(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 737 855**

(51) Int. CI.:

G06F 9/455 (2008.01)
G06F 9/50 (2006.01)
G06F 9/46 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **30.09.2009 PCT/US2009/059105**

(87) Fecha y número de publicación internacional: **08.04.2010 WO10039887**

(96) Fecha de presentación y número de la solicitud europea: **30.09.2009 E 09818470 (8)**

(97) Fecha y número de publicación de la concesión europea: **15.05.2019 EP 2335157**

(54) Título: **Virtualización del espacio de configuración**

(30) Prioridad:

03.10.2008 US 245543

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
16.01.2020

(73) Titular/es:

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

(72) Inventor/es:

OSHINS, JACOB;
ALLSOP, BRANDON y
THORNTON, ANDREW JOHN

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

DESCRIPCIÓN

Virtualización del espacio de configuración

Campo de tecnología

El contenido en la presente memoria divulgado versa sobre el campo de la informática y, más en particular, acerca de una virtualización de ordenador, aunque la virtualización es simplemente un campo ejemplar y no limitante.

Antecedentes

La mayoría de dispositivos de entrada/salida (I/O) se diseñan suponiendo que existe un elemento de soporte lógico fiable que configura todos los dispositivos de I/O en el sistema. Normalmente, también se supone que esos dispositivos de I/O están controlados en última instancia por controladores de dispositivos que son módulos enchufables que abstraen las diferencias de los dispositivos individuales. Además, se supone que estos controladores están todos contenidos en un único núcleo.

Sin embargo, en el contexto de las máquinas virtuales, las anteriores suposiciones pueden ya no ser válidas. Cada máquina virtual contiene, normalmente, su propio núcleo del sistema operativo, en el que pueden confiar o no todos los otros núcleos que se ejecutan en todas las otras máquinas virtuales. La configuración y el control de los dispositivos en un host físico normalmente implican cierta autoridad central que tiene la capacidad para imponer directrices relativas a cómo pueden afectar las acciones de una máquina virtual a otras máquinas virtuales. En algunos sistemas, esta autoridad central se encuentra en un sistema operativo host. En otros sistemas la autoridad puede encontrarse en un hipervisor, y en otros adicionales, la autoridad puede encontrarse en una de las máquinas virtuales que se ejecutan sobre un hipervisor.

Cuando se construye un sistema de virtualización, un enfoque puede ser mantener un control completo de todos los dispositivos de I/O en la autoridad central descrita anteriormente. Por lo tanto, cuando una máquina virtual necesita servicios de I/O, la máquina virtual puede pasar una solicitud (directa o indirectamente) a la autoridad central que controla la I/O. Este enfoque puede funcionar pero adolece de dos problemas. En primer lugar, la I/O opera más lentamente de lo que lo haría en comparación con un sistema operativo que se ejecuta en un soporte físico real en vez de en una máquina virtual. En segundo lugar, la gama de dispositivos de I/O expresada a las máquinas virtuales puede estar limitada por el soporte lógico de virtualización. Sería deseable asignar cada uno de los dispositivos en un ordenador físico a una o más de las máquinas virtuales que se ejecutan en el mismo. De esta forma, la I/O no adolecería de la penalización del rendimiento asociada con la indirección y cualquier dispositivo que pueda ser enchufado en el ordenador puede ser utilizado por una máquina virtual sin requerir que las capas de virtualización entiendan completamente su función interna.

Las unidades de gestión de la memoria de I/O (IOMMU) se divulgan en “The price of safety: Evaluating IOMMU performance” de M. Ben-Yehuda et al, Proceedings of the Linux Symposium, 2007, páginas 9 a 20. Las IOMMU son dispositivos de soporte físico que traducen las direcciones DMA del dispositivo a las direcciones de la máquina y que pueden ser utilizadas para restringir un dispositivo, de forma que solo pueda acceder a partes de la memoria a las que se haya concedido acceso de forma explícita. Se utilizan tablas de entidad de control de la traducción para traducir una dirección de I/O a una dirección de la memoria host.

En consecuencia, se necesitan otras técnicas en la especialidad para solucionar los problemas descritos anteriormente.

Sumario

En la presente memoria se divulgan diversos procedimientos y sistemas para acotar el comportamiento de una máquina virtual no privilegiada (una máquina virtual que no es propietaria una directriz en todo el sistema para el dispositivo) que interactúa con un dispositivo creando o recibiendo una descripción del dispositivo que indica a una autoridad privilegiada (por ejemplo, un hipervisor u otro aspecto privilegiado de un sistema de virtualización) (1) qué operaciones en el dispositivo pueden tener efectos en todo el sistema y (2) qué operaciones tienen efectos locales al dispositivo. La autoridad privilegiada puede permitir o denegar, entonces, estas acciones. La autoridad privilegiada también puede traducir estas acciones en otras acciones con consecuencias benignas.

En una realización, para cada dispositivo, se puede construir una asignación del espacio de configuración, en la que cada bit en la configuración puede tener una o más de las siguientes propiedades:

i. Solo lectura.
ii. Siempre 0 en lectura.
iii. Siempre 1 en lectura.
iv. Lectura-escritura.
v. La escritura de 1 borra, la escritura de 0 se deja tal cual.
vi. La escritura de 1 establece, la escritura de 0 se deja tal cual.
vii. La escritura de 0 borra, la escritura de 1 se deja tal cual.

viii. La escritura de 0 establece, la escritura de 1 se deja tal cual.
ix. Borrar a 0 después de la primera lectura.
x. Establecer a 1 después de la primera lectura.

Los anteriores comportamientos son ejemplares, y se pueden incluir comportamientos adicionales para acotar las acciones permitidas en ubicaciones de la memoria. También se puede construir una asignación del espacio MMIO, en la que se puede asignar cada página en la máquina virtual. Para páginas que se excluyen de la asignación de la máquina virtual, la autoridad privilegiada puede escoger cargar en esa página una imagen estática que se asemeja al dispositivo. De forma alternativa, la autoridad privilegiada puede escoger recibir interceptaciones y gestionar las interceptaciones utilizando el espacio de configuración con una asignación aplicada para estas páginas específicas.

En realizaciones adicionales, se divulga un procedimiento para intercambiar información para contener un dispositivo de forma segura. En algunas realizaciones se puede embeber una representación de las asignaciones en un fichero de instalación de controlador. Los ficheros de instalación pueden estar firmados digitalmente por la entidad que los produce. En consecuencia, un administrador de una máquina puede permitir que la autoridad privilegiada procese los ficheros de instalación sin instalar realmente un controlador para el dispositivo en la autoridad privilegiada. El controlador puede instalarse en la máquina virtual y el dispositivo puede ser funcional en la máquina virtual y estar contenido de manera que el dispositivo no afecte a otras máquinas virtuales o a la autoridad privilegiada.

Además de lo anterior, se describen otros aspectos en las reivindicaciones, en los dibujos y en el texto que forman una parte de la presente divulgación. Un experto en la técnica puede apreciar que uno o más aspectos diversos de la divulgación pueden incluir, sin limitación, circuitería y/o programación para efectuar los aspectos a los que se hace referencia en la presente memoria de la presente divulgación; la circuitería y/o la programación pueden ser virtualmente cualquier combinación de soporte físico, de soporte lógico y/o de soporte lógico inalterable configurados para efectuar los aspectos a los que se hace referencia en la presente memoria dependiendo de las decisiones de diseño del diseñador del sistema.

Se debería hacer notar que se proporciona este sumario para presentar una selección de conceptos de forma simplificada que se describen adicionalmente a continuación en la descripción detallada. No se concibe que este sumario identifique características claves o características esenciales del contenido reivindicado, ni se prevé que sea utilizado como una ayuda para determinar el alcance del contenido reivindicado.

Breve descripción de los dibujos

El anterior sumario, al igual que la siguiente descripción detallada, se comprende mejor cuando es leído junto con los dibujos adjuntos. Para ilustrar la presente divulgación, se ilustran diversos aspectos de la divulgación. Sin embargo, la divulgación no está limitada a los aspectos específicos mostrados. Se incluyen las siguientes figuras:

La **Figura 1a** ilustra un entorno de máquina virtual, con una pluralidad de máquinas virtuales, que comprende una pluralidad de procesadores virtuales y correspondientes sistemas operativos invitados; las máquinas virtuales son mantenidas por una capa de virtualización que puede comprender un planificador y otros componentes, virtualizando la capa de virtualización un soporte físico para la pluralidad de máquinas virtuales;
la **Figura 1b** ilustra un diagrama que representa la estratificación lógica de la arquitectura del soporte físico y del soporte lógico para un entorno virtualizado en un sistema de ordenador;
la **Figura 1c** muestra un sistema ejemplar de ordenador en el que se pueden implementar aspectos de la presente divulgación;
la **Figura 2** ilustra un sistema informático virtualizado ejemplar;
la **Figura 3** ilustra un sistema informático virtualizado alternativo;
la **Figura 4** muestra un diagrama del sistema ejemplar que ilustra un espacio IO y MMIO relacionado con los dispositivos de PCI;
la **Figura 5** es un diagrama ejemplar que ilustra un espacio de direcciones en el que puede residir la RAM;
la **Figura 6** ilustra un ejemplo de un procedimiento operativo para gestionar los efectos globales y locales de transacciones entre una máquina virtual no privilegiada y un dispositivo de I/O;
la **Figura 7** ilustra un ejemplo de un procedimiento operativo para gestionar los efectos global y local de transacciones entre una máquina virtual no privilegiada y un dispositivo de I/O;
la **Figura 8** ilustra un ejemplo de un procedimiento operativo para gestionar los efectos global y local de una máquina virtual no privilegiada que interactúa con un dispositivo;
la **Figura 9** muestra un procedimiento operativo ejemplar para acotar el comportamiento de una máquina virtual no privilegiada que interactúa con un dispositivo;
la **Figura 10** ilustra un medio legible por un ordenador que tiene instrucciones ejecutables por un ordenador expuesto con respecto a las anteriores Figuras 1-9.

Descripción detallada

Máquinas virtuales en términos generales

En la siguiente descripción y en las figuras se definen ciertos detalles específicos para proporcionar una comprensión exhaustiva de diversas realizaciones de la invención. Ciertos detalles bien conocidos a menudo asociados con la tecnología informática y de soporte lógico no se definen en la siguiente divulgación para evitar ofuscar innecesariamente las diversas realizaciones de la invención. Además, las personas con un nivel normal de dominio de la técnica relevante comprenderán que pueden poner en práctica otras realizaciones de la invención sin uno o más de los detalles descritos a continuación. Finalmente, aunque se describen diversos procedimientos con referencia a etapas y a secuencias en la siguiente divulgación, la descripción como tal es para proporcionar una implementación clara de realizaciones de la invención, y no se debería que las etapas y las secuencias de etapas sean imprescindibles para poner en práctica la presente invención.

Se debería comprender que las diversas técnicas descritas en la presente memoria pueden ser implementadas en conexión con soporte físico o soporte lógico o, cuando sea apropiado, con una combinación de ambos. Por lo tanto, los procedimientos y el aparato de la invención, o ciertos aspectos o porciones de los mismos, pueden adoptar la forma de código de programa (es decir, instrucciones) implementado en medios tangibles, tales como disquetes, CD-ROM, unidades de disco duro o cualquier otro medio de almacenamiento legible por una máquina en los que, cuando se carga el código de programa en una máquina —tal como un ordenador—, y es ejecutado por la misma, la máquina se convierte en un aparato para poner en práctica la invención. En el caso de la ejecución del código de programa en ordenadores programables, el dispositivo informático incluye, en general, un procesador, un medio de almacenamiento legible por el procesador (incluyendo memoria volátil y no volátil y/o elementos de almacenamiento), al menos un dispositivo de entrada y al menos un dispositivo de salida. Uno o más programas que pueden implementar o utilizar los procedimientos descritos en conexión con la invención, por ejemplo, mediante el uso de una API, controles reutilizables o similares. Preferentemente, tales programas se implementan en un lenguaje de programación de procedimientos u orientada a objetos de alto nivel para comunicarse con un sistema de ordenador. Sin embargo, el o los programas pueden implementarse en un lenguaje ensamblador o de máquina, si se desea. En cualquier caso, el lenguaje puede ser un lenguaje compilado o interpretado, y combinado con implementaciones de soporte físico.

La Figura 1a ilustra un entorno 100 de máquina virtual, con una pluralidad de máquinas virtuales 120, 121, que comprende una pluralidad de procesadores virtuales 110, 112, 114, 116, y sistemas operativos invitados correspondientes 130, 132. Las máquinas virtuales 120, 121 son mantenidas mediante una capa 140 de virtualización que puede comprender un planificador 142 y otros componentes (no mostrados), virtualizando la capa 140 de virtualización soporte físico 150 para la pluralidad de máquinas virtuales 120, 121. La pluralidad de procesadores virtuales 110, 112, 114, 116 pueden ser los homólogos virtuales de procesadores físicos 160, 162 de soporte físico subyacente.

