

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 687 767**

51 Int. Cl.:

G06F 15/16 (2006.01)

G06F 9/44 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.12.2012 PCT/US2012/068326**

87 Fecha y número de publicación internacional: **20.06.2013 WO13090126**

96 Fecha de presentación y número de la solicitud europea: **07.12.2012 E 12857917 (4)**

97 Fecha y número de publicación de la concesión europea: **27.06.2018 EP 2791815**

54 Título: **Almacenamiento previo en memoria caché de CDN accionado por aplicación**

30 Prioridad:

16.12.2011 US 201113328444

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

29.10.2018

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, WA 98052, US**

72 Inventor/es:

**TULIANI, JONATHAN ROSHAN;
HOLT, NICHOLAS LEONARD y
HUANG, CHENG**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 687 767 T3

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

DESCRIPCIÓN

Almacenamiento previo en memoria caché de CDN accionado por aplicación

Antecedentes

5 Una red de entrega de contenido (CDN, *content delivery network*) es una red de ordenadores que contiene copias de datos colocadas en varios nodos de red. Una CDN proporciona una interfaz para datos entre servidores de origen y ordenadores de usuario final. Los servidores de origen son las fuentes primarias de contenido, y los servidores de las copias de memoria caché de CDN del contenido que tienen la demanda más alta. Los servidores de la CDN se pueden disponer estratégicamente para encontrarse más cerca de los ordenadores de usuario final que los servidores de origen. Los ordenadores de usuario final pueden acceder a los datos con alta demanda en los servidores de la CDN en lugar de tener que acceder a datos directamente desde los servidores de origen. En ese sentido, una CDN mejora el acceso a los datos al aumentar el ancho de banda de acceso, aumentar la redundancia y reducir la latencia de acceso.

15 El servicio Bing® Maps proporcionado por Microsoft Corporation de Redmond, Washington, es un ejemplo de una aplicación en línea que proporciona un contenido usando una CDN. Esta aplicación tiene una gran cantidad de contenido estático en la forma de teselas de mapa (imágenes de porciones de un mapa) almacenadas en servidores de origen que se entrega a los usuarios finales a través de la CDN. Por ejemplo, un usuario puede usar un navegador web para navegar por un mapa en su dispositivo informático, tal como al dar un vistazo panorámico de una parte a otra del mapa, acercarse o alejar una porción del mapa, etc. A medida que el usuario navega por el mapa, el navegador transmite solicitudes de nuevas teselas de mapa a la CDN.

20 Se han desarrollado varias técnicas para posibilitar que el contenido (por ejemplo, las teselas de mapa) se proporcione más rápidamente a un navegador web en el dispositivo informático de usuario. De acuerdo con una primera técnica, un servidor de origen para una aplicación en línea puede predecir solicitudes de contenido futuras. El contenido futuro predicho se puede almacenar previamente en memoria caché en la CDN de tal modo que el dispositivo informático de usuario pueda acceder más rápidamente al mismo si realmente se solicita el contenido futuro predicho. De acuerdo con una segunda técnica, un navegador web de cliente en un dispositivo informático de usuario puede predecir que el contenido puede ser deseado en el futuro, y puede capturar previamente el contenido en el navegador web. De acuerdo con una tercera técnica, un servidor de CDN puede predecir solicitudes de contenido futuras por un dispositivo informático de usuario, y puede almacenar previamente en memoria caché el contenido futuro en la CDN.

30 No obstante, estas técnicas tienen desventajas. Por ejemplo, tener el contenido futuro de predicción de servidores de origen tiene una desventaja que el contenido que se va a almacenar previamente en memoria caché se determina por adelantado. Con respecto a capturar previamente un contenido a un navegador web, en los casos en los que el usuario incurre a un coste por transferencias de datos o tiene ancho de banda limitado (tal como en aplicaciones móviles), tal almacenamiento previo en memoria caché puede no ser aceptable y / o puede no ofrecer desempeño deseado. Además, con respecto a hacer que la CDN prediga y almacene previamente en memoria caché el contenido futuro, la CDN tendría que ser muy sofisticada para entender los tipos de solicitudes de contenido que la CDN recibe para ser capaz de inferir solicitudes de contenido futuras.

40 En ese sentido, las técnicas de almacenamiento en memoria caché actuales son poco deseables. En el caso del servicio Bing® Maps, debido al número de teselas de mapa que están disponibles y las diferentes áreas de atención de los diferentes usuarios, la relación de "aciertos de memoria caché" (relación de datos solicitados que se almacenan en memoria caché frente a la cantidad total de datos disponibles) es indeseablemente baja (por ejemplo, menor que un 50 %). Esto da como resultado una latencia relativamente alta en cuanto a la carga de mapas, debido a que, en lugar de recuperar las teselas del mapa a partir de la caché en la CDN, las teselas de mapa frecuentemente tienen que recuperarse a partir de los servidores de origen para los usuarios.

45 El documento US 2008/208789 A1 se refiere a la aceleración de HTTP mediante predicción y captura previa. La aceleración a través de predicción y captura previa mejorará de forma significativa el tiempo para recuperar objetos. Para lograr la aceleración a través de predicción y captura previa, un servidor representante web recibe solicitudes a partir de los clientes. Los clientes pueden no tener conocimiento alguno de la presencia del servidor representante web. Cuando el servidor representante web recibe en primer lugar una solicitud web a partir del cliente, el servidor representante web analiza la solicitud para determinar qué contenido se está solicitando y, por lo general, busca en una memoria caché local para ver si el contenido solicitado está almacenado en memoria caché y sigue siendo válido. En el caso de que el contenido solicitado sea necesario, el servidor representante web crea una solicitud web para ese objeto. Cuando se recibe la respuesta que se corresponde con un objeto de página principal, un manipulador de respuestas del servidor representante web devuelve la respuesta al cliente, y también proporciona la respuesta a una lógica de análisis. Por lo general, el cliente comenzará a solicitar los objetos enumerados en la respuesta de página principal, usando el número limitado de conexiones web un tiempo que se corresponde con una solicitud a lo largo de cada conexión permitida. No obstante, en lugar de aguardar las solicitudes de cliente, la lógica de análisis comienza la captura previa de los objetos enumerados en el objeto de página principal que la lógica de análisis predice que el cliente estará solicitando subsiguientemente.

60 El documento US 2002/062384 A1 se refiere a las plataformas en red. Un dispositivo de red comprende un servidor

representante que incluye una memoria caché, un módulo de soporte lógico de agente de captura previa y un módulo de soporte lógico de análisis. En lugar de simplemente enviar los URL solicitados a un cliente, en el servidor representante remoto el módulo de soporte lógico de análisis analiza los archivos de HTML y crea una lista de URL que estaban enlazados o embebidos en el archivo de HTML. El servidor representante puede capturar entonces alguno o la totalidad de los URL en esa lista y almacenar los mismos en su memoria caché, sin aguardar en primer lugar a que el cliente emita solicitudes para esos URL. Se pueden usar módulos de soporte lógico de directivas para equilibrar los efectos no deseables de la captura previa con sus beneficios. En un ejemplo, en lugar de capturar previamente todos los URL enlazados en una página de HTML, solo se capturan previamente los URL embebidos. En otro ejemplo, una directiva de captura previa incluye una captura previa selectiva basándose en la probabilidad de que un usuario acceda en realidad al URL enlazado, para reducir el ancho de banda y el espacio de memoria caché que se usan de forma innecesaria. Esta directiva se podría poner en práctica al examinar cómo de popular es un enlace. Esta información se puede obtener, por ejemplo, mediante la puesta en práctica de una ampliación especial del protocolo HTTP en la que el módulo de soporte lógico de agente de captura previa obtiene, cuando se ejecuta, una información de señal acerca de la probabilidad de que se acceda a cualquier enlace en la página a partir del servidor de contenido. En un ejemplo, un dispositivo de cliente también incluye un soporte lógico de representante local, de tal modo que, por lo general, las operaciones que son llevadas a cabo por el servidor representante remoto pueden descargarse al dispositivo de cliente. El representante local en el dispositivo de cliente puede poner en práctica un análisis, tal como se ha descrito previamente.

El documento US 2003/126232 A1 se refiere a una mejora de desempeño de la captura previa en un entorno de red. Se describe un modelo de predicción de captura previa que tiene unos parámetros de uso de energía para predecir el impacto de los archivos especificados de captura previa en el uso de energía de un sistema. Un motor de predicción de captura previa utiliza el modelo de predicción de captura previa para evaluar los archivos especificados con respecto a los criterios de captura previa, incluyendo criterios de captura previa de eficiencia energética, y genera una decisión de captura previa con respecto a cada archivo de los archivos especificados. En un ejemplo, un mensaje de solicitud que es enviado por un cliente al servidor incluye un encabezado y un cuerpo. El encabezado almacena información acerca del mensaje tal como su destino y codificación. En un ejemplo, un encabezado también incluye un campo adicional para almacenar sugerencias de cliente. Las sugerencias de cliente son un subconjunto de los parámetros del modelo de predicción de captura previa del cliente o parámetros derivados al menos en parte del modelo de predicción de captura previa del cliente. El modelo de eficiencia de captura previa de lado de servidor incluye varios tipos de parámetros. Algunos de los parámetros en el modelo de eficiencia de captura previa de lado de servidor tienen unos valores que cambian de forma dinámica de acuerdo con los cambios en el entorno operativo del sistema informático de servidor. El modelo de eficiencia de captura previa incluye, entre otras cosas, parámetros de atributos de cliente que son enviados por el cliente a partir del modelo de predicción de captura previa del cliente.

35 **Sumario**

El objeto de la presente invención es reducir el espacio de memoria que se usa en, y el ancho de banda que es usado por, los servidores de memoria caché de la técnica anterior.

El presente objeto se soluciona por medio de la materia objeto de las reivindicaciones independientes.

Se definen algunas formas de realización preferidas por medio de las reivindicaciones dependientes.

40 El presente sumario se proporciona para presentar una selección de conceptos en una forma simplificada que se describe adicionalmente a continuación en la descripción detallada. No se tiene por objeto que el presente sumario identifique características clave o características esenciales de la materia objeto que se reivindica, ni se tiene por objeto que se use para limitar el alcance de la materia objeto que se reivindica.

45 Se proporcionan procedimientos, sistemas y productos de programa informático para el almacenamiento en memoria caché de un contenido antes de que el contenido se solicite en realidad. Una aplicación de cliente puede predecir el contenido que puede ser solicitado subsiguientemente al contenido deseado (en lugar de que se haga la predicción en el servidor de origen o en un servidor de red de entrega de contenido). Además, el contenido predicho se puede almacenar en memoria caché en un servidor de almacenamiento en memoria caché de una red de entrega de contenido (en lugar de almacenarse en memoria caché en el navegador web). Tales técnicas pueden aumentar las relaciones de aciertos de memoria caché, reducir la latencia y / o tener beneficios adicionales.

50 En una puesta en práctica de procedimiento en un servidor de almacenamiento en memoria caché, se recibe una primera solicitud de contenido deseado a partir de una aplicación de cliente. La primera solicitud también puede indicar un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente. No se indica en la actualidad que se desee el consumo del contenido adicional indicado, sino que se predice que, potencialmente, su consumo va a ser deseado en el futuro. El contenido deseado y el contenido adicional indicado se recuperan a partir de un servidor de origen. El contenido deseado se transmite a la aplicación de cliente en un dispositivo de usuario, y el contenido adicional se almacena en memoria caché en el servidor de almacenamiento en memoria caché. Subsiguientemente, se puede recibir una segunda solicitud a partir de la aplicación de cliente que incluye una solicitud del contenido adicional (por ejemplo, debido a que, en la actualidad, se desea el consumo del contenido adicional en la aplicación de cliente). El contenido adicional, que se almacena en

memoria caché en el servidor de almacenamiento en memoria caché en respuesta a la primera solicitud, se sirve a la aplicación de cliente por medio del servidor de almacenamiento en memoria caché en respuesta a la segunda solicitud.

5 La primera solicitud de contenido deseado puede incluir una indicación de probabilidad que indica una probabilidad de que el contenido adicional pueda ser solicitado subsiguientemente por la aplicación de cliente. Una solicitud del contenido adicional que se indica en el servidor de origen se puede priorizar basándose en la indicación de probabilidad (en relación con otras solicitudes de contenido) y / u otra información.

10 En una puesta en práctica de sistema, un servidor de almacenamiento en memoria caché puede incluir un analizador de solicitudes de contenido, un módulo de recuperación de contenido y un módulo de provisión de contenido. El analizador de solicitudes de contenido recibe una solicitud de contenido deseado a partir de una aplicación de cliente en un dispositivo de usuario. La solicitud indica un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente. El módulo de recuperación de contenido transmite al menos una solicitud del contenido deseado en el contenido adicional indicado a un servidor de origen, recibe un contenido deseado y el contenido adicional a partir del servidor de origen, y almacena en memoria caché el contenido adicional en almacenamiento. El módulo de provisión de contenido transmite el contenido deseado a la aplicación de cliente en el dispositivo de usuario. El analizador de solicitudes de contenido recibe una segunda solicitud a partir de la aplicación de cliente que incluye una solicitud del contenido adicional. El módulo de provisión de contenido sirve el contenido adicional de almacenamiento en memoria caché a la aplicación de cliente en respuesta a la segunda solicitud.

20 Además, el módulo de recuperación de contenido puede incluir una unidad de priorización de solicitudes que prioriza la transmisión de una solicitud del contenido adicional indicado al servidor de origen basándose en una indicación de probabilidad.

25 En una puesta en práctica de aplicación de cliente, la aplicación de cliente puede incluir una interfaz de comunicación y una unidad de predicción de contenido adicional. La unidad de predicción de contenido adicional recibe una indicación de contenido actualmente deseado por la aplicación de cliente, y predice el contenido adicional que puede ser solicitado subsiguientemente por la aplicación de cliente. La interfaz de comunicación genera una primera solicitud del contenido deseado que también indica el contenido adicional predicho. La interfaz de comunicación recibe el contenido deseado a partir de un servidor de almacenamiento en memoria caché que recuperó el contenido deseado a partir de un servidor de origen en respuesta a la primera solicitud. Subsiguientemente, la interfaz de comunicación genera una segunda solicitud, que es una solicitud del contenido adicional predicho como un contenido actualmente deseado. El contenido adicional predicho es recibido a partir del servidor de almacenamiento en memoria caché. El servidor de almacenamiento en memoria caché recuperó previamente el contenido adicional a partir de un servidor de origen y ha almacenado en memoria caché el contenido adicional en respuesta a la primera solicitud.

