duración GOP 182, un codificador de medios 184, un segmentador 186, un emisor 188, un programador 190, un generador de tramas 192, y un excitador/amplificador 194. El sistema 180, y por lo tanto cada uno de estos componentes del sistema 180, pueden incluirse en un dispositivo fuente, tal como el dispositivo servidor 60 de la FIG. 1 o el dispositivo de preparación de contenidos 20 de la FIG. 1. Nuevamente, un dispositivo servidor y un dispositivo de preparación de contenidos pueden estar integrados funcionalmente en un solo dispositivo.

[0099] En el ejemplo de sistema 180, la memoria intermedia de medios de duración GOP 182 recibe y almacena temporalmente datos de medios a codificar. El codificador de medios 184 (que de hecho puede representar una pluralidad de codificadores de medios, tal como codificador de audio 26 de la FIG. 1 y el codificador de video 28 de la FIG. 1) codifica datos de medios (por ejemplo, datos de audio, datos de video, y/o datos de texto temporizado), encapsula los datos de medios en las unidades de capa de abstracción de red (NAL), y proporciona los datos de medios codificados, en forma de unidad de NAL 202, al segmentador 186. Además, el codificador de medios 184 envía datos que representan los identificadores de MDE explícitos 200 al segmentador

186.

[00100] El segmentador 186 encapsula las unidades NAL recibidas en Segmentos respectivos. Así, en general, el segmentador 186 no proporciona información explícita indicativa de las ubicaciones de los MDE dentro de los Segmentos. Efectivamente, sin embargo, el segmentador 186 proporciona los MDE al emisor 188, en forma de intervalos de byte dentro de los Segmentos (206). Además, el segmentador 186 también determina los requisitos de tiempo de transmisión 204 para los MDE y envía datos representativos de los requisitos de tiempo de transmisión de 204 al emisor 188 emisor.

[00101] El emisor 188 representa un ejemplo de una unidad de envío de protocolo basado en archivos de acuerdo con las técnicas de la presente divulgación. En este ejemplo, el emisor 188 envía los datos de los Segmentos recibidos de acuerdo con ROUTE. En otros ejemplos, el emisor 188 puede enviar datos de los Segmentos recibidos de acuerdo con FLUTE u otro tipo de protocolos basados en archivos. De acuerdo con las técnicas de la presente divulgación, el emisor 188 envía los datos de MDE (encapsulados dentro de unidades de datos con base en la red, por ejemplo, IP, UDP, ROUTE, ALP, o similares) 210 y los requisitos de tiempo de transmisión 208 al programador 190 y generador de tramas 192. El generador de tramas 192 forma tramas de red que encapsulan los datos de red al excitador/amplificador 194 de acuerdo con datos de programación determinados por el programador 190, donde las tramas de red incluyen además una descripción de banda base de los eventos de entrega de datos de capa física (DDE) 212. El excitador/amplificador 194 transmite activamente los datos, por ejemplo, a través de la capa de red física, tal como a través de una emisión de ATSC OTA, Ethernet, o similares.

[00102] Un sistema de emisión, tal como un sistema de Comité de Sistemas de Televisión Avanzada (ATSC), puede soportar el uso de los intervalos de bytes conscientes de medios implementados con MDE dentro de ROUTE u otro protocolo de entrega de objetos, que incluyen los intervalos de bytes conscientes de medios. La identificación (definición) de los MDE individuales para el sistema puede requerir comportamiento, por lo que el sistema puede entregar los MDE. Es decir, el sistema debe ser capaz de determinar qué intervalos de bytes conscientes de medios representan un MDE (u objeto de datos similar) con el fin de empaquetar y enviar el MDE al Programador con el etiquetado de tiempo apropiado.

[00103] Un ejemplo de definición de un MDE es un intervalo de bytes conscientes de medios que sea significativo para un códec de medios de recepción. Debido a la

estructura de trama de un códec, un MDE puede o no puede tener múltiples muestras en el uso de ISO BMMF del término muestra. Una trama de audio en Codificación de Audio Avanzada de Alta Eficiencia (HE-AAC) (ISO/IEC JTC1/SC29/WG11/N7016 (11 de enero de 2005)), que contiene múltiples muestras de audio, es por sí misma una muestra individual en un contexto BMFF ISO. Una muestra ISO BMFF en un contexto HEVC puede ser un cuadro de datos de video único. Un MDE puede incluir una o varias muestras ISO BMFF. Esta posibilidad existe debido a la interdependencia entre los cuadros de video sucesivos. Hay un MDE más pequeño posible, pero por conveniencia o para otros fines, estos MDE atómicos, es decir, MDE más pequeños posibles, se pueden agregar en un MDE más grande. En términos generales, los MDE no se superponen.

[00104] Los elementos del sistema 180 de la FIG. 5 que pueden ser en primer lugar conscientes de MDE son el codificador de medios 184, el segmentador 186, y el emisor 188. Las entidades funcionales en la FIG. 5 son conceptuales. Esta construcción lógica de la cascada es por conveniencia de consideración del sistema y no es necesaria. Las funciones podrían integrarse funcionalmente en una sola entidad, o posiblemente subdividirse entre unidades adicionales o alternativas de una manera diferente. Una ubicación lógica de la formación de MDE está dentro del alcance del codificador de medios 184, el segmentador 186, o sección de entrada del emisor 188.

[00105] El alcance y la definición de un MDE se define de acuerdo con el tipo de medio manipulado. Por ejemplo, el audio se puede empaquetar en Segmentos separados o multiplexar dentro de los Segmentos, es decir, archivos/objetos ISO BMFF. MPEG DASH (INTERNATIONAL STANDARD ISO/IEC 23009-1 Segunda edición 2014-05-01 Information Technology—Dynamic Adaptive Streaming Over HTTP (DASH) Part 1: Media Presentation Description and Segment Formats), por ejemplo, define Segmentos multiplexados y no multiplexados. Esto es simplemente un aspecto de configuración. Como se muestra con anterioridad en la FIG. 5, puede haber una función de configuración del sistema que suministra una configuración para las diferentes funciones.