La Figura 1b es un diagrama que representa la estratificación lógica de la arquitectura de soporte físico y de soporte lógico para un entorno virtualizado en un sistema de ordenador. En la Fig. 1b, un programa 180 de virtualización se ejecuta directa o indirectamente en la arquitectura 182 de soporte físico real. El programa 180 de virtualización puede ser (a) un monitor de máquina virtual que se ejecuta junto con un sistema operativo host, (b) un sistema operativo host con un componente hipervisor, llevando a cabo el componente hipervisor la virtualización, (c) soporte físico o (d) microcódigo. El programa de virtualización también puede ser un hipervisor que se ejecuta por separado de cualquier sistema operativo. En otras palabras, el programa de virtualización del hipervisor no necesita ejecutarse como parte de ningún sistema operativo, y no necesita ejecutarse junto con ningún sistema operativo. El programa de virtualización del hipervisor puede ejecutarse, en cambio, “debajo” de todos los sistemas operativos, incluyendo la “partición raíz”. El programa 180 de virtualización virtualiza una arquitectura 178 de soporte físico invitado (mostrada como líneas discontinuas para ilustrar el hecho de que este componente es una “partición” o una “máquina virtual”), es decir, soporte físico que no existe realmente sino que es virtualizado, en cambio, por el programa 180 de virtualización. Un sistema operativo invitado 176 se ejecuta en la arquitectura 178 de soporte físico invitado, y una aplicación 174 de soporte lógico puede ejecutarse en el sistema operativo invitado 176. En el entorno operativo virtualizado de la Fig. 1b, la aplicación 174 de soporte lógico puede ejecutarse en un sistema de ordenador incluso si la aplicación 174 de soporte lógico está diseñada para ejecutarse en un sistema operativo que incompatible, en general, con un sistema operativo host y con la arquitectura 182 del soporte físico.

Normalmente, una máquina virtual contiene un sistema operativo completo y un conjunto de aplicaciones, que constituyen conjuntamente muchos procedimientos, pudiendo denominarse a la totalidad de los mismos “carga de trabajo” o “proceso” en el contexto de máquinas virtuales. En la presente divulgación, se pueden utilizar las expresiones “proceso” y “carga de trabajo” de forma intercambiable en el contexto de máquinas virtuales, y los expertos en la técnica comprenderán fácilmente que “proceso” puede hacer referencia a múltiples procesos incluyendo todos los sistemas y las aplicaciones que pueden ser instanciadas en una máquina virtual.

A continuación, la Fig. 2 ilustra un sistema informático virtualizado que comprende una capa (204 de soporte lógico de sistema operativo host (SO host) que se ejecuta directamente sobre el soporte físico 202 de ordenador físico, proporcionando el SO host 204 acceso a los recursos del soporte físico 202 de ordenador físico exponiendo las interfaces a particiones A 208 y B 210 para su uso por los sistemas operativos A y B, 212 y 214, respectivamente. Esto permite que el SO host 204 pase desapercibido por las capas 212 y 214 del sistema operativo que se ejecutan sobre él. De nuevo, para llevar a cabo la virtualización, el SO host 204 puede ser un sistema operativo diseñado

especialmente con capacidades nativas de virtualización o, de forma alternativa, puede ser un sistema operativo estándar con un componente hipervisor incorporado para llevar a cabo la virtualización (no mostrada).

Con referencia de nuevo a la Fig. 2, encima del SO host 204 hay dos particiones, la partición A 208, que puede ser, por ejemplo, un procesador Intel 386 virtualizado, y la partición B 210, que puede ser, por ejemplo, una versión virtualizada de uno de la familia de procesadores Motorola 680X0. En cada partición 208 y 210 existen sistemas operativos invitados (SO invitado) A 212 y B 214, respectivamente. Ejecutándose sobre el SO invitado A 212 hay dos aplicaciones, la aplicación A1 216 y la aplicación A2 218, y ejecutándose sobre el SO invitado B 214 hay la aplicación B1 220.

Con respecto a la Fig. 2, es importante hacer notar que la partición A 208 y la partición B 214 (que se muestran con líneas discontinuas) son representaciones de soporte físico de ordenador virtualizado que pueden existir únicamente como construcciones de soporte lógico. Se hace que sean posibles debido a la ejecución de uno o más soportes lógicos especializados de virtualización que no solo presentan la partición A 208 y la partición B 210 al SO invitado A 212 y al SO invitado B 214, respectivamente, pero que también llevan a cabo todas las etapas del soporte lógico necesarias para el SO invitado A 212 y el SO invitado B 214 para interactuar indirectamente con el soporte físico real 202 de ordenador físico.

La Figura 3 ilustra un sistema informático virtualizado alternativo en el que se lleva a cabo la virtualización mediante un monitor 204' de máquina virtual (VMM) que se ejecuta junto con el sistema operativo host 204". En ciertos casos, el VMM 204' puede ser una aplicación que se ejecute sobre el sistema operativo host 204" e interactúa con el soporte físico 202 de ordenador únicamente mediante el sistema operativo host 204". En otros casos, según se muestra en la Fig. 3, el VMM 204' puede comprender, en cambio, un sistema de soporte lógico parcialmente independiente que en algunos niveles interactúa indirectamente con el soporte físico 202 de ordenador mediante el sistema operativo host 204", pero en otros niveles el VMM 204' interactúa directamente con el soporte físico 202 de ordenador (similar a la forma en la que el sistema operativo host interactúa directamente con el soporte físico de ordenador). Y en otros casos adicionales, el VMM 204' puede comprender un sistema de soporte lógico completamente independiente que en todos los niveles interactúa directamente con el soporte físico 202 de ordenador (similar a la forma en la que el sistema operativo host interactúa directamente con el soporte físico de ordenador) sin utilizar el sistema operativo host 204" (aunque aún interactúa con el sistema operativo host 204" para coordinar el uso del soporte físico 202 de ordenador y evitar conflictos y similares).

La Figura 4 muestra un diagrama del sistema ejemplar que ilustra un espacio de IO y MMIO relacionado dispositivos de PCI. El diagrama incluye un bus 400 del sistema, una memoria física 410, un procesador 420, un dispositivo 430 de PCI con un registro 460 y un dispositivo 440 de puente host-PCI. Hay un bus 450 de PCI fijado al dispositivo 440 de puente host-PCI, y el dispositivo 430 de PCI está fijado al bus de PCI. El dispositivo 430 de PCI contiene al menos un registro 460 en una ubicación de la memoria que debe ser leída y escrita desde los procesadores del sistema para controlar el dispositivo. Puede verse que los espacios de dirección de memoria física pueden ser distintos del espacio del puerto de IO, que puede ser un espacio separado de dirección. Los recursos de IO pueden ser traducidos en recursos de MMIO, que es una razón por la que el espacio de dirección del puerto de IO puede ser asignado mediante accesos de puerto de IO asignados con la memoria.

Con referencia a la Figura 5, se muestra un diagrama que ilustra el espacio 500 de dirección en el que puede residir la RAM. Según se muestra, las áreas 510 de entrada/salida (MMIO) asignadas con la memoria también pueden residir en el mismo espacio de dirección. Las interfaces modernas típicas de control del soporte físico basadas en un espacio de dirección residen en la porción de MMIO de este espacio de dirección. En general, el espacio de dirección física del sistema hace referencia al espacio 500 de dirección física del sistema de ordenador físico, igual que el "espacio de dirección física invitada", que también hace referencia al espacio 500 de dirección "física" de un sistema de ordenador virtual. Normalmente, el espacio 500 de dirección de la memoria física está separado del espacio del puerto de IO. Se puede utilizar el espacio separado del puerto de IO para controlar dispositivos más antiguos, y también puede ser utilizado para disponer y configurar dispositivos más nuevos, dado que se accede al espacio de configuración de la PCI, normalmente, por medio de un espacio de puerto de IO. Además, las direcciones del espacio de puerto de IO tienen normalmente 16 bits en vez de 32 bits o 64 bits.

Todas estas variaciones para implementar las particiones mencionadas anteriormente son únicamente implementaciones ejemplares, y no se debería interpretar que nada de la presente memoria sea limitante de la divulgación de cualquier aspecto particular de virtualización.

Virtualización del espacio de configuración

La mayoría de dispositivos de entrada/salida (I/O) están diseñados con la suposición de que existe un elemento de soporte lógico fiable que configura todos los dispositivos de I/O en el sistema. También se supone, normalmente, que esos dispositivos de I/O están controlados en última instancia por medio de controladores del dispositivo que son módulos enchufables que abstraen diferencias individuales del dispositivo. Además, se supone que estos controladores están todos contenidos en un único núcleo.

Sin embargo, en el contexto de las máquinas virtuales, las anteriores suposiciones pueden ya no ser válidas. Cada máquina virtual contiene, normalmente, su propio núcleo del sistema operativo, en el que pueden confiar o no todos los otros núcleos que se ejecutan en todas las otras máquinas virtuales. La configuración y el control de los dispositivos en una máquina normalmente implica cierta autoridad central que tiene la capacidad de imponer directrices relativas a cómo las acciones de una máquina virtual pueden afectar a otras máquinas virtuales. En algunos sistemas, esta autoridad central se encuentra en un sistema operativo host. En otros sistemas, la autoridad puede encontrarse en un hipervisor, y en otros adicionales, la autoridad puede encontrarse en una de las máquinas virtuales que se ejecutan sobre un hipervisor.

Cuando se construye un sistema de virtualización, un planteamiento puede ser mantener un control completo de todos los dispositivos de I/O en la autoridad central descrita anteriormente. Por lo tanto, cuando una máquina virtual necesita servicios de I/O, la máquina virtual puede pasar una solicitud (directa o indirectamente) a la autoridad central que controla la I/O. Este enfoque puede ser aceptable pero adolece de dos problemas. En primer lugar, la I/O opera más lentamente de lo que lo haría en comparación con un sistema operativo que se ejecuta en soporte físico real en vez de en una máquina virtual. En segundo lugar, la gama de dispositivos de I/O expresada a las máquinas virtuales puede estar limitada por el soporte lógico de virtualización. Sería deseable asignar cada uno de los dispositivos en un ordenador físico a una o más de las máquinas virtuales que se ejecutan en las mismas. De esta forma, la I/O no adolecería de la penalización del rendimiento asociada con la indirección. Además, cualquier dispositivo que pueda ser enchufado en el ordenador puede ser utilizado por una máquina virtual sin requerir que las capas de virtualización comprendan completamente sus funciones internas.

Por ejemplo, si se enchufa un controlador de interfaz de red (NIC) en una máquina física, puede ser razonable suponer que el soporte lógico de virtualización puede controlar y manipular el NIC. Los NIC son comunes y los proveedores de NIC pueden desear proporcionar un soporte lógico de control del dispositivo tanto para sistemas operativos populares como para sistemas de virtualización. Por otra parte, si se enchufa un dispositivo más esotérico en un ordenador, no es probable que haya disponible un soporte lógico correspondiente de virtualización. Por lo tanto, sería deseable permitir que una máquina virtual tenga un acceso directo al dispositivo incluso sin ninguna comprensión de cómo utilizará la máquina virtual el dispositivo.

Por desgracia, la configuración y la disposición del dispositivo esotérico pueden tener consecuencias en todo el sistema que pueden tener un impacto sobre la función de otras máquinas virtuales. Por ejemplo, la activación del dispositivo puede provocar un pico de corriente de irrupción que puede provocar que todo el ordenador tenga un apagado parcial si la irrupción se produce simultáneamente con otro pico de irrupción. En otro ejemplo, la configuración del dispositivo puede implicar dan instrucciones al dispositivo para que reivindique intervalos del espacio de dirección de la memoria que puede ser ocupado por otros dispositivos o la memoria principal. Como ejemplo final, un dispositivo puede estar empaquetado en un chip con muchos otros dispositivos. En términos de la especificación de la interconexión de componentes periféricos (PCI), tales dispositivos son denominados “funciones” y el chip es denominado “paquete”. El paquete puede tener una conexión con el bus (o en el caso de la PCI Express, una conexión con el tejido) y cada función en el paquete puede compartir parte del soporte físico asociado con la conexión al resto del sistema. En este caso, la configuración de la función de menor número (nº 0) puede tener, a menudo, efectos secundarios visibles en la operación de funciones de mayor número. Si la función 0 está bajo el control de una máquina virtual, las opciones realizadas en esa máquina virtual pueden tener un impacto en otras funciones en el mismo paquete. Esto puede provocar que otras máquinas virtuales reciban ningún servicio o un servicio deficiente de las funciones que se encuentran bajo el control de otras máquinas.

En diversas realizaciones divulgadas en la presente memoria, el comportamiento de una máquina virtual no privilegiada que interactúa con un dispositivo puede ser acotado creando una descripción del dispositivo que indica a una autoridad privilegiada (por ejemplo, un hipervisor u otro aspecto privilegiado de un sistema de virtualización) (1) qué operaciones en el dispositivo pueden tener efectos en todo el sistema y (2) qué operaciones tienen efectos locales al dispositivo. Una máquina virtual no privilegiada hace referencia a una máquina virtual que no es propietaria de una directriz en todo el sistema para el sistema o el dispositivo. En otras palabras, una máquina virtual no privilegiada no es el hipervisor o un precursor/raíz/SO host. Entonces, la autoridad privilegiada puede permitir o denegar estas acciones. La autoridad privilegiada también puede traducir estas acciones a otras acciones con consecuencias benignas.

Aunque puede ser posible implementar algunos de los anteriores procedimientos cargando el controlador de un dispositivo en el contexto de la autoridad privilegiada, tal planteamiento es normalmente no deseable debido al código adicional requerido en la autoridad privilegiada. Minimizar la cantidad de código en la autoridad privilegiada es a menudo importante para hacer un sistema de virtualización tanto seguro como eficaz.

Además, cuando se permite que toda una función de PCI (en vez de, por ejemplo, únicamente un subconjunto de un dispositivo) se encuentre bajo el control de una máquina virtual no privilegiada, puede no existir ningún código en la autoridad privilegiada para subasignar los recursos del dispositivo a muchas máquinas virtuales. Este procedimiento de subasignación es común cuando se comparte un dispositivo entre muchas máquinas virtuales. En cambio, la presente divulgación describe procedimientos para colocar un dispositivo diferenciado completo bajo el control de una máquina virtual.