35 Además, la unidad de predicción de contenido adicional puede incluir una unidad de priorización de contenido adicional. La unidad de priorización de contenido adicional genera una indicación de probabilidad que indica una probabilidad de que el contenido adicional pueda ser solicitado subsiguientemente por la aplicación de cliente.

40 También se describen en el presente documento unos productos de programa informático para aplicaciones de cliente que predicen solicitudes de contenido futuras, para servidores de almacenamiento en memoria caché que almacenan en memoria caché un contenido futuro predicho, así como unas formas de realización adicionales.

45 Características y ventajas adicionales de la invención, así como la estructura y el funcionamiento de varias formas de realización de la invención, se describen con detalle a continuación con referencia a los dibujos adjuntos. Se hace notar que la invención no se limita a las formas de realización específicas que se describen en el presente documento, sino que se define por medio de las reivindicaciones. Tales formas de realización se presentan en el presente documento solo para fines a modo de ejemplo. Algunas formas de realización adicionales les resultarán evidentes a los expertos en la materia o materias relevantes basándose en las enseñanzas contenidas en el presente documento.

Breve descripción de los dibujos / figuras

50 Las figuras adjuntas, que se incorporan en el presente documento y forman una parte de la memoria descriptiva, ilustran la presente invención y, junto con la descripción, sirven adicionalmente para explicar los principios de la invención y para posibilitar que un experto en la materia pertinente haga y use la invención.

La figura 1 muestra un diagrama de bloques de una red de entrega de contenido que entrega un contenido a dispositivos de usuario, de acuerdo con una forma de realización a modo de ejemplo.

55 La figura 2 muestra un diagrama de bloques de una red de entrega de contenido en el que se recupera un contenido predicho y es almacenado en memoria caché por un servidor de almacenamiento en memoria caché, de acuerdo con una forma de realización a modo de ejemplo.

La figura 3 muestra un diagrama de flujo que proporciona un procedimiento para que un servidor de almacenamiento en memoria caché almacene en memoria caché un contenido que puede ser solicitado en el

futuro por una aplicación de cliente, de acuerdo con una forma de realización a modo de ejemplo.

La figura 4 muestra un diagrama de flujo que proporciona un procedimiento para proporcionar un contenido previamente almacenado en memoria caché a una aplicación de cliente, de acuerdo con una forma de realización a modo de ejemplo.

5 La figura 5 muestra un diagrama de flujo que proporciona un procedimiento para que una aplicación de cliente solicite y reciba un contenido deseado y para indicar un contenido que puede ser solicitado en el futuro, de acuerdo con una forma de realización a modo de ejemplo.

La figura 6 muestra un diagrama de bloques de un dispositivo de usuario que tiene una aplicación de cliente configurada para solicitar y recibir un contenido deseable y para indicar un contenido que puede ser solicitado en el futuro, de acuerdo con una forma de realización a modo de ejemplo.

10 La figura 7 muestra un diagrama de bloques de una unidad de predicción de contenido adicional que incluye una unidad de priorización de contenido adicional, de acuerdo con una forma de realización a modo de ejemplo.

La figura 8 muestra un procedimiento para indicar una probabilidad de que el contenido pueda ser solicitado por una aplicación de cliente en el futuro, de acuerdo con una forma de realización a modo de ejemplo.

15 La figura 9 muestra un diagrama de bloques de un servidor de almacenamiento en memoria caché, de acuerdo con una forma de realización a modo de ejemplo.

La figura 10 muestra un diagrama de bloques de un módulo de recuperación de contenido que incluye una unidad de priorización de solicitudes, de acuerdo con una forma de realización a modo de ejemplo.

20 La figura 11 muestra un procedimiento para priorizar las solicitudes a un servidor de origen para un contenido que puede ser solicitado por una aplicación de cliente en el futuro, de acuerdo con una forma de realización a modo de ejemplo.

La figura 12 muestra un diagrama de bloques de un ordenador a modo de ejemplo que puede ser usado para poner en práctica las formas de realización de la presente invención.

25 Las características y ventajas de la presente invención se harán más evidentes a partir de la descripción detallada que se describe a continuación cuando se tome junto con los dibujos, en los que caracteres de referencia similares identifican elementos correspondientes de principio a fin. En los dibujos, números de referencia similares indican elementos idénticos, funcionalmente similares y / o estructuralmente similares. El dibujo en el cual aparece en primer lugar un elemento se indica por el dígito o dígitos a la izquierda en el número de referencia correspondiente.

Descripción detallada

I. Introducción

La presente memoria descriptiva describe una o más formas de realización que incorporan las características de la invención. La forma o formas de realización que se describen simplemente ilustran la invención. El alcance de la invención no se limita a la forma o formas de realización que se describen. La invención se define por medio de las reivindicaciones adjuntas al presente documento.

35 Las referencias en la memoria descriptiva a “una forma de realización”, “forma de realización”, “una forma de realización a modo de ejemplo”, etc., indican que la forma de realización que se describe puede incluir un aspecto, estructura o característica particular, pero que no toda forma de realización tiene que incluir, necesariamente, el aspecto, estructura o característica particular. Además, tales expresiones no están, necesariamente, haciendo referencia a la misma forma de realización. Además, cuando se describe un aspecto, estructura o característica particular en conexión con una forma de realización, se afirma que efectuar tal aspecto, estructura o característica se encuentra dentro del conocimiento de un experto en la materia, en conexión con otras formas de realización, ya se describan, o no, de forma explícita.

45 Numerosas formas de realización a modo de ejemplo de la presente invención se describen tal como sigue. Se hace notar que no se tiene por objeto que encabezado alguno de sección / subsección que se proporcione en el presente documento sea limitante. Las formas de realización se describen a través de este documento, y cualquier tipo de forma de realización se puede incluir bajo cualquier sección / subsección.

II. Formas de realización a modo de ejemplo

50 Una red de entrega de contenido (CDN, *content delivery network*) es una red de ordenadores que contienen copias de datos colocadas en varios nodos de red. Una CDN proporciona una interfaz para datos entre servidores de origen y ordenadores de usuario final. Los servidores de origen son las fuentes primarias de contenido, y los servidores de las copias de memoria caché de CDN del contenido de los servidores de origen que tienen la demanda más alta. Los servidores de la CDN se pueden disponer estratégicamente para encontrarse más cerca de los ordenadores de usuario final que los servidores de origen. Los ordenadores de usuario final pueden acceder a los datos con alta demanda en los servidores de la CDN en lugar de tener que acceder a los datos directamente desde los servidores de origen. En ese sentido, una CDN mejora el acceso a los datos al aumentar el ancho de banda de acceso, aumentar la redundancia y reducir la latencia de acceso. Los tipos de contenido que se pueden almacenar en memoria caché en las CDN incluyen objetos web (texto, gráficos, URL y secuencias de comandos), objetos descargables (archivos de medios, soporte lógico, documentos), aplicaciones, medios de transmisión por secuencias en directo y consultas de base de datos.

Se han desarrollado varias técnicas para posibilitar que el contenido (por ejemplo, teselas de mapa) se proporcione más rápidamente a un navegador web en el dispositivo informático de usuario. De acuerdo con una primera técnica, un servidor de origen para una aplicación en línea puede predecir solicitudes de contenido futuras. El contenido futuro predicho se puede almacenar previamente en memoria caché en la CDN de tal modo que el dispositivo informático de usuario pueda acceder más rápidamente al mismo si el contenido futuro predicho se solicita realmente. De acuerdo con una segunda técnica, un navegador web de cliente puede predecir qué contenido puede ser deseable en el futuro, y puede capturar previamente el contenido al navegador web en el dispositivo informático de usuario (por ejemplo, usando técnicas de AJAX (JavaScript asincrónico)). Por ejemplo, un usuario puede usar un navegador web para ver una serie de imágenes en un dispositivo informático. Cada vez que se ve una imagen, se puede predecir en el servidor de origen una imagen siguiente que se va a observar, y la imagen siguiente predicha se puede capturar previamente en el navegador web para posibilitar una transición uniforme e instantánea de una imagen actual a una imagen siguiente. De acuerdo con una tercera técnica, un servidor de CDN puede predecir solicitudes de contenido futuras por un dispositivo informático de usuario, y puede almacenar previamente en memoria caché el contenido futuro predicho en la CDN.

No obstante, estas técnicas tienen desventajas. Por ejemplo, tener el contenido futuro de predicción de servidores de origen tiene una desventaja que el contenido que se va a almacenar previamente en memoria caché se determina por adelantado. Con respecto a capturar previamente un contenido a un navegador web, en los casos en los que el usuario incurre a un coste por transferencias de datos o tiene ancho de banda limitado (tal como en aplicaciones móviles), tal almacenamiento previo en memoria caché puede no ser aceptable y / o puede no ofrecer desempeño deseado. Además, con respecto a hacer que la CDN prediga y almacene previamente en memoria caché el contenido futuro, la CDN tendría que ser muy sofisticada para entender los tipos de solicitudes de contenido que la CDN recibe para ser capaz de inferir solicitudes de contenido futuras.

Las formas de realización de la presente invención superan tales desventajas. En una forma de realización, la lógica para predecir un contenido que puede ser deseable en el futuro se puede poner en práctica en una aplicación de cliente en el dispositivo informático de usuario (por ejemplo, en un navegador web u otra aplicación de cliente). En ese sentido, la lógica se puede adaptar más específicamente a la aplicación de cliente, y puede evolucionar (por ejemplo, a través de actualizaciones) a medida que evoluciona la aplicación de cliente.

Además, en una forma de realización, el contenido futuro predicho se puede almacenar en memoria caché en el almacenamiento en la CDN, en lugar de almacenarse previamente en memoria caché en el navegador web. El almacenamiento previo en memoria caché para su almacenamiento en la CDN no aumenta de forma significativa los datos transferidos al navegador web y, por lo tanto, no afecta de forma adversa al desempeño de los navegadores web, y no presenta problemas en los que las transferencias de datos tienen un coste o en los que se encuentra presente un ancho de banda limitado.

En ese sentido, algunas formas de realización posibilitan el almacenamiento previo en memoria caché del contenido en servidores de almacenamiento en memoria caché de una CDN en lugar de en las propias aplicaciones de cliente. Se posibilita que la lógica de negocios resida dentro de la aplicación de cliente que decide qué contenido se va a almacenar previamente en memoria caché. La aplicación de cliente puede sugerir a una CDN en lo que respecta a qué contenido almacenar previamente en memoria caché directamente basándose en el contexto / comportamiento de usuario y / u otros factores. La sugerencia de almacenamiento previo en memoria caché puede incluir, de forma opcional, una indicación de una probabilidad de que el contenido se use en el futuro, permitiendo que la CDN priorice solicitudes de almacenamiento previo en memoria caché de acuerdo con la capacidad disponible.

Las formas de realización se pueden poner en práctica en cualquier tipo de CDN. Por ejemplo, la figura 1 muestra un diagrama de bloques de una CDN 100 que entrega un contenido a dispositivos de usuario, de acuerdo con una forma de realización a modo de ejemplo. La CDN 100 se muestra como un tipo a modo de ejemplo de red de entrega de contenido, y no se tiene por objeto que sea limitante. Tal como se muestra en la figura 1, la CDN 100 incluye un servidor de origen 102 y un primer y un segundo servidores de almacenamiento en memoria caché 104a y 104b. Además, tal como se muestra en la figura 1, la CDN 100 entrega un contenido a un primer a un cuarto dispositivos de usuario 106a - 106d. A pesar de que, en la figura 1, se muestran para fines a modo de ejemplo un servidor de origen 102 individual, dos servidores de almacenamiento en memoria caché 104a y 104b y cuatro dispositivo de usuario 106a - 106d, se puede encontrar presente cualquier número de estas características de la figura 1, incluyendo uno o más servidores de origen adicionales, uno o más servidores de almacenamiento en memoria caché adicionales y / o uno o más dispositivos de usuario adicionales, incluyendo decenas, cientos, miles, e incluso unos números más grandes de servidores y / o dispositivos de usuario. En una forma de realización, los servidores de almacenamiento en memoria caché 104a y 104b pueden estar o no incluidos en una agrupación de servidores de almacenamiento en memoria caché (así como, de forma opcional, servidores de almacenamiento en memoria caché adicionales), y se puede encontrar presente cualquier número de agrupaciones de servidores de almacenamiento en memoria caché.

Cada uno de los dispositivos de usuario 106a - 106d puede ser cualquier tipo de dispositivo informático estacionario o móvil, incluyendo un ordenador de escritorio (por ejemplo, un ordenador personal, etc.), o un ordenador o dispositivo informático móvil (por ejemplo, un dispositivo Palm®, un dispositivo BlackBerry® de RIM, un asistente digital personal (PDA, *personal digital assistant*), un ordenador portátil, un ordenador ultraportátil, un ordenador de

tipo tableta (por ejemplo, un iPad™ de Apple), un ultraportátil, etc.), un teléfono inteligente (por ejemplo, un iPhone de Apple, un teléfono Android™ de Google, un teléfono de Microsoft Windows®, etc.), u otro tipo de dispositivo informático.

5 Tal como se muestra en la figura 1, el servidor de origen 102 y los servidores de almacenamiento en memoria caché 104a y 104b están acoplados de forma comunicativa a través de una red 108a, y los servidores de almacenamiento en memoria caché 104a y 104b y los dispositivos de usuario 106a - 106d están acoplados de forma comunicativa a través de una red 108b. Las redes 108a y 108b pueden ser unas redes separadas o se pueden incluir en una red individual. Los ejemplos de las redes 108a y 108b incluyen una LAN (*local area network*, red de área local), una WAN (*wide area network*, red de área extensa), o una combinación de redes, tales como Internet. Los ejemplos de los enlaces de comunicación que se pueden incluir en las redes 108a y 108b incluyen enlaces inalámbricos de la norma 802.11 de IEEE de LAN inalámbrica (WLAN, *wireless LAN*), enlaces de interoperabilidad mundial para acceso por microondas (Wi-MAX, *Worldwide Interoperability for Microwave Access*), enlaces de red celular, enlaces de red de área personal (PAN, *personal area network*) inalámbrica (por ejemplo, enlaces de Bluetooth™), enlaces de Ethernet, enlaces de USB (*universal serial bus*, bus serie universal), etc. Cada uno del servidor de origen 102 y los servidores de almacenamiento en memoria caché 104a y 104b puede ser cualquier tipo de dispositivo informático que se describa en el presente documento o que se conozca de otro modo.

