

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 015 283**

51 Int. Cl.:

G06F 9/30 (2008.01)

G06F 13/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.01.2020 PCT/EP2020/050757**

87 Fecha y número de publicación internacional: **06.08.2020 WO20156797**

96 Fecha de presentación y número de la solicitud europea: **14.01.2020 E 20700224 (7)**

97 Fecha y número de publicación de la concesión europea: **26.02.2025 EP 3918467**

54 Título: **Manejo de una instrucción de almacenamiento de entrada/salida**

30 Prioridad:

31.01.2019 EP 19154735

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

30.04.2025

73 Titular/es:

**INTERNATIONAL BUSINESS MACHINES CORPORATION (100.00%)
New Orchard Road
Armonk, New York 10504, US**

72 Inventor/es:

**RAISCH, CHRISTOPH;
KRAEMER, MARCO;
LEHNERT, FRANK;
KLEIN, MATTHIAS;
BRADBURY, JONATHAN;
JACOBI, CHRISTIAN;
DRIEVER, PETER y
BELMAR, BRENTON**

74 Agente/Representante:

ELZABURU, S.L.P

ES 3 015 283 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

Manejo de una instrucción de almacenamiento de entrada/salida

La presente invención se refiere en general a sistemas de procesamiento de datos, en particular a un procedimiento para gestionar una instrucción de almacenamiento de entrada/salida para múltiples dispositivos externos, así como a un producto de programa informático y a un sistema de procesamiento de datos.

Antecedentes

Un entorno informático puede incluir uno o más tipos de dispositivos de entrada/salida, incluidos varios tipos de adaptadores. Un tipo de adaptador es un adaptador de interconexión de componentes periféricos (PCI) o interconexión expresa de componentes periféricos (PCle). Este adaptador incluye uno o más espacios de direcciones que se utilizan para comunicar datos entre el adaptador y el sistema al que está conectado el adaptador.

En algunos sistemas, una parte de un espacio de direcciones de la unidad central de procesamiento (CPU) acoplada al adaptador se asigna a un espacio de direcciones del adaptador, lo que permite que las instrucciones de la CPU que acceden al almacenamiento manipulen directamente los datos del espacio de direcciones del adaptador.

La comunicación con los adaptadores, como los adaptadores PCI o PCle, se puede facilitar mediante instrucciones de control diseñadas específicamente para la comunicación de datos hacia y desde los adaptadores y utilizadas para la comunicación.

En el estado de la técnica, una instrucción de almacenamiento para almacenar datos en un adaptador incluye, por ejemplo, obtener una instrucción de máquina para su ejecución, definiéndose la instrucción de máquina para su ejecución por ordenador según una arquitectura de ordenador, la instrucción de máquina incluyendo, por ejemplo, un campo de código de operación que identifica una instrucción de almacenamiento a adaptador. Un primer campo identifica una primera ubicación que incluye datos que se almacenarán en un adaptador. Un segundo campo identifica una segunda ubicación, cuyo contenido incluye un identificador de función que identifica el adaptador, una designación de un espacio de direcciones dentro del adaptador donde se van a almacenar los datos y un desplazamiento dentro del espacio de direcciones. La instrucción de la máquina se ejecuta, la ejecución incluye el uso del identificador de funciones para obtener una entrada de tabla de funciones asociada al adaptador. Se obtiene una dirección de datos del adaptador utilizando al menos una de las informaciones de la entrada de la tabla de funciones y el desplazamiento. Los datos se almacenan desde la primera ubicación en una ubicación específica en el espacio de direcciones identificado por la designación del espacio de direcciones, la ubicación específica identificada por la dirección de datos del adaptador.

Una característica existente en un sistema multiprocesador de gran tamaño es la capacidad de poner en modo inactivo todos los procesadores de una zona diana. Las funciones de inactividad funcionan para pausar o alterar temporalmente el estado de un procesador o grupo de procesadores para su ejecución, por ejemplo, actualizaciones del sistema o copias de seguridad. En algunos casos, una interrupción de inactividad es aplicable solo a un subconjunto de los recursos del sistema. En tales casos, el sistema se puede dividir en diferentes zonas. En el caso de una operación de inactividad aplicable a una zona (una zona de destino), los procesadores que se encuentran fuera de la zona de destino pueden seguir ejecutándose, aunque es posible que se bloqueen las nuevas traslaciones. Normalmente, al menos un controlador del sistema u otro mecanismo difunde la inactividad a todos los procesadores físicos del sistema, gestiona la recopilación de información sobre el estado de inactividad e indica al procesador solicitante cuándo se han iniciado todos los procesadores o ignoran (filtran) la solicitud de inactividad.

Un controlador de inactividad se puede acoplar comunicativamente a un procesador en un sistema multiprocesador y a una máquina de estado de inactividad configurada para recibir una solicitud de inactividad. El sistema informático está configurado para realizar un procedimiento que incluye recibir una solicitud de inactividad en el controlador de inactividad desde un procesador solicitante, siendo el procesador solicitante uno de una pluralidad de procesadores en un sistema multiprocesador, y determinar que la solicitud de inactividad no se acepta según estado de la máquina en estado de inactividad. El procedimiento también incluye, según que la solicitud no se acepte, generar un mensaje de rechazo configurado para indicar que la solicitud de inactividad ha sido rechazada, retener el mensaje de rechazo hasta que se transmita una orden de inactividad al sistema multiprocesador, la orden de inactividad basada en una solicitud de inactividad diferente y enviar el mensaje de rechazo al procesador solicitante según la transmisión de la orden de inactividad detectada por el controlador de inactividad.

El documento US 2011/0320764 A1 describe que se facilita la comunicación con los adaptadores de un entorno informático. Se proporcionan instrucciones que se dirigen explícitamente a los adaptadores. La información proporcionada en una instrucción se usa para dirigir la instrucción a una ubicación apropiada dentro del adaptador.

Además, el documento US 2015/0378737 A1 describe un procedimiento y un aparato para enviar paquetes usando secuencias de escritura PIO optimizadas sin frases. Las secuencias de instrucciones de escritura de entrada/salida programadas (PIO - Programmed Input/Output) para escribir paquetes de datos en una memoria de envío de PIO se reciben en un procesador que admite una ejecución desordenada. Las instrucciones de escritura de PIO se reciben en un orden original y se ejecutan fuera de orden, y cada instrucción de escritura de PIO escribe una unidad de

almacenamiento de datos en un búfer de almacenamiento o un bloque de datos de almacenamiento en el búfer de almacenamiento. Se proporciona lógica para que el búfer de almacenamiento detecte cuándo se llenan los bloques de almacenamiento, lo que da como resultado que los datos en esos bloques de almacenamiento se agoten a través de escrituras publicadas de PCIe que se escriben en bloques de envío en la memoria de envío PIO en direcciones definidas por las instrucciones de escritura PIO. La lógica se emplea para detectar el tamaño de llenado de los paquetes y cuándo se han llenado los bloques de envío de un paquete, lo que permite que los datos del paquete sean elegibles para la salida.

Mientras que el documento US 6,286,095 B1 describe un aparato informático que incorpora instrucciones especiales para forzar la ejecución de las operaciones de carga y almacenamiento en el orden del programa. La presente invención proporciona una instrucción de almacenamiento nueva y novedosa que se suspende hasta que una CPU asociada haya completado todas las instrucciones de almacenamiento anteriores. Además, se proporciona una nueva instrucción de carga que bloquea la ejecución de cualquier instrucción de carga posterior hasta que una CPU asociada haya completado esta instrucción de carga. Estas instrucciones permiten implementar sistemas informáticos de alta eficiencia que optimizan el rendimiento de las instrucciones mediante la ejecución de instrucciones posteriores mientras se espera que se complete una instrucción anterior.

El documento US 7,827, 443 B1 describe que los circuitos de recuperación reaccionan ante los errores en un núcleo de procesador esperando que se complete sin errores cualquier instrucción condicional de almacenamiento pendiente o una carga inhibida por la caché antes de dejar de comprobar o respaldar el progreso de un núcleo de procesador. Los circuitos de recuperación eliminan el núcleo del procesador de la configuración lógica del sistema multiprocesador simétrico, lo que puede reducir la propagación de errores a otras partes del sistema. El núcleo del procesador se restablece y los valores de los puntos de control pueden restaurarse en los registros del núcleo del procesador. El procesador central no solo puede reanudar la ejecución justo antes de las instrucciones que no se ejecutaron correctamente la primera vez, sino que también puede operar en un modo de ejecución reducida para un número de grupos preprogramado. Si el número preprogramado de grupos de instrucciones se ejecuta sin errores, se permite que el núcleo del procesador reanude la ejecución normal.

Compendio

Se propone un sistema de procesamiento de datos para gestionar una instrucción de almacenamiento de entrada/salida, que comprende un nido de sistema acoplado comunicativamente a al menos un bus de entrada/salida mediante un controlador de bus de entrada/salida. El sistema de procesamiento de datos comprende además al menos una unidad de procesamiento de datos que comprende un núcleo,

un firmware del sistema y una interfaz núcleo-nido asíncrona. La unidad de procesamiento de datos se acopla comunicativamente al nido del sistema a través de un búfer de agregación. El nido del sistema está configurado para cargar de forma asíncrona y/o almacenar datos en al menos un dispositivo externo que está acoplado comunicativamente al bus de entrada/salida. La interfaz núcleo-nido asíncrona comprende una matriz de estado de entrada/salida con varios búferes de estado de entrada/salida, así como una lógica de acceso y administración de matrices.

El sistema de procesamiento de datos está configurado para realizar: (i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos que emite la instrucción de almacenamiento de entrada/salida que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos a transferir y/o un puntero a los datos a transferir, y una longitud de los datos; (ii) la unidad de procesamiento de datos está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida; (iii) la unidad de procesamiento de datos está configurado para verificar si el acceso a la función de entrada/salida está permitido en un espacio de direcciones y a nivel de instancia invitada, donde el huésped se ejecuta en el sistema de procesamiento de datos; (iv) la unidad de procesamiento de datos está configurado para completar la instrucción de almacenamiento de entrada/salida antes de que se complete la ejecución de la instrucción de almacenamiento de entrada/salida en el nido del sistema; (v) el firmware del sistema está configurado para notificar al sistema operativo mediante una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida la unidad de procesamiento de datos detecta un error al transmitir los datos de la ejecución asíncrona fallida; (vi) la lógica de administración y acceso a la matriz recopila la finalización de la instrucción de almacenamiento y actualiza los búferes de estado de entrada/salida según los mensajes de finalización recibidos; y (vii) la unidad de procesamiento de datos retrasa la ejecución de la instrucción de almacenamiento, hasta que esté disponible un búfer de estado de entrada/salida para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes.

De este modo, se pueden permitir favorablemente múltiples instrucciones de almacenamiento asíncronas pendientes al mismo tiempo para reducir los ciclos por instrucción de las instrucciones de almacenamiento asíncronas repetidas. Un pedido se define entre instrucciones de almacenamiento asíncronas e instrucciones de carga/almacenamiento sincrónicas. La compatibilidad con varias instrucciones de almacenamiento asíncronas pendientes se basa en la contabilidad de varios mensajes de estado y en la correlación de las respuestas con las entradas de estado.

El sistema de procesamiento de datos según una primera realización de la invención comprende instrucciones que se cargan y almacenan en al menos un dispositivo externo del sistema de procesamiento de datos a través de un bus de

entrada/salida. Las instrucciones asíncronas se completan antes de que los datos se almacenen en el dispositivo externo, mientras que las instrucciones sincrónicas se completan después de que los datos se almacenen en el dispositivo externo. Dentro de las formas de realización aquí descritas, la PCI se utilizará indistintamente para cualquier otra tecnología de entrada/salida, sin restringir así la forma de realización de la invención a la PCI.

5 Las realizaciones de la invención describen la ejecución de una instrucción de almacenamiento de entrada/salida de una manera estrictamente ordenada, tal como se puede observar desde arriba del límite de la arquitectura, mientras que la ejecución real puede estar fuera de orden dentro del hardware de la unidad de procesamiento de datos (CPU).

10 Según algunas realizaciones de la invención, una instrucción de almacenamiento PCI puede ejecutarse con una ejecución asíncrona del efecto de almacenamiento PCIe y una gestión de estado asíncrona. La ejecución fiable asíncrona se basa en mecanismos de reenvío fiables en la microarquitectura del sistema de procesamiento de datos de la invención.

Una instrucción PCI de almacenamiento y bloque de almacenamiento existente suele ser sincrónica hasta el punto en que los datos de almacenamiento PCI se entregan a la interfaz PCIe y se devuelven su finalización a una unidad de procesamiento.

15 El estándar PCI solo requiere un comando de envío asíncrono de información PCI, que normalmente se implementa a través de una cola de almacenamiento en el procesador que agrega datos con un envío asíncrono.

Ventajosamente, según las realizaciones de la invención, se puede lograr una mejora con respecto a los ciclos por instrucción reemplazando una instrucción PCI sincrónica por un procedimiento de envío asíncrono fiable de una instrucción de almacenamiento de entrada/salida.

20 Alternativa o adicionalmente a los datos que se van a transferir, la instrucción de almacenamiento según una realización de la invención también puede especificar un puntero a una memoria principal que debe usarse para extraer datos, en lugar de contener los datos directamente.