[00106] Los ejemplos de algunos aspectos potencialmente configurados incluyen:

- Para DASH
 - Duración del Segmento
 - Composición del Segmento, es decir, tipos de medios contenidos en un objeto de medios específico
 - Cadencia de tipo del Segmento, por ejemplo, un Segmento de medios

específico contiene un Punto de Acceso de Flujo (SAP) o no

- Para un codificador de video
 - Duración de GOP (la nomenclatura específica de códec varía para tales estructuras)
 - Tipos de GOP, por ejemplo, abiertos o cerrados
 - Cadencia de tipo de Trama, por ejemplo, I, B, B, P, B, B, P...
 - Frecuencia de cuadro, resolución, progresiva vs. entrelazada, etc.
 - Límites de MDE a través de coincidencia de patrón o estructuras de medios permitidos, por ejemplo
 - la trama de audio es una muestra atómica de MDE e ISO BMFF
 - la trama individual I (IDR) es un MDE atómico
 - los grupos de cuadros de video son MDE atómicos como se representa en la FIG. 3
 - Tal determinación puede realizarse o conocerse en el Codificador o determinarse por medio del Segmentador
- Para un codificador de audio
 - Tasa de muestra
 - Selección de herramientas de codificación
 - Detalles de construcción de trama de audio de acuerdo con códec especificado

[00107] En la medida en que estos aspectos son configuraciones estáticas, las definiciones de MDE asociadas pueden proporcionarse como definiciones estáticas, es decir, la configuración del codificador es conocida para el Segmentador y/o Emisor, dado que es un conjunto configurado estático de ajustes. Suponiendo que el video ejecuta una única cadencia definida de tipos de cuadros de video, los límites de MDE pueden definirse por medio de la configuración. Véase la FIG. 7 a continuación para un ejemplo de una cadencia de video específica y sus MDE atómicos. Cabe señalar que esta figura representa el orden de presentación de los cuadros de video.

[00108] La FIG. 6 es un diagrama conceptual que ilustra una composición de alto nivel de MDE de ejemplo.

[00109] La FIG. 7 es un diagrama conceptual que ilustra una cadencia de tipo de trama de ejemplo y MDE definidos correspondientes. En particular, la FIG. 7 ilustra los cuadros de video 250-296. Los diversos cuadros de video están etiquetados, ya sea como I, P, o B, para representar si los marcos son intra-predichos (I), inter-predichos unidireccionalmente (P), o inter-predichos bidireccionalmente (B).

[00110] Los cuadros de video representados en la FIG. 7 se ilustran en orden de visualización. Si la figura se representara en orden entrega aparecerían los agrupamientos de trama a medida que un byte contiguo oscila en el flujo fuente a partir del codificador de video y dentro del contenedor ISO BMFF, es decir, un Segmento de Medios para DASH.

[00111] La identificación de los límites de MDE para una cadencia tal definida se puede lograr a través de tablas, que, por ejemplo, definen los primeros cuadros de video N de GOP como MDE 1, próximos cuadros M como MDE 2, etc., o cadencia específica de tipos de cuadro, por ejemplo, coincidencia de patrón, que se reconoce en la sintaxis de códec. El uso del método de la tabla es simple y conveniente, si la cadencia de los tipos de trama es estática en función de cada Segmento. Este es potencialmente un método conveniente para el tipo de medios de audio para los que puede haber un número fijo de tramas de códec de audio por MDE y por códec. (Este es también un ejemplo de una regla trivial.) Estas tablas que pueden contener una descripción explícita de un patrón o reglas pueden expresarse como XML. La FIG. 7 muestra que algunos MDE atómicos pueden agregarse para formar un MDE más grande. Esto puede determinarse a través de regla o coincidencia de patrón o combinación de los mismos. Estos métodos pueden implementarse dentro del codificador de medios y pasarse adelante como una descripción explícita al Segmentador y Emisor, por ejemplo, ROUTE o MMT.

[00112] De acuerdo con lo mencionado brevemente con anterioridad la tabla de configuración puede estar formada de acuerdo con una regla. La lógica de una regla de este tipo puede permitir la operación de la función de identificación de MDE (medios de intervalos de bytes conscientes) como adaptativa. Esto es conveniente, por ejemplo, si se multiplexa un Segmento (Objeto ISO BMFF), es decir, contiene varios tipos de medios. La función de identificación MDE del Segmentador puede analizar, por ejemplo, subtítulos opcionales, audio y video. Además, la cadencia para los tipos de cuadros de video de ejemplo puede ser adaptativa.

[00113] Los ejemplos de aspectos adaptativos de video desde una perspectiva MDE incluyen:

- Cadencia de tipo de cuadros de video para habilitar GOP abiertos y cerrados, en una secuencia de Segmentos
- Recuento y cadencia de tipo de cuadros de video en una base por Segmento o Período
 - Por ejemplo, 24 fps comerciales y contenido de 30 fps intercalado

- Tal cambio en la frecuencia de cuadro de video puede causar un cambio en la cadencia de tipo de cuadro de video y, como consecuencia, la cadencia de MDE.

[00114] Se describen técnicas para el soporte de identificación de MDE dentro de los Segmentos, es decir, los medios de intervalos de bytes conscientes que son contextualmente relevante para un códec de medios asociados en soporte de protocolo de intervalos de bytes conscientes, por ejemplo, ROUTE o MMT.

