

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 971 138**

51 Int. Cl.:

G06F 1/02

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.08.2018** **PCT/US2018/046590**

87 Fecha y número de publicación internacional: **21.03.2019** **WO19055164**

96 Fecha de presentación y número de la solicitud europea: **14.08.2018** **E 18762704 (7)**

97 Fecha y número de publicación de la concesión europea: **27.12.2023** **EP 3682331**

54 Título: **Escrituras persistentes para memoria no volátil**

30 Prioridad:

15.09.2017 US 201715706530

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

03.06.2024

73 Titular/es:

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

72 Inventor/es:

RAMANUJAN, RAJ;
BAINS, KULJIT SINGH;
WANG, LIYONG y
QUEEN, WESLEY

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 971 138 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

Escrituras persistentes para memoria no volátil

5 Campo de divulgación

Los aspectos divulgados están dirigidos a sistemas de memoria. Más particularmente, aspectos de ejemplo están dirigidos a operaciones de escritura persistente y sus protocolos para memoria no volátil.

10 Antecedentes

La memoria de clase de almacenamiento (SCM) generalmente se refiere a la memoria de alta capacidad que también puede tener un alto rendimiento. SCM puede usarse en aplicaciones tales como servidores u otros sistemas de procesamiento en donde un conjunto operativo de datos para un procesador o unidad central de procesamiento puede almacenarse en la SCM, mientras que el conjunto de datos completo puede almacenarse en una memoria de respaldo o en una unidad de disco duro (HDD). Una expectativa importante de la SCM es la persistencia de las escrituras, lo que significa que la información escrita en la SCM no se perderá si, por ejemplo, el servidor falla o se queda sin potencia. Sin embargo, es posible que la memoria no volátil convencional, que puede cumplir con las expectativas relacionadas con la persistencia, no pueda cumplir con las métricas de capacidad y rendimiento que se pueden desear de SCM. Por lo tanto, tecnologías tal como la memoria de cambio de fase (PCM), la memoria de acceso aleatorio magnético de par de transferencia de giro (STT MRAM), la RAM resistiva (ReRAM), etc., se están volviendo más populares en las implementaciones de SCM.

Cuando se utiliza SCM, una aplicación puede utilizar operaciones de escritura en memoria para actualizar la memoria persistente correspondiente. Para que una escritura en la SCM sea persistente, la aplicación que solicita la operación de escritura puede esperar una confirmación explícita de que la operación de escritura ha alcanzado la memoria persistente. Por el contrario, las operaciones de escritura en memoria no persistente (tal como la memoria dinámica de acceso aleatorio (DRAM) u otra memoria volátil) se consideran convencionalmente completadas o publicadas, desde la perspectiva de la aplicación una vez que la operación de escritura y los datos asociados se han transferido a la memoria y no se requiere confirmación explícita de que los datos se han escrito. Por tanto, para aplicaciones que utilizan SCM con una expectativa de persistencia, son deseables técnicas de alto rendimiento que proporcionen confirmación explícita de operaciones de escritura en la memoria persistente, en donde las técnicas de alto rendimiento también son compatibles con diferentes tamaños de datos para maximizar la eficiencia.

Hay dos tipos de esquemas convencionales para operaciones de memoria persistente. Un primer esquema supone que todo el sistema de memoria (por ejemplo, un módulo de memoria dual en línea (DIMM) que comprende una serie de circuitos integrados DRAM, como se conoce en la técnica) está respaldado por energía. En este caso, una operación de escritura en un regulador intermedio en el extremo receptor del DIMM puede ser suficiente para satisfacer las expectativas de persistencia. En una implementación, una vez que se completa con éxito una operación de escritura a través de una interfaz de canal entre la aplicación que solicita la operación de escritura y el DIMM, la operación de escritura puede considerarse persistente. Sin embargo, la implementación de dichos esquemas puede implicar el uso de dispositivos de almacenamiento de energía, tales como supercondensadores o baterías, que proporcionan potencia/carga para descargar los reguladores intermedios del DIMM cuando se detecta un corte de potencia. Pero es posible que estos dispositivos de almacenamiento de energía no estén disponibles en todos los DIMM y, además, incluso si estuvieran disponibles, tienen un coste elevado.

En un segundo esquema, todas las operaciones de escritura anteriores se pueden descargar en la memoria persistente mientras la aplicación espera un estado de finalización del DIMM. Sin embargo, este esquema puede generar un coste de rendimiento significativo. Por ejemplo, en los casos en donde la aplicación puede estar solicitando escrituras persistentes de granularidad fina en el DIMM pero puede haber otras operaciones de escritura simultáneas pero independientes que se transmiten al DIMM, descargar todas las operaciones de escritura anteriores en la memoria persistente en espera de un estado de finalización puede ralentizar no solo las solicitudes de escritura persistente sino también las operaciones de escritura simultáneas.

En consecuencia, existe una necesidad en la técnica de operaciones de escritura persistente de alto rendimiento y alta eficiencia que soporten diferentes granularidades o tamaños de las escrituras persistentes, evitando al mismo tiempo los inconvenientes antes mencionados de los enfoques convencionales. El documento US 2014/136786 describe almacenes persistentes asíncronos para transacciones. El documento US 2017/147207 describe un regulador no volátil para operaciones de memoria. El documento US 2016/034225 describe una jerarquía de memoria no volátil multiversionada para memoria persistente. El documento US 2010/070631 describe el sistema y el método para la gestión de datos compartidos de forma remota. El documento US 10 061 719 describe finalizaciones de alambre empacado.

Resumen

La invención se expone en el conjunto de reivindicaciones adjunto. Las reivindicaciones dependientes establecen realizaciones particulares.

Los aspectos de ejemplo de la invención incluyen sistemas y métodos para operaciones persistentes. Un sistema de memoria, al recibir un comando de escritura persistente y datos de escritura asociados desde un servidor, realiza una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente. El sistema de memoria recibe una identificación de escritura (WID) asociada con el comando de escritura persistente del servidor y proporciona, al finalizar exitosamente la escritura persistente, una indicación de finalización de escritura persistente junto con la WID asociado al servidor.

Por ejemplo, un aspecto de ejemplo está dirigido a un método para realizar operaciones persistentes, el método que comprende recibir, en un sistema de memoria, un comando de escritura persistente y datos de escritura asociados desde un servidor, y realizar una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente.

Otro aspecto de ejemplo está dirigido a un método para realizar operaciones persistentes, el método que comprende proporcionar, desde un servidor a un sistema de memoria, un comando de escritura persistente y datos de escritura asociados, en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil.

Otro aspecto de ejemplo está dirigido a un aparato que comprende un sistema de memoria configurado para recibir un comando de escritura persistente y datos de escritura asociados desde un servidor, y realizar una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente.

Otro aspecto de ejemplo está dirigido a un aparato que comprende un servidor configurado para proporcionar un comando de escritura persistente y datos de escritura asociados a un sistema de memoria, en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil.

Otro aspecto de ejemplo está dirigido a un aparato que comprende medios para almacenar datos, que comprende medios para recibir un comando de escritura persistente y datos de escritura asociados desde un servidor, y medios para realizar una escritura persistente de los datos de escritura en una memoria no volátil en los medios para almacenar, con base en el comando de escritura persistente.

Otro aspecto de ejemplo está dirigido a un aparato que comprende un medio para procesar, que comprende medios para proporcionar un comando de escritura persistente y datos de escritura asociados a un sistema de memoria, en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil.

Otro aspecto de ejemplo está dirigido a un medio de almacenamiento no transitorio legible por ordenador que comprende código, que, cuando es ejecutado por un procesador, hace que el procesador realice operaciones persistentes, el medio de almacenamiento transitorio legible por ordenador que comprende código para recibir, en un sistema de memoria, un comando de escritura persistente y datos de escritura asociados desde un servidor, y código para realizar una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente.

Otro aspecto de ejemplo está dirigido a un medio de almacenamiento no transitorio legible por ordenador que comprende código, que, cuando es ejecutado por un procesador, hace que el procesador realice operaciones persistentes, el medio de almacenamiento transitorio legible por ordenador que comprende código para proporcionar, desde un servidor a un sistema de memoria, un comando de escritura persistente y datos de escritura asociados, en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil.

Breve descripción de los dibujos

Los dibujos adjuntos se presentan para ayudar en la descripción de aspectos de la invención y se proporcionan únicamente para ilustrar los aspectos y no limitarlos.

La figura 1 ilustra un sistema de procesamiento de acuerdo con aspectos de esta divulgación.

Las figuras 2A-C ilustran transacciones para manejar escrituras persistentes, de acuerdo con diversos aspectos de esta divulgación.

La figura 3 ilustra un ejemplo de codificación para un comando de escritura persistente de acuerdo con esta divulgación.

Las figuras 4A-B ilustran secuencias de eventos pertenecientes a métodos de ejemplo para realizar escrituras persistentes, de acuerdo con aspectos de esta divulgación.

La figura 5 representa un dispositivo informático de ejemplo en el que se puede emplear ventajosamente un aspecto de la divulgación.

Descripción detallada

Aspectos de la invención se divulgan en la siguiente descripción y dibujos relacionados dirigidos a aspectos específicos de la invención. Se pueden idear aspectos alternativos sin apartarse del alcance de la invención. Además, los elementos bien conocidos de la invención no se describirán en detalle o se omitirán para no oscurecer los detalles relevantes de la invención.

La expresión "de ejemplo" se utiliza en este documento con el significado de "que sirve como ejemplo, instancia o ilustración". Cualquier aspecto descrito en este documento como "de ejemplo" no necesariamente debe interpretarse como preferido o ventajoso sobre otros aspectos. Asimismo, la expresión "aspectos de la invención" no requiere que todos los aspectos de la invención incluyan la característica, ventaja o modo de operación discutidos.

