

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 693 341**

51 Int. Cl.:

G06F 12/084 (2006.01)
G06F 12/08 (2006.01)
G06F 12/12 (2006.01)
G06F 11/34 (2006.01)
G06F 12/0846 (2006.01)
G06F 12/0842 (2006.01)
G06F 12/0831 (2006.01)
G06F 12/128 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **09.11.2015** **PCT/US2015/059677**
87 Fecha y número de publicación internacional: **02.06.2016** **WO16085641**
96 Fecha de presentación y número de la solicitud europea: **09.11.2015** **E 15797546 (7)**
97 Fecha y número de publicación de la concesión europea: **15.08.2018** **EP 3224728**

54 Título: **Proporcionar control de asignación de memoria caché compartida en sistemas de memoria caché compartida**

30 Prioridad:

25.11.2014 US 201462084480 P
22.09.2015 US 201514861025

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.12.2018

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)
5775 Morehouse Drive
San Diego, CA 92121-1714, US

72 Inventor/es:

HOWER, DEREK ROBERT y
CAIN, III, HAROLD WADE

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 693 341 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

Proporcionar control de asignación de memoria caché compartida en sistemas de memoria caché compartida

5 **ANTECEDENTES****I. Campo de la divulgación**

10 **[0001]** La tecnología de la divulgación se refiere en general a sistemas de memoria caché compartidos y, en particular, al control de asignación de memoria caché compartida.

II. Antecedentes

15 **[0002]** Un número creciente de unidades de hardware (por ejemplo, unidades de procesamiento central (CPU), unidades de procesamiento de gráficos (GPU), unidades de procesamiento de señal digital (DSP) y/o motores de acceso directo a memoria (DMA), como ejemplos no limitativos) para compartir recursos del sistema de memoria tales como memorias memoria caché, memoria, ancho de banda de interconexión y ancho de banda de memoria caché. La interferencia de los recursos y los conflictos entre las unidades de hardware podrían tener consecuencias negativas, como la falta de un plazo en tiempo real en un sistema móvil en chip (SoC) o la violación de un acuerdo de nivel de servicio (SLA) en un servidor consolidado, como ejemplos no limitativos. Además, los flujos de referencia asociados con algunas unidades de hardware pueden tener poca localidad temporal, lo que conduce a la contaminación de la memoria caché y un impacto negativo en el rendimiento general si no se controla. En este sentido, puede ser deseable permitir a los usuarios controlar el uso de recursos compartidos.

25 **[0003]** Sin embargo, a diferencia de la mayoría de otros componentes críticos de rendimiento de un sistema (por ejemplo, uso de CPU, uso de memoria, ancho de banda de red y ancho de banda de disco), la asignación de una memoria caché compartida convencional permanece fuera del control de los usuarios. Como resultado, tales sistemas convencionales de memoria caché compartida pueden permanecer subutilizados para proteger contra el peor rendimiento en caso de interferencia de memoria caché. Además, la calidad de servicio (QoS) de la memoria caché y los esquemas de partición existentes pueden carecer de la flexibilidad necesaria para abordar casos de uso múltiple, como la asignación estricta de capacidad y/o la maximización del rendimiento, como ejemplos no limitativos. El documento US 2008/0235457 A1 divulga una calidad dinámica de servicio que proporciona un procedimiento para distribuir el espacio de memoria caché, que implica asociar un indicador de prioridad con datos almacenados en una entrada de una memoria caché compartida por un núcleo/procesador que ejecuta un hilo.

35 **[0004]** El artículo titulado "CQoS: a framework for enabling QoS in shared caches of CMP platforms" ["CQoS: un marco para habilitar la QoS en memorias memoria caché compartidas de plataformas CMP"] de Iyer R. Ed - Asociación para Maquinaria de Computación, DOI: 10.1145/1006209.1006246, ISBN: 978-1-58113-839-9 divulga la partición dinámica de una memoria caché compartida, basada en clases de QoS.

40 **SUMARIO DE LA DIVULGACIÓN**

45 **[0005]** La presente invención está definida en las reivindicaciones independientes adjuntas, a las cuales se debería hacer referencia. Las características ventajosas están expuestas en las reivindicaciones dependientes adjuntas.

50 **[0006]** Los aspectos divulgados en la descripción detallada incluyen proporcionar control de asignación de memoria caché compartida en un sistema de memoria caché compartida. A este respecto, en un aspecto, se proporciona un controlador de memoria caché de una memoria caché compartida. El controlador de memoria caché admite una pluralidad de clases de calidad de servicio (QoS), cada una de las cuales corresponde a un identificador de QoS (QoSID). El controlador de memoria caché proporciona un circuito de asignación de memoria caché, que permite que las líneas de memoria caché de la memoria caché compartida se asignen selectivamente basándose en un QoSID de una solicitud de acceso a memoria recibida por el controlador de memoria caché. El circuito de asignación de memoria caché comprende una máscara de bits de mapeado mínimo, que mapea una clase de QoS a una partición mínima de las líneas de memoria caché de la memoria caché compartida. La partición mínima representa la porción más pequeña de la memoria caché compartida que puede asignarse a la clase QoS. De forma similar, el circuito de asignación de memoria caché también comprende una máscara de bits de mapeado máximo, que mapea la clase de QoS a una partición máxima de las líneas de memoria caché de la memoria caché compartida. La partición máxima representa la mayor porción de la memoria caché compartida que puede asignarse a la clase QoS. Al recibir una solicitud de acceso a memoria, el circuito de asignación de memoria caché determina si la solicitud de acceso a memoria corresponde a una línea de memoria caché de una pluralidad de líneas de memoria caché. Si no (es decir, si se produce una falta de memoria caché), el circuito de asignación de memoria caché selecciona como partición objetivo la partición mínima asignada a la clase QoS correspondiente al QoSID por la máscara de bits de mapeado mínimo o la partición máxima asignada a la clase QoS correspondiente al QoSID por la máscara de bits de mapeado máximo. El circuito de asignación de memoria caché luego asigna una línea de memoria caché dentro de la partición objetivo para una operación de llenado de memoria

caché. De esta manera, la memoria caché compartida puede dividirse en varias clases de QoS para permitir un control preciso y flexible sobre la asignación de las líneas de memoria caché de la memoria caché compartida.

[0007] En otro aspecto, se proporciona un controlador de memoria caché de un sistema de memoria caché compartida. El sistema de memoria caché compartida comprende una pluralidad de líneas de memoria caché. El controlador de memoria caché comprende un circuito de asignación de memoria caché. El circuito de asignación de memoria caché comprende una máscara de bits de mapeado mínimo para mapear una clase de QoS de una pluralidad de clases de QoS a una partición mínima de la pluralidad de líneas de memoria caché. El circuito de asignación de memoria caché comprende además una máscara de bits de mapeado máximo para mapear la clase de QoS de la pluralidad de clases de QoS a una partición máxima de la pluralidad de líneas de memoria caché. El circuito de asignación de memoria caché está configurado para recibir una solicitud de acceso a memoria que comprende un identificador de QoS (QoSID) correspondiente a la clase de QoS de la pluralidad de clases de QoS. El circuito de asignación de memoria caché está configurado además para determinar si la solicitud de acceso a memoria corresponde a una línea de memoria caché de la pluralidad de líneas de memoria caché. El circuito de asignación de memoria caché también está configurado para responder que la solicitud de acceso a memoria no corresponde a la línea de memoria caché de la pluralidad de líneas de memoria caché, para seleccionar, como una partición objetivo, una de la partición mínima mapeada a la clase de QoS correspondiente al QoSID por la máscara de bits de mapeado mínimo y la partición máxima mapeada a la clase de QoS correspondiente al QoSID por la máscara de bits de mapeado máximo. El circuito de asignación de memoria caché está configurado adicionalmente para asignar una línea de memoria caché dentro de la partición objetivo para una operación de llenado de memoria caché.

[0008] En otro aspecto, se proporciona un controlador de memoria caché de un sistema de memoria caché compartida. El controlador de memoria caché comprende un medio para recibir una solicitud de acceso a memoria que comprende un QoSID correspondiente a una clase de QoS de una pluralidad de clases de QoS. El controlador de memoria caché comprende además un medio para determinar si la solicitud de acceso a memoria corresponde a una línea de memoria caché de una pluralidad de líneas de memoria caché del sistema de memoria caché compartida. El controlador de memoria caché también comprende un medio para seleccionar, como una partición objetivo, una de una partición mínima asignada a la clase de QoS correspondiente al QoSID mediante una máscara de bits de mapeado mínimo y una partición máxima mapeada a la clase de QoS correspondiente al QoSID por una máscara de bits de mapeado máximo, que responde a la determinación de que la solicitud de acceso a memoria no se corresponde con la línea de memoria caché de la pluralidad de líneas de memoria caché. El controlador de memoria caché comprende adicionalmente un medio para asignar la línea de memoria caché de la pluralidad de líneas de memoria caché dentro de la partición objetivo para una operación de llenado de memoria caché, que responde a la determinación de que la solicitud de acceso a memoria no corresponde a la línea de memoria caché de la pluralidad de líneas de memoria caché.

[0009] En otro aspecto, se proporciona un procedimiento para asignar líneas de memoria caché de un sistema de memoria caché compartida. El procedimiento comprende recibir, mediante un circuito de asignación de memoria caché de un controlador de memoria caché, una solicitud de acceso a memoria que comprende un QoSID correspondiente a una clase de QoS de una pluralidad de clases de QoS. El procedimiento comprende además determinar si la solicitud de acceso a memoria corresponde a una línea de memoria caché de una pluralidad de líneas de memoria caché del sistema de memoria caché compartida. El procedimiento también comprende, en respuesta a que la solicitud de acceso a memoria no corresponde a la línea de memoria caché de la pluralidad de líneas de memoria caché, seleccionar, como una partición objetivo, una de una partición mínima mapeada a la clase de QoS correspondiente al QoSID por una máscara de bits de mapeado mínimo y una partición máxima mapeada a la clase de QoS correspondiente al QoSID por una máscara de bits de mapeado máximo. El procedimiento comprende adicionalmente asignar la línea de memoria caché de la pluralidad de líneas de memoria caché dentro de la partición objetivo para una operación de llenado de memoria caché.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0010]

La figura 1 es un diagrama de bloques que ilustra un sistema de memoria caché compartida ejemplar que incluye un controlador de memoria caché que comprende un circuito de asignación de memoria caché para proporcionar un control de asignación de memoria caché compartida;

Las figuras 2A-2C son diagramas de bloques que ilustran las máscaras de bits de mapeado mínimo y máximo ejemplares empleadas por el circuito de asignación de memoria caché de la Figura 1 para imponer la asignación del sistema de memoria caché compartida de la Figura 1;

Las figuras 3A-3C son diagramas de flujo que ilustran operaciones ejemplares del controlador de memoria caché de la Figura 1 para proporcionar control de asignación de memoria de memoria caché compartida; y

La figura 4 es un diagrama de bloques de un sistema basado en procesador ejemplar que puede incluir el controlador de memoria caché de la Figura 1 para proporcionar control de asignación de memoria caché compartida.