20 Tal como se muestra en la figura 1, cada uno de los dispositivos de usuario 106a - 106d incluye una correspondiente de las aplicaciones de cliente 110a - 110d. Las aplicaciones de cliente 110a - 110d son unas aplicaciones que se ejecutan en los dispositivos de usuarios 106a - 106d que acceden a un contenido a través de la CDN 100. Los ejemplos de las aplicaciones de cliente 110a - 110d incluyen navegadores web, reproductores de medios (por ejemplo, reproductores de vídeo, visualizadores de imagen, reproductores de audio, etc.), y otros tipos de aplicaciones de cliente. Tal como se muestra en la figura 1, el servidor de origen 102 almacena el contenido 112 en el almacenamiento al que se puede desear que accedan las aplicaciones de cliente 110a - 110d. El contenido 112 puede ser cualquier tipo de contenido, incluyendo objetos web (por ejemplo, texto, gráficos / imágenes / vídeos, URL (*uniform resource locator*, localizador de recursos uniformes), secuencias de comandos, etc.), objetos descargables (por ejemplo, archivos de medios, soporte lógico, documentos, etc.), aplicaciones, medios de transmisión por secuencias en directo y datos de base de datos. En algunos casos, las aplicaciones de cliente 110a - 110d pueden acceder al contenido 112 directamente en el servidor de origen 102. Además, cada uno de los servidores de almacenamiento en memoria caché 104a - 104b puede almacenar en memoria caché unas porciones de contenido 112 como el contenido almacenado en memoria caché 114a y el contenido almacenado en memoria caché 114b, de forma respectiva. En ese sentido, en algunos casos las aplicaciones de cliente 110a - 110d pueden acceder al contenido 112 en los servidores de almacenamiento en memoria caché 104a y 104b como el contenido almacenado en memoria caché 114a y 114b, en lugar de tener que obtener el contenido 112 directamente a partir del servidor de origen 102, lo que puede ser un cuello de botella.

35 Por ejemplo, tal como se muestra en la figura 1, la aplicación de cliente 110a del dispositivo de usuario 106a puede transmitir una solicitud de contenido 116 al servidor de almacenamiento en memoria caché 104a. La solicitud de contenido 116 indica un contenido que se predice que va a ser solicitado en el futuro por la aplicación de cliente 110a y un contenido inmediatamente deseado por la aplicación de cliente 110a. Por lo tanto, la aplicación de cliente 110a comunica al servidor de almacenamiento en memoria caché 104a el contenido que puede ser solicitado en el futuro y, por lo tanto, que puede ser deseable que sea almacenado en memoria caché por el servidor de almacenamiento en memoria caché 104a. Obsérvese que, en la forma de realización de la figura 1, la solicitud de contenido 116 indica tanto el contenido deseado por la aplicación de cliente 110a como el contenido que se predice que va a ser solicitado en el futuro por la aplicación de cliente 110a. En otra forma de realización, la solicitud de contenido 116 puede indicar el contenido que se predice que va a ser solicitado en el futuro por la aplicación de cliente 110a, y el contenido deseado por la aplicación de cliente 110a se puede indicar en una solicitud separada transmitida de la aplicación de cliente 110a al servidor de almacenamiento en memoria caché 104a.

50 El servidor de almacenamiento en memoria caché 104a puede generar una solicitud de contenido 118 que se transmite al servidor de origen 102, que solicita el contenido futuro predicho que se indica en la solicitud de contenido 116. Si el contenido deseado solicitado por la aplicación de cliente 110a no está almacenado aún en memoria caché en el servidor de almacenamiento en memoria caché 104a, el servidor de almacenamiento en memoria caché 104a puede indicar el contenido deseado en la solicitud de contenido 118, o puede transmitir una solicitud separada al servidor de origen 102 (u otro servidor de origen) para el contenido deseado. El servidor de origen 102 puede transmitir el contenido del contenido 112 al servidor de almacenamiento en memoria caché 104a en una respuesta 120 a la solicitud de contenido 118. La respuesta 120 puede incluir el contenido deseado y el contenido futuro predicho. Como alternativa, el contenido deseado y el contenido futuro predicho se pueden transmitir del servidor de origen 102 al servidor de almacenamiento en memoria caché 104a en unas transmisiones separadas, o el contenido deseado puede ser un contenido futuro predicho que se puede transmitir a partir de diferentes servidores de origen. El servidor de almacenamiento en memoria caché 104a puede almacenar en memoria caché el contenido deseado y el contenido futuro predicho como el contenido almacenado en memoria caché 114a, y puede transmitir el contenido deseado a la aplicación de cliente 110a como el contenido deseado 122. Si la aplicación de cliente 110a solicita subsiguientemente, de hecho, un contenido a partir del servidor de almacenamiento en memoria caché 104a que fue el contenido futuro predicho, el servidor de almacenamiento en memoria caché 104a puede transmitir el contenido almacenado en memoria caché 114a a la aplicación de cliente

110a, que incluye un contenido futuro predicho. Las aplicaciones de cliente 110b - 110d pueden interactuar de forma similar con cualquiera o ambos de los servidores de almacenamiento en memoria caché 104a y 104b para recibir un contenido del servidor original 102 en una forma más eficiente que en las técnicas convencionales, debido al almacenamiento en memoria caché de un contenido futuro predicho en el servidor de almacenamiento en memoria caché 104a y 104b.

Tal como se ha descrito en lo que antecede, en una forma de realización, las aplicaciones de cliente pueden predecir el contenido al que se va a acceder en el futuro, y el contenido predicho se puede almacenar en memoria caché en los servidores de almacenamiento en memoria caché para una respuesta más rápida. En una forma de realización, las aplicaciones de cliente 110a - 110d de la figura 1 pueden predecir un contenido al que se va a acceder en el futuro, y pueden indicar el contenido futuro predicho a los servidores de almacenamiento en memoria caché 104a y 104b para su almacenamiento en memoria caché. Por ejemplo, la figura 2 muestra un diagrama de bloques de una CDN 200, de acuerdo con una forma de realización a modo de ejemplo. Tal como se muestra en la figura 2, la CDN 200 incluye un servidor de origen 102, un servidor de almacenamiento en memoria caché 202 y un dispositivo de usuario 106a. El servidor de almacenamiento en memoria caché 202 es un ejemplo de uno o más servidores de almacenamiento en memoria caché 104a y 104b de la figura 1. Tal como se muestra en la figura 2, el dispositivo de usuario 106a incluye una aplicación de cliente 204 y la aplicación de cliente 204 incluye una unidad de predicción de contenido adicional 208. La aplicación de cliente 204 es un ejemplo de una de las aplicaciones de cliente 110a - 110d de la figura 1. En la figura 2, la aplicación de cliente 204 está configurada para predecir el contenido que se puede solicitar subsiguientemente y el contenido predicho es recuperado y almacenado en memoria caché por el servidor de almacenamiento en memoria caché 202.

La figura 2 se describe tal como sigue con respecto a la figura 3. La figura 3 muestra un diagrama de flujo 300 que proporciona un procedimiento para que un servidor de almacenamiento en memoria caché almacene en memoria caché el contenido que puede ser solicitado en el futuro por una aplicación de cliente, de acuerdo con una forma de realización a modo de ejemplo. En una forma de realización, el diagrama de flujo 300 puede ser llevado a cabo por el servidor de almacenamiento en memoria caché 202 de la figura 2. En una forma de realización, el diagrama de flujo 300 puede ser llevado a cabo conjuntamente por múltiples servidores de almacenamiento en memoria caché que comparten el almacenamiento en memoria caché. Además, las comunicaciones en el diagrama de flujo 300 con respecto a un servidor de origen se pueden llevar a cabo con un servidor de origen individual o múltiples servidores de origen. Algunas formas de realización estructurales y operativas adicionales les resultarán evidentes a los expertos en la materia o materias relevantes basándose en el siguiente análisis con respecto al diagrama de flujo 300 y el servidor de almacenamiento en memoria caché 202.

El diagrama de flujo 300 comienza con la etapa 302. De la etapa 302, se recibe una solicitud de contenido deseado a partir de una aplicación de cliente en un dispositivo de usuario, indicando la solicitud un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente. Por ejemplo, tal como se muestra en la figura 2, el servidor de almacenamiento en memoria caché 202 puede recibir una primera solicitud de contenido 210 deseada a partir de la aplicación de cliente 204 del dispositivo de usuario 106a. La primera solicitud de contenido 210 deseada es una solicitud de contenido deseado tal como una tesela de mapa de un mapa que es observado por un usuario que usa la aplicación de cliente 204, una imagen en una serie de imágenes que son observadas por un usuario que usa la aplicación de cliente 204, una trama de vídeo de un objeto de vídeo (por ejemplo, un archivo de vídeo) que es observado por un usuario que usa la aplicación de cliente 204, una trama de audio de un objeto de audio (por ejemplo, un archivo de audio) que es reproducido por un usuario que usa la aplicación de cliente 204, un artículo de contenido de una página web que va a ser observada por un usuario que usa la aplicación de cliente 204, y / u otro contenido que se describa en cualquier otra parte en el presente documento o que se conozca de otro modo.

Además, una primera solicitud de contenido 210 incluye una indicación de un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente 204. En una forma de realización, la unidad de predicción de contenido adicional 208 de la aplicación de cliente 204 predice un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente 204. Por ejemplo, el contenido adicional predicho puede ser una o más teselas de mapa adicionales del mapa que es observado por el usuario que se predice que va(n) a ser observada(s) subsiguientemente por el usuario, una o más imágenes adicionales en la serie de imágenes que son observadas por el usuario que se predice que va(n) a ser observada(s) subsiguientemente por el usuario, una o más tramas de vídeo adicionales del objeto de vídeo que es observado por el usuario que se predice que va(n) a ser observada(s) subsiguientemente por el usuario, una o más tramas de audio adicionales del objeto de audio que es reproducido por el usuario que se predice que van a ser reproducidas subsiguientemente, uno o más artículos de contenido adicionales a los que hace referencia una página web que se predice que va(n) a ser observado(s) subsiguientemente, etc. Una indicación del contenido adicional predicho es incluida en la primera solicitud de contenido deseado 210 por la aplicación de cliente 204.

Obsérvese que la primera solicitud de contenido deseado 210 (y la respuesta o respuestas a la misma) se puede incluir en una o más conexiones de comunicación (por ejemplo, una conexión de TCP) entre la aplicación de cliente 204 y el servidor de almacenamiento en memoria caché 202. Se puede formar cualquier número de conexiones entre la aplicación de cliente 204 y el servidor de almacenamiento en memoria caché 202, y cada conexión puede

incluir una solicitud de contenido deseado y/o puede indicar uno o más artículos de contenido adicionales predichos. En algunos casos, el contenido deseado puede estar ya almacenado en memoria caché en el servidor de almacenamiento en memoria caché 202, mientras que el contenido futuro predicho no está aún almacenado en memoria caché. En tal caso, solo el contenido futuro predicho puede ser solicitado por el servidor de almacenamiento en memoria caché 202 a partir del servidor de origen (es decir, no es necesario que el contenido deseado, que está ya almacenado en memoria caché en el servidor de almacenamiento en memoria caché 202, se solicite a partir del servidor de origen).

Haciendo referencia de nuevo a la figura 3, en la etapa 304, se transmite una solicitud de al menos el contenido adicional indicado a un servidor de origen. Por ejemplo, tal como se muestra en la figura 2, el servidor de almacenamiento en memoria caché 202 puede transmitir una solicitud de servidor 212 que indica el contenido adicional que se indica en la primera solicitud de contenido deseado 210. La solicitud de servidor 212 es recibida por el servidor de origen 102. El servidor de almacenamiento en memoria caché 202 puede transmitir, de forma opcional, una solicitud al servidor de origen 102 para el contenido deseado que se indica en la primera solicitud de contenido deseado 210 si el contenido deseado aún no ha sido almacenado en memoria caché por el servidor de almacenamiento en memoria caché 202. En una forma de realización, la solicitud de servidor 212 puede incluir tanto una solicitud del contenido adicional predicho como una solicitud del contenido deseado.

En la etapa 306, se recibe un contenido adicional a partir del servidor de origen. Por ejemplo, en respuesta a la solicitud de servidor 212, el servidor de origen 102 puede transmitir una respuesta de servidor 214 que incluye el contenido deseado, si se solicita. Por ejemplo, el servidor de origen 102 puede acceder al contenido deseado en el almacenamiento asociado con el servidor de origen 102. El servidor de almacenamiento en memoria caché 202 recibe la respuesta de servidor 214 a partir del servidor de origen 102. Además, el servidor de origen 102 puede acceder al contenido adicional predicho en un almacenamiento asociado con el servidor de origen 102, y puede transmitir el contenido adicional predicho al servidor de almacenamiento en memoria caché 202 en la respuesta de servidor 214 o en una respuesta separada.

En la etapa 308, el contenido deseado se transmite a la aplicación de cliente en el dispositivo de usuario. Por ejemplo, tal como se muestra en la figura 2, el servidor de almacenamiento en memoria caché 202 puede transmitir el contenido deseado solicitado a la aplicación de cliente 204 en una primera respuesta de contenido deseado 216.

En la etapa 310, el contenido adicional se almacena en memoria caché. Por ejemplo, tal como se muestra en la figura 2, el servidor de almacenamiento en memoria caché 202 puede incluir el almacenamiento 222, y el contenido adicional predicho que se recibe a partir del servidor de origen 102 en la respuesta de servidor 214 se puede almacenar en memoria caché en el almacenamiento 222 como el contenido almacenado en memoria caché 224. El almacenamiento 222 puede incluir uno o más de cualquier tipo de mecanismo de almacenamiento para almacenar en memoria caché el contenido, incluyendo un disco magnético (por ejemplo, en una unidad de disco duro), un disco óptico (por ejemplo, en una unidad de disco óptico), una cinta magnética (por ejemplo, en una unidad de cinta), un dispositivo de memoria tal como un dispositivo de RAM (*random access memory*, memoria de acceso aleatorio), etc., y/o cualquier otro tipo adecuado de medio de almacenamiento. Además, en una forma de realización, el almacenamiento 222 se puede compartir entre múltiples servidores de almacenamiento en memoria caché.