El nivel de instancia invitada también puede significar que un único huésped o anfitrión puede estar ejecutándose en el sistema de procesamiento de datos.

25 La dirección del desplazamiento de la función de entrada/salida en sí misma puede ser una dirección lógica, física y virtual. Las direcciones virtuales y lógicas generalmente se traducen a través de una unidad de administración de memoria (MMU - memory management unit) a una dirección física, y la dirección física permite identificar a qué función y desplazamiento se refieren.

30 En este contexto, la dirección física significa «la dirección más baja en la jerarquía de traducción de direcciones accesible desde un sistema operativo/huésped».

35 Ventajosamente, los búferes de estado de entrada/salida pueden recopilar los estados devueltos del nido del sistema y/o del controlador del bus de entrada/salida, en particular un mensaje de finalización del nido del sistema. Estos búferes de estado de entrada/salida pueden recopilar los estados devueltos actuando como un búfer de mensajes del sistema asíncrono que soporta el procedimiento de transmisión asíncrona. Ventajosamente, los búferes de estado de entrada/salida pueden integrarse directamente en la interfaz asíncrona núcleo-nido para una respuesta rápida.

40 Según una realización favorable del sistema de procesamiento de datos de la invención, la unidad de procesamiento de datos puede retrasar la ejecución de instrucciones de almacenamiento asíncronas y/o sincrónicas, hasta que esté disponible un búfer de estado de entrada/salida para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes. Por lo tanto, un procedimiento de pedido de un múltiplo de instrucciones de almacenamiento a un múltiplo de dispositivos de entrada/salida puede gestionarse de una manera eficiente, lo que ofrece la capacidad de ahorrar tiempo de procesamiento.

45 Según una realización favorable del sistema de procesamiento de datos de la invención, los búferes de estado de entrada/salida pueden recopilar estados de mensajes del nido del sistema y/o del controlador del bus de entrada/salida, en particular un estado de finalización del nido del sistema. De esta manera, la información sobre el estado de finalización de las diferentes instrucciones de almacenamiento puede manejarse de manera ordenada y eficiente.

Según una realización favorable del sistema de procesamiento de datos de la invención, los estados de los mensajes y/o el estado de finalización pueden numerarse mediante un índice de búfer de estado de entrada/salida. La numeración permite gestionar los mensajes y, en particular, los estados de finalización de forma ordenada y eficiente para procesar posteriormente otras instrucciones de almacenamiento.

50 Según una realización favorable del sistema de procesamiento de datos de la invención, el búfer de agregación puede acoplarse comunicativamente a la interfaz asíncrona núcleo-nido a través de un bus asíncrono. Por lo tanto, el búfer de agregación puede gestionar los datos enviados directamente por la interfaz asíncrona núcleo-nido de forma consecutiva hasta que todos los datos que se transferirán al dispositivo externo se almacenen en el búfer de agregación. De esta manera, el mecanismo de transmisión asíncrona para la transferencia de datos desde la interfaz

asíncrona núcleo-nido puede soportarse favorablemente.

5 Según una realización favorable del sistema de procesamiento de datos de la invención, los datos pueden transferirse mediante la instrucción de almacenamiento de entrada/salida a través de un mecanismo de transmisión asíncrono con un mensaje de finalización temprana en múltiples paquetes de datos al búfer de agregación, si la longitud de los datos de origen supera los ocho bytes, de lo contrario, los datos pueden transferirse en un paquete de datos. El mecanismo de transmisión asíncrona es favorable porque el dispositivo emisor puede reutilizarse libremente en un estado anterior.

10 Según una realización favorable del sistema de procesamiento de datos de la invención, el firmware del sistema puede comprender un código controlador de entrada/salida asíncrono para gestionar la instrucción de almacenamiento de entrada/salida. Por lo tanto, se puede usar un mecanismo de transmisión asíncrono para transferir datos desde la unidad de procesamiento de datos al dispositivo externo.

15 Según una realización favorable del sistema de procesamiento de datos de la invención, el núcleo puede comprender un código de configuración asíncrono para gestionar los requisitos de memoria para la información de estado del código controlador de entrada/salida asíncrono. Este código de configuración asíncrono puede facilitar aún más el mecanismo de transmisión asíncrona a través del búfer de agregación hasta el nido del sistema y el controlador del bus de entrada/salida.

20 Según una realización favorable del sistema de procesamiento de datos inventivo, la interfaz asíncrona núcleo-nido puede comprender un componente asíncrono de reenvío de interfaz núcleo-nido para reenviar los datos con finalización local. Este componente se puede implementar en el hardware de la interfaz núcleo-nido asíncrona. Por lo tanto, se puede soportar un modo de transmisión asíncrono favorable para enviar los datos en paquetes de datos al búfer de agregación.

Según una realización favorable del sistema de procesamiento de datos de la invención, el búfer de agregación puede comprender una lógica de finalización temprana para entregar un mensaje gratuito para su reutilización después de enviar una solicitud. Esto permite una continuación temprana del procedimiento de transmisión de los datos a través del búfer de agregación al nido del sistema y al controlador del bus de entrada/salida.

25 Según una realización favorable del sistema de procesamiento de datos de la invención, el firmware del sistema puede comprender una lógica de gestión de matrices, que asigna/desasigna los búferes de estado de entrada/salida en la matriz de estado de entrada/salida y/o inicia el inicio de una nueva instrucción de almacenamiento. Por lo tanto, los búferes de estado inactivo pueden atribuirse a instrucciones de almacenamiento adicionales. Un procesamiento ordenado de las instrucciones de almacenamiento se puede gestionar de forma eficiente y ahorrando tiempo.

30 Según una realización favorable del sistema de procesamiento de datos de la invención, un mensaje del sistema puede comprender: una dirección diana física jerárquica, el origen de un subprocedimiento SMT (subprocesamiento múltiple simultáneo) o un identificador de búfer agregado, una longitud de datos, una dirección de bus de entrada/salida o un índice de búfer de estado de entrada/salida. De este modo, se puede garantizar una transmisión ventajosa de la información relevante a través del sistema de procesamiento de datos.

35 Además, se propone un procedimiento para gestionar una instrucción de almacenamiento de entrada/salida para al menos un dispositivo externo de un sistema de procesamiento de datos, comprendiendo el sistema de procesamiento de datos un nido de sistemas acoplado comunicativamente a al menos un bus de entrada/salida mediante un controlador de bus de entrada/salida. El sistema de procesamiento de datos comprende además al menos una unidad de procesamiento de datos que comprende un núcleo, un firmware del sistema y una interfaz núcleo-nido asíncrona.
40 La unidad de procesamiento de datos se acopla comunicativamente al nido del sistema a través de un búfer de agregación. El dispositivo externo está acoplado comunicativamente al bus de entrada/salida. La interfaz núcleo-nido asíncrona comprende una matriz de estado de entrada/salida con varios búferes de estado de entrada/salida, así como una lógica de acceso y administración de matrices.

45 El procedimiento comprende: (i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos que emite la instrucción de almacenamiento de entrada/salida que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos que se transferirán y/o un puntero a los datos que se transferirán, y una longitud de los datos; (ii) la unidad de procesamiento de datos está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida; (iii) la unidad de procesamiento de datos está configurado para verificar si el acceso a la función de entrada/salida está permitido
50 en un espacio de direcciones y a nivel de instancia invitada, donde el huésped funciona en el sistema de procesamiento de datos; (iv) la unidad de procesamiento de datos está configurado para completar la instrucción de almacenamiento de entrada/salida antes de que se complete la ejecución de la instrucción de almacenamiento de entrada/salida en el nido del sistema; (v) el firmware del sistema está configurado para notificar al sistema operativo mediante una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida la unidad de procesamiento de datos detecta un error al transmitir los datos de la ejecución asíncrona fallida; (vi) la lógica de administración y acceso a la matriz recopila la finalización de la instrucción de almacenamiento y actualiza los búferes de estado de entrada/salida según los mensajes de finalización recibidos; y (vii) la unidad de procesamiento de datos retrasa la ejecución de la instrucción de almacenamiento, hasta que esté disponible un búfer de estado de
55

entrada/salida para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes.

De este modo, se pueden permitir favorablemente múltiples instrucciones de almacenamiento asíncronas pendientes al mismo tiempo para reducir los ciclos por instrucción de las instrucciones de almacenamiento asíncronas repetidas. Un pedido se define entre instrucciones de almacenamiento asíncronas e instrucciones de carga/almacenamiento sincrónicas. La compatibilidad con varias instrucciones de almacenamiento asíncronas pendientes se basa en la contabilidad de varios mensajes de estado y en la correlación de las respuestas con las entradas de estado.

El procedimiento según una realización adicional de la invención comprende la carga y el almacenamiento de instrucciones en un dispositivo externo del sistema de procesamiento de datos a través de un bus de entrada/salida. Las instrucciones asíncronas se completan antes de que los datos se almacenen en el dispositivo externo, mientras que las instrucciones sincrónicas se completan después de que los datos se almacenen en el dispositivo externo. Dentro de las realizaciones aquí descritas, la PCI se utilizará indistintamente para cualquier otra tecnología de entrada/salida, sin restringir así la forma de realización de la invención a la PCI.

Las realizaciones de la invención describen la ejecución de una instrucción de almacenamiento de entrada/salida de una manera estrictamente ordenada, tal como se puede observar desde arriba del límite de la arquitectura, mientras que la ejecución real puede estar fuera de orden dentro del hardware de la unidad de procesamiento de datos (CPU).

Según algunas realizaciones del procedimiento de la invención, una instrucción de almacenamiento PCI puede ejecutarse con una ejecución asíncrona del efecto de almacenamiento PCIe y una gestión de estado asíncrona. La ejecución fiable asíncrona se basa en mecanismos de reenvío fiables en la microarquitectura del sistema de procesamiento de datos de la invención.

Una instrucción PCI de almacenamiento y bloque de almacenamiento existente suele ser sincrónica hasta el punto en que los datos de almacenamiento PCI se entregan a la interfaz PCIe y se devuelven su finalización a una unidad de procesamiento.

El estándar PCI solo requiere un comando de envío asíncrono de información PCI, que normalmente se implementa a través de una cola de almacenamiento en el procesador que agrega datos con un envío asíncrono.

Ventajosamente, según las realizaciones del procedimiento de la invención, se puede lograr una mejora con respecto a los ciclos por instrucción reemplazando una instrucción PCI sincrónica por un procedimiento de envío asíncrono fiable de una instrucción de almacenamiento de entrada/salida.

Alternativa o adicionalmente a los datos que se van a transferir, la instrucción de almacenamiento según una realización de la invención también puede especificar un puntero a una memoria principal que debe usarse para extraer datos, en lugar de contener los datos directamente.

El nivel de instancia invitada también puede significar que un único huésped o anfitrión puede estar ejecutándose en el sistema de procesamiento de datos.

La dirección del desplazamiento de la función de entrada/salida en sí puede ser una dirección virtual, física o lógica. Las direcciones virtuales y lógicas generalmente se traducen a través de una unidad de administración de memoria (MMU - memory management unit) a una dirección física, y la dirección física permite identificar a qué función y desplazamiento se refieren.

En este contexto, la dirección física significa «la dirección más baja en la jerarquía de traducción de direcciones accesible desde un sistema operativo/huésped».

Según una realización favorable del procedimiento de la invención, la unidad de procesamiento de datos puede retrasar la ejecución de instrucciones de almacenamiento asíncronas y/o sincrónicas, hasta que esté disponible un búfer de estado de entrada/salida para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes. Por lo tanto, un procedimiento de pedido de un múltiplo de instrucciones de almacenamiento a un múltiplo de dispositivos de entrada/salida puede gestionarse de una manera eficiente, lo que ofrece la capacidad de ahorrar tiempo de procesamiento.

Según una realización favorable del procedimiento de la invención, los búferes de estado de entrada/salida pueden recopilar estados de mensajes del nido del sistema y/o del controlador del bus de entrada/salida, en particular un estado de finalización del nido del sistema, en donde los estados del mensaje y/o el estado de finalización se numeran mediante un índice de búfer de estado de entrada/salida. De esta manera, la información sobre el estado de finalización de las diferentes instrucciones de almacenamiento puede manejarse de manera ordenada y eficiente. La numeración permite gestionar los mensajes y, en particular, los estados de finalización de forma ordenada y eficiente para procesar posteriormente otras instrucciones de almacenamiento.

Según una realización favorable del procedimiento inventivo, el firmware del sistema puede comprender una lógica de gestión de matriz, que asigna/desasigna búferes de estado de entrada/salida en la matriz de estado de entrada/salida y/o inicia un inicio de una nueva instrucción de almacenamiento. Por lo tanto, los búferes de estado inactivo pueden

atribuirse a instrucciones de almacenamiento adicionales. Un procesamiento ordenado de las instrucciones de almacenamiento se puede gestionar de forma eficiente y ahorrando tiempo.