5 **DESCRIPCIÓN DETALLADA**

[0011] Con referencia a continuación a las figuras de los dibujos, se describen varios aspectos a modo de ejemplo de la presente divulgación. El término «a modo de ejemplo» se usa en el presente documento para indicar que «sirve de ejemplo, caso o ilustración». Cualquier aspecto descrito en el presente documento como «a modo de ejemplo» no debe interpretarse necesariamente que es preferente o ventajoso con respecto a otros aspectos.

[0012] A este respecto, la Figura 1 se proporciona para ilustrar una estructura de un sistema de memoria caché compartida 100. El sistema de memoria de memoria caché compartida 100 puede proporcionarse en una matriz de semiconductores, como un ejemplo no limitativo. En algunos aspectos, el sistema de memoria caché compartida 100 puede ser un memoria caché de nivel 1 (L1), un memoria caché de nivel 2 (L2), o un memoria caché de nivel 3 (L3), entre otros, en una jerarquía de memorias (no mostrada). 106(L)106(L)106(L) En el ejemplo de la figura 1, el sistema de memoria caché compartida 100 es una matriz de memoria organizada en bancos 102(0)-102(X). Cada uno de los bancos 102(0)-102(X) comprende uno o más conjuntos 104(0)-104(Y), con cada uno de los conjuntos 104(0)-104(Y) compuesto por un subconjunto de líneas de memoria caché 106(0)-106(L). El número Z de las líneas de memoria caché 106(0)-106(L) en cada uno de los conjuntos 104(0)-104(Y) se denomina "asociatividad" del sistema de memoria caché compartida 100, y el grupo de las líneas de memoria caché 106(0)-106(L) situadas en una posición particular desde 0-Z dentro de cada uno de los conjuntos 104(0)-104(Y) se denomina "camino" 108(0)-108(Z) correspondiente. Cada camino 108(0)-108(Z) se puede visualizar así como un "corte" vertical de las líneas de memoria caché 106(0)-106(L) a través de todos los conjuntos 104(0)-104(Y).

[0013] Debe entenderse que los aspectos descritos en este documento no están restringidos a ninguna disposición particular de elementos, y las técnicas divulgadas pueden extenderse fácilmente a diversas estructuras y disposiciones del sistema de memoria caché compartida 100. La configuración ilustrada en la Figura 1 es solo para fines ilustrativos. En algunos aspectos, el sistema de memoria caché compartida 100 puede comprender menos o más bancos 102(0)-102(X), menos o más conjuntos 104(0)-104(Y), menos o más líneas de memoria caché 106(0)-106(L), y/o menos o más caminos 108(0)-108(Z) que los aquí ilustrados.

[0014] Con referencia continua a la Figura 1, un controlador de memoria caché 110 está acoplado de manera conectiva a cada banco 102(0)-102(X). En la operación convencional, un agente solicitante (no mostrado), tal como una aplicación de software de ejecución, puede solicitar una instrucción o valor (no mostrado) almacenado en una dirección de memoria (no mostrada). Si la instrucción o el valor solicitado no se encuentra en el sistema de memoria caché compartida 100, se produce un error de memoria caché. En respuesta, el controlador de memoria caché 110 puede iniciar una operación de llenado de memoria caché, que provoca que una porción de la memoria del sistema (no mostrada) sea igual al tamaño de una o más de las líneas de memoria caché 106(0)-106(L) y que contiene la instrucción o valor solicitado, para ser recuperado y almacenado en una de las líneas de memoria caché 106(0)-106(L).

[0015] El sistema de memoria de memoria caché compartida 100 proporciona soporte para múltiples clases de QoS 112(0)-112(N) para implementar soporte de QoS para el uso de memoria compartida. Cada una de las clases de QoS 112(0)-112(N) está asociada con un único QoSID 114(0)-114(N). Para asignar más eficientemente las líneas de memoria caché 106(0)-106(L) entre múltiples agentes solicitantes, el controlador de memoria caché 110 en algunos aspectos puede proporcionar un circuito de asignación de memoria caché 116. El circuito de asignación de memoria caché 116 puede habilitar el control del usuario de los recursos de memoria caché del sistema de memoria caché compartida 100 asociando cada uno de los QoSID 114(0)-114(N) con un agente solicitante en particular, y especificando una asignación de las líneas de memoria caché 106(0)-106(L) para ese QoSID 114(0)-114(N).

[0016] Aunque un controlador de memoria caché convencional puede permitir la asignación de las líneas de memoria caché 106(0)-106(L), el controlador de memoria caché convencional puede no proporcionar la flexibilidad necesaria para abordar diferentes casos de uso. Por ejemplo, en un caso de uso del entorno de computación en la nube, puede ser deseable proporcionar una asignación estricta de las líneas de memoria caché 106(0)-106(L), de modo que múltiples máquinas virtuales de ejecución (VM) tengan acceso exclusivo a un subconjunto del líneas de memoria caché 106(0)-106(L). En otros casos de uso, puede ser deseable una asignación más flexible. Utilizando el mecanismo de asignación flexible, los usos de la memoria caché se aproximan a los valores medios especificados bajo la plena utilización del sistema de memoria caché compartida 100. En condiciones de baja contención, o cuando un consumidor de memoria caché se ejecuta solo, el uso debe acercarse a una asignación máxima establecida. Cuando un consumidor de memoria caché no se está ejecutando, el uso debe acercarse a un mínimo establecido. Finalmente, cuando un consumidor de memoria caché no requiere suficientes líneas de memoria caché 106(0)-106(L) para cumplir con la asignación promedio, la capacidad adicional debe ser utilizada por otra clase.

[0017] A este respecto, el circuito de asignación de memoria caché 116 proporciona una máscara de bits de mapeado mínimo (no mostrada) y una máscara de bits de mapeado máximo (no mostrada) para cada una de una o más de las clases de QoS 112(0)-112(N). Cada máscara de bits de mapeado mínimo se usa para mapear la clase de QoS correspondiente 112(0)-112(N) a una partición mínima de la pluralidad de líneas de memoria caché 106(0)-106(L) (es decir, una porción mínima de la pluralidad de líneas de memoria caché 106(0)-106(L) que pueden asignarse a esa clase QoS 112(0)-112(N)). Asimismo, cada máscara de mapeado máximo mapea su clase de QoS correspondiente 112(0)-112(N) a una partición máxima de la pluralidad de líneas de memoria caché 106(0)-106(L), que representa una porción máxima de la pluralidad de líneas de memoria caché 106(0)-106(L) que pueden asignarse a esa clase de QoS 112(0)-112(N). En algunos aspectos, las máscaras de bits de mapeado mínimo pueden almacenarse en una tabla de mapeado mínimo (no mostrada), mientras que las máscaras de bits de mapeado máximo pueden almacenarse en una tabla de mapeado máximo (no mostrada).

[0018] Para cada máscara de bits de mapeado mínimo y máscara de bits de mapeado máximo compuesta por un número N de bits, cada bit corresponde a una partición 1/N disjunta de la pluralidad de líneas de memoria caché 106(0)-106(L) del sistema de memoria caché compartida 100. Cada una de las máscaras de bits de mapeado mínimo y las máscaras de bits de mapeado máximo representa, por tanto, una vista lógica de una partición de memoria caché, con aspectos específicos que determinan cómo las máscaras de bits de mapeado mínimo y las máscaras de bits de mapeado máximo se traducen en una partición realizable. Como se analiza en mayor detalle a continuación, algunos aspectos también pueden proporcionar que se especifique un valor objetivo de asignación promedio (no mostrado) para cada una de las clases de QoS 112(0)-112(N). En tales aspectos, un valor objetivo de asignación promedio puede especificarse como un porcentaje de las líneas de memoria caché 106(0)-106(L) del sistema de memoria caché compartida 100, y puede usarse como una asignación objetivo cuando las máscaras de bits de mapeado mínimo y las máscaras de bits de mapeado máximo asociadas con cada clase de QoS 112(0)-112(N) son diferentes.

[0019] El circuito de asignación de memoria caché 116 está configurado además para recibir una solicitud de acceso a memoria 118 desde un agente solicitante (no mostrado). La solicitud de acceso a memoria 118 incluye un QoSID 120 que corresponde a una de las clases de QoS 112(0)-112(N). En el ejemplo de la Figura 1, el QoSID 120 corresponde al QoSID 114(0) de la clase de QoS 112(0). El circuito de asignación de memoria caché 116 determina si la solicitud de acceso a memoria 118 corresponde a una de la pluralidad de líneas de memoria caché 106(0)-106(L) del sistema de memoria caché compartida 100. Si la solicitud de acceso a memoria 118 no corresponde a ninguna de la pluralidad de líneas de memoria caché 106(0)-106(L) (es decir, una falta de memoria caché), el circuito de asignación de memoria caché 116 selecciona la partición mínima mapeada a la clase de QoS 112(0) correspondiente al QoSID 120 mediante la máscara de bits de mapeado mínimo o la partición máxima mapeada a la clase de QoS 112(0) correspondiente al QoSID 120 mediante la máscara de bits de mapeado máximo. Como se analiza con mayor detalle a continuación, en algunos aspectos, el circuito de asignación de memoria caché 116 puede seleccionar de manera determinista la máscara de bits de mapeado mínimo o la máscara de bits de mapeado máximo como sea necesario para asegurar que el uso de memoria caché se aproxime al valor objetivo de asignación promedio o puede seleccionar probabilísticamente la máscara de bits de mapeado mínimo o la máscara de bits mapeado máximo.

[0020] El circuito de asignación de memoria caché 116 entonces asigna una de las líneas de memoria caché 106(0)-106(L) dentro de la partición objetivo para una operación de llenado de memoria caché. En algunos aspectos, el circuito de asignación de memoria caché 116 puede asignar una de las líneas de memoria caché 106(0)-106(L) identificando primero una o más líneas de memoria caché 106(0)-106(L) dentro de la partición objetivo como uno o más candidatos de expulsión (a los que se hace referencia en este documento como "candidatos de expulsión 106(0)-106(L)"). El circuito de asignación de memoria caché 116 puede expulsar entonces a un candidato de expulsión 106(0)-106(L) del uno o más candidatos de expulsión 106(0)-106(L) del sistema de memoria caché compartida 100.