En una realización, cada dispositivo de PCI (o PCI-X, o PCI Express) puede implementar dos o tres espacios de dirección. El primer espacio de dirección puede describirse como un espacio de I/O asignado con la memoria y puede comportarse de forma similar al direccionamiento de RAM. Las lecturas y escrituras en un dispositivo pueden llevarse a cabo como lecturas y escrituras en RAM pero utilizando distintas direcciones. Con referencia a la Figura 5, el espacio 520 de dirección de la RAM puede ocupar, por ejemplo, los primeros 2 GB de espacio de dirección de la memoria con dispositivos de I/O que ocupan el espacio 510 de dirección entre 3 GB (dirección 3221225472) y 4 GB (dirección 4294967296). El espacio de dirección de MMIO puede ser utilizado para una interacción momento a momento con el dispositivo por el controlador del dispositivo. El acceso al espacio de dirección de MMIO es normalmente rápido y es normalmente llevado a cabo por el controlador del dispositivo (que es suministrado, normalmente, por el proveedor del dispositivo) para el dispositivo. Cuando se desactiva el dispositivo, el dispositivo normalmente no decodifica ningún espacio de MMIO.

El segundo espacio de dirección que puede implementarse es el espacio de configuración implementado por dispositivos de PCI. Este espacio de configuración puede estar poblado con mecanismos (por ejemplo, registros) para configurar el dispositivo. Tales mecanismos pueden incluir desactivar y activar el dispositivo, asignar recursos y similares. Normalmente, el espacio de configuración se decodifica si el dispositivo está activado o desactivado. La especificación de PCI identifica los comportamientos de algunos de los registros en este espacio. Los registros permiten un elemento genérico de soporte lógico de configuración (no suministrado por el proveedor del dispositivo) configure el dispositivo, asigne recursos al dispositivo (tal como un intervalo asignado de direcciones de espacio de MMIO), y active el dispositivo. El espacio de configuración puede contener, y normalmente lo hace, registros específicos del dispositivo sin el significado definido por la especificación de PCI. Normalmente, tales registros solo pueden ser manipulados por el controlador de dispositivo para el dispositivo. Por último, se pueden añadir nuevas características a la especificación de PCI definiendo nuevos intervalos en el espacio de configuración denominado “estructuras de capacidad”.

El tercer espacio de dirección que puede utilizar un dispositivo de PCI es denominado espacio de “I/O” y es fundamentalmente histórico. En general, el espacio de I/O tiene las propiedades del espacio de MMIO.

Una autoridad privilegiada tal como un hipervisor u otro intermediario de virtualización puede necesitar decidir qué partes del espacio de configuración pueden ser puestas bajo el control de una máquina virtual no privilegiada. En las realizaciones divulgadas a continuación, se describirá un hipervisor como la autoridad privilegiada. Sin embargo, debería ser inmediatamente evidente para los expertos en la técnica que se pueden implementar realizaciones divulgadas en conexión con cualquier otro intermediario de virtualización.

El hipervisor puede intentar, además, contener las partes del espacio de MMIO y de I/O a las que puede acceder la máquina virtual no privilegiada. En diversas realizaciones la presente divulgación describe procedimientos para contener el comportamiento de la máquina virtual no privilegiada. En una realización, para cada dispositivo se puede construir una asignación del espacio de configuración, en la que cada bit en la asignación tiene una o más de las siguientes propiedades:

i. Solo lectura.
ii. Siempre 0 en lectura.
iii. Siempre 1 en lectura.
iv. Lectura-escritura.
v. La escritura de 1 borra, la escritura de 0 se deja tal cual.
vi. La escritura de 1 establece, la escritura de 0 se deja tal cual.
vii. La escritura de 0 borra, la escritura de 1 se deja tal cual.
viii. La escritura de 0 establece, la escritura de 1 se deja tal cual.
ix. Borrar a 0 después de la primera lectura.
x. Establecer a 1 después de la primera lectura.

Los anteriores comportamientos son ejemplares, y se pueden incluir comportamientos adicionales para acotar las acciones permitidas en ubicaciones de la memoria. También se pueden asignar comportamientos con ubicaciones de la memoria a mayores niveles de granularidad, tales como bytes o segmentos más grandes de memoria tales como páginas.

Se puede construir una asignación del espacio de MMIO, en la que cada página puede estar bien asignada en la máquina virtual o bien no asignada en la máquina virtual. La asignación puede estar construida con una granularidad de página más que con una granularidad de bit. Si se utiliza una granularidad de bit, puede haber potencialmente numerosos bits de espacio de MMIO, siendo el resultado de que la asignación puede volverse ilógicamente grande. Además, los procesadores normalmente proporcionan al hipervisor la capacidad de interceptar únicamente en la granularidad de página; así, construir una asignación de nivel de bit implicaría que el hipervisor tendría que interceptar cada operación de MMIO y aplicar el filtro apropiado implicado por la asignación. Tal interferencia constante con la operación del dispositivo probablemente tendría un impacto negativo sobre la operación del dispositivo.

Algunos dispositivos pueden asignar registros de sus espacios de configuración una segunda vez en sus espacios de I/O o de MMIO. Esto puede hacerse debido a que el acceso al espacio de configuración es normalmente lento y puede

ser conveniente proporcionar un acceso a un registro antes de que se configure el dispositivo, en cuyo caso la asignación debería estar en el espacio de configuración. También se debería proporcionar el acceso al registro posteriormente en el momento de la ejecución mediante una vía jerarquizada al registro, en cuyo caso la asignación también debería estar en el espacio de memoria. En consecuencia, uno de los comportamientos para una página del espacio de MMIO es que la página puede estar configurada como un pseudónimo del espacio de configuración, en el que cualquier acceso a la página debería ser atrapado y redirigido al código que gestiona el espacio de configuración. Además de asignar una página completa de esta forma, se pueden marcar bits individuales en una página como pseudónimos de bits específicos en el espacio de configuración.

Para una página que se excluye de la asignación de la máquina virtual, el hipervisor puede escoger cargar en la página excluida una imagen estática que se asemeja al dispositivo. De forma alternativa, el hipervisor puede decidir aceptar interceptaciones y gestionar las interceptaciones como el espacio de configuración con una asignación aplicada para estas páginas específicas. En otras palabras, una asignación del espacio de MMIO puede tener dos niveles. Un nivel puede ser para la lista de páginas del espacio de MMIO del dispositivo que están asignadas en la máquina virtual. La asignación de la segunda capa puede definir, opcionalmente, los bits en las páginas excluidas.

El espacio de I/O para el dispositivo puede ser tratado como el espacio de configuración. De forma alternativa, el espacio de I/O puede ser excluido de la máquina virtual.

Según se ha expuesto anteriormente, las diversas realizaciones divulgadas pueden permitir que una autoridad, tal como un hipervisor, contenga de forma segura un dispositivo para el cual no emplea un controlador del dispositivo. Por lo tanto, es posible que el hipervisor no tenga la información para poblar tal asignación. Por lo tanto, es necesaria una forma de obtener esta información del proveedor del dispositivo. En una realización, se puede crear una representación de las asignaciones que pueden estar embebidas en un fichero de instalación del controlador. En una realización, el fichero de instalación del controlador puede denominarse INF. Los INF pueden estar contenidos en paquetes de instalación del controlador. Además, los INF pueden estar firmados digitalmente por la entidad que produce los paquetes. En consecuencia, un administrador de máquina puede decidir permitir que el hipervisor procese el INF suministrado por el fabricante del dispositivo sin instalar realmente un controlador para el dispositivo. Entonces, se puede instalar el controlador en la máquina virtual y el dispositivo puede volverse funcional en la máquina virtual y contenido de forma que el controlador no afecte a otras máquinas virtuales o al propio hipervisor.

Los aspectos en la presente memoria divulgados pueden implementarse como sistemas, procedimientos, instrucciones ejecutables por un ordenador que residen en medios legibles por un ordenador, etcétera. Por lo tanto, cualquier divulgación de cualquier sistema, procedimiento o medio legible por un ordenador particular no está limitada a los mismos, sino que más bien se extiende a otras formas de implementar el contenido divulgado.

Las Figuras 6 a 8 muestran un ejemplo de un procedimiento operativo para gestionar comunicaciones entre una máquina virtual y un dispositivo de I/O. El procedimiento puede incluir operaciones 600, 602, 604, 605 y 606. Con referencia a la Figura 6, la operación 600 pone en marcha el procedimiento operativo y en la operación 602 se puede construir una representación del espacio de configuración para el dispositivo de I/O indicando acciones que pueden llevarse a cabo sobre el dispositivo de I/O mediante la máquina virtual. Este espacio de configuración puede ser poblado con mecanismos (por ejemplo, registros) para configurar el dispositivo. Se puede construir 603 una representación del espacio de I/O asignado con la memoria en la que cada página del espacio de I/O asignado con la memoria está asignada en la máquina virtual o excluida de la máquina virtual. La operación 604 ilustra el control del acceso a dicho dispositivo de I/O según dicha representación del espacio de configuración y dicha representación del espacio de I/O asignado con la memoria. La operación 605 ilustra que dicha construcción de una representación del espacio de configuración comprende, además, asociar cada bit en dicha representación del espacio de configuración con al menos una operación de lectura y de escritura. La operación 606 ilustra que para cualquier memoria excluida de dicha representación del espacio de configuración o para cualquier memoria excluida de dicha representación del espacio de I/O asignado con la memoria, se puebla dicha cualquier memoria con datos representativos de dicho dispositivo de I/O. Para una página que está excluida de la asignación de la máquina virtual, por ejemplo, el hipervisor puede decidir cargar en la página excluida una imagen estática que se asemeja al dispositivo.

Con referencia a la FIG. 7, la operación 706 ilustra que las operaciones de lectura y de escritura comprenden: solo lectura 708, siempre 0 en lectura 710, siempre 1 en lectura 712, lectura-escritura 714, la escritura de 1 borra / la escritura de 0 se deja tal cual 716, la escritura de 1 establece / la escritura de 0 se deja tal cual 718, la escritura de 0 borra / la escritura de 1 se deja tal cual 720, la escritura de 0 establece / la escritura de 1 se deja tal cual 722, borrar a 0 tras la primera lectura 724 o establecer a uno tras la primera lectura 726.

Con referencia a la Figura 8, la operación 802 ilustra la definición de bits en páginas para la memoria excluida. La operación 804 ilustra la recepción de interceptaciones y el procesamiento de las interceptaciones utilizando páginas con bits definidos. Por ejemplo, el hipervisor puede decidir aceptar interceptaciones y gestionar las interceptaciones como el espacio de configuración con una asignación aplicado para estas páginas específicas. En una realización 806, la información puede ser recibida para construir dichas asignaciones, en la que dicha información es recibida en un fichero proporcionado por un proveedor de dicho dispositivo de I/O. En otra realización, el fichero está firmado

digitalmente por dicho proveedor 808. La operación 810 ilustra la construcción de las representaciones según la información recibida del proveedor.

La gestión puede llevarse a cabo por una capa 813 de virtualización utilizando las páginas con bits definidos. Un controlador puede estar instalado en la máquina virtual y el dispositivo puede volverse funcional en la máquina virtual y contenido de forma que el controlador no afecte a otras máquinas virtuales o al propio hipervisor.

En la operación 814, se construye una representación del espacio de I/O. La operación 815 ilustra la introducción de datos en la representación del espacio de I/O en función de la información recibida. La operación 825 ilustra el control del acceso a dicho dispositivo de I/O según la representación del espacio de I/O. La operación 830 ilustra la introducción de datos tanto en dicha asignación del espacio de configuración como en dicha asignación del espacio de I/O asignado con la memoria en función de la información recibida. El espacio de I/O de la máquina virtual puede ser excluido en la operación 835.

La Figura 9 muestra un procedimiento operativo ejemplar para gestionar las comunicaciones entre una máquina virtual y un dispositivo incluyendo las operaciones 900, 902, 904, 906, 908, 910, 912 y 914. Con referencia a la Figura 9, la operación 900 inicia el procedimiento operativo y la operación 902 ilustra la recepción de una descripción del dispositivo, comprendiendo la descripción información relativa a qué operaciones en el dispositivo tienen efectos en todo el sistema y cuáles tienen efectos que son locales al dispositivo. La operación 904 ilustra la creación de una representación de la descripción. La operación 906 ilustra la incorporación de la representación en un fichero de instalación para un controlador para el dispositivo, en la que la representación permite la construcción de una asignación del espacio de configuración para el dispositivo y una asignación del espacio de I/O asignado con la memoria, en la que se puede utilizar la asignación del espacio de configuración y la asignación del espacio de I/O asignado con la memoria para acceder al dispositivo.

La operación 908 ilustra que cada bit en dicha asignación o página asociado con la asignación del espacio de configuración y con la asignación del espacio de I/O asignado con la memoria comprende al menos una de las siguientes propiedades: solo lectura 910, siempre 0 en lectura 912, siempre 1 en lectura 914, lectura-escritura 916, la escritura de 1 borra / la escritura de 0 se deja tal cual 918, la escritura de 1 establece / la escritura de 0 se deja tal cual 920, la escritura de 0 borra / la escritura de 1 se deja tal cual 922, la escritura de 0 establece / la escritura de 1 se deja tal cual 924, borrar a 0 tras la primera lectura 926 o establecer a 1 tras la primera lectura 928. La operación 930 ilustra que el fichero de instalación es un INF y la operación 932 ilustra la firma digital del INF.