La terminología utilizada en este documento tiene el propósito de describir aspectos particulares únicamente y no pretende ser limitativa de aspectos de la invención. Como se utilizan en este documento, las formas singulares "un", "uno, una" y "el, la" pretenden incluir también las formas plurales, a menos que el contexto indique claramente lo contrario. Se entenderá además que las expresiones "comprende", "que comprende", "incluye" y/o "que incluye", cuando se utiliza en este documento, especifica la presencia de características, números enteros, pasos, operaciones, elementos y/o componentes indicados, pero no excluye la presencia o adición de una o más características diferentes, números enteros, pasos, operaciones, elementos, componentes y/o grupos de los mismos.

Además, muchos aspectos se describen en términos de secuencias de acciones a realizar, por ejemplo, por elementos de un dispositivo informático. Se reconocerá que diversas acciones descritas en este documento pueden realizarse mediante circuitos específicos (por ejemplo, circuitos integrados de aplicación específica (ASIC)), mediante instrucciones de programa ejecutadas por uno o más procesadores, o mediante una combinación de ambos. Además, se puede considerar que esta secuencia de acciones descritas en este documento está incorporada en su totalidad dentro de cualquier forma de medio de almacenamiento legible por ordenador que tenga almacenado en él un conjunto correspondiente de instrucciones de ordenador que, tras su ejecución, provocarían que un procesador asociado realice la funcionalidad descrita en este documento. Por lo tanto, los diversos aspectos de la invención pueden realizarse en un número de formas diferentes, todas las cuales se han contemplado dentro del alcance del objeto reivindicado. Además, para cada uno de los aspectos descritos en este documento, la forma correspondiente de cualquiera de dichos aspectos puede describirse en este documento como, por ejemplo, "lógica configurada para" realizar la acción descrita.

Los aspectos de ejemplo de esta divulgación están dirigidos a operaciones de escritura persistente eficientes y de alto rendimiento para memoria no volátil tal como DIMM no volátil (o NVDIMM). De manera correspondiente, se divulga un NVDIMM persistente o NVDIMM-P como un sistema de memoria de ejemplo que soporta operaciones de escritura persistente de acuerdo con aspectos de ejemplo. Un dispositivo servidor puede configurarse para proporcionar solicitudes/comandos de ejemplo, por ejemplo, para operaciones persistentes, y los datos correspondientes a un sistema de memoria de ejemplo, y el sistema de memoria puede configurarse para realizar las operaciones de escritura persistente solicitadas y proporcionar la señalización correspondiente al dispositivo host, como se analizará con más detalle en las siguientes secciones.

Con referencia ahora a la figura 1, se muestra un sistema 100 de procesamiento de ejemplo que comprende un servidor 120 y un sistema 130 de memoria. El servidor 120 puede comprender uno o más elementos de procesamiento tales como una unidad central de procesamiento (CPU), un procesador de señales digitales (DSP), procesador multimedia, procesador de sistema, unidad de procesamiento gráfico (GPU), modulador-demodulador (módem), procesador de aplicaciones, etc., aunque no hayan sido ilustrados explícitamente. Estos elementos de procesamiento pueden realizar solicitudes para acceder al sistema 130 de memoria. Un controlador de memoria (no mostrado) puede estar presente en el servidor 120 para controlar estas solicitudes de acceso.

El sistema 130 de memoria puede ser una memoria persistente, por ejemplo, un NVDIMM-P de acuerdo con esta divulgación. Se muestra que el sistema 130 de memoria incluye el bloque 132 de entrada/salida (I/O) y el banco 134 de memoria. El banco 134 de memoria puede incluir memoria flash, DRAM, etc.

Se muestra la interconexión 110 entre el servidor 120 y el sistema 130 de memoria, con el bus de datos (DQ) 112, el bus de comando y dirección (CA) 114 y la respuesta 116 identificados por separado. El servidor 120 puede proporcionar comandos y direcciones relacionadas para solicitudes de acceso a memoria a través de CA 114 y enviar/recibir datos a través de DQ 112 (mostrado como un bus bidireccional). La respuesta 116, aunque se muestra por separado, puede configurarse como parte del CA 114 y puede implementarse como un bus bidireccional en algunos casos. La respuesta 116 se puede utilizar para proporcionar información tal como el estado de las escrituras persistentes en algunos aspectos de ejemplo. También pueden estar presentes otros buses/cables en la interconexión 110, aunque no se han identificado por separado. En algunos casos, el sistema 130 de memoria puede usar buses separados para respuestas deterministas y no deterministas, que se explicarán más adelante.

En una implementación en donde el sistema 130 de memoria puede configurarse como un NVDIMM, con soporte adicional para una configuración de NVDIMM persistente (NVDIMM-P) para al menos algunas operaciones, el servidor 120 puede ser capaz de proporcionar uno o más de los siguientes comandos de ejemplo al sistema 130 de memoria, por ejemplo, en CA 114:

- Comando leer (por ejemplo, con codificación de longitud en múltiplos de 64B), junto con una identificación de lectura (RID);
- Comando escribir (por ejemplo, un comando de escritura convencional);
- Comando escribir-p (por ejemplo, un comando de escritura persistente, junto con una identificación de escritura (WID) para escrituras persistentes, junto con un bit de persistencia que indica cuándo las escrituras con un WID determinado necesitan una señal de escritura persistente completa (W_PER) (por ejemplo, que se proporcionará en la respuesta 116) desde el sistema 130 de memoria);
- Comando ADRx: direccionamiento extendido;
- Comando enviar (por ejemplo, un comando para que el sistema 130 de memoria proporcione el estado de una solicitud de lectura de datos);
- Comando enviar_estado (por ejemplo, un comando para que el sistema 130 de memoria proporcione lecturas de errores, WID, etc. Relacionados con operaciones persistentes desde el sistema 130 de memoria);
- Comando descargar (para descargar las escrituras anteriores y enviarlas a la memoria persistente)
- NOP (sin operación);
- Comando de lectura especulativa (por ejemplo, utilizado para leer la memoria caché); y
- Otros comandos de almacenamiento en caché, que pueden ser específicos de la implementación.

Como se mencionó anteriormente, se pueden proporcionar buses separados en la interconexión 110 para respuestas deterministas y no deterministas desde el sistema 130 de memoria al servidor 120. Las respuestas deterministas incluyen metadatos, información de error/paridad tal como cifrado de control de errores (ECC) relacionada con la lectura de datos enviados en el DQ 112 al servidor 120, etc., que pueden multiplexarse en buses que emanan de clavijas acopladas a la I/O 132, tales como clavijas de bit de verificación.

Entre los bits ECC, puede haber medios ECC específicos para implementaciones del sistema 130 de memoria (por ejemplo, como NVDIMM) y bits ECC específicos de canal en DQ 112, por ejemplo, que pueden estandarizarse para permitir la compatibilidad cruzada entre diversas implementaciones.

Los bits de metadatos pueden incluir RID retrasados para solicitudes de lectura enviadas fuera del orden del programa (en donde, para operaciones en orden, la RID puede establecerse en un estado de "no importa"). Los bits de metadatos también pueden incluir un crédito de escritura (WC), que se refiere a una cuota no utilizada para operaciones de escritura asignadas a ciertos servidores o elementos de procesamiento del servidor 120. Los bits de metadatos pueden incluir además bits de envenenamiento de datos de un equipo de usuario como se conoce en la técnica, y otros bits definidos por el usuario.

Las respuestas no deterministas de acuerdo con esta divulgación pueden pertenecer a operaciones persistentes y pueden enviarse a través de señalización dedicada tal como la respuesta 116 desde el sistema 130 de memoria al servidor 120, y pueden indicar lo siguiente:

- R_RDY: una señal desde el sistema 130 de memoria al servidor 120 para indicar que los datos leídos están disponibles;
- Wr_Per: una señal desde el sistema 130 de memoria al servidor 120 para indicar que se ha completado una escritura persistente; y
- Error: una señal desde el sistema 130 de memoria al servidor 120 para indicar condiciones de error tales como verificación CRC, violación de crédito, tiempo de espera de medios, etc.

Para una implementación del sistema 130 de memoria como NVDIMM-P, se pueden definir las siguientes clavijas de protocolo en la I/O 132, por ejemplo. Usando la modulación de ancho de pulso, se pueden implementar las siguientes configuraciones de clavijas y señales de respuesta. Por ejemplo, en una implementación de NVDIMM-P de tasa 5 de datos doble (DDR5), se puede usar un solo alambre etiquetado como RSP_n (uno dedicado por subcanal) para proporcionar la siguiente señalización: 2 pulsos de reloj bajos para R_RDY, 4 pulsos de reloj bajos para W_PER y 6 pulsos de reloj bajos para mensaje. Cada pulso bajo puede ir seguido de al menos 2 pulsos altos de reloj. Si se necesita una señal de error separada, se puede definir como un pulso bajo de 8 relojes.

Para una implementación DDR4: se pueden usar dos clavijas para solucionar problemas de rendimiento con una sola clavija (ODT1 y CKE1), en donde ODT1 representa un ancho de pulso bajo de 2 relojes para R_RDY y un pulso bajo de 4 relojes para mensaje, y CKE1 representa un pulso bajo de 2 relojes para W_PER. Cada pulso bajo puede ir seguido de al menos 2 pulsos altos de reloj, y si se necesita una señal de error separada, entonces puede definirse como un pulso bajo de 6 relojes en ODT1.

En implementaciones de ejemplo de escrituras persistentes, se pueden configurar combinaciones adecuadas de hardware, software, firmware, etc., (por ejemplo, aplicaciones, accionadores, etc.) para permitir que se proporcionen notificaciones al servidor 120 desde el sistema 130 de memoria cuando una o más solicitudes de escritura desde el servidor 120 al sistema 130 de memoria logran persistencia. Estas notificaciones pueden ser específicas de la implementación, como se explica a continuación.