[00115] Los ejemplos de técnicas para soportar dicha definición de MDE, es decir, medios de intervalos de bytes conscientes de códec, pueden incluir:

- Métodos tabulares que describen secuencias específicas, tales como:
 - Tipo de cuadro de video
 - Secuencias de tipos de cuadro de video
 - Número de, por ejemplo, cuadros de video o tramas de audio
 - Combinaciones de estructuras descritas
 - Coincidencia de patrón de secuencia específica de tipos de cuadros de video de ejemplo
 - La implementación de estos métodos en un codificador de medios, Segmentador o Emisor
- Detección con base en MDE con base en regla
 - Por ejemplo, una configuración específica de un códec HEVC puede permitir estas construcciones lógicas de secuencias de tipo de trama:
 - El inicio de la secuencia de tramas (MDE) cumple este primer conjunto de condiciones.
 - El fin de la secuencia de tramas (MDE) cumple con un segundo conjunto de condiciones.
 - La secuencia de tipos de trama se encuentra con un tercer conjunto de condiciones.
- Los métodos de patrón y regla descritos se pueden combinar, por ejemplo:
 - Este patrón se descubre
 - Entonces, debe aplicarse esta regla
- Un codificador de medios puede tener una biblioteca de modos operativos definidos con construcciones específicas de MDE.
 - El codificador de medios comunica dicho modo operativo en tiempo real con los medios codificados pasados al Segmentador
- Véase el Apéndice A, por ejemplo, de XML o lógica con base en regla de

archivo y la lógica con base en patrón.

[00116] Cada MDE se puede definir con un tiempo de transmisión más temprano y/o más posterior. Esto es para el beneficio del Programador, cuya tarea es asignar los medios a las ranuras de entrega, por ejemplo, Tramas de FEC específicos en ATSC 3.0 en la capa física. Estos atributos más tempranos y más posteriores pueden limitarse por diferentes aspectos, tales como:

- Tiempo de transmisión más temprano: Este es literalmente el instante en que el primer bit de un MDE dado (o Segmento) puede irradiarse. Este bit radiado más temprano incluye el protocolo de encapsulación relacionado.
- Tiempo de transmisión más posterior: Este es literalmente el último instante en que el último bit de un MDE (o Segmento) tiene que irradiarse. Este último bit incluye todo protocolo de encapsulación relacionado.

[00117] El tiempo de transmisión más temprano puede verse limitado por la configuración del sistema. Por ejemplo, la cadencia de trama de capa física podría tener marcos iniciados en la totalidad de los segundos límites de tiempo de capa física ATSC (48b TAI segundos, 10b mseg o, posiblemente, otra fracción de segundo). Esta cadencia de trama de capa física puede o no puede estar directamente relacionada con la duración de Segmento de Medios. Un período dado de un programa, por ejemplo, podría tener que entregarse en su totalidad antes del final de la totalidad del segundo anuncio que precede a un anuncio y que es un reemplazo del mismo. El anuncio de reemplazo de salida, por ejemplo, no inicia la entrega hasta, al menos, el comienzo de la totalidad del próximo segundo anuncio. Para tiempos fuera de los límites de tiempo, el tiempo de transmisión más temprano podría estar en, ejemplo, la totalidad de un segundo conjunto.

[00118] El tiempo de transmisión más posterior puede estar relacionado con el tiempo de decodificación de los últimos medios decodificados contenidos dentro de un MDE etiquetado. El etiquetado aplicado a la entrega de nivel de Segmento de Medios puede determinarse por medio de la línea de tiempo de disponibilidad nominal de DASH/Segmento. El tiempo de transmisión más posterior puede ser consistente con el cumplimiento del plazo de entrega para el Segmento. El requisito de tiempo de transmisión más posterior para el último MDE (intervalo de bytes conscientes de medios) de un Segmento puede ser igual que el del Segmento completado.

[00119] La FIG. 8 es un diagrama conceptual que representa una vista simple del tiempo de sistema. Es decir, la FIG. 8 muestra una vista de nivel superior del tiempo de sistema. Este es el caso, o bien para la reproducción de nivel de Segmento o de la

reproducción basados en intervalos de bytes. Una diferencia para la reproducción de MDE es que el retardo del receptor estático puede ser notablemente más corto que para la reproducción de Segmentos. El tiempo del Reloj de Pared de la Estación puede entregarse al receptor a través de la capa física y el protocolo de soporte. El Reloj de Pared del receptor en cualquier instante, después de la sincronización, es el tiempo del Reloj de Pared de la Estación actual menos el tiempo de vuelo de señal vigente en ese momento del emisor dominante para el receptor.

[00120] La FIG. 9 es un diagrama de bloques que representa el sistema 300, que representa una infraestructura de transmisión que está configurada con una relación de tiempo. Es decir, la FIG. 9 proporciona una representación conceptual de cómo los retardos de tiempo pueden afectar el proceso de transmisión del sistema 300. En este ejemplo, el sistema 300 incluye la memoria intermedia de medios de duración GOP 306, el codificador de medios 308, el segmentador 310, el control de tasa 304, la unidad de gestión de sistema 302, el emisor 312, el programador 314 y el excitador/amplificador 316.

[00121] En este ejemplo, la memoria intermedia de medios de duración GOP 306 almacena (es decir, almacena temporalmente) datos de medios brutos a codificar. El codificador de medios 308 recupera los datos de medios a codificar de la memoria intermedia de medios de duración GOP 306, codifica los datos de medios, encapsula los datos de medios codificados en forma de unidad NAL, y suministra las unidades NAL 320 al segmentador 310. Debe comprenderse que los datos de medios codificados también incluyen MDE, por ejemplo, como se muestra en la FIG. 7. Además, como se explicó con anterioridad, cada uno de los MDE pueden tener requisitos de tiempo de transmisión, tal como los requisitos de tiempo de transmisión más tempranos y/o más posteriores.