[0021] Asignando las líneas de memoria caché 106(0)-106(L) entre los QoSID 114(0)-114(N) (asignando efectivamente las líneas de memoria caché 106(0)-106(L) entre las clases de QoS correspondientes 112(0)-112(N)), el sistema de memoria caché compartida 100 puede mejorar el rendimiento de la memoria caché en una serie de escenarios de casos de uso. Por ejemplo, en algunos aspectos, el sistema de memoria caché compartida 100 puede permitir el mantenimiento de las restricciones en tiempo real para garantizar que las operaciones críticas de latencia llevadas a cabo por los agentes solicitantes tengan suficiente memoria caché para satisfacer los requisitos en tiempo real. Para proporcionar un servicio receptivo a las restricciones en tiempo real, el circuito de asignación de memoria caché 116 puede configurarse para garantizar que se preservará un conjunto de trabajo mínimo de las líneas de memoria caché 106(0)-106(L) para un agente solicitante de tiempo crítico en el sistema de memoria caché compartida 100. La asignación de las líneas de memoria caché 106(0)-106(L) por el circuito de asignación de memoria caché 116 también puede ser útil para aplicar Acuerdos de Nivel de Servicio (SLA) de computación en la nube, y/o para restringir la asignación de memoria caché de las líneas de memoria caché 106(0)-106(L) para ciertos agentes solicitantes para mejorar el rendimiento general de un sistema de procesamiento informático. Las máscaras de bits de mapeado mínimo y máximo a modo de ejemplo empleadas por el circuito de asignación de memoria caché 116 de la figura 1 para imponer la asignación del sistema de

memoria caché compartida 100 de la figura 1 en casos de uso diferentes se analizan con mayor detalle a continuación con respecto a las figuras 2A-2C.

[0022] En algunos aspectos, el circuito de asignación de memoria caché 116 también puede proporcionar un valor objetivo de asignación promedio para cada una de las clases de QoS 112(0)-112(N). Cuando posteriormente se selecciona la partición mínima o la partición máxima como la partición objetivo para la clase de QoS 112(0)-112(N) correspondiente al QoSID 120, el circuito de asignación de memoria caché 116 puede basar su selección en la partición objetivo que cause una asignación real de las líneas de memoria caché 106(0)-106(L) asignadas a la clase QoS 112(0)-112(N) para acercarse al valor objetivo promedio de asignación. Así, cuando la clase de QoS 112(0)-112(N) se está ejecutando activamente en un sistema disputado, el circuito de asignación de memoria caché 116 intentará mantener la asignación real de la pluralidad de líneas de memoria caché 106(0)-106(L) para la clase de QoS 112(0)-112(N) cerca del valor objetivo promedio de asignación. Si la clase de QoS 112(0)-112(N) no está programada, o si la clase de QoS 112(0)-112(N) está utilizando una asignación de la pluralidad de líneas de memoria caché 106(0)-106(L), eso es más pequeño que la asignación promedio, la asignación real se aproximará a un nivel mínimo representado por la máscara de bits de mapeado mínimo correspondiente. Si la clase de QoS 112(0)-112(N) se ejecuta aisladamente o con otras cargas de trabajo que tienen una ligera presión de memoria caché, la asignación real se aproximará a un nivel máximo representado por la máscara de bits de mapeado máximo correspondiente.

[0023] Según algunos aspectos, el circuito de asignación de memoria caché 116 puede usar una indicación 122 de una asignación de la pluralidad de líneas de memoria caché 106(0)-106(L) entre la pluralidad de clases de QoS 112(0)-112(N) desde un monitor de uso de memoria caché 124 para guiar la selección de la partición mínima o la partición máxima como la partición objetivo. Como un ejemplo no limitativo, la indicación 122 puede representar una medida aproximada del uso de la memoria caché por una clase de QoS dada 112(0)-112(N). En base a la indicación 122, el circuito de asignación de memoria caché 116 en algunos aspectos puede seleccionar la partición mínima o la partición máxima según sea necesario para provocar que la asignación de memoria caché para esa clase de QoS 112(0)-112(N) tienda hacia un valor objetivo de asignación promedio.

[0024] Utilizando la máscara de bits de mapeado mínimo o la máscara de bits de mapeado máximo para determinar la partición durante las operaciones de llenado de memoria caché, la asignación resultante de la pluralidad de líneas de memoria caché 106(0)-106(L) oscilará naturalmente entre los usos mínimo y máximo. Por ejemplo, si el circuito de asignación de memoria caché 116 selecciona la máscara de bits de mapeado mínimo, el uso de memoria caché para la clase de QoS correspondiente 112(0)-112(N) disminuirá gradualmente hacia su partición mínima. Si el circuito de asignación de memoria caché 116 selecciona la máscara de bits de mapeado máximo, el uso de memoria caché para la clase de QoS correspondiente 112(0)-112(N) aumentará hacia la partición máxima. Por consiguiente, en algunos aspectos, el circuito de asignación de memoria caché 116 puede seleccionar determinísticamente la máscara de bits de mapeado mínimo cuando el uso de memoria caché actual está por encima del valor de objetivo de asignación promedio, y la máscara de bits de mapeado máximo cuando el uso de memoria caché actual es inferior al valor objetivo de asignación promedio.

[0025] En algunos aspectos, puede ser deseable que el circuito de asignación de memoria caché 116 seleccione probabilísticamente la máscara de bits de mapeado mínimo o la máscara de bits de mapeado máximo, de manera que el circuito de asignación de memoria caché 116 seleccione periódicamente la máscara de bits "incorrecta" (es decir, la máscara de bits que da como resultado que el uso actual de la memoria caché se aleje del valor objetivo de asignación promedio). Al seleccionar probabilísticamente, el circuito de asignación de memoria caché 116 puede ser más efectivo explorando asignaciones de memoria caché que favorecen los patrones de memoria caché natural para cada una de las clases de QoS 112(0)-112(N), lo que resulta en un uso más eficaz y efectivo del sistema de memoria caché compartida 100.

[0026] Según algunos aspectos, el circuito de asignación de memoria caché 116 puede configurarse para seleccionar de forma probabilística la máscara de bits de mapeado mínimo o la máscara de bits de mapeado máximo basándose en una función de probabilidad (no mostrada) proporcionada por un circuito proveedor de función de probabilidad 126. Algunos aspectos pueden proporcionar que el circuito proveedor de función de probabilidad 126 se pueda basar en una función de probabilidad lineal y/o una función de probabilidad sigmoidea, como ejemplos no limitativos. Por ejemplo, una función de probabilidad lineal puede ser preferible en aspectos en los que la facilidad de implementación es importante. Según algunos aspectos, el uso de una función de probabilidad sigmoidea puede dar como resultado una tendencia más rápida hacia el valor objetivo de asignación promedio, permitiendo así que el circuito de asignación de memoria caché 116 sea más receptivo a medida que se producen los cambios (por ejemplo, circunstancias en las que una clase de QoS 112(0)-112(N) tiene cambios de fase frecuentes, como un ejemplo no limitativo). En aspectos que emplean una función de probabilidad sigmoidea, el circuito proveedor de función de probabilidad 126 puede implementarse como una tabla de búsqueda de hardware prepopulada (no mostrada) junto con un registro de desplazamiento de retroalimentación lineal (no mostrado) como fuente de aleatoriedad. A continuación se muestra una función sigmoidea ejemplar que toma el parámetro de forma μ :

$$x = \frac{\text{promedio} - \text{uso}}{\text{promedio} - \text{mín}}, y = \frac{\text{uso} - \text{promedio}}{\text{máx} - \text{promedio}}$$

$$P(m) = \begin{cases} \text{uso} < \text{promedio} : \frac{1}{2} - \frac{\ln(xe^{\mu} + 1/\mu)}{2\mu} \\ \text{uso} \geq \text{promedio} : \frac{1}{2} + \frac{\ln(ye^{\mu} + 1/\mu)}{2\mu} \end{cases}$$

[0027] Para ilustrar las máscaras de bits de mapeado mínimo y máximo ejemplares que pueden emplearse por el circuito de asignación de memoria caché 116 de la figura 1 para imponer la asignación del sistema de memoria caché compartida 100, se proporcionan las Figuras 2A a 2C. En particular, la Figura 2A ilustra ejemplos de máscaras de bits de mapeado mínimo y máscaras de bits de mapeado máximo que pueden usarse para aislar estrictamente las clases de QoS 112(0) y 112(1) entre sí. La Figura 2B ilustra ejemplos de máscaras de bits de mapeado mínimo para mapear las clases de QoS 112(0), 112(1) a particiones mínimas que son exclusivas para cada clase de QoS 112(0), 112(1) y máscaras de bits de mapeado máximo para mapear las clases de QoS 112(0), 112(1) a particiones máximas que no son exclusivas para cada clase de QoS 112(0), 112(1). La Figura 2C ilustra ejemplos de máscaras de bits de mapeado mínimo y máscaras de bits de mapeado máximo que pueden usarse para reducir el impacto de las clases de QoS 112(0)-112(N) que pueden dar como resultado la pérdida de la memoria caché. En las Figuras 2A-2C, las clases de QoS 112(0)-112(2) están asociadas con las máscaras de bits de mapeado mínimo 200(0)-200(2), las máscaras de bits de mapeado máximo 202(0)-202(2) y valores objetivo de asignación promedio 204(0)-204(2), respectivamente, mediante el circuito de asignación de memoria caché 116. Las máscaras de bits de mapeado mínimo 200(0)-200(2) mapean particiones mínimas 206(0)-206(2) para las clases de QoS 112(0)-112(2), mientras que las máscaras de bits de mapeado máximo 202(0)-202(2) mapean particiones máximas 208(0)-208(2) para las clases de QoS 112(0)-112(2).

[0028] En el ejemplo de la Figura 2A, las máscaras de bits de mapeado mínimo 200(0), 200(1) y las máscaras de bits de mapeado máximo 202(0), 202(1) mapean las clases de QoS 112(0), 112(1) a una misma partición objetivo de la pluralidad de líneas de memoria caché 106(0)-106(L) que es exclusiva para cada clase de QoS 112(0), 112(1). Esto da como resultado que las clases de QoS 112(0), 112(1) estén estrictamente aisladas entre sí dentro del sistema de memoria caché compartida 100, lo que puede ser deseable en casos de uso tales como aplicaciones de computación en la nube. Debido a que las máscaras de bits de mapeado mínimo 200(0), 200(1) y las máscaras de bits de mapeado máximo 202(0), 202(1) son las mismas para una clase de QoS dada 112(0), 112(1), las clases de QoS 112(0), 112(1) nunca excederán el uso especificado por las máscaras de bits de mapeado mínimo 200(0), 200(1) y las máscaras de bits de mapeado máximo 202(0), 202(1). En algunos aspectos en los que las máscaras de bits de mapeado mínimo 200(0), 200(1) y las máscaras de bits de mapeado máximo 202(0), 202(1) son iguales, el circuito de asignación de memoria caché 116 puede usar una sola de las máscaras de bits de mapeado mínimo 200(0), 200(1) y las máscaras de bits de mapeado máximo 202(0), 202(1) cuando se selecciona una partición objetivo.