Se puede implementar cualquiera de los aspectos mencionados anteriormente en procedimientos, sistemas, medios legibles por un ordenador o cualquier tipo de fabricación. Por ejemplo, según la Fig. 10, un medio legible por un ordenador puede almacenar en el mismo instrucciones ejecutables por un ordenador para controlar el acceso a un dispositivo de PCI, PCI-X o PCI-Express, en el que el dispositivo está acoplado de forma comunicativa con una máquina física que alberga máquinas virtuales. Tales medios pueden comprender un primer subconjunto de instrucciones para recibir un fichero de instalación para el dispositivo, en el que el fichero de instalación comprende información relativa a qué operaciones en el dispositivo tienen efectos en todo el sistema y cuáles tienen efectos que son locales al dispositivo 1010; un segundo subconjunto de instrucciones para construir al menos una asignación de atributos para el espacio de configuración, el espacio de I/O asignado con la memoria y el espacio de I/O para el dispositivo, en la que cada página o cada bit asociado con la al menos una asignación está asignado en la máquina virtual y en la que se puede presentar una página estática de bits en una máquina virtual como el estado del dispositivo 1012; un tercer subconjunto de instrucciones para poblar la al menos una asignación en función de dicho fichero recibido 1014 de instalación, y un cuarto conjunto de instrucciones para utilizar la al menos una asignación para gestionar el acceso al dispositivo 1016. Los expertos en la técnica apreciarán que se pueden utilizar conjuntos adicionales de instrucciones para capturar los otros diversos aspectos divulgados en la presente memoria, y que los tres subconjuntos en la presente memoria divulgados pueden variar en detalle según la presente divulgación.

Por ejemplo, las instrucciones pueden comprender, además, instrucciones 1020 en las que cada bit en dicha asignación o página asociada con la al menos una asignación contiene una de las siguientes propiedades: siempre 0 en lectura, siempre 1 en lectura, lectura-escritura, la escritura de 1 borra / la escritura de 0 se deja tal cual, la escritura de 1 establece / la escritura de 0 se deja tal cual, la escritura de 0 borra / la escritura de 1 se deja tal cual, la escritura de 0 establece / la escritura de 1 se deja tal cual, borrar a 0 tras la primera lectura o establecer a 1 tras la primera lectura.

De nuevo, a modo de ejemplo, las instrucciones pueden comprender, además, instrucciones para: poblar dicha cualquier memoria con datos predeterminados para cualquier memoria excluida de la asignación del espacio de configuración o para cualquier memoria excluida de la asignación del espacio de I/O asignado con la memoria 1021; los datos predeterminados se corresponden con un dispositivo predeterminado 1022; definir bits en páginas para la memoria excluida 1023; recibir interceptaciones y procesar las interceptaciones utilizando las páginas con bits definidos 1024; y el fichero de instalación es un INF proporcionado por un proveedor del dispositivo y, opcionalmente, puede estar firmado digitalmente 1026.

Según se ha divulgado anteriormente, los aspectos de la invención pueden ejecutarse en un ordenador programado. Se concibe que la FIG. 1c y la siguiente exposición proporcionen una breve descripción de un entorno informático

adecuado en el que se pueden implementar esos aspectos. Un experto en la técnica puede apreciar que el sistema de ordenador de la FIG. 1c puede efectuar, en algunas realizaciones, diversos aspectos de las Figuras 1a y 1b. En estas realizaciones ejemplares, el servidor y el cliente pueden incluir algunos de los componentes, o todos ellos, descritos en la FIG. 1c y, en algunas realizaciones, cada uno del servidor y del cliente puede incluir circuitería configurada para ejemplificar aspectos específicos de la presente divulgación.

El término circuitería utilizado en toda la divulgación puede incluir componentes especializados de soporte físico. En las mismas realizaciones, o en otras, la circuitería puede incluir microprocesadores configurados para llevar a cabo la o las funciones mediante soporte lógico inalterable o conmutaciones. En las mismas realizaciones ejemplares, o en otras, la circuitería puede incluir una o más unidades de procesamiento de uso general y/o unidades de procesamiento de múltiples núcleos, etc., que pueden estar configuradas cuando se cargan en la memoria, por ejemplo RAM y/o memoria virtual, instrucciones de soporte lógico que implementan lógica operable para llevar a cabo la o las funciones. En realizaciones ejemplares en las que la circuitería incluye una combinación de soporte físico y de soporte lógico, un implementador puede escribir código fuente que implementa lógica y el código fuente puede ser compilado en código legible por una máquina que puede ser procesado por la o las unidades de procesamiento de uso general.

La FIG. 1c muestra un ejemplo de un sistema informático que está configurado con aspectos de la divulgación. El sistema informático puede incluir un ordenador 20 o similar, incluyendo una unidad 21 de procesamiento, una memoria 22 del sistema y un bus 23 del sistema que acopla diversos componentes del sistema, incluyendo la memoria del sistema, a la unidad 21 de procesamiento. El bus 23 del sistema puede ser cualquiera de varios tipos de estructuras de bus incluyendo un bus de memoria o controlador de memoria, un bus periférico y un bus local utilizando cualquiera de una variedad de arquitecturas de bus. La memoria del sistema incluye memoria 24 de solo lectura (ROM) y memoria 25 de acceso aleatorio (RAM). Un sistema básico 26 de entrada/salida (BIOS), que contiene las rutinas básicas que ayudan a transferir información entre elementos en el ordenador 20, tal como durante la puesta en marcha, se almacena en la ROM 24. El ordenador 20 puede incluir, además, una unidad 27 de disco duro para leer un disco duro, no mostrado, y escribir en el mismo, una unidad 28 de disco magnético para leer un disco magnético extraíble 29, o escribir en el mismo y una unidad 30 de disco óptico para leer un disco óptico extraíble 31, o escribir en el mismo, tal como un CD ROM u otros medios ópticos. En algunas realizaciones ejemplares, se pueden almacenar instrucciones ejecutables por un ordenador, que implementan aspectos de la invención, en una ROM 24, un disco duro (no mostrado), una RAM 25, un disco magnético extraíble 29, un disco óptico 31 y/o una memoria intermedia de la unidad 21 de procesamiento. La unidad 27 de disco duro, la unidad 28 de disco magnético y la unidad 30 de disco óptico están conectadas con el bus 23 del sistema por medio de una interfaz 32 de unidad de disco duro, de una interfaz 33 de unidad de disco magnético y de una interfaz 34 de unidad óptica, respectivamente. Las unidades y sus medios asociados legibles por un ordenador proporcionan un almacenamiento no volátil de instrucciones legibles por un ordenador, estructuras de datos, módulos de programa y otros datos para el ordenador 20. Aunque el entorno descrito en la presente memoria emplea un disco duro, un disco magnético extraíble 29 y un disco óptico extraíble 31, los expertos en la técnica deberían apreciar que también se pueden utilizar en el entorno operativo otros tipos de medios legibles por un ordenador que pueden almacenar datos que son accesibles por un ordenador, tales como casetes magnéticos, tarjetas de memoria *flash*, discos de vídeo digital, cartuchos Bernoulli, memorias de acceso aleatorio (RAM), memorias de solo lectura (ROM) y similares.

Se puede almacenar un número de módulos de programa en el disco duro, en el disco magnético 29, en el disco óptico 31, en la ROM 24 o en la RAM 25, incluyendo un sistema operativo 35, uno o más programas 36 de aplicación, otros módulos 37 de programa y datos 38 de programa. Un usuario puede introducir instrucciones e información en el ordenador 20 a través de dispositivos de entrada, tales como un teclado 40 y un dispositivo 42 de puntero. Otros dispositivos (no mostrados) de entrada pueden incluir un micrófono, una palanca de juego, un mando para juegos, una antena parabólica, un escáner o similares. Estos y otros dispositivos de entrada están conectados a menudo con la unidad 21 de procesamiento por medio de una interfaz 46 de puerto serie que está acoplada con el bus del sistema, pero pueden estar conectados por otras interfaces, tales como un puerto paralelo, un puerto para juegos o un bus serie universal (USB). También se puede conectar un medio 47 de visualización u otro tipo de dispositivo de visualización con el bus 23 del sistema mediante una interfaz, tal como un adaptador 48 de vídeo. Además del medio 47 de visualización, los ordenadores normalmente incluyen otros dispositivos periféricos (no mostrados) de salida, tales como altavoces e impresoras. El sistema de la FIG. 1 también incluye un adaptador host 55, un bus 56 de interfaz de sistemas informáticos pequeños (SCSI) y un dispositivo 62 de almacenamiento externo conectado con el bus 56 de SCSI.

El ordenador 20 puede operar en un entorno en red utilizando conexiones lógicas a uno o más ordenadores remotos, tal como un ordenador remoto 49. El ordenador remoto 49 puede ser otro ordenador, un servidor, un dispositivo de encaminamiento, un PC de red, un dispositivo del mismo nivel u otro nodo común de red y, normalmente, puede incluir muchos de los elementos, o todos ellos, descritos anteriormente con respecto al ordenador 20, aunque solo se ha ilustrado un dispositivo 50 de almacenamiento de memoria en la FIG. 1c. Las conexiones lógicas mostradas en la FIG. 1 pueden incluir una red 51 de área local (LAN) y una red 52 de área amplia (WAN). Tales entornos de red son comunes en oficinas, en redes empresariales de ordenadores, intranets e Internet.

Cuando se utiliza en un entorno de red LAN, el ordenador 20 puede estar conectado con la LAN 51 mediante un adaptador 53 o interfaz de red. Cuando se utiliza en un entorno de red WAN, el ordenador 20 puede incluir,

normalmente, un módem 54 u otros medios para establecer comunicaciones por la red 52 de área amplia, tal como Internet. El módem 54, que puede ser interno o externo, puede estar conectado con el bus 23 del sistema a través de la interfaz 46 del puerto serie. En un entorno de red, se pueden almacenar los módulos de programa mostrados relativos al ordenador 20, o porciones de los mismos, en el dispositivo remoto de almacenamiento de memoria. Se apreciará que las conexiones de red mostradas son ejemplos y se pueden utilizar otros medios para establecer un enlace de comunicaciones entre los ordenadores. Además, aunque se prevé que numerosas realizaciones de la invención sean particularmente aptas para sistemas de ordenador, no se prevé que nada en el presente documento limite la divulgación a tales realizaciones.

La anterior descripción detallada ha expuesto diversas realizaciones de los sistemas y/o de los procedimientos mediante ejemplos y/o diagramas operativos. Dado que los diagramas de bloque y/o ejemplos contienen una o más funciones y/u operaciones, los expertos en la técnica comprenderán que cada función y/u operación en tales diagramas de bloque, o ejemplos, puede ser implementada, individual y/o colectivamente, mediante una amplia gama de soporte físico, de soporte lógico, de soporte físico inalterable o casi cualquier combinación de los mismos.

En último lugar, aunque se ha descrito la presente divulgación en conexión con los aspectos preferentes, según se ilustra en las diversas figuras, se comprenderá que se pueden utilizar otros aspectos similares o se pueden realizar modificaciones y adicionales a los aspectos descritos para llevar a cabo la misma función de la presente divulgación sin desviarse de los mismos. Por ejemplo, en diversos aspectos de la divulgación, se divulgaron diversos mecanismos para acotar el comportamiento de una máquina virtual no privilegiada que interactúa con un dispositivo. Sin embargo, también se contemplan por las enseñanzas de la presente memoria otros mecanismos equivalentes a estos aspectos descritos. Por lo tanto, la presente divulgación no debería estar limitada a ningún aspecto individual, sino más bien interpretada en el ámbito y el alcance según las reivindicaciones adjuntas.

La invención está definida por las reivindicaciones independientes adjuntas. Las reivindicaciones dependientes definen realizaciones preferentes.

REIVINDICACIONES

1. Un procedimiento para gestionar las comunicaciones entre una máquina virtual (120) y un dispositivo (430) de I/O mediante una capa (140) de virtualización, en el que la capa (140) de virtualización mantiene la máquina virtual (120),
recibiendo una descripción del dispositivo de I/O, comprendiendo la descripción información relativa a qué acciones en el dispositivo de I/O tienen efectos en todo el sistema y que tienen efectos que son locales al dispositivo de I/O;
en función de dicha descripción, construir (602) una representación de un espacio de configuración del dispositivo de I/O que indica las acciones que puede llevar la máquina virtual a cabo sobre el dispositivo de I/O para las ubicaciones de memoria en dicho espacio de configuración;
en función de dicha descripción, construir (603) una representación de un espacio de I/O asignado con la memoria del dispositivo de I/O, en el que cada página del espacio de I/O asignado con la memoria está asignada en dicha máquina virtual o es excluida de dicha máquina virtual;
instalar un controlador para el dispositivo de I/O en la máquina virtual (120); y
controlar el acceso (604) a dicho dispositivo de I/O según dicha representación del espacio de configuración y dicha representación del espacio de I/O asignado con la memoria, en el que dicha información es recibida en un fichero proporcionado por un proveedor de dicho dispositivo de I/O, y en el que dicha construcción de una representación del espacio de configuración y dicha construcción de una representación del espacio de I/O asignado con la memoria comprende, además, construir las representaciones según dicha información.

2. El procedimiento según la reivindicación 1, en el que dicha construcción de una representación del espacio de configuración comprende, además, asociar cada bit en dicha representación del espacio de configuración con al menos una operación (605) de lectura y de escritura.