[00122] El Segmentador 310 recibe las unidades NAL 320 del codificador de medios 308 y encapsula uno o más MDE (es decir, una o más de las unidades NAL) en Segmentos respectivos. Nuevamente, los Segmentos pueden corresponder a archivos independientemente entregables asociados con direcciones URL respectivas, donde las direcciones URL pueden identificar en forma única el Segmento correspondiente. El MDE puede corresponder a los respectivos intervalos de bytes de los Segmentos. Además, el Segmentador 310 puede determinar requisitos de tiempo de transmisión para cada uno de los MDE, y pasar los requisitos de tiempo de transmisión 322 y MDE (en forma de intervalos de bytes de los Segmentos) 324 al emisor 312.

[00123] El emisor 312 puede operar de acuerdo con un protocolo de entrega basado

en archivos, tal como ROUTE. Por lo tanto, el emisor 312 puede encapsular porciones respectivas de los Segmentos (por ejemplo, los MDE) en unidades de entrega respectivas con base en la red 328, tales como paquetes o tramas, al programador 314 para la transmisión. Además, el emisor 312 proporciona los requisitos de tiempo de transmisión 326 al programador 314, de manera tal que el programador 314 pueda entregar las unidades de entrega con base en red de un dispositivo cliente (no se muestra en la FIG. 9) de acuerdo con los requisitos de tiempo de transmisión. Más en particular, el programador 314 forma una descripción de banda base de los eventos de entrega de datos de capa física (DDE) 330 y pasa la descripción de banda base de DDE de capa física 330 al excitador/amplificador 316. Luego, el excitador/amplificador 316 transmite datos a la capa física al, por ejemplo, dispositivo cliente.

[00124] La FIG. 9 también representa varios retardos introducidos por el sistema 300 al enviar datos de medios a un dispositivo cliente. Por ejemplo, el tiempo entre los datos de medios que se almacenan en memoria intermedia a los datos de medios programados para la transmisión se denomina retardo de duración de análisis 332. El retardo de duración de análisis incluye particularmente retardo de solo simple para una muestra específica 334, que es el tiempo entre el codificador de medios 308 que codifica la muestra y la encapsulación de la muestra en una unidad NAL y proporciona la unidad NAL al segmentador 310, como se muestra en la FIG. 9. Además, el tiempo entre los datos programados para la transmisión y los datos realmente siendo transmitidos se denomina retardo de transmisor y STL 336.

[00125] La definición del tiempo para los datos de los medios es una cuestión de considerar hasta qué punto tienen que estar operando antes del tiempo de radiación los distintos bloques funcionales de la infraestructura de transmisión del sistema 300 y cuánto tiempo después de la recepción se entregarán los medios fuera del decodificador de comunicación. Por consiguiente, esta discusión se divide antes de la radiación y después de la recepción. ¿Por qué es esencial el tiempo de radiación real? Para lograr una conmutación síncrona potencial entre los flujos fuente de una aplicación tal como inserción de anuncios personalizados. El tiempo real que limita la transmisión de los medios puede verse limitado. El tiempo de entrega al receptor se sincroniza con el tiempo de radiación de la característica de arranque de capa física. El lapso de tiempo en que un período específico está presente en la capa física tiene que ser entenderse y controlarse. Un MDE o Segmento de Medios está etiquetado con un intervalo de tiempo al cual se le permite irradiarse, de manera tal que todo proceso antes de la emisión real se desplace en el tiempo, típicamente, mediante un retardo

constante asociado con un bloque o función específica. La FIG. 9 representa una organización plausible de dominios de tiempo dentro de la infraestructura de transmisión. Esta partición no es un requisito, es meramente conveniente desde el punto de vista conceptual. A continuación, se muestra una lista de algunos de los posibles dominios de tiempo y sus duraciones plausibles.

[00126] Retardo de enlace de Transmisor y Transmisor de Estudio (STL) 336: El retardo definido a través de estos bloques es lo suficientemente extenso para que una trama de capa física dada comenzando a entregarse a STL puede comenzar a emitirse por el transmisor, cuando ha transcurrido este período. Este retardo incluye nominalmente al menos una duración de trama de capa física, una duración de intercalador y un retardo de transmisión más largo de la red de distribución SFN. Se podría considerar la ejecución más corta de este retardo, por ejemplo, sobre una base de subtrama. Las subtramas de soporte potencialmente pueden causar un tipo de pico aumentado en el estudio en el enlace de transmisor de estudio, es decir, el enlace de transmisor de estudio debe mantener el máximo rendimiento permitido de capa física por lo menos una duración de subtrama. La restricción en una aplicación de imagen máxima es el pico permitido de asignación de velocidad medido sobre una base de trama completa, que puede ser menor que por subtrama, aunque puede ser igual si todos los PLP fueran de la misma capacidad.

[00127] Retardo de Duración de Análisis 332: Una de las tareas del programador 314 es asignar medios a la capa física de una manera que cumpla con sus plazos de tiempo definidos, y cumple con las limitaciones de capacidad por las definiciones de PLP actuales de capa física. Con el fin de lograr esto, el programador 314 debe analizar al menos la duración más larga de un valor de trama de capa física de los medios. El programador 314 en última instancia podría fallar si no programa con éxito al menos un tiempo de trama de capa física de los medios codificados por tiempo de trama de capa física transcurrido. Esta es una tarea acumulada. Una trama de capa física con menos de la duración de su tiempo en los medios puede compensarse con una trama de capa física que representa más tiempo de medios.

[00128] La duración de una trama de capa física puede ser más corta que la duración del Segmento de los medios más largo o duración GOP. Para tal caso, el límite de limitación está relacionado con al menos la duración más larga del Segmento de medios o la duración de GOP es el GOP que se compone de múltiples Segmentos de medios. El programador 314 puede asegurar que la duración más larga del Segmento de Medios/GOP está programada con éxito. El programador 314 puede generar

adicionalmente al menos un segundo de Segmento de Medios codificado/GOP por segundo de medio. De acuerdo con lo descrito en la presente, el análisis de duración de retardo 332 no es constante, dado que el programador 314 tiene una flexibilidad Tiempo de Radiación más posterior - Tiempo de Radiación Más Temprano en el tiempo de radiación real.