[0029] En la Figura 2B, cada una de las máscaras de bits de mapeado máximo 202(0), 202(1) se especifica como un superconjunto de la respectiva máscara de bits de mapeado mínimo 200(0), 200(1) para una clase de QoS dada 112(0), 112(1). En consecuencia, las máscaras de bits de mapeado mínimo 200(0), 200(1) mapean la clase de QoS 112(0), 112(1) a las particiones mínimas 206(0), 206(1) que son exclusivas de las clases de QoS 112(0), 112(1). De manera similar, las máscaras de bits de mapeado máximo 202(0), 202(1) mapean la clase de QoS 112(0), 112(1) a las particiones máximas 208(0), 208(1) que no son exclusivas de las clases de QoS 112(0), 112(1) (es decir, porciones de las particiones máximas 208(0), 208(1) son compartidas entre las clases de QoS 112(0), 112(1)). Esta configuración puede permitir que un usuario se asegure de que las clases de QoS 112(0), 112(1) tengan cada una acceso exclusivo a una porción de la pluralidad de líneas de memoria caché 106(0)-106(L), pero también permitan que las clases de QoS 112(0), 112(1) usen más de la pluralidad de líneas de memoria caché 106(0)-106(L) bajo contención ligera. Como se ve en la Figura 2B, las clases de QoS 112(0), 112(1) tienen garantizada una asignación mínima del 25 %, pero se les permite usar hasta el 75 % de la pluralidad de líneas de memoria caché 106(0)-106(L) cuando no se están ejecutando aplicaciones de baja prioridad o cuando esas aplicaciones de baja prioridad no usan más del 25 % de la pluralidad de líneas de memoria caché 106(0)-106(L). Debe tenerse en cuenta que en esta configuración, la ubicación de los bits es importante. En particular, las clases de QoS 112(0), 112(1) reciben cada una el 25 % de la pluralidad de líneas de memoria caché 106(0)-106(L) porque los bits en sus máscaras de bits de mapeado mínimo 200(0), 200(1) no se superponen.

[0030] Con referencia ahora a la Figura 2C, para maximizar el rendimiento, las máscaras de bits de mapeado mínimo 200(0)-200(2) y las máscaras de bits de mapeado máximo 202(0)-202(2) se pueden configurar para reducir el impacto de las aplicaciones que hacen perder la memoria caché. En el ejemplo de la Figura 2C, la máscara de bits de mapeado mínimo 200(0) y la máscara de bits de mapeado máximo 202(0) limitan estrictamente la asignación para la clase de QoS 112(0). Las máscaras de bits de mapeado mínimo 200(1), 200(2) y la máscara de bits de mapeado máximo 202(1), 202(2) permiten a las clases de QoS 112(1), 112(2) compartan el resto de la pluralidad de líneas de memoria caché 106(0)-106(L), asegurando una asignación mínima para las clases de QoS

112(1), 112(2). Como resultado, cualquier pérdida de la memoria caché causada por aplicaciones asociadas con las clases de QoS 112(1), 112(2) se limitará a las particiones superpuestas de las particiones máximas 208(1), 208(2) y no afectará el resto de la memoria caché.

[0031] Las Figuras 3A-3C son diagramas de flujo que ilustran operaciones a modo de ejemplo del circuito de asignación de memoria caché 116 de la Figura 1 para proporcionar un control de asignación de la memoria caché compartida. Por motivos de claridad, los elementos de las Figuras 1 y 2A-2C se mencionan al describir las Figuras 3A-3C. En algunos aspectos, las operaciones comienzan con el circuito de asignación de memoria caché 116 que proporciona un valor objetivo de asignación promedio 204(0) (bloque 300). Por consiguiente, el circuito de asignación de memoria caché 116 se puede denominar aquí como "un medio para proporcionar un valor objetivo de asignación promedio". El circuito de asignación de memoria caché 116 recibe una solicitud de acceso a memoria 118 que comprende un QoSID 120 correspondiente a una clase de QoS 112(0) de una pluralidad de clases de QoS 112(0)-112(N) (bloque 302). De este modo, el circuito de asignación de memoria caché 116 puede denominarse en la presente memoria como "un medio para recibir una solicitud de acceso a memoria que comprende un QoSID correspondiente a una clase de QoS de una pluralidad de clases de QoS".

[0032] El circuito de asignación de memoria caché 116 determina entonces si la solicitud de acceso a memoria 118 corresponde a una línea de memoria caché 106(0) de la pluralidad de líneas de memoria caché 106(0)-106(L) del sistema de memoria caché compartida 100 (bloque 304). A este respecto, el circuito de asignación de memoria caché 116 se puede denominar aquí "un medio para determinar si la solicitud de acceso a memoria corresponde a una línea de memoria caché de una pluralidad de líneas de memoria caché del sistema de memoria caché compartida". Si el circuito de asignación de memoria caché 116 determina en el bloque de decisión 304 que la solicitud de acceso a memoria 118 corresponde a una línea de memoria caché 106(0) de la pluralidad de líneas de memoria caché 106(0)-106(L), el circuito de asignación de memoria caché 116 devuelve la línea de memoria caché 106(0) (es decir, un acierto de memoria caché) (bloque 306). Sin embargo, si la solicitud de acceso a memoria 118 no corresponde a una línea de memoria caché 106(0) de la pluralidad de líneas de memoria caché 106(0)-106(L) (es decir, una falta de memoria caché), el procesamiento se reanuda en el bloque 308 de la Figura 3B.

[0033] Con referencia ahora a la Figura 3B, el circuito de asignación de memoria caché 116 selecciona, como una partición objetivo, una de una partición mínima 206(0) asignada a la clase de QoS 112(0) correspondiente al QoSID 120 mediante una máscara de bits de mapeado mínimo 200(0) y una partición máxima 208(0) mapeada a la clase de QoS 112(0) correspondiente al QoSID 120 mediante una máscara de bits de mapeado máximo 202(0) (bloque 308). En consecuencia, el circuito de asignación de memoria caché 116 puede denominarse aquí como "un medio para seleccionar, como una partición objetivo, una de una partición mínima asignada a la clase de QoS correspondiente a la QoSID mediante una máscara de bits de mapeado mínimo y una partición máxima asignada a la clase de QoS correspondiente al QoSID mediante una máscara de bits de mapeado máximo". En algunos aspectos, las operaciones del bloque 308 para seleccionar la de la partición mínima 206(0) y la partición máxima 208(0) incluyen seleccionar la de la partición mínima 206(0) y la partición máxima 208(0) a causa una asignación real de las líneas de memoria 106(0)-106(L) asignadas a la clase de QoS 112(0) correspondiente al QoSID 120 para aproximarse al valor objetivo de asignación promedio 204(0) (bloque 310). El circuito de asignación de memoria caché 116 puede denominarse así como "un medio para seleccionar uno de la partición mínima y la partición máxima para provocar una asignación real de líneas de memoria caché asignadas a la clase de QoS correspondiente al QoSID para aproximarse al valor objetivo de asignación promedio". Algunos aspectos pueden proporcionar que las operaciones del bloque 310 para seleccionar la una de la partición mínima 206(0) y la partición máxima 208(0) se basan en una indicación 122 de una asignación de la pluralidad de líneas de memoria caché 106(0)-106(L) entre la pluralidad de clases de QoS 112(0)-112(N) proporcionadas por un monitor de uso de memoria caché 124 (bloque 312). Como un ejemplo no limitativo, la indicación 122 puede representar una medida aproximada del uso de la memoria caché por una clase de QoS dada 112(0)-112(N). En base a la medición aproximada de uso de memoria caché indicada por la indicación 122, el circuito de asignación de memoria caché 116 en algunos aspectos puede seleccionar la partición mínima o la partición máxima según sea necesario para causar la asignación de memoria caché para esa clase de QoS 112(0)-112(N) para tender hacia el valor objetivo de asignación promedio 204(0). A este respecto, el circuito de asignación de memoria caché 116 se puede denominar aquí como un medio para seleccionar la una de la partición mínima y la partición máxima basándose en una indicación de una asignación de la pluralidad de líneas de memoria caché entre la pluralidad de clases de QoS proporcionadas por un monitor de uso de memoria caché".

[0034] En algunos aspectos, las operaciones del bloque 308 para seleccionar la una de la partición mínima 206(0) y la partición máxima 208(0) mediante el circuito de asignación de memoria caché 116 incluyen la selección probabilística de una de la partición mínima 206(0) y la partición máxima 208(0) (bloque 314). En consecuencia, el circuito de asignación de memoria caché 116 se puede denominar aquí como "un medio para seleccionar probabilísticamente una de las particiones mínimas y las particiones máximas". Algunos aspectos pueden proporcionar que las operaciones del bloque 314 para seleccionar probabilísticamente una de la partición mínima 206(0) y la partición máxima 208(0) se basan en un circuito proveedor de función de probabilidad 126 (bloque 316). El procesamiento se reanuda en el bloque 318 de la figura 3C.

[0035] Pasando ahora a la figura 3C, el circuito de asignación de memoria caché 116 luego asigna una línea de memoria caché 106(0) dentro de la partición objetivo para una operación de llenado de memoria caché (bloque 318). A este respecto, el circuito de asignación de memoria caché 116 puede denominarse aquí "un medio para asignar una línea de memoria caché dentro de la partición objetivo para una operación de llenado de memoria caché, que responde a la determinación de que la solicitud de acceso a memoria no corresponde a la línea de memoria caché de la pluralidad de líneas de memoria caché". Según algunos aspectos, las operaciones del bloque 318 para asignar la línea de memoria caché 106(0) dentro de la partición objetivo para la operación de llenado de memoria caché incluyen el circuito de asignación de memoria caché 116 que identifica una o más líneas de memoria caché 106(0)-106(L) dentro de la partición objetivo como uno o más candidatos de expulsión 106(0)-106(L) (bloque 320). En consecuencia, el circuito de asignación de memoria caché 116 se puede denominar aquí como "un medio para identificar una o más líneas de memoria caché dentro de la partición objetivo como uno o más candidatos de expulsión". El circuito de asignación de memoria caché 116 puede expulsar entonces a un candidato de expulsión 106(0) del uno o más candidatos de expulsión 106(0)-106(L) del sistema de memoria caché compartida 100 (bloque 322). El circuito de asignación de memoria caché 116 se puede denominar de este modo como "un medio para expulsar a un candidato de expulsión del uno o más candidatos de expulsión del sistema de memoria caché compartida".

[0036] Proporcionar un control de asignación de la memoria caché compartida en los sistemas de memoria caché compartida puede proporcionarse o integrarse en cualquier dispositivo basado en procesador. Los ejemplos, sin limitación, incluyen un módulo de conexión, una unidad de entretenimiento, un dispositivo de navegación, un dispositivo de comunicaciones, una unidad de datos de ubicación fija, una unidad de datos de ubicación móvil, un teléfono móvil, un teléfono celular, un ordenador, un ordenador portátil, un ordenador de sobremesa, un asistente digital personal (PDA), un monitor, un monitor de ordenador, un televisor, un sintonizador, una radio, una radio por satélite, un reproductor de música, un reproductor de música digital, un reproductor de música portátil, un reproductor de vídeo digital, un reproductor de vídeo, un reproductor de discos de vídeo digital (DVD) y un reproductor de vídeo digital portátil.