Según una realización favorable, el procedimiento puede comprender además: (i) el sistema operativo que emite la instrucción de almacenamiento de entrada/salida; (ii) el firmware del sistema (10) asigna un índice de búfer de estado de entrada/salida libre; si no hay ningún índice de búfer de estado de entrada/salida libre disponible, entonces espera a que aparezca un índice de búfer de estado de entrada/salida libre; (iii) el firmware del sistema que inyecta la instrucción de almacenamiento en el motor de envío asíncrono; si está bloqueado por otra instrucción de almacenamiento, espera hasta que se complete la instrucción de almacenamiento; (iv) según la longitud de los datos: si la longitud de los datos excede ocho bytes, el firmware del sistema emite repetidamente un mensaje del sistema para enviar un paquete de datos al búfer de agregación hasta que todos los datos de un bloque de almacenamiento se hayan enviado al búfer de agregación, mientras que el firmware del sistema espera hasta que los datos hayan sido enviados por el mensaje del sistema; de lo contrario, el firmware del sistema emite un mensaje del sistema para enviar los datos al búfer de agregación; además, independientemente de la longitud de los datos, (v) el firmware del sistema envía un mensaje de sistema al búfer de agregación para reenviar los datos de forma asíncrona como un mensaje de nido único al controlador del bus de entrada/salida, mientras espera a que el búfer de agregación envíe un mensaje de finalización; (vi) el búfer de agregación que inyecta el mensaje del nido en el nido del sistema, en donde el búfer de agregación está libre para su reutilización inmediatamente después de la operación de envío, enviando una señal al firmware del sistema; a continuación, el búfer de agregación envía un mensaje de libre reutilización; (vii) el sistema reenvía el mensaje a la ubicación de destino; (viii) el controlador del bus de entrada/salida que recibe el mensaje y reenvía los datos en una trama de datos al bus de entrada/salida; (ix) el controlador de bus de entrada/salida envía un mensaje de finalización al nido del sistema; (x) el nido del sistema reenvía el mensaje de finalización al búfer de agregación de origen; (xi) el búfer de agregación reenvía la finalización a la interfaz asíncrona núcleo-nido; (xii) la interfaz asíncrona núcleo-nido almacena el estado de finalización en el búfer de estado de entrada/salida para el índice del búfer de estado de entrada/salida y la señalización de la finalización de la operación al firmware del sistema; (xiii) el firmware del sistema actualiza un búfer de estado de entrada/salida que rastrea mediante la entrada/ índice de búfer de estado de salida; y (xiv) el firmware del sistema que señala asincrónicamente los defectos al sistema operativo en caso de error.

Etapas (ii) depende de la longitud de los datos y es diferente para la longitud de los datos que supere los ocho bytes y para la longitud de los datos que no supere los ocho bytes.

Según la realización del procedimiento de la invención, los datos se transmiten en segmentos al búfer de agregación hasta que todos los datos de un bloque de almacenamiento se envíen al búfer de agregación, en donde el firmware del sistema espera hasta que los datos hayan sido enviados por la interfaz núcleo-nido asíncrona.

Por lo tanto, si los datos tienen menos de ocho bytes, se puede omitir el procedimiento de llenado del búfer de agregación en segmentos con paquetes de datos y el procedimiento de transmisión de los datos al dispositivo externo se puede completar en una sola etapa.

Según una realización favorable del procedimiento inventivo, los datos pueden ser transferidos por la instrucción de almacenamiento de entrada/salida a través de un mecanismo de transmisión asíncrono con un mensaje de finalización temprana en múltiples paquetes de datos al búfer de agregación, si la longitud de los datos excede ocho bytes. El mecanismo de transmisión asíncrona es favorable porque el dispositivo emisor puede reutilizarse libremente en un estado anterior.

Según una realización favorable del procedimiento inventivo, el firmware del sistema puede utilizar un código de controlador de entrada/salida asíncrono para manejar la instrucción de almacenamiento de entrada/salida. Por lo tanto, se puede usar un mecanismo de transmisión asíncrono para transferir datos desde la unidad de procesamiento de datos al dispositivo externo.

Según una realización favorable del procedimiento inventivo, el núcleo puede utilizar un código de configuración asíncrono para gestionar los requisitos de memoria para la información de estado del código de controlador de entrada/salida asíncrono. Este código de configuración asíncrono puede facilitar aún más el mecanismo de transmisión asíncrona a través del búfer de agregación hasta el nido del sistema y el controlador del bus de entrada/salida.

Ventajosamente, la interfaz de núcleo y nido asíncrona puede usar un componente de reenvío de interfaz núcleo-nido asíncrona para reenviar los datos con finalización local. Por lo tanto, se puede soportar un modo de transmisión asíncrono favorable para enviar los datos en paquetes de datos al búfer de agregación.

Ventajosamente, el búfer de agregación puede usar una lógica de finalización temprana para entregar un mensaje gratuito para su reutilización después de enviar una solicitud. Esto permite una continuación temprana del procedimiento de transmisión de los datos a través del búfer de agregación al nido del sistema y al controlador del bus de entrada/salida.

Ventajosamente, los búferes de estado de entrada/salida pueden recopilar los estados devueltos del nido del sistema y/o del controlador del bus de entrada/salida, en particular un mensaje de finalización del nido del sistema. Estos búferes de estado de entrada/salida pueden recopilar los estados devueltos actuando como un búfer de mensajes del

sistema asíncrono que soporta el procedimiento de transmisión asíncrona.

Según una realización favorable del procedimiento inventivo, un mensaje de sistema puede comprender uno de los siguientes: - una dirección de destino física jerárquica, - la obtención de un hilo SMT o un identificador de búfer agregado, - una longitud de datos, - una dirección de bus de entrada/salida, o - un índice de búfer de estado de entrada/salida. De este modo, se puede garantizar una transmisión ventajosa de la información relevante a través del sistema de procesamiento de datos.

Además, se propone un producto de programa informático favorable para manejar una instrucción de almacenamiento de entrada/salida a al menos un dispositivo externo de un sistema de procesamiento de datos, comprendiendo el sistema de procesamiento de datos un nido de sistema acoplado comunicativamente a al menos un bus de entrada/salida mediante un controlador de bus de entrada/salida. El sistema de procesamiento de datos comprende además al menos una unidad de procesamiento de datos que comprende un núcleo, un firmware del sistema y una interfaz núcleo-nido asíncrona. La unidad de procesamiento de datos se acopla comunicativamente al nido del sistema a través de un búfer de agregación. El dispositivo externo está acoplado comunicativamente al bus de entrada/salida. La interfaz núcleo-nido asíncrona comprende una matriz de estado de entrada/salida con varios búferes de estado de entrada/salida, así como una lógica de acceso y administración de matrices.

El producto de programa informático comprende un medio de almacenamiento legible por ordenador que tiene instrucciones de programa incorporadas en el mismo, siendo las instrucciones de programa ejecutables por el sistema informático para hacer que el sistema informático realice un procedimiento que comprende: (i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos que emite la instrucción de almacenamiento de entrada/salida que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos que se transferirán y/o un puntero a los datos que se transferirán, y una longitud de los datos; (ii) la unidad de procesamiento de datos está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida; (iii) la unidad de procesamiento de datos está configurado para verificar si el acceso a la función de entrada/salida está permitido en un espacio de direcciones y a nivel de instancia invitada, donde el huésped funciona en el sistema de procesamiento de datos; (iv) la unidad de procesamiento de datos está configurado para completar la instrucción de almacenamiento de entrada/salida antes de que se complete la ejecución de la instrucción de almacenamiento de entrada/salida en el nido del sistema; (v) el firmware del sistema está configurado para notificar al sistema operativo mediante una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida la unidad de procesamiento de datos detecta un error al transmitir los datos de la ejecución asíncrona fallida; (vi) la lógica de administración y acceso a la matriz recopila la finalización de la instrucción de almacenamiento y actualiza los búferes de estado de entrada/salida según los mensajes de finalización recibidos; y (vii) la unidad de procesamiento de datos retrasa la ejecución de la instrucción de almacenamiento, hasta que esté disponible un búfer de estado de entrada/salida para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes.

Además, se propone un sistema de procesamiento de datos para la ejecución de un programa de procesamiento de datos, que comprende instrucciones de programa legibles por ordenador para realizar el procedimiento descrito anteriormente.

Breve descripción de las varias vistas de los dibujos

La presente invención, junto con los objetos y ventajas mencionados anteriormente y otros, se puede entender mejor a partir de la siguiente descripción detallada de las realizaciones, pero no se limita a las realizaciones.

La Figura 1 representa un diagrama de bloques de un sistema de procesamiento de datos para gestionar una instrucción de almacenamiento de entrada/salida a un dispositivo externo según una realización de la invención.

La Figura 2 representa un diagrama de secuencias de mensajes de un procedimiento para gestionar una instrucción de almacenamiento de entrada/salida a un dispositivo externo según una realización de la invención.

La Figura 3 representa una primera parte de un diagrama de flujo para gestionar una instrucción de almacenamiento de entrada/salida a un dispositivo externo según una realización de la invención.

La Figura 4 representa una segunda parte de un diagrama de flujo para gestionar una instrucción de almacenamiento de entrada/salida a un dispositivo externo según una realización de la invención.

La Figura 5 representa un ejemplo de realización de un sistema de procesamiento de datos para ejecutar un procedimiento según la invención.

Descripción detallada

En los dibujos, los elementos similares se denominan con números de referencia iguales. Los dibujos son simplemente representaciones esquemáticas, no pretenden representar parámetros específicos de la invención. Además, los dibujos pretenden representar solo las realizaciones típicas de la invención y, por lo tanto, no deben considerarse como limitantes del alcance de la invención.

Las realizaciones ilustrativas descritas en esta invención proporcionan un sistema de procesamiento de datos y un procedimiento para gestionar una instrucción de almacenamiento de entrada/salida, que comprende un nido de sistema acoplado comunicativamente a al menos un bus de entrada/salida mediante un controlador de bus de entrada/salida. El sistema de procesamiento de datos comprende además al menos una unidad de procesamiento de datos que comprende un núcleo, un firmware del sistema y una interfaz núcleo-nido asíncrona. La unidad de procesamiento de datos se acopla comunicativamente al nido del sistema a través de un búfer de agregación. El nido del sistema está configurado para cargar y/o almacenar datos de forma asíncrona en al menos un dispositivo externo que está acoplado comunicativamente al bus de entrada/salida. La interfaz núcleo-nido asíncrona comprende una matriz de estado de entrada/salida con varios búferes de estado de entrada/salida, así como una lógica de acceso y administración de matrices.

Las realizaciones ilustrativas pueden usarse para el procedimiento que comprende: (i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos emite la instrucción de almacenamiento de entrada/salida que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos a transferir y/o un puntero a los datos a transferir, y una longitud de los datos; (ii) la unidad de procesamiento de datos está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida; (iii) la unidad de procesamiento de datos está configurado para verificar si el acceso a la función de entrada/salida está permitido en un espacio de direcciones y a nivel de instancia invitada, donde el huésped se ejecuta en el sistema de procesamiento de datos; (iv) la unidad de procesamiento de datos está configurado para completar la instrucción de almacenamiento de entrada/salida antes de que se complete la ejecución de la instrucción de almacenamiento de entrada/salida en el nido del sistema; (v) el firmware del sistema está configurado para notificar al sistema operativo mediante una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida la unidad de procesamiento de datos detecta un error al transmitir los datos de la ejecución asíncrona fallida; (vi) la lógica de administración y acceso a la matriz recopila la finalización de la instrucción de almacenamiento y actualiza los búferes de estado de entrada/salida según los mensajes de finalización recibidos; y (vii) la unidad de procesamiento de datos retrasa la ejecución de la instrucción de almacenamiento, hasta que esté disponible un búfer de estado de entrada/salida para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes.

Alternativa o adicionalmente a los datos que se van a transferir, la instrucción de almacenamiento según una realización de la invención también puede especificar un puntero a una memoria principal que debe usarse para extraer datos, en lugar de contener los datos directamente.

El nivel de instancia invitada también puede significar que un único huésped o anfitrión puede estar ejecutándose en el sistema de procesamiento de datos.

La dirección del desplazamiento de la función de entrada/salida en sí puede ser una dirección virtual, física o lógica. Las direcciones virtuales y lógicas generalmente se traducen a través de una unidad de administración de memoria (MMU) a una dirección física, y la dirección física permite identificar a qué función y desplazamiento se refieren.

En este contexto, la dirección física significa «la dirección más baja en la jerarquía de traducción de direcciones accesible desde un sistema operativo/huésped».

La Figura 1 muestra un diagrama de bloques de un sistema de procesamiento de datos para manejar una instrucción de almacenamiento de entrada/salida a al menos un dispositivo externo según una realización de la invención. El sistema de procesamiento de datos comprende un nido de sistema acoplado comunicativamente a un bus de entrada/salida mediante un controlador de bus de entrada/salida, una unidad de procesamiento de datos que comprende un núcleo, un firmware de sistema y una interfaz núcleo-nido asíncrona. El controlador de bus de entrada/salida también puede estar acoplado a través de múltiples buses de entrada/salida a múltiples dispositivos externos.

La unidad de procesamiento de datos está acoplado comunicativamente al nido de sistema a través de un búfer de agregación. El nido de sistema está configurado para cargar de forma asíncrona y/o almacenar datos en el dispositivo externo que está acoplado comunicativamente al bus de entrada/salida a través de la interfaz de controlador de bus de entrada/salida de búfer como parte del nido de sistema y el controlador de bus de entrada/salida.