3. El procedimiento según la reivindicación 1, en el que para cualquier memoria excluida de dicha representación del espacio de configuración o para cualquier memoria excluida de dicha representación del espacio de I/O asignado con la memoria, se puebla dicha cualquier memoria con datos representativos de dicho dispositivo (606) de I/O.

4. El procedimiento según la reivindicación 2, en el que dichas operaciones (706) de lectura y de escritura comprenden:
solo lectura, siempre cero en lectura, siempre uno en lectura, lectura-escritura, la escritura de uno borra / la escritura de cero se deja tal cual, la escritura de uno establece / la escritura de cero se deja tal cual, la escritura de cero borra / la escritura de uno se deja tal cual, la escritura de cero establece / la escritura de uno se deja tal cual, borrar a cero tras la primera lectura, y establecer a uno tras la primera lectura.

5. El procedimiento según la reivindicación 3, que comprende, además:

definir bits en páginas para la memoria excluida (802); y
recibir interceptaciones y procesar las interceptaciones utilizando páginas con los bits definidos (804).

6. El procedimiento según la reivindicación 1, que comprende, además:

construir una representación del espacio de I/O (815);
poblar dicha representación del espacio de I/O en función de dicha información recibida (815); y
controlar el acceso a dicho dispositivo de I/O según dicha representación del espacio de I/O (825).

7. El procedimiento según la reivindicación 1, que comprende, además, poblar tanto dicha asignación del espacio de configuración como dicha asignación del espacio de I/O asignado con la memoria en función de la información recibida (830).

8. El procedimiento según la reivindicación 1, que comprende, además, excluir el espacio de I/O de dicha máquina virtual (835).

9. Un sistema adaptado para gestionar las comunicaciones entre una máquina virtual (120) y un dispositivo (430) de I/O que implementa un espacio de configuración y un espacio de I/O asignado con la memoria, que comprende:

al menos un procesador; y
al menos una memoria acoplada de forma comunicativa con dicho al menos un procesador, teniendo la memoria almacenada en la misma instrucciones ejecutables por un ordenador con capacidad para implementar el procedimiento según una de las reivindicaciones precedentes.

10. El sistema de la reivindicación 9, que comprende embeber dicha representación en un fichero de instalación para un controlador para dicho dispositivo de I/O.

11. Un medio (1000) de almacenamiento legible por un ordenador que almacena en el mismo instrucciones ejecutables por un ordenador para controlar el acceso a un dispositivo de I/O, en el que el dispositivo de I/O está acoplado de forma comunicativa con una máquina física que alberga máquinas virtuales, que comprende instrucciones para implementar el procedimiento según una de las reivindicaciones 1 a 8.

12. El medio de almacenamiento legible por un ordenador de la reivindicación 11, que comprende, además, instrucciones para:

recibir un fichero de instalación para el dispositivo, en el que el fichero de instalación comprende información relativa a qué operaciones sobre el dispositivo tienen efectos en todo el sistema y cuáles tienen efectos que son locales al dispositivo (1010);
construir al menos una asignación de atributos para el espacio de configuración del dispositivo, el espacio de I/O asignado con la memoria y el espacio de I/O del dispositivo, en la que cada página o cada bit asociado con la al menos una asignación está asignado en dicha máquina virtual, y en la que se puede proporcionar una página estática de bits a una máquina virtual como el estado del dispositivo (1012);
poblar la al menos una asignación en función de dicho fichero recibido (1014) de instalación; y
utilizar la al menos una asignación para gestionar el acceso al dispositivo (1014).

13. El medio de almacenamiento legible por un ordenador de la reivindicación 12, en el que cada bit en dicha asignación o página asociada con la al menos una asignación contiene uno de los siguientes atributos (1020); solo lectura, siempre cero en lectura, siempre uno en lectura, lectura-escritura, la escritura de uno borra / la escritura de cero se deja tal cual, la escritura de uno establece / la escritura de cero se deja tal cual, la escritura de cero borra / la escritura de uno se deja tal cual, la escritura de cero establece / la escritura de uno se deja tal cual, borrar a cero tras la primera lectura o establecer a uno tras la primera lectura.

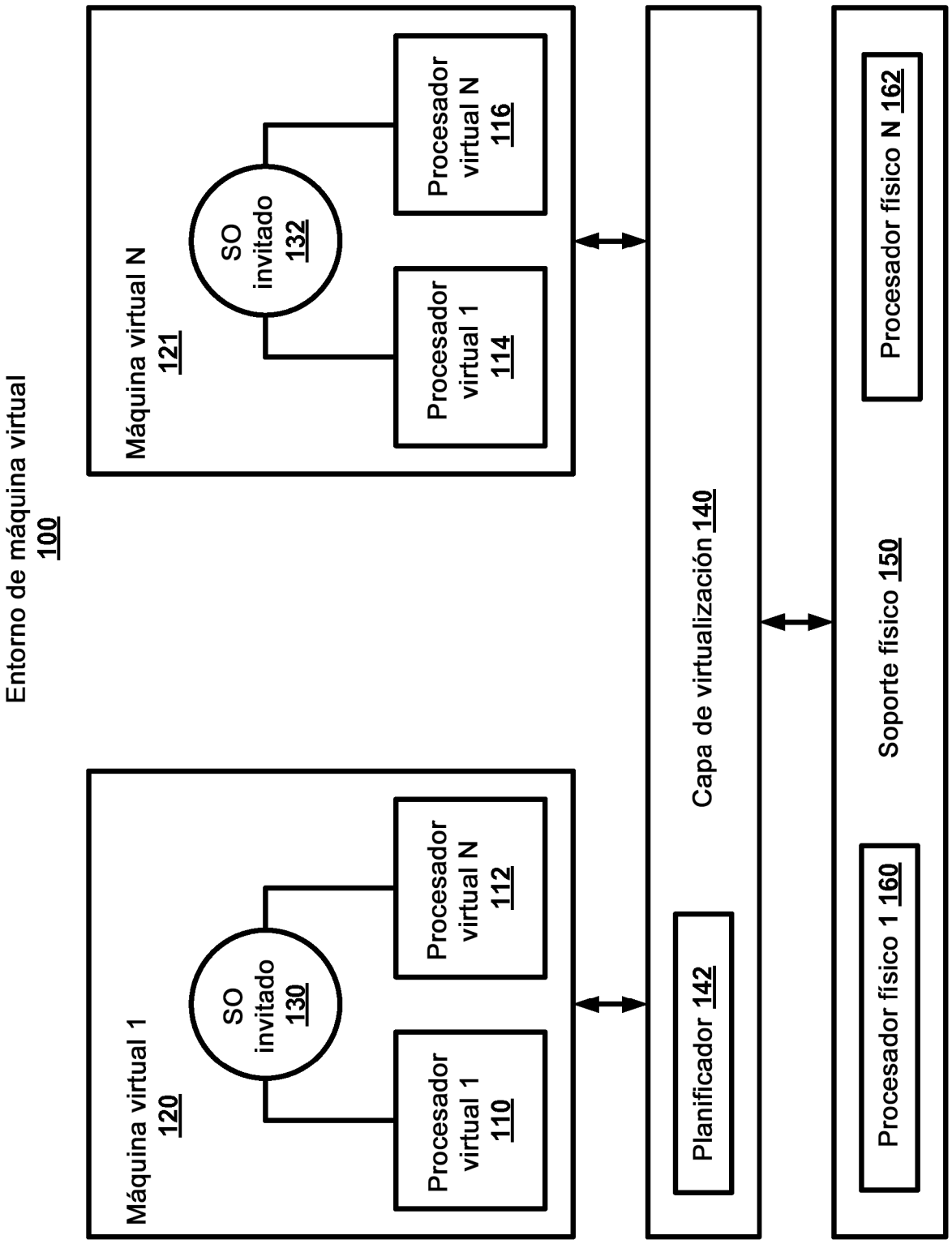
Entorno de máquina virtual
100
Máquina virtual 1
120
SO invitado
130
Procesador virtual 1
110
Procesador virtual N
112
Máquina virtual N
121
SO invitado
132
Procesador virtual 1
114
Procesador virtual N
116
Capa de virtualización 140
Planificador 142
Soporte físico 150
Procesador físico 1 160
Procesador físico N 162

FIG. 1a

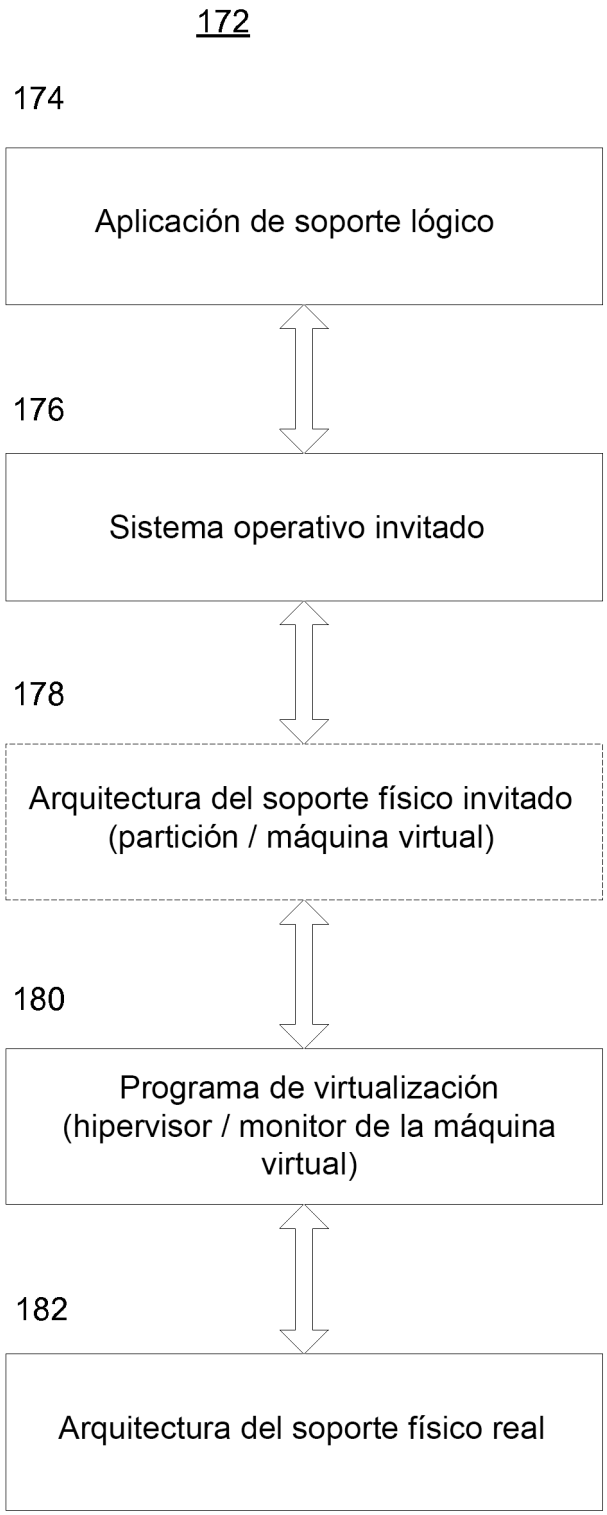
172
174
Aplicación de soporte lógico
176
Sistema operativo invitado
178
Arquitectura del soporte físico invitado (partición / máquina virtual)
180
Programa de virtualización (hipervisor / monitor de la máquina virtual)
182
Arquitectura del soporte físico real