[0037] A este respecto, la Figura 4 es un diagrama de bloques de un sistema 400 basado en procesador ejemplar que puede incluir el sistema de memoria caché compartida (SCMS) 100 que tiene el controlador de memoria caché 110 de la Figura 1. En este ejemplo, el sistema basado en procesador 400 incluye una o más CPU 402, que incluyen cada una uno o más procesadores 404. La o las CPU 402 pueden ser un dispositivo maestro. La o las CPU 402 puede(n) tener una memoria caché 406 acoplada al/a los procesador(es) 404 para un rápido acceso a datos almacenados temporalmente. En algunos aspectos, la memoria caché 406 puede comprender el sistema de memoria caché compartida 100 y/o el circuito de asignación de memoria caché 116 de la Figura 1. La(s) CPU 402 está(n) acoplada(s) a un bus de sistema 408 y pueden interconectar dispositivos maestros y esclavos, incluidos en el sistema 400 basado en procesadores. Como es bien sabido, la o las CPU 402 se comunican(n) con estos otros dispositivos intercambiando información de dirección, control y datos por el bus del sistema 408. Por ejemplo, la(s) CPU 402 puede(n) comunicar solicitudes de transacciones de bus a un controlador de memoria 410, como un ejemplo de un dispositivo esclavo.

[0038] Otros dispositivos maestros y esclavos pueden conectarse al bus del sistema 408. Según se ilustra en la Figura 4, estos dispositivos pueden incluir un sistema de memoria 412, uno o más dispositivos de entrada 414, uno o más dispositivos de salida 416, uno o más dispositivos de interfaz de red 418 y uno o más controladores de visualización 420, como ejemplos. El/los dispositivo(s) 414 puede(n) incluir cualquier tipo de dispositivo de entrada, incluyendo, pero sin limitarse a, teclas de entrada, conmutadores, procesadores de voz, etc. El/los dispositivo(s) de salida 416 puede(n) incluir cualquier tipo de dispositivo de salida, incluyendo, pero sin limitarse a, audio, vídeo, otros indicadores visuales, etc. El/los dispositivo(s) de interfaz de red 418 puede(n) ser cualquier dispositivo configurado para permitir el intercambio de datos a y desde una red 422. La red 422 puede ser cualquier tipo de red, incluyendo, pero sin limitarse a, una red cableada o inalámbrica, una red privada o pública, una red de área local (LAN), una red de área local inalámbrica (WLAN), una red de área amplia (WAN), una red BLUETOOTH™ e Internet. El/los dispositivo(s) de interfaz de red 418 puede(n) configurarse para admitir cualquier tipo de protocolo de comunicaciones deseado. El sistema de memoria 412 puede incluir una o más unidades de memoria 424(0)-424(N).

[0039] La(s) CPU 402 también puede(n) configurarse para acceder al/a los controlador(es) de visualización 420 por el bus de sistema 408, para controlar información enviada a una o más pantallas 426. El/los controlador(es) de visualización 420 envía(n) información a la(s) pantalla(s) 426, para que se visualice(n) a través de uno o más procesadores de vídeo 428, que procesan la información que vaya de visualizarse a un formato adecuado para la(s) pantalla(s) 426. La o las pantallas 426 pueden incluir cualquier tipo de pantalla, incluyendo, pero sin limitarse a, una pantalla de tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, etc.

[0040] Los expertos en la materia apreciarán, además, que los diversos bloques lógicos, módulos, circuitos y algoritmos ilustrativos descritos en relación con los aspectos divulgados en el presente documento pueden implementarse como hardware electrónico. Los dispositivos descritos en el presente documento pueden emplearse en cualquier circuito, componente de hardware, circuito integrado (IC) o chip de IC, como ejemplos. La

memoria divulgada en el presente documento puede ser una memoria de cualquier tipo y tamaño y puede configurarse para almacenar cualquier tipo de información deseada. Para ilustrar claramente esta intercambiabilidad, anteriormente se han descrito diversos componentes, bloques, módulos, circuitos y etapas ilustrativos, generalmente en términos de su funcionalidad. Cómo se implementa dicha funcionalidad depende de la aplicación en particular, de las elecciones de diseño y/o de las restricciones de diseño que se imponen en el sistema general. Los expertos en la técnica pueden implementar la funcionalidad descrita de varias maneras para cada aplicación particular, pero no se debería interpretar que dichas decisiones de implementación suponen apartarse del alcance de las reivindicaciones adjuntas.

[0041] Los diversos bloques lógicos ilustrativos, módulos y circuitos descritos con relación a los modos de realización divulgados en el presente documento pueden implementarse o realizarse con un procesador, un procesador de señales digitales (DSP), un circuito integrado específico de la aplicación (ASIC), una matriz de puertas programables por campo (FPGA) u otro dispositivo lógico programable, compuerta discreta o lógica de transistor, componentes discretos de hardware o cualquier combinación de los mismos diseñada para realizar las funciones descritas en el presente documento. Un procesador puede ser un microprocesador pero, como alternativa, el procesador puede ser cualquier procesador, controlador, microcontrolador o máquina de estado convencional. Un procesador también puede implementarse como una combinación de dispositivos informáticos, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores junto con un núcleo de DSP o cualquier otra configuración de este tipo.

[0042] También se observa que las etapas operativas descritas en cualquiera de los modos de realización ejemplares en el presente documento están descritas para proporcionar ejemplos y análisis. Las operaciones descritas pueden realizarse en numerosas secuencias diferentes distintas de las secuencias ilustradas. Además, las operaciones descritas en una única etapa operativa pueden realizarse realmente en varias etapas diferentes. Adicionalmente, pueden combinarse una o más etapas operativas analizadas en los aspectos a modo de ejemplo. Se entenderá que las etapas operativas ilustradas en los diagramas de flujo pueden someterse a numerosas modificaciones diferentes, como resultará inmediatamente evidente para un experto en la técnica. Los expertos en la técnica también entenderán que la información y las señales pueden representarse usando cualquiera entre una amplia variedad de distintas tecnologías y técnicas. Por ejemplo, los datos, las instrucciones, los comandos, la información, las señales, los bits, los símbolos y los chips que puedan haberse mencionado a lo largo de la descripción anterior pueden representarse mediante tensiones, corrientes, ondas electromagnéticas, campos o partículas magnéticos, campos o partículas ópticos o cualquier combinación de los mismos.

[0043] La descripción anterior de la divulgación se da a conocer para permitir que cualquier experto en la técnica realice o use la divulgación. Varias modificaciones a la divulgación, cayendo dichas modificaciones dentro del alcance de las reivindicaciones adjuntas, serán evidentes para los expertos en la técnica. Por tanto, la divulgación no está prevista para limitarse a los ejemplos y diseños descritos en el presente documento, sino que se le concede el alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un controlador de memoria caché (110) de un sistema de memoria caché compartida (100), que comprende:

medios para recibir una solicitud de acceso a memoria que comprende un identificador de calidad de servicio, QoS, QoSID, correspondiente a una clase de QoS (112(0), 112(1)) de una pluralidad de clases de QoS;

medios para determinar si la solicitud de acceso a memoria corresponde a una línea de memoria caché de una pluralidad de líneas de memoria caché del sistema de memoria caché compartida;

medios para seleccionar, como una partición objetivo, una de una partición mínima asignada a la clase de QoS correspondiente al QoSID mediante una máscara de bits de mapeado mínimo (200(0), 200(1)) y una partición máxima asignada a la clase de QoS correspondiente al QoSID mediante una máscara de bits de mapeado máximo (202(0), 202(1)), que responde a la determinación de que la solicitud de acceso a memoria no corresponde a la línea de memoria caché de la pluralidad de líneas de memoria caché; y

medios para asignar la línea de memoria caché de la pluralidad de líneas de memoria caché dentro de la partición objetivo para una operación de llenado de memoria caché, que responde a la determinación de que la solicitud de acceso a memoria no corresponde a la línea de memoria caché de la pluralidad de líneas de memoria caché.

2. El controlador de memoria caché del sistema de memoria caché compartida de acuerdo con la reivindicación 1, que comprende además un circuito de asignación de memoria caché (116) que comprende:

la máscara de bits de mapeado mínimo para mapear la clase de calidad de servicio, QoS, de la pluralidad de clases de QoS a la partición mínima de la pluralidad de líneas de memoria caché; la máscara de bits de mapeado máximo para mapear la clase de QoS de la pluralidad de clases de QoS a la partición máxima de la pluralidad de líneas de memoria caché; por lo que el circuito de asignación de memoria caché comprende además:

los medios para recibir la solicitud de acceso a memoria que comprenden el identificador de QoS,

los medios para determinar,

los medios para seleccionar y

los medios para asignar.

3. El controlador de memoria caché de la reivindicación 2, en el que el circuito de asignación de memoria caché está configurado para asignar la línea de memoria caché dentro de la partición objetivo de la siguiente manera:

identificar una o más líneas de memoria caché de la pluralidad de líneas de memoria caché dentro de la partición objetivo como uno o más candidatos de expulsión; y

expulsar a un candidato de expulsión del uno o más candidatos de expulsión del sistema de memoria caché compartida

4. El controlador de memoria caché de la reivindicación 2, en el que:

el circuito de asignación de memoria caché está configurado además para proporcionar un valor objetivo de asignación promedio;

el circuito de asignación de memoria caché está configurado para seleccionar, como la partición objetivo, la una de la partición mínima y la partición máxima para hacer que una asignación real de líneas de memoria caché asignadas a la clase de QoS correspondiente al QoSID se acerque al valor objetivo de asignación promedio y opcionalmente en el que

el controlador de memoria caché comprende además un monitor de uso de memoria caché para proporcionar una indicación de una asignación de la pluralidad de líneas de memoria caché entre la pluralidad de clases de QoS; y

el circuito de asignación de memoria caché está configurado para seleccionar, como la partición objetivo, la una de la partición mínima y la partición máxima en función de la indicación proporcionada por el monitor de uso de memoria caché.

- 5 **5.** El controlador de memoria caché de la reivindicación 2, en el que el circuito de asignación de memoria caché está configurado para seleccionar, como la partición objetivo, la una de la partición mínima y la partición máxima seleccionando de manera probabilística la una de la partición mínima y la partición máxima, y opcionalmente comprendiendo además un circuito proveedor de función de probabilidad en el que el circuito de asignación de memoria caché está configurado para seleccionar de manera probabilística la una de la partición mínima y la partición máxima en base al circuito proveedor de función de probabilidad.
- 10
- 15 **6.** El controlador de memoria caché de la reivindicación 2, en el que la máscara de bits de mapeado mínimo y la máscara de bits de mapeado máximo mapean la clase de QoS de la pluralidad de clases de QoS en una misma partición de la pluralidad de líneas de memoria caché que es exclusiva de la clase de QoS.
- 20 **7.** El controlador de memoria caché de la reivindicación 2, en el que:

 la máscara de bits de mapeado mínimo mapea la clase de QoS de la pluralidad de clases de QoS a la partición mínima que es exclusiva de la clase de QoS; y

 la máscara de bits de mapeado máximo mapea la clase de QoS de la pluralidad de clases de QoS a la partición máxima que no es exclusiva de la clase de QoS.
- 25 **8.** El controlador de memoria caché de la reivindicación 1 integrado en un circuito integrado, IC y/o un dispositivo seleccionado de un grupo que consiste en: un decodificador; una unidad de entretenimiento; un dispositivo de navegación; un dispositivo de comunicaciones; una unidad de datos de ubicación fija; una unidad de datos de ubicación móvil; un teléfono móvil; un teléfono celular; un ordenador; un ordenador portátil; un ordenador de sobremesa; un asistente digital personal, PDA; un monitor; un monitor de ordenador; un televisor; un sintonizador; una radio; una radio por satélite; un reproductor de música; un reproductor de música digital; un reproductor de música portátil; un reproductor de vídeo digital; un reproductor de vídeo; un reproductor de disco de vídeo digital, DVD; y un reproductor de vídeo digital portátil.
- 30
- 35 **9.** Un procedimiento para asignar líneas de memoria caché de un sistema de memoria caché compartida, que comprende:

 recibir, mediante un circuito de asignación de memoria caché de un controlador de memoria caché, una solicitud de acceso a memoria que comprende un identificador de calidad de servicio, QoS, QoSID correspondiente a una clase de QoS de una pluralidad de clases de QoS;

 determinar si la solicitud de acceso a memoria corresponde a una línea de memoria caché de una pluralidad de líneas de memoria caché del sistema de memoria caché compartida; y

 en respuesta a la determinación de que la solicitud de acceso a memoria no se corresponde con la línea de memoria caché de la pluralidad de líneas de memoria caché:

 seleccionar, como una partición objetivo, una de una partición mínima asignada a la clase de QoS correspondiente a la QoSID mediante una máscara de bits de mapeado mínimo y una partición máxima asignada a la clase de QoS correspondiente al QoSID mediante una máscara de bits de mapeado máximo; y

 asignar la línea de memoria caché de la pluralidad de líneas de memoria caché dentro de la partición objetivo para una operación de llenado de memoria caché.
- 40
- 45
- 50
- 55 **10.** El procedimiento de la reivindicación 9, en el que la asignación de la línea de memoria caché de la pluralidad de líneas de memoria caché dentro de la partición objetivo comprende:

 identificar una o más líneas de memoria caché dentro de la partición objetivo como uno o más candidatos de expulsión; y

 desalojar a un candidato de expulsión del uno o más candidatos de expulsión del sistema de memoria caché compartida.
- 60
- 65 **11.** El procedimiento de la reivindicación 9, que comprende además proporcionar un valor objetivo de asignación promedio;

- 5 en el que la selección, como la partición objetivo, de la una de la partición mínima y la partición máxima comprende seleccionar la una de la partición mínima y la partición máxima para hacer que una asignación real de líneas de memoria caché asignadas a la clase de QoS correspondiente al QoSID se acerque al valor objetivo de asignación promedio y opcionalmente en el que la selección, como la partición objetivo, la una de la partición mínima y la partición máxima comprende seleccionar la una de la partición mínima y la partición máxima en base a una indicación de una asignación de la pluralidad de líneas de memoria caché entre la pluralidad de clases de QoS proporcionadas por un monitor de uso de memoria caché.
- 10 **12.** El procedimiento de la reivindicación 9, en el que la selección, como partición objetivo, de la una de la partición mínima y la partición máxima comprende seleccionar de manera probabilística la una de la partición mínima y la partición máxima y, opcionalmente, en el que la selección de forma probabilística de la una de la partición mínima y la partición máxima se basa en un circuito proveedor de función de probabilidad.
- 15 **13.** El procedimiento de la reivindicación 9, en el que la máscara de bits de mapeado mínimo y la máscara de bits de mapeado máximo mapean la clase de QoS de la pluralidad de clases de QoS en una misma partición de la pluralidad de líneas de memoria caché que es exclusiva de la clase de QoS.
- 20 **14.** El procedimiento de la reivindicación 9, en el que:
- la máscara de bits de mapeado mínimo mapea la clase de QoS de la pluralidad de clases de QoS a la partición mínima exclusiva de la clase de QoS; y
- 25 la máscara de bits de mapeado máximo mapea la clase de QoS de la pluralidad de clases de QoS a la partición máxima que no es exclusiva de la clase de QoS.

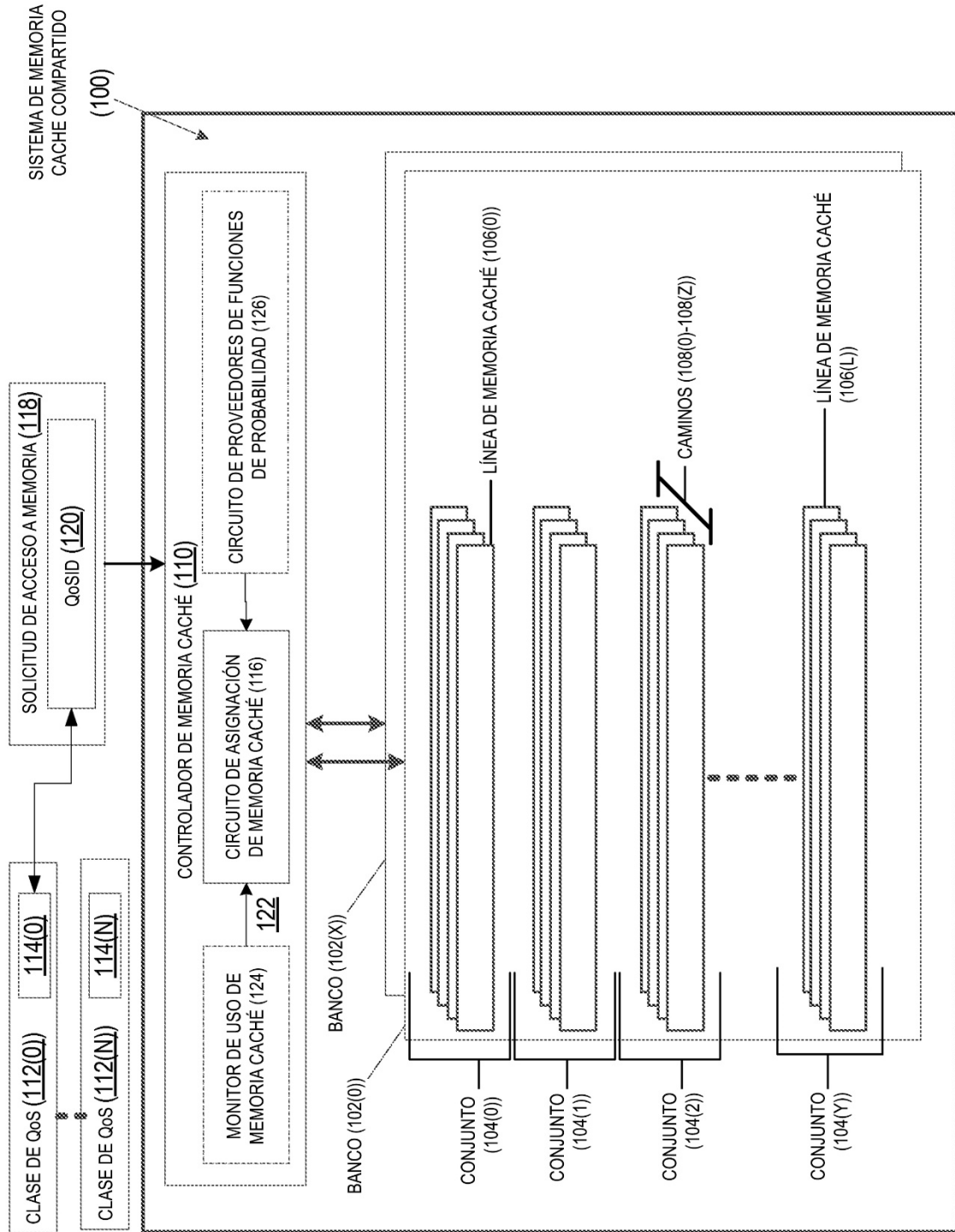
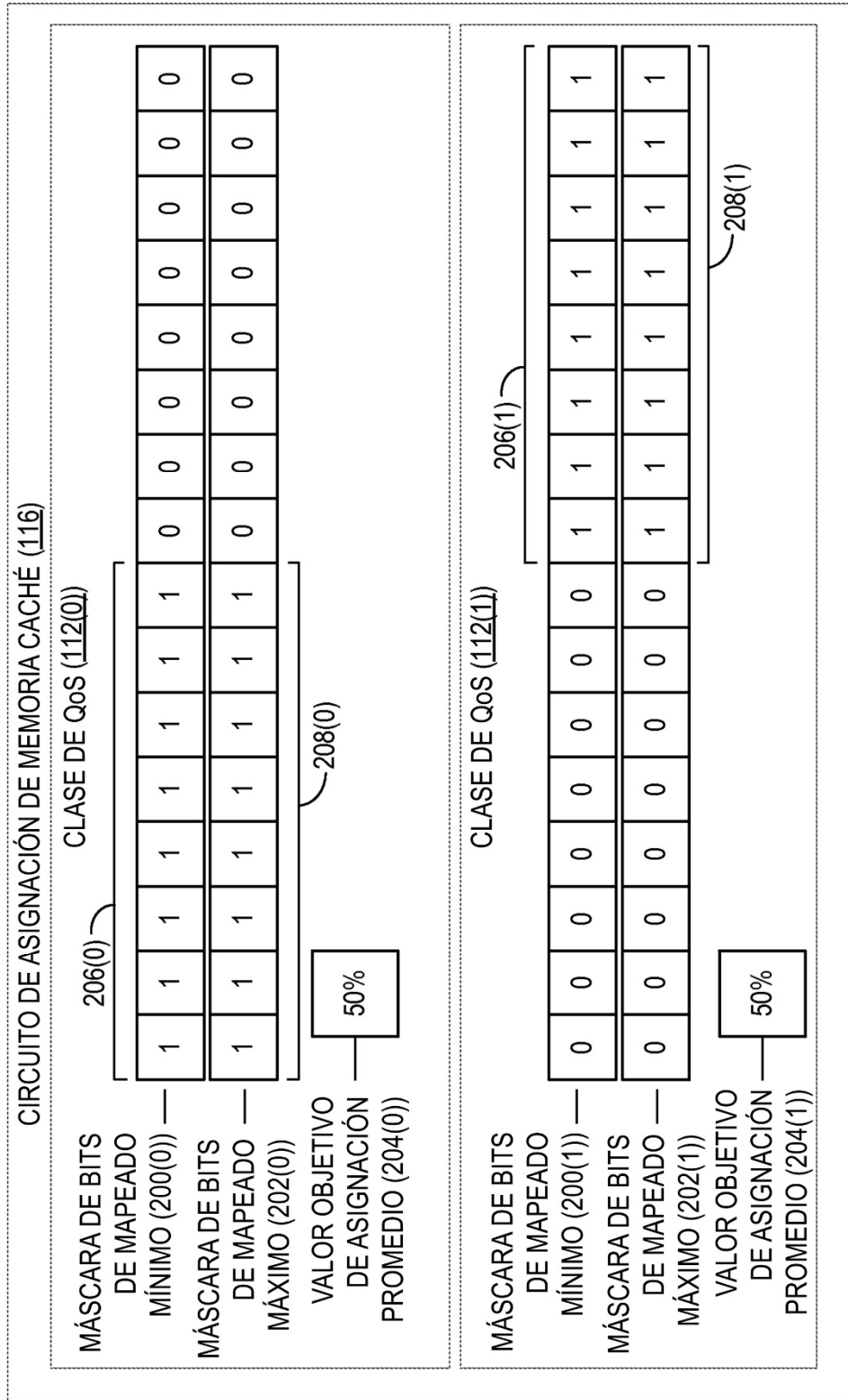
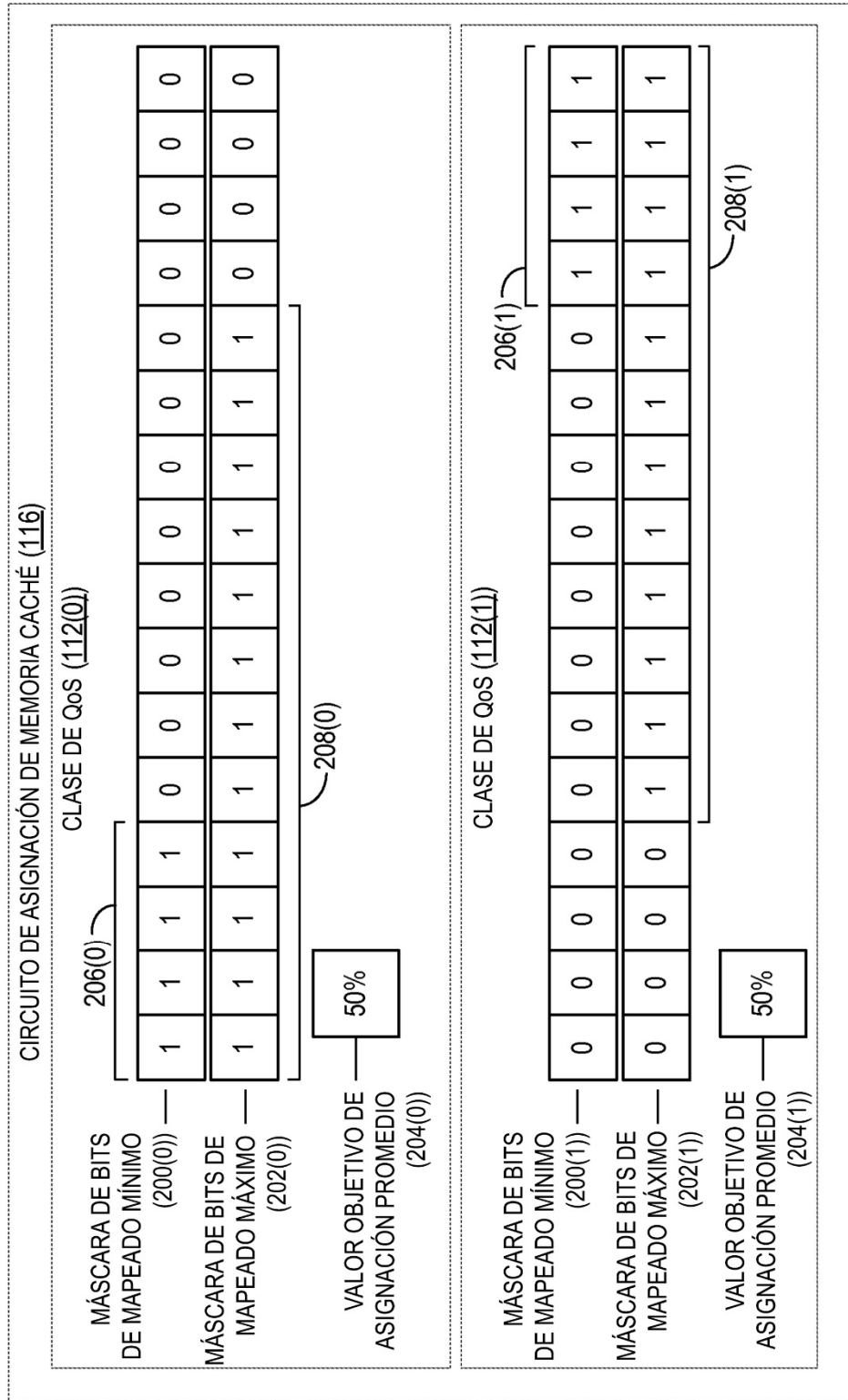