De acuerdo con el diagrama de flujo 300, una aplicación de cliente ha informado a la CDN de una predicción de contenido adicional que se puede desear en el futuro, y el contenido adicional predicho se almacena previamente en memoria caché en la CDN (por ejemplo, en el servidor de almacenamiento en memoria caché 402 de la CDN). Este almacenamiento previo en memoria caché aumenta la relación de aciertos de memoria caché para un contenido que realmente es solicitado subsiguientemente, mejorando de ese modo la experiencia general del usuario final. Obsérvese que no necesariamente se usará todo el contenido almacenado previamente en memoria caché y, de forma opcional, unas solicitudes de almacenamiento previo en memoria caché se pueden priorizar por debajo de unas solicitudes de contenido de CDN normales. A continuación se describen adicionalmente unas formas de realización a modo de ejemplo de tal priorización. Un beneficio del almacenamiento previo en memoria caché es un mejor aprovechamiento de la capacidad de reserva, en donde exista la misma, para mejorar la experiencia general del usuario. El uso de una priorización (por ejemplo, indicadores de probabilidad) ayuda a optimizar el almacenamiento previo en memoria caché.

Por ejemplo, la figura 4 muestra un diagrama de flujo 400 que proporciona un procedimiento para proporcionar un contenido previamente almacenado en memoria caché a una aplicación de cliente, de acuerdo con una forma de realización a modo de ejemplo. Por ejemplo, en una forma de realización, el diagrama de flujo 400 puede seguir el diagrama de flujo 300 de la figura 3. El diagrama de flujo 400 se describe tal como sigue con respecto a la figura 2.

El diagrama de flujo 400 comienza con la etapa 402. En la etapa 402, se recibe una segunda solicitud a partir de la aplicación de cliente que incluye una solicitud del contenido adicional. Por ejemplo, tal como se muestra en la figura 2, el servidor de almacenamiento en memoria caché 202 puede recibir una segunda solicitud de contenido deseado 218 a partir de la aplicación de cliente 204 del dispositivo de usuario 106a. La segunda solicitud de contenido deseado 218 es una solicitud de un contenido para la aplicación de cliente 204 (por ejemplo, una tesela de mapa, imágenes, vídeo, audio, páginas web, etc.) que se predijo previamente que iba a ser solicitada potencialmente por la aplicación de cliente 204 en el futuro, y se indicó en la primera solicitud de contenido deseado 210 (u otra solicitud

de contenido deseado previa).

En la etapa 404, el contenido adicional almacenado en memoria caché se sirve a la aplicación de cliente en el dispositivo de usuario. En una forma de realización, el servidor de almacenamiento en memoria caché 202 puede analizar una solicitud de contenido deseado 218 para determinar si cualquier contenido que se solicite en la misma está ya almacenado en memoria caché en el servidor de almacenamiento en memoria caché 202 (por ejemplo, en el almacenamiento 222). Por ejemplo, en una forma de realización, el servidor de almacenamiento en memoria caché 202 puede acceder a un mapa de contenido almacenado en memoria caché o a otra estructura de datos que pone en correspondencia unos identificadores de contenido (por ejemplo, números de identificación, tales como valores de troceo, etc.) con un contenido almacenado en memoria caché en el almacenamiento 222. El servidor de almacenamiento en memoria caché 202 puede comparar un identificador de contenido recibido en la segunda solicitud de contenido deseado 218 para el contenido deseado con los identificadores de contenido en el mapa de contenido almacenado en memoria caché para determinar si el contenido deseado está ya almacenado en memoria caché. Si el identificador de contenido para la cantidad deseada se encuentra en el mapa de contenido almacenado en memoria caché, el contenido deseado está ya almacenado en memoria caché en el servidor de almacenamiento en memoria caché 202. En tal caso, el servidor de almacenamiento en memoria caché 202 puede acceder al contenido almacenado en memoria caché en el almacenamiento 222 (por ejemplo, como el contenido almacenado en memoria caché 224), y puede servir el contenido almacenado en memoria caché a la aplicación de cliente 204 en una transmisión de una respuesta de contenido almacenado en memoria caché 220.

En el caso de que el contenido deseado que se indica en la segunda solicitud de contenido deseado 218 no esté aún almacenado en memoria caché en el servidor de almacenamiento en memoria caché 202, el servidor de almacenamiento en memoria caché 202 puede solicitar el contenido deseado a partir del servidor de origen 102 tal como se ha descrito en lo que antecede con respecto a las etapas 304 y 306 del diagrama de flujo 300 (la figura 3). Además, la segunda solicitud de contenido deseado 218 puede incluir, de forma opcional, un contenido adicional más que la unidad de predicción de contenido adicional 208 prediga que potencialmente va a ser deseado subsiguientemente, en una forma similar a la de la primera solicitud de contenido deseado 210. En tal caso, el servidor de almacenamiento en memoria caché 202 puede solicitar el contenido adicional más que se indica a partir del servidor de origen 102, y puede almacenar en memoria caché el contenido adicional más que se indica en el almacenamiento 222 para las solicitudes de contenido subsiguientes, en una forma similar a como se ha descrito en lo que antecede con respecto a las etapas 304, 306 y 310 del diagrama de flujo 300.

En las siguientes subsecciones se describen algunas formas de realización a modo de ejemplo para aplicaciones de cliente y servidores de almacenamiento en memoria caché.

A. Formas de realización de aplicación de cliente a modo de ejemplo

Tal como se ha descrito en lo que antecede, en una forma de realización, una aplicación de cliente (por ejemplo, la aplicación de cliente 204) puede predecir el contenido al que se va a acceder en el futuro, y el contenido predicho se puede almacenar en memoria caché en los servidores de almacenamiento en memoria caché para una respuesta más rápida. Por ejemplo, la figura 5 muestra un diagrama de flujo 500 que proporciona un procedimiento para que una aplicación de cliente solicite y reciba un contenido deseado e indique el contenido que puede ser solicitado en el futuro para su almacenamiento en memoria caché, de acuerdo con una forma de realización a modo de ejemplo. En una forma de realización, el diagrama de flujo 500 puede ser llevado a cabo por la aplicación de cliente 204 de la figura 2. El diagrama de flujo 500 se describe con respecto a la figura 6. La figura 6 muestra un diagrama de bloques de un dispositivo de usuario 600 que incluye una forma de realización a modo de ejemplo de la aplicación de cliente 204. Tal como se muestra en la figura 6, la aplicación de cliente 204 incluye una interfaz de comunicación 602 y una unidad de predicción de contenido adicional 208, y la interfaz de comunicación 602 incluye un formateador de solicitudes 604. Algunas formas de realización estructurales y operativas adicionales les resultarán evidentes a los expertos en la materia o materias relevantes basándose en el siguiente análisis con respecto al diagrama de flujo 500 y la aplicación de cliente 204 de la figura 6.

El diagrama de flujo 500 comienza con la etapa 502. En la etapa 502, se genera una primera solicitud de contenido deseado que indica un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente. Por ejemplo, tal como se muestra en la figura 6, la aplicación de cliente 204 puede generar una primera solicitud de contenido deseado 210 que, tal como se ha descrito en lo que antecede, es una solicitud de contenido deseado. Además, tal como se ha descrito en lo que antecede, la primera solicitud de contenido deseado 210 incluye una indicación de un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente 204.

En una forma de realización, la unidad de predicción de contenido adicional 208 de la aplicación de cliente 204 predice un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente 204. Por ejemplo, tal como se muestra en la figura 6, la unidad de predicción de contenido adicional 204 puede recibir una indicación de contenido deseado 606. La indicación de contenido deseado 606 se puede generar internamente con respecto a la aplicación de cliente 204 y / o puede ser introducida por un usuario en una interfaz de usuario que es proporcionada por la aplicación de cliente 204. La indicación de contenido deseado 606 indica un contenido que se desea presentar o reproducir, o con el que se desea interactuar de otra forma. Por

ejemplo, la indicación de contenido deseado 606 puede indicar uno o más de una tesela de mapa, una trama de vídeo, una trama de audio, una imagen, una página web, etc. (por ejemplo, por medio de un URL, un nombre de archivo, etc.) Una unidad de predicción de contenido adicional 208 puede predecir un contenido adicional basándose en el contenido deseado que es indicado por la indicación de contenido deseado 606, y puede enviar el contenido adicional predicho como la indicación de contenido adicional predicho 608.

Por ejemplo, cuando la indicación de contenido deseado 606 indica una tesela de mapa de un mapa visto como un contenido deseado, la unidad de predicción de contenido adicional 208 puede predecir que una o más teselas de mapa adicionales del mapa van a ser vistas subsiguientemente por el usuario. En tal caso, la unidad de predicción de contenido adicional 208 puede generar la indicación de contenido adicional predicho 608 para indicar una o más teselas de mapa que son espacialmente adyacentes a la tesela de mapa indicada, que son teselas de mapa dentro de la tesela de mapa indicada (acercada), que son teselas de mapa que incluyen la tesela de mapa indicada (alejada), y / o que son teselas de mapa del mapa que es probable que vayan a ser vistas subsiguientemente por el usuario.

En otro ejemplo, cuando la indicación de contenido deseado 606 indica una imagen de una serie de imágenes como un contenido deseado, la unidad de predicción de contenido adicional 208 puede predecir que una o más imágenes adicionales de la serie de imágenes van a ser observadas subsiguientemente por el usuario, tal como una o más imágenes subsiguientes de la secuencia de imágenes (incluyendo la totalidad del resto de la secuencia de imágenes) que se encuentran adyacentes (temporal o espacialmente) a la imagen indicada, una imagen que es una porción de la imagen indicada (por ejemplo, una imagen acercada), una imagen que incluye la totalidad de la imagen indicada (por ejemplo, una imagen alejada), etc.

En otro ejemplo, cuando la indicación de contenido deseado 606 indica una trama de vídeo de un vídeo como un contenido deseado, la unidad de predicción de contenido adicional 208 puede predecir que una o más tramas de vídeo adicionales del vídeo van a ser observadas subsiguientemente por el usuario, tal como una o más tramas de vídeo subsiguientes de vídeo (incluyendo la totalidad del resto del vídeo).

En otro ejemplo más, cuando la indicación de contenido deseado 606 indica una trama de audio de un audio como un contenido deseado, la unidad de predicción de contenido adicional 208 puede predecir que una o más tramas de audio adicionales del audio van a ser reproducidas subsiguientemente por el usuario, tal como una o más tramas de audio subsiguientes del audio (incluyendo la totalidad del resto del audio).

En otro ejemplo más, cuando la indicación de contenido deseado 606 indica una página web como un contenido deseado, la unidad de predicción de contenido adicional 208 puede predecir que una o más páginas web adicionales u otros objetos web (por ejemplo, imágenes, vídeos, audio, etc.) van a ser observados subsiguientemente por el usuario, tal como una o más páginas web u otros objetos web enlazados dentro de la página web indicada, una o más páginas web de un sitio web que incluye la página web indicada, etc.

Tal como se muestra en la figura 6, la interfaz de comunicación 602 recibe una indicación de contenido deseado 606 y una indicación de contenido adicional predicho 608. En una forma de realización, el formateador de solicitudes 304 genera una o más solicitudes para solicitar el contenido deseado que es indicado por la indicación de contenido deseado 606 y el contenido adicional predicho indicado por la indicación de contenido adicional predicho 608. La interfaz de comunicación 602 está configurada para transmitir la solicitud o solicitudes generadas por el formateador de solicitudes 604 a partir de la aplicación de cliente 204 como la primera solicitud de contenido deseado 210. En algunas formas de realización, el formateador de solicitudes 604 puede generar la solicitud o solicitudes para que tengan cualquier formato, según se desee para una técnica de comunicación particular.

Por ejemplo, en una forma de realización, el formateador de solicitudes 604 puede generar la solicitud en la forma de un mensaje de solicitud de HTTP (*hypertext transfer protocol*, protocolo de transferencia de hipertexto). En tal forma de realización, el mensaje de solicitud de HTTP puede estar configurado para solicitar el contenido deseado, y el contenido adicional predicho se puede indicar en el mensaje de solicitud de HTTP de cualquier forma adecuada, tal como indicarse en un encabezado. Por ejemplo, a continuación se muestra una porción de un mensaje de solicitud de HTTP a modo de ejemplo:

```

GET/images/logo.png HTTP/1.1
...
HINTS: <URL1 >=20; <URL2>=60
...

```

En el presente ejemplo, el mensaje de solicitud de HTTP incluye una línea de solicitud que solicita un objeto web "/images/logo.png" a partir de un servidor como un contenido deseado. Además, el mensaje de solicitud de HTTP a modo de ejemplo incluye un encabezado de http de "HINTS" ("sugerencias") que indica dos URL - el "URL1" y el "URL2", como un contenido adicional predicho. En el presente ejemplo, cada uno del "URL1" y el "URL2" se puede reemplazar en el encabezado de "HINTS" con unos URL completos al contenido adicional predicho correspondiente (por ejemplo, <http://tv.msn.com/tv/article.aspx?news=xyz>, etc.) A pesar de que en el encabezado a modo de ejemplo anterior se muestran dos URL, se puede indicar cualquier número de artículos de contenido adicional predicho en un

encabezado de esta forma por medio de un URL u otro identificador.

Por lo tanto, en una forma de realización, el formateador de solicitudes 604 puede generar un mensaje de solicitud de HTTP que indica un contenido deseado (por ejemplo, usando una instrucción de "GET" ("obtener")), y que indica un contenido adicional predicho en un encabezado (por ejemplo, un tipo de encabezado de "HINTS" u otro tipo predeterminado para un contenido adicional predicho). En otras formas de realización, el contenido deseado y el contenido adicional predicho pueden indicarse en otras formas en una solicitud por formateador de solicitudes 604. Obsérvese que un encabezado de "HINTS" se puede encontrar presente en un mensaje de solicitud de HTTP tal como se ha mostrado en el ejemplo anterior, o se pueden encontrar presentes múltiples encabezados de "HINTS". Lo que es más, en una forma de realización, un mensaje de solicitud de HTTP puede indicar un contenido adicional predicho sin indicar contenido deseado alguno (es decir, solo se especifica / solicita un contenido adicional predicho). En tal mensaje de solicitud de HTTP, se puede usar una instrucción de "GET" para especificar el contenido adicional predicho, y puede que no se use un encabezado de "HINTS" o de otro tipo, o se puede encontrar presente un encabezado (por ejemplo, un encabezado de "CACHEONLY" ("solo memoria caché")) para indicar al servidor de almacenamiento en memoria caché que el contenido solicitado se ha de almacenar en memoria caché pero no se ha de devolver aún al cliente.

La interfaz de comunicación 602 puede transmitir primera solicitud de contenido deseado 210 a un servidor de almacenamiento en memoria caché. En algunas formas de realización, la interfaz de comunicación 602 puede estar configurada para transmitir la primera solicitud de contenido deseado 210 en un mensaje de HTTP, y / o se puede configurar para comunicar solicitudes en otras formas, tal como sería conocido por los expertos en la materia o materias relevantes.