Cuando los datos que se van a escribir para una operación de escritura alcanzan un regulador protegido contra fallas de potencia en un controlador de medios (por ejemplo, una memoria protegida contra fallas de potencia del sistema 130 de memoria), la operación de escritura puede considerarse persistente durante las operaciones normales. Sin embargo, en ciertos casos poco frecuentes o cuando los reguladores del controlador de medios no están protegidos contra fallas de potencia, el software garantizará que las escrituras se envíen hasta el medio NVM.

Para una implementación del sistema 130 de memoria como NVDIMM-P, los DIMM respaldados por energía implican configuraciones en donde los reguladores antes mencionados están protegidos contra fallas de potencia, lo que significa que el comando de escritura NVDIMM-P se puede utilizar incluso cuando se requiere persistencia en los casos normales. Además, se puede utilizar un comando NVDIMM-P descargar, como se define en este documento, para descargar todas las escrituras en los reguladores del controlador de medios a la memoria no volátil. En el caso del comando descargar, solo se garantiza que las escrituras que ocurrieron antes del descargar se harán persistentes en la memoria no volátil. Los comandos implementados por software, tales como las operaciones de barrera apropiadas, se pueden usar después de la última escritura y antes de que se emita el comando descargar para garantizar que se mantenga el orden correcto de las escrituras (por ejemplo, cuando el servidor 120 puede configurarse para enviar solicitudes de escritura persistente fuera del orden del programa).

Aunque los DIMM sin respaldo de energía pueden usarse con menos frecuencia que los DIMM con respaldo de energía discutidos anteriormente, el comando de escritura persistente NVDIMM-P también se puede usar cuando se requiere persistencia para los DIMM sin respaldo de energía. Se puede configurar un controlador de memoria del servidor 120, como se mencionó anteriormente, para determinar cuándo emitir el comando de escritura persistente. En este caso, se espera que el sistema 130 de memoria proporcione una notificación explícita cuando se complete la escritura persistente, como se explicará con referencia a la figura 2A. Además, también se puede utilizar un comando NVDIMM-P descargar como antes para descargar todas las escrituras (incluso las escrituras no persistentes) en la memoria no volátil.

Con referencia ahora a la figura 2A, se muestra un conjunto de transacciones de ejemplo entre el servidor 120 y el sistema 130 de memoria para ilustrar aspectos del comando de escritura persistente. Hay algunas características del comando de escritura persistente (Wr_Per) de ejemplo (o simplemente, "escritura persistente") que pueden ser comunes al comando de lectura descrito anteriormente desde el servidor 120. Estos incluyen una identificación de escritura común (WID), que puede ser un identificador de varios bits para identificar instrucciones de escritura específicas. Se muestra un conjunto de ejemplo de WID 210, que pueden tener hasta 16 bits de ancho cada uno que incluye un bit válido "Vld" (en consecuencia, pueden estar presentes hasta 31 WID en un paquete de comando de 64 bytes enviado en CA 114, por ejemplo). El comando de escritura persistente también puede tener un campo reservado en la codificación WID para las devoluciones de estado del comando descargar, que se explicará con más detalle en los siguientes pasajes.

En un aspecto, el servidor 120 puede configurarse para emitir una escritura persistente sólo cuando el servidor 120 tiene créditos de escritura persistente asociados disponibles. Los créditos de escritura persistente (similares a los créditos de lectura conocidos en la técnica) pueden determinarse durante la configuración y gestionarse por el servidor 120, y pueden reflejar un número de escrituras persistentes pendientes que el servidor 120 puede emitir.

Una vez emitido, el servidor 120 puede configurarse para realizar un seguimiento de las escrituras persistentes pendientes con base en sus respectivas WID 210. En la figura 2A (con referencia combinada a la figura 1), se muestran dos escrituras persistentes (escritura-P 1 con una primera dirección y WID, y escritura-P 2 con una segunda dirección y WID) etiquetadas 202a y 204a, emitidas desde el servidor 120 al sistema 130 de memoria en CA 114, junto con los datos respectivos, escribir los datos 202b y escribir los datos 204b en DQ 112, por ejemplo.

El sistema 130 de memoria está configurado para emitir una respuesta "Wr_Per" en la respuesta 116, para una escritura persistente particular, una vez que todos los datos para esa escritura persistente se han escrito en la memoria no volátil en el sistema 130 de memoria. Wr_Per 202c y Wr_Per 204c se muestran para las respectivas escrituras 202a y 204a persistentes. Sin embargo, se muestra que Wr_Per 202c y Wr_Per 204c se envían en un orden diferente al del sistema 130 de memoria que recibió las escrituras 202a y 204a persistentes para ilustrar que las respuestas no necesitan estar en el orden del programa o en el orden en que se reciben las solicitudes de escritura persistente desde el servidor 120. En un aspecto, el sistema 130 de memoria puede afirmar la señal "Req" en la respuesta 116 junto con la codificación apropiada para el mensaje "escritura_Rdy" para las respuestas Wr_Per.

Además, el servidor 120 también puede configurarse para emitir un comando "enviar estado para WID" designado con el numeral 206a de referencia, en cualquier momento, para determinar el estado de sus escrituras persistentes pendientes. En respuesta, el sistema 130 de memoria puede configurarse para emitir un paquete de estado con WID

de escrituras persistentes completadas, por ejemplo, en una longitud de ráfaga de 8 o transferencia "BL8" a través de DQ 112.

Como se mencionó anteriormente, se pueden empaquetar hasta 31 WID 210 en cada paquete de estado 64B, en donde para cada WID 210 pueden haber 16 bits asignados para el WID de 15 bits y el bit válido, combinados. Además, el sistema 130 de memoria también puede usar el campo de metadatos mencionado anteriormente para devolver el estado de otras escrituras. El servidor 120 puede usar los WID 210 devueltos en el paquete 206b de estado de WID, por ejemplo, para terminar el seguimiento de escrituras persistentes pendientes.

En algunos aspectos, se pueden agrupar dos o más escrituras persistentes. Por ejemplo, se puede agrupar un conjunto de escrituras persistentes de 64 B para confirmar (o escribir en memoria no volátil) en el caso de módulos DIMM sin respaldo de energía, por ejemplo. Una implementación de ejemplo puede implicar que se emita un bloque de escrituras persistentes al sistema 130 de memoria desde el servidor 120, en donde el sistema 130 de memoria puede configurarse para recopilar hasta el bloque de escrituras persistentes en un regulador y confirmar todo el bloque de escrituras persistentes a la vez, lo que puede conducir a una eficiencia mejorada. Se entenderá, sin embargo, que no es necesario agrupar escrituras persistentes y confirmarlas en un bloque para DIMM respaldados por energía en donde los reguladores están protegidos contra fallas de potencia.

Se pueden realizar las siguientes modificaciones al comando de escritura persistente para implementar las confirmaciones de grupo analizadas anteriormente. El servidor 120 puede elegir un único WID (de los WID 210, por ejemplo) para un conjunto de dos o más escrituras. Se puede agregar un bit adicional denominado "Persistir" al comando de escritura persistente cuando se envía en CA 114, por ejemplo. El bit de Persistencia puede usarse para determinar cuándo se ha enviado todo el grupo de escrituras persistentes al sistema 130 de memoria.

Por ejemplo, tres escrituras persistentes 64B se pueden agrupar de la siguiente manera usando WID = 5 en un ejemplo ilustrativo. Se puede enviar una primera escritura persistente (WID=5, persistencia =0), una segunda escritura persistente (WID=5, persistencia =0) y una tercera escritura persistente (WID=5, persistencia =1) en CA 114. El sistema 130 de memoria puede configurarse para recopilar las escrituras persistentes con WID=5 en un regulador mientras el bit de persistencia es 0, y cuando llega la última escritura persistente con el bit de persistencia establecido en 1, iniciar los procesos de confirmación de persistencia.

En una implementación, sólo se puede configurar una escritura persistente con un bit de persistencia establecido en 1 para obtener una respuesta Wr_Per del sistema 130 de memoria (por ejemplo, solo la tercera escritura persistente en el ejemplo anterior) para el grupo de escrituras persistentes. Esto puede reducir el tráfico en la respuesta 116.

En algunos aspectos, las escrituras persistentes con diferentes WID pueden entrelazarse, por ejemplo, en CA 114. En consecuencia, la agrupación de escrituras persistentes para confirmación persistente no implica que las escrituras persistentes en un grupo con el mismo WID se envíen consecutivamente desde el servidor 120.

En algunos aspectos, para abordar las condiciones de carrera que pueden surgir en las respuestas Wr_Per a escrituras persistentes, se puede usar un método de estado de ID de grupo de escritura (WGID) para agrupar estados de una o más escrituras persistentes, usando diferentes mapas de bits, tal como un mapa de bits completo con WGID y un mapa de bits pendiente con WGID, como se explicará con referencia a la figura 2B a continuación. Teniendo en cuenta las escrituras persistentes con los WID respectivos, el sistema 130 de memoria puede afirmar un Wr_Per respectivo (denominado "W_PER" en este caso) para cada escritura persistente con persistencia = 1 y para cada finalización de descarga. El servidor 120 puede usar otro comando enviar-W_PER-estado después de recibir uno o más W_PER (en donde, el servidor 120 también puede mantener un recuento de los W_PER, denominado W_PER-recuento). El sistema 130 de memoria puede devolver el estado WGID completado con bits completados únicamente con base en los W_PER ya afirmados. A su vez, el servidor 120 puede actualizar una lista para el WGID, o "lista WGID" y disminuir el W_PER-recuento con base en el número de finalizaciones.