FIG. 1





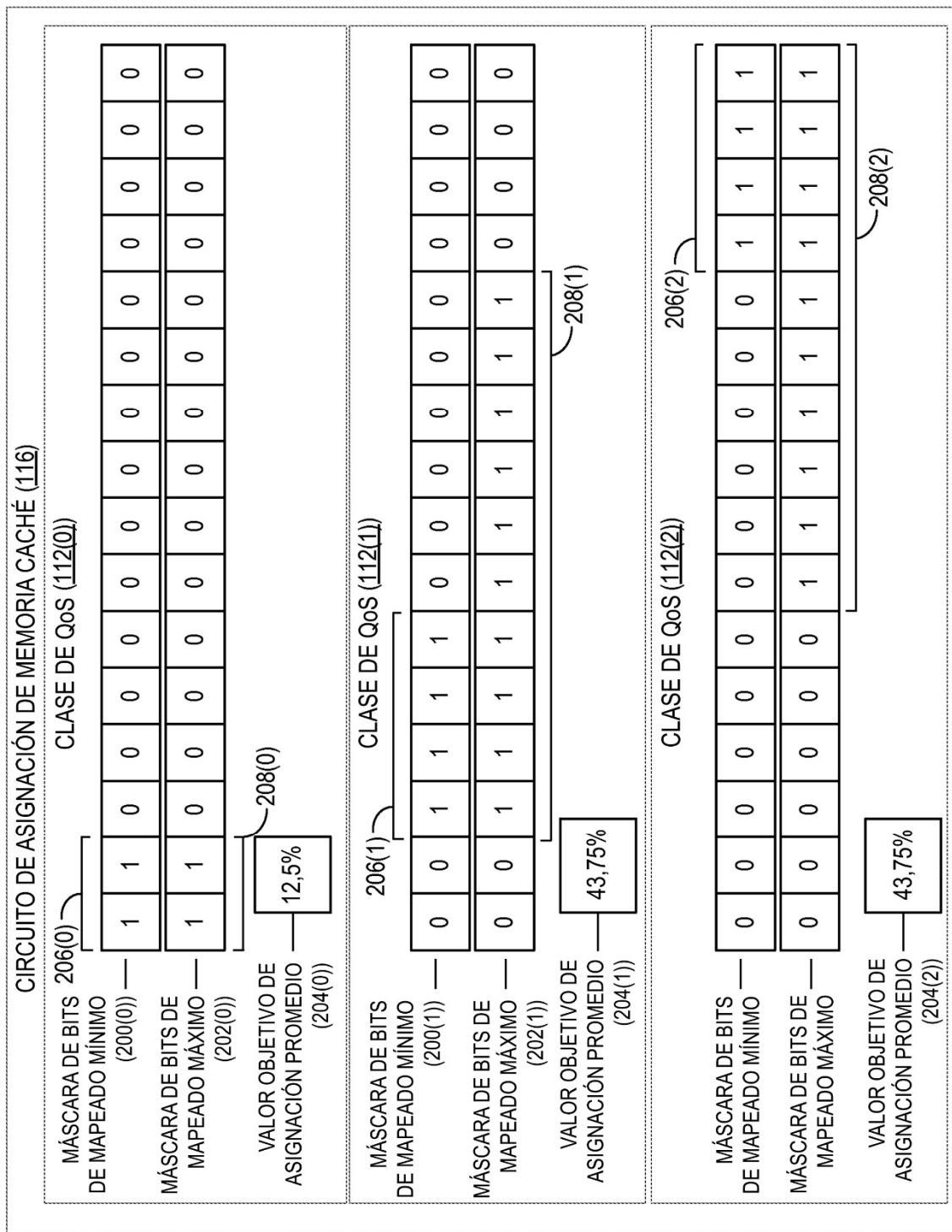


FIG. 2C

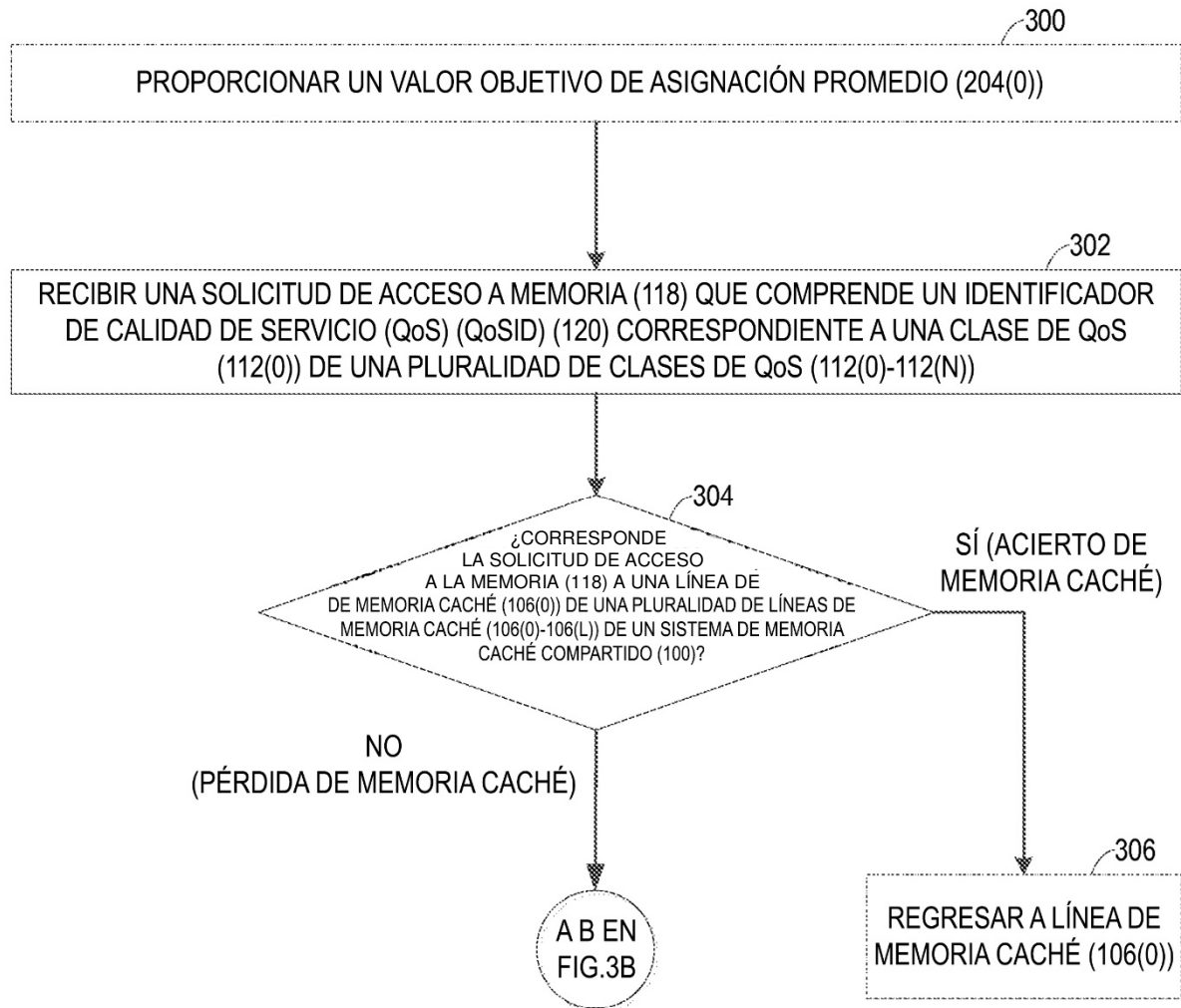


FIG. 3A