[00129] Retardo de Paso Individual 334: Si los medios de entrada se transmitieran directamente sobre medios de codificador 308 y segmentador 310, habría un retardo de decodificación que es estático para una configuración dada de entrada de una primera trama de GOP al tiempo de decodificación de la primera trama decodificada, como se describe en el primer Segmento resultante. Esta es una construcción conceptual.

[00130] Puede ser el caso de que medios de codificador 308 (que pueden representar múltiples codificadores) probablemente produzcan más de un tiempo de medio de contenido codificado por tiempo de Segmento de Medios. Es probable que las limitaciones reales se definan por el número de “pases” utilizado por el esquema de codificación. En este contexto, “pase” es una instancia de una codificación de los medios. Si los medios de codificador 308 fueron capaces de cumplir con todos los objetivos en una sola pasada de codificación, entonces el requisito sería al menos una duración de Segmento de Medios de medio por el tiempo de Segmento de Medios transcurrido. Es bastante desafiante obtener el tamaño de datos deseado en la primera pasada, por lo que un Programador puede intentar limitar las pasadas a codificación de dos pasadas. En la codificación de dos pasadas, una primera pasada puede proporcionar una calibración de la complejidad de los datos de medios a codificar, y una segunda pasada se puede utilizar para ajustar las decisiones tomadas durante la primera pasada para satisfacer la capacidad de la capa física. Si la pasada de codificación inicial da lugar a tamaños de datos que son sustancialmente diferentes del tamaño de datos deseado/requerido final, y luego se puede realizar una tercera pasada u otro ajuste a la tasa de datos. Cabe señalar que el Programador tiene que trabajar con el tamaño de datos real a entregar en la capa física, que incluye toda encapsulación.

[00131] Suponiendo que los medios se transmiten a los medios de codificador 308 en tiempo real, y que este es un esquema de codificación de pasadas múltiples, potencialmente se requiere un retardo completo de GOP para capturar la transmisión en tiempo real antes de los medios de codificador. Los tiempos de decodificación proporcionados de los medios codificados en el receptor deben reflejar todo el retardo

de la infraestructura y el receptor. Se espera que estos se basen en la reproducción de Segmento de Medios. Las oportunidades de MDE pueden calcularse a partir del punto de puesta en marcha de RAP MDE en relación con lo que habría ocurrido durante la reproducción de nivel de Segmento.

[00132] La FIG. 10 es un diagrama conceptual que ilustra múltiples Segmentos por GOP con una ubicación RAP escalonada. Un tiempo de Segmento RAP escalonado como se ilustra en la FIG. 10 (es decir, múltiples duraciones de análisis por GOP) puede reducir al mínimo el tiempo de captura requerido antes de la programación de los medios y, por lo tanto, una menor latencia general. Esto tiene beneficios de multiplexación estático algo limitados porque, como se ilustra, la tasa de datos para una trama de capa física dada está dominada por un solo tamaño de datos de Segmento RAP. El etiquetado de tiempo de los medios se captura con los medios.

[00133] De acuerdo con lo descrito con anterioridad, el retardo de transmisión, o retardo a la radiación, puede expresarse de la siguiente manera:

Duración de Retardo de Análisis - Retardo de Pasada única a Muestra Específica + Retardo de Transmisor y STL

(Tiempo de Trama + Profundidad de Intercalador + Distribución de SFN)

[00134] La FIG. 11 es un diagrama conceptual que ilustra un retardo mínimo para decodificar tiempo para la reproducción de medios. Es decir, la FIG. 11 representa los retardos mínimos que afectan a los tiempos de decodificación de medios para el Segmento de medios y reproducción de MDE. Una de diferencia entre los Segmentos de Medios y reproducción de MDE es de manera tal que el tiempo de inicio de la reproducción de MDE en la demora de nivel de Segmento deba tener en cuenta el retardo apilado más largo más un retardo de presentación sugerida, mientras que la reproducción de nivel de MDE es el retardo real del dispositivo estático individual más un retardo para asegurar que la plenitud de la memoria intermedia de ROUTE es suficiente. El valor total de retardo de MDE puede ser menor que un tiempo de Segmento.

[00135] El retardo de decodificación de Segmento de extremo a extremo final se puede expresar de la siguiente manera:

Duración de Retardo de Análisis - Retardo de Pasada única a Muestra Específica + Retardo de Transmisor y STL

(Tiempo de Trama + Profundidad de Intercalador + Distribución de SFN) +
Tiempo de Vuelo + Retardo en Disponibilidad de Tiempo de Inicio +
MPD@suggestedPresentationDelay + Retardo dentro del Segmento de

Muestra Específica

[00136] El retardo de decodificación de MDE de extremo a extremo final se puede expresar de la siguiente manera:

Duración de Retardo de Análisis - Retardo de Pasada única a Muestra Específica + Retardo de Transmisor y STL

(Tiempo de Trama + Profundidad de Intercalador + Distribución de SFN) +
Tiempo de Vuelo + Llegada con Retardo en ROUTE o Receptor de MMT
Similar + Tiempo de Mantenimiento en Desconexión de Intervalo de Byte en
Streaming + Retardo dentro del Segmento de Muestra Específica

[00137] Ciertos detalles de los métodos de metadatos de tiempo para la programación se han analizado con anterioridad. En la presente se analizan más detalles sobre métodos de metadatos de tiempo para programación. La construcción de los metadatos de tiempo para MDE se pueden basar en un punto de anclaje definido por el tiempo de transmisión más posterior requerido debido a un inicio de Segmento. El tiempo requerido más posterior para un MDE puede ser lineal con el fin del primer tiempo de decodificación de cada MDE respectivo después del inicio del Segmento. El tiempo más posterior puede verse limitado por el tiempo de decodificación del flujo a decodificar en muestra en el MDE, en relación con la primera muestra decodificada del Segmento, que también es el primer MDE decodificado del Segmento. Este tipo de método es lineal en el tiempo de decodificación, que es solamente una aproximación a la definición de metadatos de tiempo de transmisión más posterior.