En algunos casos, puede ocurrir un error incorregible (UE) en las transacciones, lo cual se discutirá con referencia a la figura 2C. Cuando hay un UE en enviar-W_PER-estado, el servidor 120 puede dejar de emitir nuevas escrituras/descargas persistentes y enviar-W_PER-estado. El servidor 120 puede enviar un comando de lectura de estado denominado estado de lectura X al sistema 130 de memoria. El sistema 130 de memoria, a su vez, recopila todas las escrituras persistentes antes de recibir el estado de lectura X para devolver el estado pendiente de WGID al servidor 120 (los paquetes de estado cubren la aserción W_PER antes de recibir un RD_RDY) y el sistema 130 de memoria puede continuar emitiendo W_PER durante las lecturas de estado. El servidor 120 puede actualizar la lista WGID mantenida por el servidor 120 y disminuir W_PER-recuento con base en las escrituras pendientes. El servidor 120 puede entonces comenzar a volver a emitir las escrituras/descargas persistentes.

Para implementaciones DIMM respaldadas por energía del sistema 130 de memoria, en un protocolo normal, el servidor 120 puede emitir comandos de escritura persistente (con persistencia = 0/1) y descargar, pero el sistema 130 de memoria no afirmará W_PER para cada escritura persistente con persistencia =1, pero el sistema 130 de memoria afirmará W_PER para el comando de lavado cuando se complete el lavado. En el caso de implementaciones WGID, el manejo de W_PER por parte del sistema 130 de memoria sigue siendo el mismo que el protocolo normal solo para

descargas. Un mapa de bits de estado WGID completado proporcionado por el sistema 130 de memoria tendrá bits de descarga WGID establecidos cuando se completen. Cuando hay un UE en enviar-W_PER-estado, la operación sigue siendo la misma que en el caso normal, excepto que el estado pendiente de WGID solo se aplica para descargas.

Los créditos por implementaciones de WGID pueden manejarse de la siguiente manera. Se pueden mantener créditos separados para escrituras de estado o escrituras X y para escrituras persistentes, en donde el servidor 120 puede determinar cómo el sistema 130 de memoria puede asignar un conjunto de créditos. La devolución de crédito incremental puede ser proporcionada por metadatos de lectura P en donde se puede usar un esquema de codificación para devolver créditos escritura X o escritura persistente. Las devoluciones de estado de lectura X pueden estar disponibles para las ranuras del regulador escritura X y escritura persistente con base en la asignación de crédito.

En una implementación, por ejemplo, que se describirá con referencia a las figuras 2B-C, el sistema 130 de memoria puede completar escrituras persistentes (denominadas en este documento escritura p) y descargas en cualquier orden. Para conservar una escritura P específico en los medios, el servidor 120 puede emitir una escritura P para un WGID determinada con persistencia = 1 o emitir una escritura P con persistencia = 0 seguido de cualquiera de los tipos de descarga. El sistema 130 de memoria puede emitir W_PER para cada escritura P completada que tenga persistencia =1 en el comando así como cada descarga completada. Si se agrupan múltiples escritura P con un único WGID con persistencia =1 solo en la última escritura P que termina el grupo, el sistema 130 de memoria puede emitir W_PER solo cuando se complete todo el grupo de escritura p.

Haciendo referencia ahora a la figura 2B, se describirá el manejo de W_PER para un caso normal. El sistema 250 se muestra con el servidor 120 y el sistema 130 de memoria. Si tanto los créditos de escritura como el WGID libre están disponibles, entonces el servidor 120 puede emitir uno o más comandos escritura P o descarga mostrados como 252a, 254a. El servidor 120 puede rastrear los comandos escritura P o descarga emitidos 252a, 254a en una lista pendiente de WGID del servidor (no se muestra, pero puede mantenerse dentro del servidor 120).

En consecuencia, el sistema 130 de memoria puede aceptar y rastrear los comandos 252a, 254a escritura P o descarga pendientes en una lista pendiente de DIMM-WGID (no mostrada). El sistema 130 de memoria puede ejecutar los comandos 252a, 254a escritura P o descarga pendientes y afirmar los W_PER 254b y 252b correspondientes (nota, mostrados en orden inverso a los comandos 252a, 254a escritura P o descarga recibidos) al servidor 120 después de la finalización respectiva de cada comando recibido.

El sistema 130 de memoria puede recopilar los comandos recibidos completos escritura P o descarga 252a, 254a en el mapa 260 de bits completado WGID, al cual se muestran diversas actualizaciones 260a, 260b, 260c, etc. El sistema 130 de memoria también puede eliminar los comandos 252a, 254a escritura P o descarga completados de la lista pendiente de DIMM-WGID.

El servidor 120 puede mantener un recuento de eventos W_PER recibidos, por ejemplo, para recibir los W_PER 254b, 252b, denominados W_PER-recuento. Al mismo tiempo, el servidor 120 puede manejar los eventos W_PER recibidos de la siguiente manera: si W_PER-recuento>0, entonces el servidor 120 puede emitir una solicitud de estado que se muestra como enviar-W_PER-estado 256a. Después de un tiempo predefinido, denominado tiempo T enviar, el sistema 130 de memoria puede enviar una instantánea del mapa 260 de bits completado WGID en esa instancia de tiempo (260b en este caso) en la respuesta mostrada como estado 256b WGID al servidor 120. La instantánea puede incluir finalizaciones de W_PER emitidas hasta el inicio de la transferencia estado 256b WGID al servidor 120.

En algunos aspectos, se registra una finalización como mínimo en la instantánea. El sistema 130 de memoria borra las posiciones de bits en el mapa 260 de bits completado WGID con base en las finalizaciones enviadas en estado 256b WGID, mostrado por la transición del mapa 260b de bits completado WGID al mapa 260c de bits completado WGID después del restablecimiento o borrado de las posiciones de bits.

El servidor 120 recibe el estado 256b WGID y puede extraer información sobre los WGID completados. En consecuencia, el servidor 120 puede liberar WGID completados de la lista pendiente de WGID del servidor y disminuir W_PER-recuento en el número de finalizaciones recibidas en el mapa 260 de bits completado WGID (por ejemplo, disminuir un recuento de 2 con base en los dos W_PER recibidos como lo indica el mapa 260b de bits completado WGID). El servidor 120 puede repetir el proceso anterior comenzando con la monitorización de W_PER-recuento y si W_PER-recuento>0, luego emitiendo otra solicitud de estado enviar-W_PER-estado 256a al sistema 130 de memoria.

En implementaciones de ejemplo, el servidor 120 y el sistema 130 de memoria pueden continuar emitiendo y ejecutando nuevos escritura P mientras el procesamiento del evento W_PER está en marcha. Aunque W_PER-recuento y las listas pendientes tal como la lista pendiente de WGID del servidor, la lista pendiente de DIMM-WGID, etc., se han analizado para una implementación de ejemplo, se pueden usar estructuras alternativas para lograr la funcionalidad descrita anteriormente sin desviarse del alcance de esta divulgación.

Con referencia ahora a la figura 2C, se muestra el sistema 270 para manejar el error incorregible (UE) del canal después de enviar-W_PER-estado desde el servidor 120, lo que resulta en la pérdida de las finalizaciones enviadas en estado 256b WGID (explicado en la figura 2B anterior) desde el sistema 130 de memoria. Además, se observa que

el sistema 130 de memoria puede haber borrado las finalizaciones anteriores del mapa 260 de bits completado WGID en la figura 2B.

En consecuencia, en un protocolo para recuperarse de tales errores en el sistema 270, el servidor 120 puede iniciar el proceso de recuperación deteniendo la emisión de nuevos comandos escritura P o descarga (por ejemplo, escritura P-3 o descarga-3 272a no se emite, se muestra en líneas discontinuas para indicar el cronograma en el que se habrían emitido si no se hubiera producido el error), mientras que el sistema 130 de memoria puede continuar emitiendo eventos RD_RDY y/o W_PER para lecturas completadas o comandos escritura P o descarga (por ejemplo, se muestra que W_PER 254b se emite mientras que 252b no se emite hasta después de la recuperación del error). El servidor 120 también puede continuar emitiendo enviar y actualizando W_PER-recuento.

Después de un retraso de tiempo mínimo preespecificado para una señal de habilitación de escritura, denominado retraso_TWE después de la última escritura P, el servidor 120 emite estado 274a de lectura X al sistema 130 de memoria, y el sistema 130 de memoria puede preparar un paquete de estado completo con una instantánea del mapa 280 de bits WGID pendiente, que es otro mapa de bits proporcionado además del mapa 260 de bits completado WGID discutido anteriormente, en donde el mapa 280 de bits WGID pendiente incluye el estado de todos los escritura P/descarga pendientes. El sistema 130 de memoria puede afirmar RD_RDY 276b, y el servidor 120 puede emitir enviar 278a en respuesta.

El sistema 130 de memoria puede entonces devolver el paquete 278b de estado preparado del cual el servidor 120 puede extraer y procesar el mapa 280 de bits WGID pendiente recibido en el paquete 278b de estado. El servidor 120 puede liberar los WGID apropiados de su lista de seguimiento pendiente de WGID del servidor y disminuir W_PER-recuento en el número de WGID liberados. El servidor 120 puede entonces repetir los procesos comenzando con la emisión de nuevos comandos escritura P/descarga y procesar los W_PER pendientes en este momento según la página anterior.

En algunos aspectos, el paquete 278b de estado está configurado para indicar si tiene el mapa 260 de bits completado WGID o el mapa 280 de bits WGID pendiente. Los paquetes de estado de respuesta W_PER contienen el mapa 260 de bits completado WGID, mientras que todos los demás paquetes de estado contienen el mapa 280 de bits pendiente de WGID. El tiempo de retraso_TWE está configurado para tener en cuenta el tiempo para recibir la notificación de error del sistema 130 de memoria para la última escritura P emitida desde el servidor 120 y el tiempo de espera desde la detección del UE antes de que estado 274a de lectura X se emita desde el servidor 120 puede variar dependiendo de cuándo se emitió la última escritura P.