Haciendo referencia de nuevo a la figura 5, en la etapa 504, se recibe el contenido deseado a partir del servidor de almacenamiento en memoria caché que recuperó el contenido deseado a partir de un servidor de origen en respuesta a la primera solicitud del contenido deseado. Por ejemplo, tal como se muestra en la figura 6, la interfaz de comunicación 602 puede recibir la primera respuesta de contenido deseado 216 a partir de un servidor de almacenamiento en memoria caché en respuesta a la primera solicitud de contenido deseado 210. La primera respuesta de contenido deseado 216 incluye el contenido deseado solicitado en la solicitud 210. La aplicación de cliente 204 puede presentar, reproducir y / o posibilitar de otra forma que un usuario en el dispositivo de usuario 600 interactúe con el contenido deseado recibido.

En la etapa 506, se genera una segunda solicitud del contenido adicional. Por ejemplo, tal como se muestra en la figura 6, la aplicación de cliente 204 puede generar una segunda solicitud de contenido deseado 218. Tal como se ha descrito en lo que antecede, la segunda solicitud de contenido deseado 218 es una solicitud de contenido para la aplicación de cliente 204 (por ejemplo, teselas de mapa, imágenes, vídeo, audio, páginas web, etc.) que previamente se predijo como que posiblemente fuera a ser solicitada subsiguientemente por la aplicación de cliente 204, y se indicó en ese sentido en la primera solicitud de contenido deseado 210. La segunda solicitud de contenido deseado 218 se puede generar de una forma similar a la de la primera solicitud de contenido deseado 210. Por ejemplo, la unidad de predicción de contenido adicional 208 puede recibir una segunda indicación de contenido deseado 606, que indica un contenido que se desea presentar o reproducir, o con el que se desea interactuar de otra forma. La unidad de predicción de contenido adicional 208 puede predecir, de forma opcional, un contenido adicional basándose en el contenido deseado que es indicado por un enlace o una indicación de contenido deseado 606, y puede enviar el contenido adicional predicho como la segunda indicación de contenido adicional predicho 608. La interfaz de comunicación 602 recibe la segunda indicación de contenido deseado 606 y, de forma opcional, recibe la segunda indicación de contenido adicional predicho 608. En una forma de realización, un formateador de solicitudes 604 genera una solicitud del segundo contenido deseado y, de forma opcional, incluye el segundo contenido adicional predicho en la solicitud generada. La interfaz de comunicación 602 se configura para transmitir la solicitud a partir de la aplicación de cliente 204 como una segunda solicitud de contenido deseado 218.

En la etapa 508, se recibe el contenido adicional a partir del servidor de almacenamiento en memoria caché que recuperó el contenido adicional a partir de un servidor de origen y almacenó en memoria caché el contenido adicional en respuesta a la primera solicitud del contenido deseado. Por ejemplo, tal como se muestra en la figura 6, la interfaz de comunicación 602 puede recibir la respuesta de contenido almacenada en memoria caché 220, que incluye un contenido almacenado en memoria caché a partir de un servidor de almacenamiento en memoria caché como el contenido deseado que se indica en la segunda solicitud de contenido deseado 218. El servidor de almacenamiento en memoria caché obtuvo previamente el contenido deseado a partir de un servidor de origen, y almacenó en memoria caché el contenido deseado en el caso de que se pudiera solicitar subsiguientemente.

Tal como se ha descrito en lo que antecede, se pueden priorizar las solicitudes de contenido para su almacenamiento en memoria caché, de forma opcional, por debajo de unas solicitudes de contenido de CDN normales. De esta forma, el contenido que se está solicitando realmente puede ser solicitado antes de que un contenido se solicite simplemente para fines de almacenamiento en memoria caché. Asimismo, el contenido que se está solicitando para fines de almacenamiento en memoria caché que es más probable que se solicite en realidad se puede almacenar en memoria caché antes que el contenido que se está solicitando para fines de almacenamiento en memoria caché que es menos probable que se solicite en realidad.

Por ejemplo, la figura 7 muestra un diagrama de bloques de una unidad de predicción de contenido adicional 208 que incluye una unidad de priorización de contenido adicional 702, de acuerdo con una forma de realización a modo de ejemplo. La unidad de priorización de contenido adicional 702 está configurada para analizar un contenido adicional predicho para determinar cómo de probable es que realmente sea solicitado subsiguientemente por la aplicación de cliente. Por ejemplo, en una forma de realización, la unidad de priorización de contenido adicional 702 puede operar de acuerdo con una etapa 802 que se muestra en la figura 8. En la etapa 802, se genera una indicación de probabilidad que indica una probabilidad de que el contenido adicional pueda ser solicitado subsiguientemente por la aplicación de cliente. En una forma de realización, para cada artículo de contenido adicional predicho, la unidad de priorización de contenido adicional 702 puede generar una indicación de probabilidad correspondiente que indica una probabilidad de que el contenido adicional pueda ser solicitado subsiguientemente por la aplicación de cliente 204. La indicación de probabilidad generada puede estar asociada con el artículo de contenido adicional predicho correspondiente, e incluirse en la solicitud de contenido deseado transmitida al servidor de almacenamiento en memoria caché. El servidor de almacenamiento en memoria caché puede usar la indicación de probabilidad para priorizar el almacenamiento en memoria caché del artículo de contenido adicional predicho en relación con otro contenido. Además, el servidor de almacenamiento en memoria caché puede transmitir la indicación de probabilidad al servidor de origen cuando se solicita el contenido predicho de tal modo que el servidor de origen pueda priorizar la provisión del contenido predicho al servidor de almacenamiento en memoria caché.

La unidad de priorización de contenido adicional 702 puede generar indicaciones de probabilidad de cualquier forma, incluyendo basándose en el contenido que se está observando y / o en el comportamiento de usuario. Por ejemplo, se pueden generar indicaciones de probabilidad de forma diferente para diferentes tipos de contenido (por ejemplo, asignando unos valores de indicación de probabilidad más altos a las tramas de vídeo en relación con las teselas de mapa, etc.) En otra forma de realización, las indicaciones de probabilidad pueden ser generadas por la unidad de priorización de contenido adicional 702, para un contenido adicional predicho, de forma diferente basándose en la proximidad (por ejemplo, en el tiempo y / o en el espacio) con respecto al contenido que se solicitó realmente. Por ejemplo, en un ejemplo de mapa, a las teselas de mapa que están inmediatamente adyacentes a una tesela de mapa que se observa en la actualidad se les pueden asignar unas indicaciones de probabilidad que son de un valor mayor en relación con las teselas de mapa que están separadas de la tesela de mapa observada por una o más teselas de mapa intermedias. A las tramas de vídeo inmediatamente siguientes de una secuencia de vídeo se les pueden asignar unas indicaciones de probabilidad que tienen unos números relativamente altos en relación con las tramas de vídeo posteriores de la secuencia de vídeo (por ejemplo, unas tramas de vídeo que son posteriores en el tiempo y la colocación en la secuencia de vídeo). Las imágenes de una secuencia de imagen y / o tramas de audio en una secuencia de audio se pueden tratar de forma similar y / o diferente, dependiendo de la puesta en práctica real. En una página web, a los artículos de contenido (por ejemplo, el URL) que están localizados cerca de la parte superior de la página web se les pueden asignar unos valores más altos de indicaciones de probabilidad en relación con los artículos de contenido que están localizados cerca de la parte inferior de la página web. En algunas formas de realización, las indicaciones de probabilidad se pueden generar de varias formas, y pueden tener cualquier valor y rango de valores (por ejemplo, rangos numéricos, rangos textuales (por ejemplo, "alto", "medio", "bajo", etc.)) que sea adecuado, según se desee para unas puestas en práctica particulares. Con respecto al comportamiento de usuario, cuando un usuario está dando un vistazo panorámico en una dirección particular en un mapa, la probabilidad de que una tesela de mapa en esa dirección se solicite en el futuro es más alta que una probabilidad de otras teselas y, en ese sentido, se le pueden asignar unas indicaciones de probabilidad mayores. Cuando un usuario hace avanzar rápidamente un vídeo o audio, la trama "siguiente" se puede encontrar múltiples tramas hacia adelante y, en ese sentido, a las tramas que se encuentran múltiples tramas hacia adelante en la secuencia se les pueden asignar unas indicaciones de probabilidad que tienen unos valores relativamente altos en relación con las tramas más inmediatas.

Por ejemplo, en el mensaje de HTTP a modo de ejemplo que se ha mostrado en lo que antecede, la indicación de probabilidad que tiene un valor de 20 pudo haber sido generada para el URL1, y una indicación de probabilidad que tiene un valor de 60 pudo haber sido generada para el URL2. En tal ejemplo, las indicaciones de probabilidad pueden tener unos valores en el rango de 1 - 100, significando unos valores de probabilidad más bajos que el artículo de contenido adicional predicho correspondiente tiene una probabilidad más baja de ser solicitado realmente por la aplicación de cliente en el futuro, y significando unos valores de probabilidad más altos que el artículo de contenido adicional predicho correspondiente tiene una probabilidad mayor de ser solicitado realmente por la aplicación de cliente en el futuro. En ese sentido, en el presente ejemplo, el URL2 tiene una probabilidad más alta de ser solicitado por la aplicación de cliente en el futuro en relación con el URL1 y, por lo tanto, se puede almacenar en memoria caché antes que el URL1.

B. Formas de realización de servidor de almacenamiento en memoria caché a modo de ejemplo

La figura 9 muestra un diagrama de bloques de un servidor de almacenamiento en memoria caché 902, de acuerdo con una forma de realización a modo de ejemplo. El servidor de almacenamiento en memoria caché 902 es un ejemplo del servidor de almacenamiento en memoria caché 202 que se muestra en la figura 2. Tal como se muestra en la figura 2, el servidor de almacenamiento en memoria caché 902 incluye un módulo de recuperación de contenido 904, un analizador de solicitudes de contenido 906, un módulo de provisión de contenido 908 y un almacenamiento 222. El servidor de almacenamiento en memoria caché 902 puede almacenar previamente en

memoria caché un contenido futuro predicho de varias formas. Por ejemplo, el servidor de almacenamiento en memoria caché 902 puede llevar a cabo los diagramas de flujo 300 y 400, en algunas formas de realización.

Por ejemplo, el analizador de solicitudes de contenido 906 puede llevar a cabo la etapa 302 del diagrama de flujo 300, en la que se recibe una solicitud de contenido deseado a partir de una aplicación de cliente en un dispositivo de usuario, indicando la solicitud un contenido adicional en relación con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente. El analizador de solicitudes de contenido 906 puede recibir y analizar la primera solicitud de contenido deseado 210 para identificar cualquier contenido deseado y cualquier contenido adicional predicho. El analizador de solicitudes de contenido 906 transmite un primer contenido deseado identificado y un contenido adicional predicho 912 al módulo de recuperación de contenido 904.

El módulo de recuperación de contenido 904 puede llevar a cabo la etapa 304 del diagrama de flujo 300, en la que se transmite una solicitud del contenido adicional indicado a un servidor de origen. Tal como se muestra en la figura 9, el módulo de recuperación de contenido 904 puede transmitir la solicitud de servidor 212 al servidor de origen, que indica el contenido adicional predicho, y puede indicar el contenido deseado (si no está todavía almacenado en memoria caché). El módulo de recuperación de contenido 904 puede incluir tanto la solicitud del contenido adicional predicho como la solicitud del contenido deseado en la solicitud de servidor 212, o los pueden transmitir en comunicaciones separadas.

El módulo de recuperación de contenido 904 puede llevar a cabo la etapa 306 del diagrama de flujo 300, en la que el contenido deseado y el contenido adicional se reciben a partir del servidor de origen. El módulo de recuperación de contenido 904 puede recibir la respuesta de servidor 214 a partir del servidor de origen que incluye el contenido deseado (si se solicita) y el contenido adicional predicho, o puede recibir el contenido deseado y el contenido adicional predicho en comunicaciones separadas.

El módulo de provisión de contenido 908 puede llevar a cabo la etapa 308 del diagrama de flujo 300, en la que el contenido deseado se transmite a la aplicación de cliente en el dispositivo de usuario. Tal como se muestra en la figura 9, el módulo de recuperación de contenido 904 puede transmitir el contenido deseado recuperado 914, que incluye el contenido deseado recibido a partir del servidor de origen en la respuesta de servidor 214. El módulo de provisión de contenido 908 puede recibir el contenido deseado recuperado 914, y puede transmitir el contenido deseado a la aplicación de cliente en la primera respuesta de contenido deseado 216.

El módulo de recuperación de contenido 904 puede llevar a cabo la etapa 310 del diagrama de flujo 300, en la que se almacena en memoria caché el contenido adicional. Tal como se muestra en la figura 9, el módulo de recuperación de contenido 904 puede almacenar en memoria caché el contenido adicional predicho en el almacenamiento 222 como el contenido almacenado en memoria caché 224.

El analizador de solicitudes de contenido 906 puede llevar a cabo la etapa 402 del diagrama de flujo 400 (la figura 4), en la que se recibe una segunda solicitud a partir de la aplicación de cliente que incluye una solicitud del contenido adicional. Tal como se muestra en la figura 9, el analizador de solicitudes de contenido 906 puede recibir una solicitud de contenido deseado 218 a partir de una aplicación de cliente. El analizador de solicitudes de contenido 906 puede recibir y analizar una segunda solicitud de contenido deseado 218 para identificar cualquier contenido deseado y cualquier contenido adicional predicho. El analizador de solicitudes de contenido 906 transmite un segundo contenido deseado identificado y el contenido adicional predicho 916 al módulo de recuperación de contenido 904. En el ejemplo de la figura 9, el segundo contenido deseado identificado y el contenido adicional predicho 916 incluyen el contenido adicional predicho del primer contenido deseado identificado y el contenido adicional predicho 912 como un contenido deseado.