[00138] Hay aspectos adicionales a considerar, por ejemplo, componentes de medios posiblemente de carga frontal, tales como audio y subtítulos. Dado que los componentes de estos medios son pequeños en relación con el video, pero la presentación completa puede no reproducirse sin ellos, posiblemente es más simple enviarlos delante del RAP de video, pero después de los objetos de metadatos posiblemente requeridos tal como MPD e IS. Esto es, en general, un método que puede utilizarse para asegurar que todos los pequeños objetos de datos requeridos se entreguen primero.

[00139] La FIG. 12 es un diagrama conceptual que ilustra la sincronización de capa física opcional y metadatos relacionados. La relación de las tramas de capa física a T-RAP es un aspecto relevante con respecto a la primera transmisión. Por ejemplo, véase el ejemplo de la FIG. 12, que muestra el tiempo más temprano de la libertad de moverse por delante del inicio de la trama física nominal. Esto puede afectar el tiempo de cambio de canal, a menos que haya una secuencia completa de metadatos de

comienzo en todas las capas.

[00140] El diseño de una capa física puede o no puede ser capaz de realizar la sincronización con precisión en los límites de GOP como se representa con anterioridad. La definición de la posición de sincronización puede verse limitada a otras ubicaciones por la numerología de la capa física. Un método viable es utilizar una ubicación cercana a la duración de GOP. Cercana podría ser, por ejemplo, la próxima vez disponible, es decir, antes del límite nominal del tráfico precedente que lleva la transformada, o después. La escala de tiempo para una transformada mínima puede estar en la escala de 1 mseg, mientras que el máximo puede estar en el intervalo de 10 mseg.

[00141] La función de programación puede considerar nominalmente un bloque fijo de tiempo, posiblemente relacionado con la duración de GOP de medios o la duración del Segmento de acuerdo con lo representado con anterioridad. Para el esquema de duración de GOP más eficiente, es posible que todos los medios se puedan programar y cierta capacidad significativa puede dejarse disponible en (N) para su uso potencial por el siguiente intervalo de análisis (N+1). En tal condición, el Programador puede intentar llenar esta capacidad en (N) con medios del siguiente intervalo de análisis (N+1). Esto puede activarse mediante metadatos de tiempo, lo que permite la transmisión de un intervalo anterior a través del parámetro más temprano. Una vez conocida la porción llena del intervalo posterior, el componente de capa física puede insertar un elemento de sincronización de trama opcional, proporcionado con el propósito de permitir la adquisición de medios anteriores que los posibles con el generador de tramas cuasi fijado, por ejemplo, bootstrap y preámbulo de ATSC 3.0

[00142] La FIG. 13 representa un ejemplo de un método que incluye una secuencia de pasos que corresponden a posibles eventos de sistema para permitir la adquisición de servicio para múltiples capas de datos de medios. Es decir, la FIG. 13 es un diagrama de flujo que ilustra un ejemplo de secuencia de eventos de adquisición. El método de la FIG. 13 se puede realizar mediante diversos dispositivos y sistemas mostrados en la presente, tal como el dispositivo de preparación de contenidos 20 y dispositivo fuente 60 de la FIG. 1, el sistema 180 de la FIG. 5, o el sistema 300 de la FIG. 9. Por fines de ejemplo y explicación, el método de la FIG. 13 se explica con respecto al sistema 180 de la FIG. 5.

[00143] El sistema 180 proporciona una secuencia de datos que permite el inicio de un servicio lineal o lineal basado en aplicación (350). La FIG. 13 ilustra una secuencia típica de eventos en un receptor y un orden de sistema (180) de entrega de metadatos

relacionados, otros objetos, y archivos multimedia relacionados con un inicio típico de recepción de servicio. Este comienzo del servicio se inicia típicamente (350) por medio del canal de entrada de número, la entrada de clave canal arriba/abajo, o la selección de guía de un servicio. La adquisición de capa física se inicia mediante la recepción de un patrón de sincronización de bootstrap o sistema (352). Los metadatos del sistema con relación a la capa física y la adquisición de los metadatos del sistema adicional se logran a través del preámbulo o la recepción de metadatos de descripción de capa física similar (354). Tras la identificación de la ubicación de PLP deseados u otros flujos de paquetes de capa física en la forma de onda deseada de flujos de paquetes recibidos que contienen ALP diferente a un protocolo de encapsulación similar que contiene un mapa (referencia cruzada) de ID de flujo de paquetes (ID de PLP) a IP contenido u otras ID de flujo de paquete similar o direcciones (356). El preámbulo o similar (354) contiene una bandera que identifica la presencia de metadatos de sistema (LLS) que identifica datos de servicio tales como el tiempo del sistema, una composición de servicios (SLT), una posible información de alerta (AEAT) (358). La SLT disponible en LLS (358) contiene la ubicación de servicio por detalles en SLS (360). Los PLP que transportan el servicio también pueden contener una aplicación que puede incluir una aplicación de tiempo de ejecución (362), que puede permitir al servicio de televisión para operar en una interfaz basada en un navegador. Si está presente, la aplicación puede ejecutarse antes del inicio del servicio o después. Una vez que estos pasos, activados por la recepción de datos correspondientes, se completan, la reproducción de medios puede iniciarse desde la recepción y decodificación de un IS (364) y el Segmento de Medios relacionados (366). Luego, se puede iniciar un servicio lineal o lineal basado en aplicación (368).