El búfer de agregación está acoplado comunicativamente a la interfaz asíncrona núcleo-nido. El firmware del sistema comprende un código controlador de entrada/salida asíncrono para gestionar la instrucción de almacenamiento de entrada/salida. El núcleo comprende un código de configuración asíncrono para gestionar los requisitos de memoria para la información de estado del código controlador de entrada/salida asíncrono. La interfaz núcleo-nido asíncrona comprende un componente de reenvío de interfaz núcleo-nido asíncrona para reenviar los datos con finalización local. El búfer de agregación comprende una lógica de finalización temprana para entregar un mensaje libre para su reutilización después de enviar una solicitud. El búfer de agregación está acoplado a la interfaz asíncrona núcleo-nido a través de un bus asíncrono. La interfaz núcleo-nido asíncrona comprende una matriz de estado de entrada/salida con múltiples búferes de estado de entrada/salida, así como una lógica de acceso y gestión de matriz. Los búferes de estado de entrada/salida recogen estados devueltos

desde el nido del sistema 18 y/o desde el controlador de bus de entrada/salida 20, en particular un mensaje de finalización del nido del sistema 18. Los búferes de estado de entrada/salida 24 están integrados directamente en la interfaz asíncrona núcleo-nido 14. El nido del sistema 18 puede recibir un mensaje 48 con la identificación de una entrada de matriz, por ejemplo, un mensaje de finalización a una de las búferes de estado de entrada/salida 24.

5 Según una realización del procedimiento inventivo, un sistema operativo que se ejecuta en el sistema de procesamiento de datos 210 emite la instrucción de almacenamiento de entrada/salida 30 que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos a transferir y/o un puntero a datos a transferir, y una longitud de los datos. La unidad de procesamiento de datos 216 está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de
10 entrada/salida 30. La unidad de procesamiento de datos 216 está configurado para verificar si se permite el acceso a la función de entrada/salida en un espacio de direcciones y en un nivel de instancia de invitado, ejecutándose el invitado en el sistema de procesamiento de datos 210. La unidad de procesamiento de datos 216 está configurado para completar la instrucción de almacenamiento de entrada/salida 30 antes de que se complete una ejecución de la instrucción de almacenamiento de entrada/salida 30 en el nido de sistema 18. El firmware del sistema 10 está
15 configurado para notificar al sistema operativo a través de una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida 30 se detecta un error por la unidad de procesamiento de datos 216, transmitiendo los datos de la ejecución asíncrona fallida.

La lógica de acceso y gestión de matrices 46 recopila una finalización de la instrucción de almacenamiento 30 y actualiza los búferes de estado de entrada/salida 24 según los mensajes de finalización recibidos. La unidad de
20 procesamiento de datos 216 retrasa una ejecución de la instrucción de almacenamiento 30, hasta que un búfer de estado de entrada/salida 24 esté disponible para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes 30. En particular, la unidad de procesamiento de datos 216 retrasa una ejecución de instrucciones de almacenamiento asíncronas y/o sincrónicas 30, hasta que un búfer de estado de entrada/salida 24 esté disponible para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes 30.

25 Los búferes de estado de entrada/salida 24 recogen estados de mensajes del nido de sistema 18 y/o del controlador de bus de entrada/salida 20, en particular un estado de finalización del nido de sistema 18. Los estados del mensaje y/o el estado de finalización pueden numerarse favorablemente mediante un índice de búfer de estado de entrada/salida.

30 El firmware del sistema 10 comprende una lógica de gestión de matrices 42, que asigna/desasigna los búferes de estado de entrada/salida 24 en la matriz de estado de entrada/salida 44 y/o inicia el inicio de una nueva instrucción de almacenamiento 30.

La instrucción de almacenamiento de entrada/salida 30 está ubicada en el sistema de procesamiento de datos 210 en el lado de la interfaz de usuario 40 a través del límite de la arquitectura que separa el hardware/firmware del sistema 50 del lado del usuario 40.

35 Por lo tanto, los datos se transfieren mediante la instrucción de almacenamiento de entrada/salida 30 a través de un mecanismo de transmisión asíncrono con un mensaje de finalización temprana en múltiples paquetes de datos al búfer de agregación 16, si la longitud de los datos de origen supera los ocho bytes, de lo contrario, los datos se transfieren en un paquete de datos.

40 Un mensaje de sistema según una realización del sistema de procesamiento de datos inventivo comprende una dirección de destino física jerárquica, un hilo SMT de origen o un identificador de búfer agregado, una longitud de datos, una dirección de bus de entrada/salida o un índice de búfer de estado de entrada/salida.

45 La semántica de hacer cola y hacer pedidos para gestionar las instrucciones de almacenamiento 30 a múltiples dispositivos externos 214 puede realizarse ventajosamente como se describe a continuación. Para una relación entre un hilo SMT individual y una función de entrada/salida, todas las operaciones de carga/almacenamiento de entrada/salida heredadas pueden ordenarse con respecto a un único hilo de la unidad del procesador 216. Las nuevas instrucciones de almacenamiento de entrada/salida están completamente desordenadas entre sí. Las nuevas instrucciones de almacenamiento de entrada/salida se ordenan comparándolas con las instrucciones de entrada/salida antiguas. Todas las instrucciones de entrada/salida para diferentes funciones de entrada/salida no están ordenadas unas contra otras.

50 La Figura 2 muestra un diagrama de secuencia de mensajes del procedimiento para manejar una instrucción de almacenamiento de entrada/salida 30 a un dispositivo externo 214 según una realización de la invención.

55 Como se muestra en la Figura 2, el procedimiento comienza con el sistema operativo emitiendo la instrucción 30 de almacenamiento de entrada/salida. En la etapa S101, el firmware del sistema 10 asigna un índice de búfer de estado de entrada/salida libre. Si no hay ningún índice de búfer de estado de entrada/salida disponible, el firmware del sistema 10 espera. En la etapa S103, el firmware del sistema 10 comprueba si la instrucción de almacenamiento puede inyectarse en un motor de envío asíncrono. Si esto es posible, el procedimiento continúa. Si esto no es posible, la instrucción de almacenamiento se retrasa hasta que se hayan completado las instrucciones de almacenamiento que causaron el retraso.

5 A continuación, como se indica en las etapas S100 y S104, el firmware del sistema 10 emite repetidamente, si la longitud de los datos supera los ocho bytes, un mensaje del sistema para enviar un paquete de datos al búfer de agregación 16 hasta que todos los datos de un bloque de almacenamiento se hayan reenviado al búfer de agregación 16, mientras que el firmware del sistema 10 espera hasta que los datos hayan sido enviados por el mensaje del sistema. En las etapas S102 y S106, se envía un mensaje de finalización local al firmware 10 del sistema.

Luego, en la etapa S108, el firmware del sistema 10 emite un mensaje de sistema al búfer de agregación 16 para reenviar los datos de manera asíncrona como un mensaje de anidamiento único al controlador de bus de entrada/salida 20, mientras espera que el búfer de agregación 16 envíe un mensaje de finalización.

10 A continuación, en la etapa S110, el búfer de agregación 16 inyecta el mensaje de anidamiento en el nido del sistema 18, en donde el búfer de agregación 16 está libre para su reutilización inmediatamente después de la operación de envío, enviando una señal al firmware del sistema 10. A continuación, el búfer de agregación 16 envía un mensaje de libre reutilización.

15 En la etapa S112, el nido de sistema 18 reenvía el mensaje a la ubicación diana, seguido de la etapa S114, donde el controlador de bus de entrada/salida 20 recibe el mensaje y reenvía los datos en una trama de datos al bus de entrada/salida, seguido de que el controlador de bus de entrada/salida 20 envíe un mensaje de finalización al nido de sistema 18 en la etapa S116.

20 A continuación, en la etapa S118, el nido de sistema 18 reenvía el mensaje de finalización al búfer de agregación de origen 16, seguido por el búfer de agregación 16 que reenvía la finalización a la interfaz asíncrona núcleo-nido 14 en la etapa S120. A continuación, en la etapa S122, la interfaz núcleo-nido asíncrona 14 almacena el estado en el búfer de entrada/salida 24 para el índice de búfer de estado de entrada/salida respectivo y señala la finalización de la operación al firmware del sistema 10. Finalmente, en la etapa S123, el firmware del sistema 10 actualiza el búfer de estado de entrada/salida 24 siguiendo el índice del búfer de estado de entrada/salida. El búfer de estado de entrada/salida 24 está ahora libre de nuevo.

25 En caso de que se produzca un error durante la transferencia de datos, el firmware del sistema 10 señala los defectos de forma asíncrona al sistema operativo.

En caso de que los datos a transferir sean inferiores a ocho bytes, se omite el llenado repetido del búfer de agregación 16.

La Figura 3 representa una primera parte de un diagrama de flujo para gestionar una instrucción de almacenamiento de entrada/salida 30 a un dispositivo externo 214 según una realización de la invención, mientras que la Figura 4 representa una segunda parte del diagrama de flujo.

30 El firmware del sistema de la unidad de procesamiento de datos comienza en la etapa S200. En la etapa S202, el firmware del sistema recibe mediante un mensaje una instrucción de bloque de almacenamiento de entrada/salida heredada. En la etapa S208 se comprueba si hay una nueva instrucción de almacenamiento 30 pendiente con o sin una respuesta asíncrona. Si este es el caso, se espera hasta que la interfaz núcleo-nido 14 no tenga más respuestas asíncronas pendientes. Si este no es el caso, el procedimiento continúa inmediatamente con la etapa S212 mediante el envío de la instrucción de almacenamiento de entrada/salida 30 heredada como un mensaje por la interfaz núcleo-nido 14. A continuación, el sistema espera, etapa S214, hasta que reciba un mensaje de respuesta heredado en la etapa S216. A continuación, la instrucción de almacenamiento heredado 30 finaliza en la etapa S218 con un mensaje y el firmware del sistema 10 finaliza el procedimiento en la etapa S220.

35 En paralelo, el firmware del sistema 10 recibe la nueva instrucción de almacenamiento de entrada/salida 30 mediante un mensaje en la etapa S204. En la etapa S222 se comprueba si está disponible una ranura de estado libre, es decir, un búfer de estado de entrada/salida 24. Si este es el caso, la ranura se marca en la etapa S224 como usada y el procedimiento de almacenamiento continúa con el punto de conexión A, representado en la segunda parte del diagrama de flujo representado en la Figura 4. Si este no es el caso, se espera hasta que la interfaz núcleo-nido 14 tenga una ranura libre disponible en la etapa S226.

40 En la etapa S206 se puede recibir un mensaje de ejecución completa asíncrona, seguido de una actualización del seguimiento de intervalos en la etapa S228 con el índice de búfer de estado de entrada/salida respectivo. A continuación, el firmware del sistema 10 finaliza el procedimiento en la etapa S230.

45 La segunda parte del diagrama de flujo, que comienza con el punto de conexión A, se representa en la Figura 4. En primer lugar, en la etapa S304, se comprueba si se van a transferir los más de 8 bytes. Si este es el caso, la interfaz núcleo-nido llena el búfer de agregación con un mensaje de hasta 16 bytes en la etapa S306. El firmware del sistema está esperando, etapa S308, hasta que se envíe un mensaje de finalización local en la etapa S310, volviendo a la etapa S304. Si quedan menos de 8 bytes en la comprobación de la etapa S304, el flujo continúa en la etapa S312 con la interfaz núcleo-nido enviando un mensaje de entrada/salida asíncrono, seguido de esperar en la etapa S314 una respuesta del búfer en la etapa S316. A continuación, en la etapa S318, se ejecuta una instrucción de finalización de almacenamiento del bloque y el flujo finaliza en la etapa S320 con una finalización en el firmware del sistema.

50 55 En la etapa S328, la lógica de la interfaz asíncrona núcleo-nido inicia un bucle de procedimiento de salida, seguido de

la recepción de un mensaje de finalización del búfer de agregación en la etapa S322 y un mensaje de reenvío de datos al búfer de agregación en la etapa S324, seguido de un mensaje de finalización de envío al firmware del sistema en la etapa S326. En la etapa S330, se recibe un mensaje de envío de entrada/salida asíncrono seguido de un reenvío del mensaje de envío de entrada/salida al búfer de agregación.

5 En la etapa S338, la lógica del búfer de agregación inicia un bucle de procedimiento de salida seguido de una recepción de datos en la etapa S334 y la agregación de datos en el búfer de agregación en la etapa S336. El búfer de agregación también recibe un mensaje de envío de entrada/salida en la etapa S340, seguido de reenviar datos desde el búfer de agregación con un mensaje de envío de entrada/salida en la etapa S242. A continuación, en la etapa S344, se envía un mensaje de respuesta desde el búfer de agregación a través de la interfaz núcleo-nido al firmware del sistema.

10 Haciendo referencia ahora a la Figura 5, se muestra un esquema de un ejemplo de un sistema de procesamiento de datos 210. El sistema de procesamiento de datos 210 es sólo un ejemplo de un sistema de procesamiento de datos adecuado y no pretende sugerir ninguna limitación en cuanto al alcance de uso o funcionalidad de las realizaciones de la invención descrita en este documento. De todos modos, el sistema de procesamiento de datos 210 es capaz de implementarse y/o realizar cualquiera de las funcionalidades establecidas anteriormente.