El módulo de recuperación de contenido 904 y el módulo de provisión de contenido 908 pueden llevar a cabo la etapa 404 del diagrama de flujo 400, en la que se sirve el contenido adicional almacenado en memoria caché a la aplicación de cliente en el dispositivo de usuario. El módulo de recuperación de contenido 904 puede analizar el segundo contenido deseado identificado y el contenido adicional predicho 916 para determinar si cualquier contenido solicitado en el mismo está ya almacenado en memoria caché en el almacenamiento 222. Por ejemplo, en una forma de realización, el módulo de recuperación de contenido 904 puede acceder al mapa de contenido almacenado en memoria caché o a otra estructura de datos que se ha descrito en lo que antecede que pone en correspondencia unos identificadores de contenido con un contenido almacenado en memoria caché en el almacenamiento 222. Si el módulo de recuperación de contenido 904 determina que el contenido deseado está ya almacenado en memoria caché en el almacenamiento 222, el módulo de recuperación de contenido 904 puede acceder al contenido almacenado en memoria caché en el almacenamiento 222 (por ejemplo, como el contenido almacenado en memoria caché 224), y puede proporcionar el contenido almacenado en memoria caché al módulo de provisión de contenido 908 como el contenido almacenado en memoria caché 918. El módulo de provisión de contenido 908 puede servir el contenido almacenado en memoria caché a la aplicación de cliente en la respuesta de contenido almacenada en memoria caché 220.

En el caso de que el contenido deseado que se indica en el segundo contenido deseado identificado y un contenido adicional predicho 916 no estén aún almacenados en memoria caché en el almacenamiento 222, el módulo de recuperación de contenido 904 puede solicitar el contenido deseado a partir del servidor de origen tal como se ha

descrito en lo que antecede. Además, el segundo contenido deseado identificado y un contenido adicional predicho 916 pueden indicar, de forma opcional, un contenido adicional más que se predice que se va a solicitar subsiguientemente. En tal caso, el módulo de recuperación de contenido 904 puede solicitar el contenido adicional más que se indica a partir del servidor de origen, y puede almacenar en memoria caché el contenido adicional más que se indica en el almacenamiento 222 para solicitudes de contenido subsiguientes tal como se ha descrito en lo que antecede.

Obsérvese que el módulo de recuperación de contenido 904, el analizador de solicitudes de contenido 906 y el módulo de provisión de contenido 908 se pueden configurar para generar solicitudes y recibir respuestas de acuerdo con cualquier protocolo y formatos de comunicación adecuados, incluyendo mensajes de HTTP, etc., tal como sería conocido por los expertos en la materia o materias relevantes.

Además, tal como se ha descrito en lo que antecede, de forma opcional se pueden priorizar las solicitudes de contenido para su almacenamiento en memoria caché en relación con las solicitudes de contenido de CDN normales. Por ejemplo, la figura 10 muestra un diagrama de bloques del módulo de recuperación de contenido 904 de la figura 9 incluyendo una unidad de priorización de solicitudes 1002, de acuerdo con una forma de realización a modo de ejemplo. La unidad de priorización de solicitudes 1002 está configurada para priorizar las solicitudes de contenido para su almacenamiento en memoria caché. Por ejemplo, en una forma de realización, la unidad de priorización de solicitudes 1002 puede operar de acuerdo con una etapa 1102 que se muestra en la figura 11. En la etapa 1102, la transmisión de una solicitud del contenido adicional indicado al servidor de origen se prioriza basándose en una indicación de probabilidad recibida en la solicitud. Por ejemplo, en una forma de realización, para cada artículo de contenido adicional predicho, la unidad de priorización de solicitudes 1002 puede priorizar la solicitud del artículo de contenido adicional predicho en relación con otro contenido basándose en la indicación de probabilidad correspondiente que se recibe a partir de la aplicación de cliente 204.

Por ejemplo, en una forma de realización, la unidad de priorización de solicitudes 1002 puede generar y mantener una lista de prioridad u otra estructura de datos que enumera un contenido adicional predicho (por ejemplo, por identificador) y las indicaciones de probabilidad correspondientes. La unidad de priorización de solicitudes 1002 puede ordenar la lista por las indicaciones de probabilidad, o puede estructurar de otra forma la lista, de tal modo que el módulo de recuperación de contenido 904 solicite un contenido adicional predicho de acuerdo con las indicaciones de probabilidad de una forma priorizada. El contenido adicional predicho y las indicaciones de probabilidad pueden ser enumerados para una aplicación de contenido de cliente individual o para múltiples aplicaciones de cliente. Por lo tanto, de esta manera, el módulo de recuperación de contenido 904 puede priorizar las solicitudes a un servidor de origen para una aplicación de cliente individual o para múltiples aplicaciones de cliente.

Por ejemplo, con respecto al mensaje de HTTP a modo de ejemplo que se ha mostrado en lo que antecede, la unidad de priorización de solicitudes 1002 puede mantener (por ejemplo, almacenar y actualizar según sea necesario) una lista de contenido adicional predicho que incluye el URL1 y el URL2, y sus indicaciones de probabilidad correspondientes. Debido a que el URL2 tiene un valor de indicación de probabilidad de 60, el URL2 se prioriza más alto que el URL1, que tiene un valor de indicación de probabilidad de 20. En ese sentido, en el presente ejemplo, el módulo de recuperación de contenido 904 recupera el contenido del URL2 a partir de un servidor de origen antes de recuperar el contenido del URL1 a partir del servidor de origen, de acuerdo con la prioridad que es mantenida por la unidad de priorización de solicitudes 1002.

Además, en algunas formas de realización, el módulo de recuperación de contenido 904 puede usar una información adicional y / o alternativa para priorizar las solicitudes de contenido, tales como la capacidad de almacenamiento disponible, la capacidad de red disponible, la capacidad de procesamiento disponible (por ejemplo, de CPU), etc. Lo que es más, en una forma de realización, el servidor de almacenamiento en memoria caché puede transmitir la indicación de probabilidad al servidor de origen cuando se solicita el contenido futuro predicho de tal modo que el servidor de origen pueda priorizar la transmisión del contenido predicho al servidor de almacenamiento en memoria caché en relación con otro contenido (por ejemplo, priorizar una solicitud del contenido futuro predicho por debajo de unas solicitudes de contenido inmediatamente deseado). En ese sentido, en una forma de realización, un servidor de origen puede incluir una unidad de priorización de solicitudes similar a la unidad de priorización de solicitudes 1002 que prioriza la transmisión de contenido predicho a los servidores de almacenamiento en memoria caché basándose en indicaciones de probabilidad y / o en información adicional / alternativa.

III. Formas de realización de dispositivo informático a modo de ejemplo

Las aplicaciones de cliente 110a - 110d, la aplicación de cliente 204, la unidad de predicción de contenido adicional 208, la interfaz de comunicación 602, el formateador de solicitudes 604, la unidad de priorización de contenido adicional 702, el módulo de recuperación de contenido 904, el analizador de solicitudes de contenido 906, el módulo de provisión de contenido 908, la unidad de priorización de solicitudes 1002, el diagrama de flujo 300, el diagrama de flujo 400, el diagrama de flujo 500, la etapa 802 y la etapa 1102 se pueden poner en práctica en soporte físico, soporte lógico, soporte lógico inalterable, o cualquier combinación de los mismos. Por ejemplo, las aplicaciones de cliente 110a - 110d, la aplicación de cliente 204, la unidad de predicción de contenido adicional 208, la interfaz de comunicación 602, el formateador de solicitudes 604, la unidad de priorización de contenido adicional 702, el módulo

de recuperación de contenido 904, el analizador de solicitudes de contenido 906, el módulo de provisión de contenido 908, la unidad de priorización de solicitudes 1002, el diagrama de flujo 300, el diagrama de flujo 400, el diagrama de flujo 500, la etapa 802 y/o la etapa 1102 se pueden poner en práctica como un código / instrucciones / lógica de programa informático configurados para ejecutarse en uno o más procesadores.

5 Como alternativa, las aplicaciones de cliente 110a - 110d, la aplicación de cliente 204, la unidad de predicción de contenido adicional 208, la interfaz de comunicación 602, el formateador de solicitudes 604, la unidad de priorización de contenido adicional 702, el módulo de recuperación de contenido 904, el analizador de solicitudes de contenido 906, el módulo de provisión de contenido 908, la unidad de priorización de solicitudes 1002, el diagrama de flujo 300, el diagrama de flujo 400, el diagrama de flujo 500, la etapa 802 y/o la etapa 1102 se pueden poner en práctica como sistema de circuitos eléctricos / lógica de soporte físico. Por ejemplo, en una forma de realización, una o más de las aplicaciones de cliente 110a - 110d, la aplicación de cliente 204, la unidad de predicción de contenido adicional 208, la interfaz de comunicación 602, el formateador de solicitudes 604, la unidad de priorización de contenido adicional 702, el módulo de recuperación de contenido 904, el analizador de solicitudes de contenido 906, el módulo de provisión de contenido 908, la unidad de priorización de solicitudes 1002, el diagrama de flujo 300, el diagrama de flujo 400, el diagrama de flujo 500, la etapa 802 y/o la etapa 1102 se pueden poner en práctica juntos en un sistema en chip (SoC, *system-on-chip*). El SoC puede incluir un chip de circuito integrado que incluye uno o más de un procesador (por ejemplo, un microcontrolador, un microprocesador, un procesador de señales digitales (DSP, *digital signal processor*), etc.), una memoria, una o más interfaces de comunicación y / o circuitos y / o soporte lógico inalterable incorporado adicionales para llevar a cabo sus funciones.

20 La figura 12 muestra una puesta en práctica a modo de ejemplo de un ordenador 1200 en el que se pueden poner en práctica las formas de realización de la presente invención. Por ejemplo, cada uno del servidor de origen 102, los servidores de almacenamiento en memoria caché 104a y 104b, los dispositivos de usuario 106a y 106d, el servidor de almacenamiento en memoria caché 202, el dispositivo de usuario 600 y el servidor de almacenamiento en memoria caché 902 se pueden poner en práctica en uno o más sistemas informáticos similares al ordenador 1200, incluyendo una o más características del ordenador 1200 y/o características alternativas. El ordenador 1200 puede ser un dispositivo informático de propósito general en la forma de un ordenador personal convencional, un ordenador móvil, un servidor o una estación de trabajo, por ejemplo, un ordenador 1200 puede ser un dispositivo informático de propósito especial. La descripción del ordenador 1200 que se proporciona en el presente documento se proporciona para fines de ilustración, y no se tiene por objeto que sea limitante. Las formas de realización de la presente invención se pueden poner en práctica en tipos adicionales de sistemas informáticos, tal como sería conocido por los expertos en la materia o materias relevantes.

Tal como se muestra en la figura 12, el ordenador 1200 incluye uno o más procesadores 1202, una memoria de sistema 1204 y un bus 1206 que acopla varios componentes de sistema incluyendo la memoria de sistema 1204 al procesador 1202. El bus 1206 representa uno o más de cualquiera de varios tipos de estructuras de bus, incluyendo un bus de memoria o controlador de memoria, un bus de periféricos, un puerto de gráficos acelerados y un bus de procesador o local que usa cualquiera de una variedad de arquitecturas de bus. La memoria de sistema 1204 incluye una memoria de solo lectura (ROM, *read only memory*) 1208 y una memoria de acceso aleatorio (RAM, *random access memory*) 1210. Un sistema de entrada / salida básico (BIOS, *basic input / output system*) 1212 está almacenado en la ROM 1208.

40 El ordenador 1200 también tiene una o más de las siguientes unidades: una unidad de disco duro 1214 para leer de y escribir en un disco duro, una unidad de disco magnético 1216 para leer de o escribir en un disco magnético extraíble 1218 y una unidad de disco óptico 1220 para leer de o escribir en un disco óptico extraíble 1222 tal como un CD-ROM, un DVD-ROM u otros medios ópticos. La unidad de disco duro 1214, la unidad de disco magnético 1216 y la unidad de disco óptico 1220 están conectadas al bus 1206 mediante una interfaz de unidad de disco duro 1224, una interfaz de unidad de disco magnético 1226 y una interfaz de unidad óptica 1228, de forma respectiva. Las unidades y sus medios legibles por ordenador asociados proporcionan un almacenamiento no volátil de instrucciones legibles por ordenador, estructuras de datos, módulos de programa y otros datos para el ordenador. A pesar de que se describen un disco duro, un disco magnético extraíble y un disco óptico extraíble, se pueden usar otros tipos de medios de almacenamiento legibles por ordenador para almacenar datos, tales como tarjetas de memoria flash, discos de vídeo digital, memorias de acceso aleatorio (RAM), memorias de solo lectura (ROM), y similares.

Se puede almacenar un número de módulos de programa en el disco duro, el disco magnético, el disco óptico, la ROM o la RAM. Estos programas incluyen un sistema operativo 1230, uno o más programas de aplicación 1232, otros módulos de programa 1234 y unos datos de programa 1236. Los programas de aplicación 1232 o los módulos de programa 1234 pueden incluir, por ejemplo, una lógica de programa informático (por ejemplo, un código de programa informático) para poner en práctica las aplicaciones de cliente 110a - 110d, la aplicación de cliente 204, la unidad de predicción de contenido adicional 208, la interfaz de comunicación 602, el formateador de solicitudes 604, la unidad de priorización de contenido adicional 702, el módulo de recuperación de contenido 904, el analizador de solicitudes de contenido 906, el módulo de provisión de contenido 908, la unidad de priorización de solicitudes 1002, el diagrama de flujo 300, el diagrama de flujo 400, el diagrama de flujo 500, la etapa 802 y/o la etapa 1102 (incluyendo cualquier etapa de los diagramas de flujo 300, 400 y 500), y / o formas de realización adicionales que se describen en el presente documento.

Un usuario puede introducir órdenes e información en el ordenador 1200 a través de dispositivos de entrada tales como el teclado 1238 y el dispositivo de señalamiento 1240. Otros dispositivos de entrada (que no se muestran) pueden incluir un micrófono, una palanca de mando, una almohadilla para juegos, una antena parabólica, un escáner, o similares. Estos y otros dispositivos de entrada están a menudo conectados al procesador 1202 a través de una interfaz de puerto serie 1242 que está acoplada al bus 1206, pero se pueden conectar por medio de otras interfaces, tales como un puerto paralelo, un puerto de juegos o un bus serie universal (USB, *universal serial bus*).

Un dispositivo de presentación 1244 también está conectado al bus 1206 a través de una interfaz, tal como un adaptador de vídeo 1246. Además del monitor, el ordenador 1200 puede incluir otros dispositivos de salida periféricos (que no se muestran) tales como altavoces e impresoras.

El ordenador 1200 está conectado a una red 1248 (por ejemplo, Internet) a través de un adaptador o interfaz de red 1250, un módem 1252, u otros medios para establecer comunicaciones a través de la red. El módem 1252, que puede ser interno o externo, puede estar conectado al bus 1206 a través de la interfaz de puerto serie 1242, tal como se muestra en la figura 12, o se puede conectar al bus 1206 usando otro tipo de interfaz, incluyendo una interfaz paralela.