FIG. 1b

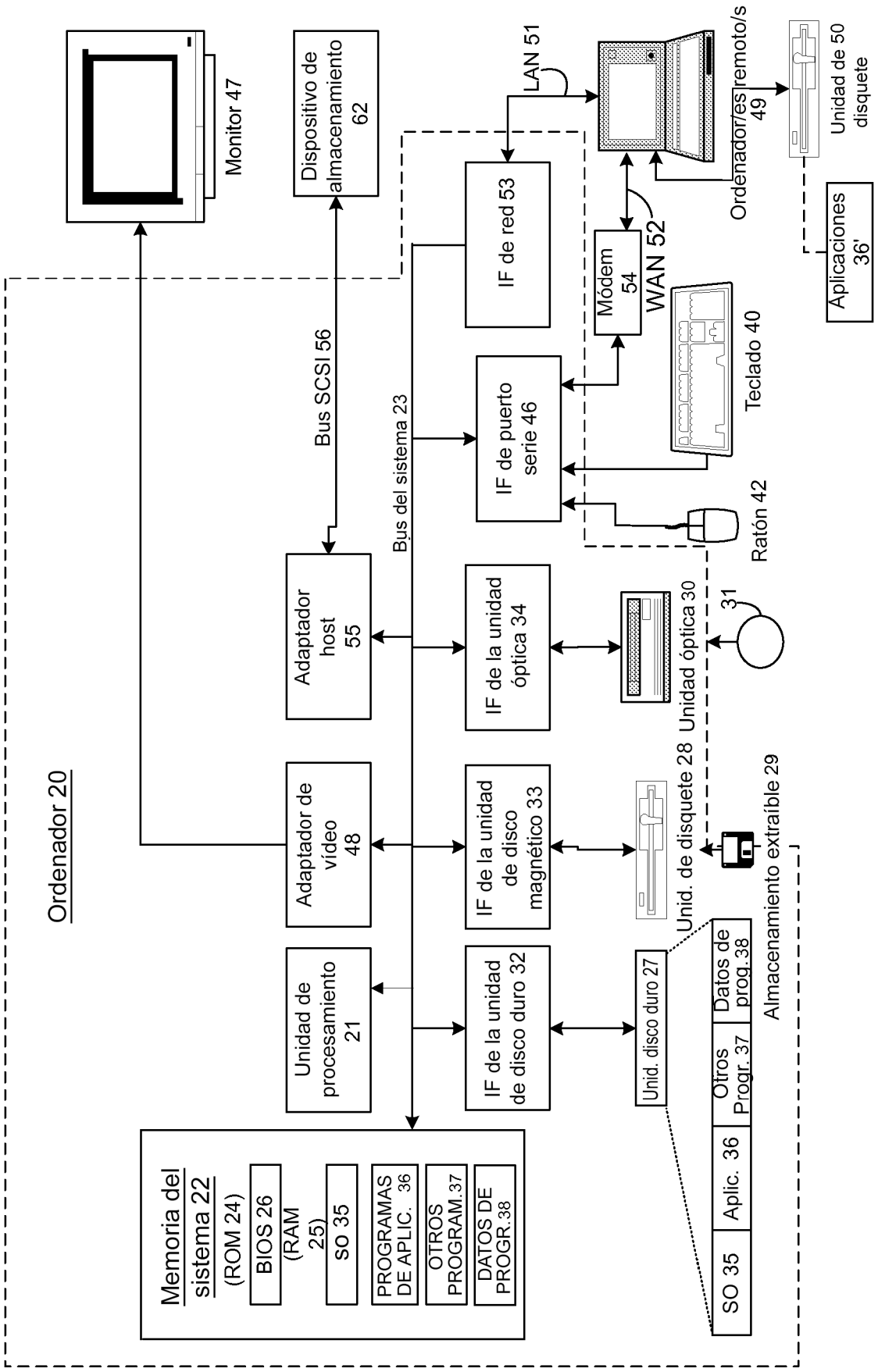
Ordenador 20
Memoria del sistema 22
(ROM 24)
BIOS 26
(RAM 25)
SO 35
PROGRAMAS DE APLIC. 36
OTROS PROGRAM.37
DATOS DE PROGR.38
Unidad de procesamiento 21
Adaptador de vídeo 48
Adaptador host 55
IF de la unidad de disco duro 32
IF de la unidad de disco magnético 33
IF de la unidad óptica 34
IF de puerto serie 46
IF de red 53
Bus del sistema 23
Bus SCSI 56
Monitor 47
Dispositivo de almacenamiento 62
Unid. disco duro 27
SO 35
Aplic. 36
Otros Progr. 37
Datos de prog.38
Unid. de disquete 28
Unidad óptica 30
31
Almacenamiento extraíble 29
Ratón 42
Teclado 40
Módem 54
WAN 52
LAN 51
Ordenador/es remoto/s 49
Unidad de disquete 50
Aplicaciones 36'

FIG. 1c

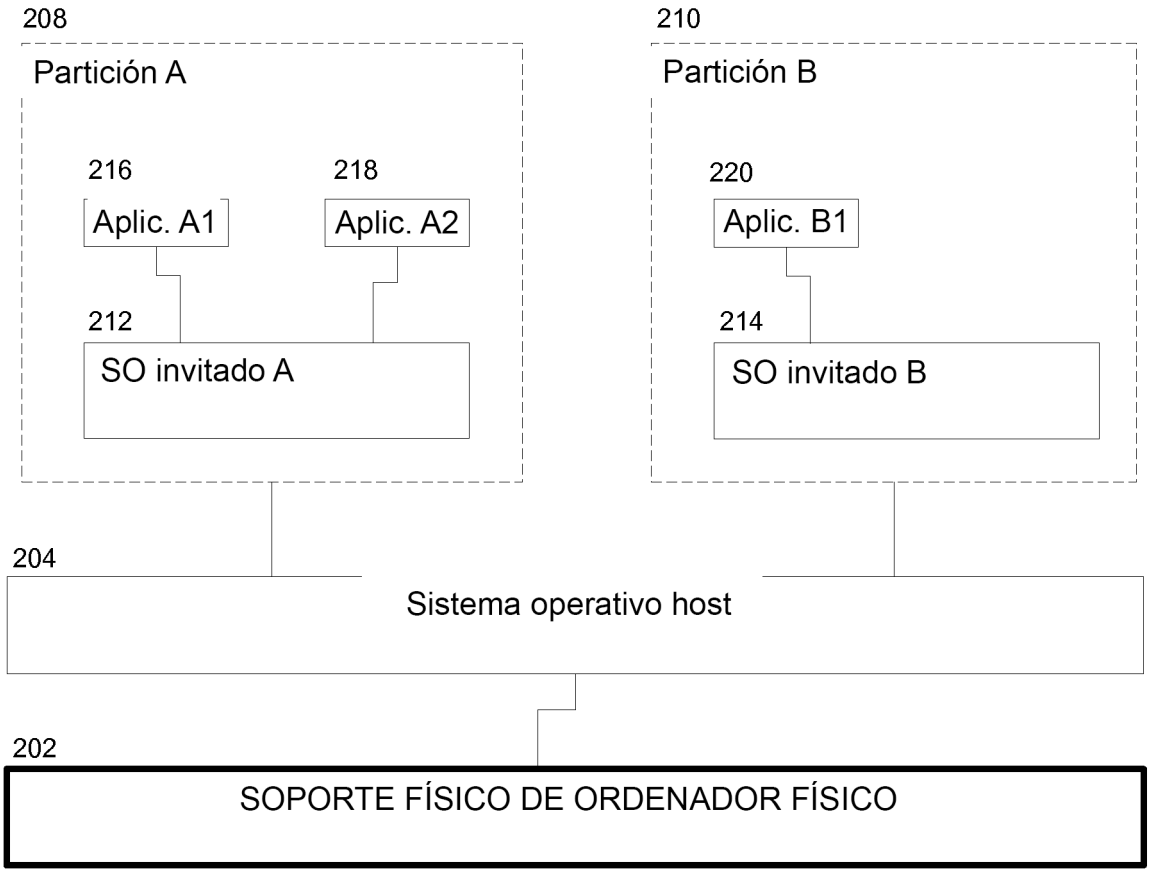
208
Partición A
216
Aplic. A1
218
Aplic. A2
212
SO invitado A
210
Partición B
220
Aplic. B1
214
SO invitado B
204
Sistema operativo host
202
SOPORTE FÍSICO DE ORDENADOR FÍSICO

FIG. 2

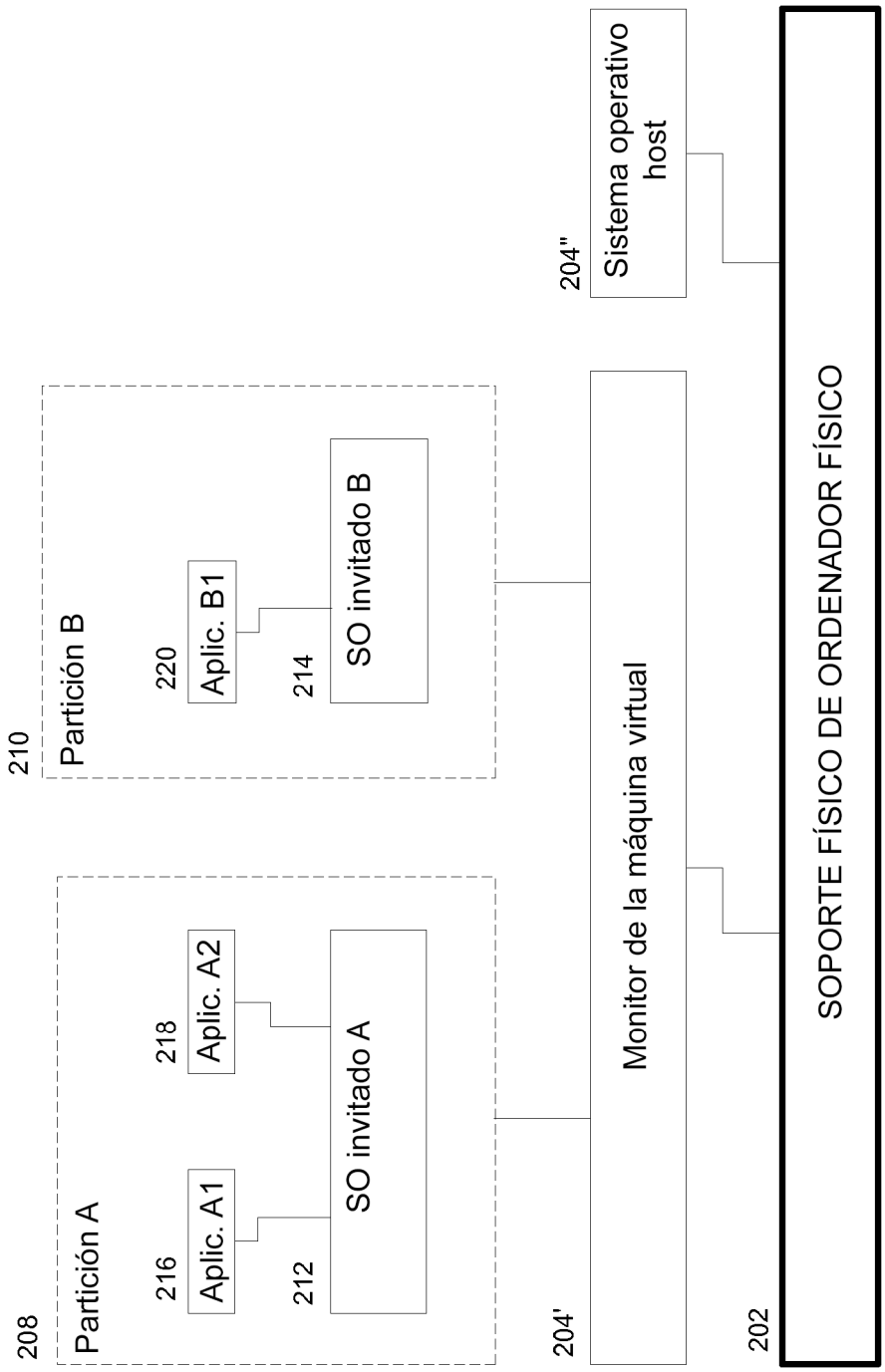
208
Partición A
216
Aplic. A1
218
Aplic. A2
212
SO invitado A
210
Partición B
220
Aplic. B1
214
SO invitado B
204'
Monitor de la máquina virtual
204"
Sistema operativo host
202
SOPORTE FÍSICO DE ORDENADOR FÍSICO

FIG. 3

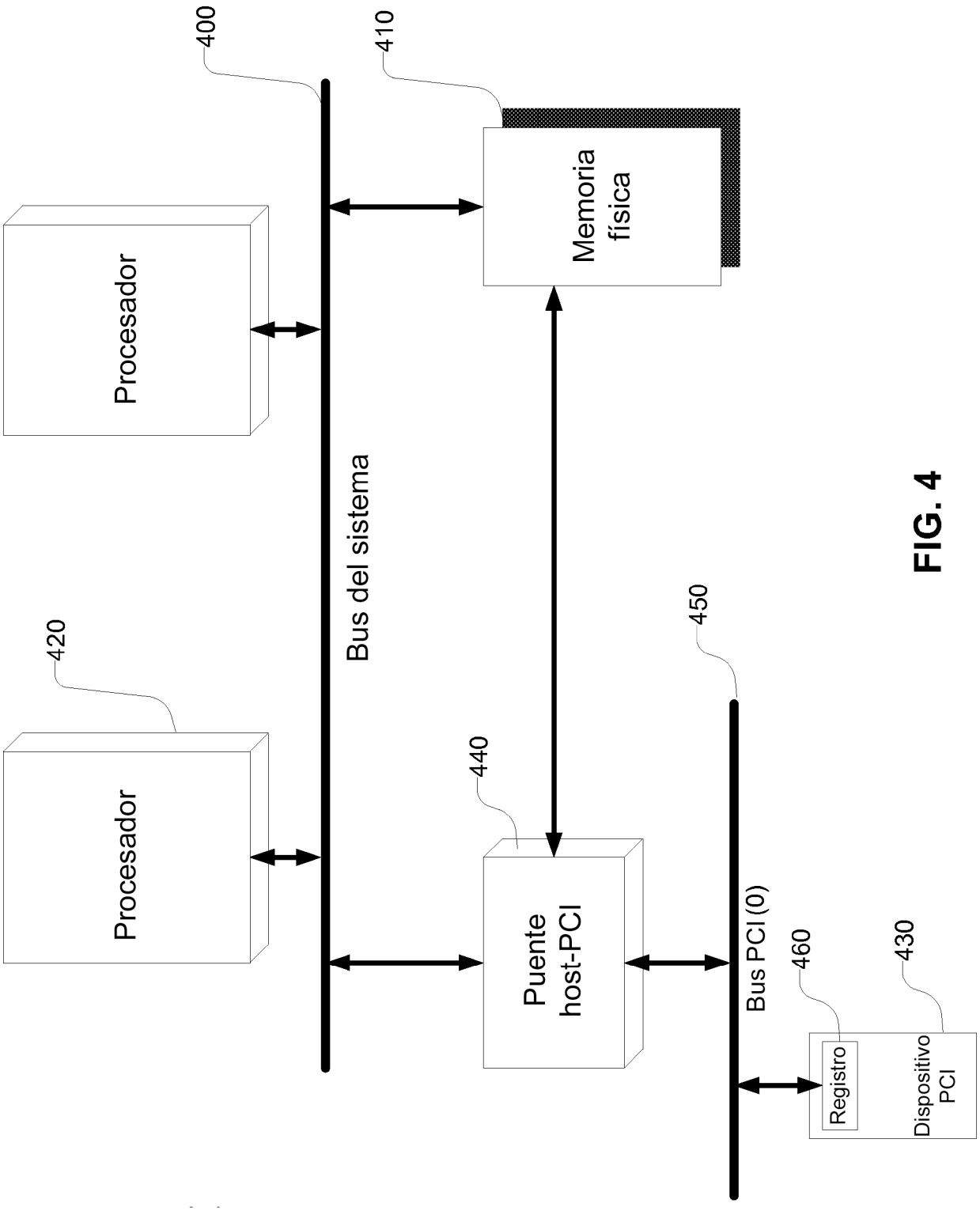
400
410
Memoria física
Procesador
Bus del sistema
420
Procesador
440
450
Puente host-PCI
Bus PCI (0)
460
430
Registro
Dispositivo PCI