Con referencia a la figura 3, se muestra un ejemplo de codificación para escrituras persistentes, por ejemplo, para una implementación DDR5 del sistema 130 de memoria. El campo CA1 se utiliza normalmente para diferenciar entre comandos 1UI y 2UI en la tecnología DDR5 y puede conservarse para implementaciones NVDIMM-P. CA 114, en algunas implementaciones, se puede configurar a velocidades DDR para DDR5 con solo 7 clavijas y, en tales casos, se puede usar una codificación de comando separada para escrituras persistentes, por ejemplo, como se muestra en la figura 3.

En la figura 3, si el bit de persistencia = 1, esto indica que el sistema 130 de memoria debe enviar todas las escrituras persistentes asociadas con el WID respectivo a la memoria no volátil. Si hay una única escritura persistente 64B en un grupo, el bit de persistencia puede establecerse en 1. Para escrituras persistentes mayores a 64B, todas las escrituras persistentes pueden tener el mismo WID, donde la última escritura persistente tiene su bit de persistencia establecido en 1, mientras que las escrituras persistentes restantes tienen sus bits de persistencia establecidos en 0.

Además de las transacciones anteriores, como se presentó en las secciones anteriores, también se puede usar otro comando en asociación con escrituras persistentes, denominado comando descargar. El comando descargar está configurado para indicar al sistema 130 de memoria que todas las escrituras anteriores almacenadas en el regulador (por ejemplo, en memoria no persistente o volátil) deben enviarse a la memoria persistente, teniendo en cuenta que es posible que las escrituras futuras no se vean afectadas de manera similar o enviadas a la memoria persistente cuando se utiliza el comando descargar.

Cuando se completa la ejecución de descargar, el sistema 130 de memoria puede volver a afirmar Wr_Per, por ejemplo, en la respuesta 116 al servidor 120, similar al caso de las escrituras persistentes analizadas anteriormente.

Además, el servidor 120 también puede proporcionar el comando enviar estado para WID (similar a escrituras persistentes) al sistema 130 de memoria en el caso del comando descargar, al cual el sistema 130 de memoria puede responder con un paquete de estado WID con un WID reservado único para indicar la finalización de la ejecución de descargar (por ejemplo, un WID con todos los bits establecidos en 1 puede ser un WID reservado usado para indicar la finalización de la ejecución de descargar).

En una implementación, sólo se puede permitir un comando descargar pendiente desde el servidor 120. Por lo tanto, en esta implementación, el servidor 120 puede tener que esperar la respuesta de finalización de descargar del sistema 130 de memoria antes de enviar otro comando de descargar. En implementaciones alternativas, los comandos

descargar pueden ir acompañados de ID descargar correspondientes (por ejemplo, seleccionados de campos WID reservados) y la respuesta correspondiente al estado de envío puede hacer que el sistema 130 de memoria devuelva ID descargar cuya ejecución de descargar se ha completado.

5 Se apreciará que los aspectos incluyen diversos métodos para realizar los procesos, funciones y/o algoritmos divulgados en este documento. Por ejemplo, la figura 4A ilustra un método 400 de ejemplo para realizar operaciones persistentes.

10 El bloque 402 comprende recibir, en un sistema de memoria (por ejemplo, el sistema 130 de memoria), un comando de escritura persistente (por ejemplo, escritura persistente 202a) y datos de escritura asociados (por ejemplo, datos 202b) desde un servidor (por ejemplo, servidor 120).

15 El bloque 404 comprende realizar una escritura persistente de los datos de escritura en una memoria no volátil (por ejemplo, en una memoria no volátil en el sistema 130 de memoria) en el sistema de memoria con base en el comando de escritura persistente. Se puede recibir una identificación de escritura (WID) asociada con el comando de escritura persistente desde el servidor y, una vez finalizada con éxito la escritura persistente, se puede proporcionar al servidor una indicación de finalización de escritura persistente (Wr_Per) junto con el WID asociado (por ejemplo, Wr_Per 202c).

20 De manera similar, la figura 4B ilustra otro método de ejemplo 450 para realizar operaciones persistentes.

El bloque 452 comprende proporcionar, desde un servidor (por ejemplo, el servidor 120) a un sistema de memoria (por ejemplo, el sistema 130 de memoria), un comando de escritura persistente (por ejemplo, escritura 202a persistente) y datos de escritura asociados (por ejemplo, datos 202b) en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil.

25 El bloque 454 comprende proporcionar una identificación de escritura (WID) (por ejemplo, WID 210) asociada con el comando de escritura persistente al sistema de memoria desde el servidor.

30 A continuación se analizará un aparato de ejemplo en el que se pueden utilizar aspectos de esta divulgación en relación con la figura 5. La figura 5 muestra un diagrama de bloques del dispositivo 500 informático. El dispositivo 500 informático puede corresponder a una implementación de ejemplo de un sistema 100 de procesamiento de la figura 1, en donde el procesador 120' puede ser uno de los elementos de procesamiento del servidor 120. Se muestra a modo de ejemplo que el procesador 120' está acoplado al sistema 130 de memoria a través de la interconexión 110, omitiéndose en esta vista detalles adicionales de la interconexión 110 en aras de la claridad. El procesador 120', la interconexión 110 y el sistema 130 de memoria pueden configurarse para realizar los métodos 400-450 como se analizó anteriormente. Se entenderá que otras configuraciones de memoria conocidas en la técnica, tales como las que implican uno o más niveles de cachés, aunque no se muestran, pueden estar presentes en el dispositivo 500 informático.

40 La figura 5 también muestra el controlador 526 de visualización que está acoplado al procesador 120' y al visualizador 528. En algunos casos, el dispositivo 500 informático puede usarse para comunicación inalámbrica y la figura 5 también muestra bloques opcionales en líneas discontinuas, tal como un codificador/decodificador (CODEC) 534 (por ejemplo, un CODEC de audio y/o voz) acoplado al procesador 120' y el altavoz 536 y el micrófono 538 pueden acoplarse al CODEC 534; y una antena 542 inalámbrica acoplada al controlador 540 inalámbrico que está acoplado al procesador 120'. Cuando uno o más de estos bloques opcionales están presentes, en un aspecto particular, el procesador 120', el controlador 526 de visualización, el sistema 130 de memoria y el controlador 540 inalámbrico se incluyen en un dispositivo 522 de sistema en paquete o sistema en chip.

50 En consecuencia, un aspecto particular, el dispositivo 530 de entrada y el suministro 544 de potencia están acoplados al dispositivo 522 de sistema en chip. Además, en un aspecto particular, como se ilustra en la figura 5, donde están presentes uno o más bloques opcionales, el visualizador 528, el dispositivo 530 de entrada, el altavoz 536, el micrófono 538, la antena 542 inalámbrica y el suministro 544 de potencia son externos al dispositivo 522 de sistema en chip. Sin embargo, cada uno del visualizador 528, el dispositivo 530 de entrada, el altavoz 536, el micrófono 538, la antena 542 inalámbrica y el suministro 544 de potencia pueden acoplarse a un componente del dispositivo 522 de sistema en chip, tal como una interfaz o un controlador.

60 Cabe señalar que aunque la figura 5 representa en general un dispositivo informático, un procesador 120' y un sistema 130 de memoria, también pueden integrarse en un decodificador, un servidor, un reproductor de música, un reproductor de vídeo, una unidad de entretenimiento, un dispositivo de navegación, un asistente digital personal (PDA), una unidad de datos de ubicación fija, un ordenador, un ordenador portátil, una tableta, un dispositivo de comunicación, un teléfono móvil u otros dispositivos similares.

65 Los expertos en la técnica apreciarán que la información y las señales pueden representarse utilizando cualquiera de una variedad de tecnologías y técnicas diferentes. Por ejemplo, los datos, instrucciones, comandos, información, señales, bits, símbolos y chips a los que se puede hacer referencia a lo largo de la descripción anterior pueden

representarse mediante voltajes, corrientes, ondas electromagnéticas, partículas o campos magnéticos, partículas o campos ópticos, o cualquier combinación de los mismos.

Además, los expertos en la técnica apreciarán que los diversos bloques lógicos, módulos, circuitos y pasos de algoritmo ilustrativos descritos en relación con los aspectos divulgados en este documento pueden implementarse como hardware electrónico, software informático o combinaciones de ambos. Para ilustrar claramente esta intercambiabilidad de hardware y software, se han descrito anteriormente diversos componentes, bloques, módulos, circuitos y pasos ilustrativos en términos generales en términos de su funcionalidad. Que dicha funcionalidad se implemente como hardware o software depende de la aplicación particular y de las restricciones de diseño impuestas al sistema general. Los expertos pueden implementar la funcionalidad descrita de diversas maneras para cada aplicación particular, pero tales decisiones de implementación no deben interpretarse como que provocan una desviación del alcance de la presente invención.

Los métodos, secuencias y/o algoritmos descritos en relación con los aspectos divulgados en este documento pueden incorporarse directamente en hardware, en un módulo de software ejecutado por un procesador o en una combinación de ambos. Un módulo de software puede residir en una memoria RAM, una memoria flash, una memoria ROM, una memoria EPROM, una memoria EEPROM, registros, un disco duro, un disco extraíble, un CD-ROM o cualquier otra forma de medio de almacenamiento conocido en la técnica. Un medio de almacenamiento de ejemplo está acoplado al procesador de manera que el procesador pueda leer información y escribir información en el medio de almacenamiento. Como alternativa, el medio de almacenamiento puede ser parte integral del procesador.

Por consiguiente, un aspecto de la invención puede incluir un medio legible por ordenador que incorpora un método para realizar escrituras persistentes. Por consiguiente, la invención no se limita a ejemplos ilustrados y cualquier medio para realizar la funcionalidad descrita en este documento se incluye en aspectos de la invención.

Si bien la divulgación anterior muestra aspectos ilustrativos de la invención, cabe señalar que se podrían realizar diversos cambios y modificaciones en este documento sin apartarse del alcance de la invención como se define en las reivindicaciones adjuntas. Las funciones, pasos y/o acciones de las reivindicaciones del método según los aspectos de la invención descritos en este documento no necesitan realizarse en ningún orden particular. Además, aunque los elementos de la invención pueden describirse o reivindicarse en singular, se contempla el plural a menos que se establezca explícitamente la limitación al singular.