15 En el sistema de procesamiento de datos 210 hay un sistema informático/servidor 212, que es operativo con numerosos otros entornos o configuraciones de sistemas informáticos de propósito general o de propósito especial. Ejemplos de sistemas, entornos y/o configuraciones informáticas bien conocidos que pueden ser adecuados para su uso con el sistema informático/servidor 212 incluyen, entre otros, sistemas informáticos personales, sistemas informáticos de servidor, clientes ligeros, clientes pesados, dispositivos portátiles o de mano, sistemas multiprocesador, sistemas basados en microprocesador, decodificadores, productos electrónicos programables para el consumidor, PC en red, sistemas de miniordenadores, sistemas informáticos mainframe y entornos informáticos distribuidos en la nube que incluyen cualquiera de los sistemas o dispositivos anteriores, y similares.

20 El sistema informático/servidor 212 puede describirse en el contexto general de instrucciones ejecutables del sistema informático, tales como módulos de programa, que son ejecutados por un sistema informático. En general, los módulos de programa pueden incluir rutinas, programas, objetos, componentes, lógica, estructuras de datos, etc., que realizan tareas particulares o implementan tipos de datos abstractos particulares. El sistema informático/servidor 212 puede utilizarse en entornos de computación en nube distribuida donde las tareas son realizadas por dispositivos de procesamiento remoto que están vinculados a través de una red de comunicaciones. En un entorno de computación en la nube distribuida, los módulos del programa pueden estar ubicados en medios de almacenamiento del sistema informático tanto local como remoto incluidos los dispositivos de almacenamiento de memoria.

25 Como se muestra en la Fig. 5, el sistema informático/servidor 212 en el sistema de procesamiento de datos 210 se muestra en forma de un dispositivo informático de propósito general. Los componentes del sistema informático/servidor 212 pueden incluir, pero no se limitan a, uno o más procesadores o unidades de procesamiento 216, una memoria de sistema 228 y un bus 218 que acopla varios componentes del sistema incluyendo la memoria 228 del sistema al procesador 216.

30 El bus 218 representa uno o más de cualquiera de varios tipos de estructuras de bus, incluyendo un bus de memoria o controlador de memoria, un puerto de gráficos acelerados, y un procesador o bus local que usa cualquiera de una variedad de arquitecturas de buses. A modo de ejemplo, y no de limitación, tales arquitecturas incluyen el bus de la Arquitectura Estándar de la Industria (ISA - Industry Standard Architecture), el bus de Arquitectura Micro Canal (MCA - Micro Channel Architecture), el bus de ISA mejorado (EISA - Enhanced ISA), el bus local de la Asociación de Estándares de Electrónica y Vídeo (VESA - Video Electronics Standards Association) y el bus de Interconexión de Componentes Periféricos (PCI - Peripheral Component Interconnect).

35 El sistema informático/servidor 212 normalmente incluye una variedad de medios legibles por el sistema informático. Dichos medios pueden ser cualquier medio disponible que sea accesible por el sistema informático/servidor 212, e incluyen tanto medios volátiles como no volátiles, medios extraíbles y no extraíbles.

40 La memoria del sistema 228 puede incluir medios legibles por el sistema informático en forma de memoria volátil, tal como la memoria de acceso aleatorio (RAM) 230 y/o la memoria caché 232. El sistema informático/servidor 212 puede incluir además otros medios de almacenamiento de sistema informático volátiles/no volátiles, extraíbles/no extraíbles. A modo de ejemplo solamente, el sistema de almacenamiento 234 se puede proporcionar para leer desde y escribir en un medio magnético no volátil y no extraíble (no mostrado y normalmente denominado «disco duro»). Aunque no se muestra, se pueden proporcionar una unidad de disco magnético para leer desde y escribir en un disco magnético extraíble, no volátil (por ejemplo, un «disco flexible») y una unidad de disco óptico para leer desde o escribir en un disco óptico extraíble, no volátil, tal como un CD-ROM, DVD-ROM u otros medios ópticos. En tales casos, cada uno puede conectarse al bus 218 mediante una o más interfaces de medios de datos. Como se representará y describirá adicionalmente a continuación, la memoria 228 puede incluir al menos un producto de programa que tiene una unidad (por ejemplo, al menos uno) de módulos de programa que están configurados para llevar a cabo las funciones de las realizaciones de la invención.

45 El programa/utilidad 240, que tiene una unidad (al menos una) de módulos de programa 242, puede almacenarse en

la memoria 228 a modo de ejemplo, y no de limitación, así como un sistema operativo, uno o más programas de aplicaciones, otros módulos de programa y datos de programa. Cada uno del sistema operativo, uno o más programas de aplicaciones, otros módulos de programa y datos de programa o alguna combinación de los mismos, pueden incluir una implementación de un entorno de interconexión de redes. Los módulos de programa 242 generalmente llevan a cabo las funciones y/o metodologías de las realizaciones de la invención como se describe en esta invención.

El sistema informático/servidor 212 también puede comunicarse con uno o más dispositivos externos 214, tales como un teclado, un dispositivo señalador, una pantalla 224, etc.; uno o más dispositivos que permiten a un usuario interactuar con el sistema informático/servidor 212; y/o cualquier dispositivo (por ejemplo, tarjeta de red, módem, etc.) que permita al sistema informático/servidor 212 comunicarse con uno o más de otros dispositivos informáticos. Tal comunicación puede ocurrir a través de las interfaces 222 de Entrada/Salida (E/S). Más aún, el sistema informático/servidor 212 puede comunicarse con una o más redes, tales como una red de área local (LAN), una red de área amplia general (WAN) y/o una red pública (por ejemplo, Internet) a través del adaptador de red 220. Como se representa, el adaptador de red 220 se comunica con los otros componentes del sistema informático/servidor 212 a través del bus 218. Se debería entender que, aunque no se muestran, se podrían utilizar otros componentes de hardware y/o software junto con el sistema informático/servidor 212. Los ejemplos incluyen, pero no se limitan a: microcódigo, controladores de dispositivo, unidades de procesamiento redundantes, matrices de unidades de disco externo, sistemas RAID, unidades de cinta y sistemas de almacenamiento de archivado de datos, etc.

La presente invención puede ser un sistema, un procedimiento y/o un producto de programa informático. El producto de programa informático puede incluir un medio (o medios) de almacenamiento legible por ordenador que tiene instrucciones de programa legibles por ordenador en el mismo para hacer que un procesador lleve a cabo aspectos de la presente invención.

El medio de almacenamiento legible por ordenador puede ser un dispositivo tangible que puede retener y almacenar instrucciones para su uso mediante un dispositivo de ejecución de instrucciones. El medio de almacenamiento legible por ordenador puede ser, por ejemplo, entre otros, un dispositivo de almacenamiento electrónico, un dispositivo de almacenamiento magnético, un dispositivo de almacenamiento óptico, un dispositivo de almacenamiento electromagnético, un dispositivo de almacenamiento semiconductor o cualquier combinación adecuada de los anteriores. Una lista no exhaustiva de ejemplos más específicos del medio de almacenamiento legible por ordenador incluye lo siguiente: un disquete de ordenador portátil, un disco duro, una memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), una memoria de solo lectura programable y borrable (EPROM o memoria Flash), una memoria de acceso aleatorio estática (SRAM), un disco compacto portátil de memoria de solo lectura (CD-ROM), un disco versátil digital (DVD), una tarjeta de memoria, un disco, un dispositivo codificado mecánicamente, tal como tarjetas perforadas o estructuras en relieve en un surco que tiene instrucciones grabadas en el mismo, y cualquier combinación adecuada de los anteriores. Un medio de almacenamiento legible por ordenador, tal como se utiliza en la presente memoria, no debe ser interpretado como señales transitorias en sí, tales como ondas de radio u otras ondas electromagnéticas que se propagan libremente, ondas electromagnéticas que se propagan a través de una guía de ondas u otro medio de transmisión (por ejemplo, pulsos de luz que pasan a través de un cable de fibra óptica) o señales eléctricas transmitidas a través de un cable.

Las instrucciones del programa legibles por ordenador descritas en la presente memoria pueden descargarse a los respectivos dispositivos informáticos/de procesamiento desde un medio de almacenamiento legible por ordenador o a un ordenador externo o dispositivo de almacenamiento externo a través de una red, por ejemplo, Internet, una red de área local, una red de área amplia y/o una red inalámbrica. La red puede comprender cables de transmisión de cobre, fibras de transmisión óptica, transmisión inalámbrica, enrutadores, firewalls, conmutadores, ordenadores de puerta de enlace y/o servidores periféricos. Una tarjeta de adaptador de red o interfaz de red en cada dispositivo informático/de procesamiento recibe instrucciones de programa legibles por ordenador desde la red y reenvía las instrucciones de programa legibles por ordenador para su almacenamiento en un medio de almacenamiento legible por ordenador dentro del dispositivo informático/de procesamiento respectivo.

Las instrucciones de programa legibles por ordenador para llevar a cabo las operaciones de la presente invención pueden ser instrucciones de ensamblador, instrucciones de arquitectura de conjunto de instrucciones (ISA), instrucciones de máquina, instrucciones dependientes de máquina, microcódigo, instrucciones de firmware, datos de ajuste de estado, o código fuente o código objeto escritos en cualquier combinación de uno o más lenguajes de programación, incluyendo un lenguaje de programación orientado a objetos como Smalltalk, C++ o similares, y lenguajes de programación procedimentales convencionales, como el lenguaje de programación «C» o lenguajes de programación similares. Las instrucciones de programa legibles por ordenador pueden ser ejecutadas completamente en el ordenador del usuario, parcialmente en el ordenador del usuario, como un paquete de software independiente, parcialmente en el ordenador del usuario y parcialmente en un ordenador remoto o completamente en el ordenador o servidor remoto. En este último escenario, el ordenador remoto puede conectarse al ordenador del usuario a través de cualquier tipo de red, incluyendo una red de área local (LAN) o una red de área amplia (WAN), o la conexión puede realizarse a un ordenador externo (por ejemplo, a través de Internet usando un Proveedor de Servicios de Internet). En algunas realizaciones, circuitos electrónicos incluyen, por ejemplo, circuitos lógicos programables, matrices de puertas programables en campo (Field-Programmable Gate Arrays, FPGA) o matrices lógicas programables (Programmable Logic Arrays, PLA) que pueden ejecutar instrucciones de programa legibles por ordenador utilizando información de estado de las instrucciones de programa legibles por ordenador para personalizar los circuitos

electrónicos, con el fin de realizar aspectos de la presente invención.

Los aspectos de la presente invención se describen en esta invención con referencia a ilustraciones de diagramas de flujo y/o diagramas de bloques de procedimientos, aparatos (sistemas) y productos de programas informáticos según las realizaciones de la invención. Se entenderá que cada bloque de las ilustraciones de diagrama de flujo y/o diagramas de bloques, y las combinaciones de bloques en las ilustraciones de diagrama de flujo y/o diagramas de bloques, pueden implementarse mediante instrucciones de programa legibles por ordenador.

Estas instrucciones de programa legibles por ordenador pueden proporcionarse a un procesador de un ordenador de propósito general, un ordenador de uso especial u otro aparato de procesamiento de datos programable para producir una máquina, de manera que las instrucciones, que se ejecutan a través del procesador del ordenador u otro aparato programable de procesamiento de datos, creen medios para implementar las funciones/acciones especificadas en el bloque o bloques del diagrama de flujo y/o diagrama de bloques. Estas instrucciones del programa legibles por ordenador también pueden almacenarse en un medio de almacenamiento legible por ordenador que puede dirigir un ordenador, un aparato de procesamiento de datos programable y/u otros dispositivos para que funcionen de una manera particular, de manera que el medio de almacenamiento legible por ordenador que tiene instrucciones almacenadas en este incluya un artículo de fabricación que incluye instrucciones que implementan aspectos de la función/acciones especificadas en el bloque o bloques en el diagrama de flujo y/o diagrama de bloques.

Las instrucciones de programa legibles por ordenador también pueden cargarse en un ordenador, otro aparato de procesamiento de datos programable u otro dispositivo para hacer que se realicen una serie de pasos operativos en el ordenador, otro aparato programable u otro dispositivo para producir un procedimiento implementado por ordenador, de tal manera que las instrucciones que se ejecutan en el ordenador, en otro aparato programable o en otro dispositivo implementan las funciones/acciones especificadas en el diagrama de flujo y/o diagrama de bloques, bloque o bloques.

El diagrama de flujo y los diagramas de bloques de las Figuras ilustran la arquitectura, funcionalidad y operación de posibles implementaciones de sistemas, procedimientos y productos de programas informáticos según varias realizaciones de la presente invención. A este respecto, cada bloque en el diagrama de flujo o los diagramas de bloques puede representar un módulo, un segmento o una parte de instrucciones, que comprende una o más instrucciones ejecutables para implementar la función o funciones lógicas especificadas. En algunas implementaciones alternativas, las funciones indicadas en el bloque pueden ocurrir fuera del orden indicado en las figuras. Por ejemplo, dos bloques mostrados en sucesión pueden, de hecho, ejecutarse de manera sustancialmente simultánea, o los bloques pueden ejecutarse a veces en orden inverso, dependiendo de la funcionalidad implicada. También se observará que cada bloque de los diagramas de bloques y/o la ilustración del diagrama de flujo, y combinaciones de bloques en los diagramas de bloques y/o la ilustración del diagrama de flujo, pueden implementarse mediante sistemas basados en hardware de propósito especial que realizan las funciones o actos especificados o llevan a cabo combinaciones de hardware de propósito especial e instrucciones de ordenador.