Tal como se usan en el presente documento, las expresiones “medio de programa informático”, “medio legible por ordenador” y “medio de almacenamiento legible por ordenador” se usan para hacer referencia generalmente a medios tales como el disco duro asociado con la unidad de disco duro 1214, el disco magnético extraíble 1218, el disco óptico extraíble 1222, así como otros medios tales como tarjetas de memoria flash, discos de vídeo digital, memorias de acceso aleatorio (RAM), memorias de solo lectura (ROM), y similares. Tales medios de almacenamiento por ordenador se distinguen de y no se solapan con medios de comunicación (no incluyen medios de comunicación). Por lo general, los medios de comunicación representan unas instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada tal como una onda portadora. La expresión “señal de datos modulada” significa una señal que tiene una o más de sus características establecidas o cambiadas de una forma tal que se codifique la información en la señal. A modo de ejemplo, y no de limitación, los medios de comunicación incluyen medios inalámbricos tales como acústicos, de RF, infrarrojos y otros medios inalámbricos. Las formas de realización también están dirigidas a tales medios de comunicación.

Tal como se ha hecho notar en lo que antecede, los programas y módulos informáticos (incluyendo programas de aplicación 1232 y otros módulos de programa 1234) se pueden almacenar en el disco duro, el disco magnético, el disco óptico, la ROM o la RAM. Tales programas informáticos también se pueden recibir a través de la interfaz de red 1250, la interfaz de puerto serie 1242 o cualquier tipo de interfaz. Tales programas informáticos, cuando son ejecutados o cargados por una aplicación, posibilitan que el ordenador 1200 ponga en práctica unas características de las formas de realización de la presente invención que se analizan en el presente documento. Por consiguiente, tales programas informáticos representan controladores del ordenador 1200.

La invención también está dirigida a productos de programa informático que comprenden soporte lógico almacenado en cualquier medio utilizable por ordenador. Tal soporte lógico, cuando se ejecuta en uno o más dispositivos de procesamiento de datos, da lugar a que un dispositivo o dispositivos de procesamiento de datos operen tal como se describe en el presente documento. Las formas de realización de la presente invención emplean cualquier medio utilizable por ordenador o legible por ordenador, conocido en la actualidad o en el futuro. Los ejemplos de los medios legibles por ordenador incluyen, pero no se limitan a, dispositivos de almacenamiento tales como RAM, unidades de disco duro, discos flexibles, CD-ROM, DVD-ROM, discos ZIP, cintas, dispositivos de almacenamiento magnético, dispositivos de almacenamiento óptico, MEM, dispositivos de almacenamiento basados en nanotecnología, y similares.

IV. Conclusión

A pesar de que en lo que antecede se han descrito varias formas de realización de la presente invención, se ha de entender que las mismas se han presentado solo a modo de ejemplo y no de limitación. Será entendido por los expertos en la materia o materias relevantes que en la misma se pueden hacer varios cambios en la forma y en los detalles sin apartarse del alcance de la invención tal como se define en las reivindicaciones adjuntas. Por consiguiente, el alcance de la presente invención se define únicamente de acuerdo con las siguientes reivindicaciones.

REIVINDICACIONES

1. Un procedimiento en un servidor de almacenamiento en memoria caché, que comprende:

5 recibir (302) una solicitud de un contenido deseado a partir de una aplicación de cliente en un dispositivo de usuario, incluyendo la solicitud una indicación de un contenido adicional relacionado con un contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente, en el que la indicación de un contenido adicional incluye una indicación de probabilidad que indica una probabilidad de que el contenido adicional pueda ser solicitado subsiguientemente por la aplicación de cliente, en el que el contenido deseado incluye una primera imagen y el contenido adicional es una segunda imagen que se encuentra espacialmente adyacente a la primera imagen, es adyacente a la primera imagen en el tiempo, es una porción de la primera imagen o incluye la primera imagen;
 10 transmitir (304) una solicitud de al menos el contenido adicional indicado a un servidor de origen, en el que dicha transmisión de la solicitud de al menos el contenido adicional indicado al servidor de origen comprende priorizar la transmisión de la solicitud del contenido adicional indicado al servidor de origen basándose en la indicación de probabilidad;
 15 recibir (306) el contenido adicional desde el servidor de origen;
 transmitir (308) el contenido deseado a la aplicación de cliente en el dispositivo de usuario; y
 almacenar en memoria caché (310) el contenido adicional.

2. El procedimiento de la reivindicación 1, que comprende adicionalmente:

20 recibir una segunda solicitud desde la aplicación de cliente que incluye una solicitud del contenido adicional; y servir el contenido adicional almacenado en memoria caché a la aplicación de cliente en el dispositivo de usuario.

3. El procedimiento de la reivindicación 2, en el que la segunda solicitud indica un segundo contenido adicional relacionado con el primer contenido adicional que puede ser solicitado subsiguientemente por la aplicación de cliente, comprendiendo adicionalmente el procedimiento:

25 transmitir una solicitud del segundo contenido adicional indicado al servidor de origen;
 recibir el segundo contenido adicional desde el servidor de origen; y
 almacenar en memoria caché el segundo contenido adicional.

4. El procedimiento de la reivindicación 1, en el que dicha transmisión de la solicitud de al menos el contenido adicional indicado al servidor de origen comprende:

30 priorizar la transmisión de la solicitud del contenido adicional indicado al servidor de origen basándose en la indicación de probabilidad y al menos una de una capacidad de almacenamiento disponible, una capacidad de red disponible o una capacidad de procesamiento disponible.

5. El procedimiento de la reivindicación 1, en el que la solicitud de un contenido deseado recibido desde la aplicación de cliente es un mensaje de HTTP, protocolo de transferencia de hipertexto, que indica el contenido adicional en un encabezado.

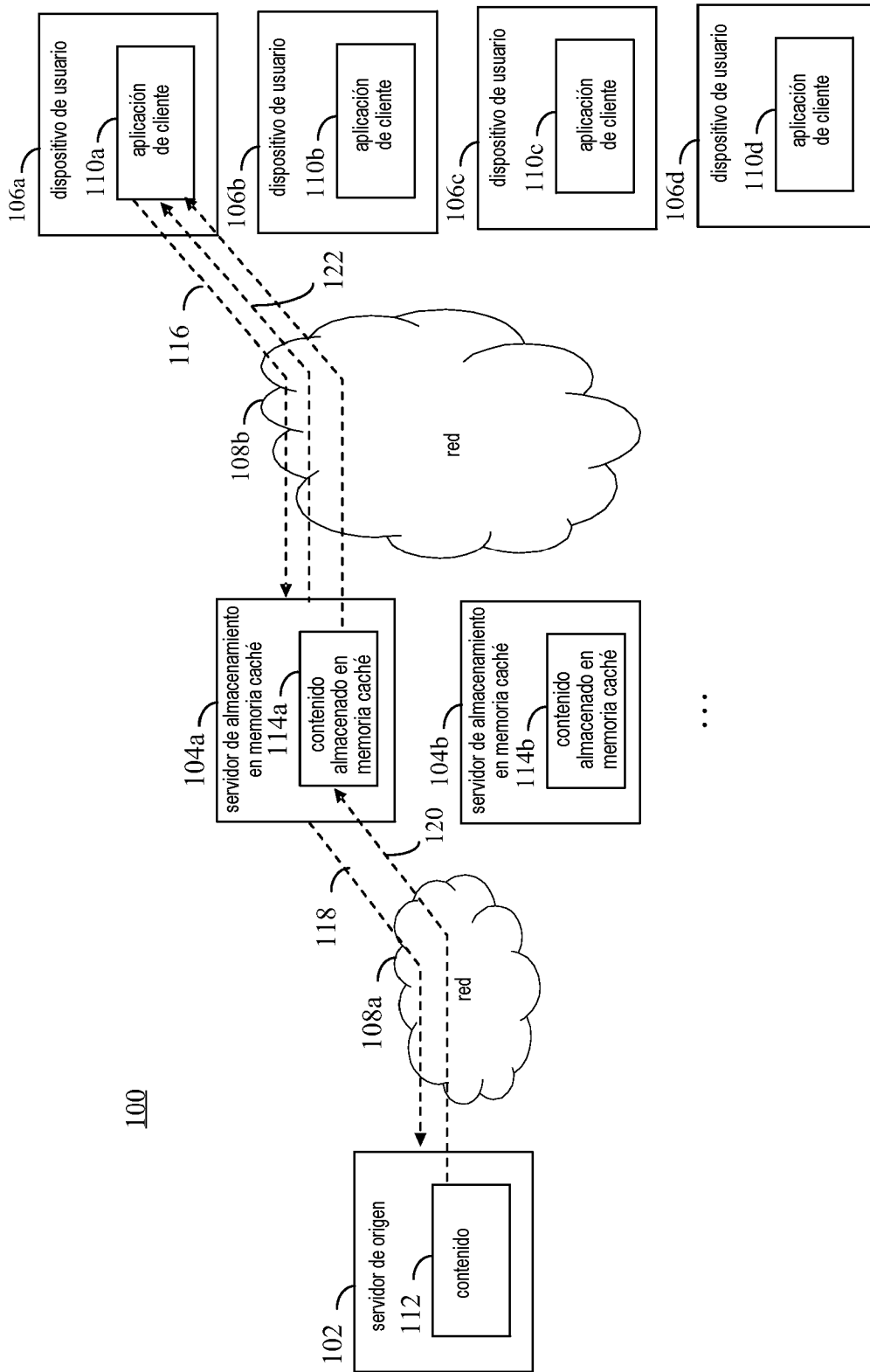
35 6. Un servidor de almacenamiento en memoria caché, que comprende:

un analizador de solicitudes de contenido (906) que está configurado para recibir una solicitud de un contenido deseado desde una aplicación de cliente en un dispositivo de usuario, incluyendo la solicitud una indicación de un contenido adicional relacionado con el contenido deseado que puede ser solicitado subsiguientemente por la aplicación de cliente, en el que la indicación de un contenido adicional incluye una indicación de probabilidad que indica una probabilidad de que el contenido adicional pueda ser solicitado subsiguientemente por la aplicación de cliente, en el que el contenido deseado incluye una primera imagen y el contenido adicional es una segunda imagen que se encuentra espacialmente adyacente a la primera imagen, es adyacente a la primera imagen en el tiempo, es una porción de la primera imagen o incluye la primera imagen;
 40 un módulo de recuperación de contenido (904) que está configurado para transmitir una solicitud de al menos el contenido adicional indicado a un servidor de origen, para recibir el contenido adicional desde el servidor de origen y para almacenar en memoria caché el contenido adicional en almacenamiento, en el que dicha transmisión de la solicitud de al menos el contenido adicional indicado al servidor de origen comprende priorizar la transmisión de la solicitud del contenido adicional indicado al servidor de origen basándose en la indicación de probabilidad; y
 45 un módulo de provisión de contenido (908) que está configurado para transmitir el contenido deseado a la aplicación de cliente en el dispositivo de usuario.

7. El servidor de almacenamiento en memoria caché de la reivindicación 6, en el que el analizador de solicitudes de contenido está configurado para recibir una segunda solicitud desde la aplicación de cliente que incluye una solicitud del contenido adicional, y el módulo de provisión de contenido está configurado para servir el contenido adicional almacenado en memoria caché a la aplicación de cliente.

8. Un producto de programa informático que comprende un medio legible por ordenador que tiene una lógica de

programa informático registrada en el mismo, que comprende:
unos medios de lógica de programa informático para posibilitar que un procesador lleve a cabo cualquiera de las reivindicaciones 1 - 5.