[00144] El retardo de decodificación de extremo a extremo de MDE es un tiempo estático anterior a la decodificación de nivel de Segmento. Este desplazamiento estático se puede lograr a través de los metadatos existentes en la MPD o una reescritura de la MPD que cambia el valor de los metadatos para permitir a la MPD reproducir los MDE. El valor del desplazamiento puede determinarse a partir de la primera reproducción de RAP a través de MDE con relación a la primera oportunidad en que RAP se reproduciría en primer lugar desde el nivel del Segmento.

[00145] Las funciones relacionadas con la programación de un MDE pueden incluir:

- La organización temporal de tipos de medios y metadatos para reducir al mínimo el tiempo de adquisición.
 - Proporcionar tipos de datos de medios más pequeños con los tiempos

más posteriores más tempranos para asegurar la entrega completa y como una simplificación en relación con el intento de intercalar con otros medios a través de numerosos objetos programados más pequeños. Este método da lugar a una mayor flexibilidad en la programación, es decir, algunos objetos más grandes con más latitud de tiempo en lugar de numerosos objetos más pequeños con requisitos más rígidos. El aumento de la flexibilidad de programación generalmente da lugar a resultados superiores en términos de eficiencia global.

- Si se desea un rendimiento más robusto para los metadatos esenciales y de audio, pueden programarse en conjunto en una tubería de capa física más robusta.
- En un caso en que se desea una intercalación específica de tipos de medios en una tubería de suministro común, el uso de Segmento multiplexado puede ser más sencillo de implementar. Las muestras en el interior del objeto multimedia multiplexado se proporcionan en orden de entrega/decodificación.
- Métodos para preservar una adquisición rápida por medio de ubicaciones de sincronización de capa física opcionales.
 - El reconocimiento de la entrega de T-RAP con anterioridad de la ubicación nominal, es decir, detrás de la sincronización de capa física nominal o intervalo de análisis nominal relacionados
 - La adaptación basada en el ciclo de programación actual que cumple todas las solicitudes de entrega, es decir, si se han cumplido todas las solicitudes de datos con la última en el período de capa física de trama nominal.
- Medios para adaptarse a la duración de tiempo no exacta de trama de capa física al intervalo de análisis alineado con GOP de medios mediante el uso de reglas que indican que el transmisor debe comenzar en un segundo total conocido y mseg. (El modo de sincronización siempre se ejecutará en segundos pares y milisegundos, por definición.) El modo de símbolos siempre mantendrá un número entero de símbolos de tiempo transcurrido de un segundo y milisegundo de tiempo de ATSC. (Las tasas de símbolos pueden cambiarse entre valores predefinidos especificados en o determinados a partir de información de configuración.)
 - La asignación de sincronización de trama a la ubicación más cercana

del símbolo disponible en tiempo de salida de posición nominal con relación al intervalo de análisis.

- La asignación de sincronización de trama a la ubicación más cercana disponible en un símbolo que conduce a una posición de intervalo de análisis nominal.
- Cualquier otro esquema con base en sincronización de capa física con base en una regla con relación a un intervalo de análisis definido.
- Un intervalo de análisis definido vinculado con GOP de medios o estructura similar.
- Un posible método de baja latencia basado en intervalos de análisis con base en las duraciones de Segmento en lugar de GOP.
 - La gestión de tasa principal con base en Segmentos RAP.
 - El ajuste de arrastre de Segmentos no RAP para compensar PSNR fallido u otra medida de calidad.

[00146] Esta divulgación describe generalmente varias técnicas, que pueden usarse solas o en cualquier combinación, para identificar las ubicaciones de MDE, medios de intervalos de bytes conscientes, dentro de los protocolos de transferencia de archivos, tal como ROUTE o MMT. Estas técnicas pueden estar basadas en la descripción explícita y la coincidencia de patrones válidos o determinación de patrones válidos por las reglas. Estas técnicas pueden combinarse. No hay ninguna restricción en el orden de combinación. De manera similar, esta divulgación también describe aspectos de tiempo como un fondo para la discusión de técnicas para la organización de la sincronización de capa física y otros aspectos de entrega. Estas técnicas pueden permitir la adquisición rápida en sistemas con ubicaciones de sincronización variables o exactas con respecto a un intervalo de análisis en relación con los medios.

[00147] La FIG. 14 es un diagrama de flujo que ilustra otro método de ejemplo de acuerdo con las técnicas de la presente divulgación. El método de la FIG. 14 se explica llevado a cabo por una unidad de envío de protocolo basado en archivos, tal como el emisor 188 de la FIG. 5 o el emisor 312 de la FIG. 9. Por fines de explicación, el método de la FIG. 14 se explica con respecto al emisor 312 de la FIG. 9, pero debe comprenderse que pueden configurarse otras unidades para realizar este método o uno similar. Como se analizó con anterioridad, el emisor 312 representa un ejemplo de una unidad de envío de protocolo basado en archivos, por ejemplo, una unidad que envía datos de acuerdo con el protocolo ROUTE o protocolo FLUTE, u otros protocolos de transmisión basados en archivos.

[00148] Inicialmente, en el ejemplo de la FIG. 14, el emisor 312 recibe uno o más Segmentos, que incluyen uno o más MDE (380). Por ejemplo, los MDE pueden corresponder a intervalos de bytes entregables de los Segmentos. El emisor 312 luego determina ubicaciones MDE (382). En general, el emisor 312 no recibe información explícita de señalización de las ubicaciones de los MDE, sino que utiliza una o más de las técnicas de la presente divulgación para procesar los datos de los Segmentos para determinar las ubicaciones de los MDE. Por ejemplo, el emisor 312 puede llevar a cabo técnicas de coincidencia de patrones o reglas que definen cómo se pueden identificar los MDE, como se analizó con anterioridad. El emisor 312, en consecuencia, puede utilizar las técnicas de coincidencia de patrones, reglas, u otras tales técnicas para identificar los MDE dentro de los Segmentos.