FIG. 4

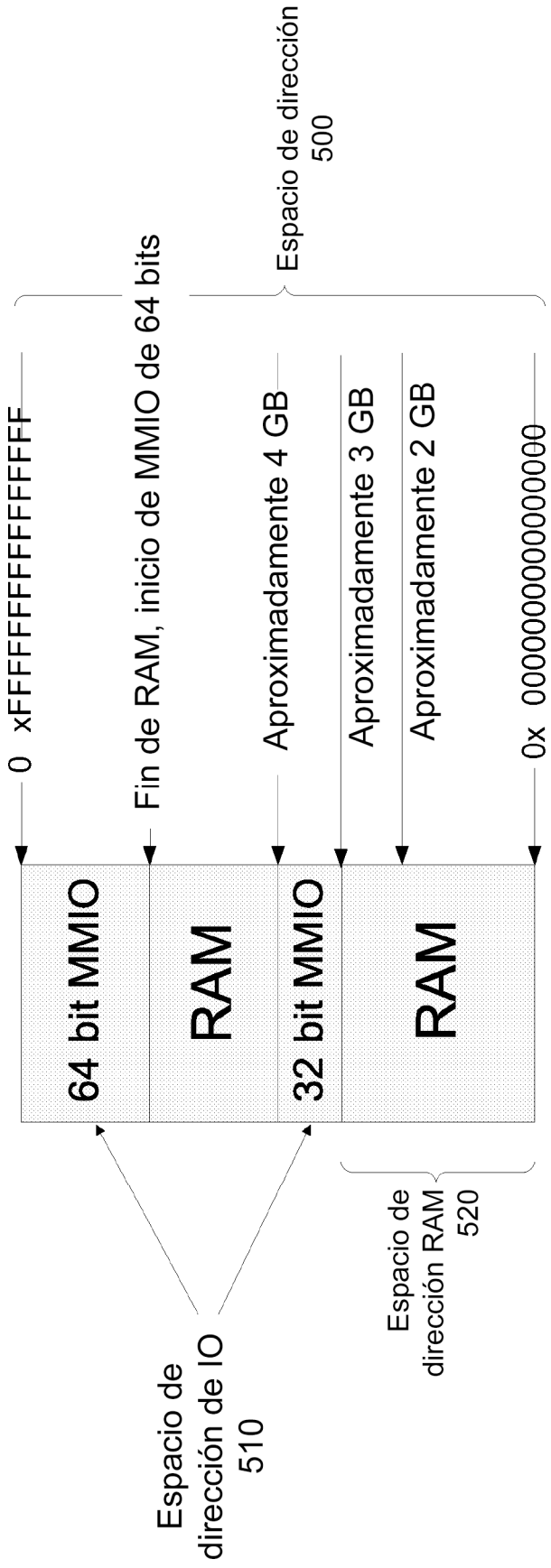
Espacio de dirección
500
0 xFFFFFFFFFFFFFFFF
Fin de RAM, inicio de MMIO de 64 bits
Aproximadamente 4 GB
Aproximadamente 3 GB
Aproximadamente 2 GB
0x 0000000000000000
64 bit MMIO
RAM
32 bit MMIO
RAM
Espacio de dirección de IO
510
Espacio de dirección RAM
520

FIG. 5

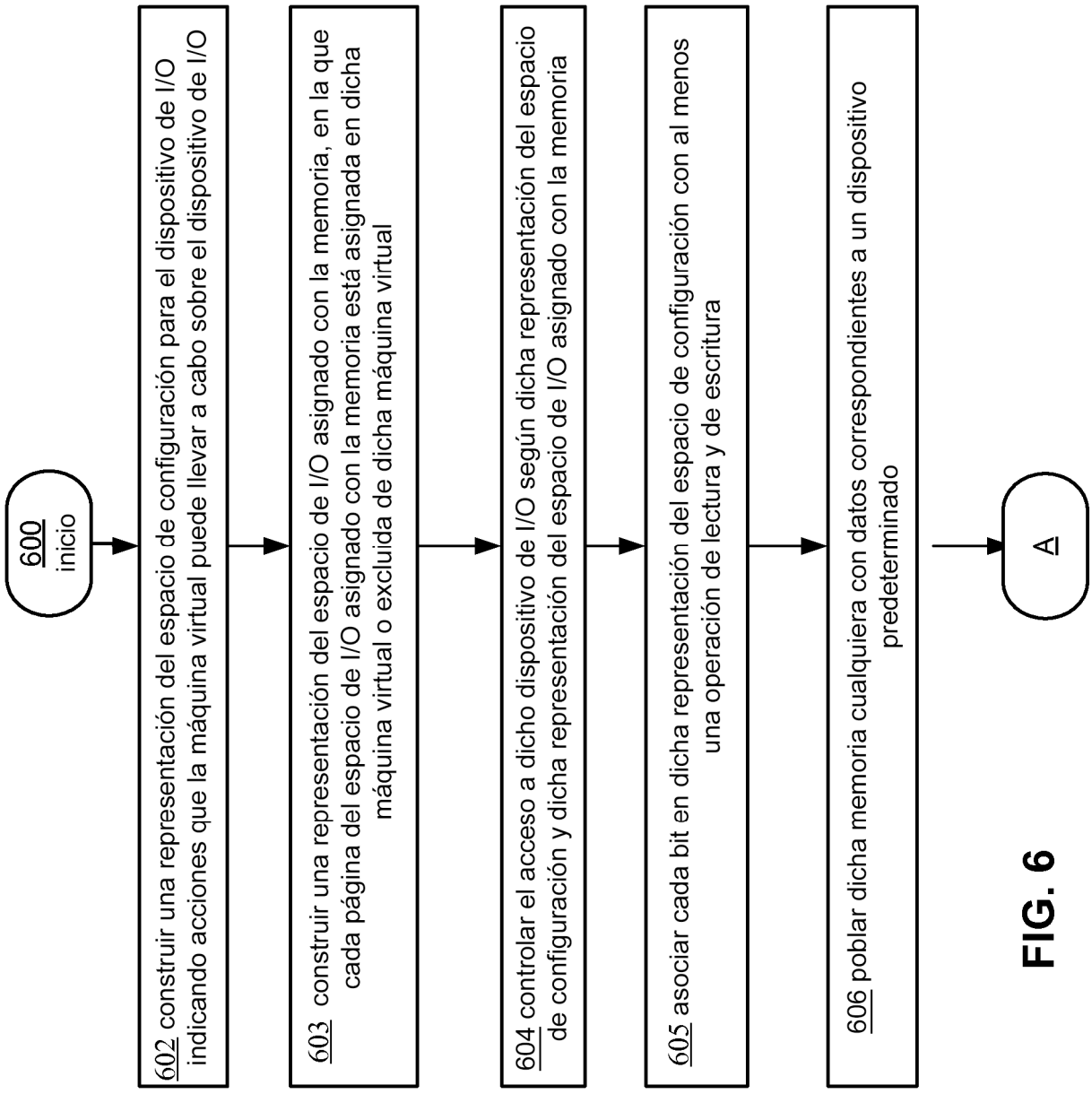
600 inicio
602 construir una representación del espacio de configuración para el dispositivo de I/O indicando acciones que la máquina virtual puede llevar a cabo sobre el dispositivo de I/O
603 construir una representación del espacio de I/O asignado con la memoria, en la que cada página del espacio de I/O asignado con la memoria está asignada en dicha máquina virtual o excluida de dicha máquina virtual
604 controlar el acceso a dicho dispositivo de I/O según dicha representación del espacio de configuración y dicha representación del espacio de I/O asignado con la memoria
605 asociar cada bit en dicha representación del espacio de configuración con al menos una operación de lectura y de escritura
606 poblar dicha memoria cualquiera con datos correspondientes a un dispositivo predeterminado
A

FIG. 6

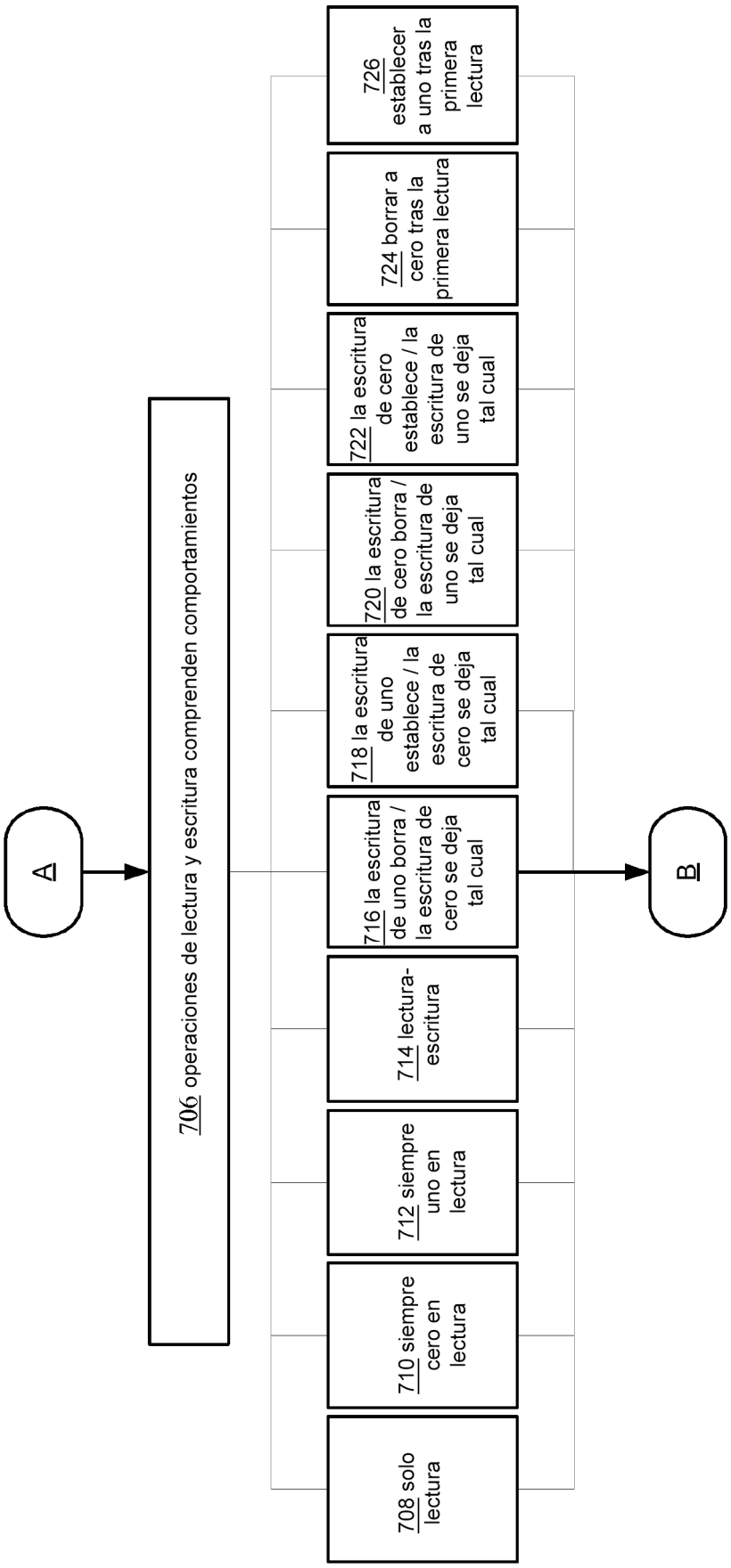
A
706 operaciones de lectura y escritura comprenden comportamientos
708 solo lectura
710 siempre cero en lectura
712 siempre uno en lectura
714 lectura-escritura
716 la escritura de uno borra / la escritura de cero se deja tal cual
718 la escritura de uno establece / la escritura de cero se deja tal cual
720 la escritura de cero borra / la escritura de uno se deja tal cual
722 la escritura de cero establece / la escritura de uno se deja tal cual
724 borrar a cero tras la primera lectura
726 establecer a uno tras la primera lectura
B