REIVINDICACIONES

1. Un método para realizar operaciones persistentes, el método que comprende:
 5 recibir (402), en un sistema de memoria, un comando de escritura persistente y datos de escritura asociados desde un servidor; y
 realizar (404) una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente, y
 recibir una identificación de escritura, WID, asociada con el comando de escritura persistente desde el servidor; y caracterizado además por comprender:
 10 recibir un grupo de dos o más comandos de escritura persistente con un WID común, teniendo el último comando de escritura persistente del grupo un bit de persistencia establecido en 1 y los comandos de escritura persistente restantes teniendo los respectivos bits de persistencia establecidos en 0 y proporcionando una indicación de finalización de escritura persistente para el último comando de escritura persistente.
- 15 2. El método de la reivindicación 1, que comprende además proporcionar, desde el sistema de memoria tras la finalización exitosa de la escritura persistente, una indicación de finalización de la escritura persistente junto con la WID asociada al servidor.
3. El método de la reivindicación 2, que comprende proporcionar dos o más indicaciones de finalización de escritura persistente al servidor en un orden diferente del orden en el que se recibieron dos o más comandos de escritura persistente correspondientes desde el servidor.
4. El método de la reivindicación 2, que comprende además recibir, desde el servidor, una solicitud para enviar el estado de una o más escrituras persistentes junto con los WID asociados.
- 25 5. El método de la reivindicación 4, que comprende además proporcionar un paquete de estado al procesador servidor, el paquete de estado que comprende WID para comandos de escritura persistente cuya ejecución se ha completado.
6. El método de la reivindicación 1, en donde el WID comprende una identificación de múltiples bits de una escritura persistente y un bit válido.
- 30 7. El método de la reivindicación 1, que comprende además recibir un comando descargar desde el servidor, en donde el comando descargar indica que todas las escrituras anteriores reguladas en medios volátiles deben enviarse a una memoria no volátil o persistente, y que comprende además proporcionar una indicación de finalización de descargar al finalizar la ejecución del comando descargar al servidor.
- 35 8. El método de la reivindicación 1, que comprende además recibir uno o más comandos de escritura persistente, mantener los estados de uno o más comandos de escritura persistente completados en un mapa de bits completado y los estados de uno o más comandos de escritura persistente pendientes en un mapa de bits pendiente, y al solicitar el estado del servidor, proporcionar el mapa de bits completo si no hay un error incorregible o el mapa de bits pendiente si hay un error incorregible.
- 40 9. El método de la reivindicación 1, en donde el sistema de memoria es un módulo de memoria dual en línea no volátil configurado para soportar escrituras persistentes, NVDIMM-P.
- 45 10. Un método para realizar operaciones persistentes, el método que comprende:
 proporcionar (452), desde un servidor a un sistema de memoria, un comando de escritura persistente y datos de escritura asociados,
 en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil, y
 50 proporcionar (454) una identificación de escritura, WID, asociada con el comando de escritura persistente al sistema de memoria desde el servidor; y caracterizado además por comprender:
 enviar desde el servidor al sistema de memoria, un grupo de dos o más comandos de escritura persistente con un WID común, con el último comando de escritura persistente del grupo con un bit de persistencia establecido en 1 y las escrituras persistentes restantes tienen los respectivos bits de persistencia establecidos en 0 y reciben del sistema de memoria, una indicación de finalización de la escritura persistente para la última escritura persistente.
- 55 11. Un aparato que comprende:
 un sistema de memoria configurado para:
 60 recibir un comando de escritura persistente y datos de escritura asociados desde un servidor; y realizar una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente y
 en donde el sistema de memoria está configurado para recibir una identificación de escritura, WID, asociada con el comando de escritura persistente desde el servidor; y en donde el sistema de memoria se caracteriza además porque
 65 está configurado para:

recibir un grupo de dos o más comandos de escritura persistente con un WID común, teniendo el último comando de escritura persistente del grupo un bit de persistencia establecido en 1 y los comandos de escritura persistente restantes teniendo los respectivos bits de persistencia establecidos en 0 y proporcionando una indicación de finalización de escritura persistente para el último comando de escritura persistente.

5 12. Un aparato que comprende:
un servidor configurado para proporcionar un comando de escritura persistente y datos de escritura asociados a un sistema de memoria,
10 en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil, y
en donde el servidor está configurado para proporcionar una identificación de escritura, WID, asociada con el comando de escritura persistente al sistema de memoria; y se caracteriza además porque está configurado para:
15 enviar desde el servidor al sistema de memoria, un grupo de dos o más comandos de escritura persistente con un WID común, teniendo el último comando de escritura persistente del grupo un bit de persistencia establecido en 1 y las escrituras persistentes restantes teniendo los respectivos bits de persistencia establecidos en 0 y recibiendo del sistema de memoria, una indicación de finalización de escritura persistente para la última escritura persistente.

20 13. Un medio de almacenamiento no transitorio legible por ordenador que comprende código, que, cuando lo ejecuta un procesador, hace que el procesador realice operaciones persistentes, el medio de almacenamiento no transitorio legible por ordenador que comprende:
código para recibir, en un sistema de memoria, un comando de escritura persistente y datos de escritura asociados desde un procesador servidor; y
código para realizar una escritura persistente de los datos de escritura en una memoria no volátil en el sistema de memoria con base en el comando de escritura persistente,
25 código para recibir una identificación de escritura, WID, asociada con el comando de escritura persistente del servidor; y caracterizado además por comprender:
código para recibir un grupo de dos o más comandos de escritura persistente con un WID común, con el último comando de escritura persistente del grupo con un bit de persistencia establecido en 1 y los comandos de escritura persistente restantes tienen los respectivos bits de persistencia establecidos en 0 y proporcionan una indicación de finalización de escritura persistente para el último comando de escritura persistente.
30

14. Un medio de almacenamiento no transitorio legible por ordenador que comprende código, que, cuando es ejecutado por un procesador, hace que el procesador realice operaciones persistentes, el medio de almacenamiento no transitorio legible por ordenador que comprende
35 código para proporcionar, desde un servidor a un sistema de memoria, un comando de escritura persistente y datos de escritura asociados,
en donde el comando de escritura persistente indica al sistema de memoria que realice una escritura persistente de los datos de escritura en una memoria no volátil.
código para proporcionar una identificación de escritura, WID, asociada con el comando de escritura persistente desde el servidor al sistema de memoria; y caracterizado además por comprender:
40 código para enviar desde el servidor al sistema de memoria, un grupo de dos o más comandos de escritura persistente con un WID común, con el último comando de escritura persistente del grupo con un bit de persistencia establecido en 1 y las escrituras persistentes restantes tienen los respectivos bits de persistencia establecidos en 0 y reciben del sistema de memoria, una indicación de finalización de la escritura persistente para la última escritura persistente.
45