100

FIG. 1

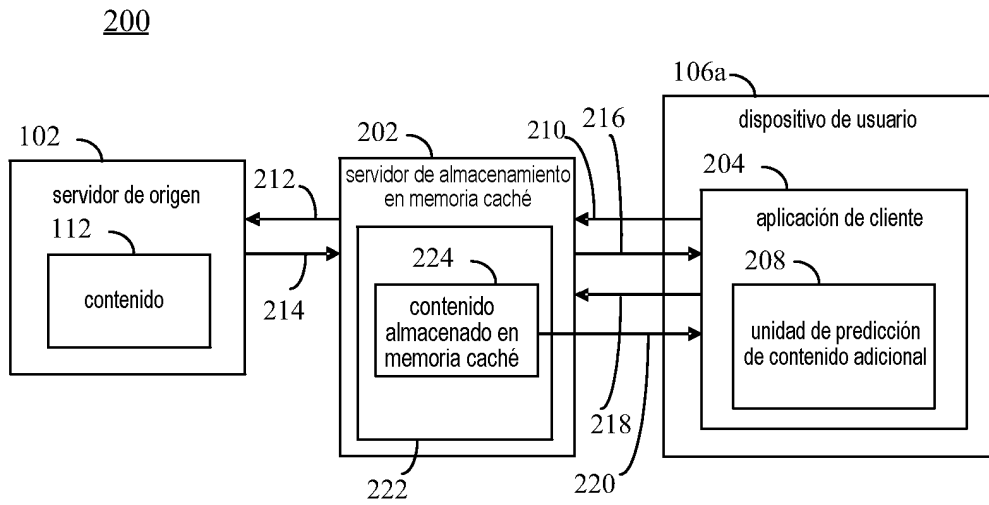


FIG. 2

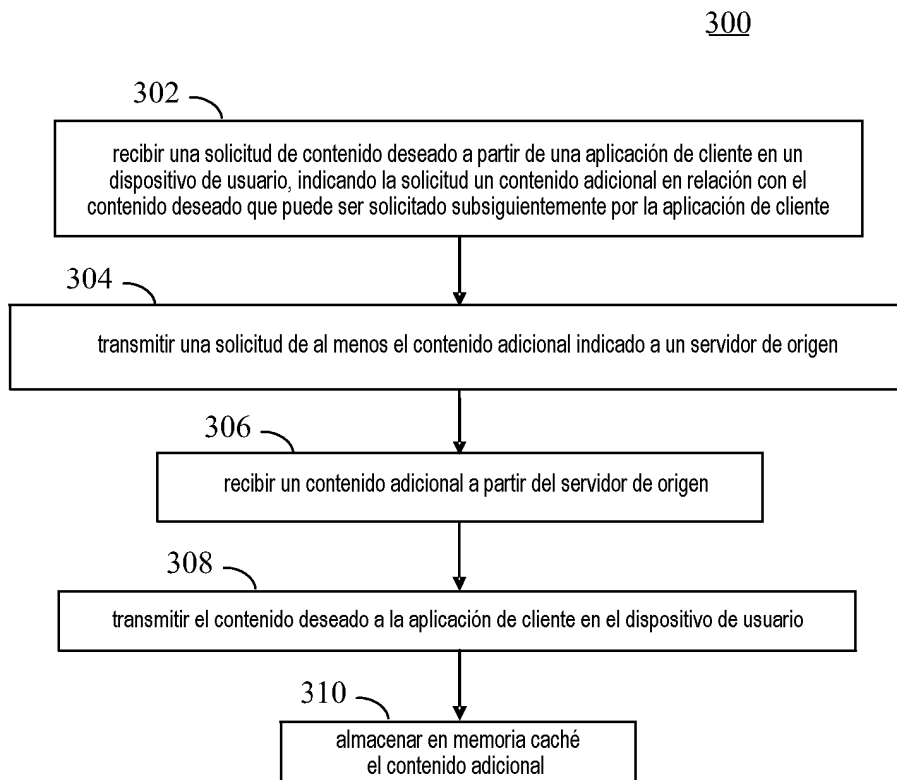


FIG. 3

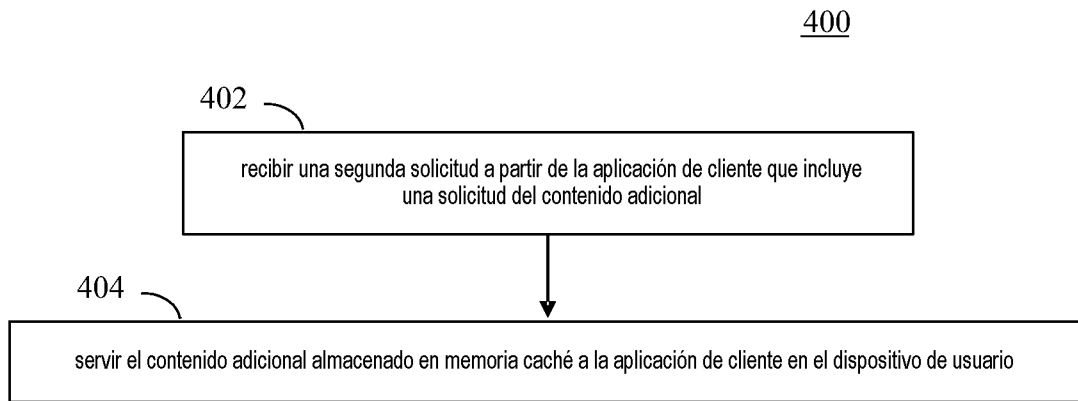


FIG. 4

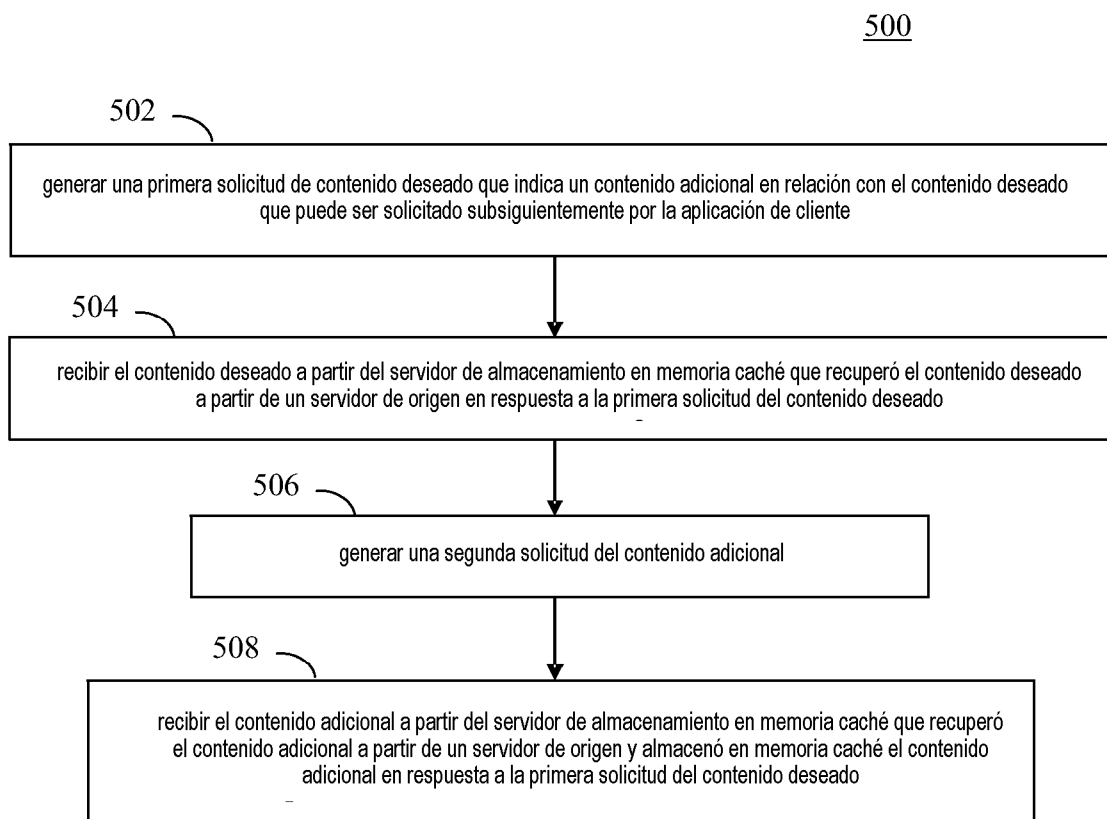


FIG. 5

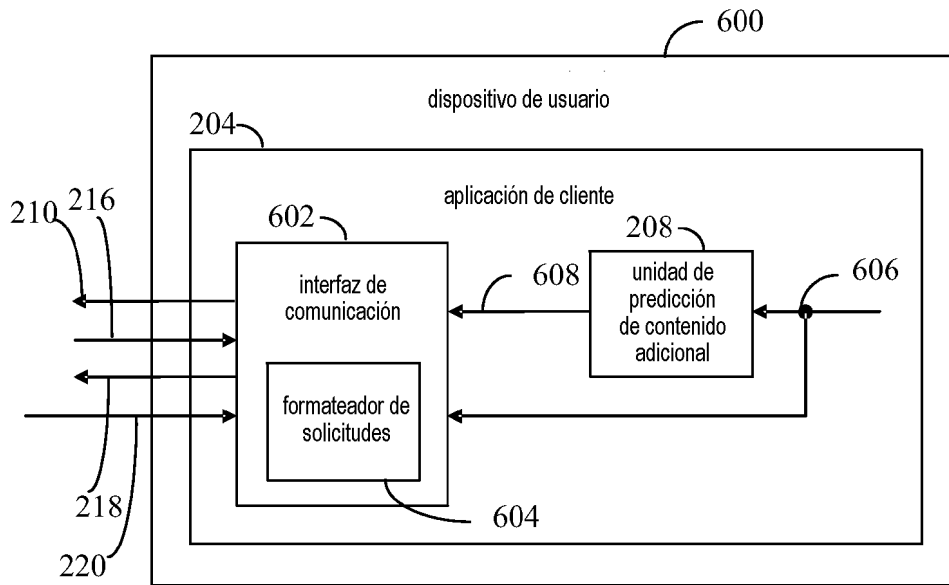


FIG. 6

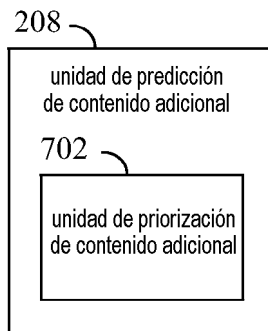


FIG. 7

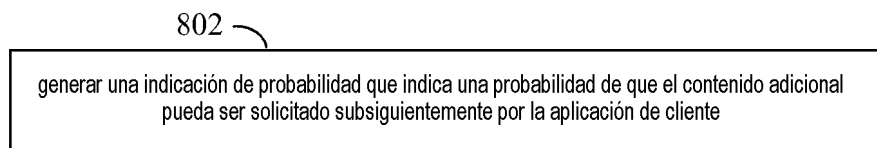


FIG. 8

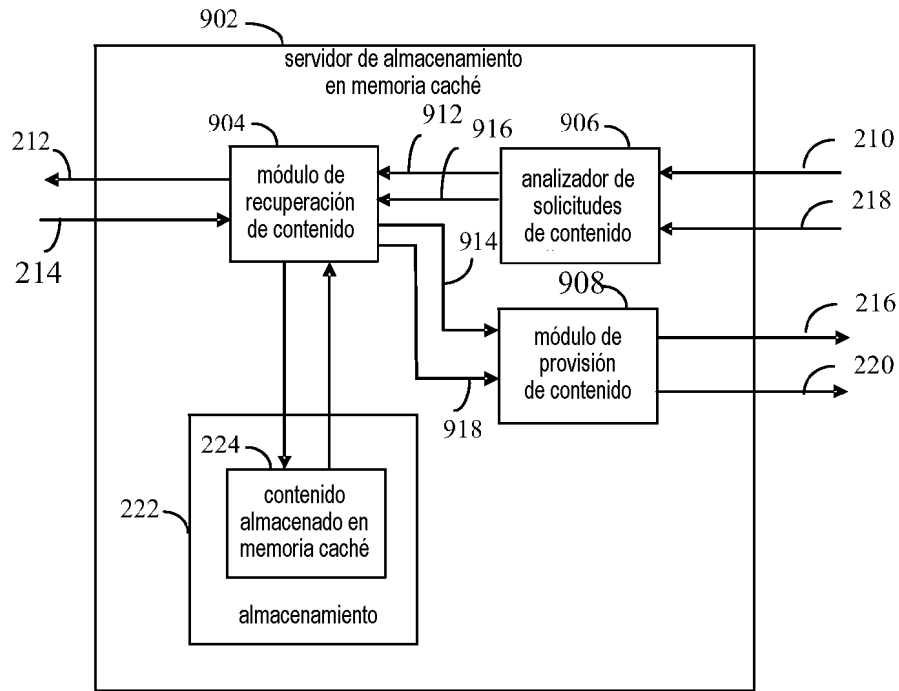


FIG. 9

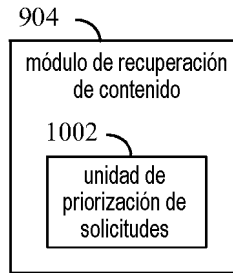


FIG. 10

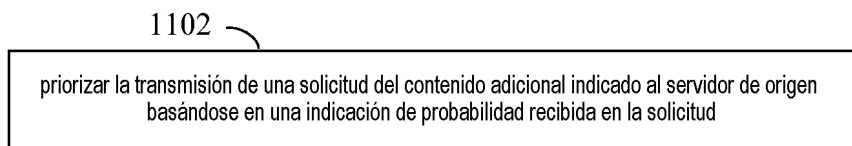


FIG. 11

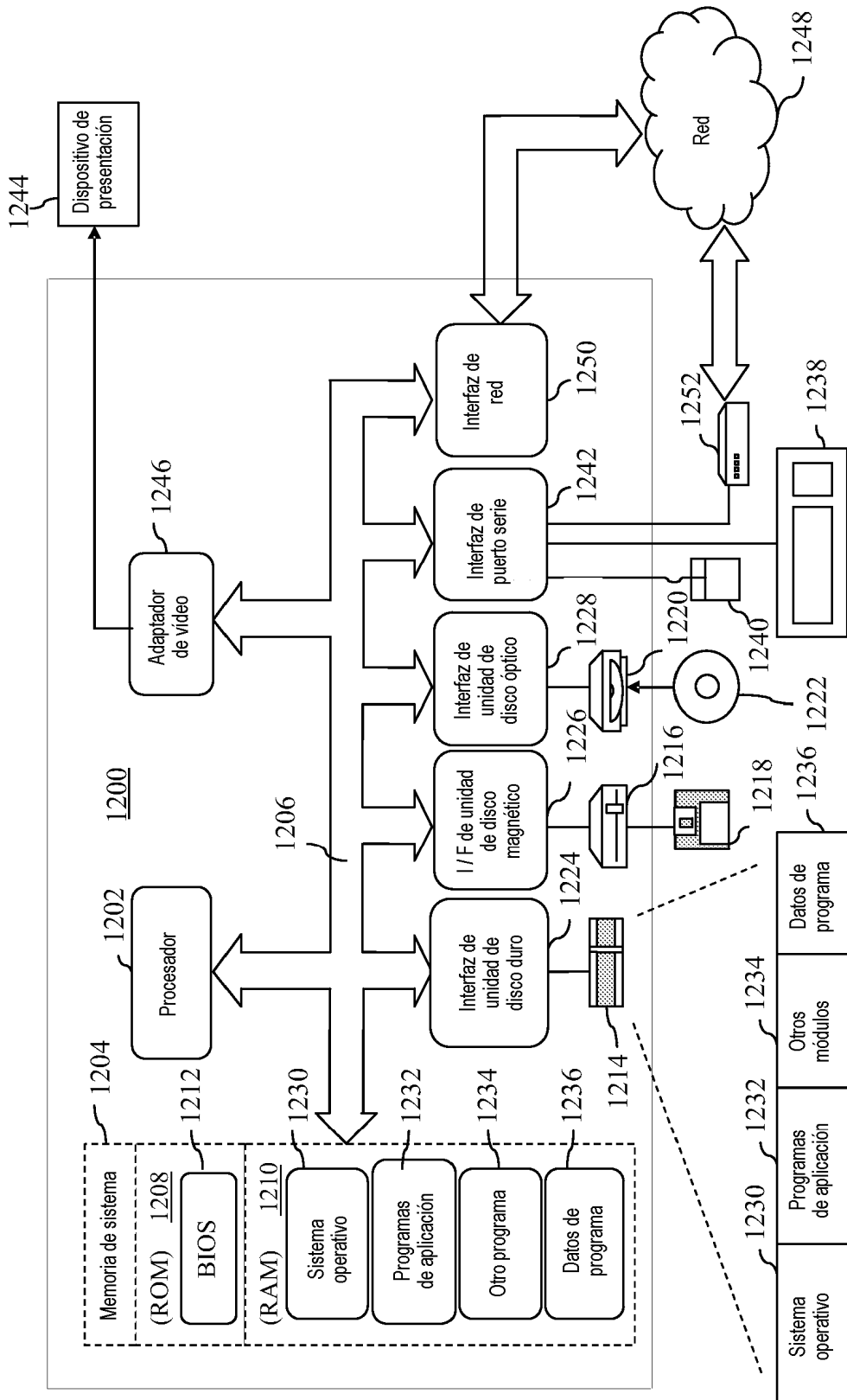


FIG. 12