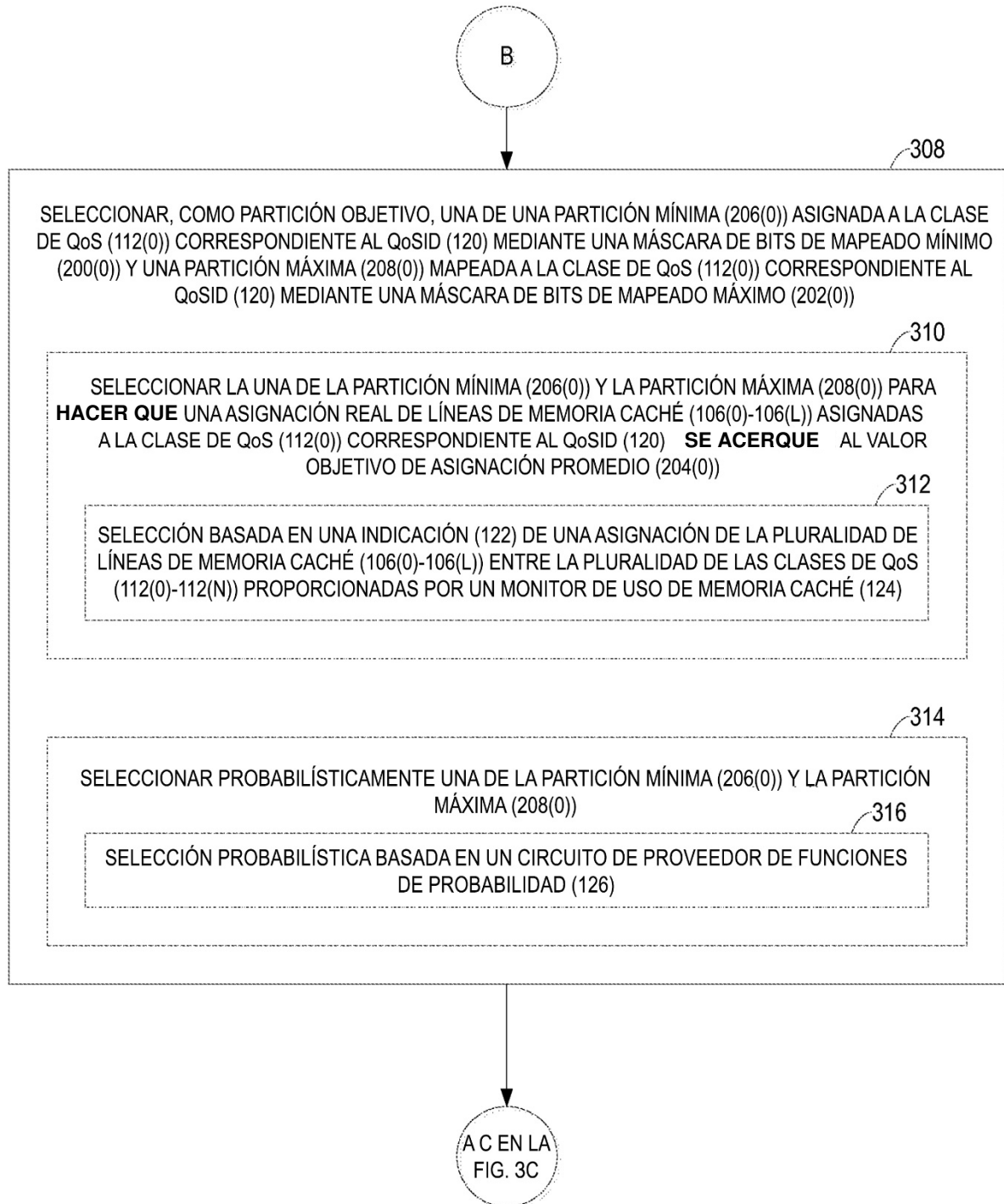


FIG. 3B

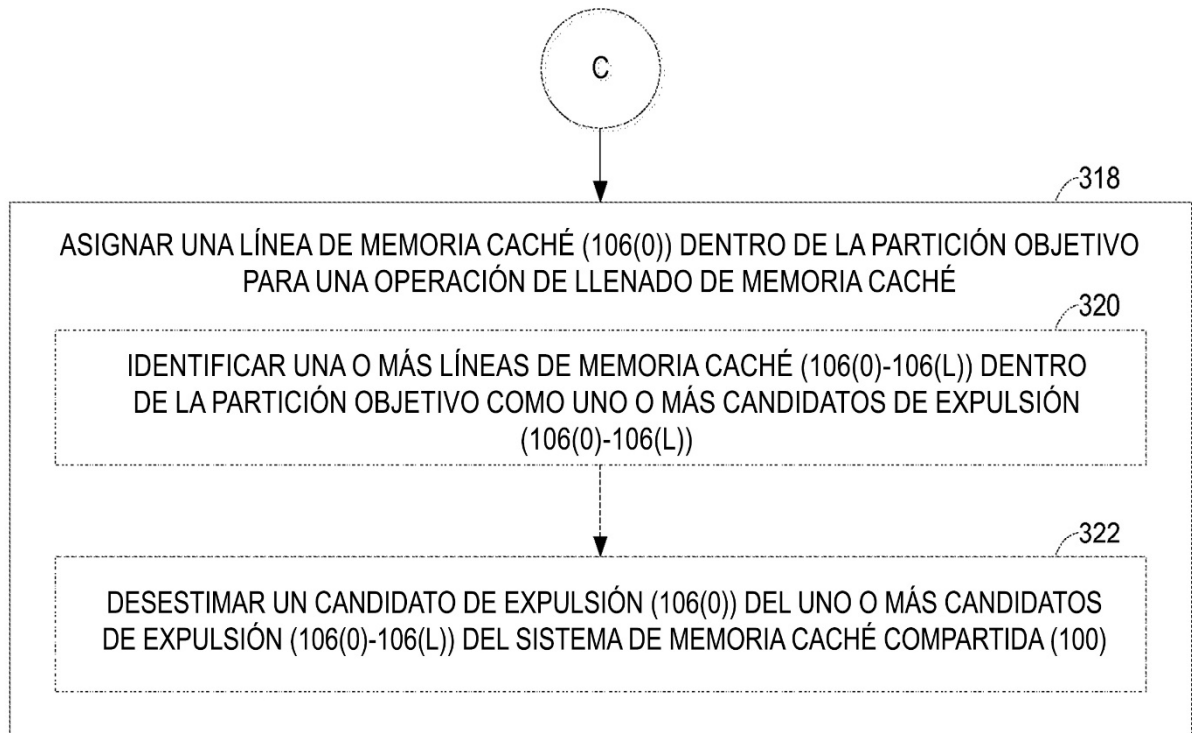


FIG. 3C

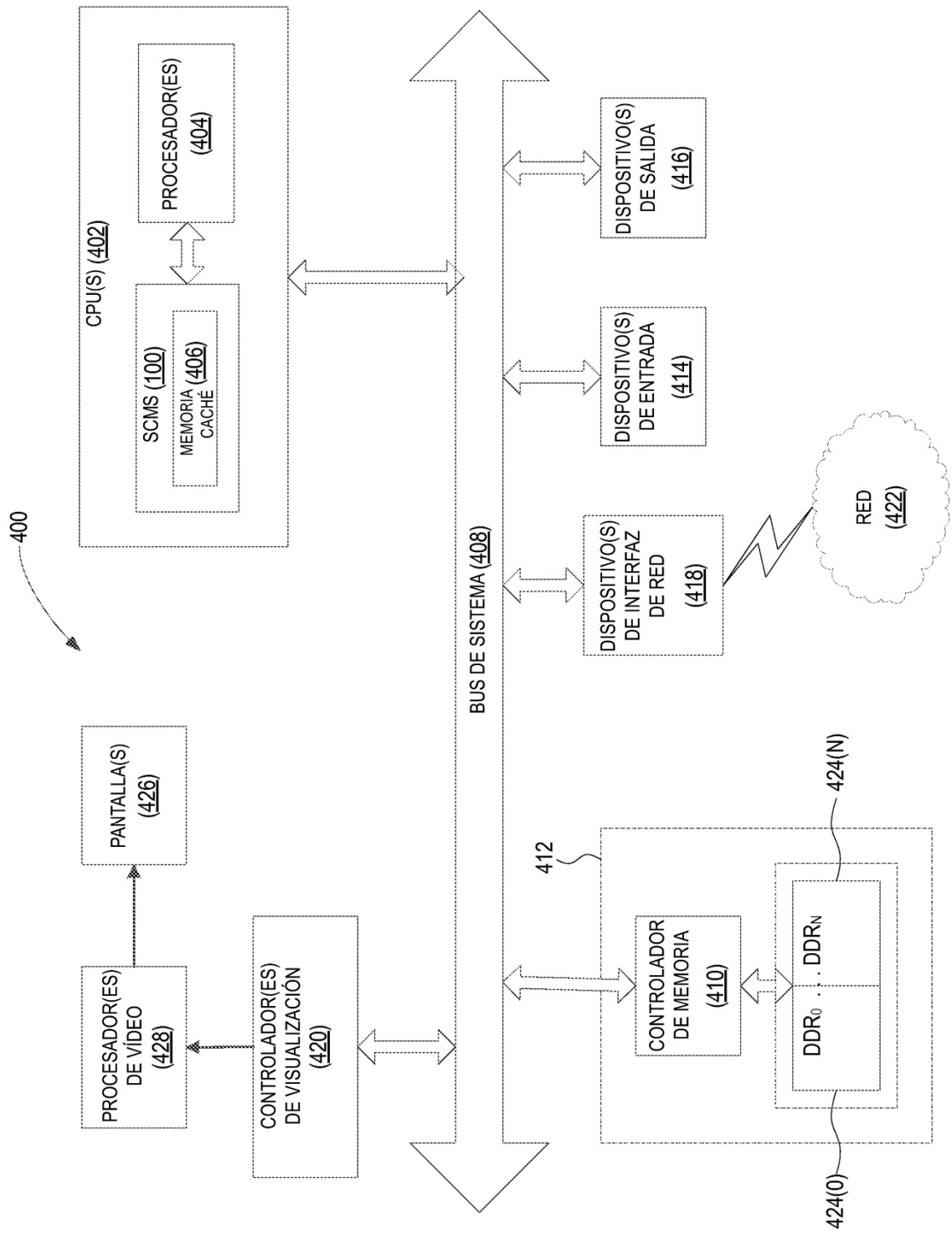


FIG. 4