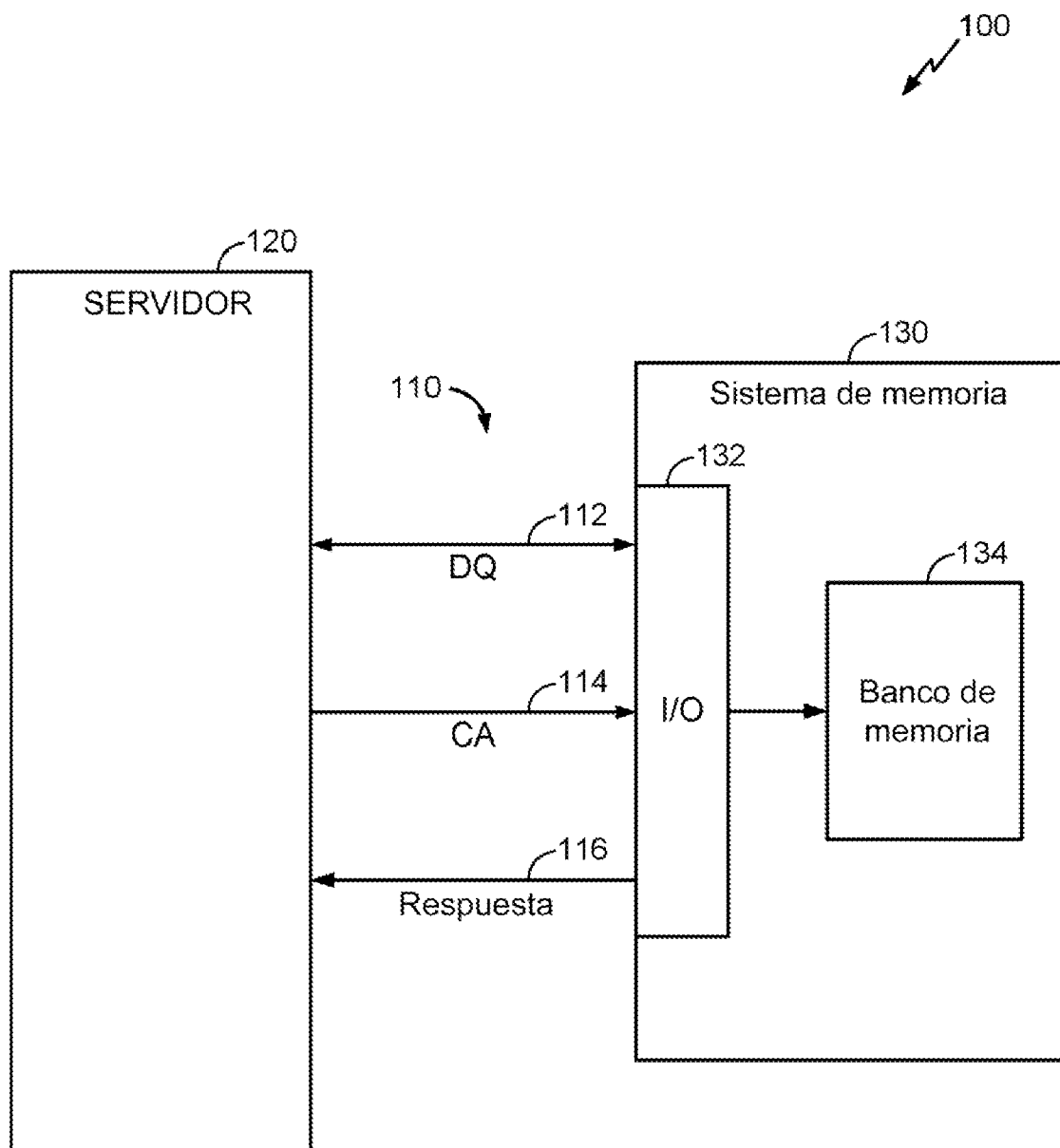


FIG. 1

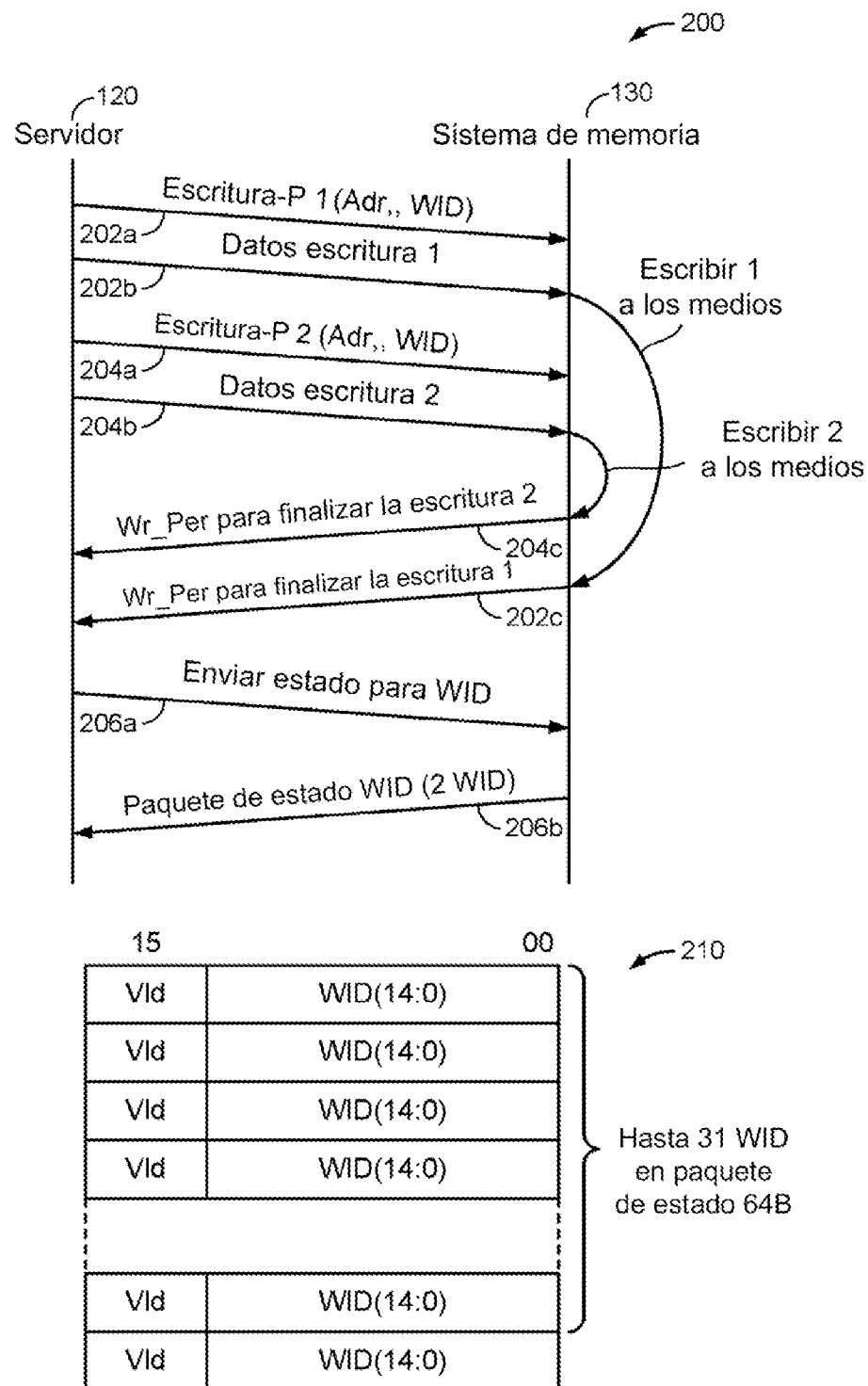


FIG. 2A

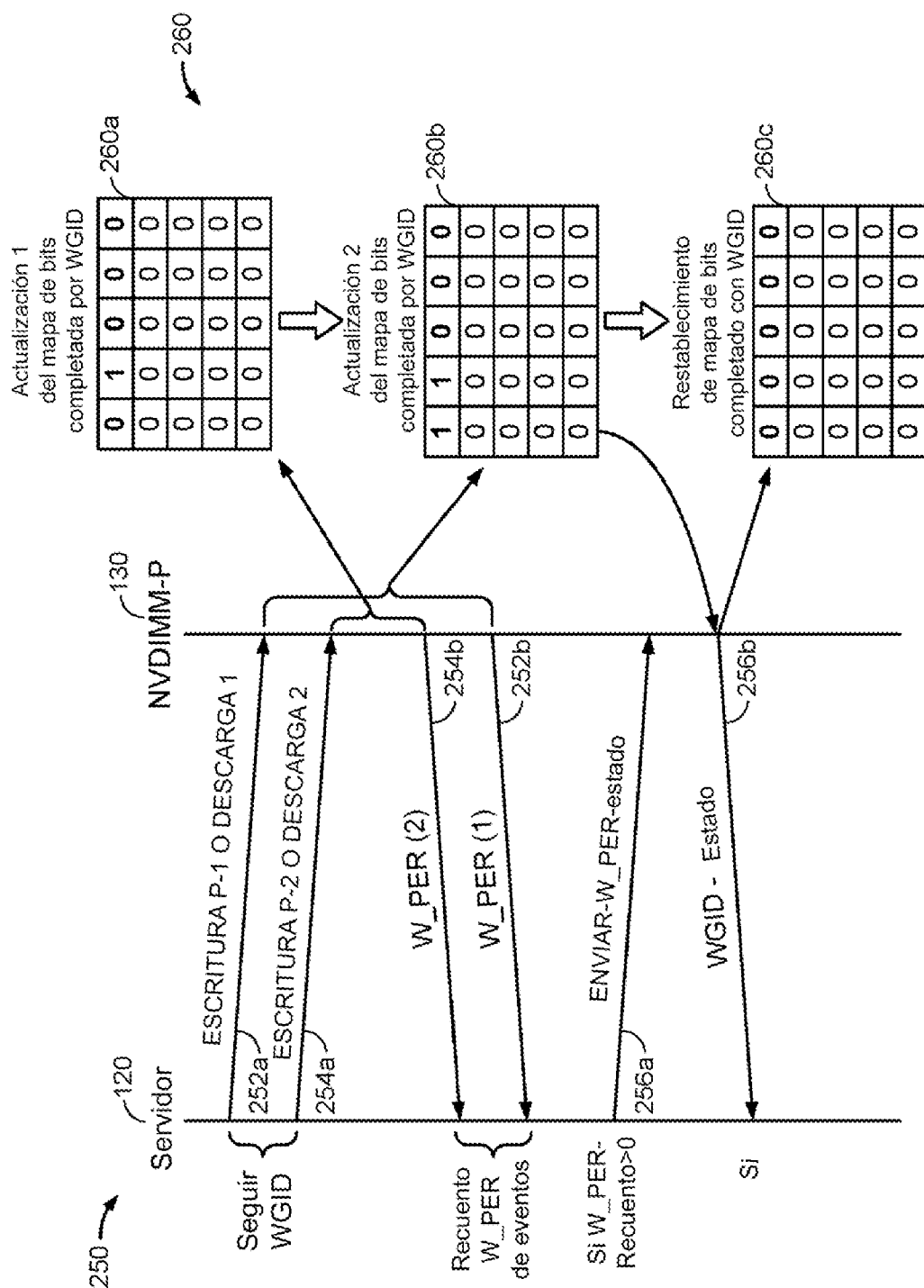


FIG. 2B

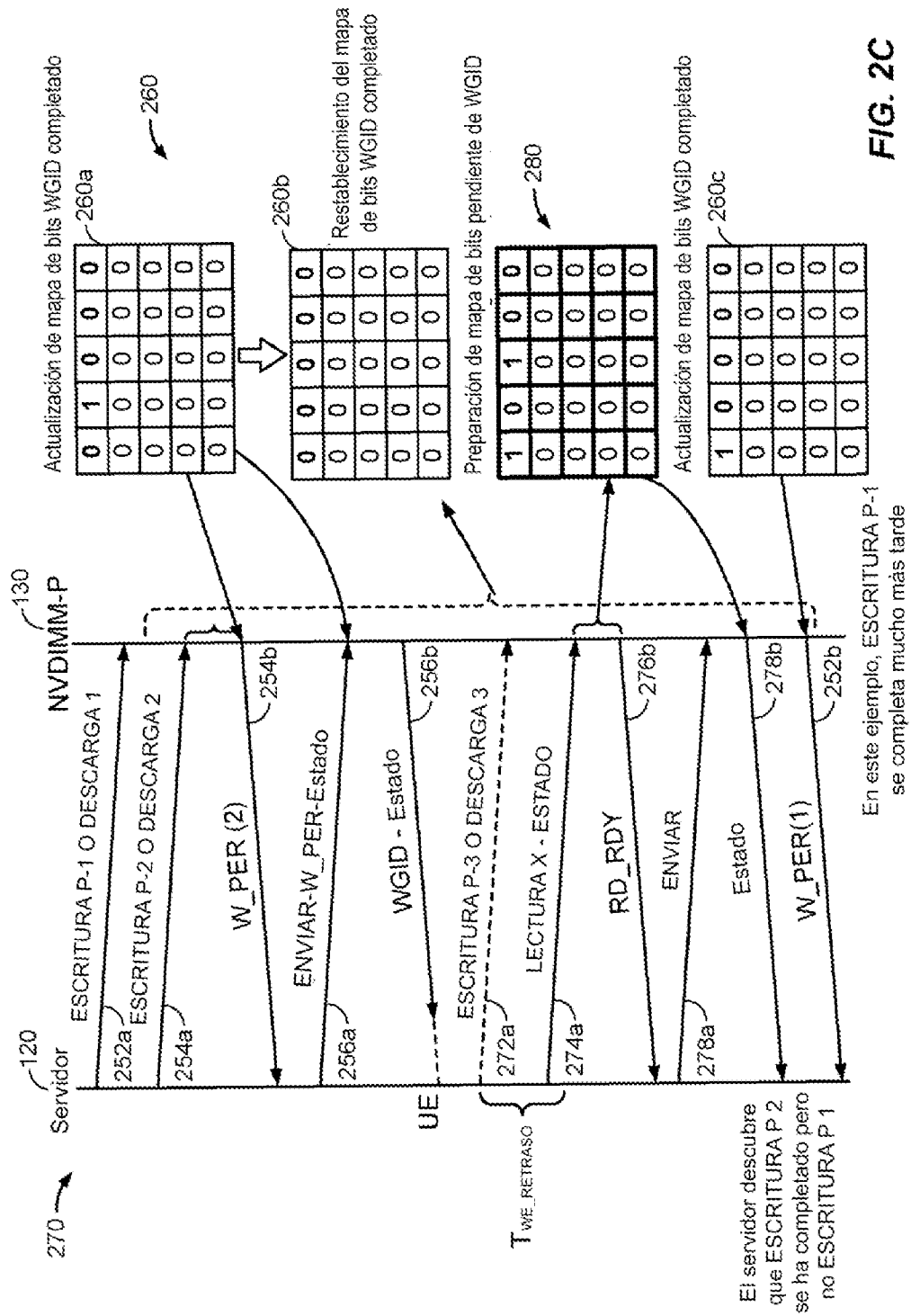


FIG. 2C

CA0	CA1	CA2	CA3	CA4	CA5	CA6	CA7
H	L	L	H	L	DIRECCIÓN [39:37]		