Las descripciones de las diversas realizaciones de la presente invención se han presentado con fines ilustrativos, pero no pretenden ser exhaustivas ni limitarse a las realizaciones descritas. Muchas modificaciones y alteraciones serán evidentes para los expertos en la materia sin apartarse del alcance de las realizaciones descritas. La terminología usada en la presente memoria se eligió para explicar mejor los principios de las realizaciones, la aplicación práctica o la mejora técnica sobre las tecnologías que se encuentran en el mercado, o para permitir que otros expertos en la materia entiendan las realizaciones descritas en la presente memoria.

Números de referencia

10 sistema FW

12 núcleo

14 IF núcleo-nido asíncrona

16 búfer de agregación

18 nido de sistema

20 controlador de bus de E/S

22 bus de E/S

24 búfer de estado de E/S

26 lógica de finalización temprana

28 controlador de bus Buffer-IO IF

30 instrucción de almacenamiento de E/S

	32 código de controlador ES asíncrono
	34 código de configuración de E/S
	36 reenvío asíncrono
	38 bus asíncrono
5	40 IF de usuario
	42 lógica de administración de matriz
	44 matriz de estado de E/S
	46 lógica de acceso y administración de matriz
	48 mensaje con identificación de entrada en matriz
10	50 sistema HW/FW
	210 sistema de procesamiento de datos
	212 sistema informático/servidor
	214 dispositivos externos
	216 CPU / unidad de procesamiento de datos
15	218 Bus ES
	220 adaptador de red
	222 interfaces ES
	224 pantalla
	228 memoria
20	230 RAM
	232 caché
	234 sistema de almacenamiento
	240 programa/utilidades
	242 módulos de programa
25	

REIVINDICACIONES

1. Un sistema de procesamiento de datos (210) para manejar una instrucción de almacenamiento de entrada/salida (30), que comprende un nido de sistema (18) acoplado comunicativamente a al menos un bus de entrada/salida (22) mediante un controlador de bus de entrada/salida (20),
- 5 comprendiendo además al menos
- una unidad de procesamiento de datos (216) que comprende un núcleo (12),
- un firmware de sistema (10) y una interfaz núcleo-nido asíncrona (14),
- en donde la unidad de procesamiento de datos (216) está acoplado comunicativamente al nido de sistema (18) a través de un búfer de agregación (16),
- 10 en donde la interfaz núcleo-nido asíncrona (14) comprende una matriz de estado de entrada/salida (44) con múltiples búferes de estado de entrada/salida (24), una lógica de acceso y gestión de matriz (46),
- y en donde
- (i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos (210) que emite la instrucción de almacenamiento de entrada/salida (30) que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos que se van a transferir y/o un puntero a los datos que se van a transferir, y una longitud de los datos;
- 15 (ii) la unidad de procesamiento de datos (216) está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida (30);
- (iii) la unidad de procesamiento de datos (216) está configurado para verificar si el acceso a la función de entrada/salida está permitido en un espacio de direcciones y a nivel de instancia invitada, donde el huésped funciona en el sistema de procesamiento de datos (210);
- 20 (iv) la unidad de procesamiento de datos (216) está configurado para completar la instrucción de almacenamiento de entrada/salida (30) antes de que se complete la ejecución de la instrucción de almacenamiento de entrada/salida (30) en el nido del sistema (18), _el nido del sistema (18) está configurado para cargar de forma asíncrona y/o almacenar datos en al menos un dispositivo externo (214) que está acoplado comunicativamente al bus de entrada/salida (22), el dispositivo externo (214) es externo al nido del sistema (18), al búfer de agregación (16) y a la unidad de procesamiento de datos (216);
- 25 (v) el firmware del sistema (10) está configurado para emitir un mensaje del sistema para enviar un paquete de datos al búfer de agregación (16) hasta que todos los datos de un bloque de almacenamiento se hayan reenviado al búfer de agregación (16);
- 30 (vi) el firmware del sistema (10) está configurado para emitir un mensaje de sistema al búfer de agregación (16) para enviar los datos de manera asíncrona como un único mensaje de anidación al controlador de bus de entrada/salida (20), mientras se espera que el búfer de agregación (16) envíe un mensaje de finalización;
- (vii) la unidad de procesamiento de datos (216) está configurado para recibir un mensaje de libre reutilización desde el búfer de agregación (16) después de enviar un mensaje al nido del sistema (18) para permitir una continuación temprana del procedimiento de transmisión de los datos a través del búfer de agregación (16) al nido del sistema (18) y al controlador del bus de entrada/salida (20);
- 35 (viii) el nido del sistema (18) está configurado para reenviar el mensaje a la ubicación diana;
- (ix) el controlador del bus de entrada/salida (20) está configurado para recibir el mensaje y reenviar los datos en una trama de datos al bus de entrada/salida y enviar un mensaje de finalización al nido del sistema (18);
- 40 (x) el nido del sistema está configurado para reenviar el mensaje de finalización al búfer de agregación (16);
- (xi) el búfer de agregación (16) está configurado para reenviar el mensaje de finalización a la interfaz asíncrona núcleo-nido (14);
- 45 (xii) la interfaz núcleo-nido asíncrona (14) está configurada para almacenar el estado en los búferes de entrada/salida (24) para el índice de búfer de estado de entrada/salida respectivo y señala la finalización de la operación al firmware del sistema (10);
- (xiii) el firmware del sistema (10) está configurado para notificar al sistema operativo, de forma asíncrona, mediante una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida (30), incluida una transferencia de datos, la unidad de procesamiento de datos (216) detecta un error al transmitir los datos de la ejecución asíncrona fallida;
- 50

(xiv) y la unidad de procesamiento de datos (216) retrasa una ejecución de la instrucción de almacenamiento (30), hasta que un búfer de estado de entrada/salida (24) esté disponible para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes (30).

5 2. El sistema de procesamiento de datos según la reivindicación 1, la unidad de procesamiento de datos (216) retrasa una ejecución de instrucciones de almacenamiento asíncronas y/o sincrónicas (30), hasta que un búfer de estado de entrada/salida (24) esté disponible para almacenar información sobre la finalización de instrucciones de almacenamiento pendientes (30).

10 3. El sistema de procesamiento de datos según la reivindicación 1 o 2, además, los búferes de estado de entrada/salida (24) recogen estados de mensajes del nido del sistema (18) y/o del controlador de bus de entrada/salida (20), en particular un estado de finalización del nido del sistema (18).

4. El sistema de procesamiento de datos de la reivindicación 3, en donde los estados del mensaje y/o el estado de finalización se numeran mediante un índice de búfer de estado de entrada/salida.

15 5. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, el búfer de agregación (16) está acoplado comunicativamente a la interfaz núcleo-nido asíncrona (14) a través de un bus asíncrono (38).

6. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, en donde los datos son transferidos por la instrucción de almacenamiento de entrada/salida (30) a través de un mecanismo de transmisión asíncrono con un mensaje de finalización temprana en múltiples paquetes de datos al búfer de agregación (16), si la longitud de los datos excede ocho bytes, de lo contrario los datos se transfieren en un paquete de datos.

20 7. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, el firmware del sistema (10) que comprende un código de controlador de entrada/salida asíncrono (32) para manejar la instrucción de almacenamiento de entrada/salida (30).

25 8. El sistema de procesamiento de datos según la reivindicación 7, el núcleo (12) que comprende un código de configuración asíncrono (34) para gestionar los requisitos de memoria para la información de estado del código de controlador de entrada/salida asíncrono (32).

9. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, la interfaz núcleo-nido asíncrona (14) que comprende un componente de reenvío de interfaz núcleo-nido asíncrona (36) para reenviar los datos con finalización local.

30 10. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, el búfer de agregación (16) que comprende una lógica de finalización temprana (26) para entregar un mensaje libre para reutilización después de enviar una solicitud.

35 11. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, en donde el firmware del sistema (10) comprende una lógica de gestión de matriz (42), que asigna/desasigna búferes de estado de entrada/salida (24) en la matriz de estado de entrada/salida (44) y/o inicia el inicio de una nueva instrucción de almacenamiento (30).

12. El sistema de procesamiento de datos según una cualquiera de las reivindicaciones anteriores, un mensaje del sistema que comprende uno de

- una dirección diana física jerárquica,
- obtener un hilo SMT o un identificador de búfer agregado,
- 40 - una longitud de datos,
- una dirección de bus de entrada/salida,
- un índice de búfer de estado de entrada/salida.

45 13. Un procedimiento para gestionar una instrucción de almacenamiento de entrada/salida (30) para al menos un dispositivo externo (214) de un sistema de procesamiento de datos (210), comprendiendo el sistema de procesamiento de datos (210)

un nido de sistema (10) acoplado comunicativamente a al menos un bus de entrada/salida (22) mediante un controlador de bus de entrada/salida (14),

y que comprende además al menos una unidad de procesamiento de datos (216) que comprende un núcleo (12), un firmware de sistema (10) y una interfaz núcleo-nido asíncrona (14),

50 en donde la unidad de procesamiento de datos (216) está acoplado comunicativamente al nido de sistema (18) a

través de un búfer de agregación (16),

en donde el dispositivo externo (214) está acoplado comunicativamente al bus de entrada/salida (22),

en donde la interfaz núcleo-nido asíncrona (14) comprende una matriz de estado de entrada/salida (44) con múltiples búferes de estado de entrada/salida (24), una lógica de acceso y gestión de matriz (46),

5 comprendiendo el procedimiento:

(i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos (210) que emite la instrucción de almacenamiento de entrada/salida (30) que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos que se van a transferir y/o un puntero a los datos que se van a transferir, y una longitud de los datos;

10 (ii) la unidad de procesamiento de datos (216) que identifica la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida (30);

(iii) la unidad de procesamiento de datos (216) que verifica si se permite el acceso a la función de entrada/salida en un espacio de direcciones y en un nivel de instancia de invitado, ejecutándose el invitado en el sistema de procesamiento de datos (210);

15 (iv) la unidad de procesamiento de datos (216) que completa la instrucción de almacenamiento de entrada/salida (30) antes de que se complete una ejecución de la instrucción de almacenamiento de entrada/salida (30) en el nido de sistema (18), el nido de sistema (18) carga de manera asíncrona desde y/o almacena datos a al menos un dispositivo externo (214) que está acoplado comunicativamente al bus de entrada/salida (22), siendo el dispositivo externo (214) externo al nido de sistema (18), al búfer de agregación (16) y a la unidad de procesamiento de datos (216);

20 (v) el firmware del sistema (10) que emite un mensaje del sistema para enviar un paquete de datos al búfer de agregación (16) hasta que todos los datos de un bloque de almacenamiento se hayan enviado al búfer de agregación (16);

25 (vi) el firmware del sistema (10) emite un mensaje de sistema al búfer de agregación (16) para enviar los datos de manera asíncrona como un único mensaje de anidación al controlador de bus de entrada/salida (20), mientras espera que el búfer de agregación (16) envíe un mensaje de finalización;

(vii) la unidad de procesamiento de datos (216) recibe un mensaje de libre para reutilización desde el búfer de agregación (16) después de que se envía un mensaje al nido de sistema (18) para permitir una continuación temprana del procedimiento de transmisión de los datos a través del búfer de agregación (16) al nido de sistema (18) y al controlador de bus de entrada/salida (20);

30 (viii) el nido del sistema (18) reenvía el mensaje a la ubicación de destino;

(ix) el controlador de bus de entrada/salida (20) recibe el mensaje y envía los datos en una trama de datos al bus de entrada/salida y envía un mensaje de finalización al nido del sistema (18);

(x) el nido del sistema reenvía el mensaje de finalización al búfer de agregación (16);

(xi) el búfer de agregación (16) reenvía la finalización a la interfaz asíncrona núcleo-nido (14);

35 (xii) la interfaz núcleo-nido asíncrona (14) que almacena el estado en el búfer de estado de entrada/salida (24) para el índice de búfer de estado de entrada/salida respectivo y señala la finalización de la operación al firmware del sistema (10);

40 (xiii) el firmware del sistema (10) notifica al sistema operativo, de forma asíncrona, a través de una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida (30), incluida una transferencia de datos, la unidad de procesamiento de datos (216) detecta un error al transmitir los datos de la ejecución asíncrona fallida;

(xiv) la unidad de procesamiento de datos (216) retrasa una ejecución de la instrucción de almacenamiento (30), hasta que un búfer de estado de entrada/salida (24) esté disponible para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes (30).

45 14. El procedimiento según la reivindicación 13, la unidad de procesamiento de datos (216) retrasa una ejecución de instrucciones de almacenamiento asíncronas y/o sincrónicas (30), hasta que un búfer de estado de entrada/salida (24) esté disponible para almacenar información sobre la finalización de instrucciones de almacenamiento pendientes (30).

50 15. El procedimiento según la reivindicación 13 o 14, además, los búferes de estado de entrada/salida (24) recogen estados de mensajes del nido de sistema (18) y/o del controlador de bus de entrada/salida (20), en particular un estado de finalización del nido de sistema (18), en donde los estados del mensaje y/o el estado de finalización se numeran

mediante un índice de búfer de estado de entrada/salida.

16. El procedimiento según la reivindicación 15, en donde el firmware del sistema (10) comprende una lógica de gestión de matriz (42), que asigna/desasigna búferes de estado de entrada/salida (24) en la matriz de estado de entrada/salida (44) y/o inicia un inicio de una nueva instrucción de almacenamiento (30).