FIG. 7

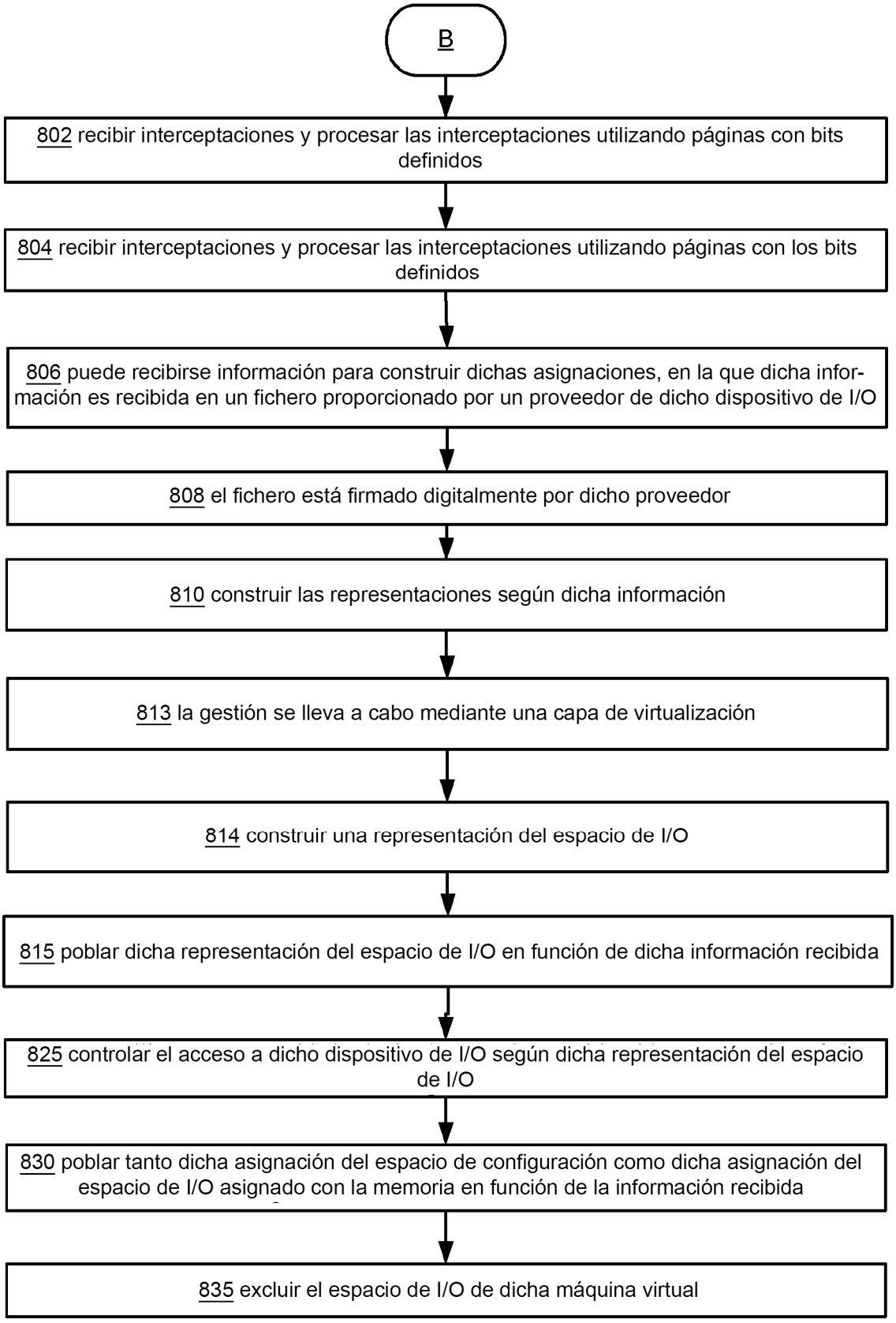
B
802 recibir interceptaciones y procesar las interceptaciones utilizando páginas con bits definidos
804 recibir interceptaciones y procesar las interceptaciones utilizando páginas con los bits definidos
806 puede recibirse información para construir dichas asignaciones, en la que dicha información es recibida en un fichero proporcionado por un proveedor de dicho dispositivo de I/O
808 el fichero está firmado digitalmente por dicho proveedor
810 construir las representaciones según dicha información
813 la gestión se lleva a cabo mediante una capa de virtualización
814 construir una representación del espacio de I/O
815 poblar dicha representación del espacio de I/O en función de dicha información recibida
825 controlar el acceso a dicho dispositivo de I/O según dicha representación del espacio de I/O
830 poblar tanto dicha asignación del espacio de configuración como dicha asignación del espacio de I/O asignado con la memoria en función de la información recibida
835 excluir el espacio de I/O de dicha máquina virtual

FIG. 8

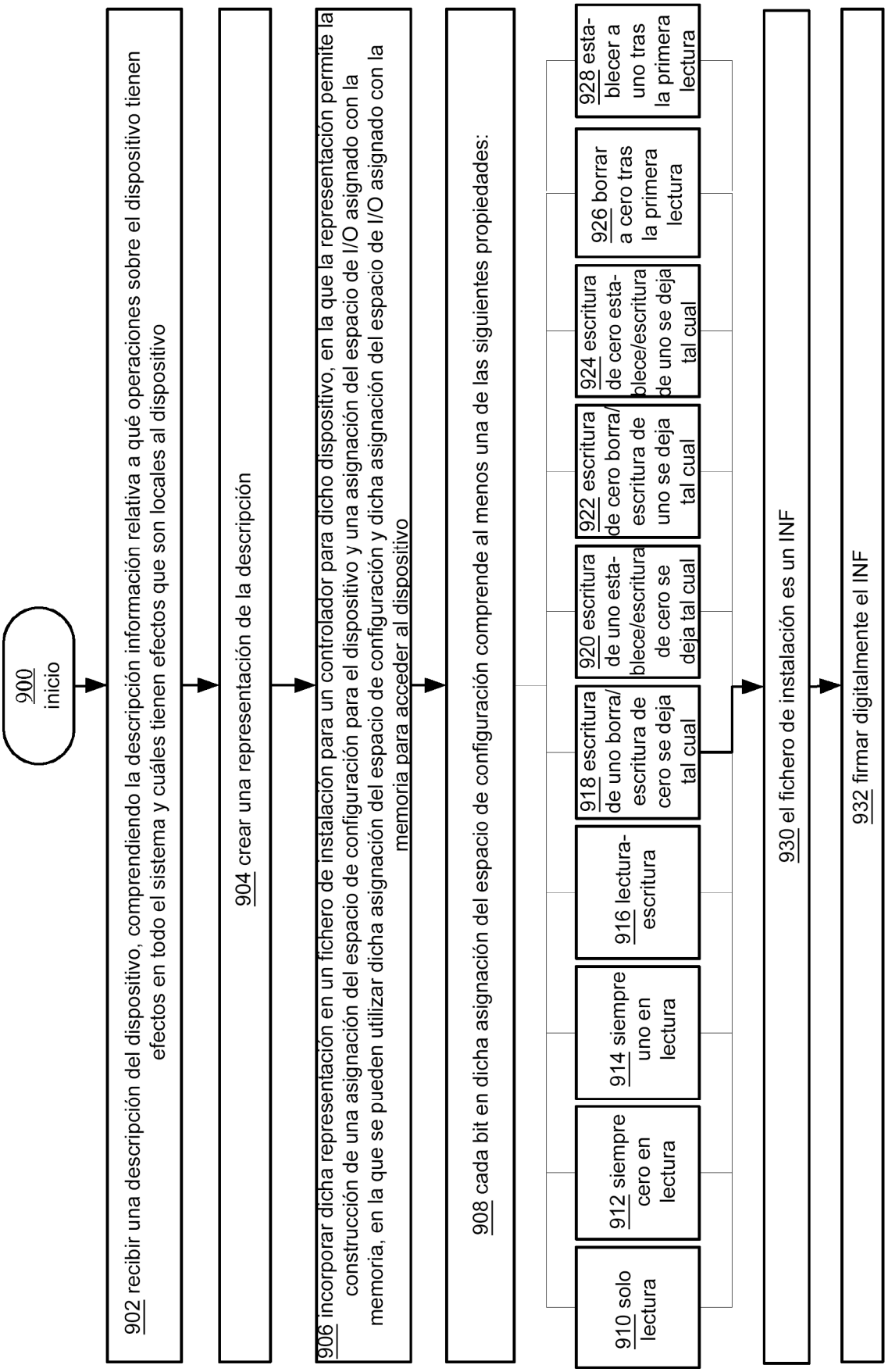
900 inicio
902 recibir una descripción del dispositivo, comprendiendo la descripción información relativa a qué operaciones sobre el dispositivo tienen efectos en todo el sistema y cuáles tienen efectos que son locales al dispositivo
904 crear una representación de la descripción
906 incorporar dicha representación en un fichero de instalación para un controlador para dicho dispositivo, en la que la representación permite la construcción de una asignación del espacio de configuración para el dispositivo y una asignación del espacio de I/O asignado con la memoria, en la que se pueden utilizar dicha asignación del espacio de configuración y dicha asignación del espacio de I/O asignado con la memoria para acceder al dispositivo
908 cada bit en dicha asignación del espacio de configuración comprende al menos una de las siguientes propiedades:
910 solo lectura
912 siempre cero en lectura
914 siempre uno en lectura
916 lectura-escritura
918 escritura de uno borra/escritura de cero se deja tal cual
920 escritura de uno establece/escritura de cero se deja tal cual
922 escritura de cero borra/escritura de uno se deja tal cual
924 escritura de cero establece/escritura de uno se deja tal cual
926 borrar a cero tras la primera lectura
928 establecer a uno tras la primera lectura
930 el fichero de instalación es un INF
932 firmar digitalmente el INF

FIG. 9

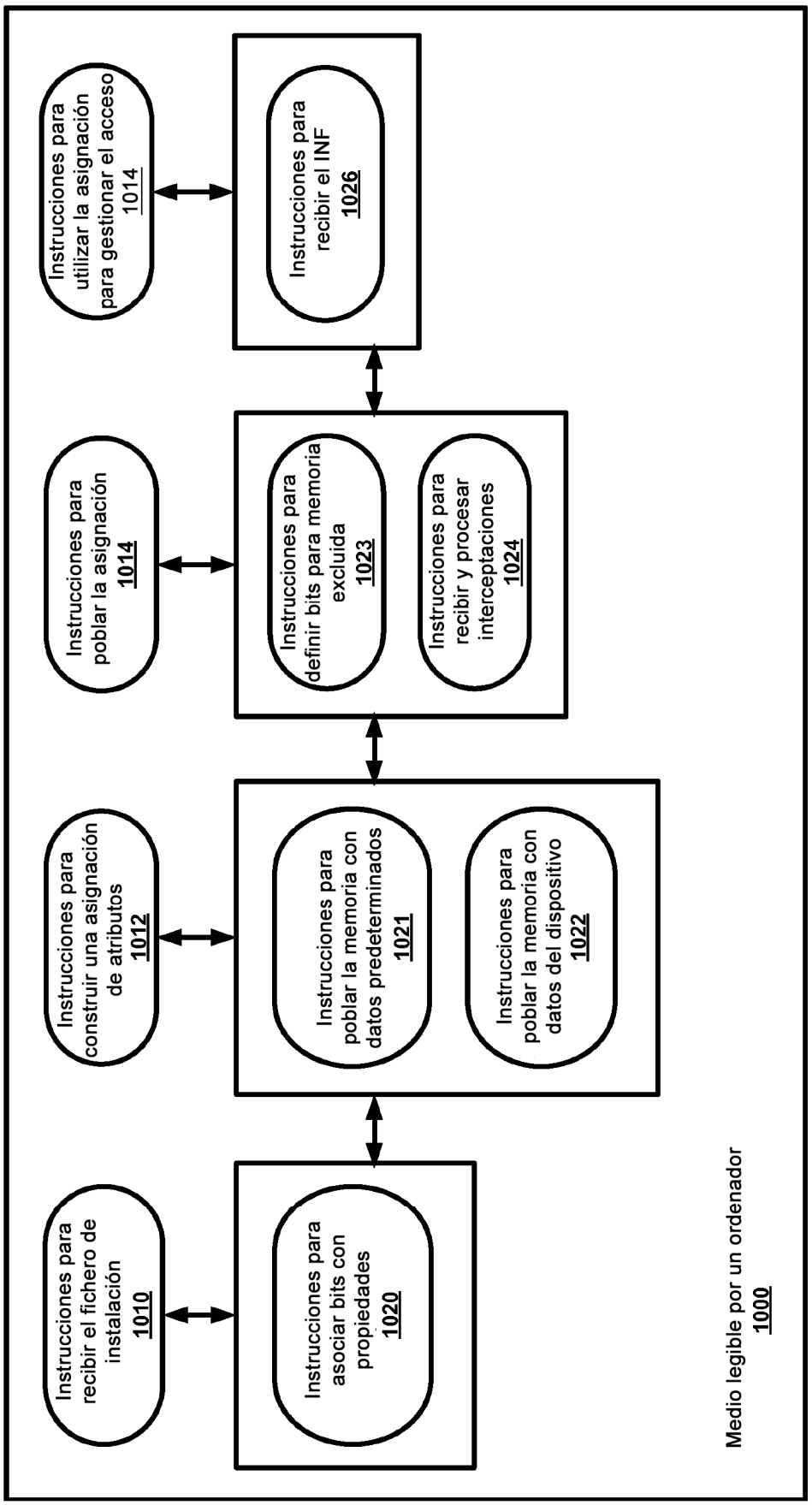
Medio legible por un ordenador
1000
Instrucciones para recibir el fichero de instalación
1010
Instrucciones para asociar bits con propiedades
1020
Instrucciones para construir una asignación de atributos
1012
Instrucciones para poblar la memoria con datos predeterminados
1021
Instrucciones para poblar la memoria con datos del dispositivo
1022
Instrucciones para poblar la asignación
1014
Instrucciones para definir bits para memoria excluida
1023
Instrucciones para recibir y procesar interceptaciones
1024
Instrucciones para utilizar la asignación para gestionar el acceso
1014
Instrucciones para recibir el INF
1026

FIG. 10