DIRECCIÓN [36:29]							
DIRECCIÓN [28]	WID [9:3]						
WID [2:0]			PERSISTIR	RFU			

ESCRITURA XP

FIG. 3

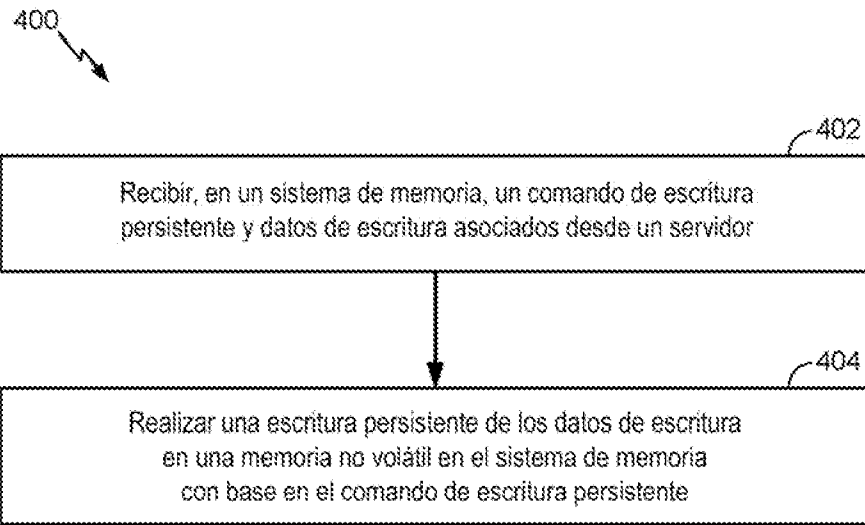


FIG. 4A

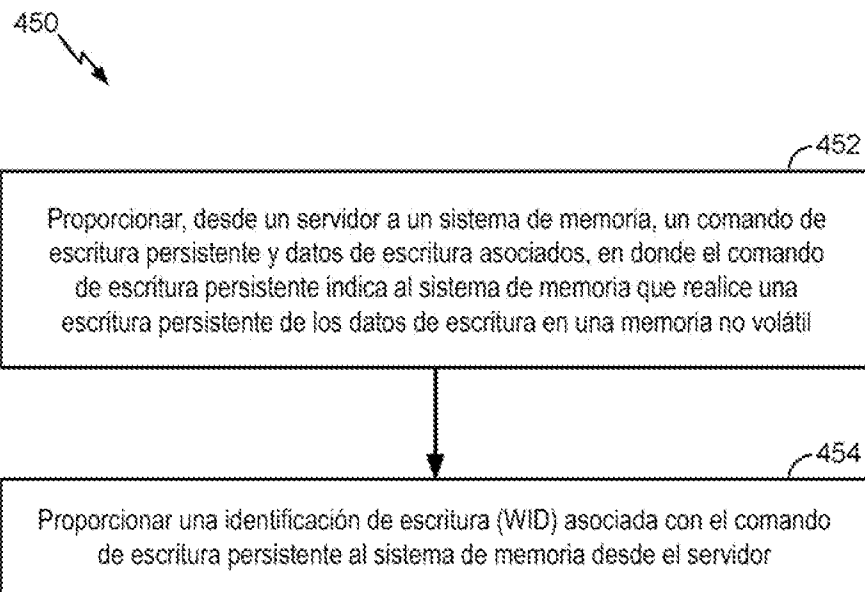


FIG. 4B

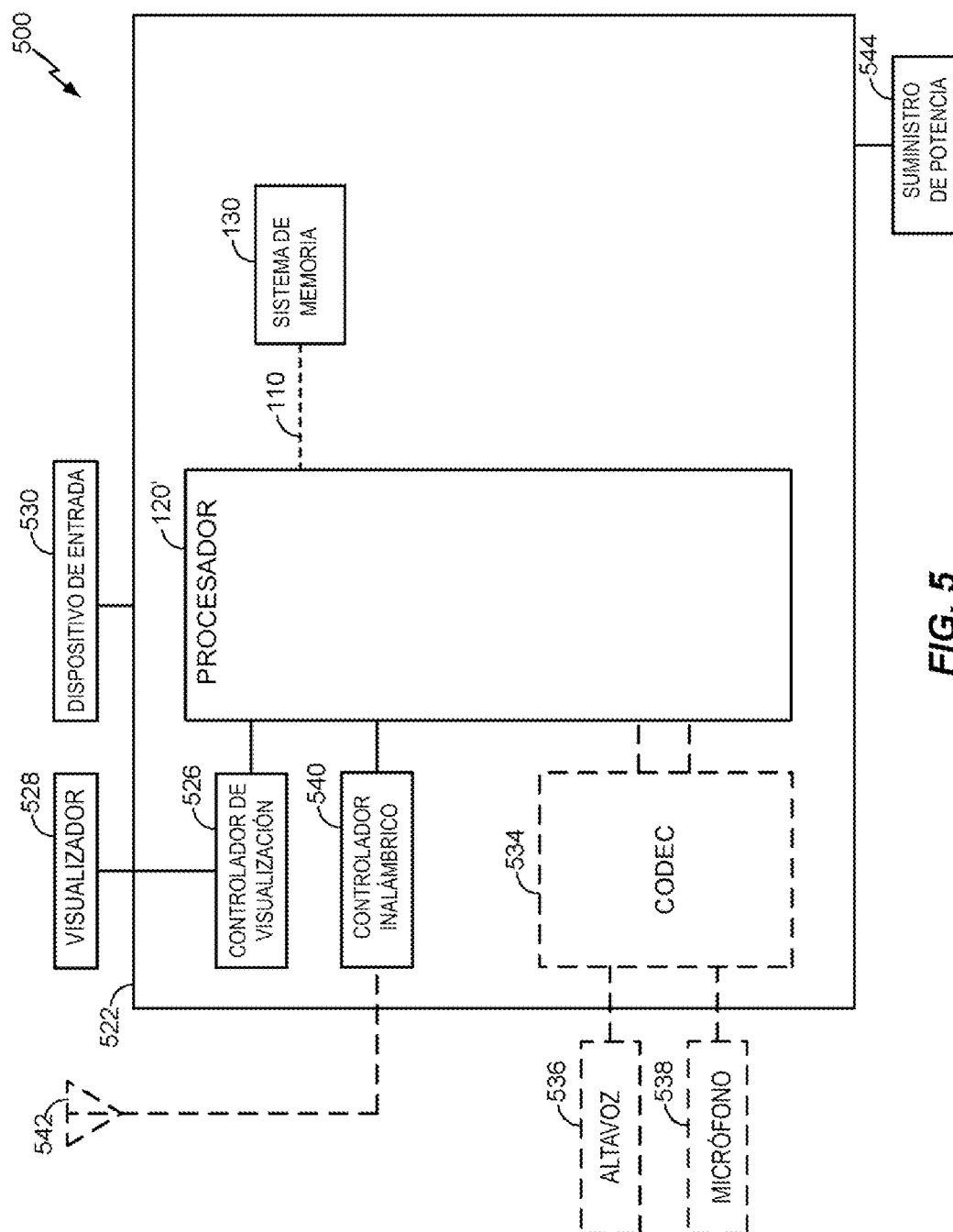


FIG. 5