- 5 17. El procedimiento de una cualquiera según las reivindicaciones 13 a 16, que comprende además
- (i) el sistema operativo que emite la instrucción de almacenamiento de entrada/salida (30);
 - (ii) el firmware del sistema (10) asigna un índice de búfer de estado de entrada/salida libre; si no hay ningún índice de búfer de estado de entrada/salida libre disponible, entonces espera a que aparezca un índice de búfer de estado de entrada/salida libre;
 - 10 (iii) el firmware del sistema (10) que inyecta la instrucción de almacenamiento (30) en el motor de envío asíncrono; si está bloqueado por otra instrucción de almacenamiento, esperando hasta que se haya completado la instrucción de almacenamiento;
 - (iv) según la longitud de los datos: si la longitud de los datos supera los ocho bytes, el firmware del sistema (10) emite repetidamente un mensaje del sistema para enviar un paquete de datos al búfer de agregación (16) hasta que todos los datos de un bloque de almacenamiento se hayan enviado al búfer de agregación (16), mientras que el firmware del sistema (10) espera hasta que los datos hayan sido enviados por el mensaje del sistema; de lo contrario el firmware del sistema (10) emite un mensaje del sistema para enviar los datos al búfer de agregación (16);
 - 15 (v) el búfer de agregación (16) que inyecta el mensaje de nido en el nido del sistema (18), en donde el búfer de agregación (16) queda libre para su reutilización inmediatamente después de la operación de envío, enviando una señal al firmware del sistema (10); luego el búfer de agregación (16) envía un mensaje de libre para su reutilización;
 - 20 (vi) el firmware del sistema (10) que actualiza un búfer de estado de entrada/salida que rastrea mediante el índice de búfer de estado de entrada/salida;
 - (vii) el firmware del sistema (10) señala de forma asíncrona los defectos al sistema operativo en caso de error.
- 25 18. El procedimiento según una cualquiera de las reivindicaciones 13 a 17, transfiriendo además los datos mediante la instrucción de almacenamiento de entrada/salida (30) a través de un mecanismo de transmisión asíncrono con un mensaje de finalización temprana en múltiples paquetes de datos al búfer de agregación (16), si la longitud de los datos excede ocho bytes.
- 30 19. El procedimiento según una cualquiera de las reivindicaciones 13 a 18, además, el firmware del sistema (10) utiliza un código de controlador de entrada/salida asíncrono (32) para manejar la instrucción de almacenamiento de entrada/salida (30).
20. El procedimiento según la reivindicación 19, además, el núcleo (12) utiliza un código de configuración asíncrono (34) para gestionar los requisitos de memoria para la información de estado del código de controlador de entrada/salida asíncrono (32).
- 35 21. El procedimiento de una cualquiera según las reivindicaciones 13 a 20, además, la interfaz núcleo-nido asíncrona (14) utiliza un componente de reenvío de interfaz núcleo-nido asíncrona (36) para reenviar los datos con finalización local.
22. El procedimiento según una cualquiera de las reivindicaciones 13 a 21, además, el búfer de agregación (16) utiliza una lógica de finalización temprana (26) para entregar un mensaje libre para reutilización después de enviar una solicitud.
23. El procedimiento según una cualquiera de las reivindicaciones 13 a 22, en donde un mensaje del sistema comprende uno de
- 40 - una dirección diana física jerárquica,
 - obtener un hilo SMT o un identificador de búfer agregado,
 - una longitud de datos,
 - una dirección de bus de entrada/salida,
 - un índice de búfer de estado de entrada/salida.
- 45 24. Un producto de programa informático para gestionar una instrucción de almacenamiento de entrada/salida (30) a al menos un dispositivo externo (214) de un sistema de procesamiento de datos (210), comprendiendo el sistema de procesamiento de datos (210)

un nido de sistema (10) acoplado comunicativamente a al menos un bus de entrada/salida (22) mediante un controlador de bus de entrada/salida (14),

y que comprende además al menos una unidad de procesamiento de datos (216) que comprende un núcleo (12), un firmware de sistema (10) y una interfaz núcleo-nido asíncrona (14),

5 en donde la unidad de procesamiento de datos (216) está acoplado comunicativamente al nido de sistema (18) a través de un búfer de agregación (16),

en donde el dispositivo externo (214) está acoplado comunicativamente al bus de entrada/salida (22),

en donde la interfaz núcleo-nido asíncrona (14) comprende una matriz de estado de entrada/salida (44) con múltiples búferes de estado de entrada/salida (24), una lógica de acceso y gestión de matriz (46),

10 el producto de programa informático que comprende un medio de almacenamiento legible por ordenador que tiene instrucciones de programa que,

cuando se ejecuta por un sistema informático (212), hacen que el sistema informático (212) lleve a cabo un procedimiento que comprende:

15 (i) un sistema operativo que se ejecuta en el sistema de procesamiento de datos (210) que emite la instrucción de almacenamiento de entrada/salida (30) que especifica al menos una función de entrada/salida con un desplazamiento a través de una dirección, datos que se van a transferir y/o un puntero a los datos que se van a transferir, y una longitud de los datos;

(ii) la unidad de procesamiento de datos (216) está configurado para identificar la función de entrada/salida mediante la dirección especificada en la instrucción de almacenamiento de entrada/salida (30);

20 (iii) la unidad de procesamiento de datos (216) está configurado para verificar si el acceso a la función de entrada/salida está permitido en un espacio de direcciones y a nivel de instancia invitada, donde el huésped funciona en el sistema de procesamiento de datos (210);

25 (iv) la unidad de procesamiento de datos (216) está configurado para completar la instrucción de almacenamiento de entrada/salida (30) antes de que se complete una ejecución de la instrucción de almacenamiento de entrada/salida (30) en el nido de sistema (18), estando configurado el nido de sistema (18) para cargar de manera asíncrona desde y/o almacenar datos en al menos un dispositivo externo (214) que está acoplado comunicativamente al bus de entrada/salida (22), siendo el dispositivo externo (214) externo al nido de sistema (18), al búfer de agregación (16) y a la unidad de procesamiento de datos (216);

30 (v) el firmware del sistema (10) está configurado para emitir un mensaje del sistema para enviar un paquete de datos al búfer de agregación (16) hasta que todos los datos de un bloque de almacenamiento se hayan reenviado al búfer de agregación (16);

(vi) el firmware del sistema (10) está configurado para emitir un mensaje de sistema al búfer de agregación (16) para enviar los datos de manera asíncrona como un único mensaje de anidación al controlador de bus de entrada/salida (20), mientras se espera que el búfer de agregación (16) envíe un mensaje de finalización;

35 (vii) la unidad de procesamiento de datos (216) está configurado para recibir un mensaje de libre reutilización desde el búfer de agregación (16) después de enviar un mensaje al nido del sistema (18) para permitir una continuación temprana del procedimiento de transmisión de los datos a través del búfer de agregación (16) al nido del sistema (18) y al controlador del bus de entrada/salida (20);

(viii) el nido del sistema (18) está configurado para reenviar el mensaje a la ubicación diana;

40 (ix) el controlador de bus de entrada/salida (20) está configurado para recibir el mensaje y reenviar los datos en una trama de datos al bus de entrada/salida y enviar un mensaje de finalización al nido del sistema (18);

(x) el nido del sistema está configurado para reenviar el mensaje de finalización al búfer de agregación (16);

(xi) el búfer de agregación (16) está configurado para reenviar la finalización a la interfaz núcleo-nido asíncrona (14);

45 (xii) la interfaz núcleo-nido asíncrona (14) que está configurada para almacenar el estado en el búfer de estado de entrada/salida (24) para el índice de búfer de estado de entrada/salida respectivo y la finalización de la señal de operación al firmware del sistema (10);

50 (xiii) el firmware del sistema (10) está configurado para notificar al sistema operativo, de forma asíncrona, mediante una interrupción, si durante la ejecución asíncrona de la instrucción de almacenamiento de entrada/salida (30), incluida una transferencia de datos, la unidad de procesamiento de datos (216) detecta un error al transmitir los datos de la ejecución asíncrona fallida;

(xiv) la unidad de procesamiento de datos (216) retrasa una ejecución de la instrucción de almacenamiento (30), hasta que un búfer de estado de entrada/salida (24) esté disponible para almacenar información sobre la finalización de las instrucciones de almacenamiento pendientes (30).