[00149] El emisor 312 además puede determinar los requisitos de tiempo de transmisión para los MDE (384). Por ejemplo, el emisor 312 puede analizar los datos de un archivo de manifiesto que corresponde a los Segmentos, y determinar los requisitos de tiempo de transmisión del archivo de manifiesto. El archivo de manifiesto puede comprender, por ejemplo, una MPD de acuerdo con DASH. Los requisitos de tiempo de transmisión pueden representar uno o ambos de un tiempo de transmisión más temprano o un tiempo de transmisión más posterior para uno o más de los MDE, de acuerdo con lo mencionado con anterioridad.

[00150] En última instancia, el emisor 312 puede proporcionar los MDE y datos que representan los requisitos de tiempo de transmisión para los MDE a una unidad de envío de capa física (386), tal como el programador 314 y el excitador/amplificador 316. De esta manera, el programador 314 puede programar la transmisión de los MDE de acuerdo con los requisitos de tiempo de transmisión, de manera tal que el excitador/amplificador 316 no transmita los datos de los MDE antes del tiempo de transmisión más temprano y/o no más tarde que el tiempo de transmisión más posterior.

[00151] De esta manera, el método de la FIG. 14 representa un ejemplo de un método que incluye, mediante una unidad de envío de protocolo basado en archivos de un dispositivo fuente, recibir un flujo de datos que comprende Segmentos de datos de medios de un segmentador del dispositivo fuente que forma los Segmentos, cada uno de los Segmentos comprende un respectivo archivo individualmente recuperable asociado con un localizador uniforme de recursos (URL) único o un identificador uniforme de recursos (URI), determinar ubicaciones de eventos de entrega de medios (MDE) en el flujo de datos de medios, donde los MDE incluyen datos para al menos

una porción de uno de los Segmentos, determinar uno o más requisitos de tiempo de transmisión para los MDE que representan tiempos en los que los MDE han de enviarse a un dispositivo cliente, y proporcionar los MDE y datos que representan los requisitos de tiempo de transmisión a una unidad de envío de capa física del dispositivo fuente de acuerdo con ranuras de entrega disponibles para la unidad de envío de capa física.

[00152] En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware, o cualquier combinación de los mismos. Si se implementan en software, las funciones pueden almacenarse en o transmitirse como una o más instrucciones o código en un medio legible por computadora y ejecutarse por medio de una unidad de procesamiento basada en hardware. Los medios legibles por computadora pueden incluir medios de almacenamiento legibles por computadora, lo que corresponde a un medio tangible tal como un medio de almacenamiento de datos, o medios de comunicación que incluyen 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 computadora generalmente pueden corresponder a (1) medios de almacenamiento legibles por computadora tangibles no transitorios o (2) un medio de comunicación tal como una onda de señal o portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se pueda acceder por medio de una o más computadoras o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en la presente divulgación. Un producto de programa informático puede incluir un medio legible por computadora.

[00153] A modo de ejemplo, y no de limitación, tales medios de almacenamiento legibles por computadora 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 pueda utilizarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder por medio de una computadora. Además, cualquier conexión se denomina correctamente un medio legible por computadora. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otra fuente remota mediante el uso de un cable coaxial, cable de fibra óptica, par trenzado, línea de abonado digital (DSL), o tecnologías inalámbricas tal como infrarroja, de radio y microondas, entonces el cable coaxial, cable de fibra

óptica, par trenzado, DSL o tecnologías inalámbricas tal como infrarroja, de radio y microondas se incluyen en la definición de medio. Debe comprenderse, sin embargo, que los medios de almacenamiento legibles por computadora y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales, u otros medios transitorios, sino que se dirigen a medios de almacenamiento tangibles no transitorios. Disco flexible y Disco, de acuerdo con lo utilizado en la presente, incluyen CD, disco láser, disco óptico, disco versátil digital (DVD), disquete y disco Blu-ray donde los discos flexibles generalmente reproducen datos magnéticamente, mientras que los discos reproducen datos ópticamente con láseres. También deben incluirse combinaciones de los anteriores dentro del alcance de los medios legibles por computadora.

[00154] Las instrucciones pueden ejecutarse por medio de uno o más procesadores, tal como uno o más procesadores de señal digital (DSP), microprocesadores de propósito general, circuitos integrados de aplicación específica (ASIC), matrices lógicas programables en campo (FPGA), u otro circuito equivalente integrado o lógica discreta. De acuerdo con ello, el término “procesador”, de acuerdo con lo utilizado en la presente puede referirse a cualquiera de la estructura anterior o cualquier otra estructura adecuada para la aplicación de las técnicas descritas en la presente. Además, en algunos aspectos, la funcionalidad descrita en la presente puede proporcionarse dentro de los módulos de hardware y/o software dedicado configurados para la codificación y decodificación, o incorporarse en un códec combinado. Además, las técnicas pueden implementarse completamente en uno o más circuitos o elementos lógicos.

[00155] Las técnicas de la presente divulgación pueden implementarse en una amplia variedad de dispositivos o aparatos, que incluyen un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chip). En esta divulgación se describen diversos componentes, módulos o unidades para enfatizar los aspectos funcionales de los dispositivos configurados para llevar a cabo las técnicas divulgadas, pero no necesariamente requieren una realización por medio de diferentes unidades de hardware. Más bien, como se ha descrito con anterioridad, diversas unidades pueden combinarse en una unidad de hardware códec o proporcionarse mediante una colección de unidades de hardware interoperativas, que incluyen uno o más procesadores como se describió con anterioridad, en conjunción con el software y/o firmware adecuado.

[00156] Se han descrito varios ejemplos. Estos y otros ejemplos están dentro del

alcance de las siguientes reivindicaciones.