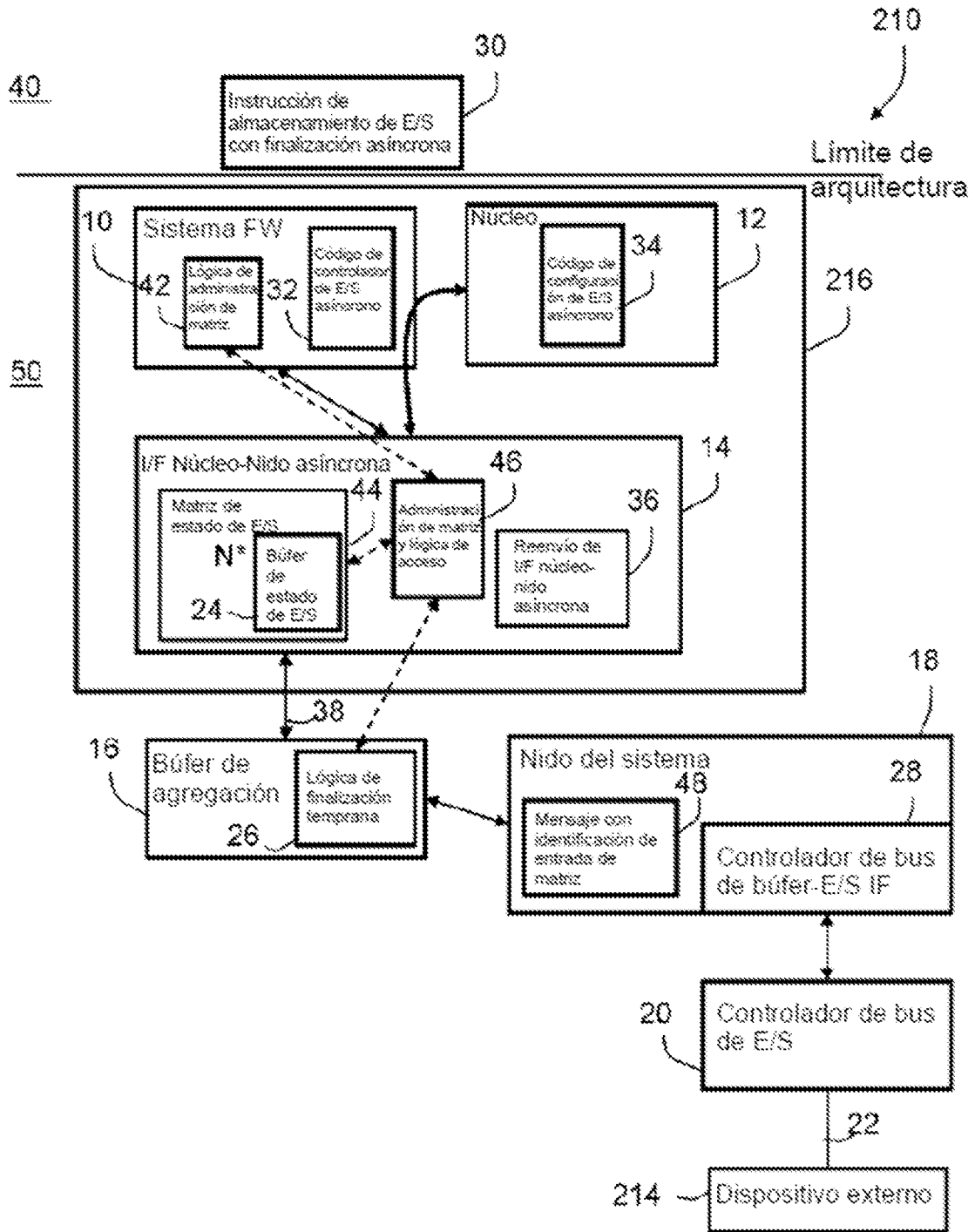


Fig. 1

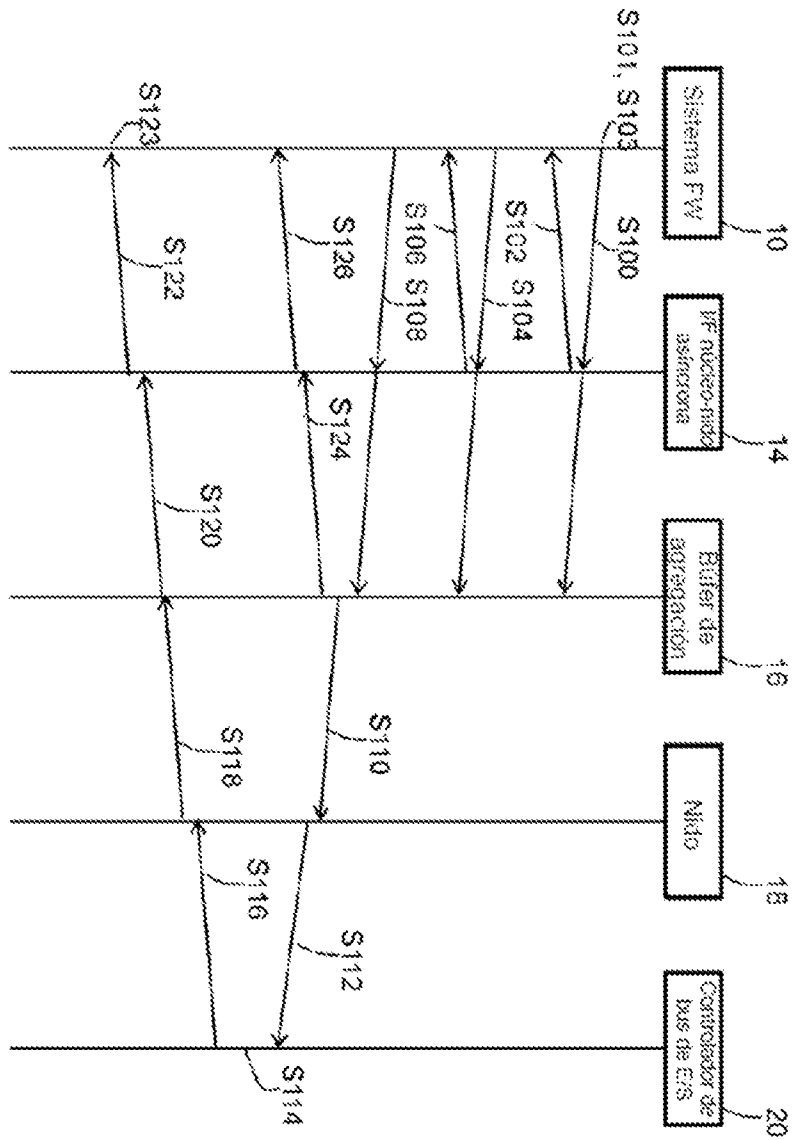


Fig. 2

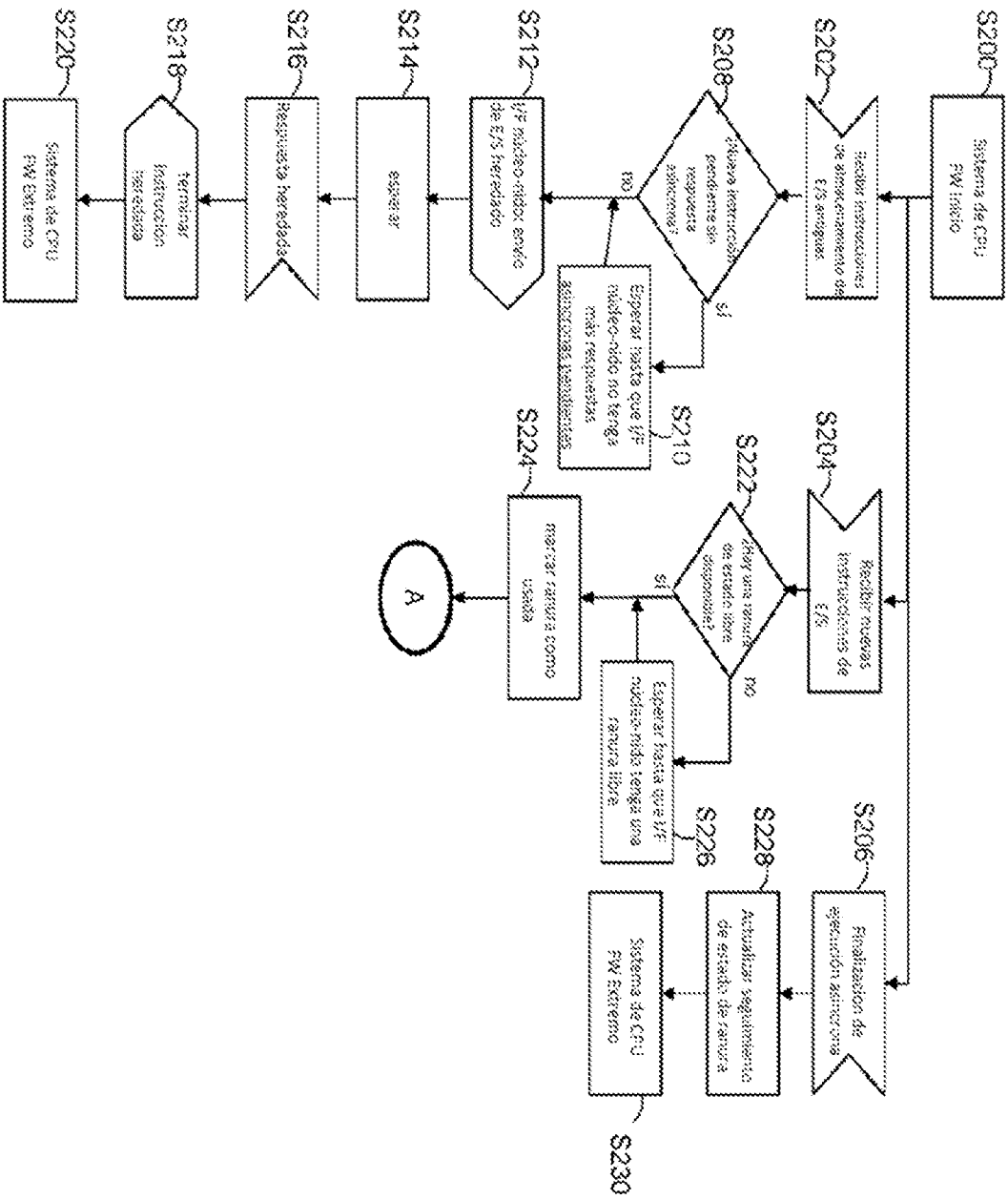


Fig. 3

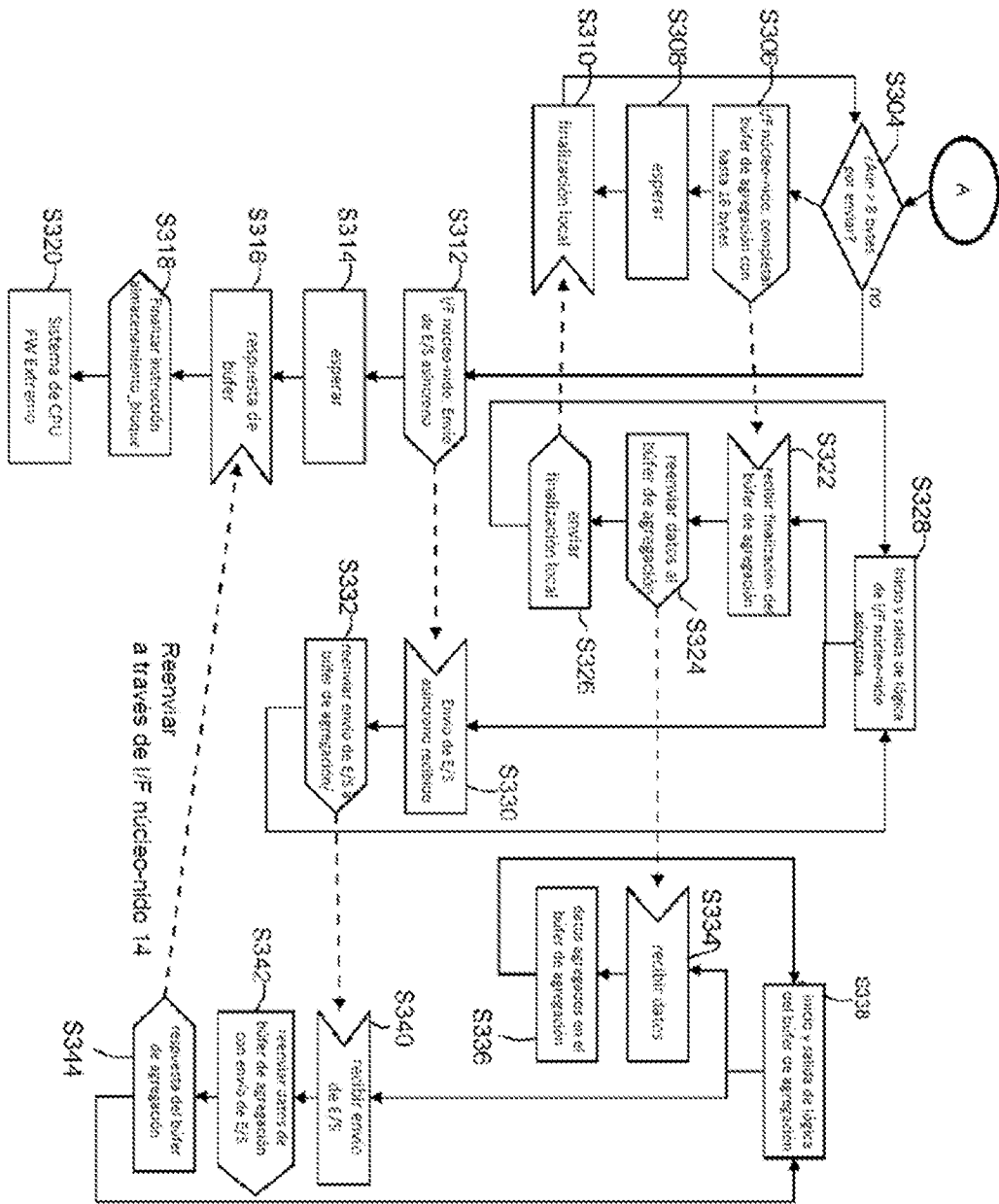


Fig. 4

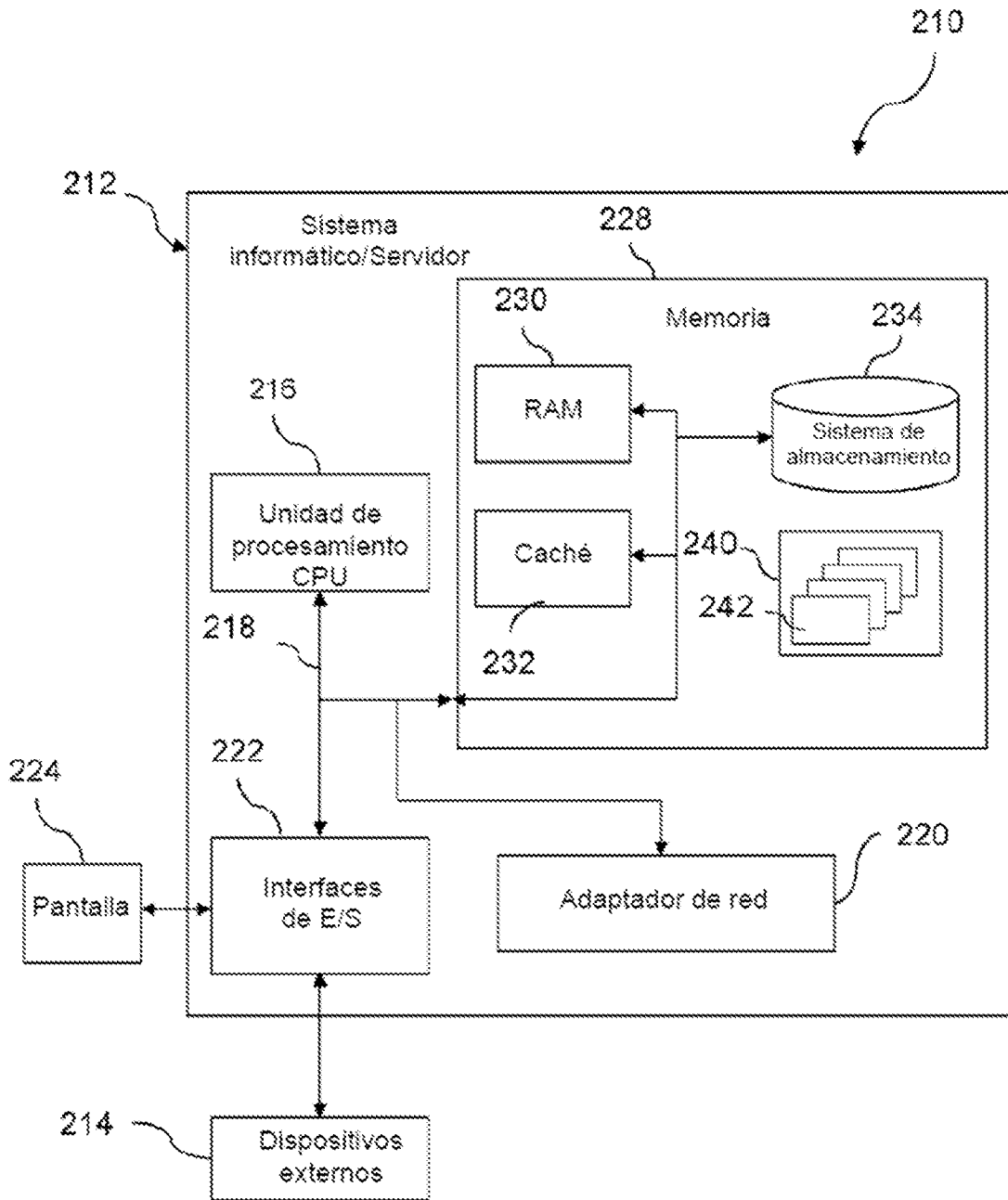


Fig. 5